Received: 17 June 2024 | Revised: 8 July 2025 | Accepted: 8 July 2025

DOI: 10.1111/1745-9125.70033

ORIGINAL ARTICLE

CRIMINOLOGY

The role of case management in misdemeanor prosecution

Lindsay Graef[1] | **Aurelie Ouss**[2]

[1]Carey Law School, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[2]Department of Criminology, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence
Lindsay Graef, Carey Law School, University of Pennsylvania, 3501 Sansom Street, Philadelphia, PA 19104, USA.
Email: lgraef@law.upenn.edu

Replication Data and Code
Replication data and code are available at https://doi.org/10.7910/DVN/KX7VIV.

Funding information
Arnold Ventures

Abstract

Despite increasing attention to prosecutors' role in shaping criminal justice outcomes, there is limited empirical research on what prosecutors do. While most theories of prosecutorial discretion emphasize overarching goals related to justice and safety, our paper shifts the focus toward the practical realities of the job, particularly in the lower courts. We propose a case management model of prosecutorial discretion, grounded in analyses of misdemeanor cases handled by the Philadelphia District Attorney's Office. Importantly, our dataset includes hearing-level prosecutor identifiers and decisions. By adapting established benchmarking methods, we identify credible counterfactuals and analyze how prosecutor actions influence case and defendant outcomes. We find that prosecutors vary in their case management skills, which are an important driver of conviction. Defendants fare better under prosecutors who secure fewer convictions, indicating negative consequences of conviction, but prosecutors who secure fewer convictions tend to be less effective in their day-to-day case management tasks. This highlights a fundamental tension in a prosecutor's work: prosecutors meeting office standards demonstrate worse outcomes in terms of future criminal justice contacts. Our findings underscore the need to rethink what makes an effective prosecutor in order to achieve policy

goals such as public safety, rehabilitation, and efficient resource use.

KEYWORDS
criminal courts, lower courts, misdemeanors, prosecutors

1 | INTRODUCTION

In recent years, there has been a growing recognition of the pivotal role played by prosecutors in shaping criminal justice outcomes (Bazelon, 2019; Pfaff, 2017). Prosecutors hold influence over several decisions that shape the trajectory of criminal cases: They determine which charges to file, extend plea offers, and can divert cases, withdraw prosecutions, or pursue conditional dismissals. At each stage of the criminal justice system, prosecutors serve as gatekeepers of the legal process. In this role, they must balance often-competing priorities such as crime control and due process, or the interests of victims and defendants, and they are given wide latitude to make decisions about these priorities as they see fit. Despite this central role in the administration of justice, empirically, the extent and nature of variation in practices among line prosecutors and their consequences on case outcomes has remained largely unexplored.

Most of the prior research on prosecutors has focused on felonies, leading to a few dominant theories of prosecutor discretion such as focal concerns and courtroom workgroups. In general, these theories argue that prosecutors seek to avoid uncertainty, balance a few high-level priorities such as public safety and defendant risk, and resolve cases in accordance with group norms. These theories have advanced our understanding of charging and sentencing especially for serious cases, yet most cases handled by prosecutors each year are for low-level offenses (Mayson & Stevenson, 2020). In addition, prior research on prosecutors has been constrained by the available data. For the most part, researchers have not had access to identifiers in order to link decisions to individual prosecutors. They also have not had access to granular data on prosecutor behaviors; instead, researchers have had to use case-level outcomes such as charges filed, dispositions, and sentences to make inferences about prosecutors' preferences.

In this article, we introduce a novel theoretical framework for understanding prosecutor decision-making by drawing attention to an overlooked aspect of prosecutor discretion: the practical realities of the job. We argue that *case management skills* play a first-order, or prerequisite, role in defining prosecutor effectiveness and shaping criminal case outcomes. In other words, case management tasks are the operational preconditions that must be met before prosecutors can exercise discretion according to their more abstract preferences on justice, rehabilitation, or public safety. Importantly, we focus on the misdemeanor courts and provide some of the first evidence on the variation in prosecutor decision-making in a high-volume, urban court system. Our case management model of prosecution fills the gaps in our understanding of both how prosecutors handle low-level cases and the specific details of prosecutor behaviors within the court process. This new perspective has important implications for theories of prosecutor decision-making, as well as for policy design.

We examine detailed administrative data from the Philadelphia District Attorney's Office on more than 67,000 misdemeanor criminal cases, handled by close to 400 prosecutors, from January 2011 to March 2020. These data allow us to identify the assistant district attorneys (ADAs)

responsible for the case at each hearing and their specific, hearing-level actions, such as their handling of discovery and witnesses. We relate these decisions to both case dispositions and defendants' future criminal justice contacts. We use these rich data to establish credible counterfactuals for each prosecutor's decisions, to make sure that differences in actions do not simply reflect differences in caseloads. Using fine-grained, hearing-level data on *individual prosecutors*, we can examine discretion and outcomes at the person level, allowing us to understand variation not just in ultimate case outcomes, but also in individual actions. We are thus able to expand on prior theoretical and empirical work that has mainly focused on jurisdiction-level differences (see, e.g., Johnson, 2018; Ulmer & Johnson, 2004) and can go beyond case outcomes, such as convictions and punishments, that are jointly produced with other workgroup members to focus on how prosecutors' actions influence those case outcomes. Although previous theories emphasize adherence to certain preferences or norms, we argue that case management skills are pivotal in shaping prosecutor actions and subsequent case outcomes. We also provide suggestive evidence that these case management skills represent a unique new dimension, distinct from those captured by traditional measures of legal actor preferences.

Our findings highlight a significant tension between efficient office work and effective policy implementation. Although prosecutors may aim to achieve high-level goals, such as delivering justice and ensuring public safety, their primary concern often lies in managing case-related tasks. Prioritizing these tasks can lead to outcomes that diverge from stated priorities. In particular, a focus on successful case management may lead to higher conviction rates; even though convictions are often considered a benchmark of prosecutorial success, they may make defendants more likely to be rearrested in the future. Our findings have significant policy implications, as district attorney's offices (DAOs) could potentially promote specific prosecutor behaviors to enhance both public safety and defendant outcomes.

2 | LITERATURE REVIEW

In this section, we first review the major theories of legal actor discretion. Next, we discuss theories of decision-making specifically focused on prosecutors and review the empirical evidence on this topic. Finally, we explore theory and evidence on misdemeanor courts and case prosecution.

2.1 | Discretion in the criminal justice system

In the criminal justice system, legal actors such as judges, prosecutors, and law enforcement officers wield significant discretionary power. This power enables them to interpret and apply laws and policies within the context of individual cases, tailoring responses to specific circumstances. However, the exercise of discretion also introduces the potential for bias, arbitrariness, and inconsistency in the application of justice. In this section, we review the major theories of discretion and evidence for each.

These theories follow a few broad themes. First, actors may primarily consider legal factors, like defendant blameworthiness and risk. Second, an individual actor's attitudes about crime and punishment may shape their decisions. Finally, they may be constrained by organizational and structural factors that may interact with or obscure the prior two points. We note that these theories are not mutually exclusive, and their relevance may vary across different jurisdictions.

One of the dominant theories of discretion is focal concerns theory. Focal concerns theory argues that judges balance a few key concerns when making sentencing decisions: defendant blameworthiness, community safety, and practical constraints (Steffensmeier et al., 1993, 1998). Although this theory was developed for judges, scholars have since argued that a range of actors prioritize these (or similar) concerns when making decisions about arrest, pretrial release, charging, and case processing (Crow & Adrion, 2011; Hartley & Tillyer, 2018; Kutateladze, 2018; Kutateladze & Lawson, 2018; Shermer & Johnson, 2010; Spohn et al., 2001; Ulmer et al., 2007; Winter & Clair, 2023).

Many of the empirical tests of focal concerns argue that legal actors rely on heuristics in which defendant risk and blameworthiness are associated with individual characteristics like race, gender, and age (Steffensmeier et al., 1993, 1998). Extensive studies in this area have found significant effects of a defendant's race, age, or gender on legal actors' decisions about arrest (Black & Reiss, 1970), pretrial detention (Chiricos & Bales, 1991; Demuth, 2003; Spohn, 2009), charging (Albonetti, 1986; Frohmann, 1997; Henning & Feder, 2005; Spohn et al., 1987), and sentencing (Hartley et al., 2007; Johnson, 2003; Johnson & Betsinger, 2009; Steffensmeier et al., 1993, 1998; Ulmer et al., 2007). Studies have also highlighted the importance of interaction effects between defendants' demographic characteristics (Hartley et al., 2007; Steffensmeier et al., 1998) or the interaction of defendant characteristics with the identities of legal actors (Ba et al., 2021; Boyd & Nelson, 2017; Johnson, 2006).

Relatedly, a large body of work has also sought to understand the effects of the individual characteristics of legal actors themselves, including judges (Boyd, 2016; Boyd & Nelson, 2017; Collins & Moyer, 2008; Harris & Sen, 2022; Lim et al., 2016; Spohn, 1991; Steffensmeier & Hebert, 1999), prosecutors (Heaton et al., 2025; Sloan, 2019), attorneys (Abrams & Yoon, 2007; Anderson & Heaton, 2012; Iyengar, 2007; King et al., 2010), and police officers (Ba et al., 2021). Underlying this work is the idea that different actors have different attitudes toward punishment, which may be correlated with their individual characteristics. Evidence is mixed on whether these characteristics influence decision-making, although a legal actor's personal characteristics seem more likely to influence their decisions in cases in which personal attributes are salient (e.g., discrimination or sexual assault cases; Rachlinski & Wistrich, 2017). Although legal actors' individual preferences may matter, they may also be constrained by organizational and structural factors.

Courtroom workgroup and courts-as-communities theories argue that legal actors operate within a bounded, constrained system. Eisenstein and Jacob (1977) first formalized the idea of a courtroom workgroup, in which repeated interactions between judges, prosecutors, defense attorneys, and court staff determine the outcome of criminal cases. Workgroups may pursue several (potentially conflicting) goals, like doing justice or disposing their caseloads. Depending on the stability, proximity, and familiarity of workgroup members, workgroups may exert influence on case outcomes through plea bargaining, information exchange, and shared norms (Eisenstein & Jacob, 1977; Haynes et al., 2010; Hester, 2017). For example, studies have found evidence of "going rates," or informal sentencing norms, that are shaped by workgroups and local legal cultures (Kim et al., 2015; Lynch, 2016; Ouss & Nguyen, 2025; Ulmer & Johnson, 2004). Kim et al. (2015) found that the effects of judge–prosecutor dyads on sentencing are consistently larger than individual prosecutor or judge effects, and that these effects vary by court district. In addition, the organization and culture of local court communities can affect the exercise of discretion by promoting certain ideologies or approaches to justice. Under these frameworks, individual actors' motivations, skills, or biases may still matter; however, their effects may be muted by workgroup norms, organizational factors, and the local community (Eisenstein & Jacob, 1977; Eisenstein et al., 1988). This perspective is supported by extensive research comparing charging and sentencing practices

across courts, in particular, across federal district courts (Hartley et al., 2007; Johnson, 2018; Johnson et al., 2008; Kautt, 2002; Lynch, 2016, 2018; Nagel & Schulhofer, 1992; Spohn & Fornango, 2009). Together, this work supports the idea of courts as communities with their own distinct cultures, norms, and organizational features that influence decision-making by legal actors.

Importantly, many theories of discretion are not inherently in conflict and may be integrated to explain jurisdictional differences (Dixon, 1995; Johnson, 2006; Spohn & Fornango, 2009). For example, Dixon (1995) empirically tested three theories of sentencing and showed that organizational context shapes criminal case outcomes by conditioning the exercise of discretion. She found that the influence of legal variables on sentencing decisions varies with the degree of bureaucratization in both judges' and prosecutors' offices. These findings demonstrate that theories of sentencing may be best understood as complementary rather than as competing, with different frameworks gaining explanatory power in different organizational environments.

2.2 | Prosecutor influence: Theory and empirical evidence

Given prosecutors' broad influence on the trajectory of criminal cases, scholars and policymakers view their discretion as both a driver of disparities and a potential lever for reform. In this subsection, we review the major theoretical frameworks that have been used to study prosecutor discretion and the empirical evidence in these areas so far. For the most part, these theories follow the frameworks that have been used to study judges and other legal actors, as reviewed above. In this article, we enhance these theories by considering how the everyday tasks that occupy the time and attention of prosecutors, as well as their varying levels of proficiency at managing these tasks, influence the outcomes of criminal cases. We integrate this new dimension into models of prosecutor decision-making, and we highlight its significance with our empirical analyses.

A key idea underlying most theories of prosecutor discretion is that prosecutors seek to minimize uncertainty and secure convictions (Albonetti, 1986). Prosecutors face unpredictability throughout the court process—from witness credibility to jury behavior and judicial rulings—which makes convictions uncertain. Because prosecutors' conviction rates may influence their career advancement (Albonetti, 1987; Eisenstein et al., 1988; Frederick & Stemen, 2012), prosecutors may make strategic decisions about charging, plea offers, and trial selection (Abrams, 2011; Abrams et al., 2024; Bushway et al., 2014; Kutateladze & Lawson, 2018). These uncertainty avoidance theories underscore the importance of legal factors, such as the strength of evidence and witness testimony, in shaping prosecutors' perceptions of the likelihood of conviction (Albonetti, 1987; Bushway et al., 2014; Davis et al., 2003; Frederick & Stemen, 2012; Hartley & Tillyer, 2018; Kutateladze et al., 2015; Spohn et al., 2001). In addition to impacting a prosecutor's perceptions of convictability, these legal factors could also be conceived of as representing a prosecutor's focal concerns.

From a focal concerns perspective, prosecutors are hypothesized to consider a few key factors when making case-processing decisions, although the specific meaning and prioritization of these focal concerns may vary. A prosecutor's focal concerns may be associated with legal (evidence strength, offense severity, presence of a witness) or extralegal (defendant race, victim age) characteristics. Previous studies have theorized that prosecutors' decision-making may be guided by factors such as the seriousness of the offense, the blameworthiness and dangerousness of the defendant, the credibility of the victim, and practical limitations such as caseload pressures (Hartley & Tillyer, 2018; Kutateladze, 2018; Kutateladze & Lawson, 2018; Shermer & Johnson, 2010; Spohn et al., 2001).

Importantly, prosecutors may emphasize different focal concerns at different stages of case processing. For example, evidence may matter more at charging (Frederick & Stemen, 2012; Kutateladze et al., 2015), whereas defendant's blameworthiness and criminal record may matter more at sentencing (Steffensmeier et al., 1998). Workgroup concerns also might impact prosecutors differently at different stages of case processing. For example, if cases are prosecuted horizontally, meaning a different prosecutor is responsible for charging decisions than for pleas and trials, the prosecutor in the charging unit may be less concerned with workgroup norms.

Prosecutors may also prioritize different concerns and face different constraints based on their local court context. Studies of prosecutors in the federal courts have shown significant variation in charging and sentencing practices across districts, even under sentencing guidelines (Hartley & Tillyer, 2018; Johnson, 2018; Johnson et al., 2008; Nagel & Schulhofer, 1992; Spohn & Fornango, 2009). Differences in prosecutorial behaviors across court communities have been linked to community size, caseload pressures, resource constraints, political context, population demographics, and community structural disadvantage (Hartley & Tillyer, 2018; Johnson, 2005; Kramer & Ulmer, 2002; Ulmer & Johnson, 2004; Wu & D'Angelo, 2014). These results suggest that not only do prosecutors in different court communities emphasize different focal concerns, but also the specific meaning of these concerns may vary by local context in ways that endure over time (Johnson, 2018; Lynch, 2016; Ulmer, 2005).

Prosecutors' court communities may also be shaped by organizational directives and legal rules. A growing body of research has examined how prosecutors respond to both internal policies and broader statutory changes. Studies of prosecutor-led initiatives, such as diversion (Nguyen & Graef, 2025), bail reform (Ouss & Stevenson, 2023), presumptive nonprosecution of low-level offenses (Agan et al., 2021), and internal sentencing policies (Lynch et al., 2021; Ouss & Nguyen, 2025), have found that prosecutors are responsive to internal policy changes. In contrast, research on external mandates like sentencing guidelines and mandatory arrest policies has suggested that prosecutors may resist policies that restrict their discretionary authority or, alternatively, shift their decision-making to take advantage of new opportunities created elsewhere in the system (Davis et al., 2003, 2008; Fischman & Schanzenbach, 2012; Lynch & Omori, 2014; Rehavi & Starr, 2014; Yang, 2013).

Despite all of these community constraints, the individual characteristics of the prosecutor may still influence decision-making. Consistent evidence exists that prosecutor experience matters, with more senior prosecutors generally being more lenient and less likely to take a case to trial (Frederick & Stemen, 2012; Heaton et al., 2025; Wright & Levine, 2014); however, Stemen and Escobar (2018) found no effects of prosecutor experience on case dismissals or pleas. Several studies have found that the effect of prosecutor characteristics depends on the identity of the defendant and the offense type, highlighting the importance of intersectional effects (Baker & Hassan, 2021; Didwania, 2022; Sloan, 2019).

We also note that although prosecutors are hypothesized to maximize convictions, empirical research has suggested that convictions may increase recidivism as the mark of a criminal record can hinder individuals from securing stable employment, housing, and social support upon release (Agan & Starr, 2017; Humphries et al., 2025; Leasure, 2019; Pager, 2003; Stewart & Uggen, 2020). If legal actors are aware of the potential negative consequences of conviction and incarceration, prosecutors seeking to maximize public safety (as focal concerns theory would suggest) may be expected to behave differently than prosecutors who are focused on securing convictions (as uncertainty avoidance theories predict). It is the *perceptions* of prosecutors that matter, however, and line prosecutors may still believe that convictions and punishments deter and incapacitate offenders, regardless of the evidence.

Overall, theories suggest that prosecutors aim to balance several goals, such as maximizing convictions and minimizing defendant risk to public safety, but the empirical evidence is still nascent. Prosecutors seem to be constrained by their environment and organizational goals as they interact with a courtroom workgroup and are influenced by internal policies. Within all of these constraints, who a prosecutor is may still matter, with their own biases and motivations subtly informing their decisions. Many studies of prosecution have used case outcomes at various stages of the court process to draw conclusions about prosecutors' decisions (e.g., modeling which cases are sentenced and inferring prosecutors' preferences). For the most part, however, researchers have been unable to answer precisely *what prosecutors do* on a daily basis to create these outcomes. In Section 3, we build on these prior theories by examining the role of mundane tasks in shaping court outcomes.

2.3 | Misdemeanor justice

Most research on prosecutors so far has focused on felonies; however, misdemeanor arrests account for more than 75% of all criminal cases filed each year (Mayson & Stevenson, 2020). As a result, we have little understanding about what prosecutors in the lower courts do and the extent to which their behaviors align with current theories of prosecutorial discretion. The present study focuses on misdemeanor prosecution in an attempt to fill this gap. In this subsection, we review the theoretical and empirical research on the misdemeanor courts.

Misdemeanor offenses are considered to be lower level offenses, ranging from drug possession to loitering to driving with an expired license.[1] Substantial differences in what counts as a misdemeanor exist depending on the jurisdiction, but misdemeanor crimes are typically defined by their potential incarceration sentence, with a common definition being that a misdemeanor is punishable by fewer than 12 months in jail (Mayson & Stevenson, 2020; Natapoff, 2011). Although misdemeanors are often minimized as "petty offenses," the consequences of a misdemeanor case can be great (Natapoff, 2018). Misdemeanor convictions still leave a visible criminal record that can significantly decrease employment prospects (Leasure, 2019; Pager, 2003) and educational opportunities (Stewart & Uggen, 2020). Individuals involved in the misdemeanor system also face heavy fines, probation, incarceration, deportation, and the loss of public assistance, housing, and other government benefits (Agan & Starr, 2017; Leasure, 2019; Natapoff, 2011; Pager, 2003; Uggen et al., 2014). Many of these consequences can accrue regardless of whether a defendant is convicted, and it is unclear which aspects of the criminal legal process contribute most to defendant outcomes (Agan et al., 2021).

Theoretical work on misdemeanors has highlighted many ways in which the lower courts may be different than the felony and federal courts, including the role that these courts play in society. In his study of New Haven, Connecticut's lower courts, Feeley (1979) argued that the "process is the punishment" for individuals accused of minor crimes: They face a burdensome string of court appearances, each requiring missed work, childcare, and transportation. Many defendants are detained in jail pretrial, incentivizing them to plead guilty in order to avoid large process costs (Feeley, 1979). More recent theoretical work has argued that the function of the lower criminal courts is not to adjudicate cases but to administer social control over "unruly" or marginalized populations (Kohler-Hausmann, 2019; Natapoff, 2018). A central implication of these social

[1] Mayson and Stevenson (2020) estimated that simple assaults, marijuana possession, petty theft, and DUIs account for approximately 60% of misdemeanor cases in most jurisdictions.

control theories is that all defendants (both innocent and guilty) receive informal, administrative punishment (Earl, 2019).[2]

Despite this renewed attention to misdemeanors, little empirical research has been done on *prosecutors* in the misdemeanor courts (Agan et al., 2021; Owusu, 2022; Roberts, 2021; Sloan, 2019). Many of the studies on prosecutor discretion that include misdemeanors have analyzed misdemeanor cases together with felonies (Berdejo, 2018; Eisenstein & Jacob, 1977; Frederick & Stemen, 2012; Haynes et al., 2010; Ulmer & Johnson, 2004). Among studies of misdemeanors, support has grown for the idea that prosecutors' decisions on misdemeanor case processing may depend on different factors than for more serious cases (Frederick & Stemen, 2012; Kutateladze & Lawson, 2018; Myers, 1982; Spohn & Cederblom, 1991). Most research on prosecutors, however, has come from state felony cases and the federal courts, leaving us with little understanding of how prosecutors handle the most common cases processed through the criminal justice system each year (Stevenson & Mayson, 2018). The following section builds a conceptual model for understanding how prosecutors impact cases in the lower courts.

3 | A CASE MANAGEMENT MODEL OF PROSECUTION

Existing theories of prosecutorial discretion argue that prosecutors make calculated decisions about their cases, even if those decisions are rapid and impressionistic. These decisions may be influenced by extraneous factors, such as defendant characteristics, or by the dynamics of repeated interactions. For example, according to focal concerns theory, a prosecutor balances their perception of a defendant's risk and blameworthiness with practical resource constraints. Theories of uncertainty avoidance argue that prosecutors weigh their likelihood of winning a conviction and triage cases accordingly. And much of the work on individual prosecutor characteristics implicitly or explicitly has argued that these characteristics reflect prosecutors' punitiveness or attitudes about crime and punishment. At the most basic level, these theories assume that prosecutors have some set of preferences for how they would like to handle their cases, and they should act according to these preferences.

We build on these theories to argue that *case management tasks* play a first-order role in defining prosecutor discretion and shaping criminal case outcomes. By "first order," we mean that case management (or operational) tasks are preconditions that must be met before prosecutors can exercise "higher order" discretion. In fact, case management tasks are so critical as to be dispositive in a large number of cases: Witness failure to appear in court is the strongest predictor of case outcomes (Graef et al., 2023), and many cases are dismissed after repeated continuances for incomplete discovery (Graef, 2025). In other words, prosecutors may prefer considering legal issues or prosecuting cases based on their preferences and beliefs, but if their witnesses do not show up or discovery is incomplete, their case might not even make it to the next hearing. Our work complements existing theories: Even though all prior theorized concerns may matter, we argue that case management tasks are first order, for both prosecutors' daily work and their influence on case outcomes, especially in the context of the lower courts. In this way, we draw a distinction between *operational discretion* (our article) and *preference-driven discretion* (much of the prior literature).

As such, this case management model may coexist with theories of prosecutorial discretion, but it is a distinct concept. Several theories do hypothesize that practical constraints influence case outcomes. For example, focal concerns theory includes the consideration of

[2] In this framework, "arrest is the gatekeeper for punishment, not conviction" (Earl, 2019, p. 28).

resource constraints, such as the impacts of sentencing on local jail populations. Courtroom workgroup theories acknowledge practical efforts to maintain good working relations with other court actors. These kinds of considerations, however, are separate from a prosecutor's *ability* to process a case: Case management tasks are the *procedural steps that must happen for a case to proceed*. These mundane tasks may be, but are not necessarily, driven by higher order considerations. We argue that prosecutors vary in their success at accomplishing these procedural tasks and that how well these tasks are accomplished impacts case outcomes independently and prior to the other concerns that prosecutors may be attempting to balance.

Our case management model of prosecutorial discretion is centered on the practical realities of their job. We use granular, hearing-level data on a prosecutor's case management tasks, or the things that must be completed for a case to move forward. These tasks include passing discovery, sending subpoenas to police officers, calling victims and witnesses, and completing return paperwork. We show that these mundane tasks and constraints matter quite a lot for case and defendant outcomes. This case management model of prosecution resonates with work in other parts of the criminal justice system. For example, Wilson (1978) found that although at a high level a police officer may be responsible for enforcing the law, they spend most of their time doing order maintenance and social service work, reshaping our understanding of police discretion. Similarly, a prosecutor's high-level objectives may matter, but these abstract ideals may influence their handling of cases much less than the practical realities of the job. As a consequence, variations in case outcomes may simply reflect variation in a prosecutor's skill at handling these case management tasks rather than a prosecutor's preferences. In fact, these constraints may be orthogonal to preferences, or even opposite.

Our goal is to explore whether and how prosecutor actions influence criminal case outcomes and future criminal justice contacts. Based on our review of the literature and the theoretical case management model of prosecution developed above, our main hypotheses are as follows:

Hypothesis 1. Even when handling similar cases, prosecutors vary in their case management skills.

Hypothesis 2. Cases handled by prosecutors with better case management skills are more likely to end in a conviction and in formal punishment.

Hypothesis 3. Prosecutors with higher conviction and formal punishment rates also see more of the defendants they worked with rearrested and charged with new offenses.

4 | THE CURRENT STUDY

4.1 | Institutional context: Philadelphia's misdemeanor courts

All criminal cases in Philadelphia begin in the municipal courts, but the path they take through the system depends on the severity of the case. With a few exceptions for cases that are appealed, misdemeanor cases are processed and disposed in the municipal courts. In contrast, felony cases receive a preliminary hearing in the municipal courts, where prosecutors must demonstrate probable cause; if they do, the case is transferred to the Court

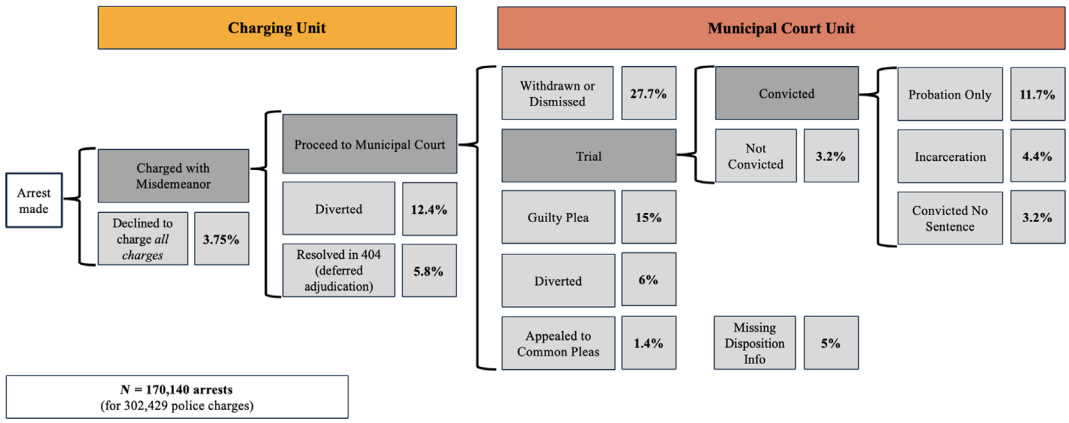

FIGURE 1 Outcomes of misdemeanor arrests in Philadelphia. *Note*: This figure charts the outcomes of 170,140 adult misdemeanor arrests made between January 1, 2011 and March 1, 2020. The percentages shown in light gray boxes are the percentage of all adult misdemeanor arrests during this time period that end at that node. We define a misdemeanor arrest by the most serious charge recommended by police. A total of 8,000 misdemeanor arrests in this time period were dropped due to missing identifiers needed to link the arrest to case information. An additional seven arrests (<0.0001%) were downgraded to summary cases at charging and therefore did not progress to municipal court. Less than 1% of arrests in this time period have not yet been disposed in the municipal courts. [Color figure can be viewed at wileyonlinelibrary.com]

of Common Pleas for resolution. This study focuses on misdemeanor cases in Philadelphia's municipal courts; however, our findings may hold for the processing of felony cases through lower court systems as well. To provide context for our research design, the following section describes the flow of misdemeanor cases through the system, as shown in Figure 1.

After an arrest in Philadelphia, a person is taken to a local police station and placed in a holding cell. The arresting police officer sends a report to the DAO, where a prosecutor in the Charging Unit reviews the case and decides what charges to file. In Philadelphia, very few cases are declined outright; approximately 96% of misdemeanor arrests are charged. After charging, some cases are diverted, and some defendants are offered deferred adjudication[3] (see Figure 1 for specific numbers); others are sent to the municipal courts. This is the focus of our study.

Cases sent to the municipal courts receive one or more scheduled hearings.[4] For misdemeanor cases, every hearing is a potential bench trial (i.e., trial by judge without a jury). ADAs in the Municipal Court Unit are responsible for reviewing the case file a few days before the scheduled hearing and for making sure the case is ready for court: Typically, this process means calling any

[3] Once charges are filed and bail is set, nondiverted misdemeanor cases are sent to a status room (in Philadelphia, typically courtroom 404). Here, prosecutors may extend a "status offer," which is a form of deferred adjudication in which charges are withdrawn if the defendant completes certain requirements (e.g., treatment, community service, restitution). Prosecutors are less likely to make such offers in cases with victims because they generally have not yet had time to speak with the victim. Status rooms are overseen by trial commissioners rather than by judges, so no pleas or trials occur. Approximately 6% of misdemeanor arrests are resolved through status offers.

[4] Currently, most cases are randomly assigned to municipal courtrooms based on the next available court date. A few types of cases, like DUIs and domestic violence cases, are assigned to specialty courtrooms. Before 2017, cases were assigned according to police districts, with courtroom numbers corresponding to specific districts.

victims and witnesses to remind them of the court date, sending subpoenas to police officers, reviewing any evidence, sharing discovery with the defense, and preparing to argue the case. Often, a case will not be ready to go to trial at the scheduled hearing, so the judge will decide whether to continue the case by scheduling another hearing date. Cases can be continued for several reasons, including because one or more necessary actors failed to appear, outstanding evidence remains to be gathered, or discovery was not shared between the Commonwealth and defense. A misdemeanor case receives an average of 3.8 hearings in the municipal courts before being disposed.

An important feature of Philadelphia's municipal courts is that cases are processed using a horizontal prosecution model, meaning that a case is passed along to a new ADA at each hearing. With the exception of specialty dockets,[5] the Municipal Court Unit supervisor creates the schedule by first placing ADAs with specially assigned cases on the calendar and then filing in the remaining ADAs with an eye to making sure each ADA gets experience working with a variety of judges and other ADAs. Typically, two to four ADAs will be assigned to a courtroom on a given day. Each day, cases are split among the ADAs in that courtroom by dividing the list from top to bottom or into alternating docket hours.[6] Although ADAs may occasionally select specific cases outside of this assignment process, such selection is likely based on features that are immediately observable in the case file—such as offense type—which are also captured in our administrative data.[7] Since our analyses account for these observable characteristics, it is unlikely that case selection is driven by unobserved factors outside our data.

Figure 2 provides a map of the municipal court process and links steps in the court process to key margins that prosecutors influence. On a granular level, ADAs are responsible for coordinating victims and witnesses, subpoenaing police officers, handling discovery, and more. In the misdemeanor courts, where each hearing is a potential bench trial, an ADA's success at these actions determines whether the case will be tried and resolved at the present hearing. Therefore, each ADA who touches a case impacts that case's time to disposition and the case's likelihood of conviction. Conditional on being convicted, ADAs may also influence sentencing. In particular, even though most convictions in the municipal courts are the result of bench trials, not guilty pleas, as in many other court systems, prosecutors can make sentencing recommendations. The primary mode of formal punishment in the municipal courts is probation (75% of misdemeanor cases ending in conviction receive probation); only ~19% of misdemeanor cases receive a sentence to incarceration.[8]

[5] For example, domestic violence and DUI cases are assigned to specialty courtrooms. Other dockets may be designated as needing special attention (via vertical prosecution by one specially assigned ADA) by the Municipal Court Unit supervisor.

[6] For example, the first ADA will take cases on the 8:00 and 10:00 docket, whereas the second ADA will take cases on the 9:00 and 11:00 docket. During the coronavirus-19 (COVID-19) pandemic, this strategy was standard practice in Philadelphia as it allowed ADAs a break between their cases to prep the victims and witnesses slated to testify on the next docket.

[7] This likelihood is supported by our observations and shadowing of prosecutors. They split up cases quickly, and once they begin working on a case, the case is theirs.

[8] We use the term "formal punishment" to denote cases that receive a sentence to confinement or probation. We do not, however, mean to diminish the administrative punishment that all defendants receive in the form of an onerous court process that may include pretrial detention, lost wages, or the hassle of showing up to multiple court dates. Some defendants receive both probation and incarceration.

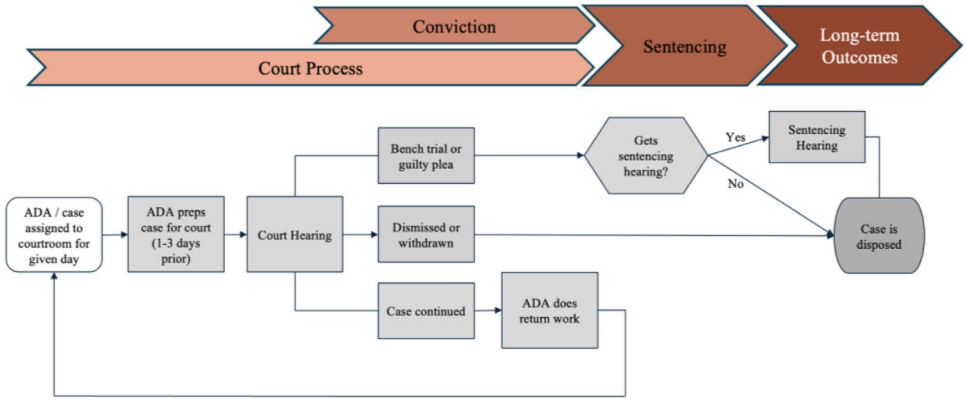

FIGURE 2 Municipal court process with sequence of prosecutor influence. *Note*: This figure shows a simplified map of the process for misdemeanor cases in Philadelphia's Municipal Court Unit. ADAs and cases are assigned to courtrooms for given days each week. For misdemeanor cases, every hearing in the municipal courts is a potential bench trial. At each hearing, three outcomes can occur: (1) The case proceeds to a bench trial or guilty plea, (2) the case is dropped, or (3) the case is continued to a future date. If a case is convicted, it may receive a separate sentencing hearing; otherwise it will be disposed that day. Above the process map are key areas of prosecutor influence: the court process itself (which we can also think of as "administrative punishment"), conviction, sentencing (formal punishment), and ultimately, long-term defendant outcomes (rearrests and new charges filed). [Color figure can be viewed at wileyonlinelibrary.com]

5 | DATA AND METHODOLOGY

5.1 | Data sources and sample construction

5.1.1 | Primary data sources

This study relies on two primary administrative data sources, maintained by the Philadelphia DAO. First, the District Attorney's Office Case Management System (DAOCMS) provides data on cases from charging to disposition. Second, the Administrative Office of Pennsylvania Courts (AOPC) provides detailed docket information on cases in Philadelphia County. We filter the dataset to cases with a misdemeanor lead charge at the charging phase.[9] The dataset is limited to cases disposed of between January 1, 2011, and March 1, 2020, providing a 2-year timeframe to assess rearrests and new charges for all defendants.

5.1.2 | Identifying ADAs

Crucially, our analysis relies on identifying the ADAs assigned to each hearing in a case. ADA identifiers are not readily available in any of our administrative datasets. We create these identifiers using docket entry comments in the AOPC data, which are detailed blocks of text entered by the stenographer and/or court clerk during court proceedings. The names of ADAs present at each hearing are extracted from this field using regular expressions; we de-duplicate records by

[9] Cases initially charged as felonies can ultimately be downgraded to misdemeanors, but they would be handled through a different court process.

linking back to a database of potential ADAs. We create a database of potential ADA names by combining internal human resources data from the DAO and a dataset of all bar-certified lawyers in Pennsylvania from the Disciplinary Board of the Supreme Court of Pennsylvania. Using these methods, we identify ADAs for 87% of hearings at which an ADA was present. We count any ADA named at the hearing as contributing to the case. Because ADAs in the Municipal Court Unit often work together while preparing their cases, this may be the most appropriate model for attributing case outcomes to ADAs.[10]

Our sample build progresses as follows. We begin with all cases from 2011 to March 2020 that have a misdemeanor as the lead charge ($N = 118{,}121$). We drop 8,516 cases that have incomplete docket entry information, which means they are missing ADA identifiers. We also exclude 41,908 cases that were diverted before reaching municipal court or were not heard in municipal court trial rooms (as these cases go through separate court processes), or where the ADA had a caseload of less than 50 misdemeanor cases in our sample period. This brings us to our final analysis sample ($N = 67{,}697$). In Table A.1, we show case characteristics for all misdemeanor cases in Philadelphia, for excluded cases, for cases where ADAs are missing, and for our primary analysis sample. Overall, our primary analysis sample is very similar to all misdemeanor arrests; however, a few notable differences are found in the proportion of driving under the influence (DUI) and drug cases. DUI and drug cases make up 22% and 37% of all misdemeanor municipal court cases, respectively, yet they each make up 31% of our main analysis sample. The differences in drug caseload are driven almost entirely by diversion programs, which divert those cases out of the municipal courts. The differences in DUI caseload are primarily driven by the fact that those cases are more likely to appear in municipal court trial rooms and by ADAs with higher caseloads (we drop cases seen in specialty courtrooms that follow different processes and cases seen by ADAs with total caseloads <50). Also, a few differences were found for cases with and without ADAs identified: For example, hearings for drug cases are slightly more likely to be missing ADA identification. Overall, however, ADA identification is not systematically related to case and defendant characteristics.

5.1.3 | Outcomes of interest

Following our hypotheses, outlined in Section 4, we classify our outcomes of interest into three categories:

- **Hearing outcomes**. We create a set of hearing-level variables to capture ADA case management skills. We use two of these measures in our main analyses. First, Commonwealth FTA is a binary variable indicating whether a key officer, victim, or civilian witness for the Commonwealth failed to appear at that hearing. Second, discovery incomplete is a binary variable for whether discovery was marked incomplete on the docket entries for a given hearing.[11]

[10] In the docket entries, sometimes only one ADA is listed at the hearing, whereas other times, multiple ADAs are reported. We count any ADA listed as contributing to the case because although we focus on individual effects, they often work collaboratively. For example, two ADAs assigned to courtroom 703 on Wednesday typically divide the docket in advance, each taking half the cases. In court, each is primarily responsible for their own set of cases; however, they may help each other (e.g., whispering tips for witness examinations during trial) or step in for each other if one ADA is temporarily busy.

[11] In Figure A.4, we consider three other measures of case management skills: (1) "Return FTAs," which is an indicator equal to one if a key Commonwealth witness failed to appear the *following* hearing, reflecting how effectively ADAs completed their return work; (2) "Commonwealth Not Ready," which is an indicator equal to one if a witness failed to

- **Case outcomes**. We also define several variables for case-level outcomes. We define a conviction as a case that ends in a guilty finding at trial or a guilty plea.[12] Case length is the amount of time in days between charging and disposition. We use a binary measure of formal punishment, which is defined as a sentence that includes any probation or incarceration.[13]
- **Defendant outcomes**. We use two outcome measures of future criminal justice contact. We define rearrest within 2 years as a binary indicator that is equal to one if the defendant has an arrest for a new incident within 2 years of the original case's open date. New charges within 2 years are similarly defined. We do not count probation violations under new arrests or new charges; we only count arrests for new incidents.

5.1.4 | Covariates

Our dataset includes fine-grained measures of case and defendant characteristics, the defendant's criminal history, elements of the courtroom workgroup, and the arrest context. Here, we describe how we classify this information in our main analyses:

- **Defendant characteristics**. We include measures of defendant age at arrest, gender, and binary indicators for Black, White, and Hispanic defendants.
- **Case characteristics**. Case covariates include indicators for the high-level charge type (e.g., violent, drug, property, DUI, public order), as well as for the specific charge, down to the subsection in the Pennsylvania criminal code. We include a separate flag for domestic violence cases and flags for whether the case includes any common "add-on" charges of resisting arrest, possession of an instrument of crime (PIC), or conspiracy. Case severity is also captured by variables for the number of current charges and the lead charge severity (e.g., M1, M2, M3).
- **Criminal history**. We measure several elements of a defendant's criminal history. First, we capture current criminal justice system involvement, using indicators for whether the defendant is currently on probation and for whether the defendant has another pending case at the time their current case is opened. We use several measures of a defendant's priors, including the number of prior arrests and indicators for whether the defendant had a prior charge in the past year, any prior misdemeanor conviction, any prior felony conviction, any prior violent conviction, and any prior incarceration. Finally, following the literature, we define pretrial detention as an indicator for whether a defendant remains in jail more than 3 days after arrest (Ouss & Stevenson, 2023).

appear, discovery was incomplete, or the docket otherwise noted the Commonwealth was unprepared; and (3) "Must Be Tried," which is an indicator equal to one if the docket noted that the case must be resolved at the following hearing, often reflecting a judge's assessment that the case has been dragging on.

[12] In our context, most cases either end in conviction (55%) or are dismissed/withdrawn (38%), with a small share acquitted at trial (4%), or diverted/appealed (3%). In our main analyses, we compare convictions to all nonconviction outcomes—including dismissals, acquittals, diversions, and appeals. To check robustness, we re-estimated our analyses with alternative codings for acquittals (e.g., treating them as their own outcome or grouping them together with convictions), and our results remain substantively unchanged.

[13] As discussed above, only 19% of all cases in our sample result in any incarceration. Approximately 14% of cases that are convicted in our sample receive no formal punishment, and 52% of cases that are convicted receive probation-only sentences.

- **Other legal actors**. We consider the judge overseeing the case at the present hearing by including a full set of judge indicator variables. Although we do not have unique identifiers for defense attorneys, we include an indicator variable for whether the case was handled by a public defender, as opposed to by a private attorney.
- **Arrest context**. Finally, our analyses are within the specific year and police district of the arrest. These are all measured as indicator variables.

5.2 | Descriptive statistics

Table 1 shows descriptive statistics for our sample. The average case in our sample takes ~7 months from charging to disposition. A total of 33% of defendants have a recent prior arrest (29% have a prior misdemeanor conviction, and 21% have a prior felony conviction). More than half of the cases are drug (31%) or DUI (31%) crimes, with the remainder split between violent (16%), property (11%), public order (8%), and other (3%) offenses. Approximately 14% of cases involved domestic violence. In 55% of cases, the defendant is convicted; most other cases end in dismissal or withdrawal (38% of cases in our sample).[14] Very few cases in our sample were acquitted at trial (4%). In terms of sentencing, in 31% of cases, the defendant gets probation, 19% are sentenced to incarceration, and 8% are convicted but get no formal punishment.[15] Regarding longer term outcomes, in 48% of cases, the defendant is rearrested within 2 years of the original case being opened, and in 41% of cases, the defendant has new charges filed against them within 2 years.

5.3 | Research design

This article seeks to quantify the amount of variation in prosecutors' granular decisions and explore how those decisions map onto case and defendant outcomes. To do so, we want to consider variation in decisions for prosecutors *handling similar cases*. Prior research has exploited quasi-random assignment of cases to judges or prosecutors (see Loeffler & Nagin, 2022 for a recent review of this literature) to estimate the causal effect of discretion on criminal case outcomes. In our context, however, no quasi-random assignment exists: We see that prosecutors tend to specialize into certain case types. We therefore constitute comparable caseloads using

[14] Our data do not allow us to meaningfully distinguish between dismissals (from judge) and withdrawals (from prosecutor). We see a case's disposition, which can be recorded as "dismissed," "nolle prossed," "withdrawn in the interest of justice," and so on. Through court observations and interviews with ADAs, however, we learned that ADAs often withdraw a case when they know the judge would otherwise dismiss it; doing so may build goodwill with the judge (because the ADA takes the "blame" for dropping the case). Given these realities, we do not believe that we can make a meaningful distinction between "dismissed" and "withdrawn" here. Throughout the article, when we refer to "dismissed" cases, it includes cases dismissed by a judge and withdrawn by a prosecutor.

[15] This happens when a judge thinks that further punishment is not in the interest of justice. For example, while observing court proceedings for misdemeanor family violence cases, we witnessed instances of judges finding a defendant guilty but ruling that the defendant had "suffered enough" through the aftermath of the incident, court proceedings, and/or financial consequences.

TABLE 1 Descriptive statistics.

Variable	Mean
Defendant and case characteristics	
Male	0.81
Black	0.60
Hispanic	0.13
Defendant age at arrest	35.52
Violent	0.16
Drugs	0.31
DUI	0.31
Property	0.11
Public order	0.08
Domestic violence	0.14
Number of current charges	2.09
Prior arrest in past year	0.33
Number of prior arrests	2.45
Pending case at open	0.19
On probation at case open	0.28
Bail posted within 3 days	0.67
Hearing outcomes	
Commonwealth FTA	0.21
Return Commonwealth FTA	0.14
Discovery incomplete	0.23
Commonwealth not ready	0.44
Marked "must be tried"	0.16
Case outcomes	
Case length days	226.26
Convicted	0.55
Formally punished	0.49
Probation only	0.31
Incarcerated	0.19
Convicted with no formal punishment	0.08
Defendant outcomes	
New arrest within 2 years	0.48
New charge within 2 years	0.41
Cases	67,697
Hearings	197,821

Note: This table shows descriptive statistics for our main sample of misdemeanor lead charge cases with at least one ADA identified at the hearing. We count any ADA who touched the case as part of the ADA team. A total of 395 ADAs were assigned to these cases/hearings.

Abbreviations: DUI, driving under the influence; FTA, failure to appear.

internal benchmarking to estimate the effect of each prosecutor (Ridgeway & MacDonald, 2009).[16] Variations on this method have been used to study racial bias in police stops and county-level incarceration rates, neighborhood-level differences in perceptions of police performance, and judges' sentencing decisions (Nguyen & Ridgeway, 2023; Ridgeway & MacDonald, 2009, 2014; Ridgeway et al., 2020).

Our goal is to compare cases handled by a given prosecutor with cases handled by other prosecutors that have similar features—their "internal benchmark." Doubly robust internal benchmarking is a two-stage process that involves (1) propensity score weighting and (2) doubly robust estimation. Once the benchmark is created, we can estimate how much the individual prosecutor differs from their benchmark in terms of case and defendant outcomes, like convictions and rearrests. We can then explore how prosecutor actions map onto case outcomes and future criminal justice contacts by exploring the association between a prosecutor's granular, hearing-level actions and ultimate outcomes.

We note a key difference with other propensity score approaches, which are typically used to determine the likelihood of receiving a treatment (Rosenbaum & Rubin, 1983). In our setting, "treatment" means that a case is handled by the "index prosecutor." All cases handled by other prosecutors get a propensity score: cases that are very similar to the index prosecutor's cases will have large propensity scores, and cases with observable characteristics that are very different from the index prosecutor's cases will have propensity scores near zero.

We estimate propensity scores using the boosted logistic regression method described in McCaffrey et al. (2004).[17] We match on a very rich set of case features, including defendant demographics (age, gender, and race), the specific lead charge down to the subsection in the PA criminal code, whether the case involved domestic violence, the defendant's criminal history (including prior arrests, prior convictions, prior incarceration, pretrial detention, and whether they had another case pending at the time of arrest), whether the case was assigned to a public defender, the judge, the police district in which the current arrest was made, and the year. This approach allows us to account for most dimensions on which prosecutors may select cases.

Creating the benchmark is an iterative process. We use the max Kolmogorov–Smirnov (KS) statistic to measure the maximum difference in case features between a prosecutor and their benchmark. If this difference is six percentage points or less (max KS $\leq$ 0.06), we consider the model converged.[18] Otherwise, we drop cases that are highly dissimilar from their peers based on the propensity scores. We repeat this process until the model converges or until fewer than 50 cases are left in a prosecutor's caseload. If convergence is not achieved with at least 50 cases, no benchmark is created for that prosecutor. This method yields benchmarks for 265 prosecutors in our sample (out of 395 total prosecutors). In practice, we are dropping prosecutors who are handling very specialized cases. Because benchmarking relies on re-weighting similar cases seen by peer prosecutors, this approach will not work on less common case types. In addition, since we iteratively drop highly dissimilar cases from a prosecutor's caseload when their model does not

[16] We note that this method was originally developed to identify outliers, for example, to detect police officers who are likely making racially biased stops (Ridgeway & MacDonald, 2009). However, we are primarily interested in the variation among prosecutors and how this variation maps onto outcomes, so we do not include the steps required to identify outliers.

[17] We remove any other prosecutors associated with the same case from being used to construct the benchmark.

[18] Prior literature (Nguyen & Ridgeway, 2023) has used the threshold of max KS $\leq$ 0.05 for convergence. We use a slightly higher threshold to include a greater number of prosecutors in our main analyses. We also run our analyses using KS $\leq$ 0.05, however, and the results remain substantively the same.

TABLE 2 Raw outcomes for ADA #362.

Variable	ADA #362	Unweighted controls	Weighted controls
ADA skills/hearing outcomes			
CW FTA	0.22	0.22	0.20
Return FTA	0.11	0.14	0.11
Disco incomplete	0.25	0.24	0.24
CW not ready	0.48	0.46	0.44
Marked MBT	0.16	0.18	0.14
Case outcomes			
Case length days	260.03	225.74	260.40
Convicted	0.66	0.54	0.67
Punished	0.59	0.49	0.65
Incarcerated	0.19	0.19	0.19
Probation only	0.41	0.30	0.46
Convicted with no formal punishment	0.08	0.08	0.06
Defendant outcomes			
New arrest within 2 years	0.50	0.48	0.52
New charge within 2 years	0.44	0.41	0.45
Observations	1276	258,189	ESS: 2419

Note: This table shows raw outcomes for a randomly selected ADA (ADA #362) compared with their unweighted controls (Column 3) and weighted controls from benchmarking (Column 4). We do not match on outcomes: This table simply shows the raw differences in outcomes for ADA #362, their unweighted peers, and their benchmark. A case is counted as part of the ADA's caseload if that ADA handled the case at any hearing in municipal court. For example, ADA #362 handled a total of 1276 hearings for 1039 unique cases. ADA skills/hearing outcomes are measures of what happened at ADA #362's hearing. Case outcomes are the final outcomes for the case, which may have occurred at subsequent hearings. Defendant outcomes measure whether the defendant in ADA #362's cases was rearrested or had a new charge filed within 2 years of the original case being opened.
Abbreviations: CW, Commonwealth; ESS, effective sample size; FTA, failure to appear.

converge, some prosecutors have a benchmark created for just a portion of their cases. As such, our estimates can be understood as the effect of each prosecutor *when handling common cases*. We cannot predict the behavior of highly specialized prosecutors or the behavior of "regular" prosecutors for highly specialized cases.

To illustrate our methodology, we demonstrate the benchmarking process for a randomly selected prosecutor, ADA #362. First, Table 2 shows the raw outcomes for prosecutor #362's cases. Overall, 66% of their cases ended in a conviction, and 50% of defendants associated with their cases were rearrested. How do we know whether a rearrest rate of 50% is high? Is the 66% conviction rate for this prosecutor's cases (12 percentage points higher than the overall mean) reflective of some special skill of this prosecutor or just the underlying features of the case? The first column in Table 3 presents characteristics of cases handled by prosecutor #362, and the second column presents the characteristics of cases handled by all other prosecutors. Their caseloads are very different: For example, prosecutor #362 saw a large proportion of drug cases (41%), whereas the proportion of drug cases seen by their peers was just 26%. To address these differences, we construct a weighted set of cases for the prosecutor's benchmark using the process described above. This weighted comparison group (the "benchmark") is shown in Column 3, and the KS statistics

TABLE 3 Caseload balance for ADA #362.

Variable ADA #362		Unweighted controls	Weighted controls	KS
Defendant and case characteristics				
Age at arrest	36.53	35.48	36.33	0.02
Male	0.82	0.81	0.82	0.00
Black	0.59	0.58	0.56	0.04
Hispanic	0.17	0.13	0.16	0.00
Violent	0.05	0.13	0.04	0.01
Drug	0.41	0.26	0.43	0.02
Property	0.17	0.09	0.17	0.00
DUI	0.24	0.44	0.27	0.02
Public order	0.10	0.07	0.08	0.02
Domestic violence	0.01	0.11	0.01	0.00
Lead charge grade M1	0.35	0.43	0.33	0.02
Number of charges	1.78	2.01	1.76	0.01
Pending case at open	0.22	0.18	0.23	0.01
On probation	0.35	0.24	0.33	0.02
Bail posted in 3 days	0.57	0.68	0.56	0.01
Prior arrest in past year	0.31	0.26	0.31	0.00
Police districts				
District 01	0.00	0.02	0.00	0.00
District 03	0.00	0.06	0.00	0.00
District 06	0.14	.04	0.13	0.01
Workgroup actors				
Has public defender	0.69	0.64	0.71	0.02
Judge A	0.02	0.03	0.02	0.00
Judge B	0.04	0.04	0.04	0.00
Judge C	0.10	0.04	0.09	0.01
Specific charges				
Simple assault	0.02	0.08	0.02	0.00
Terroristic threats	0.01	0.03	0.01	0.00
Theft by unlawful taking (a)	0.07	0.04	0.07	0.00
Observations	1276	258,189	ESS: 2419	Max KS: 0.04

Note: This table shows an example of the benchmarking process for ADA #362. The first column shows the distribution of features for ADA #362's caseload. The second column shows the raw distribution of features for hearings handled by all other ADAs. The third column shows the distribution of features for ADA #362's benchmark, created by the propensity score weighting process. The fourth column shows the maximum vertical distance between the cumulative distribution of ADA #362's caseload and their benchmark. For brevity, we truncate the list of case characteristics, years, police districts, specific charges, and judges displayed. We match on specific charges down to the subsection in the Pennsylvania criminal code.

Abbreviations: DUI, driving under the influence; ESS, effective sample size; KS, Kolmogorov–Smirnov.

are shown in Column 4. Now, the weighted comparison cases are quite similar to ADA #362's caseload—the distributions of the case features differ by a maximum of four percentage points. The difference in outcomes is also now smaller. To investigate caseload balance across all benchmarked ADAs, the plots in Figure A.1 show the difference in six case and defendant characteristics using all ADAs (in blue) or the benchmarked ADAs (in red). Although a lot of variation occurs across ADAs in their raw caseloads, the differences in benchmarked caseloads are tightly centered near zero.

Although our rich set of controls allows us to account for most dimensions on which prosecutors could select cases, we may still be concerned that prosecutors are selecting cases based on factors that are unobservable to the analyst. To explore this possibility, we investigate balance for two variables that are not explicitly controlled for in the propensity score weighting phase, and on which prosecutors could potentially select what cases to work on: arrest narrative length, which can proxy strength of the evidence,[19] and upstream charging decisions—that is, how charging ADAs (who work in a different unit) chose to charge a case, relative to the arresting officer's recommendation. Figure A.2 presents the difference in average arrest narrative length using all ADAs (in blue) or the benchmarked ADAs (in red). Figure A.3 does a similar exercise, but this time using *prior* decisions made by the charging prosecutor, before a case reaches municipal court. In Philadelphia, charging prosecutors respond to police recommended charges, adjusting them upward, downward, or keeping the same charge.[20] Both arrest narrative length and prior charging decisions could capture measures of the strength of the evidence at hand. For both outcomes, we see that a lot of variation occurs across ADAs in caseloads; however, after benchmarking, their caseloads are much closer on both dimensions. Thus, benchmarked caseloads are not only (mechanically) balanced on observable case characteristics; the fact that we find them to also be balanced on case features that we did not explicitly benchmark on suggests that ADAs face similar caseloads even for unobservable characteristics.

After creating a suitable benchmark for each prosecutor, we use doubly robust estimation to get the effect of the index prosecutor on each of our hearing, case, and defendant outcomes. Doubly robust estimation uses the propensity score weights as sampling weights in a regression model that includes the potential confounders that were matched on in the propensity score weighting stage. This approach protects against model misspecification and provides a consistent estimate of the prosecutor's effect (the "treatment") with either a correctly specified propensity score model or a correctly specified outcomes regression model (Bang & Robins, 2005; Ho et al., 2007). All but one of our outcomes of interest are binary; we use a propensity score weighted linear probability model. For case length, we use propensity score weighted ordinary least squares.

From these regression models, we extract doubly robust estimates of the treatment effect of each prosecutor, as well as z-scores that measure how much the target prosecutor deviates from their benchmark on a given outcome.[21] We scale the z-scores to have a mean of 0 and standard deviation of 1 so that the scores can be more easily compared across outcomes.Table A.2 shows a random

[19] We note that this is a crude proxy of the strength of evidence or case complexity, reflecting the extent of details police officers choose to record when making an arrest.

[20] We define upward and downward adjustments using the method developed in Heaton et al. (2025). Specifically, when looking at all charges in an arrest, a case is considered an upward (downward) adjustment if (1) the maximum charge rank increases (decreases) from the set of charges recommended by the police, (2) the overall number of charges at the maximum charge rank increases (decreases), or (3) common "add-on" charges are added (dropped).

[21] These z-scores are the equivalent of the t-statistics associated with a regression coefficient (calculated as coefficient estimate/standard error).

draw of five rows from the prosecutor-level dataset with standardized *z*-scores. Each row contains information for a single prosecutor that was benchmarked (identified by an ADA ID number). Each column holds *z*-scores for a hearing, case, or defendant outcome, which measures the magnitude to which a prosecutor's outcomes differ from their benchmark. For example, prosecutor #362 has a conviction *z*-score of 0.744, indicating that they are somewhat more likely than their peers to convict a case. Prosecutor #362 also has a *z*-score of 0.255 for Commonwealth FTA and 0.221 for incomplete discovery. In short, this prosecutor is slightly more likely to have key Commonwealth witnesses fail to appear and to secure convictions than their statistical benchmark of peers.

6 | RESULTS

Given similar caseloads, we estimate prosecutor variation in outcomes. Figure 3 displays histograms of the estimated treatment effects of prosecutors on hearing, case, and defendant outcomes (future criminal justice contacts). These estimates measure the effect of being assigned to prosecutor *i* relative to a statistical benchmark of their peers.

6.1 | Variation in prosecutors' case management skills

The hearing-level outcomes shown in the first row can be thought of as demonstrating a prosecutor's case management skills; these are the granular actions that prosecutors take at each hearing in the municipal courts. Compared with a benchmark of comparable cases handled by their peers, prosecutors vary in the extent to which they get key Commonwealth witnesses (including civilians and police officers) to appear in court. The estimated treatment effect ranges from −0.24 to 0.19, with a standard deviation of 0.06. A negative treatment effect here says that prosecutor *i* is less likely than their peers to have a key witness fail to appear at a hearing that they are responsible for. In other words, moving from the most effective to the least effective prosecutor increases the probability that a key witness fails to appear at a given hearing by 43 percentage points. More conservatively, moving from a prosecutor in the 10th percentile to the 90th percentile increases the probability of a witness failure to appear by 12.4 percentage points. Similarly, prosecutors are also responsible for passing discovery to the defense. Moving from the most effective to the least effective prosecutor increases the probability that discovery is marked incomplete at their hearing by 42 percentage points (or by 12 percentage points when moving from 10th to 90th percentile). Consistent with H1, these figures show that considerable variation occurs in how good prosecutors are at the case management aspect of their jobs: Even when faced with similar caseloads, they have different success rates in terms of getting all the necessary elements in place for a case to move forward. Fourteen out of 265 prosecutors deviate in case management outcomes by more than 2 standard deviations from the mean. Many prosecutors, however, have similar case management outcomes compared with their benchmarked peers. On the whole, a handful of prosecutors are very successful case managers, and a large chunk of their peers are slightly less successful than average at managing witnesses and discovery.

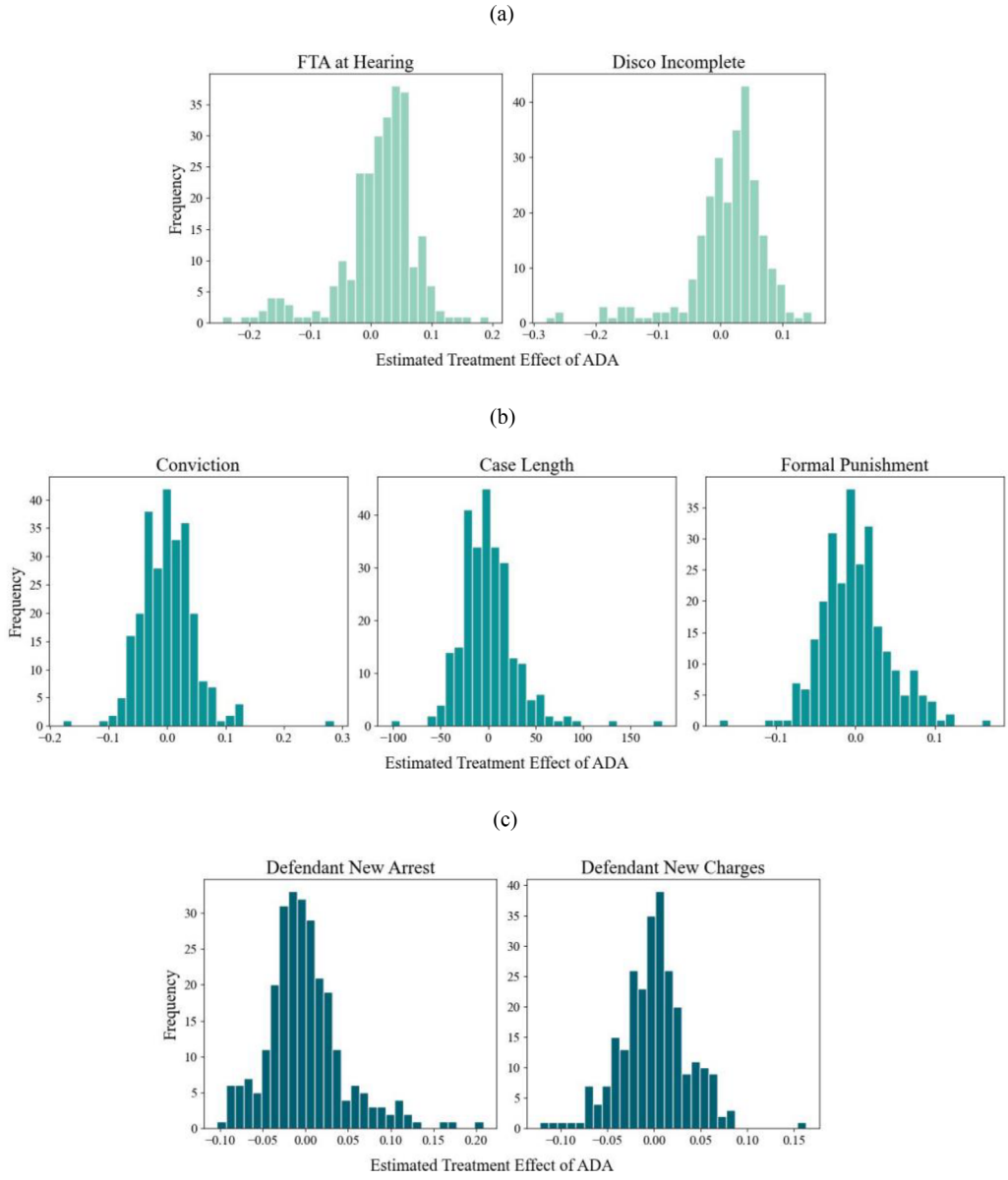

FIGURE 3 Distribution of estimated treatment effects of prosecutors on (a) Hearing outcomes. (b) Case outcomes. (c) Future criminal justice contacts. *Note*: This figure shows histograms of the estimated treatment effects of prosecutors on case management skills (hearing-level outcomes), case outcomes, and defendant outcomes. Treatment effects are estimated using propensity score weighted ordinary least-squares regressions of each outcome on an indicator for prosecutor *i* and our full set of covariates. In short, they estimate the effect of each prosecutor on a given outcome, relative to a statistical benchmark of their peers. For example, to create the leftmost plot in panel (a), we estimate the effect of being assigned to prosecutor *i* on whether a Commonwealth failure to appear occurs at prosecutor *i*'s hearing. We do this for all 265 prosecutors in our benchmarked sample and then create a histogram to show the distribution of these estimated treatment effects. All outcomes (except for case length) are binary, so the treatment effects can be interpreted as percentage point changes. Case length effects can be interpreted as the change in days to disposition for a case assigned to a given prosecutor. ADA, assistant district attorney; FTA, failure to appear. [Color figure can be viewed at wileyonlinelibrary.com]

6.2 | Prosecutors' effects on case outcomes and future criminal justice contacts

We next turn to case outcomes and future criminal justice contacts, conducting similar analyses to those for hearing outcomes. These analyses help us consider some of the costs associated with different ways that prosecutors handle cases. In the second row of Figure 3, we show the estimated treatment effects of prosecutors on the case outcomes of conviction, case length, and formal punishment. The estimated treatment effect of prosecutors on convictions ranges from −0.18 to 0.29, with a standard deviation of 0.04. Moving from a prosecutor in the 10th percentile to a prosecutor in the 90th percentile increases the probability of conviction by 10.5 percentage points. Similarly, moving from the 10th to 90th percentile increases the probability of formal punishment by 10.4 percentage points. Finally, in the third row of Figure 3, we show the estimated treatment effects of prosecutors on the defendant outcomes of new arrests and new charges filed. We find that moving from a prosecutor in the 10th percentile to a prosecutor in the 90th percentile increases the likelihood of rearrest by 10.5 percentage points and increases the likelihood of new charges by 9.2 percentage points. Overall, these results provide support for variation in prosecutor actions that maps onto disparate case and defendant outcomes.

6.3 | Better prosecutor case management skills translate into more future criminal justice contacts

To understand whether and how prosecutor skills map onto ultimate outcomes, we examine the correlations between prosecutor *z*-scores for each set of outcomes. Figures 4 and 5 display these results.

First, in Figure 4, we map hearing outcomes onto case outcomes. These hearing outcomes represent a prosecutor's case management skills of coordinating witnesses and managing discovery. We find that prosecutors who are better case managers are more likely to secure convictions and use formal punishments.[22] Prior theoretical and empirical literature has posited that variation in conviction rates among prosecutors stems from differences in how they exercise discretion to reflect their preferences. Our results unpack that prosecutorial discretion by understanding what exactly prosecutors do to convict a case.

We also explore the relationship between higher convictions and more recidivism by mapping case outcomes onto future criminal justice contacts. Figure 5 presents these correlations. We find that prosecutors who are more likely to have cases end in convictions and formal punishment are descriptively more likely to have defendants on their caseload rearrested or have new charges filed against them within 2 years. This finding aligns with prior studies that have estimated the causal effect of misdemeanor convictions on recidivism and defendant outcomes (Agan et al., 2021; Stewart & Uggen, 2020). Although we cannot parse the treatment effects of conviction versus punishment here, we show that prosecutors who have higher recidivism rates compared with their statistical benchmark have cases that end in convictions and formal punishments more frequently than their peers. The first rearrest for defendants in our sample is most likely to be for a drug offense (32%), followed by property (24.5%), violent (21.6%), public order (10.8%), DUI (7.2%), and other offenses (3.9%).

[22] Because dismissals are essentially the inverse of conviction in our sample, the results for dismissals are very similar but in the opposite direction. For this reason, we do not report dismissal-specific results here.

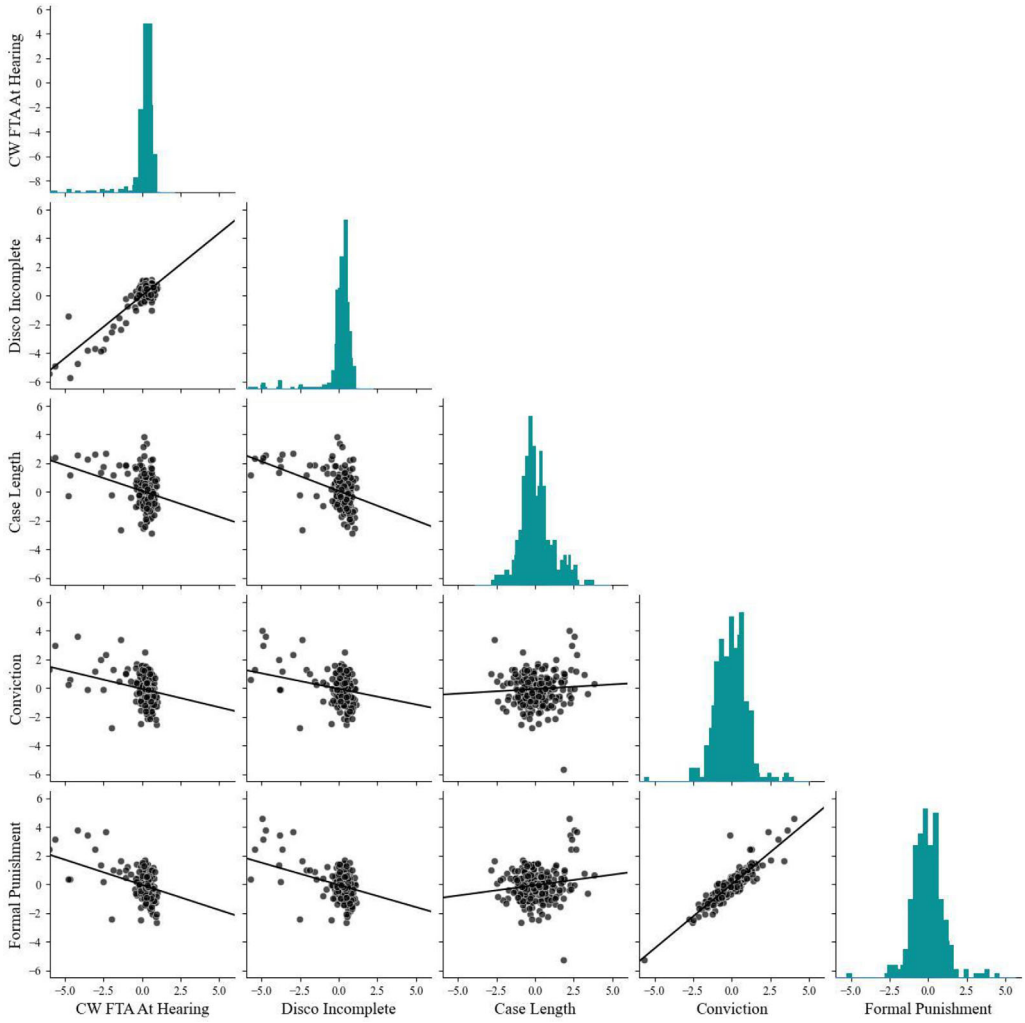

FIGURE 4 Associations between hearing outcomes and case outcomes. *Note*: These plots show the association between prosecutors' *z*-scores for hearing and case outcomes. Each point in the scatterplot corresponds to a prosecutor's standardized *z*-score. The black lines represent ordinary least-squares regression. Histograms for the standardized *z*-scores corresponding to each outcome are shown on the diagonal. CW, Commonwealth; FTA, failure to appear. [Color figure can be viewed at wileyonlinelibrary.com]

Tables A.3 and A.4 also show these results in regression format. For example, in Column 2 of Table A.3, the coefficient on Commonwealth Not Ready tells us that, on average, a prosecutor who is 1 standard deviation more likely to be "not ready" at their hearing (by having a key witness fail to appear or discovery incomplete) is 0.413 standard deviations less likely to have their case convicted.

Because we are testing the effect of several hundred prosecutors, we also calculate adjusted *z*-scores to account for multiple hypothesis testing. To do this, we perform false discovery rate corrections using the Benjamini–Yekutieli (BY) method. We obtain the *q*-values and use these to generate adjusted *z*-scores. We then re-run our regressions of case outcomes on case management

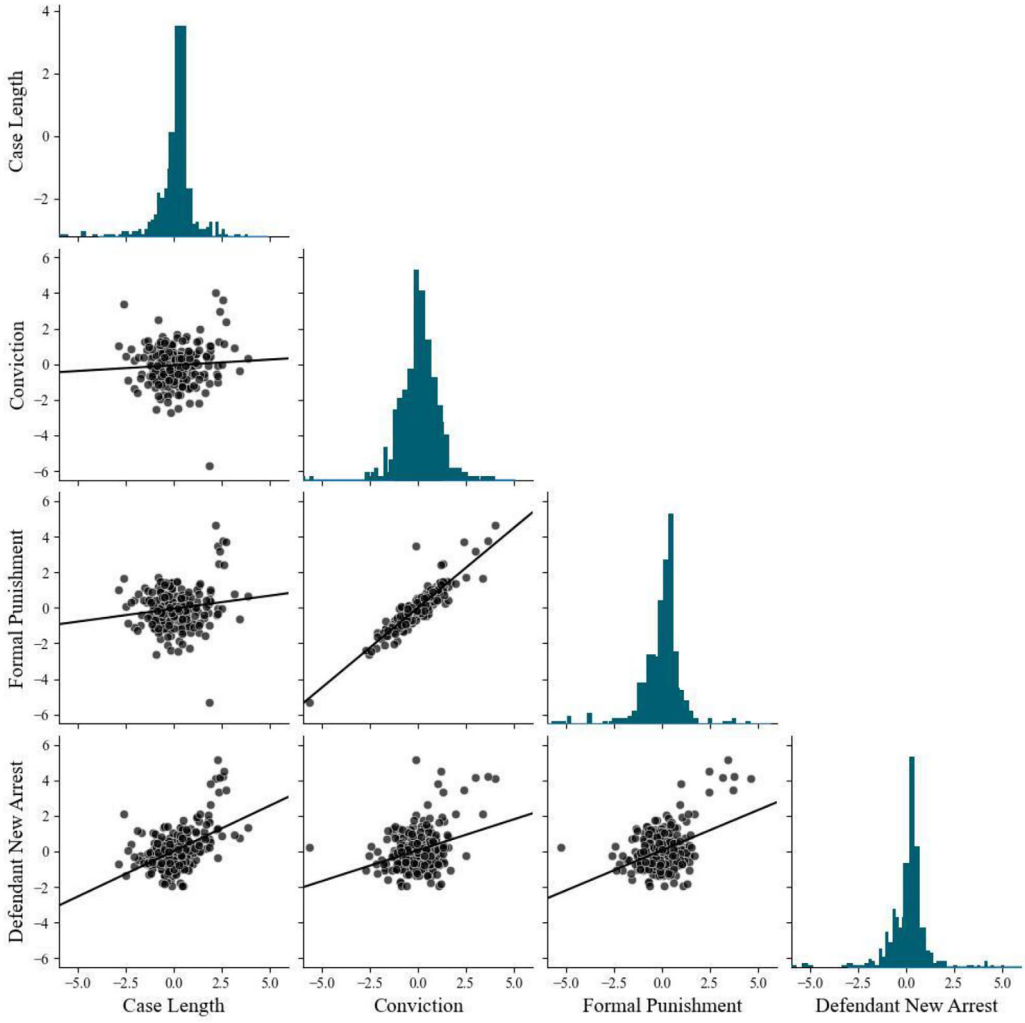

FIGURE 5 Associations between case outcomes and future criminal justice contacts. *Note*: These plots show the association between prosecutors' *z*-scores for case and defendant outcomes. Each point in the scatterplot corresponds to a prosecutor's standardized *z*-score. The black lines represent ordinary least-squares regression. Histograms for the standardized *z*-scores corresponding to each outcome are shown on the diagonal. [Color figure can be viewed at wileyonlinelibrary.com]

skills and defendant outcomes on case outcomes. These results are displayed in Tables A.5 and A.6. Substantively, our results remain the same.

6.4 | Case management skills and prosecutor preferences

We have shown so far that a prosecutor's mundane tasks and constraints influence case and defendant outcomes. Prior research has argued (explicitly or implicitly) that individual legal actors have their own preferences or "stringencies," which influence case outcomes. We present some suggestive evidence consistent with the idea that a prosecutor's high-level objectives may matter but

that these preferences may influence their handling of cases much less than the practical realities of the job. As a consequence, variations in case outcomes may reflect variation in prosecutors' proficiency at handling these case management tasks, rather than their preferences. In fact, attorney skills may not necessarily align with their preferences and could potentially yield outcomes contrary to prosecutors' inclinations.

To explore this idea, we compare the effects of prosecutor case management skill with traditional proxies for individual preferences, such as political preferences, race, or gender (Ba et al., 2021; Cohen & Yang, 2019; Myers, 1988). For example, studies have found that Republican judges are more punitive than their Democratic counterparts when judging similarly situated defendants (Cohen & Yang, 2019). We use Pennsylvania voter registration data to identify the political affiliations of prosecutors. We match prosecutors to voting rolls using full names and birth dates, where available. We can identify 66% of prosecutors in the voter rolls: This percentage corresponds to the percentage of Americans who are registered to vote (Pew Charitable Trusts, 2017). We then recreate our analyses showing the association between political affiliation as it appears on voter rolls and prosecutors' case management skills, as shown in Figure A.5.

We find no clear patterns linking conviction rates, formal punishment rates, or case management skills and a prosecutor's political leanings. This analysis has two implications. First, it suggests that case management skills represent a unique dimension, distinct from those captured by other measures of legal actor preferences. Second, it supports our hypothesis that prosecutors' case management skills are first order, taking precedence over more conservative attitudes toward public safety that might be expected to lead to more convictions and punishments. Although our analyses are suggestive, we acknowledge that political affiliation on voter rolls can be a blunt measure of political preferences. In addition, there are few Republican and unaffiliated prosecutors overall. Future research should explore the differences between individual preferences and case management skills in greater detail.

6.5 | Robustness

In this subsection, we re-run our main analyses prior to 2018, which are subject to neither changes in leadership nor COVID-19 disruptions. This approach addresses concerns that our findings in Philadelphia are being driven by the election of a progressive prosecutor, Lawrence Krasner, who took office in January 2018. Upon taking office, DA Krasner fired a large group of seasoned prosecutors and hired new line prosecutors who were judged to be more aligned with his progressive vision. Although the major structures and procedural requirements of the municipal courts did not change under Krasner, these new ADAs may have approached their case management tasks differently. To address this concern, we re-estimate our analyses using a sample of cases that were disposed by January 2018. Although our original dataset only included cases disposed by March 2020, this revised dataset also excludes cases where the recidivism window overlapped with the COVID-19 pandemic. The results of this robustness test are shown in Tables A.7 and A.8. The robustness results are very similar to those found using the main sample.

7 | DISCUSSION AND LIMITATIONS

7.1 | Discussion

It is widely held that prosecutors vary in their decisions, particularly in their conviction and punishment decisions. Prior theories have argued that this variation may be due to different levels of punitiveness, variation in prosecutors' focal concerns such as defendant blameworthiness or risk, or community-level differences in workgroup norms and organizational constraints. Many tests of these theories, however, have been restricted to examining differences in case outcomes and inferring prosecutors' preferences. We build on these prior theories by using data on prosecutors' granular, hearing-level actions. These fine-grained data allow us to test a case management model of prosecutor discretion that is grounded in the practical realities of their job.

We find support for all three hypotheses presented in Section 3. First, we confirm that prosecutors do vary in their case management skill even when handling statistically similar caseloads (H1). Some prosecutors are better at getting key witnesses to appear in court, and some are better at managing discovery in order to get everything shared with the defense. Both of these tasks are critical for cases to proceed. Thus, this variation in case management skill also maps onto variation in case outcomes.

We find that prosecutors who are better case managers are more likely to secure convictions and formal punishments, consistent with our second hypothesis (H2). This finding confirms other recent papers demonstrating the significance of failures to appear and discovery in determining case outcomes (Graef, 2025; Graef et al., 2023).

Finally, prosecutors whose caseloads result in more convictions and formal punishments are also more likely to have defendants on their caseload rearrested and recharged within 2 years (H3). Although findings from prior research on the effects of conviction and incarceration have been mixed, recent papers support the idea that misdemeanor criminal justice involvement in particular has negative effects (Agan et al., 2021). We cannot parse the distinct effects of conviction and formal punishment here; however, we estimate the causal effect of prosecutor assignment on rearrests and new charges and find that the conviction/punishment "bundle" is associated with increased future criminal justice contacts for the defendants on their cases. These future criminal justice contacts could be due to criminogenic effects of conviction and punishments, meaning those defendants are more likely to engage in future criminal behavior, or they could be due to differential system responses, as having a prior conviction may make legal actors more likely to arrest and charge someone for the same future behavior.

Taken together, these findings highlight a fundamental tension in a prosecutor's work, where prosecutors who are typically seen as good workers by office standards also have worse defendant outcomes in terms of future criminal justice contacts. Much prior work on prosecutors has focused on maximizing convictions: Federal prosecutors and felony units have traditionally embraced convictions as a way to measure success (Albonetti, 1986). In the local misdemeanor courts, often prosecutors have no clear way to track convictions; however, in this environment with particularly opaque goals and few salient performance metrics, prosecutors may embrace case-processing efficiency as a way to feel accomplished (Graef, 2025). We find that this emphasis on convictions and case processing may, in fact, perpetuate criminal justice involvement and diminish public safety. Prosecutors who have better case management skills and higher conviction rates are also more likely to have defendants on their caseload rearrested and charged with a new crime. This finding has implications for criminal justice reform. If prosecutors are primarily concerned with

(and rewarded by) efficient office work and convictions, they may also be incentivized to diminish public safety, at least in the misdemeanor case. Criminal justice reforms that seek to improve defendant outcomes and community safety should address the incentives prosecutors face and the organization of their work as the focus on case management may come at the cost of societal goals.

Prior theories demonstrate that prosecutors in some contexts do think about high-level goals and seek to balance them (Frederick & Stemen, 2012; Winter & Clair, 2023). We find, however, that a prosecutor's case management tasks are first order, particularly in the context of the lower courts. The primary way in which prosecutors influence cases in our setting is through their ability to process a case through to disposition, which they accomplish by getting witnesses to appear and making sure discovery is passed. As a consequence, observed variation in case and defendant outcomes is a product of a prosecutor's case management skills, which could be unrelated to preferences. Importantly, this case management model does not rule out other theories of prosecutorial discretion. For example, at sentencing, prosecutors in Philadelphia might consider the focal concerns of defendant blameworthiness, public safety, and resource constraints. However, cases often cannot make it to sentencing unless a prosecutor has succeeded at several case management tasks beforehand.

We also provide some suggestive evidence that case management skill is a distinct new dimension, uncorrelated with individual preferences as captured by political leanings. This analysis should be interpreted with caution, however, as voter registration is an imperfect measure of preferences, and prosecutors in Philadelphia lean heavily Democratic. Future research should investigate the extent to which case management skill correlates with other proxies for prosecutor preferences, such as age, race, gender, and experience, and which dimensions of prosecutor preference or case management skill are most salient.

Similar patterns of case management variation and influence may or may not apply to prosecutors working in other stages of the criminal justice system. For example, at the front end of the system, charging prosecutors may face fewer organizational constraints, which may allow them to consider more abstract goals in their case processing. Aside from occasional discussions with peers and police detectives, these prosecutors often work alone and may have more authority to determine a case's path. Future research should compare variation in prosecutors' granular case management actions and influence at several points in the progression of a case.

We have established that case management tasks are necessary for line prosecutors to achieve convictions; however, a question that remains is whether prosecutors' variation in case management efforts is strategic. For example, prosecutors may work harder to get a witness to appear or make extra efforts to track down discovery in cases that they feel are especially serious or deserving. Concurrent qualitative work finds that although prosecutors do occasionally use case management efforts to triage cases, this is not the norm in the municipal courts (Graef, 2025). Often, prosecutors in the Municipal Court Unit are handling an entire caseload of misdemeanor cases, so they do not have much variation in case severity to triage between (e.g., they do not need to organize their work to prioritize a murder case over a minor drug possession). Even when prosecutors handle a mixed caseload of misdemeanors and felony preliminary hearings, the rookie prosecutors who staff the municipal courts are instructed to process every case forward: The default is to try every case. More work is needed, however, to understand the use of case management as a potential triage mechanism in other jurisdictions, particularly in felony and mixed trial units.

Finally, different theories of legal actor decision-making also have different methodological implications for how we study them. For example, theories of discretion focused on individual characteristics advance the notion that legal actors hold distinct views on punishment and

exhibit differing degrees of punitiveness, or stringency. This idea has led to the development of causal research designs leveraging individual variation. Specifically, examiner leniency instrumental variables designs use the as-if random assignment of cases to legal actors who vary in the leniency of their decisions (Cunningham, 2021). This method has been used to study the effects of pretrial detention (Dobbie et al., 2018; Gupta et al., 2016; Heaton et al., 2017; Leslie & Pope, 2017; Stevenson, 2018), incarceration (Loeffler & Nagin, 2022), misdemeanor prosecution (Agan et al., 2021), and more. Although this work provides strong causal estimates, it is agnostic about what is driving differences in outcomes—which is a separately important question to understand prosecutor decision-making. We show that although individual variation exists, it may simply be due to variation in case management skills, which may be orthogonal or even opposite to an actor's stringency. Greater nuance is needed to understand the context of variation in legal actor behaviors as policy prescriptions based on ideas of individual stringency may be quite different than those based on case management skill or varying organizational constraints.

7.2 | Generalizability and limitations

Our study focuses on the Philadelphia court system, which is a large, urban court system. Although our study starts in 2010, the Philadelphia DAO has been led by a reform-oriented DA, Lawrence Krasner, since 2018. Although every court system is distinct, we believe our findings have broader implications as many urban jurisdictions share similar institutional features. For instance, larger prosecutors' offices often separate work into units (e.g., a misdemeanor unit, a felony unit, a homicide unit), with prosecutors primarily handling cases of a similar complexity or type. As such, prosecutors in these systems are not typically balancing misdemeanor cases with high-stakes homicide trials, allowing for more consistent and focused processing of cases. Second, although Philadelphia does have a higher-than-average use of bench trials, several other jurisdictions across the country report significant use of bench trials (Court Statistics Project, 2023), and case management tasks may play a different, although similarly critical, role in systems of plea bargaining as well. In addition, our results are not driven by the presence of a progressive DA: As demonstrated in Section 6.5, our findings hold even when cases handled under the current DA are excluded. Although most criminal cases in the United States are processed through large, urban systems like Philadelphia,[23] rural justice systems may have organizational and structural differences that shape case processing in different ways. Future research should explore the role of case management in rural prosecutors' offices.

By focusing on one jurisdiction, we can gather rich data on hearing-level prosecutor behaviors that has not been available to researchers before at this scale; however, focusing on one jurisdiction also means we cannot examine case management tasks across court systems or different organizational contexts. One organizational feature across which case management may vary is whether the office operates under a horizontal or vertical prosecution model. In a horizontal model, cases are passed to a new ADA at each stage in the prosecution process (or even at every hearing, as in Philadelphia). In a vertical model, the same prosecutor handles a case from charging to disposition. A vertical model may allow prosecutors to build relationships with witnesses, potentially altering that aspect of case management. It may also give prosecutors more incentive to withdraw cases instead of passing them on to a colleague. Thus, even though our findings are relevant to

[23] According to data from the US Census Bureau, approximately 80% of Americans live in large, urban areas with a population of 1 million or more (U.S. Census Bureau, 2023).

other large, urban systems, further research is needed to understand the form and effects of case management across different types of court systems and prosecutorial environments.

Our study is also limited to a focus on misdemeanors, although aspects of the case management model likely apply to felonies as well.[24] We focus on misdemeanors because even though felonies also begin with a preliminary hearing in municipal courts, they ultimately are disposed through a different process that often involves multiple DAO units. As such, we focus on misdemeanor cases that are disposed in the municipal courts in order to simplify our model. The process for handling felony preliminary hearings, however, is nearly identical to the process for misdemeanor bench trials. At each felony hearing, prosecutors must manage their witnesses and pass discovery; only when these tasks are complete can the case proceed to a preliminary hearing. And although many more felony cases are resolved through guilty pleas, no plea offers are made until a felony has passed its preliminary hearing and been "held for court"—at which point, they are no longer in municipal court. In this way, case management tasks are still first order for felony cases during the initial stages of case processing, which occur in Philadelphia's municipal courts, and are handled by the DAO's Municipal Court Unit. Beyond the municipal courts and in other settings, felony case processing may or may not follow the case management model. Future research should explore the role of case management skills in shaping the outcomes of felony cases in different contexts.

Finally, we acknowledge some potential limitations with regard to our data and methodology. Although benchmarking performs well at balancing on observable features, a possibility for omitted variable bias always exists. We note, however, that omitted variable bias is unlikely in our setting, for two reasons. First, due to the way cases are assigned to ADAs, little opportunity exists for selection on unobservables in our setting. Second, Figures A.2 and A.3 present evidence that benchmarking helps us achieve balance even on variables we did not explicitly benchmark on, such as proxies for evidence strength and prior ADA decisions. We also face a common data challenge: Namely, it is quite difficult to gather nuanced measures of subjective and contextual factors, such as case complexity and individual political preferences. Although we use arrest narrative length as a proxy for evidence strength and voter registration as a proxy for political preferences, we recognize that these measures are rough approximations of these concepts. Future work with richer data on these contextual factors and random assignment of cases to prosecutors could help address these limitations more directly.

8 | CONCLUSION

Prosecutors play a central role in the criminal justice system, yet we still know little about how their day-to-day decisions shape case outcomes. Research on this topic is difficult because of the lack of detailed data on prosecutorial behavior and because it is hard to separate the role of individual discretion from differences in caseloads. In this study, we use benchmarking methods and detailed administrative data on misdemeanor cases to shed new light on the scope and the consequences of variation across prosecutors in their daily practices.

We find that prosecutors in the misdemeanor courts vary in their granular actions in ways that impact ultimate case outcomes, supporting the case management theory of prosecutor discretion that we propose in this article. This variation may be tempered by organizational features of the misdemeanor courts, including Philadelphia's horizontal prosecution model; future research

[24] We also have sample restrictions necessary to conduct our data analyses; but as shown in Table A1, and as discussed in Section 5.1, the misdemeanors that we consider are very similar to all misdemeanors in Philadelphia.

should explore prosecutor case management variation under different models and contexts. We also examine how prosecutors' decisions map onto case and defendant outcomes, finding that defendants are better off under prosecutors who are less likely to secure a conviction. Together, these findings suggest a fundamental tension in a prosecutor's work, where good office work does not seem to translate into good policy.

Future research should expand this work to other decision points and other stages of the court system as prosecutors may face different case management challenges or be more or less constrained at various points in the system. Future work should also compare prosecutor actions and influence across these points to help determine where interventions aimed at guiding prosecutor discretion are most likely to be effective. Finally, the prominence of case management tasks and differences in variation in decision-making at different stages also means that we need to think carefully about the use and interpretation of causal research designs that leverage the assignment of cases to legal actors.

ACKNOWLEDGMENTS

We are extremely grateful to the Philadelphia District Attorney's Office and its Transparency Analytics (DATA) Lab for making this project possible. For their helpful comments, we thank C. J. Arayata, Francesca Arruda de Amaral, Yasmin Ayala-Johnson, Mariel Delacruz, Oren Gur, Paul Heaton, Mike Hollander, Charles Loeffler, John MacDonald, Tyrell Mann-Barnes, Sandy Mayson, Viet Nguyen, Greg Ridgeway, Tyler Tran, and Wes Weaver. We gratefully acknowledge funding from Arnold Ventures. This work was produced independently and does not necessarily reflect the view of the Philadelphia District Attorney's Office.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Lindsay Graef https://orcid.org/0009-0008-3221-0164

REFERENCES

Abrams, D. (2011). Is pleading really a bargain? *Journal of Empirical Legal Studies*, *8*(s1), 200–221. https://doi.org/10.1111/j.1740-1461.2011.01234.x

Abrams, D., Nguyen, V., Ouss, A., & Reinhold, J. (2024). *A (plea) offer you can refuse* [Working paper].

Abrams, D., & Yoon, A. (2007). The luck of the draw: Using random case assignment to investigate attorney ability. *University of Chicago Law Review*, *74*, 1145. https://doi.org/10.2307/20141859

Agan, A., Doleac, J. L., & Harvey, A. (2021). *Misdemeanor prosecution*. (Working paper no. 28600; Working Paper Series). National Bureau of Economic Research. https://doi.org/10.3386/w28600

Agan, A., & Starr, S. (2017). The effect of criminal records on access to employment. *American Economic Review*, *107*(5), 560–564. https://doi.org/10.1257/aer.p20171003

Albonetti, C. A. (1986). Criminality, prosecutorial screening, and uncertainty: Toward a theory of discretionary decision making in felony case processings. *Criminology*, *24*(4), 623–644. https://doi.org/10.1111/j.1745-9125.1986.tb01505.x

Albonetti, C. A. (1987). Prosecutorial discretion: The effects of uncertainty. *Law & Society Review*, *21*(2), 291–313. https://doi.org/10.2307/3053523

Anderson, J. M., & Heaton, P. (2012). How much difference does the lawyer make? The effect of defense counsel on murder case outcomes. *Yale Law Journal*, *122*, 154–217. https://doi.org/10.2139/ssrn.1884379

Ba, B. A., Knox, D., Mummolo, J., & Rivera, R. (2021). The role of officer race and gender in police-civilian interactions in Chicago. *Science*, *371*(6530), 696–702. https://doi.org/10.1126/science.abd8694

Baker, D. B., & Hassan, S. (2021). Gender and prosecutorial discretion: An empirical assessment. *Journal of Public Administration Research and Theory*, *31*(1), 73–90. https://doi.org/10.1093/jopart/muaa017

Bang, H., & Robins, J. M. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics*, *61*(4), 962–973. https://doi.org/10.1111/j.1541-0420.2005.00377.x

Bazelon, E. (2019). *Charged: The new movement to transform American prosecution and end mass incarceration*. Random House. https://www.barnesandnoble.com/w/charged-emily-bazelon/1129473779

Berdejo, C. (2018). Criminalizing race: Racial disparities in plea-bargaining. *Boston College Law Review*, *59*(4), 1189–1250.

Black, D. J., & Reiss, A. J. (1970). Police control of juveniles. *American Sociological Review*, *35*(1), 63. https://doi.org/10.2307/2093853

Boyd, C. L. (2016). Representation on the courts? The effects of trial judges' sex and race. *Political Research Quarterly*, *69*(4), 788–799.

Boyd, C. L., & Nelson, M. J. (2017). The effects of trial judge gender and public opinion on criminal sentencing decisions. *Vanderbilt Law Review*, *70*(6), 1819.

Bushway, S. D., Redlich, A. D., & Norris, R. J. (2014). An explicit test of plea bargaining in the "shadow of the trial". *Criminology*, *52*(4), 723–754. https://doi.org/10.1111/1745-9125.12054

Chiricos, T. G., & Bales, W. D. (1991). Unemployment and punishment: An empirical assessment. *Criminology*, *29*(4), 701–724. https://doi.org/10.1111/j.1745-9125.1991.tb01085.x

Cohen, A., & Yang, C. S. (2019). Judicial politics and sentencing decisions. *American Economic Journal: Economic Policy*, *11*(1), 160–191.

Collins, T., & Moyer, L. (2008). Gender, race, and intersectionality on the federal appellate bench. *Political Research Quarterly*, *61*(2), 219–227. https://doi.org/10.1177/1065912907306467

Court Statistics Project. (2023, October 9). *CSP STAT criminal*. Court Statistics Project. https://www.courtstatistics.org/court-statistics/interactive-caseload-data-displays/csp-stat-nav-cards-first-row/csp-stat-criminal

Crow, M. S., & Adrion, B. (2011). Focal concerns and police use of force: Examining the factors associated with taser use. *Police Quarterly*, *14*(4), 366–387. https://doi.org/10.1177/1098611111423740

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale University Press. https://mixtape.scunning.com/instrumental-variables.html

Davis, R. C., O'Sullivan, C. S., Farole, D. J., & Rempel, M. (2008). A comparison of two prosecution policies in cases of intimate partner violence: Mandatory case filing versus following the victim's lead. *Criminology & Public Policy*, *7*(4), 633–662. https://doi.org/10.1111/j.1745-9133.2008.00532.x

Davis, R. C., Smith, B. E., & Taylor, B. (2003). Increasing the proportion of domestic violence arrests that are prosecuted: A natural experiment in Milwaukee. *Criminology & Public Policy*, *2*(2), 263–282. https://doi.org/10.1111/j.1745-9133.2003.tb00125.x

Demuth, S. (2003). Racial and ethnic differences in pretrial release decisions and outcomes: A comparison of Hispanic, Black, and White felony arrestees. *Criminology*, *41*(3), 873–908. https://doi.org/10.1111/j.1745-9125.2003.tb01007.x

Didwania, S. H. (2022). Gender favoritism among criminal prosecutors. *The Journal of Law and Economics*, *65*(1), 77–104. https://doi.org/10.1086/718463

Dixon, J. (1995). The organizational context of criminal sentencing. *American Journal of sociology*, *100*(5), 1157–1198.

Dobbie, W., Goldin, J., & Yang, C. S. (2018). The effects of pre-trial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review*, *108*(2), 201–240. https://doi.org/10.1257/aer.20161503

Earl, J. (2019). The process is the punishment revisited: Forty-year anniversary review. In M. M. Feeley (Ed.), *A Lower criminal courts* (pp. 24–31). Routledge.

Eisenstein, J., Flemming, R. B., & Nardulli, P. F. (1988). *The contours of justice: Communities and their courts*. Little, Brown.

Eisenstein, J., & Jacob, H. (1977). *Felony justice: An organizational analysis of criminal courts*. Little, Brown.

Feeley, M. M. (1979). *The process is the punishment: Handling cases in a lower criminal court*. Russell Sage Foundation.

Fischman, J. B., & Schanzenbach, M. M. (2012). Racial disparities under the federal sentencing guidelines: The role of judicial discretion and mandatory minimums. *Journal of Empirical Legal Studies*, *9*(4), 729–764. https://doi.org/10.1111/j.1740-1461.2012.01266.x

Frederick, B., & Stemen, D. (2012). *The anatomy of discretion: An analysis of prosecutorial decision making—Technical report*. National Institute of Justice.

Frohmann, L. (1997). Convictability and discordant locales: Reproducing race, class, and gender ideologies in prosecutorial decisionmaking. *Law & Society Review*, *31*(3), 531–556. https://doi.org/10.2307/3054045

Graef, L. (2025). *Prosecutors as case managers* [Working paper].

Graef, L., Mayson, S., Stevenson, M. T., & Ouss, A. (2023). Systemic failure to appear in court. *University of Pennsylvania Law Review*, *172*(1), 1–60.

Gupta, A., Hansman, C., & Frenchman, E. (2016). The heavy costs of high bail: Evidence from judge randomization. *The Journal of Legal Studies*, *45*(2), 471–505. https://doi.org/10.1086/688907

Harris, A. P., & Sen, M. (2022). *How judges' professional experience impacts case outcomes: An examination of public defenders and criminal sentencing* [Working paper].

Hartley, R. D., Maddan, S., & Spohn, C. C. (2007). Prosecutorial discretion: An examination of substantial assistance departures in federal crack-cocaine and powder-cocaine cases. *Justice Quarterly*, *24*(3), 382–407. https://doi.org/10.1080/07418820701485379

Hartley, R. D., & Tillyer, R. (2018). Examining prosecutorial discretion in federal criminal cases: Legal and extralegal determinants of declination and charge change decisions. *Justice Quarterly*, *35*(7), 1195–1225. https://doi.org/10.1080/07418825.2018.1530367

Haynes, S. H., Ruback, B., & Cusick, G. R. (2010). Courtroom workgroups and sentencing: The effects of similarity, proximity, and stability. *Crime & Delinquency*, *56*(1), 126–161. https://doi.org/10.1177/0011128707313787

Heaton, P., Graef, L., & Ridgeway, G. (2025). *Prosecutor backgrounds and charging decisions* [Working paper].

Heaton, P., Mayson, S., & Stevenson, M. (2017). The downstream consequences of misdemeanor pretrial detention. *Stanford Law Review*, *69*, 711.

Henning, K., & Feder, L. (2005). Criminal prosecution of domestic violence offenses: An investigation of factors predictive of court outcomes. *Criminal Justice and Behavior*, *32*(6), 612–642. https://doi.org/10.1177/0093854805279945

Hester, R. (2017). Judicial rotation as centripetal force: Sentencing in the court communities of South Carolina. *Criminology*, *55*(1), 205–235. https://doi.org/10.1111/1745-9125.12132

Ho, D., Imai, K., King, G., & Stuart, E. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*, 199–236.

Humphries, J. E., Ouss, A., Stavreva, K., Stevenson, M. T., & Van Dijk, W. (2025). Conviction, incarceration, and recidivism: Understanding the revolving door. *The Quarterly Journal of Economics*, *140*(4), 2907–2962.

Iyengar, R. (2007). *An analysis of the performance of federal indigent defense counsel* (Working paper no. 13187). National Bureau of Economic Research. https://doi.org/10.3386/w13187

Johnson, B. D. (2003). Racial and ethnic disparities in sentencing departures across modes of conviction. *Criminology*, *41*(2), 449–490. https://doi.org/10.1111/j.1745-9125.2003.tb00994.x

Johnson, B. D. (2005). Contextual disparities in guidelines departures: Courtroom social contexts, guidelines compliance, and extralegal disparities in criminal sentencing. *Criminology*, *43*(3), 761–796. https://doi.org/10.1111/j.0011-1348.2005.00023.x

Johnson, B. D. (2006). The multilevel context of criminal sentencing: Integrating Judge- and county-level influences. *Criminology*, *44*(2), 259–298. https://doi.org/10.1111/j.1745-9125.2006.00049.x

Johnson, B. D. (2018). In search of the missing link: Examining contextual variation in federal charge bargains across U.S. district courts. *Justice Quarterly*, *35*(7), 1133–1165. https://doi.org/10.1080/07418825.2018.1523447

Johnson, B. D., & Betsinger, S. (2009). Punishing the "model minority": Asian-American criminal sentencing outcomes in federal district courts. *Criminology*, *47*(4), 1045–1090. https://doi.org/10.1111/j.1745-9125.2009.00169.x

Johnson, B. D., Ulmer, J. T., & Kramer, J. H. (2008). The social context of guidelines circumvention: The case of federal district courts. *Criminology*, *46*(3), 737–783. https://doi.org/10.1111/j.1745-9125.2008.00125.x

Kautt, P. M. (2002). Location, location, location: Interdistrict and intercircuit variation in sentencing outcomes for federal drug-trafficking offenses. *Justice Quarterly*, *19*(4), 633–671. https://doi.org/10.1080/07418820200095381

Kim, B., Spohn, C., & Hedberg, E. C. (2015). Federal sentencing as a complex collaborative process: Judges, prosecutors, judge–prosecutor dyads, and disparity in sentencing. *Criminology*, *53*(4), 597–623. https://doi.org/10.1111/1745-9125.12090

King, R. D., Johnson, K. R., & McGeever, K. (2010). Demography of the legal profession and racial disparities in sentencing. *Law & Society Review*, *44*(1), 1–32. https://doi.org/10.1111/j.1540-5893.2010.00394.x

Kohler-Hausmann, I. (2019). *Misdemeanorland*. Princeton University Press.

Kramer, J. H., & Ulmer, J. T. (2002). Downward departures for serious violent offenders: Local court "corrections" to Pennsylvania's sentencing guidelines. *Criminology*, *40*(4), 897–932. https://doi.org/10.1111/j.1745-9125.2002.tb00977.x

Kutateladze, B. L. (2018). Tracing charge trajectories: A study of the influence of race in charge changes at case screening, arraignment, and disposition. *Criminology*, *56*(1), 123–153. https://doi.org/10.1111/1745-9125.12166

Kutateladze, B. L., & Lawson, V. Z. (2018). Is a plea really a bargain? An analysis of plea and trial dispositions in New York City. *Crime & Delinquency*, *64*(7), 856–887. https://doi.org/10.1177/0011128717695224

Kutateladze, B. L., Lawson, V. Z., & Andiloro, N. R. (2015). Does evidence really matter? An exploratory analysis of the role of evidence in plea bargaining in felony drug cases. *Law and Human Behavior*, *39*(5), 431–442. https://doi.org/10.1037/lhb0000142

Leasure, P. (2019). Misdemeanor records and employment outcomes: An experimental study. *Crime & Delinquency*, *65*(13), 1850–1872. https://doi.org/10.1177/0011128718806683

Leslie, E., & Pope, N. G. (2017). The unintended impact of pretrial detention on case outcomes: Evidence from New York City arraignments. *The Journal of Law and Economics*, *60*(3), 529–557. https://doi.org/10.1086/695285

Lim, C. S. H., Silveira, B. S., & Snyder, J. M. (2016). Do judges' characteristics matter? Ethnicity, gender, and partisanship in Texas state trial courts. *American Law and Economics Review*, *18*(2), 302–357. https://doi.org/10.1093/aler/ahw006

Loeffler, C. E., & Nagin, D. S. (2022). The impact of incarceration on recidivism. *Annual Review of Criminology*, *5*(1), 133–152. https://doi.org/10.1146/annurev-criminol-030920-112506

Lynch, M. (2016). *Hard bargains* (1st ed.). Russell Sage Foundation.

Lynch, M. (2018). Prosecutorial discretion, drug case selection, and inequality in federal court. *Justice Quarterly*, *35*(7), 1309–1336. https://doi.org/10.1080/07418825.2018.1535083

Lynch, M., Barno, M., & Omori, M. (2021). Prosecutors, court communities, and policy change: The impact of internal DOJ reforms on federal prosecutorial practices. *Criminology*, *59*(3), 480–519. https://doi.org/10.1111/1745-9125.12275

Lynch, M., & Omori, M. (2014). Legal change and sentencing norms in the wake of Booker: The impact of time and place on drug trafficking cases in federal court. *Law & Society Review*, *48*(2), 411–445.

Mayson, S. G., & Stevenson, M. T. (2020). Misdemeanors by the numbers. *Boston College Law Review*, *61*(3), 971–1044.

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004, December). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, *9*(4), 403–425.

Myers, M. A. (1982). Common law in action: The prosecution of felonies and misdemeanors. *Sociological Inquiry*, *52*(1), 1–15. https://doi.org/10.1111/j.1475-682X.1982.tb01235.x

Myers, M. A. (1988). Social background and the sentencing behavior of judges. *Criminology*, *26*(4), 649–676. https://doi.org/10.1111/j.1745-9125.1988.tb00858.x

Nagel, I. H., & Schulhofer, S. J. (1992). A tale of three cities: An empirical study of charging and bargaining practices under the federal sentencing guidelines. *Southern California Law Review*, *66*(1), 501–566.

Natapoff, A. (2011). Misdemeanors. *Southern California Law Review*, *85*(5), 1313–1376.

Natapoff, A. (2018). *Punishment without crime*. Basic Books. https://www.hachettebookgroup.com/titles/alexandra-natapoff/punishment-without-crime/9780465093809/?lens=basic-books

Nguyen, V., & Graef, L. (2025). *The net-narrowing potential of diversion: A theoretical and empirical assessment of prosecutor-led, adult diversion* [Manuscript conditionally accepted]. University of Pennsylvania.

Nguyen, V., & Ridgeway, G. (2023). Judges on the benchmark: Developing a sentencing feedback system. *Justice Quarterly*, *41*(1), 1–37. Advance online publication. https://doi.org/10.1080/07418825.2023.2244738

Ouss, A., & Nguyen, V. (2025). *Prosecutorial policies and courtroom practices* [Manuscript conditionally accepted]. University of Pennsylvania.

Ouss, A., & Stevenson, M. (2023). Does cash bail deter misconduct? *American Economic Journal: Applied Economics*, *15*(3), 150–182. https://doi.org/10.1257/app.20210349

Owusu, F. (2022). *Presumptive declination and diversion in Suffolk County, MA*. Harvard Kennedy School Rappaport. Institute for Greater Boston. https://www.hks.harvard.edu/sites/default/files/Taubman/RIGB/Presumptive%20Declination%20and%20Diversion%20in%20Suffolk%20County%2C%20MA.pdf

Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology*, *108*(5), 937–975.

Pew Charitable Trusts. (2017, June 21). *Why are millions of citizens not registered to vote?* http://pew.org/2sQgul2

Pfaff, J. (2017). *Locked in the true causes of mass incarceration – and how to achieve real reform*. Basic Books.

Rachlinski, J. J., & Wistrich, A. J. (2017). Judging the judiciary by the numbers: Empirical research on judges. *Annual Review of Law and Social Science*, *13*(1), 203–229. https://doi.org/10.1146/annurev-lawsocsci-110615-085032

Rehavi, M. M., & Starr, S. B. (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy*, *122*(6), 1320–1354. https://doi.org/10.1086/677255

Ridgeway, G., & MacDonald, J. M. (2009). Doubly robust internal benchmarking and false discovery rates for detecting racial bias in police stops. *Journal of the American Statistical Association*, *104*(486), 661–668. https://doi.org/10.1198/jasa.2009.0034

Ridgeway, G., & MacDonald, J. M. (2014). A method for internal benchmarking of criminal justice system performance. *Crime & Delinquency*, *60*(1), 145–162. https://doi.org/10.1177/0011128713514802

Ridgeway, G., Moyer, R. A., & Bushway, S. D. (2020). Sentencing scorecards: Reducing racial disparities in prison sentences at their source. *Criminology & Public Policy*, *19*(4), 1113–1138. https://doi.org/10.1111/1745-9133.12529

Roberts, J. (2021). Prosecuting misdemeanors. In R. F. Wright, K. L. Levine, & R. M. Gold (Eds.), *The Oxford handbook of prosecutors and prosecution* (pp. 512–538). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190905422.013.26

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Shermer, L. O., & Johnson, B. D. (2010). Criminal prosecutions: Examining prosecutorial discretion and charge reductions in U.S. federal district courts. *Justice Quarterly*, *27*(3), 394–430. https://doi.org/10.1080/07418820902856972

Sloan, C. (2019). *Racial bias by prosecutors: Evidence from random assignment* [Working paper].

Spohn, C. (1991). Decision making in sexual assault cases: Do Black and female judges make a difference? *Women & Criminal Justice*, *2*(1), 83–105. https://doi.org/10.1300/J012v02n01_06

Spohn, C. (2009). Race, sex, and pretrial detention in federal court: Indirect effects and cumulative disadvantage. *Kansas Law Review*, *57*, 879–901. https://doi.org/10.17161/1808.20099

Spohn, C., Beichner, D., & Davis-Frenzel, E. (2001). Prosecutorial justifications for sexual assault case rejection: Guarding the "gateway to justice". *Social Problems*, *48*(2), 206–235. https://doi.org/10.1525/sp.2001.48.2.206

Spohn, C., & Cederblom, J. (1991). Race and disparities in sentencing: A test of the liberation hypothesis. *Justice Quarterly*, *8*(3), 305–327. https://doi.org/10.1080/07418829100091071

Spohn, C., & Fornango, R. (2009). U.S. attorneys and substantial assistance departures: Testing for interprosecutor disparity. *Criminology*, *47*(3), 813–846. https://doi.org/10.1111/j.1745-9125.2009.00163.x

Spohn, C., Gruhl, J., & Welch, S. (1987). The impact of the ethnicity and gender of defendants on the decision to reject or dismiss felony charges. *Criminology*, *25*(1), 175–192. https://doi.org/10.1111/j.1745-9125.1987.tb00794.x

Steffensmeier, D., & Hebert, C. (1999). Women and men policymakers: Does the judge's gender affect the sentencing of criminal defendants? *Social Forces*, *77*(3), 1163.

Steffensmeier, D., Kramer, J., & Streifel, C. (1993). Gender and imprisonment decisions. *Criminology*, *31*(3), 411–446. https://doi.org/10.1111/j.1745-9125.1993.tb01136.x

Steffensmeier, D., Ulmer, J., & Kramer, J. (1998). The interaction of race, gender, and age in criminal sentencing: The punishment cost of being young, Black, and male. *Criminology*, *36*(4), 763–798. https://doi.org/10.1111/j.1745-9125.1998.tb01265.x

Stemen, D., & Escobar, G. (2018). Whither the prosecutor? Prosecutor and county effects on guilty plea outcomes in Wisconsin. *Justice Quarterly*, *35*(7), 1166–1194. https://doi.org/10.1080/07418825.2018.1548637

Stevenson, M., & Mayson, S. (2018). The scale of misdemeanor justice. *Boston University Law Review*, *98*(3), 731–778.

Stevenson, M. T. (2018). Distortion of justice: How the inability to pay bail affects case outcomes. *The Journal of Law, Economics, and Organization*, *34*(4), 511–542. https://doi.org/10.1093/jleo/ewy019

Stewart, R., & Uggen, C. (2020). Criminal records and college admissions: A modified experimental audit. *Criminology*, *58*(1), 156–188. https://doi.org/10.1111/1745-9125.12229

Uggen, C., Vuolo, M., Lageson, S., Ruhland, E., & K Whitham, H. (2014). The edge of stigma: An experimental audit of the effects of low-level criminal records on employment. *Criminology*, *52*(4), 627–654. https://doi.org/10.1111/1745-9125.12051

Ulmer, J. T. (2005). The localized uses of federal sentencing guidelines in four U.S. district courts: Evidence of processual order. *Symbolic Interaction*, *28*(2), 255–279. https://doi.org/10.1525/si.2005.28.2.255

Ulmer, J. T., & Johnson, B. (2004). Sentencing in context: A multilevel analysis. *Criminology*, *42*(1), 137–178. https://doi.org/10.1111/j.1745-9125.2004.tb00516.x

Ulmer, J. T., Kurlychek, M. C., & Kramer, J. H. (2007). Prosecutorial discretion and the imposition of mandatory minimum sentences. *Journal of Research in Crime and Delinquency*, *44*(4), 427–458. https://doi.org/10.1177/0022427807305853

U.S. Census Bureau. (2023). *Metropolitan and micropolitan statistical areas population totals: 2020-2023*. https://www.census.gov/data/tables/time-series/demo/popest/2020s-total-metro-and-micro-statistical-areas.html

Wilson, J. Q. (1978). *Varieties of police behavior*. Harvard University Press.

Winter, A. S., & Clair, M. (2023). "The roughest form of social work:" How court officials justify bail decisions. *Criminology*, *61*(4), 904–928. https://doi.org/10.1111/1745-9125.12350

Wright, R. F., & Levine, K. L. (2014). The cure for young prosecutors' syndrome. *SSRN Electronic Journal*, *56*, 65. https://doi.org/10.2139/ssrn.2405137

Wu, J., & D'Angelo, J. M. (2014). Unwarranted disparity in federal sentencing: Noncitizen crime as a social/group threat. *Criminal Justice Review*, *39*(1), 58–80. https://doi.org/10.1177/0734016814522644

Yang, C. S. (2013). Have inter-judge sentencing disparities increased in an advisory guidelines regime? Evidence from Booker. *SSRN Electronic Journal*, *89*, 1268–1342. https://doi.org/10.2139/ssrn.2348140

How to cite this article: Graef, L., & Ouss, A. (2025). The role of case management in misdemeanor prosecution. *Criminology*, 1–48. https://doi.org/10.1111/1745-9125.70033

APPENDIX

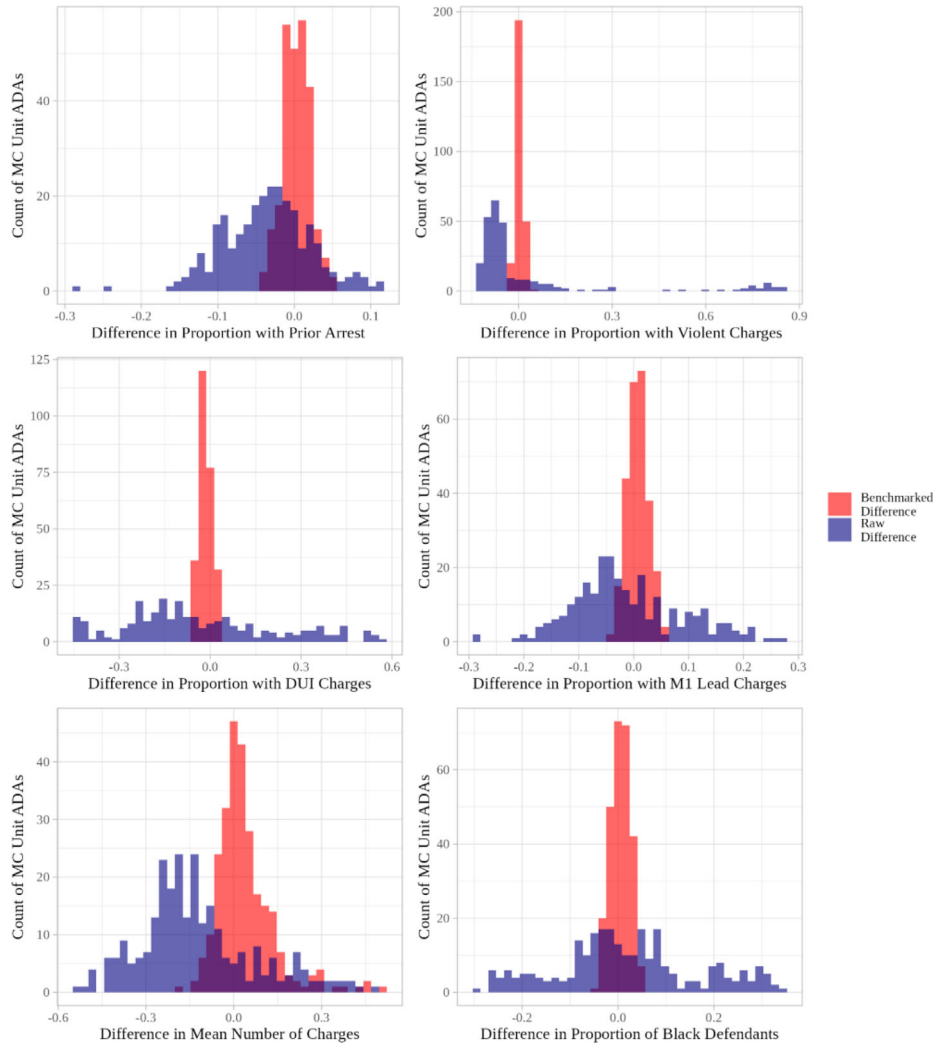

FIGURE A.1 Caseload balance by observable characteristics. *Note*: This figure shows evidence of overall balance achieved by the benchmarking process for six selected case and defendant characteristics (priors, violent, driving under the influence [DUI], M1 lead charge, number of charges, and Black). The *x*-axis shows the difference in the proportion (or mean) of a Municipal Court (MC) Unit assistant district attorney's (ADA's) caseload with that characteristic. For example, in the top right-hand plot, the blue bars show the distribution of the raw difference in the proportion of violent cases in a given ADA's caseload compared with all other cases (the unweighted control group). The red bars show the distribution of the difference in the proportion of violent cases for a given ADA's caseload compared with their individualized, internal benchmark (the weighted controls). These plots show that the benchmarking process has done a good job balancing cases for ADAs overall: Even though the blue histograms are widely spread, the red histograms are tightly centered near zero. [Color figure can be viewed at wileyonlinelibrary.com]

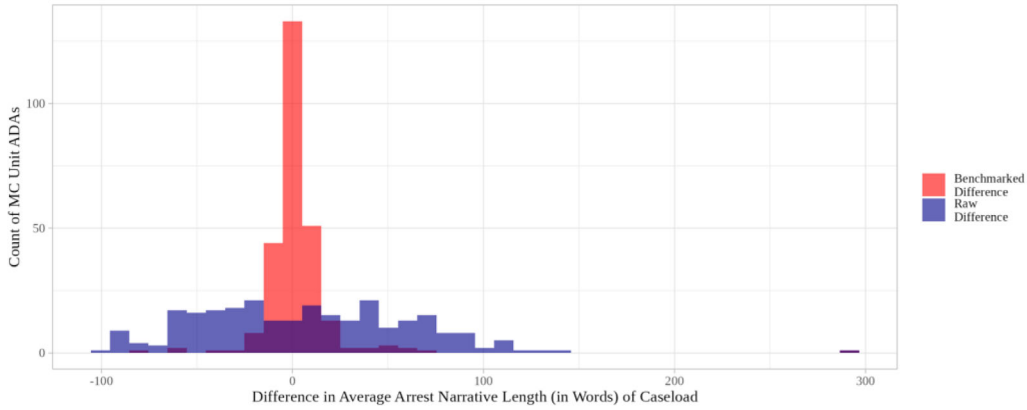

FIGURE A.2 Case balance for strength of evidence. *Note*: This figure shows balance on our proxy of strength of evidence achieved by the benchmarking process. We use arrest narrative length, or the number of words written in the police report at arrest, as a measure of strength of evidence. The *x*-axis shows the difference in the average arrest narrative length of a Municipal Court (MC) Unit assistant district attorney's (ADA's) caseload. The blue bars show the distribution of the raw difference in average arrest narrative length for a given ADA's caseload compared with all other cases (the unweighted control group). The red bars show the distribution of the difference in average arrest narrative length for a given ADA's caseload compared with their individualized, internal benchmark (the weighted controls). This plot shows that although we do not explicitly balance on arrest narrative length, the benchmarking process has also done a fair job of balancing cases on strength of evidence: Even though the blue histogram is widely spread, the red histogram is tightly centered near zero. [Color figure can be viewed at wileyonlinelibrary.com]

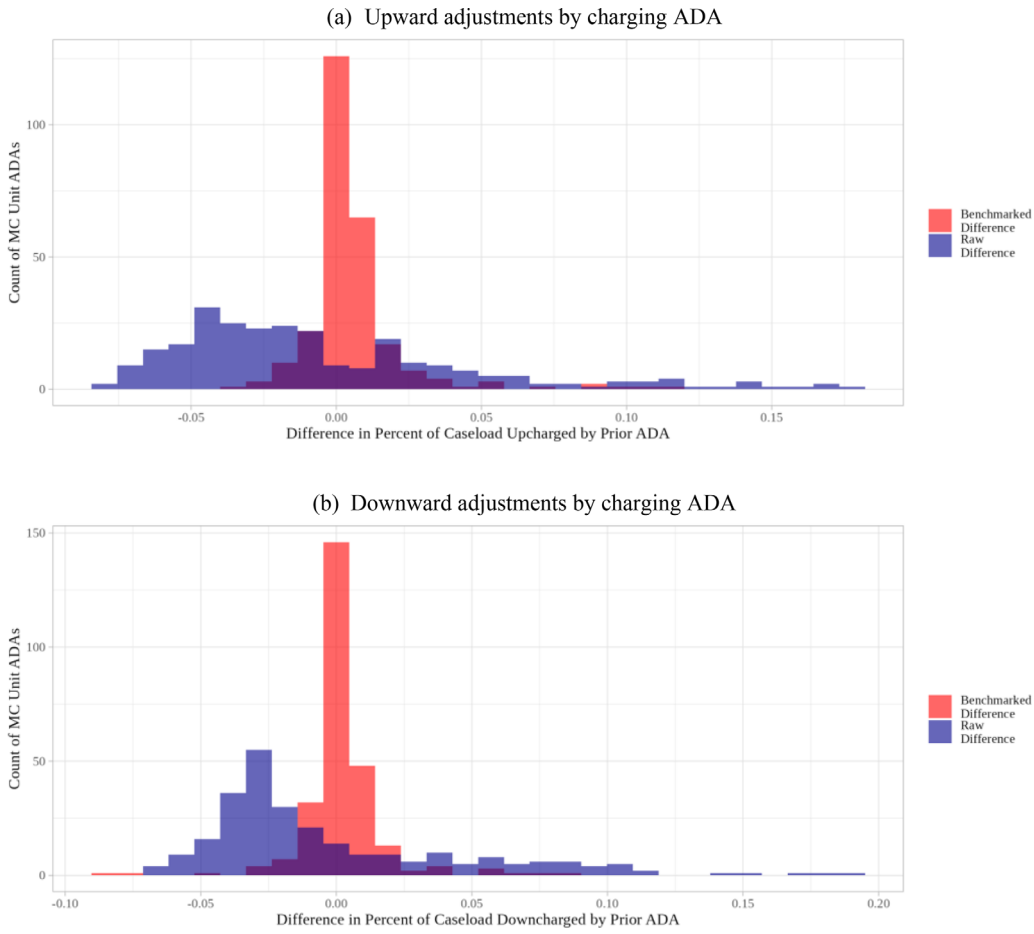

FIGURE A.3 Case balance for prior ADA decisions. *Note*: This figure shows balance on unobservables achieved by the benchmarking process. The *x*-axis shows the difference in the percentage of a Municipal Court (MC) Unit assistant district attorney's (ADA's) caseload where a prior ADA at the charging phase adjusted charges either upward or downward from the police recommended charges. The blue bars show the distribution of the raw difference in upward adjustments for a given ADA's caseload compared with all other cases (the unweighted control group). The red bars show the distribution of the difference for a given ADA's caseload compared with their individualized, internal benchmark (the weighted controls). This plot shows that although we do not explicitly balance on upward and downward adjustments, the benchmarking process has also done a fair job of balancing cases on this metric: Even though the blue histograms are widely spread, the red histograms are tightly centered near zero. [Color figure can be viewed at wileyonlinelibrary.com]

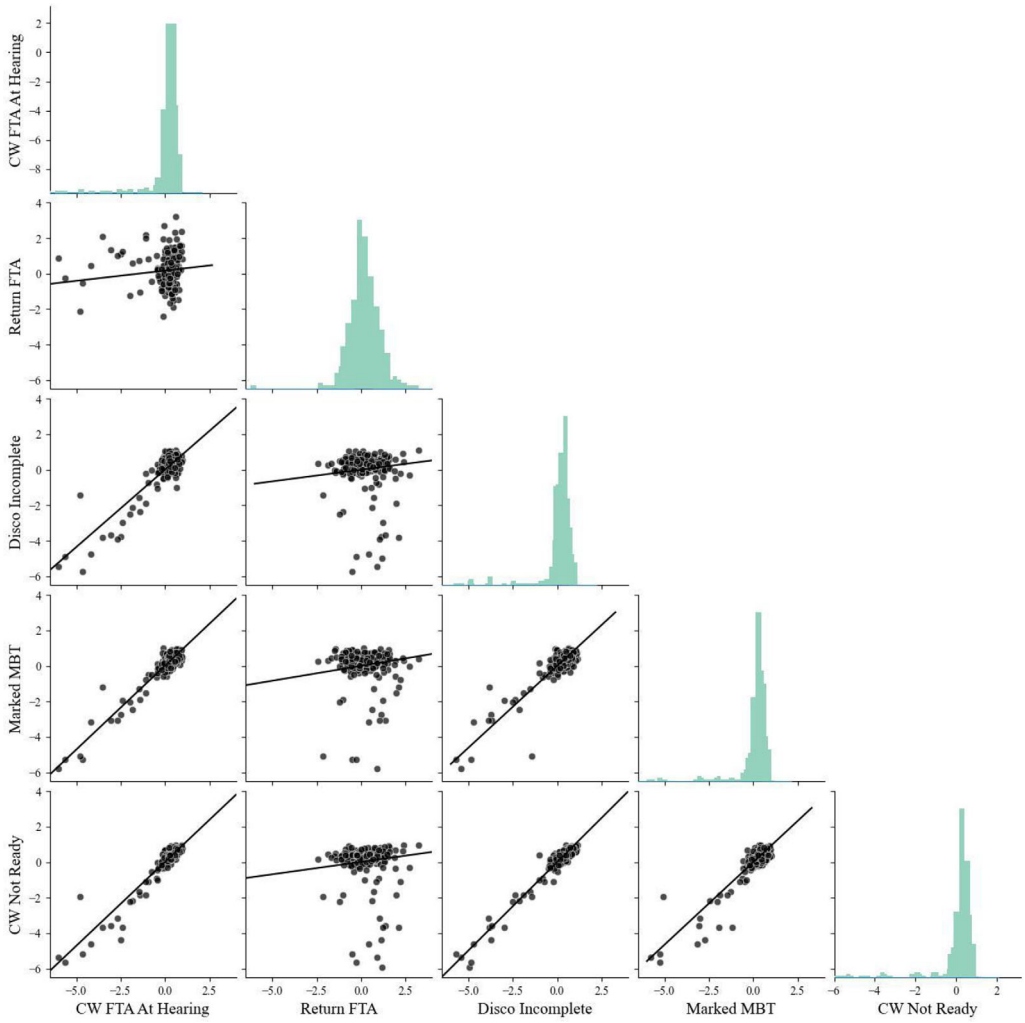

FIGURE A.4 Associations between case management skills (hearing outcomes). *Note*: These plots show the association between prosecutors' *z*-scores for case management skills: Commonwealth (CW) failure to appear (FTA) at current hearing, return FTA (CW FTA at the next hearing), discovery incomplete at current hearing, marked "Must Be Tried" (MBT) at present hearing, and Commonwealth "Not Ready" at present hearing. Each point in the scatterplot corresponds to a prosecutor's standardized *z*-score. The black lines represent ordinary least-squares regression. Histograms for the standardized *z*-scores corresponding to each outcome are shown on the diagonal. [Color figure can be viewed at wileyonlinelibrary.com]

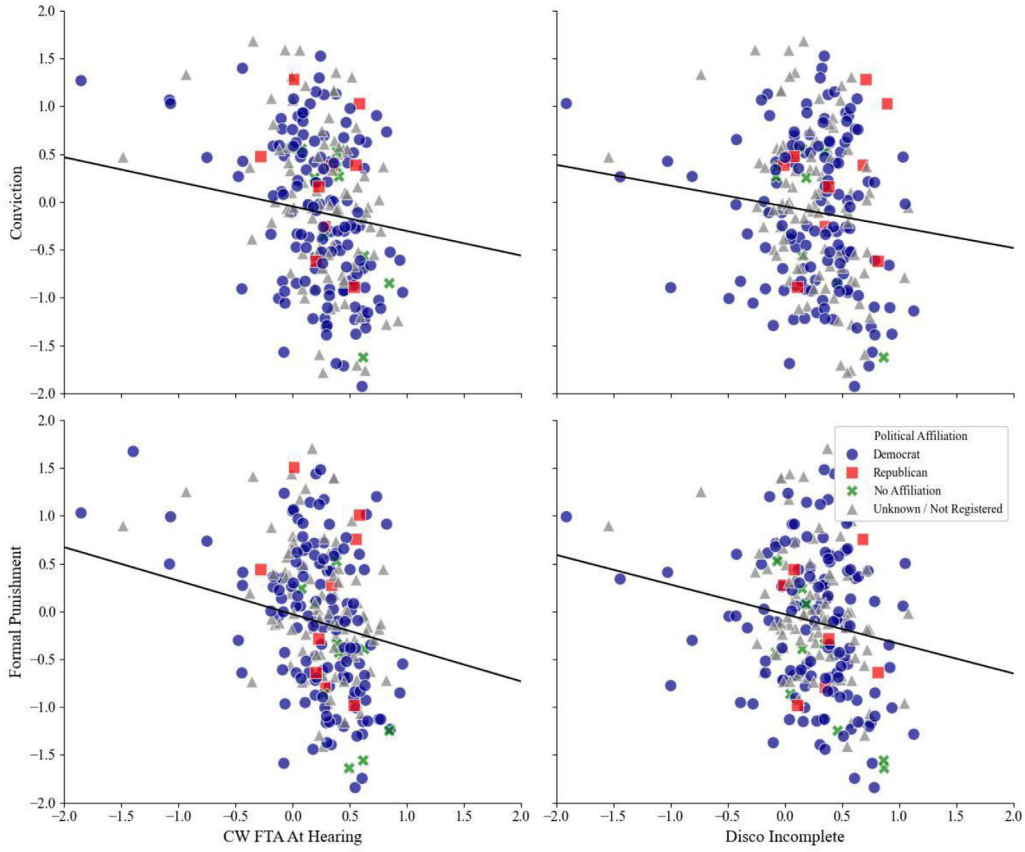

FIGURE A.5 Associations between convictions, punishments, and case management skills by prosecutors' political affiliations. *Note*: These plots show the association between prosecutors' *z*-scores for convictions (top row) and formal punishment (bottom row) and their case management skills (Commonwealth [CW] failure to appear [FTA] and discovery incomplete at current hearing). Each point in the scatterplot corresponds to a prosecutor's standardized *z*-score. The black lines represent ordinary least-squares regression. Prosecutors who are registered Democrats are colored in blue, registered Republicans are colored in red, and registered voters with no party affiliation are colored in green. Prosecutors who were not present in registered Pennsylvania voter rolls are colored in gray. Overall, we can identify political affiliations for 66% of prosecutors, which is approximately equal to the proportion of registered voters in the overall population. [Color figure can be viewed at wileyonlinelibrary.com]

TABLE A.1 Descriptive statistics by sample progression and ADA identification.

Variable	All misdemeanor cases	Excluded Cases	Missing ADA identification	Main sample
Male	0.78	0.74	0.75	0.81
Black	0.56	0.51	0.59	0.60
Hispanic	0.13	0.14	0.12	0.13
Defendant age at arrest	35.83	36.29	36.12	35.52
Violent	0.15	0.14	0.09	0.16
Drugs	0.37	0.43	0.52	0.31
DUI	0.22	0.10	0.13	0.31
Property	0.12	0.14	0.12	0.11
Public order	0.09	0.11	0.12	0.08
DV	0.15	0.19	0.08	0.14
Number of current charges	2.01	1.91	1.76	2.09
Prior arrest in past year	0.31	0.30	0.30	0.33
Prior misdemeanor conviction	0.28	0.27	0.25	0.29
Prior felony conviction	0.19	0.18	0.14	0.21
Prior violent conviction	0.08	0.09	0.05	0.08
Prior incarceration	0.11	0.11	0.08	0.12
Number of prior arrests	2.34	2.25	1.88	2.45
Pending case at open	0.18	0.16	0.17	0.19
Bail posted within 3 days	0.69	0.71	0.71	0.67
Cases	118,121	41,908	8516	67,697
Hearings	507,364	169,335	11,482	197,821
ADA hearings	—	—	11,482	264,407

Note: This table shows descriptive statistics by sample and assistant district attorney (ADA) identification. Column 1 shows descriptive statistics for all misdemeanor lead charge municipal court cases from 2011 to March 2020. Column 2 shows descriptive statistics for cases that were dropped (including cases in specialty courtrooms and cases with incomplete information). Column 3 shows descriptive statistics for hearings in municipal court trial rooms that we expect should have had a prosecutor present, but no prosecutor was identified. Column 4 shows descriptive statistics for our main analysis sample of nondiverted cases in municipal court trial rooms where ADAs are identified and have caseloads of 50 or greater. To connect these numbers with Figure 1, starting with 170,140 misdemeanor arrests, 4% of arrests are declined, 12% are diverted at charging, and 6% get deferred adjudication. This process happens before cases get to municipal court. In addition, multiple arrests can be combined into one "case" or docket number. The unit of analysis for Figure 1 is arrests, whereas the unit of analysis for our main sample of municipal court cases is dockets. Finally, the filtering for Figure 1 is whether the arrest is made between January 1, 2011 and March 1, 2020, whereas the filtering for our main analysis sample is whether the case is opened on/after January 1, 2011 and disposed by March 1, 2020. These processes account for how 170,140 misdemeanor arrests (as shown in Figure 1) become 118,121 misdemeanor municipal court cases (Column 1).

Abbreviations: DUI, driving under the influence; DV, domestic violence.

TABLE A.2 Prosecutor *z*-scores.

	Hearing outcomes		Case outcomes			Defendant outcomes	
ADA ID	CW FTA at hearing	Disco incomplete	Case length	Convicted	Punished	New arrest	New case
321	0.372	0.197	0.452	−1.089	−0.685	−0.051	−0.087
323	0.445	0.231	−0.440	−0.929	−0.675	0.177	0.187
293	−0.154	0.039	0.624	0.003	−0.136	−0.547	−0.738
362	0.255	0.221	−0.617	0.744	0.176	−1.077	−1.316
330	0.321	0.644	−1.031	−0.980	−1.156	−0.719	−1.788

Note: This table shows assistant district attorney (ADA) *z*-scores for a randomly selected sample of prosecutors. For each prosecutor who is benchmarked, we obtain a *z*-score from the regression coefficient and standard error of their estimated treatment effect. Each prosecutor has seven *z*-scores, one for each hearing, case, and defendant outcome. Our outcomes are defined as follows. "CW FTA at Hearing" refers to whether a Commonwealth failure to appear occurs at that prosecutor's hearing. Disco incomplete refers to whether discovery is marked incomplete at that prosecutor's hearing. Case length is the length of the case in days; convicted and punished are indicators for those case dispositions. New arrest and new case are indicators for whether the defendant had a new arrest or new case opened against them within 2 years of the original case's open date.

TABLE A.3 Association between hearing outcomes and case outcomes.

	Dependent variable		
	Case length	Conviction	Formal punishment
Commonwealth not ready	−0.413***	−0.248***	−0.342***
	(0.052)	(0.055)	(0.053)
Num. Obs.	265	265	265
R^2	0.195	0.071	0.139

Note: This table shows the associations between an assistant district attorney's hearing-level actions and their case outcomes. The dependent variables are the prosecutors' standardized *z*-scores for case length, conviction, and formal punishment. The independent variable is the standardized *z*-score for whether a prosecutor is "Not Ready" at their hearing, which includes both Commonwealth failures to appear (FTAs) and whether discovery was marked incomplete. We combine FTAs and incomplete discovery into one measure here because these skills are highly correlated.

Abbreviation: Num. Obs., number of observations.

†$p < 0.1$. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

TABLE A.4 Association between case outcomes and future criminal justice contacts.

	Dependent variable			
	New arrest		New case	
	(1)	(2)	(3)	(4)
Conviction	0.315***		0.263***	
	(0.053)		(0.054)	
Formal punishment		0.382***		0.287***
		(0.052)		(0.055)
Case length days	0.489***	0.454***	0.373***	0.348***
	(0.053)	(0.052)	(0.054)	(0.054)
Num. Obs.	265	265	265	265
R^2	0.331	0.368	0.223	0.233

Note: This table shows the associations between an assistant district attorney's case outcomes and outcomes for the defendants on their cases. The dependent variables are the prosecutors' standardized z-scores for whether a defendant was rearrested within 2 years or had new charges filed within 2 years. The independent variables are the standardized z-scores for a whether the case was ever convicted, whether the case received formal punishment (either probation or incarceration), and the length of the case in days.

Abbreviation: Num. Obs., number of observations.

†$p < 0.1$. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

TABLE A.5 Association between hearing outcomes and case outcomes: Adjusted z-scores.

	Dependent variable		
	Case length	Conviction	Formal punishment
Commonwealth not ready	−0.121***	−0.043**	−0.080***
	(0.024)	(0.015)	(0.017)
Num. Obs.	265	265	265
R^2	0.089	0.029	0.081

Note: This table shows the associations between an assistant district attorney's hearing-level actions and their case outcomes using z-scores that have been adjusted to account for multiple hypothesis testing. To obtain adjusted z-scores, we perform false discovery rate correction using the Benjamini–Yekutieli method and then calculate the z-score from the resulting q-value. The dependent variables in these regressions are the prosecutors' adjusted z-scores for case length, conviction, and formal punishment. The independent variable is the adjusted z-score for whether a prosecutor is "Not Ready" at their hearing, which includes both Commonwealth failures to appear (FTAs) and whether discovery was marked incomplete. We combine FTAs and incomplete discovery into one measure here because these skills are highly correlated.

Abbreviation: Num. Obs., number of observations.

†$p < 0.1$. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

TABLE A.6 Association between case outcomes and future criminal justice contacts: adjusted *z*-scores.

	Dependent variable			
	New arrest		New case	
	(1)	(2)	(3)	(4)
Conviction	0.408***		0.177***	
	(0.053)		(0.029)	
Formal punishment		0.506***		0.216***
		(0.045)		(0.025)
Case length days	0.281***	0.220***	0.124***	0.098***
	(0.033)	(0.031)	(0.018)	(0.017)
Num. Obs.	265	265	265	265
R^2	0.352	0.468	0.264	0.344

Note: This table shows the associations between an assistant district attorney's case outcomes and outcomes for the defendants associated with their cases using *z*-scores that have been adjusted to account for multiple hypothesis testing. To obtain adjusted *z*-scores, we perform false discovery rate correction using the Benjamini–Yekutieli method and then calculate the *z*-score from the resulting *q*-value. The dependent variables in these regressions are the prosecutors' adjusted *z*-scores for whether a defendant was rearrested within 2 years or had new charges filed within 2 years. The independent variables are the adjusted *z*-scores for whether the case was ever convicted, whether the case received formal punishment (either probation or incarceration), and the length of the case in days.

Abbreviation: Num. Obs., number of observations.

†$p < 0.1$. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

TABLE A.7 Robustness: Association between hearing and case outcomes for pre-Krasner sample.

	Dependent variable		
	Case length	**Conviction**	**Formal punishment**
Panel A: Failures to appear			
Main sample			
Commonwealth FTA at hearing	−0.360***	−0.257***	−0.351***
	(0.053)	(0.055)	(0.052)
Pre-Krasner sample			
Commonwealth FTA at hearing	−0.298***	−0.232***	−0.335***
	(0.054)	(0.059)	(0.058)
Panel B: Discovery			
Main sample			
Discovery incomplete	−0.411***	−0.217***	−0.310***
	(0.053)	(0.057)	(0.055)
Pre-Krasner sample			
Discovery incomplete	−0.334***	−0.138*	−0.240***
	(0.055)	(0.062)	(0.062)
Num. Obs. (main)	265	265	265
Num. Obs. (pre-Krasner)	211	211	211

Note: This table shows the associations between an assistant district attorney's hearing-level actions and their case outcomes. The dependent variables are the prosecutors' standardized *z*-scores for case length, conviction, and formal punishment. The independent variables are the standardized *z*-scores for a prosecutor's case management tasks at each hearing. In panel A, the independent variable is Commonwealth failures to appear (FTAs); we present regression results for both the main sample and the pre-Krasner robustness sample. In panel B, the independent variable is whether discovery was marked incomplete; again, we present regression results for both the main sample and the pre-Krasner robustness sample. Substantively, results do not change when looking at the pre-Krasner sample compared with the main sample. Num. Obs., number of observations.

†$p < 0.1$. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

TABLE A.8 Robustness: Association between case outcomes and future criminal justice contacts for pre-Krasner sample.

	Dependent variable			
	New arrest		New case	
	(1)	(2)	(3)	(4)
Main sample				
Conviction	0.315***		0.263***	
	(0.053)		(0.054)	
Formal punishment		0.382***		0.287***
		(0.052)		(0.055)
Case length days	0.489***	0.454***	0.373***	0.348***
	(0.053)	(0.052)	(0.054)	(0.054)
Pre-Krasner sample				
Conviction	0.338***		0.279***	
	(0.061)		(0.062)	
Formal punishment		0.387***		0.286***
		(0.059)		(0.061)
Case length (days)	0.554***	0.519***	0.430***	0.407***
	(0.065)	(0.064)	(0.065)	(0.066)
Num. Obs. (main sample)	265	265	265	265
Num. Obs. (pre-Krasner)	211	211	211	211

Note: This table shows the associations between an assistant district attorney's case outcomes and outcomes for the defendants on their cases. The dependent variables are the prosecutors' standardized z-scores for whether a defendant was rearrested within 2 years or had new charges filed within 2 years. The independent variables are the standardized z-scores for a whether the case was ever convicted, whether the case received formal punishment (either probation or incarceration), and the length of the case in days. We display results for the main sample (top panel) and the pre-Krasner robustness sample (bottom panel); substantively, our results remain the same.

Abbreviation: Num. Obs., number of observations.

†$p < 0.1$. *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

AUTHOR BIOGRAPHIES

Lindsay Graef is a postdoctoral research fellow at the University of Pennsylvania's Carey Law School. Her research focuses on pretrial justice and the courts, with an emphasis on understanding how legal actors and institutions shape the everyday administration of justice.

Aurelie Ouss is the Janice and Julian Bers Assistant Professor of Criminology at the University of Pennsylvania. Her research examines how the design of criminal justice institutions and policies can make law enforcement fairer and more efficient, using empirical approaches from applied economics.