

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Sociology

2021

Pretrial Risk Assessment Algorithms in the Courtroom

Zhang, Simone Ximeng

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp0176537449g>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

PRETRIAL RISK ASSESSMENT ALGORITHMS IN THE COURTROOM

SIMONE ZHANG

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
SOCIOLOGY
ADVISER: DALTON CONLEY

NOVEMBER 2021

© Copyright by Simone Zhang, 2021.

All Rights Reserved

Abstract

This dissertation investigates how algorithmic risk assessments used in criminal courts shape decision outcomes, courtroom debate, and the claims that judges, prosecutors, and defense attorneys advance about defendants. Focusing on a tool intended to help judges set pretrial release conditions, I build on a pre-existing randomized controlled trial in a U.S. county that randomized whether the risk assessment report for a given arrested individual was provided to the court or withheld. Using a mix of qualitative and quantitative methods, I analyze administrative data and a sample of court hearings transcripts from the county.

Chapter 2 explores the effect of access to risk assessment reports on how often judges order cash bail, with a focus on how the information influences whether prosecutors and defense attorneys secure their requested bail conditions. The results indicate that risk assessments can have an asymmetric effect, with recommendations for cash bail exerting more influence than recommendations to not require cash bail. Chapter 3 identifies two additional ways that the risk assessment tool altered the behavior of judges: how often they required that defendants be subject to pretrial supervision and how they spoke to defendants. Chapter 4 tests theoretical predictions that risk assessments might affect how defendants are evaluated and shape courtroom actors' understandings of the goals of their decision-making. It demonstrates that risk assessment tools may not lead courtroom actors to fundamentally reorient their approach to evaluating defendants, but may spur shifts in some specific domains of their practices. Some of these shifts may be regarded as counterproductive, while others may help align practices with legal ideals. Together, the dissertation contributes to our understanding of the on-the-ground impact of algorithmic decision aids and highlights some of the trade-offs that the adoption of such tools may entail.

Acknowledgements

I would first like to thank my chair, Dalton Conley, whose generosity, intellectual curiosity, and drive to get to the bottom of things have inspired me throughout my time at Princeton. Dalton helped me to navigate numerous roadblocks that I encountered as I tried to move this research forward, and I am immensely grateful for his support and feedback. Brandon Stewart helped me to grow in more ways than I can count. I thank him for building several wonderful communities from which I benefited tremendously and for serving as a model for the type of mentor and teacher I hope to become. Matt Salganik also played an important role in shaping my trajectory. He challenged me to engage broadly and generously across disciplinary boundaries and I am thankful for his words of encouragement and thoughtful advice. I am also indebted to Adam Goldstein, who offered incisive feedback on my work and helped to move it forward in productive directions.

Beyond my committee, several other faculty members contributed to my growth at Princeton. I would like to express my gratitude to Paul Starr for serving as an external reader, for many helpful conversations as I developed different research ideas, and for always encouraging me to think bigger. I also thank Viviana Zelizer and Janet Vertesi for their excellent advice and critical engagement with my work.

This dissertation could not have gotten off the ground without the generosity of many people. I thank James Greiner for sharing the Access to Justice Lab's randomization data with me despite the endless wrangling with bureaucracy that entailed and Ryan Halen for his assistance with the data transfer. At Princeton, Donna DeFrancisco, Mona Hall, Amanda Rowe, and Mellissa Williams helped me to secure the funding and data sharing agreements the project required. I am also grateful to Barry Widera, Jo Reis, and Carlo Esqueda for their help with local court records; Clark Rodgers and Maggie Thomas for patiently helping me to better understand various aspects of the Public Safety Assessment's local implementation; and McKay Bartlett for her careful research assistance. I owe a special debt of gratitude to Lori Gerlach and Jeanne Farino-Malagold who worked nights and

weekends during a difficult year to complete hundreds of transcript requests.

The research included in this dissertation was supported by the National Science Foundation (Grant Number 2001832) and the Princeton Data-Driven Social Science Initiative.

I was incredibly lucky to go through graduate school with a fantastic cohort of peers and grow as a member of several on-campus communities, including the Stewart Lab, the Biosociology Lab, the Center for the Study of Social Organization, and the Prize Fellowship in Social Sciences. Special thanks to Jeremy Cohen, Chris Felton, Gracie Himmelstein, Rebecca Johnson, Ian Lundberg, Daniela Urbina Julio, Hannah Waight, Taylor Winfield, and Janet Xu for their support and feedback throughout the dissertation writing process.

To close, I would like to thank friends and family who helped me get through the tougher parts of graduate school. My partner, Andres, has served as a sounding board for all of my ideas and I could not be more grateful for his brilliance, understanding, and endless optimism. My sisters, Keri and Melissa, have helped me stay grounded and spoiler-free on all the television shows I fell behind on while writing this dissertation. Finally, I am grateful to my parents, Xiaoying and Wencheng, whose selflessness, patience, and love have made everything possible.

Contents

Abstract	iii
Acknowledgements	iv
1 Introduction	1
1.1 Road Map	5
1.2 The Public Safety Assessment	7
1.3 Study Design Overview	12
1.3.1 Transcript Sampling and Analytic Strategy	14
1.4 The Field Site	19
1.4.1 Demographic and Political Profile of Dane County, WI	19
1.4.2 Governing Laws	20
1.4.3 Local Procedures and PSA Implementation	21
1.4.4 PSA Recommendations: Local Decision Rules and Distribution Relative to Baseline	27
2 Asymmetries in the Use of Pretrial Risk Assessments and Their Impact on Cash Bail and Procedural Outcomes	29
2.1 Decision-making in the Courtroom Workgroup	29
2.2 Do Risk Assessments Reduce or Spread Suspicion and Restriction?	31
2.3 Approach and Methods	33
2.4 Descriptive Statistics	36

2.5	Results	41
2.5.1	Discussions of the Public Safety Assessment	41
2.5.2	Effect of PSA on Hearing Outcomes	45
2.6	Reasoning Through Potential Assumption Violations	52
2.6.1	Restricting Time Between Assessment Completion and Hearing . .	57
2.7	Discussion	60
2.7.1	Summary of findings	60
2.7.2	Implications, Limitations, and Directions for Future Research . . .	61
3	The Reactive Potential of Pretrial Risk Assessments	64
3.1	Introduction	64
3.2	Expanding Assessments of the Impact of Pretrial Risk Assessments	65
3.3	Data and Methods	68
3.3.1	How Judges Speak to Defendants	68
3.3.2	Effect on Bail Monitoring	69
3.4	Results	70
3.4.1	The Rhetorical Use of Pretrial Risk Assessments	70
3.4.2	The Effect of Pretrial Risk Assessments on Bail Monitoring Referrals	76
3.5	Discussion and Implications	83
3.6	Appendix	87
4	The PSA's Impact on Decision Goals and How Defendants are Evaluated	88
4.1	Theoretical Predictions about How Actuarial Prediction Shapes the Means and Ends of Decision-Making	88
4.2	Particularities of the Pretrial Context	91
4.3	Research Questions and Expectations	95
4.4	Methods	96
4.5	Results	97

4.5.1	Results Part I: How defendants are seen: what dimensions of de- fendants are brought to bear in courtroom discussions	98
4.5.2	Results Part II: Decision goals and target outcomes	109
4.6	Conclusion and Discussion	117
4.7	Appendix	122
4.7.1	Heterogeneity by Defendant Race and Sex	122
4.7.2	Explanations of Defendant History: Contextual Factors Highlighted and Judge Responses	126
5	Conclusion	129
5.0.1	Limitations, Implications, and Directions for Future Research . . .	130

Chapter 1

Introduction

Organizations and institutions increasingly turn to predictive algorithms to help them make consequential decisions that shape how and to whom aid, resources, suspicion, and punishment are allocated (O’Neil, 2016; Eubanks, 2018; Brayne, 2017; Burrell and Fourcade, 2021). Today, algorithms are used as decision supports and decision-making mechanisms in wide-ranging applications, from selecting which calls into child maltreatment hotlines to investigate (Chouldechova et al., 2018) to which job applicants to interview (Bogen and Rieke, 2018; Raghavan et al., 2020) to where to dispatch law enforcement (Brayne, 2020). Often, these shifts come into tension with traditional approaches to rendering decisions, such as relying on human judgment or administrative rules, and raise important questions about whether predictive algorithms ameliorate, reproduce, or exacerbate the social problems and inequalities commonly attributed to the systems that they replace.

The deployment of predictive algorithms within the U.S. criminal justice system is a flash point in debates about the implications of this broader turn. The use of actuarial risk assessment algorithms in particular – computational procedures that score criminal justice subjects on their likelihood of committing crimes or violating court conditions in the future – has generated concerns that it may increase detention, expand surveillance and suspicion, widen already substantial racial and economic disparities in punishment, and

erode due process (Starr, 2014; Angwin and Larson, 2016; Eckhouse et al., 2019; Barabas et al., 2019; Mayson, 2019; Goel et al., 2021). Although advocates of these tools have suggested that they may help judges, parole and probation officers, and other criminal justice actors to make fairer, more evidence-based decisions, critics fear that the tools may instead accelerate unnecessary and unequal punishment, all while cloaking decisions in a veil of objectivity (Harcourt, 2007; Van Cleve and Mayes, 2015; Menefee, 2018; Benjamin, 2019).

In light of these debates, this dissertation examines the courtroom impact of a pretrial risk assessment used to help judges make decisions about whether and under what conditions arrested individuals may be released before their cases have concluded. Pretrial risk assessments assign defendants risk scores and generate recommendations for what judges should require of defendants so that defendants may await trial in the community rather than in jail. While judges' options vary across jurisdictions, pretrial risk assessments typically help judges choose between one of three: 1) detain the defendant until his or her case is completed, 2) set a sum of money that the defendant must deposit with the court ("cash bail") to be released pending trial, which is forfeited if the defendant fails to comply with the court's other release conditions like appearing for future court dates, or 3) allow the individual to be released without any monetary conditions, requiring instead that individual sign a promise to return to court (commonly referred to as a "recognizance bond" or "signature bond"). In some cases, pretrial risk assessments also make recommendations about whether defendants should be subject to ongoing pretrial monitoring.

Pretrial risk assessments follow a long string of efforts to improve accountability and fairness in law through quantification (Espeland and Vannebo, 2007). In theory, by helping judges to better calibrate their decisions to a defendant's "true" risk, they are supposed to help combat a host of problems in pretrial justice: the over-use of cash bail, which ties release to a defendant's economic situation; high levels of pretrial detention, which causes significant harms to defendants; and anti-Black discrimination (James, 2004; Arnold et al.,

2018; Digard and Swavola, 2019). Their ascendance as a solution to these problems is emblematic of a broader turn toward governance strategies that employ statistical techniques to characterize individuals along a population distribution and to render decisions based on probabilistic predictions of their future behavior (Simon, 1988; Hacking, 1990; Hannah-Moffat, 2013a; Hirschman and Bosk, 2020). As of 2017, more than one in four people in the United States lived in a jurisdiction that uses a pretrial risk assessment, a sharp increase from one in ten just four years earlier (PJI, 2017).

Yet despite being commonly championed as evidence-based reforms, the actual evidence base around how pretrial risk assessments impact decision-making is thin (Bechtel et al., 2017; Stevenson, 2018). In this dissertation, I address this gap by leveraging a pre-existing randomized controlled trial that randomized whether risk assessment reports were supplied to the judge, prosecutor, and defense attorney in a given case. I use this infrastructure to investigate how the addition of a risk assessment affects decision outcomes and courtroom debate. Broadly, this dissertation asks: what impact do algorithmic risk assessments have on legal practices and the reach of the criminal justice system?

I build on three streams of interdisciplinary scholarship on the consequences of pretrial risk assessments and predictive algorithms more generally. The first is research that focuses on the technical characteristics or performance of risk assessment algorithms. Studies in this vein evaluate, for instance, whether the tools' outputs exhibit biases as a result of the data, analytic choices, or methods that went into their development (Angwin and Larson, 2016; Flores and Bechtel, 2016; Chouldechova, 2016; DeMichele et al., 2020; Desmarais et al., 2021).¹

The second stream of research consists of empirical social science research. These studies encompass a diverse set of approaches, ranging from econometric methods to ethnography to experiments, but broadly center their inquiries on whether and to what extent the focal targets of the risk assessments – judges – respond to the tools acting on them as nudges

¹A sub-genre of this work proposes and evaluates the properties of different formal definitions of fairness (Corbett-Davies and Goel, 2018; Mitchell et al., 2021)

or incursions on their professional authority and autonomy (Christin, 2017; Cowgill, 2018; DeMichele et al., 2018; Stevenson, 2018; Albright, 2019; Redcross et al., 2019).

The third stream of scholarship, meanwhile, features macro-level, primarily theoretical work positing ways in which predictive algorithms in general or risk assessments in particular might reorganize social, political, and economic life in their image (Feeley and Simon, 1992; Fourcade and Healy, 2013). Broadly, this work suggests ways in which the adoption of risk assessments and other predictive tools encourages institutions to further organize their activities and understandings of their objectives around risk management, displacing other potential goals, usually to dehumanizing effect (Simon, 1988; Garland, 2001; Harcourt, 2007).

Across the chapters of this dissertation, I address two sets of challenges in these streams of research. One is the shared practice in technical and macro-theoretical scholarship to begin and end their analyses at what the tools themselves do, albeit for different reasons and in different ways. Within technical research this is perfectly reasonable, as researchers add tremendous value as they identify issues within predictive algorithms even as they operate in isolation. Within macro-level theoretical scholarship, meanwhile, there is a tendency to conflate what the tools do as technical objects with their ultimate social implications, with an assumption often that the logic of the tools translate seamlessly and hegemonically into the way institutions operate. While the latter may be closer to the case in more heavily automated sites like those in finance, predictive decision aids in most settings are merely advisory, leaving considerable space for social processes to structure their real-world impact.

A second set of challenges come from what I contend is under-explored territory in empirical social science research. On the whole, existing research tends to analyze predictive tools primarily as potential impositions or constraints and interpret null results as evidence of overt or tacit professional resistance (Christin, 2017; DeMichele et al., 2018; Brayne and Christin, 2020). Yet surveys of judges, prosecutors, and public defenders also indicate that

substantial segments of each group report using risk assessments and finding them useful (DeMichele et al., 2019; Terranova et al., 2020). This suggests that researchers may need to pay greater attention to what kinds of actions predictive tools *enable*, rather than merely constrain.

I take up these issues in different parts of this dissertation. Some derive and test the implications of macro-level theories to see if they bear out at the meso-level in courtroom practices. Others analyze risk assessment information as resources that different courtroom actors can leverage to advance their own objectives. Throughout, this dissertation also speaks to policy questions, including whether and how access to risk assessments shapes judges' decisions and the alignment between on-the-ground courtroom procedures and legal ideals.

1.1 Road Map

In the remainder of the current chapter, I introduce the risk assessment that I examine in this dissertation, provide an overview of the project's research design, and describe data sources and methods shared across the three empirical chapters of the dissertation. I also offer an overview of my research site, including laws and procedures that govern local pretrial decision-making. The chapter closes with descriptive statistics summarizing the distribution of risk assessment outputs in my research site.

Chapter 2 focuses on the impact of access to risk assessment reports on how often judges require defendants to post cash bail to be released before trial. Rather than focus on judges as the sole targeted actors of risk assessments, this chapter investigates the tools' impact through the lens of how they shape the courtroom dynamics between prosecutors, defense attorneys, and judges. Do risk assessments help prosecutors or defense attorneys secure their requested bail conditions? Do they influence how often judges render decisions that stay within the confines of prosecutor and defense attorney requests? The chapter

further tests claims that risk assessments make defendants appear riskier and, consequently, lead to increases in the use of cash bail. The results indicate that risk assessments increase the chance that prosecutors successfully secure requests for cash bail and enable judges to more often impose cash bail despite consensus requests from prosecutors and defense attorneys for a signature bond. Overall, the results suggest that risk assessments have an asymmetric effect: defendants are punished for being rated high risk but receive minimal credit for being rated low risk.

In **Chapter 3**, I identify two ways that the risk assessment tool altered the behavior of judges beyond their decisions of whether to impose cash bail. First, I show that the risk assessment leads judges to more often order that defendants deemed a medium level of risk be subject to ongoing pretrial supervision. Second, I find that judges leverage risk assessments as part of several rhetorical strategies to encourage defendants to comply with court orders. Both shifts, I argue, are important because they directly affect defendants' experiences with the criminal justice system and because they may further present mechanisms through which the use of the tool may influence the very outcomes the tool is said to predict.

Chapter 4 takes up procedural concerns and tests theoretical predictions that risk assessments can affect how defendants are evaluated – making some dimensions of defendants' circumstances and background more salient while making others less so – and shape courtroom actors' understandings of the goals and imperatives that should inform the decision task at hand. The results present a mixed bag of findings. They suggest that the adoption of a risk assessment tool does not lead to a large-scale re-orientation of how courtroom actors evaluate defendants, but does cause shifts in some specific domains of their practices. In some, the risk assessment leads courtroom actors to double down on focusing on factors that the risk assessment and past scholarship discourage for being biased or poor predictors. In others, it helps to bring practices in alignment with legal ideals. Overall, the chapter highlights ways in which the adoption of risk assessments can entail

complex trade-offs and have consequences that neither fully match the visions set forth by the proponents nor critics of these tools.

Chapter 5 concludes by highlighting limitations, lingering questions, and areas for future research.

1.2 The Public Safety Assessment

In this dissertation, I examine the impact of the Public Safety Assessment, the most popular pretrial risk assessment in use in the United States. The Public Safety Assessment (PSA) was created with the support of Arnold Ventures,² a private philanthropy that has invested heavily in criminal justice reform initiatives. The tool purports to help reduce over-reliance on cash bail and close racial disparities by helping judges better calibrate decisions to a defendant’s “objective” risk (Arnold Ventures, 2019; Desmarais and Lowder, 2019). Proponents claim that the PSA could help more defendants secure pretrial release on less restrictive conditions without increasing crime or the rate at which defendants fail to appear for court. To date, over 40 jurisdictions across the United States have adopted the tool, including five states (Arizona, Kentucky, New Jersey, Rhode Island, and Utah).

The PSA relies on administrative court records to output two scores on a scale from 1 to 6.³ One reflects the risk that a defendant will fail to appear for court (“Failure to Appear” or FTA) while the other reflects the risk that the defendant will commit another crime if released (“New Criminal Activity” or NCA). The PSA also flags defendants it predicts will commit a violent crime (“New Violent Criminal Activity” or NVCA). The two scores and NVCA flag are determined by subsets of the following nine inputs:

- Age at current arrest
- Whether current charges include a violent offense

²Formerly the Laura and John Arnold Foundation

³Most other risk assessments gather information via defendant interviews. Using administrative records instead is thought to reduce the potential for gaming on the part of arrestees and the influence of staff discretion, while also making reports less costly to produce.

- Whether defendant had any other pending charges at the time of the arrest
- Prior misdemeanor conviction
- Prior felony conviction
- Prior conviction on a violent offense
- Prior failure to appear in the past 2 years
- Prior failure to appear older than 2 years
- Prior sentence to incarceration

These inputs were identified as the most predictive of FTA, NCA, and NVCA in a dataset that pooled cases from over 300 U.S. jurisdictions (DeMichele et al., 2020). Inputs are assigned whole-number weights, which are then summed and scaled to determine the final score. The scoring system, which is shared across all implementation sites, is shown in Figure 1.1. The top panel shows the weights assigned to each input, while the bottom panel illustrates how raw scores are converted into scaled scores.

Each combination of scores is then associated with a particular decision recommendation following locally developed rubrics called Decision-Making Frameworks (DMFs). The DMF for Dane County is presented in Figure 1.2. Low FTA and NCA scores are assigned recommendations for pretrial release without monetary conditions, while higher scores are typically assigned recommendations for increasing amounts of cash bail and more restrictive non-monetary conditions. In many jurisdictions, the rubric that converts scores into recommendations is further governed by local decision rules that automatically recommend more restrictive monetary and non-monetary conditions for certain offenses. I describe Dane County's local decision rules in a later section in this chapter.

Figure 1.1: PSA Inputs, Weights, and Scaling

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

DMF Matrix

	NCA 1	NCA 2	NCA 3	NCA 4	NCA 5	NCA 6
FTA 1	Signature Bond with No Conditions	Signature Bond with No Conditions				
FTA 2	Signature Bond with No Conditions	Signature Bond with No Conditions	Signature Bond with Conditions if Appropriate	Signature Bond with PSL I - Conditions if Appropriate	Modest Cash Bail with PSL II - Conditions if Appropriate	
FTA 3		Signature Bond with Conditions if Appropriate	Signature Bond with Conditions if Appropriate	Signature Bond with PSL I - Conditions if Appropriate	Modest Cash Bail with PSL II - Conditions if Appropriate	Cash Bail, PSL III and Maximum Conditions
FTA 4		Signature Bond with Conditions if Appropriate	Signature Bond with Conditions if Appropriate	Signature Bond with PSL I - Conditions if Appropriate	Modest Cash Bail with PSL II - Conditions if Appropriate	Cash Bail, PSL III and Maximum Conditions
FTA 5		Modest Cash Bail with PSL I - Conditions if Appropriate	Modest Cash Bail with PSL I - Conditions if Appropriate	Modest Cash Bail with PSL II - Conditions if Appropriate	Moderate Cash Bail with PSL III and GPS	Cash Bail, PSL III and Maximum Conditions
FTA 6				Cash Bail, PSL III and Maximum Conditions	Cash Bail, PSL III and Maximum Conditions	Cash Bail, PSL III and Maximum Conditions

Figure 1.2: **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.

The PSA is a strategic object of study for practical and theoretical reasons. From the standpoint of practical impact, the PSA merits attention because it is currently the most widely adopted pretrial risk assessment in the United States. Following years of limited release, Arnold Ventures made the PSA available to all interested jurisdictions in 2018. Given the momentum behind this tool, it is critical to understand how it is used in practice and to what consequence.

Examining the PSA also bridges an important gap in our understanding of the social consequences of algorithms by focusing on a predictive algorithm specifically designed to overcome some of the common critiques lodged against the algorithms. The PSA's design aims to address two central concerns that dominate theoretical and empirical work on pre-

dictive algorithms. The first is that these algorithms may generate biased outputs because they rely on biased data, especially in criminal justice where administrative records encode systemic biases on the basis of race and class (Harcourt, 2007; Reiman and Leighton, 2016; Knox et al., 2020). The second is that risk assessment algorithms can be black boxes – relying on data and procedures that may be hidden for technical or proprietary reasons. This lack of transparency can make algorithms hard to contest and make democratic accountability challenging, if not impossible.

Unlike other risk assessment algorithms used in criminal justice that have been the subject of significant scrutiny, such as COMPAS (Angwin and Larson, 2016), the PSA can claim to exercise discipline over what inputs it incorporates (DeMichele et al., 2020). It excludes arrests as an input, for instance, because they are thought to be too heavily influenced by biased policing patterns.⁴ The PSA further excludes ties to the community, a traditional consideration in assessing flight risk, because they were not predictive. These design choices help proponents of the PSA market the tool as race-neutral and less biased than prevailing alternatives.⁵

The PSA was also designed to be comparatively straightforward and transparent. Risk scores are calculated by summing a limited number of inputs. Courts make no secret of how the PSA works: simple infographics that summarize the algorithm’s inputs, weights, and outputs are often posted around courthouses and on court websites. This stands in contrast to other risk assessment tools in criminal justice that often incorporate many more inputs and withhold information about weights used.

The PSA thus presents distinct *affordances*, perceived or actual properties of an object that determine how it could possibly be used (Norman, 1988), relative to black box algorithms that have been the focus of scholarly attention. The PSA is a fertile site to expand

⁴To be clear, the PSA does still include inputs related to convictions and sentencing, outcomes that are also well-documented to encode racial bias (Steffensmeier et al., 1998). However, proponents of the tool argue that those downstream outcomes are *less* tainted than arrests, which are more directly reflective of racially biased policing and uneven surveillance.

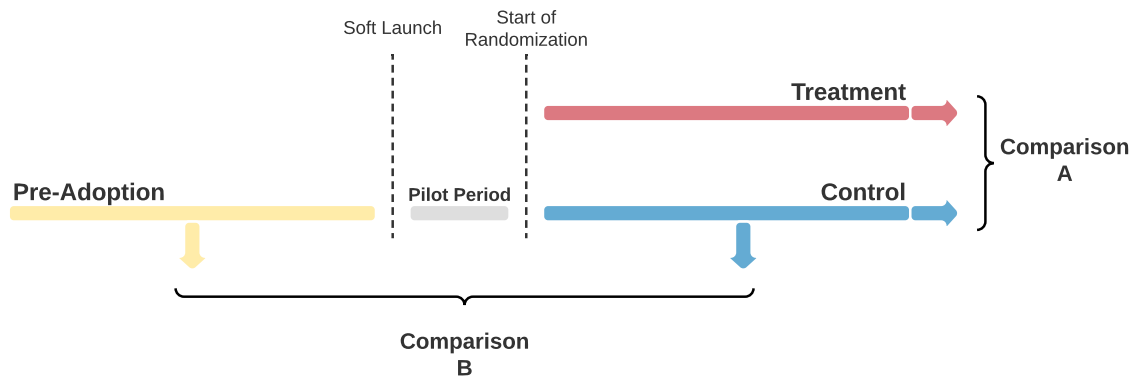
⁵Whether this is actually true is the subject of ongoing research. Scholars emphasize however that claims to colorblindness can obscure racial disproportionality (Van Cleve and Mayes, 2015; Benjamin, 2019)

our understanding of the social processes that different kinds of algorithms may set in motion. The coupling of the PSA’s claims to discipline over inputs and the PSA’s transparency and simplicity, for instance, can influence whether and how courtroom actors mobilize, interrogate, or question the assessment’s numerical evaluations of defendants (Lynch, 2019).

1.3 Study Design Overview

I study how the PSA shapes courtroom practices by examining its impact in Dane County, WI, where the PSA was implemented within the framework of a randomized controlled trial. Directed by the Access to Justice Lab at Harvard Law School, the RCT randomized whether the PSA report for a given arrestee was provided to the court. Even-numbered cases were assigned to the treatment group, while odd-numbered cases were assigned to the control group. PSA risk scores and recommendations were generated for all eligible cases,⁶ but only provided to judges, prosecutors, and defense attorneys in treatment group cases. The RCT was launched on June 1, 2017 and concluded at the end of December 2019.

Figure 1.3: **Study Design Overview**



I leverage this pre-existing randomization infrastructure to collect additional data on how access to the PSA influences courtroom debates and practices.⁷ Due to limitations in

⁶Some classes of cases, like ones involving an extradition, were excluded from the RCT

⁷The original aim of the RCT was to compare the treatment and control groups along a set of aggregate

data availability, I focus on hearings that took place in the first year of the RCT, from June 1, 2017 to May 31, 2018.

The primary data for this dissertation consist of written transcripts of hearings and administrative court records for cases in three groups – pre-adoption, treatment, and control – as depicted in Figure 1.3. I combine administrative data from the Dane County Clerk’s Office and Wisconsin Circuit Court Access, the statewide court record database, with data from the Access to Justice Lab on treatment status, PSA risk scores, and PSA recommendations. I supplement those data with field notes from preliminary fieldwork conducted in January 2020, during which I observed hearings in person and met with court staff and commissioners involved with PSA implementation.⁸

As shown in Figure 1.3, this overall design enables me to make two comparisons. The bulk of the dissertation will focus on Comparison A, in which I compare treated cases (in which the PSA was provided) and control cases (in which the PSA was withheld). The randomized availability of PSA reports creates comparison groups that, on average in the aggregate, should be similar to each other in every respect except whether the PSA report was presented. This feature enables me to isolate the causal effect of providing PSA information on courtroom outcomes.

In some parts of the dissertation, I also present results based on Comparison B, in which I compare hearings that took place six months before any PSA reports were ever generated and hearings in the control group. This comparison enables me to describe how the broader infrastructure that came with adopting the PSA – including staffing shifts, trainings for courtroom actors, and record keeping systems – might shift courtroom practices, apart from the information that individual PSA reports might provide.

Care should be taken in interpreting these latter over-time results, however, because outcomes that the tool is intended to ultimately move, such as the number of days arrested individuals spend incarcerated before trial and the rates at which defendants fail to appear and commit further crimes (Griffin, 2016)

⁸I originally planned to conduct more in-depth fieldwork and formal interviews in spring/summer 2020, but that was canceled due to COVID-19 restrictions.

several factors impede their interpretation as solely attributable to the causal impact of PSA-related infrastructure. A first is the timing of cases in the two comparison groups. Because there was a period during which the PSA was soft launched (between approximately March 28 and May 31, 2017), there is a time gap of about two months between the groups. Given the seasonality of crime, cases in the pre-adoption group, which took place during the winter months and in the control group, which took place over the course of a year, may not be directly comparable. A second factor is that Dane County implemented a variety of other policy changes in this period, such as law enforcement expanding the use of citations instead of arrests, which could alter the number and composition of cases brought to bail hearings (Greiner et al., 2020).

1.3.1 Transcript Sampling and Analytic Strategy

To construct the sample of court hearing transcripts that I examine in this dissertation, I conducted stratified random sampling. I stratified on arrestee race, sex, and judge, and over-sampled Black female and white female arrestees.

For the treatment and control groups, I drew observations from the first year of RCT implementation based on data provided by the Access to Justice Lab. For the pre-adoption period, I used administrative records provided by the Dane County Clerk's Office on all in-custody initial appearances and bail/bond hearings conducted in the six months prior to PSA adoption (all cases where the defendant was arrested, booked into the county jail, and remained in jail at the time of the hearing). This set of observations contained some observations that would not have been eligible for the RCT. To identify observations that likely would have been excluded from the RCT, I trained models to predict RCT inclusion using comparable data from the Clerk's Office covering the RCT period.⁹ The best model, a random forest, achieved an out-of-sample accuracy of 79% (a modest improvement over

⁹Most of the information needed to definitively determine which cases would have been eligible for the RCT had they taken place later resides in a separate data system that I unfortunately did not have access to. Given this significant constraint, the aim here was to approximate as best as possible using the data that I did have access to.

chance given 72% of observations in the test set were included in the RCT). I applied it to observations in the pre-adoption period to create the sampling frame for the group.

The universe that my transcript and administrative data covers differs from that of the broader RCT study in a few ways. First, as already noted, it covers a more limited time period. Second, due to resource limitations, I focus only on cases in which the highest charge was a felony rather than a misdemeanor. I chose to focus on felony cases because the vast majority of misdemeanor cases – around 86% – resulted in signature bonds historically, increasing the possibility of ceiling effects (County of Dane, 2018). Felony cases, by contrast, received signature bonds around 70% of the time. I also exclude a set of cases that were filed under temporary case numbers at the time of randomization due to delays in prosecutors filing formal charges (9% of observations). Under recently implemented rules, the Wisconsin court records system restricts public access to records kept under temporary case numbers, making it difficult to collect case attributes and outcomes. Third, my study and the broader RCT study may be affected by different sets of missing observations depending on when we pulled our data. Record availability is affected both by expungement (individual requests to remove records) and newly implemented sunseting laws that remove records of cases that resulted in a dismissal or acquittal after two years. In my analyses, I drop 5 observations due to missing records. Finally, my transcript sample only covers hearings for Black and white defendants that were conducted by the three judges that handled the vast majority of hearings.

I further note that throughout the dissertation, I treat hearings, rather than individual cases, as the unit of analysis. In a single hearing, judges may set pretrial release conditions in several cases at once if the defendant has multiple pending cases. I measure all decision outcomes – including the use of cash bail and pretrial supervision – by taking the most restrictive condition imposed across the cases processed in a hearing. That is, if a single hearing addresses two ongoing cases and the judge orders cash bail in one case, but a signature bond in another, I consider that hearing to have resulted in cash bail.

Unless otherwise indicated, I estimate the average treatment of access to risk assessment reports on outcomes by calculating differences-in-means between treated and control cases. Analyses based on transcript data are based on weighted differences to account for stratified random sampling.

These analyses rely on the randomization of hearings to the treatment and control groups, which in expectation should generate balanced comparison groups. Table 1.1 compares the treatment and control groups along observed covariates in the administrative data sample, while Table 1.2 compares the characteristics of the pre-adoption, treatment, and control observations in the transcript sample. The tables show that the samples are generally well-balanced along most observed covariates. In the administrative data sample, Failure to Appear and New Criminal Activity scores are moderately higher in the treatment group, but New Violent Criminal Activity values and PSA recommendations are comparatively more balanced. In the transcript sample, there is a similar pattern of moderate imbalance in PSA outputs. Other case-related variables, however, are comparable across the groups.

Table 1.1: **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	700	713	
Defendant race (%)			0.062
Black	303 (43.5)	309 (43.5)	
White	369 (52.9)	371 (52.2)	
American Indian or Alaskan Native	5 (0.7)	5 (0.7)	
Asian/Pacific Islander	8 (1.1)	12 (1.7)	
Hispanic	8 (1.1)	11 (1.5)	
Unknown	4 (0.6)	3 (0.4)	
Defendant sex = Male (%)	555 (79.6)	580 (81.6)	0.049
Defendant age (mean (SD))	33.75 (11.94)	33.47 (11.52)	0.024
Judge (%)			0.057
Judge A	301 (43.0)	319 (44.7)	
Judge B	321 (45.9)	309 (43.3)	
Judge C	69 (9.9)	77 (10.8)	
Other Judges	9 (1.3)	8 (1.1)	
Severity of highest felony charge (%)			0.065
A-C	55 (8.3)	49 (7.2)	
D-E	37 (5.6)	31 (4.6)	
F-G	145 (21.8)	146 (21.5)	
H-I	428 (64.4)	452 (66.7)	
# Cases (mean (SD))	1.17 (0.71)	1.20 (0.54)	0.042
# Unique charges (mean (SD))	2.49 (1.41)	2.57 (1.44)	0.054
# Violent charges (mean (SD))	0.57 (0.96)	0.50 (0.87)	0.067
# Bail jumping charges (mean (SD))	0.53 (1.18)	0.50 (1.24)	0.030
Failure to Appear Score (mean (SD))	2.93 (1.41)	3.08 (1.37)	0.113
New Criminal Activity Score (mean (SD))	3.55 (1.48)	3.71 (1.38)	0.110
New Violent Criminal (%)	135 (19.3)	137 (19.2)	0.002
Activity = True			
Recommendation (%)			0.081
Signature bond	387 (55.3)	391 (54.8)	
Low cash	78 (11.1)	71 (10.0)	
Moderate cash	24 (3.4)	35 (4.9)	
High cash	211 (30.1)	216 (30.3)	

Table 1.2: **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 pre-adoption, control, and treatment samples.

	Pre-adopt	Control	Treatment	SMD
N	174	213	216	
Defendant race (%)				0.015
Black	84 (48.3)	103 (48.4)	102 (47.2)	
White	90 (51.7)	110 (51.6)	114 (52.8)	
Defendant sex = Male (%)	129 (74.1)	158 (74.2)	158 (73.1)	0.016
Defendant age (mean (SD))	32.07 (10.88)	32.87 (11.69)	33.38 (11.69)	0.077
Judge (%)				0.032
Judge A	72 (41.4)	89 (41.8)	89 (41.2)	
Judge B	71 (40.8)	86 (40.4)	85 (39.4)	
Judge C	31 (17.8)	38 (17.8)	42 (19.4)	
Severity of highest felony charge (%)				0.069
A-C	12 (6.9)	15 (7.2)	17 (7.9)	
D-E	13 (7.5)	11 (5.3)	13 (6.0)	
F-G	36 (20.8)	46 (22.0)	46 (21.3)	
H-I	112 (64.7)	137 (65.6)	140 (64.8)	
# Cases (mean (SD))	1.19 (0.69)	1.18 (0.74)	1.22 (0.63)	0.038
# Unique charges (mean (SD))	2.63 (1.48)	2.60 (1.55)	2.75 (1.63)	0.061
# Violent charges (mean (SD))	0.59 (1.15)	0.60 (0.94)	0.55 (0.90)	0.037
# Bail jumping charges (mean (SD))	0.72 (1.31)	0.64 (1.20)	0.65 (1.74)	0.039
Failure to Appear Score (mean (SD))	-	2.97 (1.41)	3.14 (1.34)	0.121
New Criminal Activity Score (mean (SD))	-	3.62 (1.43)	3.75 (1.35)	0.094
New Violent Criminal Activity = True (%)	-	47 (22.1)	43 (19.9)	0.053
Recommendation (%)				0.123
Signature bond	-	114 (53.5)	113 (52.3)	
Low cash	-	23 (10.8)	25 (11.6)	
Moderate cash	-	6 (2.8)	11 (5.1)	
High cash	-	70 (32.9)	67 (31.0)	

For qualitative analyses, I conducted multiple rounds of coding (Charmaz, 2006). In initial rounds of coding, I generated broad codes guided by theory, prior research, and findings that emerged in my quantitative results. I then generated additional codes inductively based on emergent themes. I used a combination of Excel spreadsheets and nVivo to record the codes and identify patterns across hearings with different attributes.

1.4 The Field Site

1.4.1 Demographic and Political Profile of Dane County, WI

Dane County is the second-most populous county in Wisconsin. Just over half the county's population resides in Madison, the county seat and state capital. The county is predominantly white with around 5% of the population identifying as Black or African American and 6% of the population identifying as Hispanic or Latino as of 2019 (US Census Bureau, 2019). Around a half of the county jail population, however, is Black (Austin and Ocker, 2021).

Home to the University of Wisconsin-Madison, Dane County is a strongly Democratic jurisdiction in a purple state. It has a robust history of progressive organizing around criminal justice issues. In recent years, local citizen efforts have encouraged Dane County to institute a suite of reforms that sought to bring the local administration of pretrial justice in line with best practices and reduce racial disparities. The PSA was adopted as one of those reforms.

My conversations with a variety of actors involved with PSA implementation revealed a shared, commonly professed belief that Dane County does things differently than other counties in Wisconsin due to the county's left-leaning political landscape. Court staff and judges suggest that practices in other more conservative parts of the state are more punitive than in Dane County.

1.4.2 Governing Laws

A complex set of local rules, state statutes, case law at the state and federal levels, and the U.S. Constitution govern the administration of pretrial justice in Dane County. Rather than review these comprehensively, I highlight in this section only the key elements most pertinent to understanding the local context and interpreting this dissertation's findings.

Wisconsin state statutes primarily offer judges two options: impose cash bail (requiring defendants to post cash to be released) or grant a signature bond (allowing the defendant to be released without depositing cash with the court).¹⁰ Statewide, judges are legally instructed to operate on a presumption of release. That is, judges are to first consider release without monetary conditions as the default. Cash bail should only be imposed if and in the amount deemed necessary to assure future appearance in court. Crucially, this means that judges are not supposed to take public safety concerns into account in setting monetary conditions.

Non-monetary release conditions, meanwhile, may be imposed with public safety considerations in mind. Non-monetary release conditions can encompass an array of different behavioral expectations, such as that defendants not consume alcohol or contact alleged victims or co-defendants, and ongoing pretrial monitoring. Judges may order non-monetary conditions to protect the community from serious bodily harm or to prevent witness intimidation. By statute, all defendants are subject to the conditions that they appear for court proceedings, not intimidate witness, and not commit any crimes when released. In Wisconsin, violations of release conditions can constitute crimes and result in bail jumping charges.

The statutes further offer a long list of considerations that may be incorporated into decisions about monetary and non-monetary release conditions:

“the ability of the arrested person to give bail, the nature, number and gravity of

¹⁰Unlike in some states, judges cannot practically detain defendants prior to conviction without the possibility of pretrial release

the offenses and the potential penalty the defendant faces, whether the alleged acts were violent in nature, the defendant's prior record of criminal convictions and delinquency adjudications, if any, the character, health, residence and reputation of the defendant, the character and strength of the evidence which has been presented to the judge, whether the defendant is currently on probation, extended supervision or parole, whether the defendant is already on bail or subject to other release conditions in other pending cases, whether the defendant has been bound over for trial after a preliminary examination, whether the defendant has in the past forfeited bail or violated a condition of release or was a fugitive from justice at the time of arrest, and the policy against unnecessary detention of the defendant's pending trial." (Wisconsin Statutes §969.01(4) 2019)

Critical to note, however, is that these laws only reflect law in the books, not law in action. Practices vary widely throughout the state and often depart from these dictates (Lavigne et al., 2018). The instruction, for instance, that cash bail should only be used to assure court appearance, rather than protect public safety, is often circumvented. Prosecutors and judges, for instance, commonly accept arguments that people who are charged with serious violent crimes face significant potential incarceration, creating incentives to flee. In this way, conditions motivated by concerns about dangerousness or public safety can be repackaged as being motivated by concerns about flight risk.

1.4.3 Local Procedures and PSA Implementation

In Dane County, the PSA is used in Initial Appearances and Bail/Bond Hearings. The judicial officers that typically conduct these hearings are court commissioners. While locally only the judicial officers that handle later proceedings are officially called "judges," I use the terms "commissioner" and "judge" interchangeably throughout this dissertation.

Initial Appearances are typically the first hearings that place after an individual is arrested and a new case is initiated. They are required to occur within 48 hours of arrest. At Initial Appearances, commissioners inform arrested individuals of the charges against them, determine if there was probable cause to support the arrest, and set pretrial release conditions. Bail/Bond Hearings, meanwhile, take place when an event in an existing case triggers a review of bail conditions, such when a defendant is arrested on a warrant following a missed court appearance or if defendants exercise their right to request a re-examination of their conditions.

The PSA is only used in Initial Appearances where the defendant is still in custody at the time of the hearing. Some arrestees are released prior to the Initial Appearance. Notably, individuals charged with most misdemeanors may be released prior to the Initial Appearance if they can post the bail set in the statewide Uniform Bond Schedule, which outlines the bail amounts associated with specific misdemeanor charges. The pool of arrestees who must appear before a judge and be subject to the PSA is thus primarily made up of three groups: those charged with misdemeanors who cannot post the standard set amounts, those charged with certain serious misdemeanors, and those charged with felonies.

Figure 1.4: Courtroom in the Dane County Public Safety Building. Photo by Emily Hamer/Wisconsin State Journal



Hearings take place in a courtroom in the county jail building, as depicted in Figure

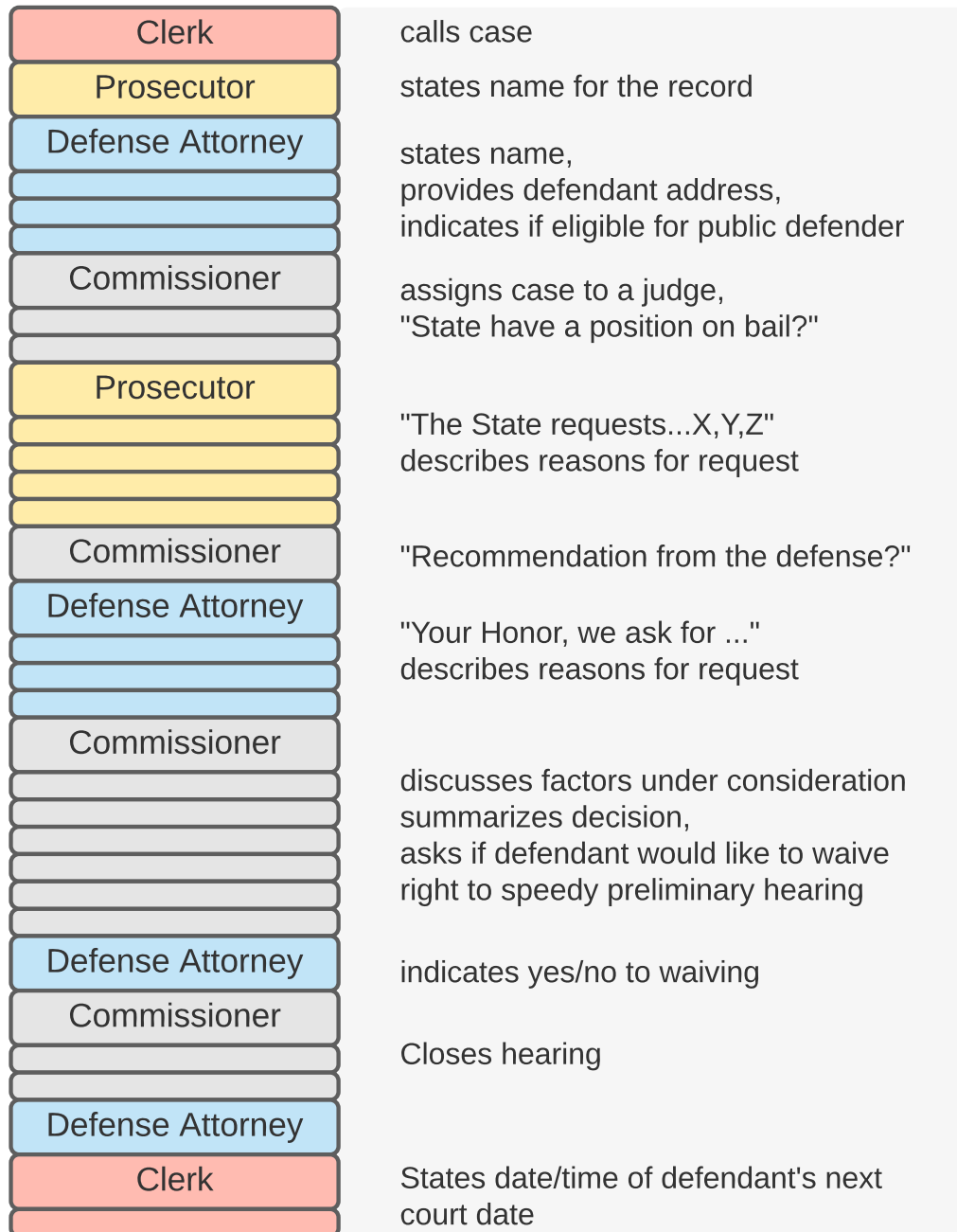
1.4.¹¹ The proceedings resemble an assembly line of justice (Kohler-Hausmann, 2013; Van Cleve, 2016), with fifteen to forty cases processed back-to-back on weekday afternoons. Hearings usually last around four to eight minutes, though some complex cases can take much longer. Arrested individuals are ushered into the cramped courtroom one-by-one. Most are represented by the public defender assigned to handle cases for the day. They are generally discouraged from speaking to decrease the chance that they say something that could be used against them. As a result, most hearings involve discussions among the prosecutor, defense attorney, and commissioner.

While not all initial appearance/bond hearings have the exact same structure, they typically follow the order of events depicted in Figure 1.5. Figure 1.5 presents 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 – which I identified by conducting a sequence analysis of sentences in transcripts (Abbott and Tsay, 2000; Gabadinho et al., 2011).¹² I have annotated what was discussed in each hearing segment.

¹¹In some jurisdictions, bail decisions are made without a formal hearing. In Kentucky, for instance, they are made via a telephone call between a judge and a pretrial services officer (Albright, 2019).

¹²I used optimal matching (Abbott and Tsay, 2000) here, but other approaches to measuring distances between sequences returned similar centro-types.

Figure 1.5: **Structure of a Typical Hearing in Dane County.** Each box represents a sentence.



At the beginning of most hearings, commissioners begin by quickly reviewing the criminal complaint, a document which summarizes the date and nature of the alleged offenses. In interviews, commissioners indicated that they usually have little or no time before hearings to read over materials pertaining to the day's docket. Consequently, pretrial decision-making usually happens live during hearings. Commissioners then provide the prosecutor and the defense attorney an opportunity to make arguments advocating for particular monetary and non-monetary pretrial release conditions. Commissioners might follow up with additional questions for the two sides or proceed directly to delivering their decision.

In Dane County, PSA reports are prepared by two pretrial assessors who draw on available administrative records to enter PSA inputs. They are usually prepared, printed out, and distributed shortly before the day's hearings to the commissioner, prosecutor, and defense attorneys working that day. As I discuss further in the next chapter, this typically means that the PSA is not prepared in time for courtroom actors to review beforehand and limits its impact on what prosecutors and defense attorneys request.

Figure 1.6 presents an example of a PSA report in Dane County. The report displays the defendant's PSA scores and summarizes the defendant's values for key PSA inputs. PSA recommendations for monetary and non-monetary pretrial release conditions are printed at the bottom.

Figure 1.6: **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]
DOB: [REDACTED]
Arrest Date: 03/25/2017

Spillman Name Number: [REDACTED]
Gender: Male
PSA Completion Date: 03/27/2017

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 -

1.4.4 PSA Recommendations: Local Decision Rules and Distribution Relative to Baseline

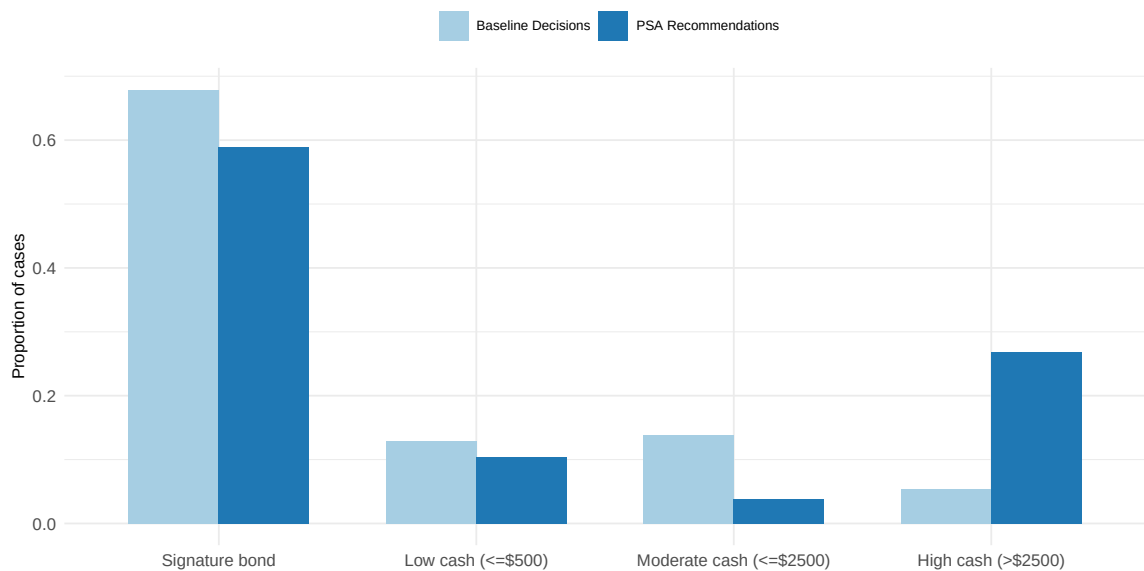
As noted earlier, the PSA outputs recommendations according to locally-developed rubrics, which deterministically map PSA scores to recommendations, and decision rules. The rubric for Dane County was shown in Figure 1.2.

Similar to other jurisdictions (Lum et al., 2020), Dane County instituted two local decision rules. First are rules that supersede the rubric and 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.

Second are 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.

Importantly, these local decision rules, which impact around 17% of recommendations, combine with the local decision rubric to recommend decisions that are systematically more restrictive than local judges were making at baseline before the PSA was adopted. As Figure 1.7 depicts, across cases, the PSA recommended high cash bail at substantially higher rates within the first year of implementation than judges were ordering cash bail before the PSA was adopted.

Figure 1.7: Distribution of PSA recommendations relative to distribution of decisions in six months prior to initial PSA implementation



Consequently, we can think of the tool as systematically offering information to courts that suggests that arrested individuals, on the whole, are more “risky” than previously thought.

Chapter 2

Asymmetries in the Use of Pretrial Risk Assessments and Their Impact on Cash Bail and Procedural Outcomes

2.1 Decision-making in the Courtroom Workgroup

To date, research on risk assessments in criminal courts has primarily understood tools like the PSA as interventions that target a sole actor: judges. Sharing the focus of past streams of scholarship on policy interventions in criminal courts, such as sentencing guidelines, researchers have aimed to interrogate how reforms shape judicial discretion and evaluate how the predictive capacity of judges and risk assessments compare (Brayne and Christin, 2020; Goel et al., 2021).

But judges do not make decisions in a vacuum. Instead, they make decisions embedded within concrete social, cultural, and organizational contexts that shape how defendants are understood and enable and constrain different lines of action. Most immediately, bail conditions are typically set through formal, interactive proceedings. In these proceedings, prosecutors and defense attorneys advocate for desired decision outcomes. They assem-

ble information, offer interpretations, and construct arguments to support their positions, guided by the unique aims and responsibilities of their positions. Prosecutors aim to protect public safety, while defense attorneys strive to advocate for the interests of their clients and safeguard their rights. This can lead at times to agreements and at times to sharp disagreements between the parties. Judges, meanwhile, respond to the arguments presented by each side and aim to reach decisions that might be understood as fair and impartial.

Thus, as decisions emerge through interactions between prosecutors, defense attorneys, and judges, it is important to analyze risk assessment algorithms not as a technology that targets a lone actor, but as one that affects the dynamics between multiple strategic actors.

What is the nature of that dynamic and how might risk assessments disrupt or reinforce it? Scholars of criminal courts have observed that key criminal court actors often interact in what has been labeled as a shared “courtroom workgroup,” in which relationships are marked less by adversarial clashes than by interdependence and coordination (Eisenstein and Jacob, 1977). Scholars suggest that, especially in high-volume and low-information settings such as in pretrial decision-making, courtroom actors develop shared routines, shortcuts, and vocabularies to resolve classic organizational concerns like reducing uncertainty and increasing efficiency. Viewed from the vantage point of organizational sociology, criminal courts present processual orders comprised of standard templates for argumentation, “going rates” for different types of cases, and norms around how cases are to be processed (Ulmer, 2019). These orders emerge over the repeated interactions and may delicately balance multiple competing institutional logics (McPherson and Sauder, 2013) variously advanced by different courtroom participants.

When a technological or policy intervention like the PSA is adopted, it may disrupt, amend, or be absorbed by these existing social arrangements. This presents a central question: how the PSA affects the balance of power and influence among courtroom actors. What parties does it support and whose position does it strengthen? This question is critical for unpacking the impact of risk assessments. It follows the lessons of scholarship

on sentencing guidelines which found, for example, that the failure of the guidelines to improve criminal justice outcomes was attributable to the ways in which the guidelines unintentionally empowered and enlarged the discretion of prosecutors (Tonry, 1996; Stith and Cabranes, 1998).

A similar possibility may exist in the case of risk assessment recommendations. While recommendations may ostensibly support either prosecutors or defense attorneys, the fact that they are only intended to be advisory creates possibilities that they may be differentially effective when mobilized for different positions. As discussed further in the next section, risk assessments advance a particular institutional logic – one built around risk management – that may upset the working balance among competing institutional logics in a local court, and tip the scales in favor of greater punitiveness and greater suspicion.

Risk assessments may, moreover, reduce how dependent judges are on prosecutors and defense attorneys, who otherwise serve as the primary parties who shape how a decision appears before a judge. Risk assessments present a seemingly independent source of information, which may disagree with both defense attorneys and prosecutors. Access to a risk assessment may thus make it easier to legitimate decisions that depart from the positions of prosecutors and defense attorneys. In this paper, I examine to what extent this may be the case.

2.2 Do Risk Assessments Reduce or Spread Suspicion and Restriction?

A critical question swirling around risk assessments in general and the PSA in particular is whether these tools ultimately help to reduce or increase unnecessary detention and restriction. Advocates of the tools have argued that the tools, on net, can help to curb overestimation of defendant risk. Some existing research supports this position, finding that courts are overly conservative in pretrial decision-making, and that risk assessments, if fol-

lowed, could facilitate the release of more individuals without increasing crime (Kleinberg et al., 2018).

But for such an outcome to be realized, courtroom actors must in practice recalibrate their risk perceptions in line with the tool when it deems defendants low risk. Critics and other scholars argue that this is unlikely to happen as proponents intend. Instead they suggest that recommendations are likely to be taken asymmetrically: defendants are unlikely to “get credit” for low risk scores, but are likely to be regarded with greater suspicion for high risk scores. Scholars have theorized, for example, that risk assessments make it functionally impossible for a defendant to be regarded as entailing no risk (Werth, 2019). Everyone is seen as a vector for bad outcomes, and no one is entirely free of taint. Critics have further argued against the tools on the basis that they formalize stigmatizing labels like “high risk” and score people based on negative outcomes – failures to appear and recidivism – rather than on successes (Robinson and Koepke, 2019). The latter can be particularly damaging because it spreads the impression that the predicted negative outcomes are common, when in practice the majority of defendants labeled “high risk” by these tools never miss court or commit crimes while on pretrial release (Barabas et al., 2019).

These arguments against risk assessments have to date largely not been investigated empirically. In this paper, I test whether the impact of the tools matches these concerns. To the extent that they do, they further shed light on whether and how risk assessments may reshape the relative positions of prosecutors, defenders, and judges. If the proponents of risk assessments are correct, we might expect to see more decisions that favor defense attorney requests for less stringent pretrial release conditions. To the extent that critics are correct, we might expect to see the PSA exert an asymmetric influence, with only PSA recommendations for more stringent conditions having an effect on outcomes.

2.3 Approach and Methods

In this paper, I analyze how access to the PSA shapes the relative positions of prosecutors, defense attorneys, and commissioners in decisions about whether a defendant is required to post cash bail to be released pretrial. I investigate to what extent PSA recommendations affect the rates at which prosecutors and defense attorneys secure their requested bail outcomes. Within this, I also examine whether access to PSA reports reduce judges' dependence on prosecutor and defender requests. To the extent that is the case, we would expect to observe judges departing from consensus requests from prosecutors and defense attorneys at higher rates. Such a result would be consistent with the PSA creating additional possibilities for judges to decide in a manner that is less constrained by the terms of debate that the prosecution and defense established.

I use data derived from court transcripts. I manually recorded the initial requests of prosecutors and defense attorneys in all treatment and control hearings in my sample.¹ I also recorded judges' final decisions.

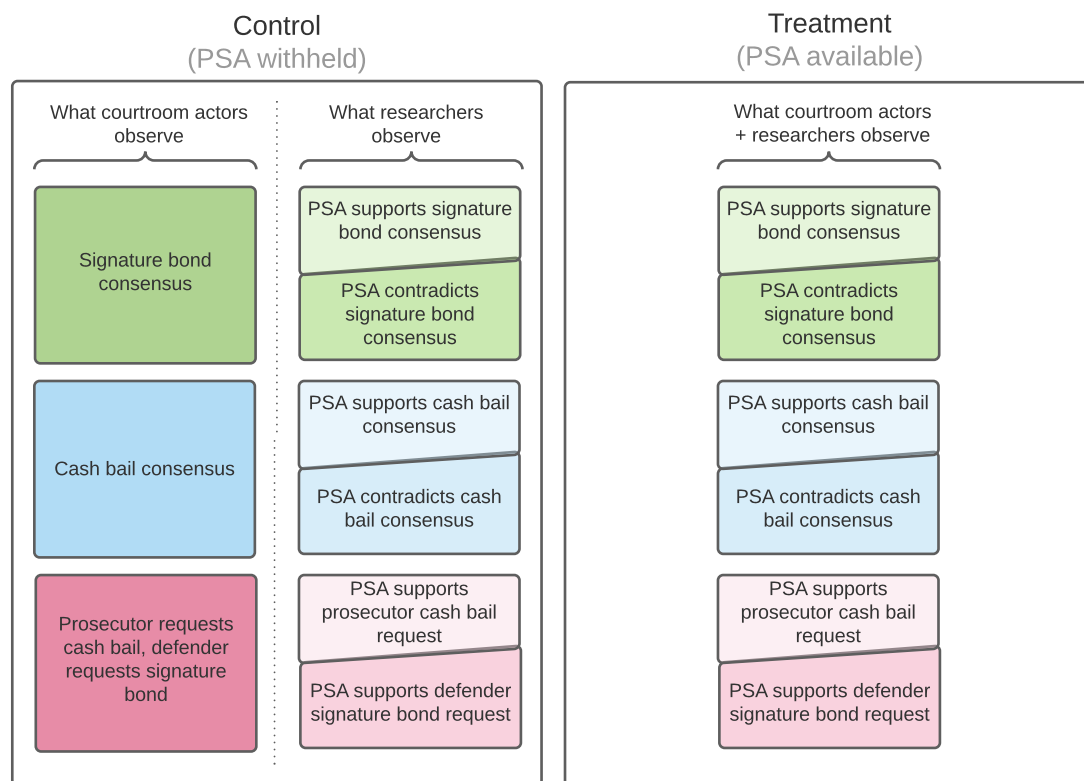
The results will proceed in two main parts. The first summarizes how often prosecutors, defense attorneys, and judges reference the PSA. For each actor, I present examples of how the PSA figures explicitly into each actor's arguments. The second half of the analysis analyzes the effect of the PSA on whether cash bail is ordered given different combinations of prosecutor requests, defense attorney requests, and PSA recommendations.

The focal analyses center on investigating how the use and effect of the PSA is moderated by how PSA recommendations interact with prosecutor and defense attorney requests. As illustrated in the treatment group panel in Figure 2.1, the PSA creates possibilities for new combinations of positions with respect to the use of cash bail. Depending on what

¹In a small subset of hearings, prosecutors and defense attorneys amended their requests over the course of the hearing in light of new information (such as a victim giving a statement) or when judges hinted that their initial requests were clearly beyond the bounds of what they would consider reasonable. I chose, however, to analyze the initial requests for the sake of consistency and as they reflected the starting positions of each party.

the PSA recommends, it can support the consensus request of the prosecution and defense, disagree with the consensus, or side with one of the two sides in a disagreement. Since defense attorneys never request anything more stringent than do prosecutors, six different combinations of prosecutor requests, defense attorney requests, and PSA recommendations are possible.

Figure 2.1: Possible Relationships between PSA Recommendation, Prosecutor Requests, and Defense Attorney Requests



In the treatment group, courtroom actors experience all six of these combinations. In the control group, courtroom actors only experience the positions of prosecutors and defense attorneys and whether they agree. Researchers, however, can identify which of the six combinations the PSA would have created were it to have been provided to the court.

These latent combinations serve as comparison groups to evaluate how access to PSA reports affects bail outcomes. It is worth clarifying, however, that in each latent group, what

the effect of PSA report represents differs. To aid with interpretation, Table 2.1 summarizes how the effect of the PSA may be understood for each subgroup.

Table 2.1: How to interpret the effect of providing vs withholding PSA reports under different combinations of prosecutor requests, defense attorney requests, and PSA recommendations

Group	Request/Recommendation			What effect of providing vs withholding the PSA represents Difference in the whether the judge in hearing <i>i</i> would order cash bail if...
	Prosecutor	Defender	PSA	
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 prosecutor cash bail request	Cash bail	Signature bond	Cash bail	a prosecutor 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 defender signature bond request	Cash bail	Signature bond	Signature bond	a defender request for signature bond that the prosecutor opposes is advanced with the support of a PSA signature bond recommendation vs if it is advanced without PSA support

The validity of these comparisons relies on a strong assumption: that seeing the PSA recommendation does not change what the prosecution or defense request. This assumption is plausible given the way the PSA was implemented in Dane County. The court staff who prepared and distributed PSA reports usually only printed them right before the afternoon's hearings, leaving almost no time for review. Prosecutors and defense attorneys typically

prepared their positions and arguments ahead of time without the input of the PSA.² While prosecutors and defense attorneys regularly wove the PSA reports on the fly into their arguments, this assumption requires that *what* they requested, whether a signature bond or cash bail, does not shift in light of the PSA report. Later in the paper, I present descriptive statistics consistent with this assumption being met, as well as an additional robustness test and sensitivity analyses that consider how the results might differ if this assumption was violated.

2.4 Descriptive Statistics

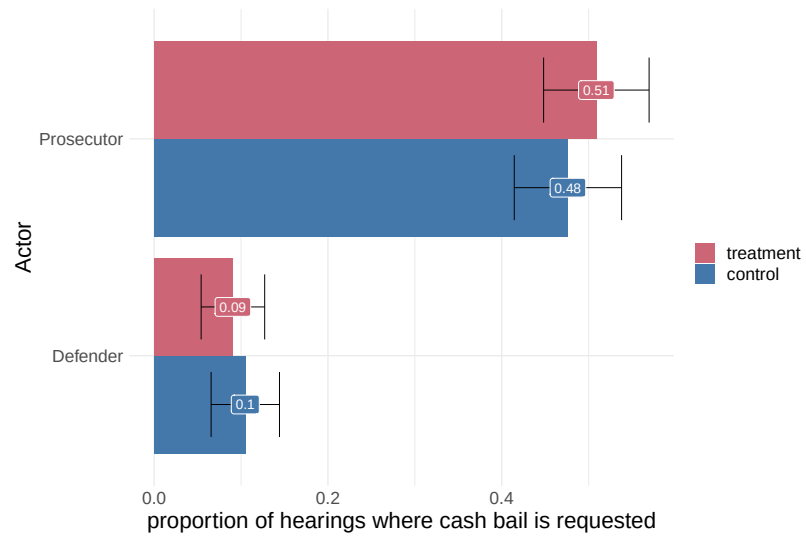
To ground the results, Figure 2.2 displays how often prosecutors and defense attorneys request cash bail rather than a signature bond, with and without the PSA. Panel A indicates that prosecutors request cash bail in around half of cases. Defense attorneys, meanwhile, request cash bail in only around 10 percent of cases. In the few cases where defense attorneys request cash bail, they request lower bail amounts than proposed by prosecutors, or decide not to contest prosecutor cash bail recommendations because the defendant will be detained for other criminal justice matters regardless of the bail outcome. Consistent with the PSA not influencing what courtroom actors request, there are no statistically significant differences in requests across treated and control hearings.

Figure 2.2, Panel B further disaggregates prosecutor and defender requests by PSA recommendation. It indicates that prosecutors and defense attorneys alike more often request cash bail in cases where the PSA recommends cash bail, regardless of whether that recommendation itself was supplied to the court. Prosecutors ask for cash bail in around 68-72 percent of hearings with a PSA cash bail recommendation and only around 30-33 percent of hearings with a signature bond recommendation. Defenders request cash bail in around 16 to 19 percent of hearings with a cash bail recommendation, and in only 3 to 4 percent of

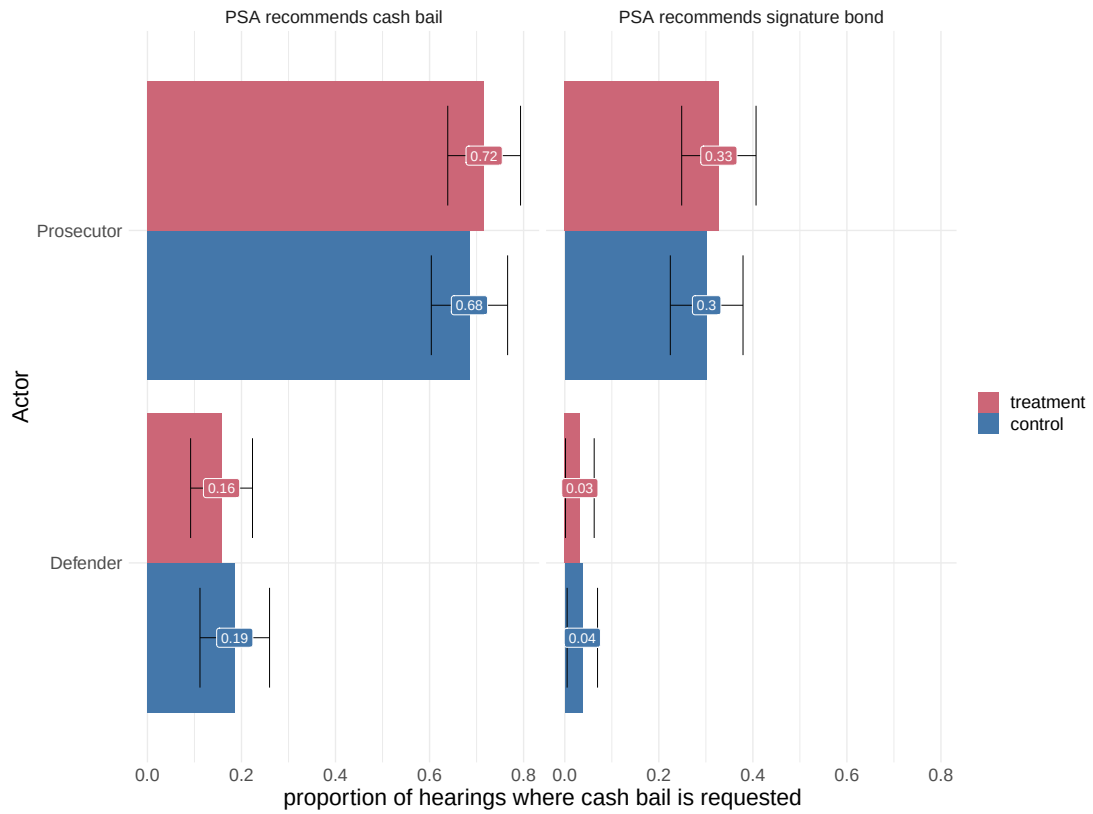
²This was more systematic in the district attorney's office, where Assistant District Attorneys were directly tasked with coming up with prosecution requests ahead of the hearings each day. The prosecutor in the courtroom usually read out these pre-set requests.

hearings with a signature bond recommendation. Again, however, there are no statistically significant differences in requests across treated and control hearings. Broadly, this indicates that the PSA recommendations *align* at least in part with the subjective evaluations of prosecutors and defenders of what an appropriate bond outcome may be. However, there is not evidence to conclude that the PSA recommendations *influence* what prosecutors and defenders request.

Figure 2.2: Cash bail request rates



(a) By actor



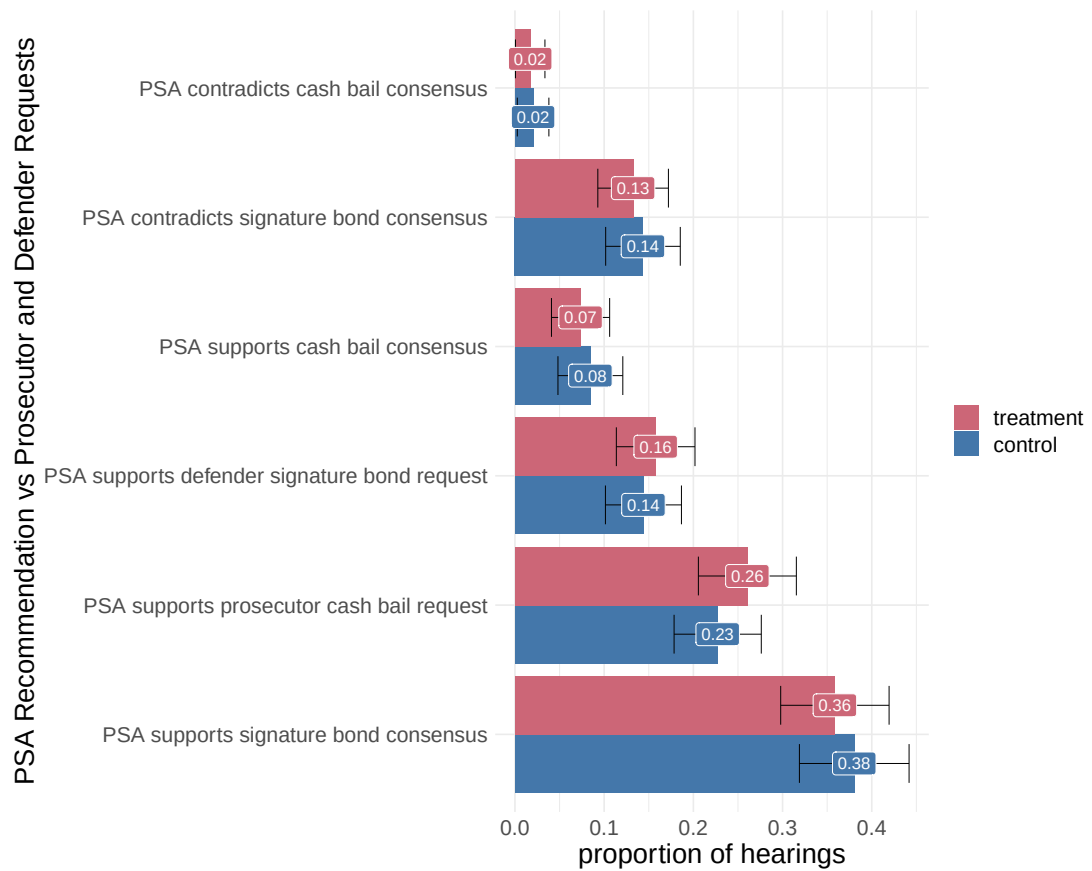
(b) By actor and by PSA recommendation

Figure 2.3 displays the prevalence of different combinations of relationships between the PSA recommendation, the prosecutor request, and the defense attorney request in a case. The PSA recommendation supports consensus requests for signature bonds in around 36-38% of cases and consensus requests for cash bail in around 7-8% of cases. In 14-16% cases, the PSA recommendation supports defenders' requests for a signature bond in the face of prosecutor requests for cash bail. In another 23-26% of cases, the PSA recommendation aligns with prosecutor requests for cash bail while contradicting defender requests for a signature bond.

Here it is valuable to highlight that Figure 2.3 indicates that, on the whole, the PSA recommendations align with the requests of prosecutors more often than they do with the requests of defenders. While the PSA aligns with the two sides equally when they request the same outcome, when there is a disagreement, the PSA sides with prosecutors at a higher rate than it does defenders.

The PSA recommendation additionally injects disagreement where there was none previously in 13-14% of cases by recommending cash bail despite joint requests from the prosecutor and defender for a signature bond. More rarely, in 2% of cases, the PSA adds disagreement by recommending a signature bond while the prosecutor and defender request cash bail. Given the rarity of this last set of cases (represented by only 3 observations in the treatment group and 4 in the control group), I will not formally test the impact of the PSA in those settings in this paper.

Figure 2.3: **Distribution of Relationship between PSA Recommendation, Prosecutor Requests, and Defense Attorney Requests**



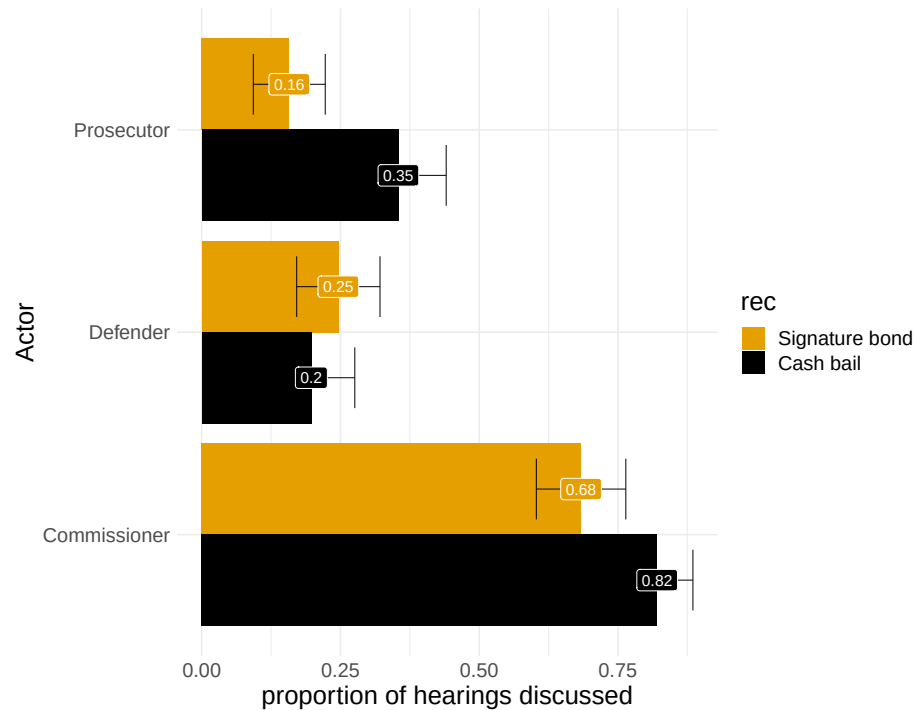
2.5 Results

2.5.1 Discussions of the Public Safety Assessment

To first examine how PSA reports intervene in the courtroom when they are supplied, it is instructive to examine how often, under what conditions, and in what manner different courtroom actors reference the PSA.

Figure 2.4 presents the rates at which prosecutors, defense attorneys, and commissioners invoke the PSA when it is available, broken out by what the PSA recommended. Generally it indicates that the PSA is far from ignored, with all actors taking up the tool explicitly in hearings to some extent. It indicates that commissioners reference the PSA substantially more often than do defenders and prosecutors, and that they reference the PSA more often when it recommends a cash bail than when it recommends a signature bond. Prosecutors similarly reference the PSA more often when there is a cash bail recommendation: they reference the PSA in 35% of such cases, compared to in only 16% of cases with a signature bond recommendation. Defenders, meanwhile, reference the PSA at more comparable rates regardless of the recommendation. These results broadly suggest that PSA cash bail recommendations are mobilized more often and discussed more frequently than are PSA signature bond recommendations, which are more likely to go unacknowledged.

Figure 2.4: **Rates of PSA Discussion in Treatment Group by Actor and by PSA Recommendation**

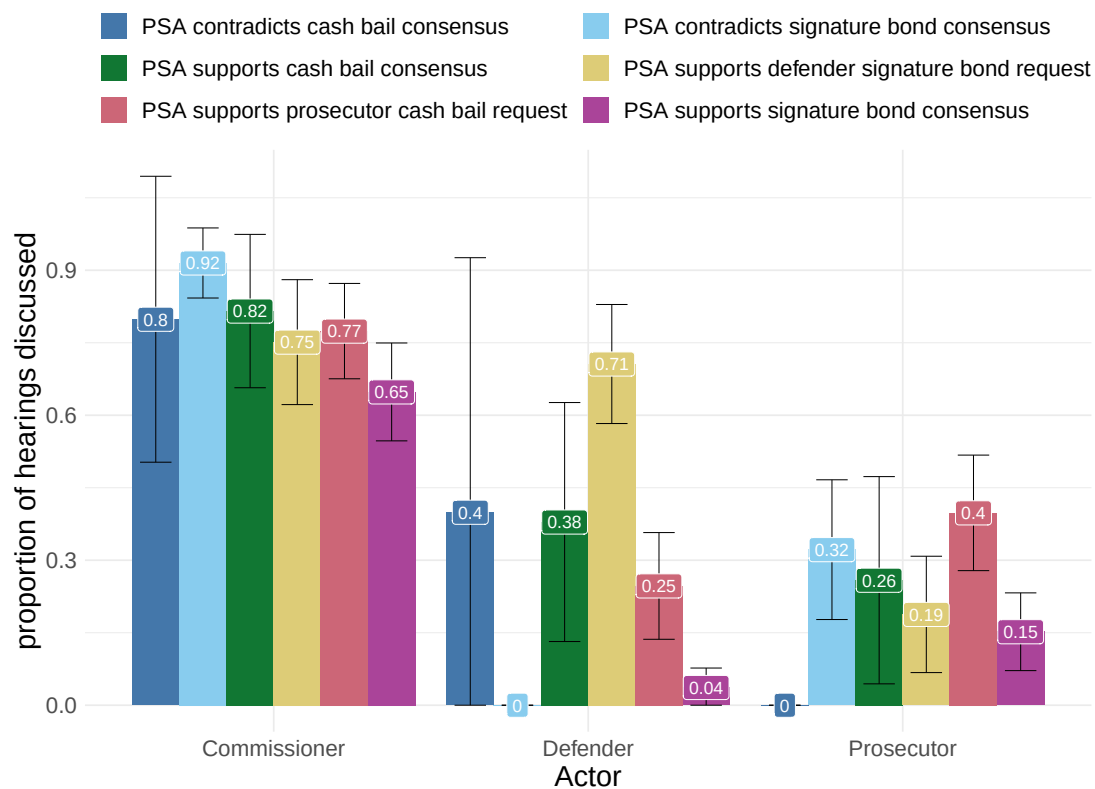


Next, Figure 2.5 presents the rates at which prosecutors, defense attorneys, and commissioners invoke the PSA when it is available, broken out by how the PSA recommendation compared with the initial requests of prosecutors and defense attorneys. It again shows that commissioners reference the PSA more often than do defenders and prosecutors on the whole. Several additional patterns are worth noting which are broadly consistent with prosecutors and defense attorneys mobilizing the PSA strategically when it strengthens their positions. First, divergence across actors in how often they reference the PSA appears to be greatest when the PSA recommends cash bail while there's a joint request from the prosecutor and defense attorney for a signature bond. In these cases, judges reference the PSA 92% of the time and prosecutors 32% of the time. Defenders, meanwhile, never once referenced the PSA in these cases.

Second, when there is disagreement between prosecutors and defense attorneys, judges reference the PSA at comparable rates regardless of which side the PSA agrees with. De-

fenders and prosecutors, meanwhile, both invoke the PSA more often when the PSA recommendation aligns with their position. Defenders, however, reference PSA signature bond recommendations that align with their requests at a considerably higher rate (71% of cases) than prosecutors invoke PSA recommendations that align with their cash bail requests (40% of cases). Indeed, defense attorney reliance on the PSA is stronger when there is disagreement than in any other setting.

Figure 2.5: Rates of PSA Discussion in Treatment Group by Relationship between PSA Recommendation, Prosecutor Requests, and Defense Attorney Requests



How the PSA is invoked by prosecutors and defense attorneys

Examining *how* prosecutors and defenders invoke or discuss the PSA reveals further differences. When prosecutor requests match the PSA, the PSA assumes a minimal supporting role in prosecutors' arguments. In the vast majority of such hearings, prosecutors mention the PSA only briefly by noting that their requests are "consistent with," "follow,"

or “agree with” the PSA’s recommendation for whether to grant a signature bond or require cash bail. They also preface their requests with “given the PSA recommendation” or conclude their requests by reiterating “the PSA does recommend a [signature bond/cash bail].” In only a small minority of hearings do prosecutors note specific elements of the PSA report or read off the defendant’s risk scores.

When defender requests match the PSA, by contrast, defenders tend to incorporate the PSA more centrally in their arguments. Beyond merely noting the PSA recommendation and how their request aligns with it, defenders also draw attention to other elements of PSA reports, such as their defendant’s printed record of making court appearances (e.g. “This assessment did not note any failures to appear within the last 2 years or failure to appear older than 2 years”) and underscore favorable scores (e.g. “He scores a one on the failure to appear score”). Importantly, defense attorneys further make statements to encourage judges to follow PSA signature bond recommendations and place authority in the tool. Defenders assure judges “you can rely on the PSA,” opine that the PSA “correctly weighs” the defendant’s risk, state that the PSA’s recommendation is “appropriate,” and explicitly ask them to follow PSA recommendations. In the whole sample, there was only one instance where the defense, represented by a private attorney who infrequently participates in initial appearances, directly questioned the PSA.

When the PSA recommendation disagrees with their requests, prosecutors and defense attorneys also discuss the PSA in different ways. When prosecutors bring up the PSA in these cases, they typically acknowledge that their requests are upward departures (requesting cash bail despite a signature bond recommendation) or downward departures (requesting signature bonds despite a cash bail recommendation). Sometimes they provide reasons for the departure. For downward departures, prosecutors give reasons like the defendant lacking a history of missing court or that the defendant is going to be under other forms of supervision, thereby reducing risk. For upward departures, prosecutors tend to highlight the severity of the charges or the defendant violating court orders in the past. It is important

to note, however, that these acknowledgments are fairly perfunctory and brief, and often take place after prompting from the judge to consider the PSA.

When defenders bring up the PSA in cases where it recommends something more stringent than they are requesting, they typically pursue one of two approaches. In the first, fairly rare approach, defenders briefly acknowledge the discrepancy without directly arguing against it (“given this record, in spite of the Public Safety Assessment, we believe that the signature bond is appropriate”). In the second more common approach, defense attorneys pick out parts of the PSA report that are favorable to their client, without drawing attention to the fact that the PSA actually recommends more stringent conditions than they are requesting. Defenders highlight the PSA reporting no missed appearances (“The PSA says zero for prior failures to appear in the last 2 years, so his recent history does not show any failures to appear”) or predicting low risk (“Going through the risk factors, it appears that he’s low risk on everything”).

In sum, these results indicate that when PSA usage tilts in favor of one side’s request over the other, defense attorneys tend to engage more often and more deeply with the PSA report than do prosecutors.

2.5.2 Effect of PSA on Hearing Outcomes

Use of cash bail

Figure 2.6 displays the estimated effect of the PSA on the use of the cash bail across all hearings in the administrative data sample (top panel) and by PSA recommendation (bottom panel). The figure suggests that access to PSA recommendations has an asymmetric effect. Among cases where the PSA would recommend cash bail, access to the PSA report increases the use of cash bail by an estimated 13 percentage points. Among cases where the PSA would recommend a signature bond, meanwhile, there is insufficient evidence to conclude that access to the PSA report has an effect. While the 95% confidence intervals

for both estimates are large, 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 offset the positive effect of PSA cash bail recommendations on the use of cash bail.

Figure 2.6: Effect of PSA on Use of Cash Bail, Overall

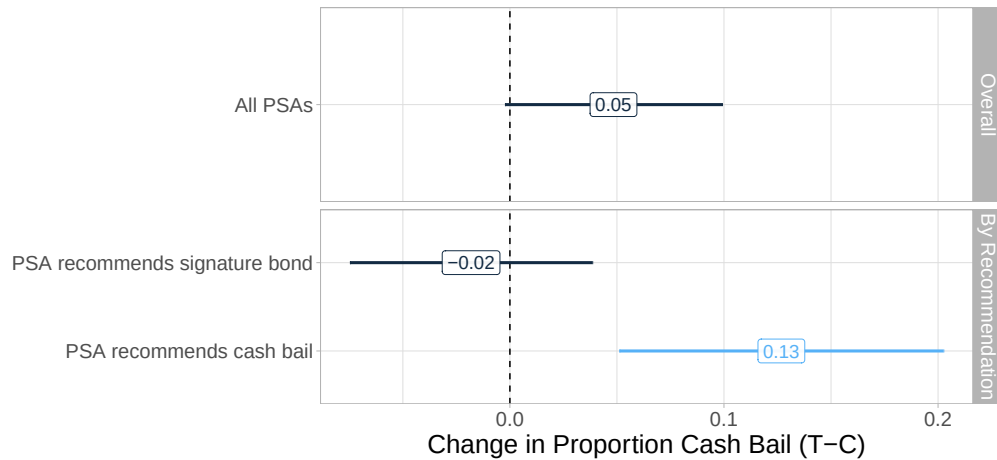


Figure 2.7: Effect of PSA on Use of Cash Bail, by Relationship with Prosecutor and Defender Requests



To disaggregate this further, Figure 2.7 draws on data from the transcript sample and presents estimates for the effect of the PSA for different subgroups of hearings defined by how the PSA recommendation compares to prosecutor and defense attorney requests. When the PSA recommendation agrees with the consensus requests of prosecutors and defenders, providing the report does not appear to have a substantial impact on the use of cash bail. However, when the PSA recommendation intervenes in a hearing where there is disagreement or when it introduces fresh disagreement, it has a larger impact.

It increases the use of cash bail in two settings. First, in cases where prosecutors and defenders jointly recommended a signature bond but the PSA recommended cash bail, providing the PSA report increased the use of cash bail by 22 percentage points (from 4% to 26%) relative to withholding the PSA report. Second, in cases where prosecutors and defenders disagreed, the PSA appears to bolster the arguments of prosecutors requesting cash bail when it recommends cash bail. In such cases, cash bail orders increased by 16 percentage points when the PSA was supplied (from 72% to 88%).

By contrast, there is insufficient evidence to indicate that there is an offsetting effect when the PSA aligns instead with a defender's request for a signature bond. Analyzing the few observations where the PSA recommended a signature bond while there was a consensus request for cash bail from the prosecutor and defender similarly suggests that signature bond recommendations are less influential: cash bail was ordered in all treatment and control hearings.

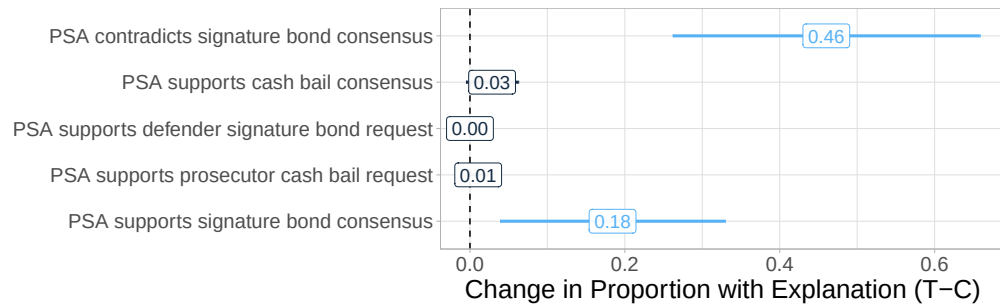
Together, these results indicate that PSA cash bail recommendations are more impactful than signature bond recommendations when there is disagreement among the PSA recommendation, the prosecutor's request, and the defense attorney's request. This asymmetric impact is consistent with the PSA encouraging judges to view defendants as more risky, on average, when the PSA is available than when it is not.

Procedural outcomes

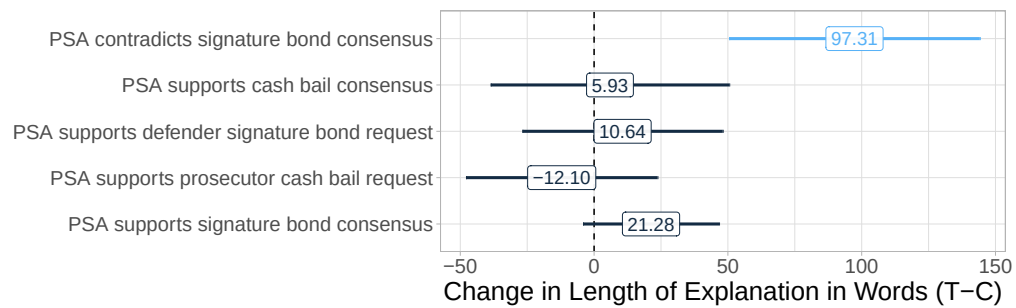
We can further examine the effect of the PSA on procedural outcomes. Figure 2.8 displays the effects of providing PSA reports on a) whether judges offer justifications/explanations for their decisions, b) the length of justifications/explanations, and c) the length of hearings. When the PSA recommendation is more restrictive than the joint requests of prosecutors and defenders, its provision leads to hearings that are 212 words longer on average (from 431 to 644 words), justifications for decisions being provided 46 percentage points more often (more than doubling the rate from 42% to 89% of hearings), and justifications that

Figure 2.8: Effect of PSA on Procedural Outcomes

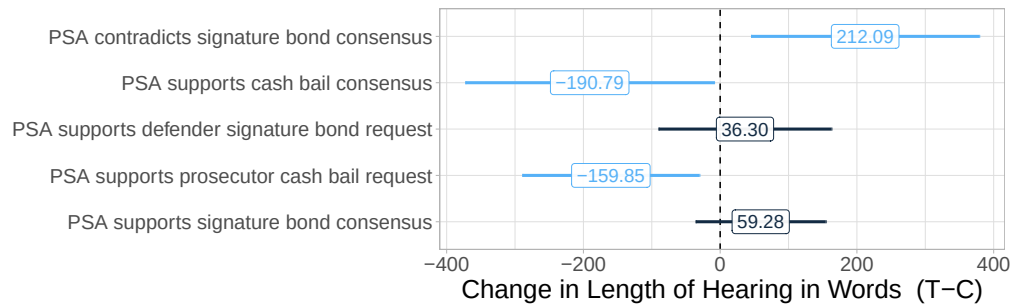
(a) Effect on Proportion with Explanation



(b) Effect on Explanation Length



(c) Effect on Hearing Length



are 97 words longer on average (more than tripling in length, from 45 to 142 words).

Reading transcripts from this subgroup, I observed two changes in how hearings proceeded that account for these effects. Both reflect ways in which the addition of a PSA cash bail recommendation introduced fresh doubts about cases that, as reflected in the processing of comparable cases in the control group, would otherwise likely have been processed as straightforward, uncontroversial signature bonds.

First, in treatment group hearings, judges began to quiz prosecutors on why there was a discrepancy between their requests for a signature bond and the PSA cash bail recom-

mendation, and to nudge the two sides to elaborate on why they thought a signature bond was sufficient. Judges would make statements like, “so I do see that there is a Public Safety Assessment that suggests modest cash bail” or ask “have you reviewed the Public Safety Assessment?” They would then follow up with “did you want to weigh in at all for the reasons for a signature bond?” or “counsel want to be heard about that?” or “Care to add anything?” These interjections hinted to both the prosecution and defense that cash bail might be on the table. Prosecutors were moved to defend their request, while defense attorneys provided more full-throated arguments in favor of a signature bond. This contrasted with comparable cases in the control group where prosecutors often simply made statements like “the State is recommending a signature bond” without offering a reason, to which defense attorneys typically replied “No objection.” Longer hearings result from more cases being subject to added scrutiny.

Second, there were a variety of changes to judges’ closing decision statements. Whereas in control group hearings in this subgroup judges often treated decisions as so self-evident as to not require any explanation, judges in treatment group hearings increasingly portrayed their final decisions as more fraught. When they ordered signature bonds, they began needing to acknowledge why they too were departing from the PSA. For example, judges might mention how they just personally “don’t see a ripe basis for setting cash bail” or note “while I respect the PSA for cash given the nature of the charge, I think the fact that he turned himself in is an overwhelming factor in favor of a signature bond.” In these cases, judges also more often expressed uncertainty and indicated that the decision was “a close call” or “right on the edge” between a signature bond and cash bail.

Meanwhile, when judges ordered cash bail despite a joint request for a signature bond, an action they increasingly took in response to PSA cash bail recommendations, they frequently cited the PSA report as part of lengthy explanations for why they thought the prosecution and defense were underestimating the true risk posed by the defendant. For instance, in one hearing the judge introduces details from the defendant’s history that were not pre-

viously brought up by the prosecutor or defender, and uses the PSA as a pivot to his read of the present 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 Surname] is on parole...for the felony version of substantial battery that relates to causing injury to a law enforcement officer. 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, obstructing, possession of THC. Ultimately he was convicted of those misdemeanors rather than any of those felonies, but 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.

These reactions are consistent with the PSA undermining the unified requests of prosecutors and defenders and introducing new doubts about defendants, placing defendants under greater suspicion than they would have been were there no PSA report.

Providing the PSA also increases the proportion of hearings where a justification is provided by 18 percentage points when the PSA joins a consensus requesting a signature bond (from 41 to 59 percent), though it does not substantially increase the length of explanations or the length of hearings overall. Examining differences between treatment and control observations in this subgroup, it is evident that this difference is attributable to judges bringing up the PSA recommendation when announcing their decisions. Around 13% of explanations in the treated group consist of brief mentions of the PSA in decision justifications,

such as “That result is consistent with the recommendation of the Public Safety Assessment as well, so that’s all good.” In another 12% of treated observations, judges reference the PSA recommendations alongside other rationales.

The PSA did not, meanwhile, have an effect on the proportion of hearings with an explanation in the remaining three subgroups: hearings in which the PSA supported a cash bail consensus, supported a defender’s signature bond request (over the objections of the prosecutor), or supported a prosecutor’s cash bail request (over the objections of the defense attorney). This result is attributable to a ceiling effect. In the treatment group sample, judges provided explanations in every single hearing. This represented little shift from the control group sample, in which judges offered explanations in more than 96% of hearings.

Finally, access to PSA reports reduces the overall length of hearings when a PSA cash bail recommendation supports a prosecutor’s lone cash bail request (by 191 words, a reduction from a mean of 882 words to a mean of 691 words) or supports a joint cash bail request (by 160 words, a reduction from a mean of 968 words to a mean of 808 words). These sets of hearings represent the two with the longest average hearing lengths in the control group. Both tend to be comprised of comparatively complex or fraught cases with a substantial likelihood of resulting in cash bail, and tend to feature extended arguments and multiple rounds of debate between courtroom participants. When the PSA report is available, it appears to streamline these normally lengthy hearings. The pattern is clearest in hearings where the PSA supports a cash bail consensus. Prosecutors, defenders, and judges all speak less in treatment group hearings. Judges also stop debates earlier, nearly halving the number of back-and-forth exchanges they allow before rendering their decision (from 11 speech turns on average without the PSA to 6 speech turns with the PSA). As they do so, they sometimes cite the PSA as an opener to signal to the courtroom that they have heard enough. For instance, judges telegraphed transitions with “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” before elaborating on their read of the case and giving their decisions. The PSA in these cases thus appears to contribute both implicitly and explicitly to judges cutting debate short.

2.6 Reasoning Through Potential Assumption Violations

As noted earlier, the preceding analyses rely on the assumption that courtroom actors do not change their requests when they have access to the PSA report. While they may frame their arguments in different ways given access to the PSA, the assumption requires that the content of their asks remains the same. To the extent that this assumption is violated, the treatment subgroups would differ systematically from the control subgroups, leading to biased estimates of the effect of the PSA.

Although the assumption is likely to hold given how the PSA was implemented in Dane County, I discuss in this section how the estimated effects of the PSA on the use of cash bail might be biased if this assumption is violated and present an additional robustness check.

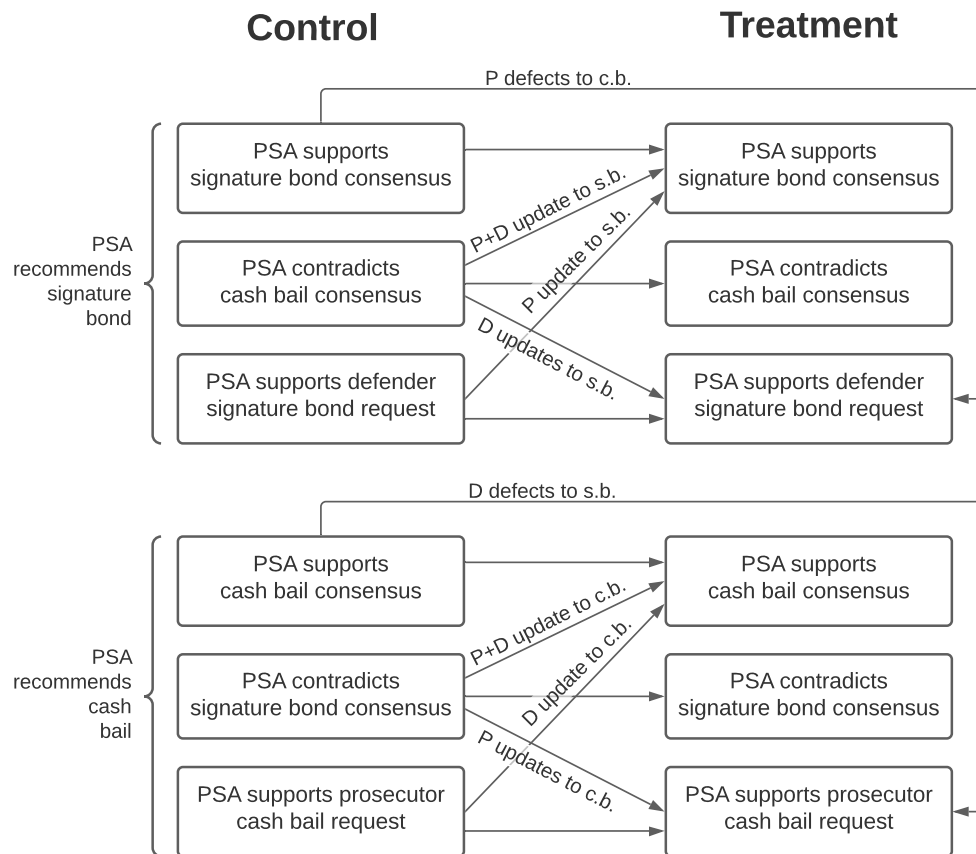
Figure 2.9 illustrates how prosecutors and defenders might respond to seeing either a PSA signature bond recommendation (top) or PSA cash bail recommendation (bottom). It shows where observations in different control subgroups might end up if courtroom actors update their requests in response to the PSA. The diagonal arrows in the center of the diagram depict the variety of ways the prosecutor and/or defender might change their requests to align with the PSA recommendation.

The diagram also presents two updates that are comparatively unlikely. The first is the prosecutor “defecting” to a cash bail request in response to a PSA signature bond recommendation, when the prosecutor would have requested a signature bond absent the PSA. The second such update is the defender “defecting” to a signature bond request in response to a PSA cash bail recommendation when they would have accepted a cash bail without the PSA. The first of these is highly improbable: prosecutors have little to gain from re-

requesting something more restrictive than a signature bond merely to defy the PSA. If they would have accepted a signature bond anyway, there is little incentive for prosecutors to adopt a position contrary to that preference, especially in a progressive county where there is public pressure to reduce, rather than increase pretrial detention.

The second form of defection is more probable than the first, but is still unlikely to affect many cases. It might take place, for example, if defenders are aiming to secure their clients a lower cash bail amount. Recognizing that a PSA that recommends cash bail will likely strengthen the prosecutor's ask, defenders might stake out a more extreme position (by requesting a signature bond when they otherwise would have accepted cash bail) to attempt to secure a lower cash bail amount as a concession. While this scenario is more plausible than the first possibility outlined, it remains unlikely because of the limited conditions under which defenders ever start off accepting cash bail. Recall that defenders only request cash bail in around 1 in 10 cases. Reviewing these cases reveals that these cases generally exhibit one of three characteristics: i) an existing parole or probation hold that prevents the defendant from being released regardless of the bail outcome, ii) very serious felony charges and/or many pending cases already with cash bail set (where all courtroom actors implicitly recognize that a signature bond is out of the question given prevailing practices) and iii) low proposed cash bail that defendants have indicated they are able to post (these are quite rare). As defenders form part of a courtroom community, which entails going along with local norms and not adding inefficiencies unnecessarily, defenders are unlikely to change their requests when they might achieve few tangible benefits for their clients (i and iii) or risk appearing unrealistic or unreasonable by requesting something beyond what is considered reasonable locally (ii).

Figure 2.9: **Map of how the treatment subgroups may be distorted relative to their control counterparts if the assumption that prosecutors and defenders do not adjust their requests is violated.** This figure depicts the range of ways that prosecutors and defenders may alter their requests after seeing the PSA recommendation. The labeled arrows point to which groups observations would flow to if request adjustments were to take place. P stands for prosecution and D stands for defense. S.b. stands for signature bond and c.b. stands for cash bail.



Aggregating all the possible ways in which prosecutors and defense attorneys may update their requests in response to the PSA, Table 2.2 summarizes how treatment subgroups would differ from their control counterparts, along with what biases might result. To reason through the implications, I adopt the simplifying idea in this table that all cases are arrayed along an underlying unidimensional scale of difficulty.³ Easy cases will be cases that exhibit no or few characteristics consistent with courtroom actors perceiving a need for stringent conditions (e.g. defendant with no criminal history charged with a low-level non-violent offense), while hard cases will strongly exhibit characteristics consistent with stringent conditions (e.g. defendant with significant recent history of intentionally missing court charged with a serious violent offense).

I employ this idea in Columns 3 and 4 in Table 2.2 to describe how the average difficulty of cases in each treated group is affected by the various possible prosecutor/defender request adjustment depicted in Figure 2.9 (with “harder” and “easier” being relative to the average difficulty of the cases that would be in the group when courtroom actors do not revise their asks in response to the PSA). Here I assume that when prosecutors and defenders change their requests, they are likely to do so for marginal cases. For instance, I assume that if a prosecutor were to modify a request from a signature bond to a cash bail in response to a PSA cash bail recommendation, they would do so in comparatively “harder” cases first.

Based on the potential shifts in Column 3, Column 4 then summarizes the overall expected net difference in treated group composition. It takes into account only the shifts that are plausible, and excludes unlikely ones such as prosecutor and defender defections in defiance of the PSA, as described earlier.

³I am intentionally avoiding conceptualizing this as “true” riskiness to allow for other considerations besides risk that actors may value.

Table 2.2: Summary of expected net impact of distortions due to courtroom actors updating their requests. Shifts due to unlikely updates are displayed in gray. “Harder” and “easier” are with respect to the average difficulty of the cases that would have been in the treatment group if courtroom actors did not adjust their requests in response to the PSA (or, alternatively, with respect to the expected difficulty of control group cases)

Group	PSA Rec	How treatment group would differ from control if actors update due to PSA Treatment group would have...	Expected net difference Treatment group contains...	Original estimate	Expected bias direction The estimated effect...
PSA supports signature bond consensus	Signature bond	Gained harder cases from - PSA contradicts cash consensus - PSA supports defender signature bond request Lost harder cases to - PSA supports defender signature bond request (unlikely)	Harder cases than control group	- 0.03	Understates negative effect
PSA contradicts signature bond consensus	Cash bail	Lost harder cases to - PSA supports prosecutor cash request - PSA supports cash consensus	Easier cases than control group	+ 0.22	Understates positive effect
PSA supports cash bail consensus	Cash bail	Gained easier cases from - PSA contradicts signature bond consensus - PSA supports prosecutor cash request Lost easier cases to - PSA supports prosecutor cash request (unlikely)	Easier cases than control group	0	Understates potential positive effect
PSA supports prosecutor cash bail request	Cash bail	Gained easier cases from - PSA contradicts signature bond consensus Lost harder cases to - PSA supports cash consensus Gained harder cases from - PSA supports cash consensus (unlikely)	Easier cases than control group	+ 0.16	Understates positive effect
PSA supports defender signature bond request	Signature bond	Gained harder cases from - PSA contradicts cash consensus Lost easier cases to - PSA supports signature bond consensus Gained easier cases from - PSA supports signature bond consensus (unlikely)	Harder cases than control group	- 0.04	Understates negative effect

This table offers several important insights. First, it suggests that if courtroom actors actually change their requests in light of the PSA, my results actually understate the increase in the use of cash bail attributable to a PSA cash bail recommendation when it contradicts a signature bond consensus and when the PSA supports a prosecutor's lone cash bail request. In other words, the true effect of the PSA on the use of cash bail in those scenarios would actually be *greater* than the 22 and 16 percentage point increases that I estimated. Second, the analysis indicates that the results might understate reductions in the use of cash bail attributable to PSA signature bond recommendations that support a signature bond consensus or a defender's lone request for a signature bond. In the main results the estimated effects of the PSA for these subgroups were not statistically different from 0. This analysis suggests that the true effect may in fact be negative. It is important to note, however, that even if PSA signature bond recommendations do in fact cause reductions in the use of cash bail, the magnitude of those reductions is likely to be smaller than the magnitude of the increases attributable to PSA cash bail recommendations.

2.6.1 Restricting Time Between Assessment Completion and Hearing

The assumption that courtroom actors do not update their requests in response to the PSA is more likely to hold when there is minimal lag between when PSA assessments are completed and when courts hold hearings. While court staff generally only provided reports right before hearings, there is a greater chance that courtroom actors might have had advance access to PSA reports if there was a longer delay between when a defendant was arrested (when a PSA was usually completed) and when the bail decision was made.

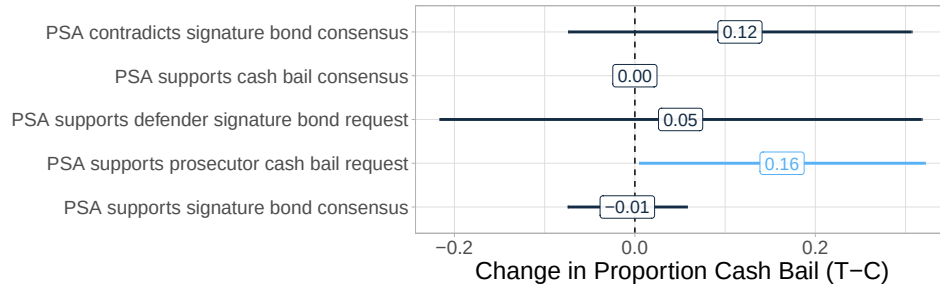
To address this potential concern, Figure 2.10 presents the main results after restricting to the 75% of observations where hearings took place within a day of assessment completion. The results of this analysis are broadly consistent with the main results. Two noteworthy divergences, though, come up with cases in which the PSA contradicts a signature bond consensus. Among these cases, restricting the sample reduces estimates of the

effects of PSA exposure on the use of cash bail (from 22 to 12 percentage points) and on hearing length (from 212 to 117 words).

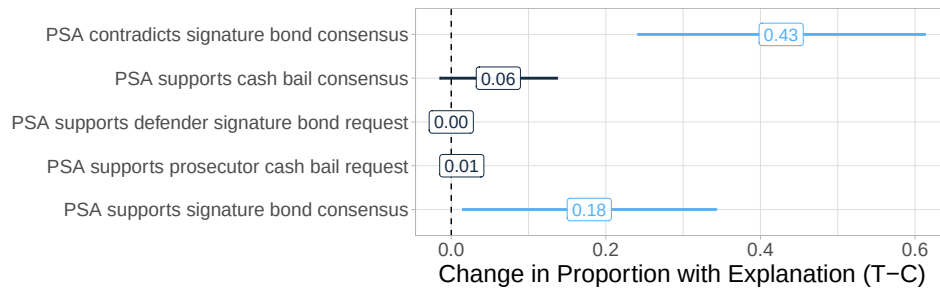
While this discrepancy may be concerning, it is also important to acknowledge that restricting to cases with quick turnarounds may distort the population under examination. There are systematic reasons why hearings are delayed, including to provide time to complete competency evaluations and to give prosecutors more time to make charging decisions. The latter may especially come into play in non-routine or serious cases, or cases involving multiple parties, where there may be greater ambiguity about how a particular defendant should be charged. Cleaving off those observations means that the estimated effects may only apply in the population of relatively straightforward cases. Thus this alternate analysis should be interpreted with caution.

Figure 2.10: Effect of PSA on Use of Cash Bail and Procedural Outcomes, Restricting to Hearings that Took Place within 1 Day of Assessment Completion

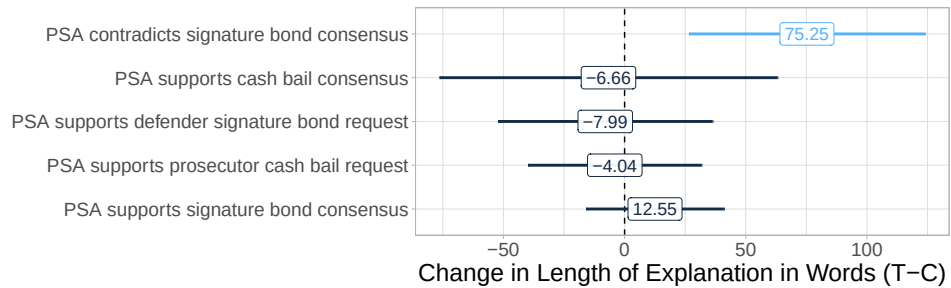
(a) Effect on Use of Cash Bail



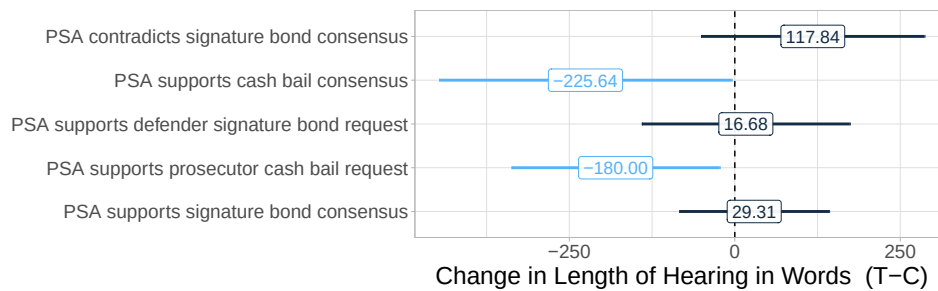
(b) Effect on Proportion with Explanation



(c) Effect on Explanation Length



(d) Effect on Hearing Length



2.7 Discussion

2.7.1 Summary of findings

This paper demonstrated that access to PSA recommendations has an asymmetric effect on the use of cash bail in the research site: PSA cash bail recommendations increase the use of cash bail while PSA signature bond recommendations do not substantially curb the use of cash bail.⁴

As part of this outcome, the PSA improves prosecutors' chances of securing cash bail decisions when they advance requests for cash bail. The PSA's contribution comes in two parts. The PSA not only sides with prosecutor requests more often than it sides with defender requests, it also substantially increases the chances that prosecutor's requests are granted. By contrast, when the PSA aligns with defense requests for signature bonds over the objection of prosecutors, the results suggest that the PSA may be less effective for the defense. This divergence in effectiveness is particularly noteworthy given that qualitative analyses indicate that prosecutors tend to reference the PSA in perfunctory ways, while defense attorneys engage more deeply with the PSA in their arguments. The value of the PSA for prosecutors comes *despite* prosecutors leaning comparatively lightly on the instrument.

The PSA additionally expands possibilities for judges to impose cash bail despite joint requests from prosecutors and defense attorneys for a signature bond, suggesting that the PSA erodes judges' reliance on the two sides to define the space of plausible decision outcomes. Presenting an alternate source of information, the PSA undermines consensus and prompts a second look at cases that would otherwise have been treated as straightforward signature bond cases. This is reflected in how access to the PSA shifts judge departures from the joint requests of prosecutors and defenders from a rare occurrence that takes place

⁴As a reminder, in this project, due to resource constraints, I opted to focus only on cases involving at least one felony charge at the time of the initial appearance/bail bond hearing. The broader RCT also includes a wider set of cases that only involve misdemeanors. In that broader sample, the PSA does not have a statistically significant positive effect on the use of cash bail.

in fewer than 1 in 20 cases to a regular occurrence that takes place in more than 1 in 4 cases. It is further consistent with the procedural outcome results, which indicate that PSA cash bail recommendations that contradict a signature bond consensus lead to longer hearings and more frequent and extensive decision explanations. Here, too, it is important to underscore that the impact of the PSA when it contradicts a consensus is asymmetric. Whereas the PSA contradicts signature bond consensus in about 13%-14% of cases, it only contradicts cash bail consensus in about 2% of cases. In the few cases fitting this latter description, all still resulted in cash bail, suggesting the PSA had no effect. Taken together, then, these results suggest that PSA cash bail recommendations increase judges' power to impose more restrictive conditions over the objections of prosecutors and defenders, an effect that PSA signature bond recommendations do not offset.

2.7.2 Implications, Limitations, and Directions for Future Research

This paper's results align with theoretical predictions and the concerns of civil rights groups that risk assessments systematically increase perceptions of defendant "riskiness" and that the tools may not live up to their promise of reducing reliance on cash bail and pretrial detention. They serve as an illustration of the ways in which analyses of theoretical or simulated policy outcomes based on perfect compliance (e.g. Kleinberg et al., 2018), may lead to overly optimistic conclusions. As advisory tools, risk assessments can be used in an asymmetric manner, with evaluations of high risk afforded significant influence and evaluations of low risk afforded less influence.⁵ Consequently, risk assessments tools like the PSA may not achieve the positive outcomes often touted as a benefit of their adoption.

Of course, it is important to acknowledge that the present study only examines the impact of the PSA as it was implemented in one moment in time in a specific county. To dig further into the reasons behind this particular result and evaluate how readily we might expect it to generalize to other sites, studies of the cognitive and social psychological

⁵To be clear, I am not advocating for courts to move to mandated compliance, which would create substantial due process concerns

mechanisms behind responses to risk assessment information may be valuable. Do these asymmetric responses reflect broader psychological phenomena like loss or risk aversion?

From an organizational perspective, it would also be valuable to consider how asymmetric implementations of risk assessment tools may be rooted in organizational efforts to manage uncertainties in their external environments and may reflect the relative tax that different types of errors can impose on organizational legitimacy. Bail decisions entail establishing a working balance between false positives and negatives – mistakenly deeming a defendant high risk vs mistakenly deeming a defendant low risk. While reasonable people can disagree about the true social costs of each type of error, stories about false negatives may receive significantly more public scrutiny and create more substantial threats to organizational legitimacy.⁶ This kind of dynamic plays out not only in courts, but also in bureaucracies like child welfare agencies (where risk assessments are also being adopted), where the costs of unnecessary investigations and out-of-home placements are usually less visible than the costs of mistakenly leaving a child in a dangerous or deadly environment.

In settings where some errors are viewed as mostly costly than others, we might well expect that risk assessment recommendations would be adopted unevenly. And to the extent that such a dynamic is endemic to a particular type of organizational decision task, we might expect risk assessment recommendations to be taken asymmetrically across many adoption sites. Indeed, it may further be the case that the kinds of organizations most likely to turn to risk assessment instruments may be those most motivated by fears of committing a particular kind of error. This would suggest that risk assessments may, in general, tend to be implemented in asymmetric ways. Future research should investigate the prevalence

⁶Stevenson and Mayson (2021) present the results of a fascinating survey in which they ask respondents in what they call a “Rawlsian cost-benefit analysis” whether they would prefer to be the victim of a crime or jailed for varying amounts of time. The idea was to empirically evaluate when an interest in public safety might outweigh an individual’s interest in liberty. The results generally show that the public highly values liberty, and that current pretrial detention rates are too high to be justified based on public opinion. It is telling, however, that the authors opted to exclude serious crimes such as murder, rape, and domestic violence from the survey, because they present “extremely severe harms” that would not work well in their cost-benefit framework. Yet those crimes, even if comparatively rare, are precisely the ones that courts might most fear and most drive public opinion outside the context of the survey.

of asymmetries in implementations of risk assessment tools in other jurisdictions, as well as evaluate whether there are any common features in the organizational settings in which risk assessment tend to be adopted.

This paper further encourages researchers to consider ways in which risk assessments may reshape the interactional dynamics between multiple parties who may inform or have a stake in consequential organizational decisions. While the express purpose of tools like the PSA is to steer a sole focal decision-maker, risk assessments often intervene in existing social contexts comprised of actors with distinct goals who are connected in vertical and horizontal relationships with one another. Here, in the case of setting bail, the results suggest how a risk assessment tool may strengthen and empower prosecutors and judges relative to defenders. These results echo previous work on the criminal justice system, such as research on sentencing guidelines in criminal courts and big data policing, that highlight how discretion and authority may shift, rather than simply be reduced in the face of efforts to increase standardization and accountability through quantification (Espeland and Vannebo, 2007; Brayne and Christin, 2020). It also builds on recent papers that highlight how other actors involved in pretrial, like law enforcement, might influence risk assessment inputs (Lum et al., 2020). This paper thus joins calls to understand how risk assessments operate as part of pre-existing social systems, alternately shaping and being shaped by the structures to which they are introduced.

Importantly, though, who gains or loses power or authority as a result a new court technology is not a mere academic or theoretical concern, but one with direct consequences for due process and the fair administration of justice. The American criminal legal system relies on the pillar that those who stand accused of crimes should be able to receive adequate representation and the benefit of a rigorous defense. To the extent that a tool may systematically undermine the position of the defense relative to the prosecution, especially in a stage as critical as pretrial, we may be concerned for how it thwarts rather than advances efforts to achieve a fairer criminal justice system.

Chapter 3

The Reactive Potential of Pretrial Risk Assessments

3.1 Introduction

In this chapter, I further build the case that risk assessments ought not be conceptualized as nudging devices that act silently on judges, but rather as resources that different courtroom actors may appropriate for various ends. While the preceding chapter aimed to expand the analytic lens to encompass how the pretrial risk assessments affect dynamics among prosecutors, defense attorneys, and judges, the present chapter aims to further illustrate how the introduction of tools like the PSA can affect defendants' experiences of the criminal justice system beyond their most heavily studied intended outcome of influencing judges' cash bail decisions.

In particular, in this chapter, I demonstrate the impact of the PSA in two previously under-examined areas: how judges talk to defendants and the rate at which judges require defendants to be subject to ongoing monitoring while on pretrial release. Drawing on courtroom observation, analyses of court transcripts, administrative records, I find that judges use the PSA in an unintended way: as characters in rhetorical strategies to encourage de-

fendants to follow pretrial release conditions. I further show that the PSA leads judges to refer select defendants to bail monitoring at higher rates.

These outcomes are important in their own right, as the former affects defendants' courtroom experiences and the latter extends the reach of the criminal justice system into the lives of individuals awaiting trial. But I also contend that shifts in these outcomes present pathways through which predictive risk scores in criminal justice can be *reactive* measures: measures that modify the phenomenon under study and change the very thing that was measured (Campbell, 1957; Espeland and Sauder, 2007). These results thus challenge common assumptions in empirical studies of risk assessments that they operate as informational interventions in otherwise inert, stable social systems, and indicate ways in which judges may leverage predictive scores in unanticipated ways that expand the tool's audience and potential impact.

3.2 Expanding Assessments of the Impact of Pretrial Risk Assessments

In the landscape of predictive algorithms, risk assessments in criminal justice like the PSA have primarily been categorized as triaging tools: devices that sort individuals or other objects in a population according to some underlying quality of institutional interest (Christin, 2020; Eubanks, 2018). This sorting in turn helps institutions understand individuals as instances of more general cases (Heimer, 2001) or classes (Fourcade and Healy, 2013), to which standardized operating procedures and heterogeneous treatment can be routinely and legitimately applied.

While there is minimal disagreement over what risk assessments analytically represent as technical objects, writings on how they intervene on social systems tend to center different groups as their subjects. Some scholarship, especially macro-level theoretical work, tends to focus on defendants as the subject *of* these algorithms, while other scholarship, es-

pecially of a more empirical variety, tend to focus on how judges are subject *to* algorithms (Simon, 1988; Feeley and Simon, 1992; DeMichele et al., 2018; Christin, 2020; Burrell and Fourcade, 2021). Concerns about agency and behavior change in response to measurement figure centrally in the latter set of work, building on broader lines of scholarship that examine how professionals and workers push back against reforms that aim to curtail their discretion or track and manage their activities (Espeland and Vannebo, 2007). Questions like “do judges accept or reject risk assessments and in what ways” seem obvious given that risk assessments are intended as devices to influence their behavior. Questions of how defendants might respond to risk assessments or otherwise be affected by their adoption, besides via how they influence bail decisions, are meanwhile sidelined or assumed away as risk assessment algorithms are deemed too opaque to shape defendant experiences directly (Christin, 2020). Risk assessments are thus thought to only assist with sorting given a population distribution, rather than, through their adoption, alter any aspect of the distribution.

In this chapter, I expand evaluations of the impact of risk assessment algorithms in two directions that challenge this view that the adoption of a pretrial risk assessment algorithm is a clearly bounded intervention without the potential to shape the world beyond the decisions judges make about cash bail. Each illuminates a potential pathway through which predictive risk scores can be reactive measures that, in providing the prediction, modify the very outcome to be predicted (Campbell, 1957; Espeland and Sauder, 2007). These pathways build on those suggested by studies of older forms of risk assessments used in other parts of the criminal justice system, such as in police, sentencing, and parole (Harcourt, 2007; Hannah-Moffat, 2013b; Werth, 2019)

First, I examine how risk assessment recommendations shape how judges speak to defendants in the courtroom. This analysis foregrounds court proceedings as socially and sociologically important in their own right in the study of risk assessment algorithms. As scores said to predict defendants’ future based on their past are stated publicly, they can lead defendants to adjust how they behave (Espeland and Sauder, 2007; Fourcade and Healy,

2013) or expose them to a more negative experience with the criminal justice system, eroding trust and fueling legal cynicism (Fagan and Tyler, 2005; Weaver and Lerman, 2010; Swaner et al., 2018). Drawing upon information in criminal records, moreover, ties people to past identities and can lead to an array of emotional, cognitive, and relational harms (Myrick, 2013; Lageson, 2016; Adams et al., 2017). Thus, how risk assessments are discussed in court hearings is consequential and holds the potential to influence recidivism and whether defendants return to court – the very outcomes the tools are usually trained to predict.

Second, in contrast to studies that have focused on how risk assessments shape the use of monetary conditions, I investigate the effect of risk assessment algorithms on judges' decisions to refer defendants to bail monitoring (also referred to as pretrial supervision). Bail monitoring programs, which are increasingly common across the country, track and monitor defendant activities on an ongoing basis while they are on pretrial release. In Dane County, defendants subject to pretrial supervision must check-in in-person with a pretrial services officer monthly or biweekly. Most are also subject to monitoring to check compliance with other non-monetary release conditions, such as remote alcohol tests for sobriety conditions, random drug testing for bans on the consumption of controlled substances, and GPS monitoring for bans from specific geographic locations.

As check-ins and technical monitoring devices can be expensive and inconvenient for daily life, bail monitoring requirements can place substantial burdens on defendants (Doleac, 2018; Kofman, 2019; Kilgour, 2020). They also expose defendants to increased surveillance and shape the very probability that a defendant is observed violating a condition or committing another offense. While few rigorous studies of the effects of pretrial supervision exist (Bechtel et al., 2017), existing evidence suggests that defendants subject to mandatory drug testing and location monitoring are more likely to be rearrested and to fail to appear for court (VanNostrand and Keebler, 2009). A randomized controlled trial that evaluated the effect of comparable forms of supervision on individuals on probation, more-

over, showed that individuals subjected to more intensive requirements were more likely to be caught committing technical violations and incarcerated at higher rates (Hyatt and Barnes, 2014).

Thus, risk assessments can have a “ratcheting effect” (Harcourt, 2007), where greater predicted risk leads to greater surveillance and monitoring and increased likelihood that additional offenses are observed by law enforcement, which in turn trigger higher risk predictions and even more surveillance. Consequently, to the extent that pretrial risk assessments influence the rate at which judges order pretrial supervision, the tools may extend the reach of the criminal justice system and further directly influence whether their evaluations of defendants are affirmed.

3.3 Data and Methods

3.3.1 How Judges Speak to Defendants

I study the impact of the PSA on how judges in Dane County speak to defendants by analyzing transcripts of judges’ decision announcements, which take place near the end of hearings. Unlike in earlier parts of hearings, which proceed primarily through discussions between the judge, prosecutor, and defense attorney in which the defendant is referred to in the third person, judges in these moments more frequently turn to address the defendant. As this is often the first time judges move to speak to the defendant directly, these moments tend to be salient departures from the rhythm of preceding discussions. Rather than evaluate the defendant as an object, judges use these opportunity to directly issue warnings, admonishments, praise, and instructions to defendants.

I conducted two rounds of coding. I first developed new codes based on emergent themes in the messages that judges directed at defendants. I paid special attention to when and how PSA information was explicitly or tacitly implicated in judges’ speeches. I then conducted a second round of coding to more uniformly apply codes identified in the first

round to transcripts in the treatment and control groups. I analyzed the codes by triangulating between the codes, the PSA recommendations, and bail outcomes to identify specific rhetorical strategies that the PSA facilitated and the contexts in which they were employed.

3.3.2 Effect on Bail Monitoring

To evaluate the effect of the PSA on referrals to bail monitoring, I linked data on cases in the randomized controlled trial to Wisconsin administrative court records. I derived information on whether bail monitoring was ordered in a given case from open-ended text boxes in case records. I then verified how well the procedure worked by checking the extent to which these administrative record-derived measures matched data on bail monitoring orders I assembled from hand-coding hearing transcripts in my sample. The record-derived measure matched the transcript-derived measure completely.

Consistent with the handling of hearings that address multiple pending cases in the preceding chapter, I pooled bail monitoring conditions across all of the cases that are addressed in a particular hearing. That is, if a single hearing addresses two ongoing cases and the judge orders bail monitoring in one but not the other, I record the hearing as having resulted in a bail monitoring order.

I estimate the effect of the PSA on these non-monetary conditions by directly comparing whether bail monitoring is imposed in treated and control hearings from the first year of RCT implementation. I present an estimate of the overall effect, as well as separate estimates for three distinct types of PSA recommendations with respect to bail monitoring: signature bonds with no pretrial supervision, signature bonds with pretrial supervision, and cash bail with pretrial supervision (note that all cash bail recommendations come automatically with a recommendation for pretrial supervision). These recommendations are associated with successively higher PSA scores in the Dane County Decision-Making Framework.

Here, it is worth clarifying the extent to which decisions about cash bail and pretrial

supervision are interrelated matters in the local context. As in other jurisdictions, pretrial supervision can be understood as a potential alternative to cash bail (Wiseman, 2014; Kofman, 2019). Defense attorneys, for example, will request a signature bond with bail monitoring for their clients when prosecutors request cash bail. However, judges also impose bail monitoring conditions when they order cash bail. While, roughly speaking, we might think of conditions as increasingly restrictive as we move from signature bonds without pretrial supervision to signature bonds with pretrial supervision to cash bail without pretrial supervision to cash bail with pretrial supervision, judges locally treat bail monitoring and cash bail as related, but distinct choices. Accordingly I analyze bail monitoring choices separately from cash bail decisions, though I also present corresponding results for the latter to aid with interpretation.

I then turned to transcript data to identify the mechanisms behind the quantitative findings. I examined the characteristics of cases where the prospect of bail monitoring was never raised. I also reviewed transcripts of hearings in which a courtroom actor discussed the possibility of bail monitoring to understand why and how it was ultimately ordered or not. Due to the relative rarity of hearings in the transcript sample that resulted in bail monitoring, I do not have sufficient sample size to conduct formal statistical tests. Rather, I analyze judges' explanations of their thinking in revealing cases to glean insights into how the PSA influences how judges reason through bail monitoring decisions.

3.4 Results

3.4.1 The Rhetorical Use of Pretrial Risk Assessments

Contrary to prevailing visions of pretrial risk assessments as tools that silently nudge judges toward particular bail decisions, I find that judges cast the PSA as an active courtroom participant in three rhetorical strategies that aim to encourage defendants to comply with pretrial release conditions. As part of each strategy, the PSA transcends being treated as

simply a decision input. An overview of these strategies is provided in Table 3.1. While I present these strategies as distinct in the analysis that follows, judges often blend them in practice.

Table 3.1: **Rhetorical Strategies** - summarizes what role the PSA plays in each rhetorical strategy and the situations in which each strategy is most commonly used

PSA role	Situation
Collaborator	PSA recommendation supports signature bond consensus
Foil	Judge departs from PSA cash bail recommendation
Cudgel	Medium to high PSA risk scores

A first strategy positions the PSA as a collaborator. In this strategy, judges refer to the PSA as though it were a member of a supportive collective that is looking out for the best interest of the defendant. For instance, in one hearing involving a young defendant, the judge repeatedly speaks using first-person plural pronouns, folding the PSA in as a collaborative partner aiming to help the defendant move on and avoid pretrial detention:

Okay. So, Mr. [redacted], our primary goal is to get you to court and so we can get your cases put behind you and – have things over with. The odds of that continuing to happen... are much higher if you're complying with law enforcement officers and you make your court dates and you stay out of trouble. *Everybody is saying that they think a signature bond is the appropriate result here, and even our Public Safety Assessment Report suggests that....* I think we'll do it. I'm going to order a signature bond... It's up to you to make good decisions while you're out there. I hope you do. You really don't want to come in here having a third felony case. At that point – pretty likely some kind of a cash bail is going to be set, and if that happens and you can't come you know with the dough, you would be stuck here. That would be a shame.

In this strategy, the judge positions the PSA as a particularly tough-to-please voice in

the room. It joins a unified front helping the defendant. The defendant is encouraged to “make good decisions” as a way to realize a collective goal to secure a good outcome.

A second strategy is used when judges decide to impose less restrictive release conditions than the PSA recommends. When judges are more lenient than the PSA, they cast the tool as a foil to themselves. Judges position themselves, and sometimes prosecutors, as more forgiving actors in comparison to the detached, pessimistic PSA. For instance, in a hearing involving a defendant with a history of substance use challenges, the judge concluded the hearing by turning to the defendant with PSA report in hand and stating:

This document that I have here is called the PSA. It’s supposed to be a scientific document, and it recommends cash bail. But what I see is that things may be improving for you, so I will take a shot at a signature bond. I hope that you can make sure that things continue to improve going forward. Good luck to you.

This strategy draws attention to the judge’s personal willingness to exercise their discretion to take a chance on the defendant. It works by simultaneously elevating the legitimacy of the PSA report as a “scientific document” and drawing a contrast between the tool’s limited impersonal gaze and the judge’s more personal understanding of the defendant. The defendant is encouraged to comply with court orders as a way to prove the PSA wrong.

A third strategy employs the PSA as a cudgel to admonish defendants and warn that they are skating on thin ice. As part of this strategy, the PSA is positioned as a trusted authority testifying to the defendant’s elevated risk. For example, in a hearing where the prosecutor and PSA recommended cash bail and the public defender argued for a signature bond, the judge reasoned through his ultimate decision to impose a signature bond in the following manner:

Well, I note the Public Safety Assessment says moderate cash bail. They allege failure to appear in the past 2 years, one older than 2 years. He’s a 5 out of 6 on both the criminal activity and failure to appear. It certainly concerns me... I

don't think there is a ton of built-in incentive for flight given the nature of the charges. It's not something that often ends in a prison sentence... so reluctantly I'm going to go along with the signature bond....

You're already on notice that the evidentiary tool we're using suggests that you're a cash bail person if you keep coming back, and that's when the State is going to start asking for cash bail and of course the Court will probably go along with it then.

In contrast to the second strategy, in which judges suggest that they see something unique in the defendant as a person that the algorithm does not, the third strategy confines the disagreement between the judge and the tool to a more technical disagreement about the particulars of the charges and whether they in fact create an incentive for flight. Despite departing from the PSA's recommendation for cash bail, the judge states the PSA's risk evaluation of the defendant as fact. He further makes clear his reluctance to go against the recommendation. In using this strategy, judges draw defendants' attention to the fact that they are only narrowly being granted pretrial release without monetary conditions. This strategy aims to motivate by drawing attention to the precarity of the defendant's position.

In some instances of this PSA-as-cudgel strategy, judges further explicitly spell out for defendants how violating pretrial release conditions can affect their future PSA results. For example, in a case in which the judge declared that the decision was "a close call," the judge prefaced his order to grant a signature bond by telling the defendant:

I'll warn you, Mr. [Defendant], just like I warned Mr. [Co-Defendant], bail jumping counts in the Public Safety Assessment Report. If you pick up a bail jumping, you may be in cash bail territory because you're inching up toward the top of the scale. You're a 4 out of 6 in both the new criminal activity and failure to appear scales.

Facilitated by the transparency of the PSA's design, judges can emphasize the impor-

tance of following conditions by laying out how negative actions, like violating release conditions, can generate numerical reactions. In so doing, judges establish the risk assessment as a potential device for intentionally influencing behavior, transforming it from a momentary measure on the page to an ongoing structure of evaluation that prescribes what actions defendants should and should not take.

Indeed, across all three strategies, judges use the PSA in an unintended way: as a means to influence what defendants do upon release. While each strategy operates by a different logic, all cast the PSA as an active presence in the courtroom. Far from a simple assessment that the judge may ignore or take under advisement as a decision input, the PSA is woven into hearings in a socially meaningful way at salient moments in which judges engage defendants directly.

Contrast with Control Group

It is worth clarifying that judges still give motivational speeches and admonish defendants in the absence of the PSA. Comparing judges' rhetoric in the treatment and control groups, however, helps to clarify what access to the PSA specifically adds or modifies. First, as explored in the first chapter, the very presence of PSA reports can provide a source of disagreement in the courtroom. It can offer a dissenting perspective against a consensus, or align with one side's request against that of the other. And further, as also illustrated in the preceding chapter, the PSA more often recommends more restrictive conditions than less restrictive conditions in comparison to other courtroom actors.

This dynamic opens new discursive possibilities by expanding the set of cases where there is at least one party that expresses doubt about whether a defendant should be subject to minimal release conditions. This furnishes judges with more occasions in which they can emphasize that the defendant is in a precarious position, as the cudgel and foil strategies tend to stress. Absent the PSA being available to be animated as a voice for greater restrictions, more hearings proceed with no conflict and little to no commentary from judges. The

following, for example, is one typical such exchange in the control group, where the PSA would have recommended high cash bail but was withheld:

PROSECUTOR: State is requesting a signature bond, with non-monetary conditions that she have no acts or threats of violence to [Alleged Victim], that she have no physical discipline to any child, and that she maintain absolute sobriety.

DEFENSE ATTORNEY: No objection.

THE COURT: We'll order a signature bond. Conditions will be that you direct no acts or threats of violence towards [Alleged Victim], that you engage in no physical discipline of any children, and that you maintain absolute sobriety. As to time limits?

Indeed, comparing the treatment and control groups indicates that judges speak directly to defendants (as opposed to referring to defendants in the third person) 6 percentage points more often when the PSA is available rather than withheld.

A second, related, contribution of the PSA is that judges' admonishments and warnings when they opt for a signature bond amidst dissenting voices hinge less on speculating about the mercy or discretion of prosecutors, and lean more on appeals to evidence and numbers. Absent the PSA, judges underscore the importance of following conditions typically by, for instance, cautioning defendants that they "don't know if the State's going to recommend a signature bond again" or that "these are the sort of charges where if you screw up on the signature bond and violate some condition of bail or pick up new charges, it's going to be a pretty easy case for them to make to ask for cash bail." Both the presence of the prosecutor and risk assessment facilitate judges' tendencies to de-emphasize their own agency in making future pretrial decisions, and to instead suggest that the defendant's fate rests on how these two harsher actors might react to the defendant's choices. Access to the PSA, however, especially accentuates this tendency, as it enables judges to ground their claims

in science and make precise statements about mechanical mathematical reactions to defendants' actions. Despite the full scope of the discretion that judges in fact enjoy, access to the PSA enhances judges' ability to rhetorically claim that their hands would be tied and that, should violations result in significant cash bail and pretrial detention, those choices, as one judge put it, "are not made by people on this side of the bench."

3.4.2 The Effect of Pretrial Risk Assessments on Bail Monitoring Referrals

In this second section of results, I first draw on administrative data to evaluate the effect of the PSA on the use of bail monitoring based on a comparison of treated and control hearings. I then turn to hearing transcripts drawn from these two groups to identify mechanisms behind the quantitative results.

Effects in first year: treated vs control results

Figure 3.1 displays estimates of the effect of the PSA in the first year of implementation, comparing hearings in which the PSA was randomized to be available to ones in which the PSA was randomized to be withheld. Figure 3.2 presents the rates at which bail monitoring was ordered in the treatment and control groups. For reference, corresponding estimates for the effect of the PSA on cash bail orders are also displayed on the right.

Figure 3.1: **Average Treatment of Effects of PSA Exposure on Use of Bail Monitoring**

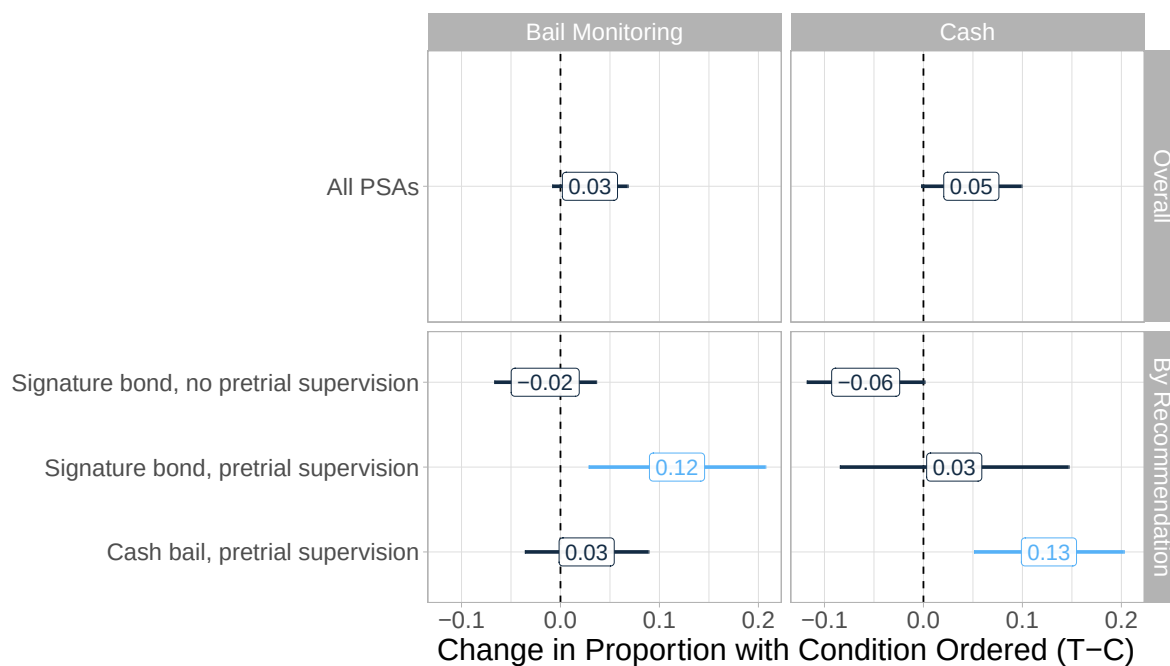
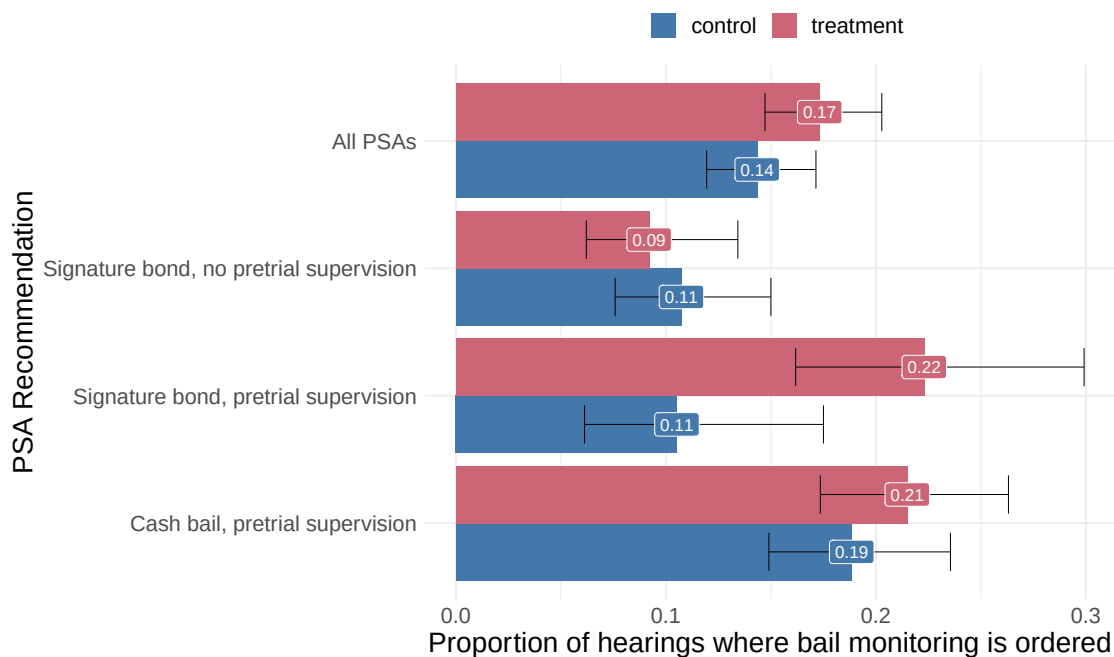


Figure 3.2: **Mean Rates of Bail Monitoring Orders in Treatment and Control Groups**



The results indicate that, overall, access to the PSA did not have a statistically significant effect on the use of bail monitoring. However, breaking out by recommendation type

reveals heterogeneity in the PSA's impact. In cases where the PSA recommends a signature bond and no pretrial supervision or cash bail with pretrial supervision, access to the PSA report does not appear to have a substantively meaningful effect on the use of bail monitoring (though it does appear to cause shifts in cash bail orders). However, in cases where the PSA recommends a signature bond with pretrial supervision, the PSA is estimated to increase the use of bail monitoring by 12 percentage points without a corresponding significant change in the use of cash bail.

As further illustrated in Figure 3.2, this result indicates that access to the PSA in these latter cases doubles the rate of bail monitoring referrals. Whereas in the control group, cases with signature bond recommendations with and without pretrial supervision recommendations were treated similarly, in the treatment group, they diverge. When PSA reports are provided to the court, cases with recommendations for signature bonds with pretrial supervision come to resemble ones with cash bail recommendations in how often they result in bail monitoring referrals.

In sum, these results suggest that providing courts access to the PSA expands the use of bail monitoring among cases designated by the tool to be appropriate for a medium degree of pretrial restriction. PSA recommendations on the extremes meanwhile have comparatively less impact on the use of bail monitoring.

Mechanisms

Hearing transcripts suggest that one mechanism behind the large effect of pretrial supervision recommendations that come with signature bond recommendations is that they may be more salient to courtroom actors. This is reflected in Table 3.2, which shows how often PSA pretrial supervision recommendations are discussed in hearings where the PSA was presented. The first column presents the proportion of hearings where the PSA's specific recommendation for pretrial services is explicitly mentioned. It indicates recommendations for pretrial services are discussed in around 26% of hearings with a signature bond recom-

mendation and only in around 16% of hearings with a cash bail recommendation.¹ This pattern appears despite the fact that, as the second column suggests, the PSA is mentioned at higher rates overall in cases with a cash bail recommendation. The third column further shows that, among discussions of the PSA, pretrial supervision recommendations are discussed twice as often when they are paired with a signature bond than when paired with a cash bail recommendation.

Table 3.2: **Salience of pretrial supervision recommendation in treated hearings**

PSA Recommendation	Prop of Hearings w/ PS Rec Mention	Prop of Hearings w/ Any PSA Mention	Prop of PSA Mentions w/ PS Rec Mention
Signature bond, no pretrial supervision	0.01	0.78	0.01
Signature bond, with pretrial supervision	0.26	0.81	0.32
Cash bail with pretrial supervision	0.16	0.97	0.16

Judges are especially prone to invite prosecutors and defense attorneys to comment on the recommendation for pretrial services in this group. Judges encourage additional consideration of pretrial supervision citing the PSA, asking prosecutors and the defense for instance “And I do see that recommendation is consistent with the PSA report, except the PSA report suggests the possibility of Pretrial Supervision. You’re not obligated to, but if you wanted to add anything related to that?” or “I do note that the Public Safety Assessment Report suggests a signature bond, which is being jointly recommended, but also suggests that pretrial supervision of some sort might make sense. I don’t know if anyone has any thoughts about that in regards to safeguards?”

When the PSA recommended cash bail and was provided to the court, judges also increasingly independently raised the possibility of bail monitoring in cases where prosecutors requested a straight signature bond without pretrial supervision, citing the PSA as they did so. This is illustrated in the exchange below, in which the PSA recommended high

¹The 1% of hearings where a pretrial supervision recommendation was discussed when it was not recommended consisted of an unusual single case where the staff member who generated the PSA stated that he may have made a data entry error that, if corrected, would have led to a more restrictive recommendation.

cash bail:

PROSECUTOR: The State's requesting a signature bond...

DEFENSE ATTORNEY: No objection.

THE COURT: Does anyone want to talk me out of at least bail monitoring? The PSA recommends cash... it looks like there's a slew of missed court dates...and he's a 5 out of 6 on the failure to appear scale.

PROSECUTOR: Given the PSA, the State would not object to bail monitoring in this case.

By diverging from the prosecutor's request, the PSA here offers a resource to judges. Judges can leverage the PSA's cash bail recommendation to undermine the prosecutor's evaluation of the "riskiness" of the defendant and legitimate a bail monitoring condition. They can do so while still appearing to take a moderate stance because the tool disproportionately recommends even more restrictive conditions. In cases where the PSA was withheld but *would have* recommended cash bail, by contrast, I did not observe any cases in which judges independently introduced a bail monitoring condition in such a setting.

But while recommendations for pretrial supervision coupled with signature bonds and with cash bail similarly enable judges to more easily independently raise the prospect of bail monitoring in hearings, transcript data further suggest that the relative composition of the cases that come with the two types of recommendations may contribute to the observed differential effects.

Cases with a cash bail recommendation are more likely to exhibit characteristics that courtroom actors think of as limiting the potential usefulness of bail monitoring. In particular, transcript data indicate that judges and prosecutors argue against bail monitoring when defendants' other entanglements with the criminal justice system, such as other ongoing cases, parole, and probation, would reduce the potential gains of added monitoring. A commonly stated reason why bail monitoring is not ordered despite being raised as a prospect by the PSA or other courtroom actors is that added monitoring for individuals who

are already under the supervision of probation or parole officers would represent a waste of scarce court resources. In a case in which the judge declined to order bail monitoring, for instance, he explained explicitly, “He’s got all kinds of other forms of supervision. If that’s not effective, I don’t think another one is going to be.” Similarly, in another case, the judge reasoned through his decision, stating

While I understand that the Public Safety Assessment is suggesting maximum level supervision, if you were to look at Bail Monitoring, they are often over-worked and over capacity. Because he’s on supervision... they will be monitoring him. They have the capability to put him on GPS and do the things that Bail Monitoring has. I don’t want to duplicate the resources under those circumstances.

The issue that the defendant is already subject to other forms of supervision is more commonly cited as a reason why bail monitoring is not ultimately implemented despite being considered in cases with cash bail recommendations. This is likely reflective of the fact that criminal history indicators are key inputs into the algorithm and are likely correlated with whether defendants are under some form of supervision. This dynamic may thus place a ceiling on the potential responsiveness of judges’ bail monitoring decisions to recommendations for pretrial supervision coupled with cash bail recommendations.

By a similar logic, a second reason why the effect of pretrial supervision recommendations may be muted among cash bail recommendation cases is that judges and prosecutors express the view that bail monitoring is unlikely to be effective for defendants with significant histories of terminations from the bail monitoring program or non-compliance with other release conditions. And, given the construction of the PSA, defendants with cash bail recommendations are substantially more likely to have such histories. Table 3.3, for instance, summarizes the proportion of cases in each group with at least one or more charges associated with violating bail conditions. It shows that a substantially larger proportion of

hearings with cash bail recommendations feature at least one case in which the defendant is charged with a bail condition violation.

Table 3.3: **Distribution of Bail Jumping Charges by PSA Recommendation**

PSA Recommendation	Proportion of Hearings w/ At Least 1 Bail Jumping Charge
Signature bond, no pretrial supervision	0.10
Signature bond, with pretrial supervision	0.25
Cash bail, with pretrial supervision	0.46

Judges express doubts about whether referrals would be an efficient use of resources when they believe that bail monitoring may not be effective, noting for instance that “the record of compliance is definitely not good, and I’m not really seeing that Bail Monitoring is going to do a whole lot of good at this point.” This matter emerges most regularly when the defense attorney requests a straight signature bond or a signature bond with pretrial supervision, in lieu of cash bail, as recommended by the prosecutor and PSA. In these contexts, judges tend to conclude that cash bail is instead the best solution as a way for defendants to get “skin in the game.” Judges make statements like “Well, given the nature of the charge, given the Bail Monitoring Program failure about 2 years ago, and taking into account that it appears as though he was just discharged from supervision within the last few weeks, I think some cash bail is appropriate.” The notion that a financial incentive is what is necessary is sometimes further spelled out, as one judge did in rejecting a defense attorney’s request for a signature bond with pretrial supervision:

Well, it looks like Mr.[Defendant Surname] was released August 21st of this year and then again on October 9 of this year, so this is the third case...He hasn’t been able to comply with Court ordered conditions. The fact that somebody is bail jumping is a factor to consider whether he’s going to show up in court.... The Public Safety Assessment recommends cash bail. He’s missed court in the past, having failed to appear at that pretrial proceeding older than 2 years. I’ll take into consideration his limited means, but I think at this point he should

have cash bail such that the State can move to forfeit that if he's going to continue to not comply with his bail conditions... I'm ordering cash bail. I'm going to set it in the amount of \$300 given his minimal ability to pay. I don't think \$700 [the prosecutor request] is necessary, but I think we have to have something that he can lose if he violates his conditions of bail.

Here and in similar cases with condition violations, judges express preferring to set a meaningfully high cash bail amount to manage risk, rather than set no or low cash bail and offset risk by introducing bail monitoring. And as such cases with repeated condition violations appear more often among cases with cash bail recommendations, they contribute to the reduced impact of pretrial supervision recommendations within this group.

Taking all these results together, then, suggests that PSA recommendations for pretrial supervision tend to expand monitoring among defendants who are otherwise not under supervision. And rather than encouraging downgrades from cash bail to signature bonds with bail monitoring, the PSA appears to spur judges to add monitoring or more restrictive financial conditions to cases that would likely have otherwise resulted in a straight signature bond.

3.5 Discussion and Implications

In this chapter, I illustrated how the PSA shaped courtroom practices in two key ways. In the first half of the analysis, I highlighted three rhetorical strategies that judges employ that cast the PSA in active roles in courtroom discussions. Far from serving as silent decision-making inputs, risk assessments are animated by judges as they seek to emphasize to defendants the importance of complying with release conditions. As judges publicly invoke the PSA in open court, its evaluations of defendants and its recommendations gain a wider audience than typically presumed in past scholarship on risk assessments.

In the second half of the analysis, I showed that the adoption of the risk assessment

increased referrals to bail monitoring among defendants that the risk assessment deemed a medium level risk. The results suggest that these shifts may be concentrated among defendants who are not otherwise under active supervision and who likely would have received a straight signature bond absent the PSA. The PSA thus contributes to extending the reach of ongoing criminal justice monitoring to a broader set of defendants. In the process, the PSA increases the chance that those defendants may later be found in violation of pretrial release conditions. Thus, through their effect on pretrial supervision, risk assessments have the potential to alter the constraints and challenges that defendants must navigate upon release, and ultimately, the outcomes that defendants realize.

I contend that the PSA's impacts on these outcomes are important in their own right because they directly affect defendants' encounters and entanglements with criminal justice system. At the same time, they suggest ways in which pretrial risk assessments can have reactive effects. With respect to this latter point, I wish to clarify that the objective and scope of the present chapter are admittedly modest: to identify potential mechanisms of reactivity that have remained hidden or been assumed away in past scholarship on pretrial risk assessments. Further research is needed to examine how these mechanisms ultimately shape outcomes like recidivism and making court dates,² as well as how they affect important intermediary outcomes like defendants' subjective experiences with and understandings of pretrial justice.

Nonetheless, this research presents several implications for scholarship on the consequences of algorithmic decision aids. First, this study suggests that scholars should take a broader view of what these tools represent when deployed within organizations and institutions. They are not mere informational nudges, but rather present resources to local actors, who can employ them to advance their own ends in unanticipated ways (Christin, 2017; Lynch, 2019). Future research should thus devote additional attention to the interactional and discursive possibilities that these tools create. Inductive and abductive approaches may

²Part of what makes studying this difficult is that, at this moment in time, those outcomes have not yet been fully realized for all cases in the sample.

be especially well suited to uncovering the range of intended and unintended ways that actors use these tools in a given institutional setting.

Attention to these practices can also help advance our understanding of whether and how algorithms are accepted by the individuals whose decisions they are supposed to influence. The embrace of the PSA observed in this study is surprising in light of past scholarship suggesting that legal professionals, who reify professional judgment and the experience of practitioners (Gorman and Sandefur, 2011; Sandefur, 2015; Greiner and Matthews, 2016), are likely to reject interventions that would curtail their discretion or undermine their expertise. It further stands in contrast to ethnographic research that found that a similar tool was largely discounted and ignored in other jurisdictions (Brayne and Christin, 2020). This research presents the possibility that an algorithm's reception depends on whether local actors find uses for these tools that advance their objectives without cannibalizing their professional authority and autonomy. Building on this insight, future research could fruitfully explore what local conditions facilitate or prevent professionals from discovering and pursuing these uses.

Second, this chapter suggests that future research should further examine how the decision-making environments and procedures into which predictive triage aids enter may expand or constrict possibilities for the tools to influence parties beyond the immediate decision-maker. In particular, scholars should consider how proximity between those making decisions and those about whom decisions are being rendered might mediate outcomes. On one of the spectrum, for example, are jurisdictions in Utah in which initial bail decisions are made by judges logging into a computer system, reviewing documents, and recording a decision. Without an interactive context, tools like the PSA may indeed function more like nudges. On the other end of the spectrum, meanwhile, may be predictive tools like those used in health in which there are more possibilities for direct engagement between the intended professional users of the tools and the patients evaluated by them.

In these latter types of interactive contexts, the use of such tools can create openings

that could enable those who are evaluated to adjust their actions in ways that support or subvert predictions, especially when the operations of the algorithm in question are transparent (Berman and Hirschman, 2018). This observation further points to the importance of additional inquiry into the social experience of being judged by an algorithm. What kinds of processes does such an experience set in motion? In what contexts are these experiences more or less likely to influence subsequent behavior? How do these experiences affect one's trust in the legal system and other formal institutions?

Finally, the argument advanced here challenges common assumptions that undergird research designs used to evaluate the predictive performance of risk assessment algorithms. Much of this research views a defendant's potential outcomes – whether they come to court or commit an offense – as set and independent of how decisions are made. For example, in research comparing the predictive performance of judges and risk assessment algorithms, researchers rely on designs in which the underlying data used were actually generated by human decision makers unaided by algorithms (Kleinberg et al., 2018). These studies compare the decisions that humans made to the decisions that a predictive algorithm *would have* made. But these comparisons are only valid if shifting to algorithmic decision-making were like flipping a switch. As documented in this chapter, however, risk assessments reconfigure courtroom discussions and defendant experiences. These experiences can, in turn, have separate direct effects on subsequent defendant outcomes. Consequently, researchers should be careful in interpreting the results of studies that rest on the assumption that *how* decisions are made is inconsequential and should design studies that account for these dynamics.

3.6 Appendix

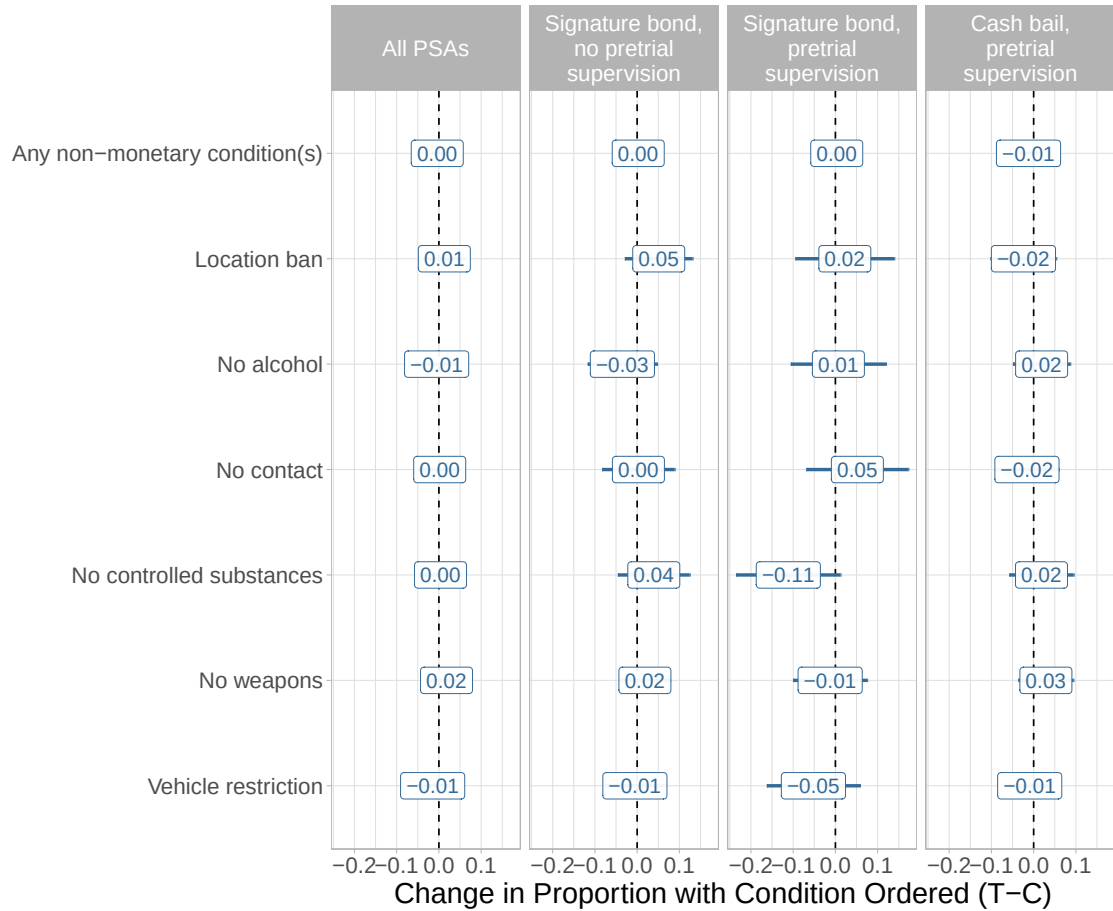


Figure 3.3: Effects of the PSA on Other Non-Monetary Conditions

Chapter 4

The PSA's Impact on Decision Goals and How Defendants are Evaluated

In this chapter, I examine the consequences of PSA adoption for what dimensions of defendants get discussed in bail hearings and how courtroom actors position competing goals that pretrial conditions may help to achieve. Unlike the preceding chapters which focused primarily on decision outcomes, this chapter centers procedural concerns and investigates the extent to which courtroom discussions come to mirror the logic of actuarial prediction that the Public Safety Assessment models.

4.1 Theoretical Predictions about How Actuarial Prediction Shapes the Means and Ends of Decision-Making

How might the adoption of predictive actuarial risk assessments shape how courtroom actors evaluate defendants and understand the goals of their decision-making? One body of macro-level theoretical scholarship suggests that algorithmic prediction and actuarial tools elevate calculations of risk as the focal basis for institutional action (Feeley and Simon, 1992; Fourcade and Healy, 2013). Within criminal justice, scholars contend that such tools

have helped to usher in a broad transformation in the means and ends of punishment. Under what has been called the “new penology” (Feeley and Simon, 1992), criminal justice institutions understand themselves as managers of risky and unruly populations (Simon, 1988; Garland, 2001; Harcourt, 2007). As part of this shift, actuarial prediction is argued to facilitate a pivot away from evaluating people as individuals and toward evaluating people in terms of their resemblance to aggregate profiles. Broadly, this body of scholarship suggests that actuarial tools like the PSA narrow conceptions of just punishment and foreclose alternate models like rehabilitation or retribution (Feeley and Simon, 1992; Garland, 2001; Harcourt, 2007).

Scholars argue that this newly dominant approach to justice and punishment is built on a corresponding set of beliefs about human nature and models for how actors should come to see and know people (Hannah-Moffat, 2013b). First, it rests on the foundation of a belief that people are largely immutable. The goal of actuarial prediction is to accurately discern a person’s future riskiness, a characteristic attributed to people as an underlying individual trait. Second, relatedly, this approach facilitates what Fourcade and Healy (2017) have described as a naive form of behavioralism. In drawing on records of past actions to predict future actions, actuarial prediction cements a presumption that observed outcomes are mechanical translations of who one is and of one’s fundamental moral character. This is important because, as a form of commensuration, “the valuation or measuring of different objects with a common metric” (Espeland and Stevens, 2008), predictive risk scoring can direct attention away from the characteristics of a person deemed irrelevant by the scoring system (Espeland and Stevens, 1998; Espeland and Vannebo, 2007). Accordingly, we might expect the use of these tools to reduce criminal justice actors’ interest in unmeasured factors boxed out by the score, such as intentions, meanings, and social context, which have traditionally been assessed to understand a person’s actions and evaluate people as moral actors. Instead, we might expect these tools to encourage criminal justice actors to privilege the measured factors that are used for scoring, like tallies of a person’s past infractions.

Finally, because algorithmic prediction requires selecting a measurable outcome as the object of prediction, predictive tools may encourage criminal justice actors to narrow their understanding of the goals of decision-making tasks. Scholars and advocates, for instance, have argued that risk assessments tools in criminal justice can displace or conflate thick, rich concepts like danger to public safety and flight risk, with thinner, imperfect proxies like likelihood of rearrest or likelihood of missing court (Feeley and Simon, 1992; Robinson and Koepke, 2019; PJI, 2020). They moreover contend that by only predicting these types of negative outcomes, they may crowd out other imperatives that should come to bear on pretrial decision-making, like whether pretrial conditions present undue burdens to defendants and whether cash bail is calibrated to a defendant's financial situation.

Other scholars, however, have challenged assertions that actuarial devices and quantification more broadly have truly reconfigured institutional actors see subjects in a deindividualized, simplistic, and mechanistic manner (Lynch, 1998; Kemshall and Maguire, 2001; Cheliotis, 2006; Werth, 2017; Lynch, 2019). Among other arguments, scholars critique macro-level theories for downplaying the agency of actors within these systems to resist, retrofit, and reinterpret risk scores (Lynch, 1998; Brayne, 2018; Brayne and Christin, 2020), trapping theories in a "dystopian cul-de-sac" (Cheliotis, 2006), and for failing to sufficiently account for on-the-ground meso and micro level processes that might accommodate multiple models of decision-making, rather than a single hegemonic one (Haney, 1996; Werth, 2017)

In particular, empirical research across an array of research sites in criminal justice and beyond testify to the endurance of individualized narratives, moral evaluation, and professional judgement as fundamental ways that institutional actors continue to try to understand people (Kiviat, 2017, 2019; Lynch, 2019). In these accounts, decision-makers make sense of what numbers say about a person's future actions by continuing to try to place them in a narrative context, rather than unproblematically regarding them as instances of a case (Heimer, 2001). In interpreting credit scores, for instance, employers infer a

sequence of events that could explain an applicant's score (Kiviat, 2017). These stories help employers make sense of scores as either reflecting an applicant's fundamental moral character, such as dishonesty or untrustworthiness, or understandable situational exigencies (Ross and Nisbett, 2011).

In the closest parallel to the case of how pretrial risk assessments are used in courts, Lynch (2019) studied the impact of sentencing guidelines, which prescribe sentences based on scores that reflect a defendant's criminal history. The study revealed a variety of strategies that prosecutors, defense attorneys, and judges use to harness quantified criminal history to tell contrasting biographies of the defendant as a moral agent. They may alternately try to undercut the score's formula, offer different interpretations of what a score really means, or otherwise make claims about what a defendant's true score should be to post-hoc justify their evaluations of defendants.

Extant research on concrete institutional practices thus would suggest that pretrial risk assessments may be unlikely to yield any substantial changes in the ways that courts come to "know" defendants. Rather than spurring dramatic shifts, numbers in an adversarial legal setting in which defendants are presented individually may be leveraged in flexible ways that accommodate status quo approaches to evaluating and making decisions about defendants. This would suggest that courts continue to grant defendants the ability to change and evolve. We might observe no systemic shift in the information about defendants that courtroom actors weigh nor in courtroom actor understandings of the objectives of pretrial decision-making.

4.2 Particularities of the Pretrial Context

While risk assessments used in pretrial decision-making bear important parallels to ones used in sentencing, probation, and parole, and in other domains of institutional life, the specific context of pretrial justice has distinctive features that I highlight in this section.

These features might influence how likely we are to see risk assessments have a similar impact on evaluation practices in pretrial as in other domains. They further lead to policy questions, unique to the pretrial context, over how to define the goals and target outcomes of judges' decisions.

In a first key contrast to decision-making at later stages of the criminal justice process, pretrial decision-making happens in a highly time-constrained, low-information environment. To give arrestees an early opportunity to be released, bail hearings in most jurisdictions take place within a few days of booking. While these time limits help reduce unnecessary detention, they leave prosecutors and defenders little time to prepare and to assemble rich, detailed information about defendants. Unlike in sentencing, parole, or probation – the primary sites of studies of the on-the-ground impact of quantification – criminal justice actors at pretrial often lack the space and information to spin rich narratives. Risk assessments at pretrial, moreover, compete less with other ways of knowing subjects, such as direct interpersonal interaction. Instead they are often competing with other heuristic, standardized, non-narrative information about defendants, like criminal justice records and answers to surveys that the clients of public defenders complete.

While scholars have found that moral judgment can still suffuse assembly-line forms of justice (Kohler-Hausmann, 2013; Van Cleve and Mayes, 2015), we know less about how risk scoring systems alter these practices. Based on the factors just cited, we might expect pretrial to be a setting within criminal justice where risk assessments encounter less resistance than in other sites and where the conditions may be most favorable for observing the actuarial dominance proposed in some theoretical accounts. At the same time, it is worth acknowledging that pretrial decision-making still proceeds in a less impersonal, distant manner than tasks in other domains like lending or insurance, in which institutions may only ever come to know subjects through standard applications. Thus, relative to other sites in criminal justice where actuarial tools are being adopted, pretrial decision-making may be one where individualized evaluation is more challenging. Relative to other institutionalized

domains, however, it may be one where it remains a greater possibility.

A second crucial distinctive feature of pretrial is fuzziness over the goals and basis of pretrial restriction and detention. Unlike in sentencing or probation, the legal basis for imposing conditions and detention is *not* punishment, as defendants are considered innocent before proven guilty. Instead, the constitutional and statutory basis for the government to deprive people of liberty is to prevent public harm (Stevenson and Mayson, 2021). Historically, the primary goal for setting pretrial conditions has been to manage the chance that a defendant flees and escapes the justice process.¹ Since the 1970s and 1980s, however, legislatures have increasingly revised federal and state bail statutes to add a second legitimate basis for pretrial restriction: dangerousness, the chance that a defendant could jeopardize community safety while awaiting trial.

The conflation of flight risk and dangerousness has been proposed as an important contributor to the overuse of cash bail and pretrial detention (Gouldin, 2016). In many state statutes, including in Wisconsin, the use of monetary pretrial release conditions is only supposed to be used to ensure court appearance. Dangerousness meanwhile is supposed to only inform the imposition of non-monetary conditions. On the ground, however, judges often fail to distinguish between these two concerns, leading judges to overestimate risk and impose cash bail at higher rates and in higher amounts than may be warranted statutorily by non-appearance risk alone.

The PSA intervenes in this context of goal ambiguity in two ways, one posited to be beneficial and one posited to be pernicious. On the beneficial side, the PSA's outputs formally distinguish between three outcomes: future non-appearance, future crime, and future violent crime. This feature was an intentional design choice on the part of the tool's creators, and sets it apart from other risk assessment tools that predict a composite "pretrial failure" outcome which pools risk of nonappearance and crime in a way that critics suggest can further legitimize overestimating defendant risk (Gouldin, 2016). The PSA's design,

¹For further discussion of the costs imposed by flight, see (Gouldin, 2018)

may thus enhance analytic clarity in courtroom discussions by succinctly and formally teasing apart concerns about flight and dangerousness. This has been posited to better align decisions with statutory restrictions on the basis for setting cash bail, enhance accountability, and provide judges political cover to not over-detain.

At the same time, the PSA's definition of target outcomes has also been presented as a liability in that it fails to distinguish between practical and morally distinct forms of nonappearance. Advocates and scholars suggest that simply measuring the outcome of non-appearance fails to account for differences in the reasons why a defendant misses a court date, such as whether or not it was intentional (Robinson and Koepke, 2019; PJI, 2020). Gouldin (2018), for instance, provides a taxonomy that distinguishes between "true flight" – a defendant intentionally fleeing the jurisdiction and going into hiding – and other forms of "local non-appearance" – a defendant staying in the area and missing court for reasons like confusion about court dates, unforeseen personal emergencies, work conflicts, or lack of transportation or childcare. Critics of generalized prediction of non-appearance argue that it lumps the latter type of non-appearance, which is often associated with socioeconomic disadvantage and can be reduced through low-cost interventions like text message reminders (Fishbane et al., 2020), together with rarer, more concerning types of true flight. Accordingly, we might be concerned that risk assessments can muddy the waters of pretrial decision-making. It may lead judges to impose harsher conditions for those who have missed court in the past by collapsing otherwise socially meaningful distinctions and effectively exacerbate punishment of social disadvantage.²

Here I want to emphasize that these proposed promises and perils of pretrial risk prediction derive from analyses of the formal technical design of the PSA itself. Analysts to date have often alternately hailed or criticized the tools by looking at what they predict and how. Implicitly built into their claims is an assumption that the conceptual definitions

²Note that these are distinctions among different forms of flight risk that have been drawn in normative legal scholarship and the work of advocacy groups; they are not formally recognized in federal law and state statutes.

and logical connections that the tool asserts among inputs and outputs translate seamlessly into practical implementation. The fears raised by critics and some theorists attribute what might be considered heroic feats to actuarial tools, suggesting that they are able to redefine rich moral concepts, redirect the attention of highly trained professionals, and reconfigure longstanding institutional scripts and routines in their image.

Yet as a rich body of scholarship on the impact of technology in organizational settings has articulated (Barley, 1986; Orlikowski, 2000), these assumptions may not be warranted. A variety of psychological, cultural, and organizational processes can mediate the impact of technologies in practice and reshape organizational practices and structures in turn. This chapter thus aims to advance the literature on risk assessments in pretrial justice by empirically evaluating the impacts of risk assessments on evaluation practices that have been proposed in past scholarship.

4.3 Research Questions and Expectations

This chapter is organized around an overarching question: to what extent does the adoption of a pretrial risk assessment shape how courtroom actors “see” and evaluate defendants? While the preceding chapters examined decision outcomes, this chapter’s focus is instead on procedural outcomes. In particular, I ask:

Research Question 1, how defendants are seen: How does access to risk assessments influence what information about defendants is discussed in hearings? Do judges start directing attention to risk assessment inputs and outputs at the cost of directing attention to other components of a defendant’s profile that they might have paid more attention to otherwise?

Research Question 2, decision goals: In a setting where decisions may serve multiple competing loosely-defined goals, how does the adoption of an actuarial risk assessment reconfigure courtroom actors’ understandings of what their collective decision tasks are to

achieve? How does it influence the arguments that different actors advance?

Table 4.1 summarizes what we might expect to observe given existing theoretical and empirical scholarship:

Table 4.1: **Predictions/expectations based on past scholarship**

	Macro-theoretical scholarship	Empirical scholarship on professional resistance and organizational practices
How defendants are seen	Debates reflect view that people don't change, risk is an inherent trait	Debates reflect view that people change, risk is situational
	Practices mirror the logic of risk assessments, courtroom actors adopt view that past actions = future actions	Courtroom actors resist the logic of risk assessments, persist in trying to understand reasons for past actions
	<u>Implications</u> - More attention paid to records of past actions - More attention paid to other risk assessment inputs - Less attention paid to intentions, social context	<u>Implications</u> - No changes in where courtroom actors direct attention - Courtroom actors contextualize risk assessment results
Decision goals	Risk assessments elevate the target outcomes that they predict, replace thick concepts with thin operationalizations	Courtroom actors deploy risk assessments in ways that accommodate old goals, contest thin operationalizations
	Risk assessments box out imperatives that they do not speak to	Courtroom actors continue to attend to other imperatives

4.4 Methods

I manually coded all speech acts in the treatment and control group transcripts to mark occasions in which courtroom actors discussed the defendant's characteristics, background, history, social context, relationships, and PSA results. I also coded instances where courtroom actors sought more information or made assertions about what aims pretrial release conditions should or could help to achieve. To develop a coding scheme and identify specific topics to code for, I first began with a preliminary review of 40 randomly selected transcripts. I then applied the coding scheme to the rest of the transcripts. As rarer top-

ics emerged that did not appear in my initial set, I marked them as “other” and returned to classify and consolidate the text in a final round of coding. In analyses that follow, I present estimates of the effect of the PSA on how frequently coded topics are discussed across hearings.

Based on my reading of these transcripts and guided by legal scholarship on legitimate bases for pretrial restriction, I also selected specific topics to examine further for changes in valence or meaning, including age, drug/alcohol addiction, and flight risk/non-appearance. I tracked how these subjects were discussed in transcripts of hearings that took place before PSA adoption and in treatment and control hearings. I present the results of those analyses in case studies.

4.5 Results

I split the results into two parts. The first examines the PSA’s impact on how often and in what ways prosecutors, defense attorneys, and judges discuss different aspects of a defendant’s current situation, history, personal background, and social context in courtroom debates. The analyses start with an overview of the effect of PSA access on how frequently different topics are discussed across hearings. Those summary results are followed by two case studies: one explores how the design of the PSA facilitates a dynamic rather than fixed model of human behavior and another investigates how the ideas embedded in the PSA influence how courtroom actors make sense of what a defendant’s age should mean for pretrial decisions.

The second half of the results examines how the PSA shapes courtroom actors’ depictions of the goals and target outcomes of pretrial release conditions. I start by evaluating the effect of access to the PSA on the aims that prosecutors and defenders appeal to in advancing their requests, as well as on the aims that commissioners invoke in rendering their decisions. I then present findings on how the PSA contributes to subsuming ancillary

goals, like protecting the health of defendants, under the banner of assuring court appearance through a case study of the handling of drug addiction.

4.5.1 Results Part I: How defendants are seen: what dimensions of defendants are brought to bear in courtroom discussions

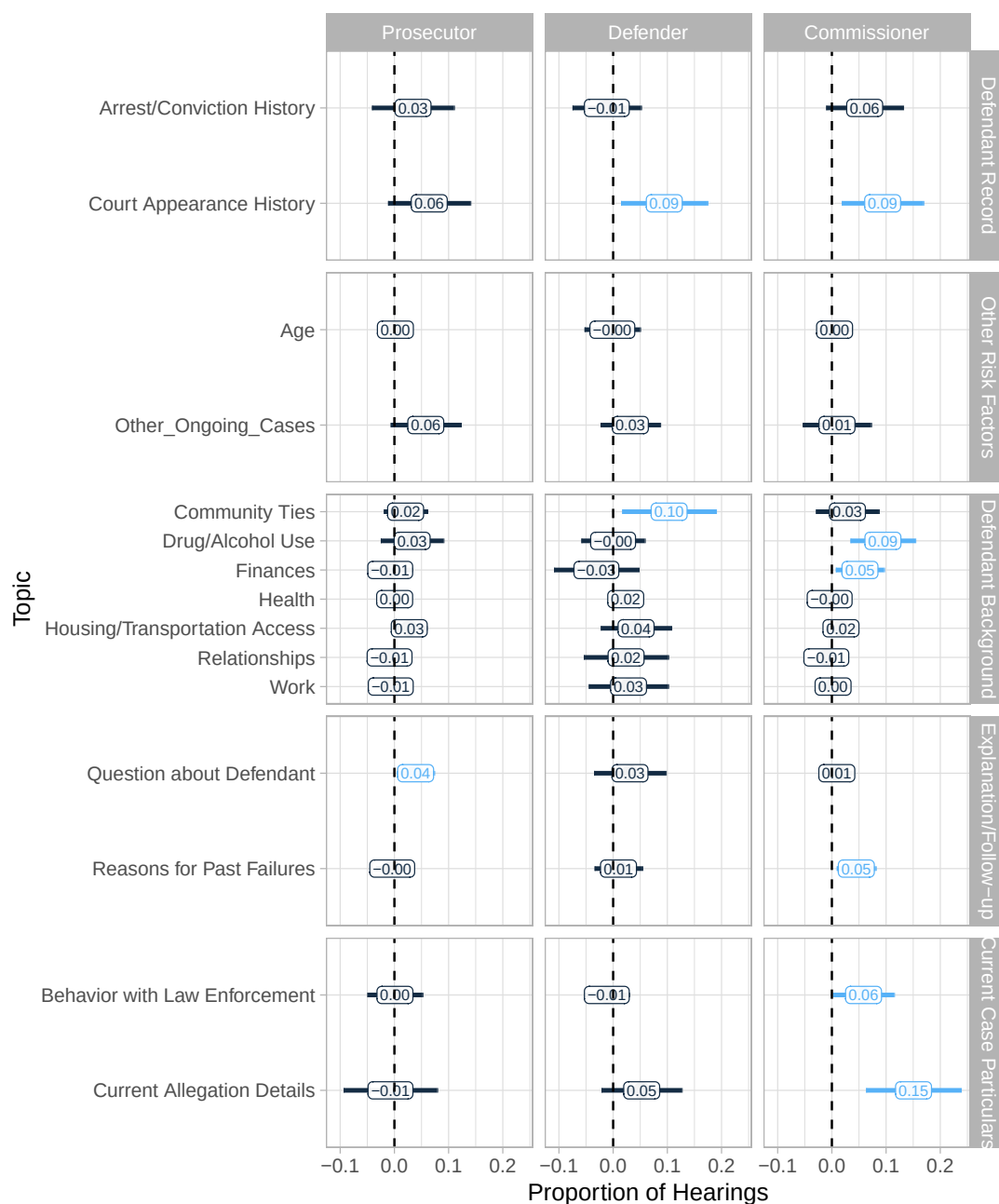
Figure 4.1 displays the average treatment effect of access to the PSA on whether different aspects of a defendant's past and current circumstances are discussed in hearings.³ The column on the left presents estimates for the topics that prosecutors discuss, while the column in the center presents estimates for the ones that defense attorneys discuss. Estimates for effects on judges' closing decision announcements, in which judges offer summaries of the factors that they deemed pertinent to their decisions, are presented on the right.

The top two sets of rows cover topics that are PSA inputs. The first set contains topics related to a defendant's past record, which macro-theoretical scholarship would posit might increasingly dominate discussions. The results indicate that access to the PSA increased how often defense attorneys and commissioners discussed a defendant's history of making court appearances by 9 percentage points. Court transcripts suggest that this shift may be attributable to the fact that PSA reports present easily digestible summaries of a defendant's court appearance history, which make the information more readily accessible than it would be otherwise. There is not evidence of a comparable effect, however, for discussions of a defendant's arrest/conviction history. Similarly, the results indicate no substantial boost in discussions of defendant age and a defendant's other ongoing cases.

The remaining sets of rows encompass topics that we might expect to receive less attention if court proceedings were to fully mirror the actuarial logic of the PSA. Broadly, the results suggest that the PSA has little impact on what aspects of a defendant's background and current allegations prosecutors tend to highlight in hearings.

³I also present additional results on heterogeneous effects by defendant race and sex in the Appendix.

Figure 4.1: Effect of PSA on topics discussed, by courtroom actor



Access to the PSA does, however, lead defense attorneys to assert the strength of their client's ties to the community 10 percentage points more frequently. This finding is notable given that ties to the community are a traditional consideration in pretrial decision-making and is recognized in many state statutes, including in Wisconsin, as a basis for setting release conditions. The PSA, however, excludes community ties as an input, a deliberate decision to align with research indicating that the community ties are poor predictors of pretrial outcomes and ethical concerns that such factors can unfairly disadvantage vulnerable and marginalized groups (Starr, 2014; Bechtel et al., 2017; DeMichele et al., 2020).

Defense attorneys' discussions of community ties are often linked to their efforts to draw greater attention their defendant's strong court appearance histories and low Failure to Appear PSA scores. They enrich the PSA's evaluations with statements like "[she] has lived in Wisconsin for the last 20 years" and "his parents, his siblings, and his friends and extended family all live here." In hearings, the PSA often serves as a jumping-off point for defense attorneys to elaborate further on how defendants' community ties make it unlikely for them to flee the jurisdiction.

Judges, for their part, do not appear to discuss a defendant's ties to the communities more often when provided access to the PSA despite defense attorneys increasingly raising the subject. Instead, judges more often highlight two aspects a defendant's background: their drug or alcohol use and their financial situation. The former appears to particularly reflect judges drawing attention to indications of opioid addiction in a defendant's current charges or past record. Judges bring up the issue as an important factor that the PSA overlooks but that they believe substantially increases the chance that a defendant will miss court, with one local judge regularly making remarks like "you obviously have an opiate addiction and people who have an opiate addiction are the worst as far as showing up in court." The latter, meanwhile, is reflective of judges increasingly signaling that they are taking defendants' financial situation into account in setting bail.

Further contrary to expectations that the risk assessment might reduce interest in con-

textual details and the circumstances surrounding a defendant's actions, the results also show that judges increasingly devote time in hearings to unpack and evaluate the reasons behind a defendant's missed court appearances, a finding I explore further in the Appendix. The results additionally suggest that the PSA causes judges to more often delve into the details of the allegations against defendants, typically to deduce clues as to their dangerousness and likelihood of making court dates, with special attention paid to defendants' interactions with law enforcement. Similar to opioid addiction, judges locally regularly assert that negative interactions with police – such as allegedly resisting arrest or giving a false name – represent revealing behaviors that are important risk factors. On one occasion, for example, a judge stated:

This would be – I think this would be a pretty solid signature bond case if it wasn't for the battery to police officer and the resisting allegations in this one Complaint. I can't think of a purer indication of someone's unwillingness to come to court than getting in a knockdown drag-out fight with the police in their parking lot of all places which is what's alleged here.

As illustrated in this quote, judges position actions toward police officers as tie-breakers that they believe reveal a defendant's general predisposition to follow rules and respect authority. In some cases, judges further explicitly point out how the PSA is incomplete for lacking the input:

One of those Complaints alleges that he was an escapee from a Juvenile Facility at the time of his apprehension – well, both at the time of his apprehension, but also at the time in January where he's alleged to have lied about his identity to Law Enforcement...I do see there is a Public Safety Assessment report. It recommends a signature bond, but that is based on the fact that it doesn't and can't possibly take into account some of these other factors that were described including his escapee status.”

Drawing on the observation that the PSA lacks this key input, judges more frequently discuss defendant interactions with law enforcement to legitimize departures from the tool.

To sum up, these results, taken as a whole, are largely inconsistent with the conjecture that the logic of the risk assessment, including what it values and what it devalues, are replicated in what courtroom actors pay attention to and discuss in courtroom proceedings. They do, however, offer indications that the PSA leads judges and, to a slightly lesser extent, defense attorneys to attend more often to topics that they believe are related to the likelihood that the defendant will appear for future court dates. We observe increases in discussions of court appearance history, as endorsed by the PSA, as well as active re-assertions of factors that the PSA excludes but that local actors believe help to predict non-appearance, such as drug/alcohol use, behavior with law enforcement, and the contexts behind past non-appearances.

Does use of risk assessments advance an understanding of riskiness as a fixed individual trait?

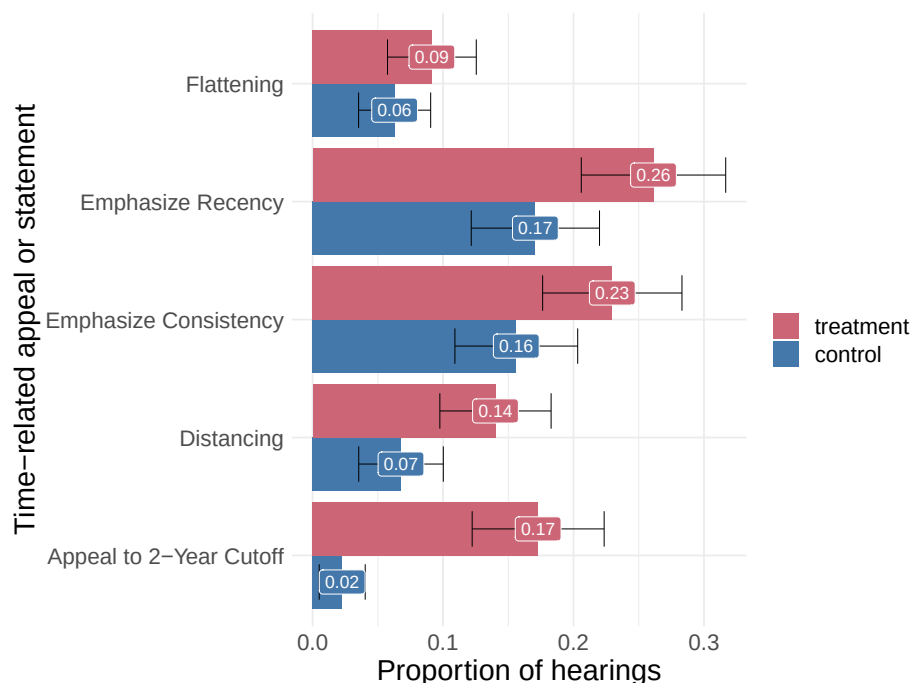
Scholars have suggested that the adoption of risk assessments perpetuates the view that a defendant's riskiness is an inherent, largely immutable individual trait. Yet the PSA presents a model for analyzing a defendant's past record that is dynamic and discounts older events. The PSA introduces a two-year threshold beyond which missed court dates are weighed less. The standard PSA report presents two separate entries for the inputs of "Prior Failure to Appear Pretrial in Past 2 Years" and "Prior Failure to Appear Pretrial Older than 2 Years", with values of yes or no for each. While both increase a defendant's score and the restrictiveness of their corresponding recommendation, missed court dates in the previous two years count more against a defendant's risk.

Comparisons of pre-adoption, control group, and treatment group transcripts suggest that courtroom actors pick up this temporal discounting that the PSA encodes as they debate and evaluate defendant histories. At a basic level, when PSA reports are available,

courtroom actors start to explicitly appeal to the two-year threshold as a standard for analyzing missed court dates. This two-year threshold appears for instance when defenders want to discount older records, such as when one noted, “the cases that were cited are from 2005 and 2006, and the Public Safety Assessment within the last past 2 years [indicates] there were no failures to appear.” Judges also regularly read off the PSA as they explain their decisions, such as by stating “given the Public Safety Assessment indicates he’s missed 2 court appearances in the last 2 years, I think arguably the State could have requested more bail.”

On a broader level, regardless of whether the two-year threshold is explicitly invoked, the PSA spurs courtroom actors to emphasize the timing of past court-related events in their arguments. As shown in Figure 4.2, PSA reports increase the rate at which courtroom actors make statements that draw attention to the recency of the defendant’s record by 9 percentage points. Commissioners for instance increasingly make statements like “I think given the pretty horrific appearance record recently with cash bail posted I’m going to set cash bail...” while prosecutors draw added attention to how little time has elapsed since a missed court appearance. Defenders, for their part, increasingly try to explain away recent missed court dates or draw attention to the recent court dates that their clients did make.

Figure 4.2: **Distribution of time-related appeals and statements across hearings**



The PSA also appears to enhance possibilities for temporal distancing – asserting that things that happened long enough ago in the past no longer reflect the defendant today – while accompanied by no substantial changes in temporal flattening – asserting that a defendant’s older records should be taken into account. When the PSA is provided, it increases the rate at which temporal distancing occurs in hearings by 7 percentage points. Judges increasingly made remarks suggesting an adverse event is “ancient,” “old,” or “dated.” Defenders further more often highlighted the elapse of time, such as by stating “the most recent case is 2005, well over ten years ago ” or by arguing “I don’t think a 2011 [six years prior] missed court date should be held against him.”

An additional consequence of the PSA is that it helps to draw attention to clean records by facilitating claims about the consistency of a defendant’s good behavior across time. The results indicate that the PSA increases the rate at which hearings feature a statement on the consistency of a defendant’s history by 7 percentage points. Defense attorneys for instance will observe “that there is no missed court at any time even recently or even be-

yond 2 years” and prosecutors and commissioners may acknowledge “he’s apparently never missed court previously” or “there doesn’t seem to be a history of failure to appear.” The ability to make such definitive statements may be attributable to the fact that, absent the clear summary that the PSA presents, it can be confusing, cumbersome, and time consuming to review a defendant’s court records thoroughly enough to arrive at these conclusions. Public defenders, in particular, regularly lack the time to assemble this information. With the provision of the PSA, that information becomes more readily accessible.

Taken together, these results suggest that the PSA facilitates discussions that feature a more dynamic, rather than fixed, model of defendants. By providing a two-year threshold and more broadly modeling a type of temporal discounting, the tool opens up additional space for discussions that follow its logic. By further providing quick summaries of a defendant’s record, it opens up more chances for courtroom discussions to draw attention to a defendant’s consistency over time, claims that would otherwise be more challenging to make definitively.

Actuarial prediction and the meaning of age

To further understand the extent to which the PSA shapes evaluation practices, I analyze in this section how the PSA influences the meanings that courtroom actors attach to age, a factor that represents a double-edged sword in criminal justice (Stevenson and Slobogin, 2018). On the one hand, youth is commonly and legally considered a mitigating factor. In sentencing, for instance, youth are granted diminished culpability and seen to be less blameworthy than older individuals. Yet, from the vantage point of risk, youth can also be an aggravating factor. Empirically, youth are systematically more likely to engage in criminal activity (Rocque et al., 2016). Indeed, the relationship between age and crime has been described as “one of the brute facts of criminology” (Hirschi and Gottfredson, 1983).

As a technical object, the PSA intervenes in this tension by endorsing the latter perspective. “Age at Current Arrest” and “Current Violent Offense and 20 Years Old or Younger”

appear under the heading “Risk Factors” on every defendant’s report. Being 22 years of age or younger adds to New Criminal Activity Scores, while being charged with a violent criminal offense at age 20 or younger contributes to New Violent Criminal Activity ratings. While the preceding quantitative analysis shows that there is no shift in how often age is discussed, does youth come to qualitatively assume a different meaning when the PSA informs pretrial release decisions?

Across pre-adoption, control, and treated hearings, there is some stability in the treatment of age. Defenders often draw attention to the youth of their clients as a factor in support of greater leniency. Sometimes they make statements like “she’s 18” as self-evident arguments for restraint in condition-setting. Other times, defenders draw more direct links between youth and other factors to support arguments for signature bonds: young people lacking the financial resources to post cash bail, young people having family that can monitor them and provide structure, and young people being especially harmed by pretrial detention because it would impede their ability to go to school and work.

Judges, on the whole, are sympathetic to these types of claims of youth as a mitigating circumstance. Judges speak of giving youthful defendants a chance on signature bonds, discuss exercising restraint in setting non-monetary conditions that would felonize youthful behavior, and suggest that they are exercising restraint in setting cash bail amounts. For instance, in one hearing where the PSA was available and recommended high cash bail, the judge noted, despite chastising the eighteen-year-old defendant for accumulating multiple pending cases that, “You know, I’m with the State that some cash bail is necessary to assure appearance in court here. If Mr.[Defendant Surname] was not an 18 – yeah, an 18-year-old guy – we’d be talking about a lot bigger number than what the State’s proposing, but I’m assuming he doesn’t have a whole lot of money.”

Yet while the PSA does not appear to spur a large-scale, systematic about-face in the use of youth as a mitigating factor, there are indications that the adoption of the PSA tool creates openings for new interpretations and arguments for how age may inform a decision.

Somewhat surprisingly, the strongest examples of age making new lines of argumentation possible come in the control group, where the PSA infrastructure is in place, but where the specific PSA report for the current case under consideration is withheld.

With the PSA in place, judges can draw upon the link that the PSA draws between age and pretrial risk as a social fact supporting their decision-making, even absent access to the report itself. This is illustrated in one hearing involving a twenty-year-old defendant, in which the prosecutor requested \$10,000 in cash bail for allegations of recklessly endangering safety and violating a court order to not contact the alleged victim in another ongoing case. The defender countered with a request for a signature bond with bail monitoring, highlighting the defendant's youth as a mitigating factor and noting that the defendant was concerned about maintaining his employment. The judge, however, retorted:

I was just looking at the Complaint. It looks like it was a situation where he has 4 pending cases... We have very serious allegations here. He's looking at significant potential penalties in prison. They're violent in nature. He's a young guy, but, you know, I think the Public Safety Assessment and the research in that would suggest that young people aren't great at following conditions of bail because of their youth they make impulsive and foolish decisions. I think it's a cash bail case. I quite frankly think the \$10,000 is reasonable given the significant exposure here and his prior history on bond.

Here, the judge references the ideas embedded in the PSA as a way to counter the defense attorney's argument and support his decision to impose a high cash bail. In the process, he goes beyond what the PSA asserts from a prediction perspective. He injects additional moral valence to the association by declaring that youth "make impulsive and foolish decisions." It is further worth noting that the judge's decision and rhetoric around this particular defendant contradicted what his *actual* PSA report would have indicated had it been provided: a low failure to appear score of 2 out of 6 and a signature bond recommendation.

A second manner in which the presence of the PSA in the local court system can perturb standard understandings of how age should inform decision-making is by undercutting arguments for leniency on the basis that the defendant's age would make pretrial detention especially burdensome or harmful. One example of this took place with an older defendant who was charged with burglary. The prosecutor sought cash bail, while the defender requested a signature bond, stating "I guess I would just indicate – also perhaps ask for some consideration for the fact – that Mr.[Defendant Surname], as far as for people in the jail, is getting to be on the older side, I think he's 59 years old, and I think he would prefer to not be in the jail for sleeping and so forth."

In response, the judge recalled a recent earlier case involving the same defendant that had been assigned to the treatment group:

I may be wrong about this, but my recollection is that the State asked for cash bail at that appearance. I granted him a signature bond despite a Public Safety Assessment report that then recommended modest cash bail. It recommended that based on a 3 out of 6 on a failure to appear score, but at that time a 5 out of 6 on the criminal activity score. *It would appear that it's evidence-based omniscience because here we are with a new criminal complaint alleging a burglary* [emphasis added]...I think we're at the point where some cash bail needs to be set...

I am sympathetic to the fact that he's older than the average person here in the County Jail, but there is really only one person that's made that choice he's made. Good luck now."

This example highlights how affirmation of the PSA's predictive power in a separate case, can legitimize imposing more stringent conditions in subsequent cases.⁴ Here, further,

⁴As a side note, this example is one among a few others indicating that there may be spillover between the treated and control groups. As you may recall, randomization happens at the level of cases, and the same defendant could have multiple treatment assignments at once. When defendants are charged in multiple cases

the judge implies that the decision is out of his hands, and ultimately places responsibility for the unfortunate outcome of detaining an older defendant on the defendant himself.

To be clear, such explicit examples occur infrequently in the sample. However, they point to ways in which the PSA can reshape pretrial reasoning or provide new lines of argumentation that were not possible in its absence. Explicit discussions of such factors can moreover represent just the perceptible component of a more subtle larger shift, one which may be more readily detectable with a larger sample of hearing transcripts.

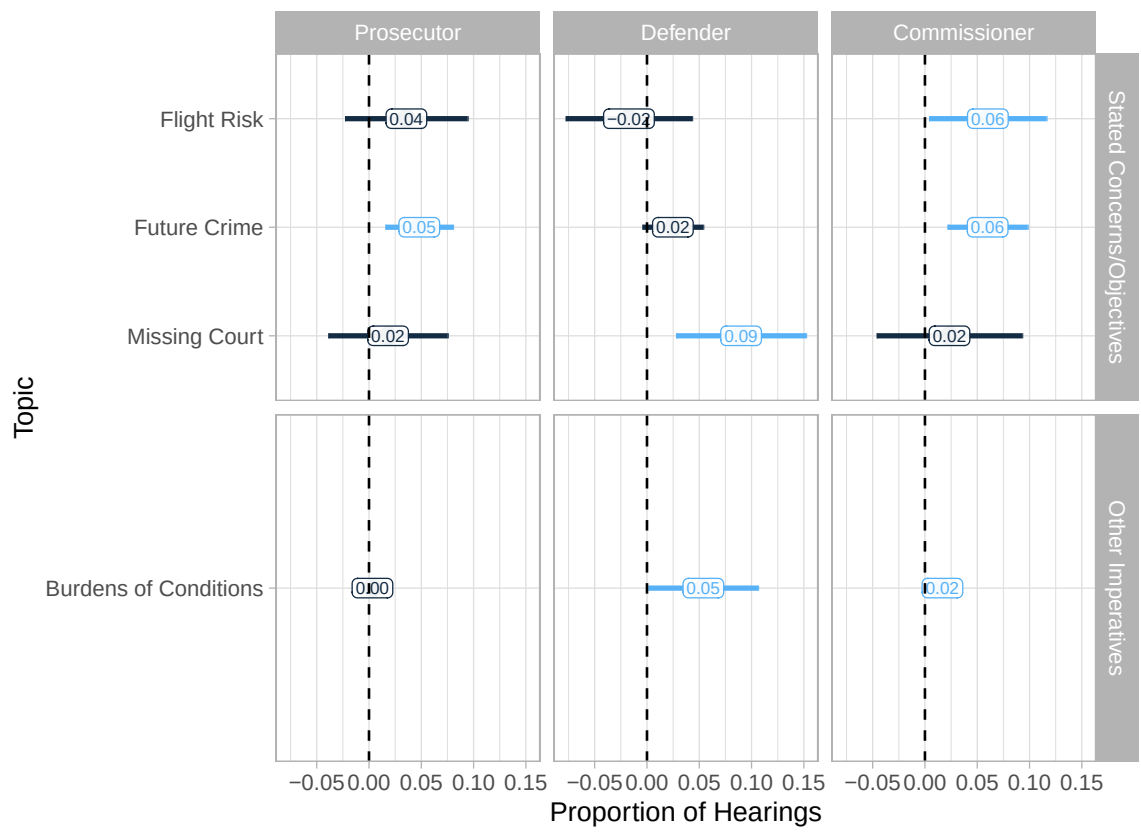
4.5.2 Results Part II: Decision goals and target outcomes

Figure 4.3 presents estimates of the effect of PSA report access on how frequently prosecutors, defense attorneys, and judges explicitly invoke different concerns that pretrial release conditions could help to resolve (top rows) and discuss other imperatives that legally and/or normatively should inform pretrial conditions (bottom row) across hearings.

In the first set, I separately coded when courtroom actors discussed how pretrial conditions might mitigate flight risk specifically, rather than mitigate missed court appearances in general. The results indicate that, given the PSA, prosecutors more frequently initiate appeals to preventing future crime by 5 percentage points, while defense attorneys more often emphasize assuring court appearance as a target outcome by 9 percentage points.

in quick succession, judges can recall the defendant's previous scores and consider them as a lower bound for what the defendant's scores in the new case would be (scores usually worsen when new cases are opened). Accordingly, judges can readily infer scores and recommendations for some control group observations.

Figure 4.3: **Effect of PSA on goal-related discussions, by courtroom actor**



The latter may be surprising, but it is important to clarify that defense attorneys almost universally highlight assuring court appearance as the primary basis for pretrial decision-making in bids to secure more favorable (less restrictive) release conditions for their clients. Consistent with findings from the first section, separate PSA scores for future criminal activity and failures to appear make it easier for defense attorneys to selectively highlight low FTA scores or clean histories for increased attention and then to assert that those factors should be what informs the judge's decision. Defense attorneys for example, preface their bail requests with "given that our primary goal is to insure [the defendant]'s appearance in court" or conclude their recounting of their client's consistent history of making court dates with a reminder that "I think that the main purpose of bail is to insure his appearance in court."

Judges, meanwhile, more often highlight mitigating future crime and flight risk specifically as the goals of pretrial conditions when exposed to the PSA. Increases in discussions of future crime are often directly tied to the contents of the PSA itself, with judges drawing attention to defendants with high new criminal activity scores. The six percentage point increase in specific appeals to flight risk, meanwhile, appears to reflect subtle redirections prompted by defense attorneys more frequently asserting that likelihood of appearing for court should be the focal consideration in setting release conditions. When judges specifically discuss flight risk, they accept the general premise, consistent with state law, that release conditions should manage non-appearance risk. But they seek to re-integrate the mental model that they hold that non-appearance is informed by defendants' rational calculations of the likelihood and severity of punishment defendants face, factors that do not figure into the PSA. In response to defense attorney declarations that their clients are unlikely to miss court based on PSA results, judges more often layer on re-evaluations of those claims by reasoning through how much prison time or "exposure" the defendant faces if convicted. Because judges believe that a defendant's prospects for punishment create incentives for flight, judges add on specific discussions of flight risk to adjust the

PSA's evaluations of non-appearance risk, which they assert misses an important part of the picture.⁵ These added discussions sometimes favor fewer restrictions, as when judges conclude "I don't think there is a strong likelihood of a lengthy prison sentence in these cases, so I don't know if there is enough incentive for him to flee." But they often do contrast with PSA evaluations. Judges will state for example that cash bail is appropriate despite fairly low Failure to Appear scores as they observe that "there is significant prison exposure given the gravity of the offenses" or that "we do have a very serious charge here and the potential penalties are pretty significant." As such, these results suggest that judges take up some of the logic set forth by the PSA and further advanced by prosecutors and defense attorneys, but tweak the underlying constructs to accommodate their status quo approaches to decision-making.

Turning finally to the question of whether the PSA crowds out considerations of other imperatives, the results suggest that the PSA in fact leads to modest increases in how often defense attorneys and judges discuss whether and how proposed monetary and non-monetary conditions might be burdensome for defendants (by 5 and 2 percentage points respectively). Defense attorneys increasingly advocate for lower cash bail amounts on the basis of a defendant's ability to pay or for reducing the scope of location bans and no-contact orders to reduce the disruption such orders would present for a defendant's daily life. And judges respond by increasingly acknowledging these burdens, albeit to a more limited extent, as they reason through their decisions. Prosecutors, for their part, never raise the issue of burdens in control group or treatment group hearings.

In sum, these results paint a mixed picture of how much courtroom actors' depictions of the goals of pretrial release decisions come to align with the PSA. The first half of these results reflect partial take-up. When the PSA is provided, defense attorneys and prosecutors elevate the goals and concerns encoded in the PSA's logic that favor their own strategic in-

⁵The aim of this paper is not to evaluate whether judges use or understand the PSA "correctly" following official guidance. However, I note that the generalized non-appearance outcome that the PSA aims to predict already encompasses true flight. Thus, when judges add an accounting for true flight, they are in effect double counting.

terests, contributing to enforcing an analytical split between recidivism and non-appearance risk. Judges respond in turn by signaling consonant concerns, but reframing them to reconcile the PSA's constructions of target outcomes with their prevailing understandings. The second half of the results, meanwhile, suggest that, although the PSA implicitly endorses risk management as the primary basis for pretrial decision-making, access to PSA reports does not box out other important imperatives. Rather, the data suggest that it leads to modest improvements in the degree to which debates also take into account concerns about imposing unnecessary burdens on defendants.

Subsuming other goals under the banner of non-appearance: the case of drug addiction

The analysis thus far has focused on the impact of the PSA on the dominant goals of cash bail – assuring court appearance and public safety – the two aims recognized formally in bail statutes and highlighted in legal scholarship. In practice, however, courtroom discussions reveal ancillary goals that courtroom actors aim to achieve through cash bail and non-monetary conditions that are not recognized or endorsed by the PSA.

One such goal that appears regularly prior to the adoption of the PSA is ensuring the health of defendants experiencing drug addiction. This aim aligns with research on the expansive role that the criminal justice system assumes in the United States as an institution for managing the health of marginalized populations in the absence of a robust social safety net (Comfort, 2007; Lara-Millan, 2014).

Prior to the adoption of the PSA, judges publicly reasoned through how their decision-making is informed by a concern that the defendant might die from a drug overdose if released with insufficient conditions in place. Judges positioned the health and safety of the defendant as an objective alongside assuring court appearance and public safety. One judge in particular often spoke of his role in the manner of a case manager entrusted with structuring access to treatment services for individuals struggling with addiction.

One such example occurred with an eighteen-year old male defendant who was charged with drug possession. The allegation was a violation of the terms of the defendant's probation and a violation of his release conditions for other pending drug-related cases. The prosecutor in the case requested cash bail, noting that the defendant was being charged as a repeater, which substantially increased potential penalties. The public defender, meanwhile, requested a signature bond, highlighting that the defendant would likely continue to be monitored by his probation agent and that cash bail may get in the way of the defendant receiving in-patient treatment. Responding to these arguments, the commissioner stated:

What I struggle with is the fact that there are a lot of "ifs" in all of that, and if any of those ifs fall away, there is a good likelihood that Mr.[Defendant Surname] is going to be dead because he's overdosed and I want to make sure that that doesn't happen to the extent that I can do that, and I also want to make sure that he's at his court dates...I don't want to stand in the way of the treatment any more than anyone else, and I agree with that concept. I'm just trying to figure out how to make sure that that happens so that something doesn't slip through the cracks where we have a signature bond here and the hold gets lifted without any sort of a response... *my goal is to make sure that Mr.[Defendant Surname] stays alive, stays safe, and gets to court. Those are all things that are important to all of us* [emphasis added]... I'm going to set cash bail here in the amount of \$1,500.

In setting cash bail in this context, the commissioner aimed to stall the immediate release of the defendant until measures could be put in place to ensure that the defendant would be monitored and receive treatment. Notably, the judge highlights keeping the defendant safe and assuring court appearance as important *separate* objectives. Under the Constitution and state statutes, however, protecting the health of the defendant is not, a sanctioned basis for preventative detention. However, judges see their decisions as matters of life and death for some defendants. Indeed, in a similar drug possession case in which

the judge ordered a significant cash bail despite requests for signature bonds from both the prosecution and defense, the judge stated, "I'm frankly concerned if I order a signature bond I'm putting into place at least one piece that would be necessary to have his death warrant signed, and I'm not going to do it."

Among control group hearings that took place after PSA adoption, we observe more sporadic examples of elaborated discussions of protecting the defendant's safety in drug possession cases, though there remain examples of concern for defendants overdosing. In another case in which the public defender suggested that cash bail, as requested by the prosecutor, would get in the way of the defendant receiving treatment, the judge reasoned through similar concerns as in pre-adoption cases, but landed ultimately on issuing a signature bond:

Commissioner: I have two minds in this case, and this is a tougher situation, because when I'm looking at a case through the standard Criminal Justice System lens, this isn't an awful case for a signature bond or for signature bond with Bail Monitoring. The other side of this is when I look at these 3 complaints, each of which involve heroin overdoses, it isn't a standard Criminal Justice System case, it's a public health disaster... and the concern I have – and I'll just articulate it – is that if I do something like a signature bond or even a signature bond with Bail Monitoring I'd be concerned that I'm just signing your death warrant, and I know you're shaking your head no and you don't think that's the case, and I hope you're totally right.

Defendant: I just want to get better. I'm sorry.

Commissioner: And I want you to get better, but I want you to be there to get better and I'm concerned at some point something is going to happen and this is going to fall apart and you're not going to be here and you're not going to be able to participate and get the treatment.

Defendant: That's not going to happen.

Commissioner: But I also don't want you to sit in jail. I'm going to go with a signature bond. [Description of conditions]. So what I'm trying to say is you need to follow through with this. If you do, obviously the hope is that your life improves and that you stay safe during the course of all of that... If it feels like I'm beating up on you, I'm not doing that because I'm trying to be mean to you. I'm concerned about you and I want to make sure that you're going to stay safe.

Defendant: I appreciate that.

Here the judge, despite expressing misgivings, ultimately exercises restraint in the use of cash bail for preventative detention.

In the treatment group hearings, where the PSA was provided, we begin to observe a more substantial shift in how concern about defendant health manifests. While judges continue to express concern for defendants in some cases, protecting defendant health increasingly pivots from being discussed as a valuable end in itself to a means to another end: assuring court appearance. Defendants overdosing shifts from an outcome to be prevented in itself to being an inconvenience that would prevent court appearance. As one judge articulated in a drug possession case, "... we have to worry about whether she's going to end up dead and she can't make her court dates if she's dead."

In these instances, two dynamics come into play. First, the PSA, through delimiting particular outcomes that pretrial decision-making is intended to address appears to constrict possibilities for discussing alternate goals like preserving defendant safety. Second, judges come to view preventing non-appearance in the most literal terms, so much so that the potential death of the defendant and the moral weight of that outcome is diminished to the point that it is only understood as a risk factor for non-appearance.

This latter dynamic emerges especially clearly as local judges try to explain why they are departing from PSA signature bond recommendations for individuals with substance

use challenges. Judges cite the goal of preventing non-appearance, as endorsed by the tool, as the basis for departing from the PSA recommendation. This is illustrated in a hearing in which the prosecutor requested a cash bail, arguing “...the defendant overdosed on heroin and had to be revived using Narcan. He’s unable to follow Court orders at this time and that makes him a poor candidate for making court appearances.... Your Honor, homeless people who are addicted to heroin are especially unreliable to appear in court due to the inability to get notice and the fact that every time they use heroin they’re likely to die as the defendant nearly did when he OD’d [two days earlier].”

The public defender, meanwhile, invoked the PSA, which gave the defendant a Failure to Appear score of 3, a New Criminal Activity score of 3, and a recommendation for signature bond, to request a signature bond. The judge, however contested the recommendation:

You know, we’re just at the point where this can’t be a signature bond right now. I don’t want to put too fine a point on it, but dead guys don’t come to court, and there is a pretty strong likelihood that that’s going to happen if he’s released to the street on a signature bond without some other intervention here... I think the least amount I could set would be \$600.

While the judge in this instance imposes cash bail in a paternalistic manner similar to the pre-adoption hearings described earlier in this section, the rationale here is ensuring that the defendant is still alive to show up for the preliminary hearing. This stands in contrast to hearings that took place before adoption and hearings in the control group, in which judges cited protecting the defendant’s health as a valuable aim on its own.

4.6 Conclusion and Discussion

This chapter presents a mixed bag of findings that bring into view the uneasy trade-offs and ambiguities that the adoption of risk assessments presents for how defendants are evaluated in court proceedings. The results neither fully support the contention that risk assessments

usher in dystopian forms of decision-making nor claims that they help to align practices with normative ideals. They further highlight distinct sites and ways in which courtroom room actors alternately accept, embrace, or resist the predictive logic encoded in the PSA.

In the column of what might be considered a negative consequence of tool adoption is the PSA spurring courtroom actors to double down on factors that they traditionally considered but that have been determined to be poor predictors and/or biased – e.g. drug/alcohol use, community ties, and defendants’ interactions with law enforcement — and to re-assert heuristic estimates of flight risk based on the punishments associated with the as-yet unproven allegations against defendants. These results align with past research uncovering resistance to predictive algorithms in criminal justice (Christin, 2017; Werth, 2017), and suggest that exclusion of these considerations from the PSA may in fact backfire as they become more salient in practice when the PSA is provided.

At first glance, these results may also appear consistent with the common refrain that risk assessments can be pernicious because they allow decision-makers to do what they were going to do anyway, but now shielded by an aura of objectivity that makes it more difficult to challenge their decisions. However, there is a key difference here in that judges persist in what they were doing previously in these areas by explicitly spotlighting their own personal experiential expertise and highlighting what they perceive to be the shortcomings of the tool. At least with respect to the continued use of these traditional factors, judges do not hide behind the PSA as an added source legitimacy. Instead, the adoption of the PSA makes judges’ subjective or idiosyncratic perspectives more, rather than less, visible. Thus, while the PSA may shift discussions in directions that might be considered undesirable, the presence of the PSA does help to make those moves more readily apparent. Future research may examine if, in the long-run, this latter impact presents a potential lever for greater accountability.

Turning to what might be regarded as positive consequences of PSA adoption, the results indicate that the PSA encourages actors to devote greater attention to factors related to

a defendant's likelihood of making court appearances and to more often weigh the burdens that conditions may place on defendants. Departing from suggestions that actuarial prediction encourages a naive behaviorist view of human nature (Fourcade and Healy, 2017), hearings moreover more often feature discussions of the context behind negative events in a defendant's record. These shifts align with Wisconsin state statutes and legal scholarship asserting that monetary pretrial release conditions should only be used if deemed necessary to assure court appearance, rather than to prevent crime, and that cash bail amounts should not exceed what is needed to achieve that goal given a defendant's economic situation. They further take place amidst no significant negative shifts in most other topics related to a defendant's background and social context, which theoretical accounts might predict courtroom actors would become less interested in when an actuarial risk assessment is adopted.

The reason for the increased discussions of burdens is hard to deduce from my current data but may be valuable for future research to explore. The result is surprising as the PSA does not consider the issue, and we might expect that the PSA would encourage actors to calibrate decisions to risk with less regard for how those decisions might impact defendants. One potential explanation might be that, as the first two chapters illustrated, the PSA expands the set of defendants who come under suspicion as "risky." Increased talk of burdens might then simply reflect more cases in which potentially burdensome conditions are broached in the first place. An alternate explanation, meanwhile, might be that the PSA, given its close association with pretrial reform efforts, might indirectly serve as an object that reminds courtroom actors of key issues highlighted by the broader movement, including critiques that judges over use cash bail and fail to calibrate amounts to a defendant's ability to pay. Judges' responses may thus be reflective of their efforts to pay deference to these concerns.

In interpreting this set of positive consequences, I additionally acknowledge that the present analyses reflect patterns in talk and do not offer a direct view into what courtroom

actors are paying attention to or thinking. Nonetheless, the setting of criminal courts is one in which talk and debate is the primary medium through which decision-making takes place. This is particularly the case in low-information environments like in bail hearings where the facets of defendants that defense attorneys and prosecutors bring up delimit and frame a significant portion of the information that judges have access to about defendants. Judges cannot pay attention to information that they do not have. All of the local judges in Dane County, moreover, routinely think out loud in hearings, noting all the factors they are weighing and talking through their thought process before they arrive at a decision. Consequently, while talk may be cheap in some settings, I would argue that the present context is one in which talk is comparatively meaningful and informative.

Finally, the results also identify a set of PSA consequences with more ambiguous valences. First, the PSA encourages judges to elevate assuring court appearance as a target outcome of pretrial conditions, a change that has been posited to facilitate judicial restraint in the use of cash bail (Gouldin, 2016). Doing so, however, can lead to dehumanizing outcomes. As exemplified by the case of how judges approach defendants experiencing drug addiction, the PSA appears to prompt judges to treat previously valued ends – like preserving a defendant’s life – as mere means to achieve the ultimate end of ensuring defendants come to court.

Second, the PSA presents a dynamic model of human behavior that courtroom actors take up in their arguments. On the one hand, this model, with its two-year cut-off, reduces the length of the shadow that negative entries in a defendant’s history cast on the defendant’s present and future. On the other hand, the modeled approach appears to prompt stronger punitive reactions to recent events. We might posit that this could draw more defendants into deeper, more punishing entanglements with the criminal justice system precisely when they are down.

These ambiguous consequences further underscore ways in which tools like the PSA can generate a host of impacts on courtroom debates and evaluation practices that place

into tension different procedural outcomes one might wish to improve. While the specific ways that risk assessment algorithms shape local practices are likely to vary from jurisdiction to jurisdiction, these results suggest that these types of tools, for better or worse, are unlikely to exert a hegemonic influence over how local actors approach decision-making. As these types of practical trade-offs are likely to appear anytime algorithmic decision aids are adopted, researchers, policymakers, and community leaders should be attentive to how their on-the-ground impact is unlikely to match the sweeping apocalyptic forecasts of their strongest critics, nor the lofty optimistic declarations of their strongest proponents.

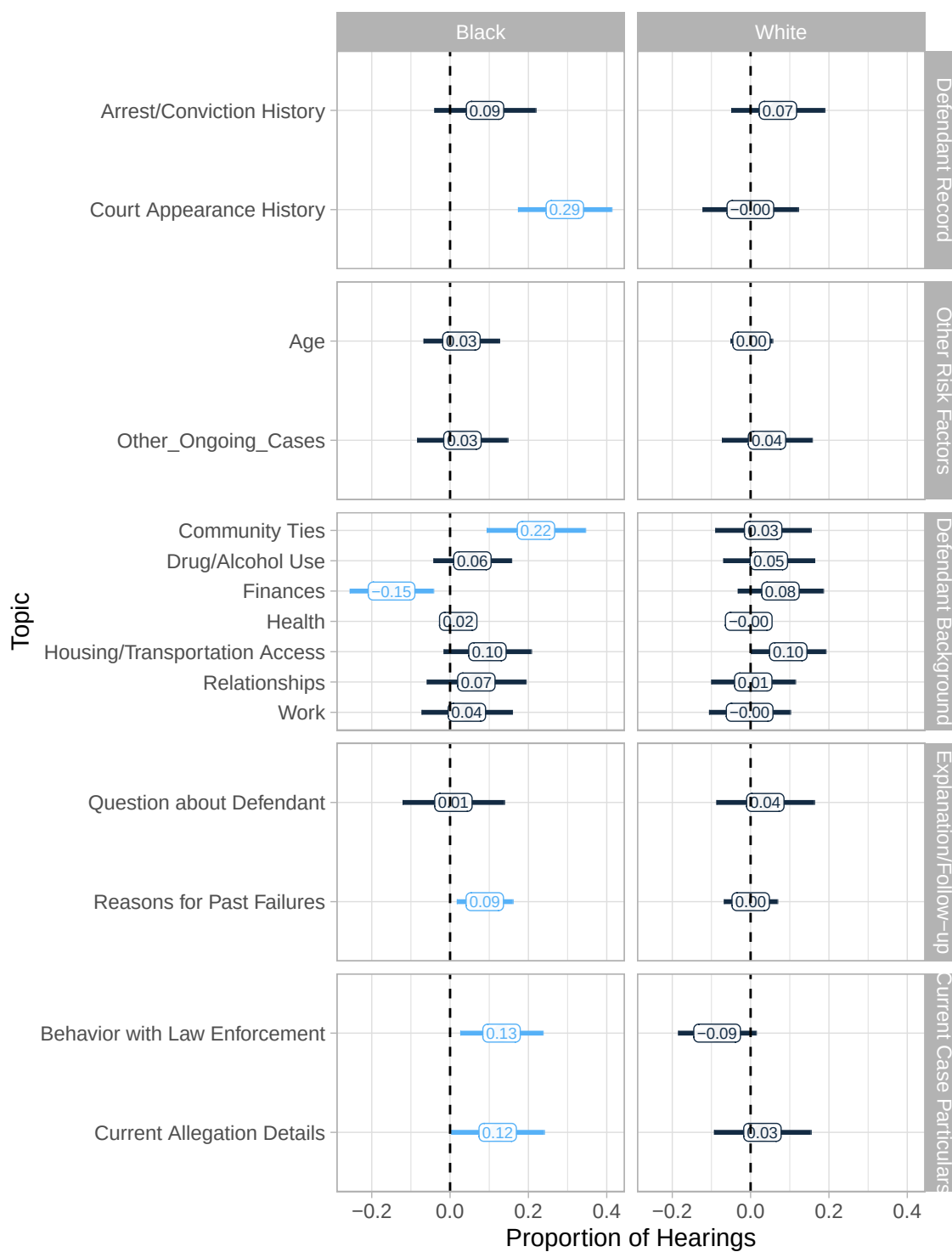
4.7 Appendix

4.7.1 Heterogeneity by Defendant Race and Sex

Race

Figure 4.4 examines whether the effect of the PSA on topics discussed varies by defendant race. Overall, the PSA appears to have had wide ranging effects on hearings involving Black defendants but comparatively little impact on hearings involving white defendants. With the PSA, hearings for Black defendants more often discussed the defendant's court appearance history (by 29 percentage points), ties to the community (by 22 percentage points), and the context of past missed court appearances (by 9 percentage points), and less often discussed the defendant's finances (by 15 percentage points). The PSA further increased discussions of the particulars of Black defendants' present cases, including the nature of the current allegations (by 12 percentage points) and the defendant's behavior with the law enforcement (by 13 percentage points).

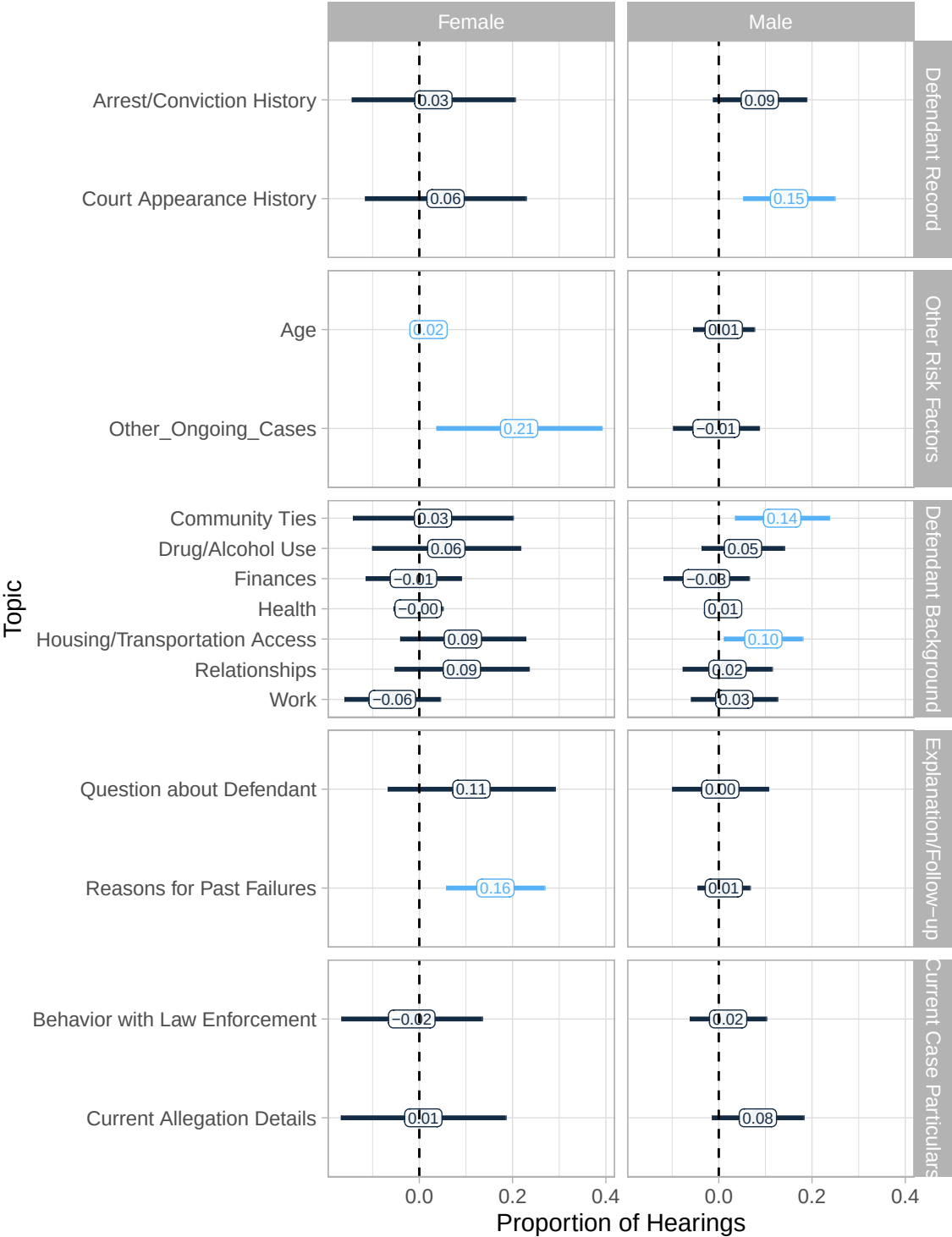
Figure 4.4: Effect of PSA on topics discussed, by race



Sex

Figure 4.5 examines whether the effect of the PSA on topics discussed varies by defendant sex. It is important to note that I have a limited sample of female defendants and as such cannot as precisely estimate treatment effects for this group. Accordingly, I will only highlight the most salient differences by sex. For female defendants but not male defendants, the PSA substantially increased discussions of the context behind past missed court dates (by 16 percentage points) and the defendant's other ongoing cases (by 21 percentage points). For male defendants but not female defendants, meanwhile, the PSA substantially increased discussions of community ties (by 14 percentage points), past court appearances (by 15 percentage points), and likelihood of making court appearances in the future (by 15 percentage points).

Figure 4.5: Effect of PSA on topics discussed, by sex



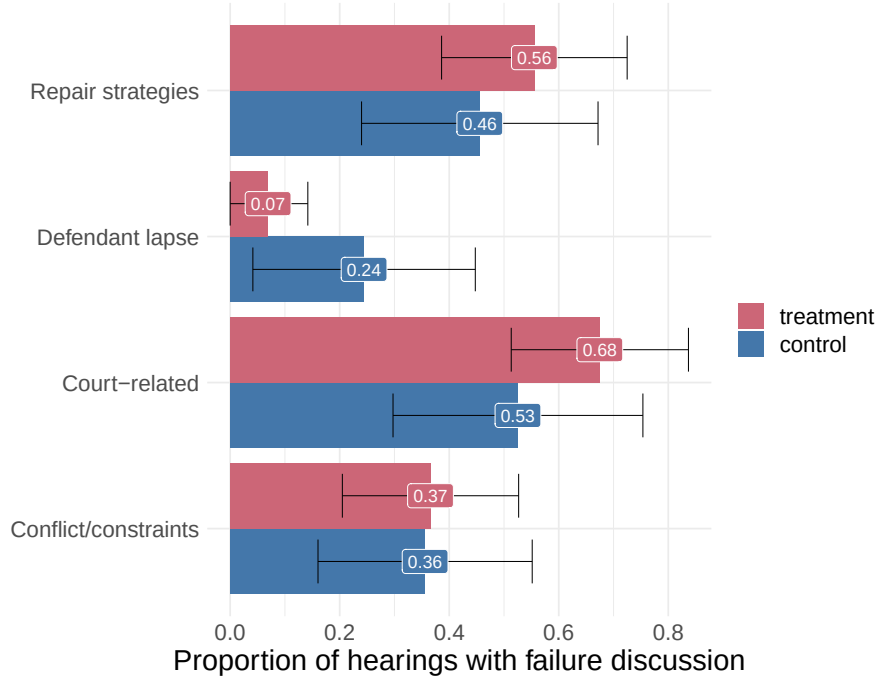
4.7.2 Explanations of Defendant History: Contextual Factors Highlighted and Judge Responses

As observed in the main results, access to the PSA increased discussions of the reasons behind a defendant's past missed court appearances. In this Appendix section, I further explore, among hearings where such discussion take place, how the PSA shifts what types of contextual factors are highlighted in discussions. Given small size issues, these supporting analyses are meant to uncover suggestive patterns rather than present definitive conclusions.

Figure 4.6 depicts the distribution of contextual factors highlighted in the treated and control groups. Contextual factors are sorted into one of four buckets:

- **Conflict and constraints:** conflicts with scheduled dates, including work, health emergencies, deaths of loved ones, care responsibilities, and lack of transportation.
- **Court-related:** these factors are related to informational gaps, such as not receiving notice of a court date, being unaware of a warrant, confusion about locations and times, and other bureaucratic mixups that keep a defendant from appearing for court. It also includes missed court due to incarceration.
- **Defendant lapse:** factors related to lapses on the part of the defendant, including fear of arrest, forgetting to attend, or accidental infractions
- **Repair strategies:** given a missed court date or condition violation, whether defendants take proactive steps to rectify the situation, such as by making arrangements with the court to reschedule, turning themselves in, and contacting their attorneys.

Figure 4.6: **Distribution of factors covered in explanations of past failures**



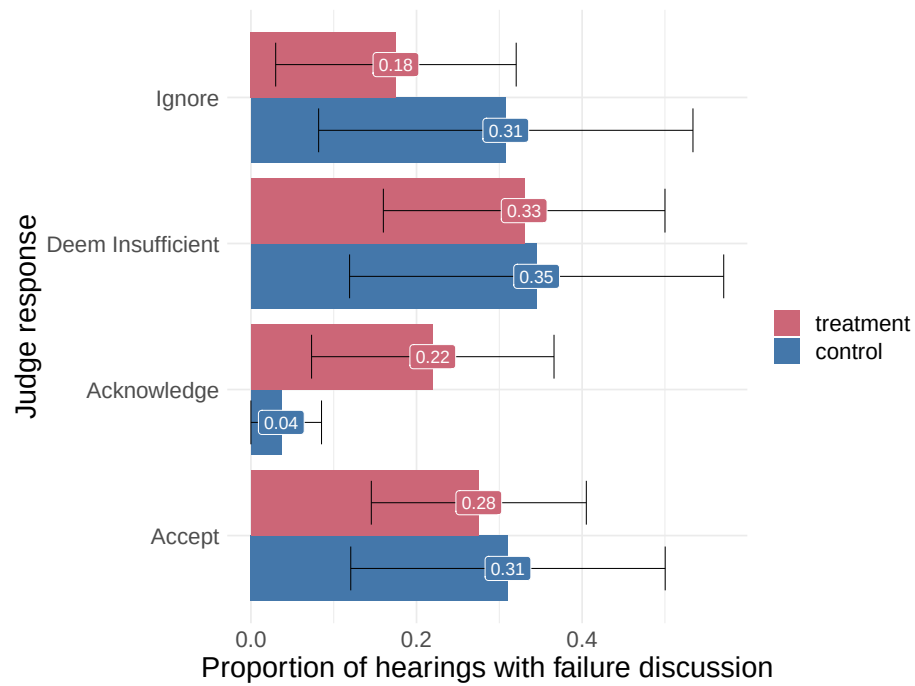
The results offer suggestive evidence of an increase in discussions of court-related factors and a decrease of discussions of defendant lapses. The increased emphasis on the former is notable given that they can absolve defendants entirely and place blame for missed appearances on courts themselves. Decreased discussions of a defendant's lapses, meanwhile, may be due to such factors becoming more deeply entrenched as the assumed reason for a missed court date.

Figure 4.7 further compares the final outcomes of failure discussions. It examines how often judges ignore explanations of failures; deem explanations insufficient to support laxer conditions; acknowledge the explanation as a mitigating circumstance but without fully accepting it; and accept the explanation and impose laxer conditions accordingly.⁶ The results indicate that judges accept explanations and deem them insufficient at comparable rates across the treatment and control conditions. However, they tend to acknowledge and give explanations greater credit in the treatment condition than in the control condition, where

⁶Here I have coded both implicit and explicit acceptance of the explanation. I considered instances where the judge imposed lax conditions, consistent with the explanation as a mitigating circumstance, directly following an explanation as implicit acceptance.

they more often ignore explanations altogether when delivering their decisions.

Figure 4.7: **Distribution of the outcomes of failure discussions**



Chapter 5

Conclusion

This dissertation examined how a pretrial risk assessment algorithm shaped practices on the ground in a criminal court. Expanding the analytic lens adopted in past research, it investigated the tool's impact in a range of areas, from decision outcomes like the use of cash bail and bail monitoring, to processes and procedures, like how courtroom actors understand the goals of decision-making and how they talk to and about defendants. The study was built on the foundation of a randomized controlled trial and combined quantitative and qualitative data. This enabled me to better isolate and unpack the consequences of the risk assessment, testing how well they aligned with theoretical expectations and commonly raised policy concerns.

The results indicated areas where practices shifted little in response to risk assessment and areas where the risk assessment amplified existing tendencies in local decision-making approaches. They also pointed to ways in which the risk assessment shifted the dynamics among courtroom actors, particularly as it served as a resource that helped prosecutors and judges make claims about defendant riskiness and shore up their positions. Overall, though, the study's mixed findings testify to the potential limits of technocratic interventions in transforming complex, fraught sites of institutional decision-making. Even as the PSA steered outcomes in some areas, it brought about neither the wide-ranging improvements

to the administration of pretrial justice that proponents of these tools hope, nor the broad distortions of justice that critics have feared. The trade-offs and ambiguities that such tools introduce present fraught choices. Ultimately, whether adopting a risk assessment is a good or bad path may be up to individual communities to evaluate.

5.0.1 Limitations, Implications, and Directions for Future Research

This dissertation raises additional questions and points to several directions for future research. To start, findings in Chapters 2 and 3 suggest that the PSA increases the restrictiveness of the conditions that judges impose on certain subsets of defendants. For those who reject cash bail and/or pretrial detention as illegitimate intrusions on the liberty of individuals, these results in themselves may be interpreted as evidence that tools like the PSA are harmful and should not be adopted. Others who have taken a more consequentialist perspective, however, may further ask whether these shifts in decision outcomes lead to improvements in ultimate outcomes that society may value, such as court appearances and public safety, and reduce potential racial disparities. This latter question is beyond the scope of the present project, but is the central research question in the broader Access to Justice Lab study. So far, preliminary research from Dane County does not find evidence of substantial differences in these outcomes (Greiner et al., 2020) and the results from other sites are mixed (Stevenson, 2018; Redcross et al., 2019; Austin and Ware, 2020). Scholars have moreover suggested that U.S. jurisdictions tend to, at baseline, over-use cash bail and over-detain (Dobbie et al., 2018; Ouss and Stevenson, 2020; Stevenson and Mayson, 2021). More research is needed, however, to conclude whether more restrictive conditions yield benefits, or merely impose unnecessary costs on defendants and their communities.

Another limitation of the present research is that it consists of a single-site study. The impact of risk assessments is likely to be moderated by myriad legal, political, and organizational factors that differ across sites. To highlight one such factor, future research may consider how the impact of a tool like the PSA might differ if implemented in more

conservative or moderate jurisdictions. On the one hand, we might expect that courts in more conservative areas face additional external pressures or that local courtroom actors might hold more cynical beliefs about defendants that lead the PSA to produce even more punitive decision outcomes. We might, for instance, observe even stronger asymmetries in the effect of the PSA, with high risk scores causing large increases in restrictive conditions and low risk scores dismissed for underestimating risk. On the other hand, it may also be the case that the impact of the tool may be less punitive than in progressive Dane County. As I noted in the introduction, there was a shared narrative among those who worked in Dane County that the county, while not perfect, was softer and less punitive than in other parts of Wisconsin. This understanding may have shaped responses to the PSA and may have in fact contributed to some of the asymmetries observed in Chapters 2 and 3. Comparative case studies, especially ones that compare the experiences of different jurisdictions governed by the same state laws, may be valuable for untangling the role of local politics further.

Two other factors idiosyncratic to the Dane County case include the presence of the randomized controlled trial and the manner in which the tool came to be adopted locally. The randomized controlled trial visibly subjected local practices to an audit and judges knew that their decisions were going to be tracked and evaluated. Given this context, we may expect take-up of the tool to be high and the impacts of the tool to be large compared to other settings where no formal accountability infrastructure was in place. The fact that the tool was adopted voluntarily, with the engagement and support of a broad cross section of local stakeholders may have further enhanced the PSA's impact. Commissioners themselves were involved in some aspects of local planning for the tool. One in fact told me that he was personally interested in the results of the RCT because he said he was a nerd who was curious if the PSA actually worked or not. This context, however, differs from that in other jurisdictions, such as counties in Utah and New Jersey, where decisions about adopting a risk assessment were made at the state level and mandated whether or not local

jurisdictions agreed. We might think that the impact of the tool may be substantially more muted in those settings, though additional research is needed to evaluate whether that is the case.

The results of the present research may be further enriched by interviews with local prosecutors, defense attorneys, and commissioners. While administrative data and court transcripts reveal patterns in behavior executed in a public setting, they do not offer insights into why courtroom actors said or did the things that they do. Interviews can offer insights into the motivations and meanings behind local practices, as well as courtroom actors' subjective understandings of how the tool affected their work and their relationships with colleagues. In the next phase of this research, I plan to conduct these interviews.¹

I also plan to further explore heterogeneity in the PSA's impact by the intersections of defendant race, gender, age, and class. Substantial bodies of scholarship document the historical and ongoing processes that generate inequalities in criminal justice outcomes along these axes of social difference (Western, 2006; Wakefield and Uggen, 2010; Van Cleve, 2016). Investigating these heterogeneities would particularly help to further address pressing questions about whether the use of risk assessments reduces or exacerbates racial disparities.

More broadly, this dissertation offers further evidence to underscore the importance of examining how decision-makers interact with algorithmic decision aids in social context. This dissertation joins a rapidly growing body of work aiming to understand in what ways these interactions accentuate or offset biases built into algorithms, or even introduce biases anew (Albright, 2019; Green and Chen, 2019; Eckhouse et al., 2019; Goel et al., 2021; Joyce et al., 2021). Adding to past work, this dissertation further moves away from investigating how decision aids intervene on atomized individuals and toward exploring how they intervene on groups of actors connected to each other via ongoing relationships of dependency and authority. This latter focus has become more commonplace in studies of how

¹This was delayed due to Covid-19.

workplace algorithms affect worker-manager relations (Shestakofsky, 2017; Kellogg et al., 2020), but has largely been absent in scholarship on algorithmic decision aids targeted at professionals to date.

Attention to the social contexts in which algorithmic decisions are used can further ground investigations of how decision-makers respond to such tools in the practical constraints that they face. For instance, studies of decision aids often center around assessing compliance with tool recommendations, with instances of decision-makers departing from recommendations interpreted as exercises of discretion. While this may be analytically true in the abstract, it may not align with decision-makers' contextual understandings of their actions. In the case of pretrial decision-making, for instance, prosecutor and defender requests ordinarily circumscribe judges' choice sets. Within the social context of bail hearings, departing from the joint recommendations of those parties, with whom judges have ongoing working relationships, is a far more socially salient exercise of discretion than departing from a risk assessment. To the extent that risk assessments make it possible for judges to credibly depart from those requests, then, we might think that risk assessments actually enhance rather than reduce the discretion of judges. Since much professional decision-making takes place in lateral teams or supervisory structures, this example illustrates the value of understanding and interpreting responses to algorithmic decision aids within the context of competing *de facto* defaults and local social constraints.

Bibliography

- Abbott, Andrew and Angela Tsay. 2000. "Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect." *Sociological Methods & Research* 29:3–33.
- Adams, Ericka B, Elsa Y Chen, and Rosella Chapman. 2017. "Erasing the mark of a criminal past: Ex-offenders' expectations and experiences with record clearance." *Punishment & Society* 19:23–52.
- Albright, Alex. 2019. "If You Give a Judge a Risk Score: Evidence from Kentucky Bail Decisions." Working paper.
- Angwin, Julia and Jeff Larson. 2016. "Machine Bias." *ProPublica* .
- Arnold, David, Will Dobbie, and Crystal S. Yang. 2018. "Racial Bias in Bail Decisions." *The Quarterly Journal of Economics* 133:1885–1932.
- Arnold Ventures. 2019. "Public Safety Assessment FAQs ("PSA 101")."
- Austin, James and Roger Ocker. 2021. "Dane County Jail Race and Ethnicity Disparity Analysis." Technical report.
- Austin, James and Wendy Ware. 2020. "Why Bail Reform is Safe and Effective: The Case of Cook County." *SSRN Electronic Journal* .
- 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, Ruha Benjamin, John Bowers, Meredith Broussard, Joy Buolamwini, Sasha Constanza-Chock, Kate Crawford, Karthik Dinakar, Colin Doyle, Timnit Gebru, and et al. 2019. "Technical Flaws of Pretrial Risk Assessments Raise Grave Concerns."
- 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* pp. 78–108.
- Bechtel, Kristin, Alexander M. Holsinger, Christopher T. Lowenkamp, and Madeline J. Warren. 2017. "A Meta-Analytic Review of Pretrial Research: Risk Assessment, Bond Type, and Interventions." *American Journal of Criminal Justice* 42:443–467.
- Benjamin, Ruha. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Medford, MA: Polity Press.
- Berman, Elizabeth Popp and Daniel Hirschman. 2018. "The Sociology of Quantification: Where Are We Now?" *Contemporary Sociology* 47:257–266.
- Bogen, Miranda and Aaron Rieke. 2018. "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias." Technical report, Upturn, Washington, DC.
- Brayne, Sarah. 2017. "Big Data Surveillance: The Case of Policing." *American Sociological Review* 82:977–1008.
- Brayne, Sarah. 2018. "The Criminal Law and Law Enforcement Implications of Big Data." *Annual Review of Law and Social Science* 14:293–308.
- Brayne, Sarah. 2020. *Predict and Surveil: Data, Discretion, and the Future of Policing*. New York: Oxford University Press.
- Brayne, Sarah and Angele Christin. 2020. "Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts." *Social Problems* p. spaa004.

- Burrell, Jenna and Marion Fourcade. 2021. "The Society of Algorithms." *Annual Review of Sociology* 47:23.1–23.25.
- Campbell, Donald T. 1957. "Factors relevant to the validity of experiments in social settings." *Psychological Bulletin* 54:297–312.
- Charmaz, Kathy. 2006. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. SAGE.
- Cheliotis, Leonidas K. 2006. "How iron is the iron cage of new penology?: The role of human agency in the implementation of criminal justice policy." *Punishment & Society* 8:313–340.
- Chouldechova, Alexandra. 2016. "Fair prediction with disparate impact: A study of bias in recidivism prediction instruments." *arXiv:1610.07524 [cs, stat]* arXiv: 1610.07524.
- 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, Angele. 2017. "Algorithms in practice: Comparing web journalism and criminal justice." *Big Data & Society* 4:1–14.
- Christin, Angele. 2020. "What Data Can Do: A Typology of Mechanisms." *International Journal of Communication* 14:20.
- Comfort, Megan. 2007. "Punishment Beyond the Legal Offender." *Annual Review of Law and Social Science* 3:271–296.
- Corbett-Davies, Sam and Sharad Goel. 2018. "The Measure and Mismeasure of Fairness: A Critical Review of Fair Machine Learning." *arXiv:1808.00023 [cs]* arXiv: 1808.00023.

- County of Dane, WI. 2018. “Frequency of Signature Bonds in Dane County Criminal Cases: 2012-2016.” Technical report, Madison, WI.
- Cowgill, Bo. 2018. “The Impact of Algorithms on Judicial Discretion: Evidence from Regression Discontinuities.” Working paper.
- 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. 2018. “The Intuitive-Override Model: Nudging Judges Toward Pretrial Risk Assessment Instruments.” *SSRN Electronic Journal* .
- 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.
- Digard, Leon and Elizabeth Swavola. 2019. “Justice Denied: The Harmful and Lasting Effects of Pretrial Detention.” Technical report, Vera Institute of Justice, New York.
- 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.

- Doleac, Jennifer L. 2018. "Study after study shows ex-prisoners would be better off without intense supervision." Brookings Institution Blog.
- Eckhouse, Laurel, Kristian Lum, Cynthia Conti-Cook, and Julie Ciccolini. 2019. "Layers of Bias: A Unified Approach for Understanding Problems With Risk Assessment." *Criminal Justice and Behavior* 46:185–209.
- Eisenstein, James and Herbert Jacob. 1977. *Felony Justice: An Organizational Analysis of Criminal Courts*. Boston: Little, Brown, and Co.
- 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 Mitchell L. Stevens. 2008. "A Sociology of Quantification." *European Journal of Sociology / Archives Européennes de Sociologie* 49:401–436.
- 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.
- Fagan, Jeffrey and Tom R. Tyler. 2005. "Legal Socialization of Children and Adolescents." *Social Justice Research* 18:217–241.
- Feeley, Malcolm M. and Jonathan Simon. 1992. "The New Penology: Notes on the Emerging Strategy of Corrections and Its Implications." *Criminology* 30:449–474.
- Fishbane, Alissa, Aurelie Ouss, and Anuj K. Shah. 2020. "Behavioral nudges reduce failure to appear for court." *Science* 370.

- Flores, Anthony W and Kristin Bechtel. 2016. "False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks."." *Federal Probation* 80:38–46.
- Fourcade, Marion and Kieran Healy. 2013. "Classification situations: Life-chances in the neoliberal era." *Accounting, Organizations and Society* 38:559–572.
- Fourcade, Marion and Kieran Healy. 2017. "Seeing like a market." *Socio-Economic Review* 15:9–29.
- Gabadinho, Alexis, Gilbert Ritschard, Nicolas S. Müller, and Matthias Studer. 2011. "Analyzing and Visualizing State Sequences in R with TraMineR." *Journal of Statistical Software* 40:1–37.
- Garland, David. 2001. *The Culture of Control: Crime and Social Order in Contemporary Society*. Chicago, IL: University of Chicago 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*.
- Gorman, Elizabeth H. and Rebecca L. Sandefur. 2011. "'Golden Age,' Quiescence, and Revival: How the Sociology of Professions Became the Study of Knowledge-Based Work." *Work and Occupations* 38:275–302.
- 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. 2019. "Disparate Interactions: An Algorithm-in-the-Loop Analysis of Fairness in Risk Assessments." In *Proceedings of the Conference on Fairness, Accountability, and Transparency - FAT* '19*, pp. 90–99, Atlanta, GA, USA. ACM Press.
- 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.
- Greiner, D. James and Andrea Matthews. 2016. "Randomized Control Trials in the United States Legal Profession." *Annual Review of Law and Social Science* 12:295–312.
- Griffin, Christopher L. 2016. "Dane County PSA Study Analysis Plan."
- Hacking, Ian. 1990. *The Taming of Chance*. Cambridge, England: Cambridge University Press.
- Haney, Lynne. 1996. "Homeboys, Babies, Men in Suits: The State and the Reproduction of Male Dominance." *American Sociological Review* 61:759–778.
- Hannah-Moffat, Kelly. 2013a. "Actuarial Sentencing: An "Unsettled" Proposition." *Justice Quarterly* 30:270–296.
- Hannah-Moffat, Kelly. 2013b. "Punishment and Risk." In *The SAGE Handbook of Punishment and Society*, pp. 129–151.
- 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.
- Hirschi, Travis and Michael Gottfredson. 1983. "Age and the Explanation of Crime." *American Journal of Sociology* 89:552–584.

- Hirschman, Daniel and Emily Adlin Bosk. 2020. "Standardizing Biases: Selection Devices and the Quantification of Race." *Sociology of Race and Ethnicity* 6:348–364.
- Hyatt, Jordan M. and Geoffrey C. Barnes. 2014. "An Experimental Evaluation of the Impact of Intensive Supervision on the Recidivism of High-Risk Probationers." *Crime & Delinquency* .
- James, Doris J. 2004. "Profile of Jail Inmates, 2002." Special Report, Bureau of Justice Statistics.
- 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* 7:2378023121999581.
- Kellogg, Katherine C., Melissa A. Valentine, and Angèle Christin. 2020. "Algorithms at Work: The New Contested Terrain of Control." *Academy of Management Annals* 14:366–410.
- Kemshall, Hazel and Mike Maguire. 2001. "Public Protection, Partnership and Risk Penalty: The Multi-Agency Risk Management of Sexual and Violent Offenders." *Punishment & Society* 3:237–264.
- Kilgour, Lauren. 2020. "Covid-19 has led to a worrisome uptick in the use of electronic ankle monitors." *MIT Technology Review*
- 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.

- 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.
- Knox, Dean, Will Lowe, and Jonathan Mummolo. 2020. "Administrative Records Mask Racially Biased Policing." *American Political Science Review* 114:619–637.
- Kofman, Ava. 2019. "Digital Jail: How Electronic Monitoring Drives Defendants Into Debt." *The New York Times Magazine* .
- Kohler-Hausmann, Issa. 2013. "Misdemeanor Justice: Control without Conviction." *American Journal of Sociology* 119:351–393.
- Lageson, Sarah Esther. 2016. "Found Out and Opting Out: The Consequences of Online Criminal Records for Families." *The ANNALS of the American Academy of Political and Social Science* 665:127–141.
- Lara-Millan, Armando. 2014. "Public Emergency Room Overcrowding in the Era of Mass Imprisonment." *American Sociological Review* 79:866–887.
- Lavigne, Michele, Bonnie Hoffman, and Renee Spence. 2018. "The Wisconsin Bail Manual." Technical report, National Association of Criminal Defense Lawyers, Washington, DC.
- Lum, Kristian, Chesa Boudin, and Megan Price. 2020. "The impact of overbooking on a pre-trial risk assessment tool." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 482–491, Barcelona Spain. ACM.
- Lynch, Mona. 1998. "Waste Managers? The New Penology, Crime Fighting, and Parole Agent Identity." *Law & Society Review* 32:839–870.
- Lynch, Mona. 2019. "The Narrative of the Number: Quantification in Criminal Court." *Law & Social Inquiry* 44:31–57.

- Mayson, Sandra. 2019. "Bias In, Bias Out." *The Yale Law Journal* 128:2122–2473.
- McPherson, Chad Michael and Michael Sauder. 2013. "Logics in Action Managing Institutional Complexity in a Drug Court." *Administrative Science Quarterly* 58:165–196.
- Menefee, Michael R. 2018. "The role of bail and pretrial detention in the reproduction of racial inequalities." *Sociology Compass* 12:e12576.
- Mitchell, Shira, Eric Potash, Solon Barocas, Alexander D'Amour, and Kristian Lum. 2021. "Algorithmic Fairness: Choices, Assumptions, and Definitions." *Annual Review of Statistics and Its Application* 8:141–163.
- Myrick, Amy. 2013. "Facing Your Criminal Record: Expungement and the Collateral Problem of Wrongfully Represented Self." *Law & Society Review* 47:73–104.
- Norman, Don. 1988. *The Design of Everyday Things*. Basic Books.
- O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown/Archetype.
- Orlikowski, Wanda J. 2000. "Using technology and constituting structures: A practice lens for studying technology in organizations." *Organization Science* 11:404–428.
- Ouss, Aurelie and Megan T. Stevenson. 2020. "Bail, Jail, and Pretrial Misconduct: The Influence of Prosecutors." SSRN Scholarly Paper ID 3335138, Social Science Research Network, Rochester, NY.
- PJI. 2017. "The State of Pretrial Justice in America." Technical report, Pretrial Justice Institute, Baltimore, Maryland.
- PJI. 2020. "The Case Against Pretrial Risk Assessment Instruments." Technical report, Pretrial Justice Institute, Baltimore, MD.

- Raghavan, Manish, Solon Barocas, Jon Kleinberg, and Karen Levy. 2020. "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 469–481, Barcelona Spain.
- Redcross, Cindy, Brit Henderson, Luke Miratrix, and Erin Valentine. 2019. "Evaluation of Pretrial Justice System Reforms That Use the Public Safety Assessment: Effects in Mecklenburg County, NC." Technical report, MDRC.
- Reiman, Jeffrey and Paul Leighton. 2016. *The Rich Get Richer and the Poor Get Prison: Ideology, Class, and Criminal Justice*. New York, NY: Routledge.
- Robinson, David G and Logan Koepke. 2019. "Civil Rights and Pretrial Risk Assessment Instruments." Technical report, Upturn.
- Rocque, Michael, Chad Posick, and Justin Hoyle. 2016. "Age and Crime." In *The Encyclopedia of Crime & Punishment*, pp. 1–8. John Wiley & Sons, Inc.
- Ross, Lee and Richard Nisbett. 2011. *The Person and the Situation: Perspectives of Social Psychology*. London: Pinter & Martin Ltd, 2 edition edition.
- Sandefur, Rebecca L. 2015. "Elements of Professional Expertise: Understanding Relational and Substantive Expertise through Lawyers' Impact." *American Sociological Review* 80:909–933.
- Shestakofsky, Benjamin. 2017. "Working Algorithms: Software Automation and the Future of Work." *Work and Occupations* 44:376–423.
- Simon, Jonathan. 1988. "The Ideological Effects of Actuarial Practices." *Law & Society Review* 22:30.
- Starr, Sonja B. 2014. "Evidence-Based Sentencing and the Scientific Rationalization of Discrimination." *Stanford Law Review* 66:803–72.

- Steffensmeier, Darrell, Jeffery Ulmer, and John Kramer. 1998. "The Interaction of Race, Gender, and Age in Criminal Sentencing: the Punishment Cost of Being Young, Black, and Male." *Criminology* 36:763–798.
- 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." Working paper.
- Stevenson, Megan T and Christopher Slobogin. 2018. "Algorithmic Risk Assessments and the Double-Edged Sword of Youth." *Washington University Law Review* 96:27.
- Stith, Kate and Jose A. Cabranes. 1998. *Fear of Judging: Sentencing Guidelines in the Federal Courts*. Chicago: University of Chicago Press.
- Swaner, Rachel, Cassandra Ramdath, Andrew Martinez, Josephine Hahn, and Sienna Walker. 2018. "What Do Defendants Really Think? Procedural Justice and Legitimacy in the Criminal Justice System." Technical report, Center for Court Innovation, New York.
- 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.
- Tonry, Michael H. 1996. *Sentencing Matters*. Oxford University Press.
- Ulmer, Jeffery T. 2019. "Criminal Courts as Inhabited Institutions: Making Sense of Difference and Similarity in Sentencing." *Crime and Justice* 48:483–522.
- US Census Bureau. 2019. "Quick Facts: Dane County, Wisconsin." Technical report.
- Van Cleve, Nicole Gonzalez. 2016. *Crook County: Racism and Injustice in America's Largest Criminal Court*. Stanford University Press, 1 edition.

- Van Cleve, Nicole Gonzalez and Lauren Mayes. 2015. "Criminal Justice Through "Color-blind" Lenses: A Call to Examine the Mutual Constitution of Race and Criminal Justice." *Law & Social Inquiry* 40:406–432.
- VanNostrand, Marie and Gena Keebler. 2009. "Pretrial Risk Assessment in the Federal Court." *Federal Probation* 73:26.
- Wakefield, Sara and Christopher Uggen. 2010. "Incarceration and Stratification." *Annual Review of Sociology* 36:387–406.
- Weaver, Vesla M. and Amy E. Lerman. 2010. "Political Consequences of the Carceral State." *American Political Science Review* 104:817–833.
- Werth, Robert. 2017. "Individualizing Risk: Moral Judgement, Professional Knowledge and Affect in Parole Evaluations." *The British Journal of Criminology* 57:808–827.
- Werth, Robert. 2019. "Theorizing the Performative Effects of Penal Risk Technologies: (Re)producing the Subject Who Must Be Dangerous." *Social & Legal Studies* 28:327–348.
- Western, Bruce. 2006. *Punishment and Inequality in America*. New York, NY: Russell Sage Foundation.
- Wiseman, Samuel R. 2014. "Pretrial Detention and the Right to Be Monitored." *The Yale Law Journal* p. 62.
- 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.