Received: 27 June 2025 | Revised: 24 February 2026 | Accepted: 2 March 2026

DOI: 10.1111/1745-9125.70045

CRIMINOLOGY

ORIGINAL ARTICLE

The racialized packaging of punishment: An instrumental variables approach to incarceration, probation, and monetary sanctions

Ryan Larson[1] | **Robert Stewart**[2] | **Veronica Horowitz**[3] | **Christopher Uggen**[4]

[1]Division of Social Science, Hamline University, St Paul, Minnesota, USA

[2]Department of Criminology & Criminal Justice, University of Maryland, College Park, Maryland, USA

[3]Department of Sociology, State University of New York at Buffalo, Buffalo, New York, USA

[4]Department of Sociology, University of Minnesota, Twin Cities, Minneapolis, Minnesota, USA

Correspondence
Ryan Larson, Division of Social Science, Hamline University, 306-A East Hall, 1556 Hewitt Ave, St Paul, MN 55104, USA.
Email: rlarson21@hamline.edu

Ryan Larson and Robert Stewart share co-lead authorship.

A version of this paper was presented at the 2020 American Sociological Association Annual Meetings.

Funding information
Arnold Ventures, Grant/Award Number: UWSC8574 BPO 27420 AMD 6; Arnold Ventures, Grant/Award Number: #23-09712

Abstract

Research on racial disparities in the criminal legal system generally examines isolated sentencing decisions, rather than the "package" of punishment that defendants experience. Using Minnesota court administrative data from 2004 to 2017, we specify multivariate and instrumental variables models to simultaneously estimate the outcomes of three elements of racialized punishment: incarceration, probation, and monetary sanctions. We instrument incarceration using jail capacity, which accounts for confounding and the simultaneity of incarceration and other punishment forms. Our results show racial patterning in the "mix" of punishment for similarly situated defendants. Before accounting for this mix, Black, Hispanic, and Native American defendants appear to receive less probation and lower monetary sanctions, but longer incarceration than White defendants. After accounting for instrumented incarceration, monetary sanctions and probation are racialized beyond incarceration in complex ways: Black, Hispanic, and Native American defendants receive lower monetary sanctions as compared to White defendants, and probation for Black and Native American defendants is higher

after adjustment. The contours of this racialized package depend critically upon whether the state guidelines recommend a prison sentence. These results show that punishment can be modeled as experienced—as a constitutive package of costs, surveillance, and confinement constrained by structural features of state sentencing guidelines.

KEYWORDS
crime, instrumental variables, punishment, race

1 | INTRODUCTION

Decades of research link race and racism to criminal sanctions. Defendant race (e.g., Light, 2022) and racial demography (e.g., Ulmer & Johnson, 2004) shape diverse sentencing outcomes, including incarceration likelihoods and probabilities (Britt, 2000; Light, 2022), probation lengths (Harrington & Spohn, 2007), sentencing guideline departures (Johnson, 2003), and monetary sanctions (Harris et al., 2011; O'Neill et al., 2022), among others. Yet, while research points to the racialized character of outcomes beyond imprisonment—such as misdemeanor justice (Kohler-Hausmann, 2013) and arrest (Sampson & Neil, 2024)—the vast majority of these studies consider each of these sentencing elements in isolation. From the perspective of those who recommend and impose those criminal sentences and those who are subject to them, punishment is conceived of and *experienced* as an interwoven "package"—a combination of punishment elements and severities that collectively constitute a criminal sentence (Pattillo & Kirk, 2020; Pleggenkuhle et al., 2021).

In this article, we develop a comprehensive account of race and punishment by modeling this package along three distinct axes of racialized social control: incarceration, probation, and monetary sanctions. We conceptualize criminal sentences as representing a matrix of confinement, surveillance, and costs that are applied and received *in tandem*, with each constitutive axis of punishment being endogenous to defendant race. Each mode of punishment interacts with the others in ways that impact individual defendants, often compounding the financial and familial strain of criminal punishment (Comfort, 2008; Harris et al., 2010).

Investigating the packaging of punishment is important for advancing both substantive understanding of penal processes and the methodological and statistical tools used to empirically assess the social drivers of punishment. On the substantive side, packaging has implications for understanding of sentencing decision-making, the experience of punishment for defendants, and broader understanding of the contours of American social control. A core idea of the "packaging of punishment" is that the presence and severity of each punishment element may be related to the presence and intensity of other elements. For example, a judge who gives a lengthy incarceration sentence may reduce (or increase) the fine amount they order. Moreover, qualitative work shows how the experience of one form of punishment directly impacts the response to other forms, as is the case when a person who is sentenced to incarceration has a difficult time paying off fines due to their confinement (Horowitz et al., 2024; Pattillo & Kirk, 2020; Pleggenkuhle et al., 2021). Different punishment components may also be racialized in different ways and are likely to be

correlated with one another in the sentencing process (Slavinski & Pettit, 2022). Understanding and unpacking these packages is especially relevant today, given the expansion of penal power through the use of noncarceral punishments in the "shadow carceral state" (Beckett & Murakawa, 2012; Murakawa & Beckett, 2024), including the dramatic expansion of court-imposed monetary sanctions (also called "legal financial obligations," or LFOs) and the growth of community supervision. Thus, a close examination of how punishment is packaged for individual defendants may reveal a more complete picture of its shape, reach, impact, and social contours.

The conceptual expansion of punishment as a "package" also holds methodological and statistical implications for empirical inquiry into the racialization of punishment. Standard residual race estimates—typically based on regression models that adjust for legal factors such as offense severity, criminal history, and/or presumptive sentence—fail to account for potential interdependencies between the punishment axes and are therefore subject to estimation bias if race is also related to another sentencing form. For example, if race is central in patterning incarceration lengths as previous studies suggest (e.g., Light, 2022) and negatively associated with court-imposed LFOs, it may appear that a racial group is subject to lower LFOs on average—when in fact that decrease is due to the racialized differences in confinement that garner lower LFOs. In any empirical case, racial influences may be over- or underestimated, depending upon the relationship between race and forms of punishment left unspecified. Such analyses not only obscure part of the punishment picture for individuals and communities, but they can also mischaracterize how race operates in the sentencing process. We may risk erroneous conclusions about how a certain punishment axis is racialized if we do not consider that component's interrelationship to other modes of punishment, which are also raced.

We model this punishment process as it is generally constructed and experienced: as a simultaneously composed "package of punishment" that incorporates multiple axes of punitive practice. We develop an instrumental variables (IV) design to explore the interrelationships between the modalities of punishment, and we compare this approach against the "standard" approach typically found in the literature. We draw on an analysis of criminal court administrative data in Minnesota, a state characterized by relatively low rates of incarceration, relatively high racial disparities, and lengthy probation terms (Sentencing Project, 2021). We argue that modeling punishment packages provides a more comprehensive picture of racial disparities in sentencing across punishment modalities, providing further insight into how racial stratification is (re)produced in the sentencing stage of criminal legal system processing.

2 | PUNISHMENT AND RACIALIZED SOCIAL CONTROL

Criminal punishment is strongly racialized in the United States and plays an integral role in the production of racialized inequality (Western, 2006). The steep rise in American jail and prison populations from the mid-1970s, for example, is largely a story of a racialized criminal punishment system, with Black men experiencing especially high rates of imprisonment and felony conviction (Pettit & Western, 2004; Shannon et al., 2017; Western, 2006). Researchers have attributed such disparities to diverse causes, including overt discrimination, color-blind racism, political fear-mongering, criminalization in response to racial threat, rising violence in impoverished communities, structural economic shifts, and various combinations of these mechanisms (Alexander, 2012; Campbell & Schoenfeld, 2013; Forman, 2017; Garland, 2001; Tonry, 2011; Wacquant, 2001). Yet, although the imprisonment rate has declined in recent years, considerable racial disproportionality remains. In 2023, 33% of the state or federal prison population was Black, 23% was

Hispanic, and 31% was non-Hispanic White (Mueller & Kluckow, 2025). Black and Hispanic people are incarcerated at considerably higher rates than non-Hispanic White people (Mueller & Kluckow, 2025). Racial disparities also persist in who has a felony record: 8% of American adults had at least one prior felony conviction in 2010, while 33% of African American men had a felony record (Shannon et al., 2017).

The demography of U.S. imprisonment and its sharp and unparalleled rise are the two components of mass imprisonment (Garland, 2001) and racialized mass incarceration (Bobo & Thompson, 2006). These accounts emphasize the coupling of criminal justice and poverty governance, where rollbacks—or redirections—of the social assistance sector coincided with the immense growth of the criminal legal apparatus (Soss et al., 2011; Wacquant, 2009). Although this scholarship has historically emphasized incarceration, recent research similarly identifies financial sanctions (Page & Soss, 2025) and probation (Phelps & Ruhland, 2022) as defining features of racialized social control in American society (Hinton & Cook, 2021).

Within this social context, racialized punishment disparities "can be located at a variety of decision-making points" (Baumer, 2013, p. 236), including selection into arrest (Chen et al., 2025) and postarrest racial biases in criminal legal processing (National Academies, 2023). Yet, only about 55% of the racial disparity in incarceration rates can be attributed to racial variation at the arrest stage, leaving a substantial portion of the disparity to postarrest factors such as racial variation in sentencing (Baumer, 2013; Beck & Blumstein, 2018; Frase, 2009; Tonry & Melewski, 2008). Criminological research finds robust, albeit time-varying (Light, 2022) racial patterns in criminal sentencing, which vary by the punishment decision under consideration and offense categorization (Kutateladze et al., 2014), the mode of conviction (Johnson, 2003), and broader courtroom, social, and contextual factors (Johnson, 2006; Ulmer & Johnson, 2004).

Several studies have carefully examined how cumulative racial advantage and disadvantage can build over a case's journey through the criminal legal system. For example, Kutateladze and colleagues (2014) consider pretrial detention, plea offers, and incarceration, finding that Black and Latino individuals were more likely to be detained, to receive custodial plea offers, and to receive incarceration sentences—but these defendants were also more likely to receive case dismissals. As the authors explain, "If we had examined only case dismissals, as some prior work has done, we would have mistakenly concluded that [Black defendants] and [Latino/a defendants] were treated more leniently than [White defendants], even though they received more severe outcomes at all other stages of the system" (Kutateladze et al., 2014, p. 538). Other research corroborates these patterns in terms of both disadvantage that accumulates across different stages of case processing (Sutton, 2013; Stolzenberg et al., 2013) and how prior punishment status shapes future punishment decision-making (Hickert et al., 2022). This scholarship illustrates (a) how different aspects of the criminal legal system are racialized in different ways to varying degrees and (b) how expanding the scope of analyses to consider broader punishment forms can yield a more robust understanding of these racial disparities within institutions of punishment.

2.1 | Race and forms of punishment

2.1.1 | Race and incarceration

In many ways, racialized mass incarceration has been a product of changes and enduring inequities in the sentencing stages of punishment. Scholarship on race and imprisonment has primarily examined race as an individual-level extralegal variable (Baumer, 2013). Studies examining

a binary incarceration outcome generally find that Black, Hispanic, and Native American defendants are more likely to be sentenced to incarceration (Britt, 2000; Franklin & Henry, 2020; King & Light, 2019; Spohn, 2017; Spohn & Holleran, 2000; Steffensmeier & Demuth, 2000; Unnever, 1982). Analyses of sentence length and those that combine multiple measures of sentence severity yield less consistent results but often conclude that in certain contexts (Peterson & Hagan, 1984; Ulmer & Johnson, 2004) or among specific subgroups (Demuth & Steffensmeier, 2004; Lizotte, 1978), Black, Indigenous, and People of Color (BIPOC) defendants are punished more harshly than White defendants. Nevertheless, this work shows significant geographic variation, as punishment is also shaped by county-level racial demographic characteristics (Ulmer & Johnson, 2004). Recent research has also noted a shrinking racial gap in federal sentencing, due in part to decreasing Black sentences and increasing White sentences, changes in case characteristics (e.g., criminal history), and shifts in prosecutorial use of mandatory minimums (Light, 2022).

Much of this scholarship explores only a few racial and ethnic groups (White, Black, and Hispanic). Among these categories, Black defendants are treated the most punitively, followed by Hispanic defendants, with White defendants the least punitively on average (Frase, 2009; Unnever, 1982), though more comprehensive sentencing analyses have found that Native Americans are often subject to even greater penalties than other groups (Franklin, 2013; Ulmer & Bradley, 2018). When age and employment are considered, research further indicates that men of color who are younger or with worse employment circumstances face lengthier incarceration sentences (Steffensmeier & Demuth, 2000). These empirical patterns are consistent with predictions within conflict and racial threat theories of punishment, in which marginalized groups deemed "threatening" either economically or culturally are subject to greater punishment (Crawford et al., 1998; Duxbury, 2021; Jacobs et al., 2005). Further, scholarship situates modern racialized incarceration patterns as part and parcel of a broader U.S. historical project of racial social control, with specific emphasis on the policing and confinement of Black bodies (Alexander, 2012; Hattery & Smith, 2021; Wacquant, 2009).

Consistent with this body of research, our first research hypothesis is as follows:

Hypothesis 1. Black, Native American, and Hispanic defendants will receive longer incarceration sentences than White defendants.

2.1.2 | Race and probation

The hyper-surveillance of racial minority groups is hardly a new social phenomenon. From practices in chattel slavery such as branding and runaway slave notices, to modern overpolicing and police violence, surveillance of BIPOC Americans is an ever-present feature of social life that plays a role in both race-making and racial domination (Browne, 2015). Community supervision is today a key facet of racialized surveillance, given the disparate rates of exposure to probation between racial groups (Phelps, 2018). Fewer studies have considered the effects of defendant race on probation supervision as compared to incarceration, but an emerging line of scholarship finds evidence that BIPOC defendants receive harsher or longer probation sentences. Most of this literature has focused on the likelihood of receiving probation versus a custodial sentence. For instance, Freiburger and Hilinski (2013) and Harrington and Spohn (2007) found that Black defendants were more likely to receive a sentence of jail rather than probation. Likewise, Black men are less likely than White individuals and women to receive probation rather than incarceration (Freiburger & Sheeran, 2020). Among those receiving probation and assigned to a low-risk

category by an assessment, Lowder and colleagues (2019) similarly found that White defendants received shorter probation lengths than Black defendants.

Taking a different approach by focusing specifically on the components of probation sentences, Anat Kimchi (2017, 2019) assessed probation conditions as "patterned sentencing packages" in a sample of Minnesota probation sentences in the Twin Cities' Hennepin and Ramsey counties. Compared to White defendants, Kimchi found that Black defendants were likely to receive more numerous probation conditions that were more restrictive, while Hispanic defendants were more likely to face probation conditions that incurred more financial burdens.

Research also indicates that disparities in probation supervision rates are generally smaller than disparities for incarceration rates (Phelps, 2018). Among those whose race is known, the Bureau of Justice Statistics (Kaeble, 2025, p. 7) reports that the 2023 U.S. probation population was approximately 51% non-Hispanic White, 31% non-Hispanic Black, 16% Hispanic, and 2% multiracial or of other races. Accounting for the endogeneity between probation and prison sentences is important if judges are more likely to incarcerate Black and Native American defendants. Racial advantage could then take the form of both a short probation sentence (relative to a long probation sentence) or a long probation sentence (relative to a prison sentence).

Given the extant research on race and probation sentencing, coupled with scholarship documenting racial surveillance as a key facet of modern racialized social control, this leads us to an additional research hypothesis:

Hypothesis 2. Black, Native American, and Hispanic defendants will receive longer probation sentences than White defendants, after accounting for confounding and the endogenous simultaneity of confinement.

2.1.3 | Race and monetary sanctions

A third form of criminal punishment, monetary sanctions or LFOs, affects far more U.S. residents than incarceration or probation (Martin et al., 2018). These monetary sanctions are financial penalties imposed by the state in criminal cases, such as fines, fees, surcharges, and restitution. Scholarship has increasingly documented the important role of these LFOs in racialized resource extraction. From early slave patrols and military policing that enforced chattel slavery and Native dispossession to modern forms of predatory extraction such as monetary sanctions, bail, and forfeitures, the criminal legal apparatus has buttressed racialized, predatory resource extraction (Friedman, 2020; Page & Soss, 2025; Watters et al., 2024). Today, LFOs represent a key mechanism by which racial inequality and dominance is strengthened via wealth depletion, hardship, and other adverse consequences of predatory extraction.

Alexes Harris and colleagues (Harris, 2016; Harris et al., 2010, 2011) systematically investigated the relationship between race/ethnicity, context, and monetary sanctions. Their analysis of Washington state sentencing data, for example, finds that Latino defendants, but not Black defendants, receive higher fines and fees on average as compared to White defendants (Harris et al., 2011). They also find that racialized crime scripts, tied to both moral outrage and cultural stereotypes, impact LFO amounts. For example, Black defendants charged with violent crimes are levied higher fines and fees than similarly situated Black defendants charged with nonviolent crimes. Similarly, Native Americans in Minnesota are subject to the highest monetary sanctions overall, particularly in areas proximal to tribal lands (Stewart et al., 2022). Further, the financial consequences are also racialized: in a panel study of Washington administrative court data, LFOs

increase future levels of poverty, particularly in less White neighborhoods (O'Neill et al., 2022). In short, these findings suggest that monetary sanctions do not impose an equal burden on all racial and ethnic groups.

Although court-ordered LFOs have much in common with other forms of punishment, they are distinct in an important way: the experience of monetary sanctions is explicitly patterned by defendants' "ability to pay." Ability to pay, in turn, is strongly patterned by race (Harris, 2016), and judges can take factors like income into consideration when imposing a financial sentence. Thus, it follows that racial patterns in criminal financial penalties could be endogenous to the large racial differences in wealth, income, employment, and other indicators of financial standing (e.g., Thomas et al., 2020), which are, in part, patterned by racialized criminal histories (Larson et al., 2022; Pager, 2003; Western & Sirois, 2019). This relatively new body of research highlights the salience of race in sentencing decisions *beyond* confinement or community supervision and motivates another research hypothesis:

Hypothesis 3. Black, Native American, and Hispanic defendants will receive larger monetary sanction orders than White defendants, after accounting for confounding and the endogenous simultaneity of confinement.

2.1.4 | Linking incarceration to other forms of punishment

This paper's primary goal is to assess the racial impacts of different axes of punishment *within the sentence*. Although we are not assessing racial impacts across different decision points, as Kutateladze et al. (2014) and others have done, a similar logic nonetheless applies: distinct aspects of the criminal justice system are differentially racialized and likely interrelated. Thus, more complex and nuanced analyses are needed to assess the racial disparities across different modes of punishment. Key to this idea is the interrelationships between axes of sentencing, particularly in terms of the main outcomes of the criminal sentence of incarceration, probation, and monetary sanctions—which we refer to as the "package of punishment." This concept anticipates that judges do not determine how much of each form of punishment—incarceration, probation, and monetary sanctions—to mete out through a series of independent processes (like selecting items a la carte from a restaurant menu); instead, this concept assumes that judges approach this as more or less a single determination process in which the character of each punishment is contingent on the others (akin to selecting side dishes to pair with an entrée).

How these punishment modalities are coconstitutive within punishment packages remains unclear. For example, are there substitution effects between punishment modes? Or is there a positive relationship between modes, representing a "piling on" (Uggen & Stewart, 2014), in which higher amounts of one sanction are accompanied by higher amounts or likelihoods of other sanctions? In an early study of fraud defendants in the federal courts, Waldfogel (1995) finds evidence of substitution, observing a negative association between monetary penalties and the likelihood of prison and the severity of prison sentences. A more recent study of federal White-collar sentencing similarly finds that defendants who pay fines are likely to receive less prison time, but restitution has a positive, albeit diminishing, impact on incarceration length (Schanzenbach & Yager, 2006). In a comprehensive analysis of federal sentencing data, Martin (2012) finds evidence that monetary sanctions function as both a complement and substitute to incarceration. Specifically, Martin finds that receiving a prison sentence is associated with reduced monetary penalties, suggestive of a "trade-off," but that prison lengths are associated with higher monetary loads. Although the

research in this area offers mixed results, the preponderance of the evidence suggests a trade-off between incarceration and other forms of punishment. Thus, our final hypothesis is as follows:

Hypothesis 4. Incarceration will be negatively associated with both probation length and monetary sanction orders.

2.2 | The racialized packaging of punishment

The current study aims to show how race simultaneously influences three distinct but interrelated punishment elements—incarceration, probation, and monetary sanctions—and to interrogate the differences in racialization across these axes of punishment. This study extends research on race and punishment in two important ways. First, we examine state court data to simultaneously model incarceration, probation, and monetary sanction outcomes, showing the similarities and differences in how case and defendant characteristics impact each axis of sentencing. We expect certain case or defendant characteristics to weigh more heavily in certain parts of the punishment process, especially regarding the disproportionate incarceration of people of color (e.g., Wacquant, 2009). In considering each component in isolation, much of the extant research has made the implicit assumption that these outcomes are independent of one another, which may mischaracterize how judges make decisions and distort the broader punishment picture. By relaxing this potentially untenable assumption, we examine how the racialization of punishment varies across modes of punishment, providing a more complete picture of the racial patterns in criminal sentencing within a single U.S. state.

Second, although some existing work has focused on Black–White disparities and narrow racial categorizations (Baumer, 2013), research has increasingly considered disparities in relation to other racial/ethnic groups including Latinx (Demuth & Steffensmeier, 2004; Durante, 2021; Light et al., 2014; Omori & Petersen, 2020; Ulmer & Konefal, 2019; Ulmer et al., 2016), Asian (Franklin & Fearn, 2015; Johnson & Betsinger, 2009; Wu, 2023), and Native American (Franklin, 2013; Stewart et al., 2022) defendants. In concert with these trends, this study uses a more expansive racial and ethnic categorization, with sufficient sample sizes to adequately estimate the punishment patterns for smaller groups, such as Native Americans. This analysis thus moves beyond the traditional White–Black dichotomy, expanding scholarly examination of racialized criminal sentences to racial groups that are less often considered in studies of sentencing disparities.

2.2.1 | Correlations within the punishment matrix

From a statistical perspective, models that examine only a single aspect of punishment may be misspecified if the component under examination influences another aspect of the punishment defendants receive. In statistical parlance, each mode of punishment may be endogenous to the other forms, such that failing to explicitly model such interrelationships would result in model misspecification and omitted variables bias. To the extent that punishment outcomes influence one another and are correlated with key exposure variables of interest, omitting other forms of punishment from the model could bias the estimated effects, potentially leading to erroneous conclusions about the sentencing process.

This mutual influence is important to consider in the estimation of racial impact, as punishment outcomes are likely dependent on the intensities of other punishment forms, which previous

research indicates *are also raced* themselves. Race is further embedded in the structure and operation of sentencing because decisions about LFOs, probation, and prison are guided by criteria that reproduce advantage and disadvantage. For example, racial disparities often widen between arrest, conviction, and incarceration, due in part to the "strongly disparate impact" of sentencing policies that place heavy emphasis on criminal history scores to determine who goes to prison and for how long (Frase, 2009, p. 264).

Thus, a key innovation of our analysis is to account for the racial patterns in the other modes of punishment, to get a more complete accounting of how race shapes criminal sentencing. This approach allows us to examine the extent to which racial patterns in one sentencing component may "bleed into" other sentencing components and how punishment components are raced above and beyond the racial patterns among the other axes. Analyses that do not account for these interdependencies risk mischaracterizing the influence of race in the criminal sentencing process. Consider, for example, a standard regression analysis examining the association between defendant racial identity and LFOs that controls for incarceration, probation, and other case characteristics. Assume the analysis finds that some racial identity is associated with lower levels of LFOs on average, net of individual defendant characteristics, legal factors, and the social context of sentencing. However, if there are significant racial disparities in incarceration, and incarceration is negatively correlated with monetary sanctions, the unadjusted estimates for race in the monetary sanctions model—which does not include an adjustment for incarceration decisions—may be biased downward as race is collinear with incarceration but left unspecified—a form of omitted variable bias. Moreover, it remains to be seen as to whether each form of punishment is racialized above and beyond the disparities induced by the racial distribution of another punishment mode. Are each racialized in similar ways or their own particular ways? Or are disparities in one merely, or in part, the consequence of another mode's racialized character?

Given these substantive and statistical concerns, our study leverages an IV design that provides two key theoretical and analytical advantages. First, the IV approach—when certain assumptions are met—improves the internal validity of our estimates of the interrelationship between incarceration and other primary punishment forms. Further, this allows us to properly estimate the one-way directional effect (i.e., removing any reciprocal causality) in our estimation of the associations between the punishment variables. This design allows for more valid estimation of these interdependencies, giving analytical leverage to test theoretical questions surrounding substitution and augmentation effects.

Second, the IV approach aids in minimizing the bias that could result from key individual extralegal variables (i.e., defendant race) being collinear with incarceration sentences. In studies where causal inference is the primary goal, the focus is on the relationship between a cause and an outcome. In the current study, our overarching research question concerns how race shapes sentencing outcomes, but our fundamental causality concern pertains to the relationship between incarceration and the other forms of punishment. In a broader sense, we treat race not as a simple demographic indicator but as a marker of dynamic social and institutional processes, whereas incarceration represents a central mechanism through which those processes shape sentencing outcomes (Zuberi, 2000). Modeling the package as interrelated and contingent provides a more revealing window into how race operates in the sentencing process across different modes of punitive practice. In our case, estimates of racial impact are potentially biased due to the endogeneity of incarceration (e.g., aspects that determine sentences associated with race could be endogenous with incarceration, thereby inducing a form of posttreatment bias in models that simply adjust for incarceration), which may be related to the other forms of punishment. Our IV approach statistically adjusts for the interrelationships between race and incarceration to better estimate the

racial determinants of monetary sanctions and probation after accounting for extant disparities in incarceration decision-making.

Although we cannot completely disentangle the causal ordering of these sentences with our current data, previous research indicates that incarceration is generally considered the most severe of these penalties by court actors. For instance, in a scale of sentencing severity derived from assessments by judges (Buchner, 1979), imprisonment was considered more severe than probation—even when the duration of probation was far lengthier than the prison term (monetary sanctions were not considered in this scale). Clair (2020) more recently confirmed that courtroom actors such as public defense attorneys also consider confinement paramount, even though some defendants disagree with this assumption.

Previous research thus suggests that incarceration decisions are of primary interest to the involved parties and may represent a key decision point, to which sanctions such as fine amounts and probation lengths are endogenous. Our study thus begins from this empirical reality that incarceration is a primary decision point in sentencing. The linkages between LFOs and probation are far less discussed in the literature, but research shows that many individuals who are sentenced to community supervision are levied fees associated with these services (Harris, 2016). As part of a larger, multi-method study of monetary sanctions, we conducted interviews with attorneys that similarly showed how questions about confinement were prioritized before probation and monetary sanctions. For instance, as one prosecutor described:

> "Well, those [monetary sanctions] are so minimal. I mean, we're usually arguing about the terms and conditions of probation typically, right? That's the most important thing, are they going to prison? If they're not going to prison, what's the term of the probation, right? That's really the meat of it. The other things are just more kind of throwaways. I think they're just part of maybe what a judge does routinely, because the law says they're supposed to."

Likewise, another Minnesota prosecutor explained:

> "…If there's going to be a plea agreement, for example, which is the vast majority of cases. We will talk to defense counsel about things like duration of probation, potential incarceration, conditions of probation and then obviously the offense, that will be the offense of conviction. We almost always say we'll leave the fine up to the court."

As these examples demonstrate, considerations about incarceration are prioritized in the courts in our research site *above* probation and monetary sanctions.[1] While this could justify treating incarceration as exogenous, we treat it as endogenous to account for any simultaneity between confinement and other modes of punishment (i.e., judges weighing multiple axes of punishment simultaneously). By explicitly modeling the contingent nature of the punishment package, we hope to bring the racial impacts on multiple aspects of punishment into sharper focus.

[1] Our courtroom observations indicated these statements by legal actors were reflected during in-court sentencing practices as well.

3 | DATA AND METHODS

To model punishment packages, we draw on the sentencing data and practices in one U.S. state, Minnesota. Minnesota's judiciary is structured as a unified district court system in which all state court matters within a county, including all civil/criminal violations from minor traffic citations to murder, are adjudicated by a single county-level district court. Our analysis focuses on the two most serious offense levels in Minnesota statute: gross misdemeanors (maximum penalty of 1 year in jail and $3000 fine) and felonies (maximum incarceration terms, ranging from 1 year and 1 day to life, and fines prescribed by statute). Judges may stay all or part of a sentence and impose a probation terms for a maximum of 6 years for gross misdemeanors or up to the maximum incarceration term prescribed for the specified felony offense (e.g., as the statutory maximum for controlled substance possession in the first degree is 30 years, the maximum probation term for that offense is also 30 years).[2] Available sentencing options and the character of the process are moderated by the case severity level. For nonfelony cases, judges have considerable discretion within the statutory maximums to determine the specific mix of fines, probation, and incarceration they impose - a feature we explicitly incorporate into our modeling strategy.

For felony cases, sentencing is regulated by the Minnesota Sentencing Guidelines. Minnesota's guidelines are structured as a set of offense grids with two axes, offense severity and defendant criminal history. The cells at the intersections of the two axes specify the presumptive disposition (probation or prison) and the recommended prison sentence length (for presumptive probation, the prison sentence would be stayed).[3] Presumptive prison cells include a recommended sentence length and a permissible range—typically from 15% below to 20% above the recommended length—to allow judges some discretion in choosing an appropriate sentence. Judges may also depart from the guidelines by choosing a different disposition (e.g., probation when the guidelines recommend prison, which occurred in 34% of presumptive prison cases from 2004 to 2017) and/or a sentence length outside the presumptive cell range. When a judge does depart, however, they must justify their departure on the record. Ultimately, sentencing in Minnesota is the purview of the judge: although prosecutors and other court actors provide recommendations regarding sentencing, judges choose the components of the punishment package for each case.

3.1 | Jail capacity and sentencing

Previous research indicates that judges consider multiple legal and extralegal factors as they construct the sentencing package, including available incarceration capacity, which is perhaps most salient for local jail sentences (D'Alessio & Stolzenberg, 1997; Johnson, 2006; Strange et al., 2024; Wang & Mears, 2010). In a recent study, Strange and colleagues (2024) found that judges are particularly attuned to local resources; they are less likely to sentence defendants to jail when local jail facilities are closer to or over capacity and more likely to sentence felony defendants to prison.

[2] These maximum probation lengths were reduced in 2023 to 5 years for most felonies, 2–4 years (depending on the offense) for gross misdemeanors, and 1–2 years (depending on the offense) for misdemeanors. Minnesota Session Laws, ch. 52, art. 6, § 13 (2023) (S.F. 2909).

[3] Approximately two thirds of cases sentenced from 2004 to 2017 had a recommended presumptive probation sentence, and approximately 88% of those sentenced to probation were also sentenced to serve at least some time in jail (based on the authors' analysis of public sentencing data produced by the Minnesota Sentencing Guidelines Commission for cases sentenced from 2004 to 2017).

Jails and prisons throughout Minnesota frequently struggled with overcrowding during the 1990s–2000s. Some counties adapted by using temporary beds in cells or on dayroom floors to expand their capacity—at least in the short term—but even these makeshift efforts were inadequate (Roepke, 2004). Many counties contracted with jurisdictions that did have available capacity, but at a premium daily rate, and nearly half of Minnesota counties began or considered building new or expanding existing facilities during this period (Gunderson, 2005; Roepke, 2004; Stanoch & Wieland, 2018).

Anecdotal evidence from the period suggests that judges and local county officials were keenly aware of local overcrowding issues. For example, during a community forum in 2004, a rural judge reported that the local jail population was more than 40% over its approved capacity (Ruzek, 2004). He explained it "isn't necessarily the best thing for public safety" as his sentencing options were limited, noting that the use of electronic home monitoring—an alternative to short-term jail sentences—was much higher than previous years. Further, in a study of driving-while-intoxicated (DWI) sentences conducted by the Minnesota House of Representatives, judges interviewed throughout the state often identified the degree of jail crowding/overcrowding as an important factor in determining the length of jail time for a defendant convicted of DWI (Cleary, 2000).

Based on this research and contemporary reports, we anticipate that judges will be mindful of overextending local resources. When resources are approaching or over capacity, we expect judges to be less likely to impose lengthy jail sentences and less likely to depart from presumptive prison sentences.

3.2 | Minnesota State Court Administrative Data

Our analysis examines administrative criminal court data from the Minnesota State Court Administrator's Office (SCAO). The SCAO data extract includes granular data points for all criminal cases filed from 2004 to 2015 that had a case resolution as of August 2018 in Minnesota criminal courts. Data points include charge-level and case-level details and events as well as defendant demographics and residential information. To identify unique persons in the data, probabilistic deduplication (Gregg & Eder, 2022) via a partially supervised algorithm with defendant full names and birthdays was used to create unique person identifiers, in order to identify cases within individuals.[4] These data include cases beyond felony-level cases (e.g., Minnesota Sentencing Guidelines Commission, 2015) wherein significant packaging of punishment occurs.

We restrict our analyses to cases that contain at least a gross misdemeanor and/or felony-level convicted charge, which is where different forms of punishment are most likely to coincide and where the most complete defendant information is available. Although petty and misdemeanor cases represent a significant aspect of the criminal legal system, all petty misdemeanors and a substantial proportion of misdemeanor cases are assigned a fine amount based on a fine schedule and payable without an appearance (i.e., defendants can simply pay a fine to close the case). Thus, incarceration and probation sentences are systematically less common in these cases.

Further, as they do not require an in-person appearance, the availability of race information is strongly correlated with the severity level of the conviction; race data are missing for 66.5% of petty misdemeanor cases, 36.5% of misdemeanor cases, 13.2% of gross misdemeanor cases, and

[4] This algorithm calculates record similarity using the affine gap distance on the full-name and birthday string fields and then uses human-classified matched outcomes to estimate the probability of a match given the observed record similarity.

9.2% of felony cases. Although restricting our sample to gross misdemeanor and felony cases substantially reduces missingness on defendant race and sex, we also impute race and gender information within defendants across cases by leveraging information from defendants' other cases to impute race and gender in the focal case. Because these measures are recorded at the case level, within-person racial and gender self-identification may change over time (Liebler et al., 2017). We therefore impute the modal demographic category when the auxiliary cases are not homogenous across cases within defendants. After this auxiliary information imputation strategy, the missingness on race is reduced to 9.6% and 5.8% for gross misdemeanor and felony cases, respectively.

3.2.1 | Endogenous measures

We have three key endogenous variables. First, total monetary sanctions includes all fines and fees (except for restitution) levied against a defendant in a case. We aggregate the total monetary sanctions ordered to the case level and adjust all amounts to January 2018 dollars to adjust for inflation. Second, incarceration length is measured as total sentenced days in prison and/or jail, minus any jail credit or stayed time. This effectively quantifies the amount of time that judges impose from the bench at sentencing as part of the prospective punishment package rather than overall sentenced incarceration length, as the latter would misrepresent the actual time that would be served. Third, probation length is the length of the probation sentence measured in days. As Martin (2012) notes, models that only consider the likelihood rather than the severity of each sanction may mischaracterize the relationships among various sentencing outcomes. We therefore model each key endogenous variable as a continuous amount.

3.2.2 | Exogenous measures

Self-reported modal racial identity of the defendant is our focal variable. We collapsed the SCAO racial categories into six groups: White, Black, Hispanic, Asian/Pacific Islander, Native American, and Other Race (which includes "Other," "Multiracial," "Unavailable," "Unknown," and "Refused"). We include a dichotomous indicator for missing race to adjust our estimates for the presence of missing defendant race data. Our models also include age and a binary gender indicator. Other legal and case characteristics are constructed from SCAO charging information. We include controls for the number of prior convictions, the amount of credit for time served, and an indicator for whether the case went to trial. As a proxy for ability to pay and defendant socioeconomic status, we include an indicator for whether a public defender appeared at any point in the case. In Minnesota, eligibility for public defenders is broadly conceived, and financial need and receipt of any means-tested benefit will typically suffice. Therefore, the public defender indicator represents a proxy for socioeconomic status that captures both extreme to moderate socioeconomic disadvantage.

Cases can contain charges at multiple levels and types, so we treat our charging measures as indicators, which allows each case to have multiple charge types and levels. We include case-level flags for "felony" and "gross misdemeanor" that indicate whether each case included at least one felony or at least one gross misdemeanor charge, respectively, that resulted in an adverse disposition, such as a conviction, a diversion, or some other outcome that was not a dismissal or an acquittal. We construct dichotomous indicators for offense type by flagging whether a

case included a charge in the following categories: violent, drug, and alcohol/driving under the influence (DUI). We include controls for judicial district and sentence year to adjust for between-district heterogeneity in sentencing practices and common year-to-year shifts in sentencing practices across districts. Finally, we include an indicator from the Minnesota Sentencing Guidelines for whether felony cases have a presumptive prison sentence, which serves as a key moderating variable in our IV models below (see Section 3.3 for more details).

Our instrumental variable, the jail operating capacity ratio, comes from the Vera Institute's Incarceration Trends Dataset (Vera Institute of Justice, 2020), which is merged at the county-year level to the SCAO extract. We calculate jail capacity ratios at the county and year level by taking the average daily number of confined persons in the jail divided by each jail's rated capacity for that year. Values above 1 indicate jail overcrowding. Minnesota includes a handful of regional jails, or jails that regularly house defendants from multiple counties. Vera (2020) accounts for this scenario by providing a flag indicating whether a facility is a regional jail and subsequently allocating the operating capacity of these regional jails proportionally across participating county jurisdictions based on population. This instrumental variable is described in further detail in Section 3.3. Descriptive statistics for all variables in the SCAO and Vera data are included in Table 1. Our sample is approximately 61% White, 16% Black, 6% Native American, and 6% Hispanic. After listwise deletion of missing values, this leaves us with 192,155 felony and gross misdemeanor level criminal cases.

3.3 | Analytic strategy

Our analysis proceeds in three steps. First, we describe the "packaging" patterns in punishment and examine the racial distribution of each of our focal endogenous punishment variables (confinement length, probation length, and monetary sanctions). This will show the raw, unadjusted racial differences in sentencing across multiple axes of punishment. We then estimate the effect of race on the package of punishment with with multivariate regression and instrumental variable modes.

3.3.1 | The "standard" approach: Multivariate model of punishment

We refer to the classic multivariate regression-based approach (common in quantitative administrative designs assessing the extra-legal effects of defendant demographic variables) as the standard approach. Under this specification, racial impact is estimated conditional on other demographic, legal, and cases variables, with the remaining unexplained variation between racial categories and punishment attributed to "defendant race" (Baumer, 2013). This approach effectively adjusts for selection into different types of cases (e.g., charge severity, criminal history) and makes statistical comparisons of "like" cases between defendants of different racial groups. Importantly, the race coefficient estimate represents not only the impact of a simple demographic attribute in the sentencing phase but also reflects the broader constellation of constructed race meanings and structures within the criminal legal system, such as social inequalities, relationships to social institutions, and other social interactional processes that play out within social institutions (Sewell, 2016; Zuberi, 2000; see Appendix A for expanded discussion). We use the standard approach to estimate a fixed-effects multivariate linear regression model, treating total fine/fee order, incarceration days, and probation days as outcome variables. This

TABLE 1 Descriptive statistics for variables in the State Court Administrator's Office, Minnesota Sentencing Guidelines, and Vera jail data.

Statistic	*N*	Mean	SD	Min	Median	Max
Incarceration days	192,155	198.950	835.039	0	10	62,050
Total fine/fee order	192,155	433.390	917.983	0.000	250.160	318,082.700
Probation days	192,155	761.887	882.966	0	730	7300
White	192,155	0.609	0.488	0	1	1
Black	192,155	0.162	0.368	0	0	1
Hispanic	192,155	0.057	0.233	0	0	1
Asian	192,155	0.021	0.143	0	0	1
Native American	192,155	0.059	0.236	0	0	1
Other race	192,155	0.013	0.112	0	0	1
Missing race	192,155	0.079	0.269	0	0	1
Male	192,155	0.785	0.411	0	1	1
Age	192,155	33.455	11.247	15	31	100
Trial	192,155	0.0004	0.019	0	0	1
Priors	192,155	2.977	4.015	0	2	84
Public defender	192,155	0.595	0.491	0	1	1
Percent credit	192,155	14.251	30.374	0.000	0.274	100.000
Percent stayed	192,155	43.680	45.792	0.000	0.000	100.000
Felony	192,155	0.350	0.477	0	0	1
Gross misdemeanor	192,155	0.618	0.486	0	1	1
Violent	192,155	0.197	0.397	0	0	1
Drug	192,155	0.131	0.337	0	0	1
Alcohol/DUI	192,155	0.314	0.464	0	0	1
Presumptive prison	192,155	0.077	0.266	0	0	1
Judicial district 1	192,155	0.142	0.349	0	0	1
Judicial district 2	192,155	0.081	0.273	0	0	1
Judicial district 3	192,155	0.093	0.290	0	0	1
Judicial district 4	192,155	0.217	0.413	0	0	1
Judicial district 5	192,155	0.065	0.247	0	0	1
Judicial district 6	192,155	0.061	0.239	0	0	1
Judicial district 7	192,155	0.099	0.298	0	0	1
Judicial district 8	192,155	0.032	0.175	0	0	1
Judicial district 9	192,155	0.083	0.276	0	0	1
Judicial district 10	192,155	0.128	0.334	0	0	1
Year 2004	192,155	0.039	0.193	0	0	1
Year 2005	192,155	0.082	0.274	0	0	1
Year 2006	192,155	0.127	0.333	0	0	1
Year 2007	192,155	0.155	0.362	0	0	1
Year 2008	192,155	0.052	0.222	0	0	1
Year 2009	192,155	0.008	0.090	0	0	1
Year 2010	192,155	0.002	0.048	0	0	1

(Continues)

TABLE 1 (Continued)

Statistic	*N*	Mean	SD	Min	Median	Max
Year 2011	192,155	0.029	0.169	0	0	1
Year 2012	192,155	0.033	0.178	0	0	1
Year 2013	192,155	0.121	0.326	0	0	1
Year 2014	192,155	0.172	0.378	0	0	1
Year 2015	192,155	0.103	0.304	0	0	1
Year 2016	192,155	0.069	0.254	0	0	1
Year 2017	192,155	0.008	0.090	0	0	1
County-level capacity ratio	192,155	0.793	0.221	0.200	0.811	2.267

method simultaneously estimates the effects on multiple dependent variables using the same set of covariates.

Our standard approach model is parameterized as follows:

$$Y_{ik} = \alpha_k + \sum \pi_{jk} Race_{ij} + \sum \beta_{jk} X_{ij} + \theta_{dk} + \lambda_{tk} + \epsilon_{ik},$$

where Y_{ik} is the kth punishment response for the ith observation, and π_{jk} represent the focal race estimates, with White defendants as the referent category. β_{jk} is the jth control's regression slope for the kth response. θ_{dk} and λ_{tk} represent fixed effects for judicial districts and sentence years, respectively, for each outcome equation. The inclusion of these terms restricts our comparisons to those occurring *within* sentencing districts and sentence years, as sentencing patterns may vary between judicial districts and across time (Stewart, 2019). These fixed effects net out time-constant unobserved heterogeneity between districts (e.g., localized court culture and sentencing practices between districts) and time variation in sentencing patterns shared across districts (e.g., state-wide policy changes). X_{ij} represents the jth case- or defendant-level control characteristics for the ith observation. Because the punishment variables are either zero or strictly positive and are positively skewed, they are expressed using the $\log(X + 1)$ transformation when included in the multivariate analyses presented here, so the coefficients below are interpreted in their exponentiated form, representing percentage changes in punishment in relation to unit changes in the independent variables. We will then compare outcomes produced using this "standard" approach for modeling racial impact to our IV approach. Each approach models the relationship between defendant race and punishment, but the standard approach treats the three primary sentence components as conditionally exogenous once controls are added, whereas our proposed IV approach treats incarceration as endogenous and uses jail capacity to isolate exogenous variation.

We provide a more robust discussion of our model justifications and specifications in Appendix A. In brief, existing research using the standard approach typically examines a single sanction or treats each punishment axis as a separate outcome rather than modeling them together. From a statistical perspective, models that examine only a single aspect of punishment may be misspecified if that component influences another aspect of the punishment defendants receive. To the extent that punishment outcomes influence one another and are also correlated with race, omitting other forms of punishment from the model could bias the estimated impact of race, resulting in a risk of ill-informed conclusions about the role of defendant race in sentencing.

3.3.2 | IV model of punishment

The approach we propose to gain purchase on the endogeneity of punishment forms involves modeling the effect of the incarceration component on both the monetary sanction and probation components, adjusting for multiple forms of bias through an instrumental variable design (see Appendix A for details). We instrument incarceration lengths with time-varying county-level available jail capacity ratios to provide an estimate of the causal effect of incarceration on decisions regarding the other axes of punishment.

IV estimation allows for consistent estimation of causal effects in the presence of endogeneity by statistically isolating the variation in the endogenous predictor that is uncorrelated with the error term. In other words, the design partitions the variation in incarceration into plausibly exogenous variation induced by the instrument, thereby allowing for causal inference. We use the county-level time-varying change in jail capacity ratios as an instrument for confinement length. This model specifies that changes in jail populations relative to capacity will be associated with changes in the length of incarceration sentences through judge awareness of local overcrowding or the availability of jail resources; in contrast, the capacity ratio will have no direct impact on monetary sanctions or probation apart from its indirect path through confinement length, net of other observed variables, thereby identifying the effect of incarceration on another axis of the punishment decision. As we expect heterogeneity in the effect of the capacity ratio on incarceration (i.e., available local jail capacity likely affects the judicial decision-making process in different ways for gross misdemeanor and felony cases), we formulate our first stage to include interaction terms between the county-level jail capacity ratio and felony and gross misdemeanor indicators, which allows the effect of capacity ratio on incarceration rates to vary across sentencing-level contexts.

We estimate the IV models using two-stage least squares (2SLS), which uses a system of equations with a first and second stage to estimate the causal impact of incarceration on the other axes of punishment. The "first stage" models, each of which regresses incarceration on the instrument, the county-level jail capacity ratio, and other covariates are specified as follows:

$$I_i = \alpha + \phi_1 CR_d + \phi_{1F}\left(CR_d \times F_{ij}\right) + \phi_{1GM}\left(CR_d \times GM_{ij}\right) + \sum_j \beta_j X_{ij} + \theta_d + \lambda_t + \epsilon_i,$$

where I_i represents the sentenced incarceration amount, CR_d represents the county-level jail capacity ratio, and ϕ_1 is the effect of the capacity ratio on incarceration amounts, net of the observed controls in the model. Akin to the multivariate model above, we include fixed effects for judicial district (θ_d) and sentence year (λ_t) to adjust for spatiotemporal heterogeneity in sentencing patterns.

The reduced form ("second stage") models are specified as follows:

$$Y_{ik} = \alpha + \phi_{2k} I_i + \phi_{2Fk}\left(I_i \times F_{ij}\right) + \phi_{2GMk}\left(I_i \times GM_{ij}\right) + \sum_j \pi_{jk} Race_{ij} + \sum_j \beta_{jk} X_{ij} + \theta_{dk} + \lambda_{tk} + \epsilon_{ik},$$

where Y_{ik}, LFOs or probation amounts, are specified as a function of the instrument (CR_d), the county-level jail capacity ratio, the observed case and defendant characteristics ($\beta_{jk} X_{ij}$), as well as the fixed effects for judicial district (θ_{dk}) and sentence year (λ_{tk}). If certain statistical assumptions

are met,[5] the local average treatment effect (LATE) of the impact of incarceration on each of the other punishment outcomes (P_{ik}) can be obtained via the ratio of the reduced form coefficients over the first-stage coefficients:

$$\delta_k\left(F_{ij}, GM_{ij}\right) = \frac{\phi_{2k} + \phi_{2Fk}F_{ij} + \phi_{2GMk}GM_{ij}}{\phi_1 + \phi_{1F}F_{ij} + \phi_{1GM}GM_{ij}},$$

where δ_k is the estimated causal effect of incarceration on each of the alternative punishment outcomes. This design alleviates race effect identification concerns with posttreatment and collider bias by isolating variation in incarceration determined by the instrument and the observed covariates. The race effect, after adjusting for instrumented incarceration, can be interpreted as the direct effect of race on other punishment outcomes independent of exogenously determined incarceration differences and ot, effectively removing the indirect pathway from race to probation and monetary sanctions that operate through incarceration. All models are estimated using the "ivreg" function in the R package "AER" (Kleiber & Zeileis, 2008).

Because Minnesota has a robust sentencing guidelines system at the felony level, we also estimate specifications with interaction terms between the endogenous punishment and focal race coefficients with the presumptive prison indicator. This allows us to examine potential heterogeneity due to this key structural feature and to examine the robustness of packaging patterns for different defendants. Importantly, this can also illuminate different mechanisms of racial advantage and disadvantage. For example, if a racial group is associated with lower probation at the non-presumptive level, this may reflect advantage, whereas a similar association at the presumptive level would imply racial disadvantage, as lower probation amounts in that context mechanically reflect lower likelihoods of downward departures from the sentencing guidelines.

4 | RESULTS

4.1 | Descriptive analyses

Figure 1 displays the mean punishment amounts for each axis by race. These means are unadjusted for offense mix and severity (among other variables) and represent the pure bivariate relationship between race/ethnicity and sentenced amounts of each form of punishment. The left-most panel of Figure 1 shows the mean incarceration days sentenced for each defendant racial group. Black defendants are sentenced to the most incarceration on average, with a mean of 348 days, followed by Hispanic (249 days) and Native American defendants (237 days). Conversely, White defendants are sentenced, on average, to lower incarceration lengths, with a mean of about 154 days.

A somewhat different pattern emerges for monetary sanctions, with White and Hispanic defendants sentenced to relatively higher LFO amounts per case. White defendants are sentenced to $490 of fines and fees on average, and Hispanic defendants on average face $437 in fines and fees. Native American defendants trail slightly behind Hispanic defendants at $371 and Black defendants are assessed comparatively less with mean fine and fee amounts at $212. Asian and Other Race defendants are ordered $344 and $359 per case, respectively. To summarize, Black and Native

[5] We provide a more robust discussion and empirical tests of IV model assumptions of instrument relevance, the exclusion restriction, and monotonicity in Appendix A.

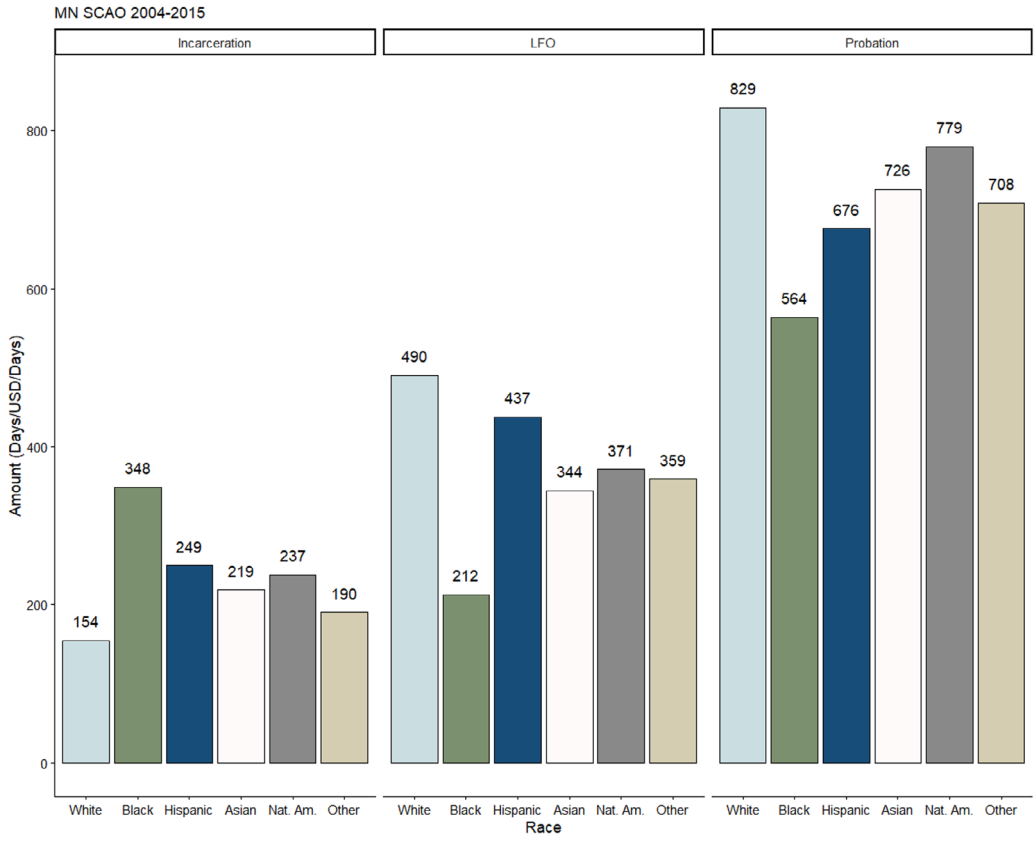

FIGURE 1 Punishment amounts by defendant race (MN SCAO 2004–2015). [Color figure can be viewed at wileyonlinelibrary.com]

American defendants receive comparatively more incarceration and lower LFOs, whereas White defendants receive more LFOs and less incarceration on average.

Figure 1 also displays the race-specific means of our third dependent variable, probation length. White and Native American defendants have the highest average probation sentences at 829 and 779 days, respectively. In comparison, Hispanic (676 days), Asian (726 days), and Other Race (708 days) defendants have lower average probation lengths, with Black defendants having the lowest mean probation length (564 days). In sum, the overall racialized pattern in sentenced punishment reveals a bifurcation along the lines of race, with White defendants receiving longer community surveillance and larger monetary punishments, and BIPOC defendants receiving longer confinement sentences. Overall, the unadjusted disparities provide preliminary evidence for Hypothesis 1 but evidence *contrary* to H2 and H3, which predicted higher LFO and probation amounts for Black, Native American, and Hispanic defendants than for White defendants.

4.2 | Multivariate model of the package of punishment

Table 2 displays results from a multivariate linear regression of incarceration length, total LFO (fine and fee) orders, and probation length—which we refer to as the standard approach to

TABLE 2 Multivariate model of punishment, Minnesota 2004-2017.

	Punishment outcome		
	log(Incarceration)	log(LFO)	log(Probation)
	Coef(SE)	Coef(SE)	Coef(SE)
Black	0.124*** (0.014)	−0.631*** (0.014)	−0.209*** (0.018)
Hispanic	0.158*** (0.020)	−0.239*** (0.020)	−0.339*** (0.026)
Asian	−0.079* (0.033)	−0.072* (0.032)	0.082* (0.041)
Native American	0.327*** (0.021)	−0.616*** (0.020)	−0.230*** (0.026)
Other race	−0.103* (0.041)	−0.230*** (0.041)	−0.012 (0.052)
Missing race	−0.314*** (0.018)	−0.058*** (0.017)	0.042 (0.022)
Male	0.700*** (0.011)	−0.011 (0.011)	−0.591*** (0.014)
log(Age)	0.339*** (0.015)	−0.061*** (0.014)	−0.447*** (0.018)
Prior convictions	0.069*** (0.001)	−0.049*** (0.001)	−0.048*** (0.002)
Public defender	0.384*** (0.010)	−0.756*** (0.010)	−0.012 (0.013)
Percent credit	−0.030*** (0.0002)	0.002*** (0.0002)	0.010*** (0.0002)
Percent stayed	−0.019*** (0.0001)	0.010*** (0.0001)	0.019*** (0.0002)
Trial	2.166*** (0.237)	0.491* (0.234)	−0.928** (0.298)
Felony	0.129*** (0.016)	0.260*** (0.016)	1.155*** (0.020)
Gross misdemeanor	−1.054*** (0.016)	0.553*** (0.016)	0.962*** (0.020)
Violent	0.390*** (0.014)	−0.003 (0.014)	0.281*** (0.017)
Drug	−0.212*** (0.016)	0.268*** (0.016)	0.903*** (0.020)
Alcohol/DUI	0.956*** (0.012)	0.541*** (0.012)	1.487*** (0.015)
Constant	1.792*** (0.059)	5.792*** (0.058)	2.605*** (0.074)
District FE	Yes	Yes	Yes
Sentence year FE	Yes	Yes	Yes
Observations	192,155	192,155	192,155
R^2	0.331	0.248	0.376
Adjusted R^2	0.331	0.248	0.376
F-statistic (df = 40; 192,114)	2374.771***	1588.104***	2893.814***

Note: All tests are two-tailed.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

estimating how race influences punishment amounts. As discussed in the methodology section, each dependent variable is log transformed, so we exponentiate each race coefficient, which then represents the multiplicative factor for each mode of punishment as compared to White defendants.

The first column in Table 2 shows the estimates in the model predicting incarceration days. Relative to White defendants, Black, Native American, and Hispanic defendants receive 13% ((exp(0.124)−1) × 100 = 13.2), 39%, and 17% longer incarceration sentences, respectively, net of the other covariates, which provides support for H1. In contrast, the coefficients for Asian and Other Race defendants are statistically significant and negative, indicating that, on average, Asian and Other Race incarceration lengths tend to be shorter as compared to White defendants, net of other factors.

Estimates in the second column predict the amount of LFOs, logged and standardized to January 2018 dollars. In contrast to Model 1, Black, Hispanic, Asian, Native American, and Other Race defendants receive *lower* monetary sanctions, net of other factors, as compared to White defendants, consistent with the bivariate results in Figure 1. Specifically, Black defendants are associated with about 47% ($(\exp(-0.631) - 1) \times 100$) lower LFO orders as compared to White individuals, and Native Americans, Asian, Other Race, and Hispanic defendants receive 46%, 7%, 21%, and 21% lower LFOs, respectively. This provides evidence contrary to H2, as Black, Hispanic, and Native American defendants have *lower* fine and fee amounts. This pattern of results largely parallels the descriptive patterns in Figure 1, albeit the disparities are attenuated when adjusted for legal and other factors. Although Black, Hispanic, and Native American defendants receive lower fines and fees on average, they are sentenced to longer incarceration sentences, net of other observed measures. These results suggest that the "punishment mix" for similarly situated cases is patterned in a bifurcated manner by defendant race.

The third column displays the multivariate estimates of probation length. The racial patterns in probation are generally opposite in direction to those observed for incarceration length. All non-White defendants receive significantly lower probation lengths relative to White defendants, net of other factors, apart from Asian defendants who receive statistically larger amounts of probation on average. In particular, Black, Hispanic, and Native American defendants are associated with 19%, 29%, and 21% shorter probation sentences, respectively. This model provides no support for H3, as Black, Hispanic, and Native American defendants have lower probation loads for similarly situated cases.

In sum, the model in Table 2 suggests that Black, Native American, and Other Race defendants receive less punitive sentences in terms of financial punishment and community supervision, but more punitive sentences in terms of incarceration, as compared to White defendants. Overall, these results suggest that more carceral forms of punishment, such as incarceration, are levied more heavily against BIPOC defendants, whereas White defendants have the highest probation loads overall.

4.3 | IV models of the package of punishment

It is clear from the results in Table 2 that race is associated with incarceration lengths, LFO amounts, and probation sentences. Further, these results show clear variations in racialization at the sentencing stage, at least prior to adjustment for intercorrelations within the punishment package matrix. If incarceration causes LFOs or probation, however, then this relationship, left unspecified, will bias the residual race estimates in the models and potentially lead to erroneous conclusions about race and sentencing. To investigate this possibility regarding the confounding character of incarceration, we next present the IV estimates of monetary sanction amounts and probation days adjusted for instrumented confinement days.

The first two models in Table 3 introduce our IV estimates of LFO amounts ordered and probation length, and Figure 2 summarizes the race coefficients across the axes of punishment and compares the estimates between the standard approach and our IV design in a coefficient plot for the overall models. According to the IV estimates, the effect of incarceration days on monetary sanctions is negative, suggestive of a substitution effect, with a 1% increase in incarceration days resulting in about a 2.7% decrease in monetary sanctions. This is indicative of a "trade-off" or "substitution" relationship between the two forms of punishment, wherein sentencing

TABLE 3 IV 2SLS models of punishment, Minnesota 2004–2017.

	Punishment outcome			
	log(LFO)	log(Probation)	log(LFO)	log(Probation)
	Coef(SE)	Coef(SE)	Coef(SE)	Coef(SE)
log(Incarceration)	−0.267*** (0.051)	−2.259*** (0.108)	−0.144*** (0.036)	−2.008*** (0.069)
Black	−0.598*** (0.016)	0.071* (0.033)	−0.745*** (0.015)	−0.106*** (0.029)
Hispanic	−0.197*** (0.022)	0.018 (0.047)	−0.250*** (0.021)	−0.129** (0.041)
Asian	−0.093** (0.033)	−0.096 (0.072)	−0.111** (0.034)	−0.095 (0.066)
Native American	−0.528*** (0.027)	0.509*** (0.057)	−0.621*** (0.023)	0.438*** (0.045)
Other race	−0.258*** (0.042)	−0.244** (0.090)	−0.286*** (0.043)	−0.262** (0.083)
Missing race	−0.141*** (0.024)	−0.668*** (0.051)	−0.108*** (0.022)	−0.624*** (0.042)
Male	0.176*** (0.037)	0.990*** (0.080)	0.069** (0.022)	0.695*** (0.043)
log(Age)	0.030 (0.023)	0.320*** (0.049)	−0.036* (0.015)	0.084** (0.030)
Prior convictions	−0.030*** (0.004)	0.109*** (0.008)	−0.041*** (0.002)	0.076*** (0.004)
Public defender	−0.653*** (0.022)	0.855*** (0.047)	−0.698*** (0.016)	0.727*** (0.031)
Percent credit	−0.006*** (0.002)	−0.056*** (0.003)	−0.002 (0.001)	−0.043*** (0.002)
Percent stayed	0.005*** (0.001)	−0.024*** (0.002)	0.008*** (0.001)	−0.015*** (0.001)
Trial	1.068*** (0.264)	3.966*** (0.564)	0.455 (0.235)	2.049*** (0.459)
Felony	0.294*** (0.018)	1.445*** (0.038)	0.200*** (0.025)	0.948*** (0.049)
Gross misdemeanor	0.272*** (0.056)	−1.419*** (0.119)	0.411*** (0.039)	−1.138*** (0.076)
Violent	0.101*** (0.024)	1.162*** (0.052)	0.013 (0.015)	0.868*** (0.029)
Drug	0.212*** (0.019)	0.424*** (0.042)	0.244*** (0.017)	0.489*** (0.033)
Alcohol/DUI	0.796*** (0.050)	3.646*** (0.107)	0.641*** (0.030)	3.253*** (0.058)
Presumptive prison (PP)			−0.904*** (0.083)	−3.307*** (0.161)
log(Incarceration) × PP			0.148*** (0.029)	0.893*** (0.057)
Black × PP			0.965*** (0.040)	0.183* (0.078)
Hispanic × PP			0.337*** (0.072)	0.254 (0.141)
Asian × PP			0.426*** (0.118)	0.022 (0.231)
Native American × PP			0.542*** (0.069)	−0.829*** (0.135)
Other race × PP			0.591*** (0.152)	0.448 (0.297)
Constant	6.270*** (0.108)	6.653*** (0.232)	6.165*** (0.111)	6.831*** (0.216)
District FE	Yes	Yes	Yes	Yes
Sentence year FE	Yes	Yes	Yes	Yes
IV *F*(Incar.)	132.89***	132.89***	2304.596***	2304.596***
IV Wu–Hausman	20.92***	955.64***	7.40***	1959.61***
Observations	192,155	192,155	192,155	192,155

Note: All tests are two-tailed. IV = county-level jail capacity ratio interacted with felony and gross misdemeanor indicators.
Abbreviations: 2SLS, two-stage least squares; IV, instrumental variables.
*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

packages with higher incarceration lengths are compensated with lower monetary sanction costs. This coefficient provides strong evidence for H4.

Importantly, the introduction of instrumented incarceration days to the model does not alter the substantive interpretation of the race estimates for LFOs as compared to the multivariate model in Table 2. Black, Hispanic, and Native American defendants, net of incarceration and other fac-

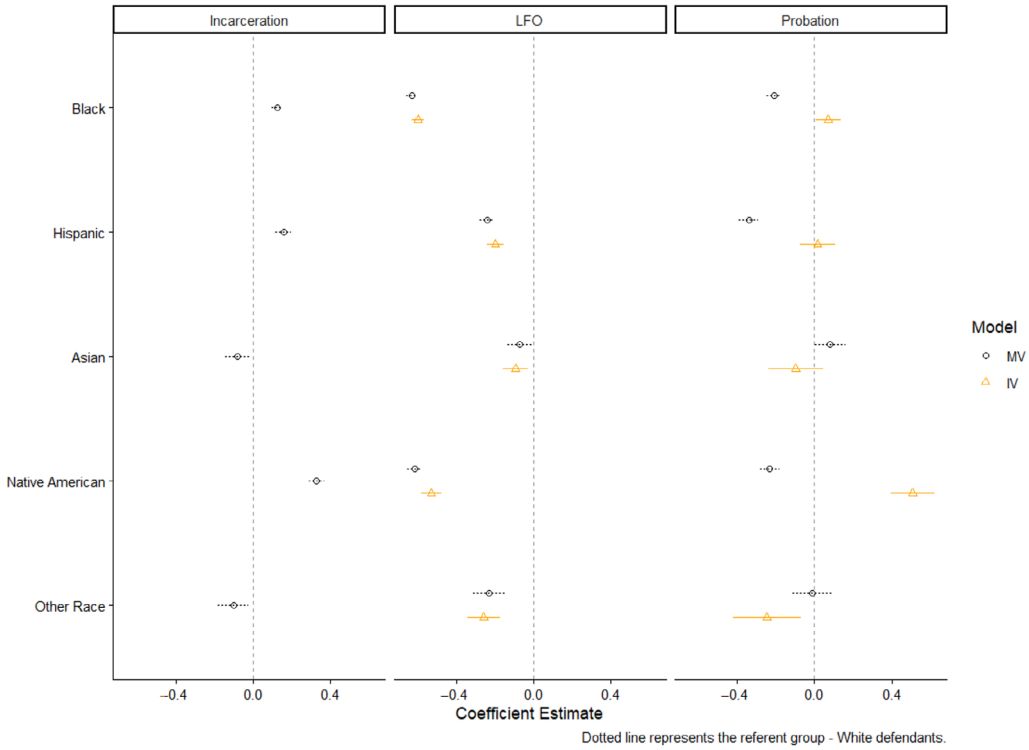

FIGURE 2 Coefficient plots for MV and IV punishment models. Dotted line represents the referent group—White defendants. MV, multivariate; IV, instrumental variable. [Color figure can be viewed at wileyonlinelibrary.com]

tors, still receive lower LFOs than White defendants, suggesting that LFOs are racialized above and beyond extant racial disparities in incarceration, with BIPOC defendants being sentenced to lower LFOs on average. Specifically, Black defendants are associated with 45% lower LFOs on average as compared to White defendants, net of incarceration, and Hispanic (−18%) and Native American (−41%) defendants also receive lower LFOs than White defendants. These findings suggest that the racial disparities in financial punishment in Table 2 are not a simple function of racial differences in incarceration length. However, the coefficients in Table 3 are modestly attenuated as compared to those in Table 2, which suggests that part of the lower LFO amounts for non-White defendants can be attributed to the accompanied longer prison sentences. Taken as a whole, this characterizes average punishment packages for BIPOC Minnesotans as consisting of more confinement and less financial punishment net of ordered confinement lengths.

The second column in Table 3 presents an IV specification for probation days, showing that, like the LFO model, incarceration has a statistically significant and negative influence on probation amounts, with a 1% increase in confinement length associated with a 2.3% decrease in probation days, the patterns of which also map onto H4. This further illustrates the fundamental packaging relationships across the full sample: more incarceration is associated with lower monetary and probation penalties, revealing multiple substitution relationships that characterize the packaging of punishment in Minnesota.

Black and Native American defendants receive higher probation sentences as compared to White defendants net of instrumented incarceration, and the direction of these coefficients runs

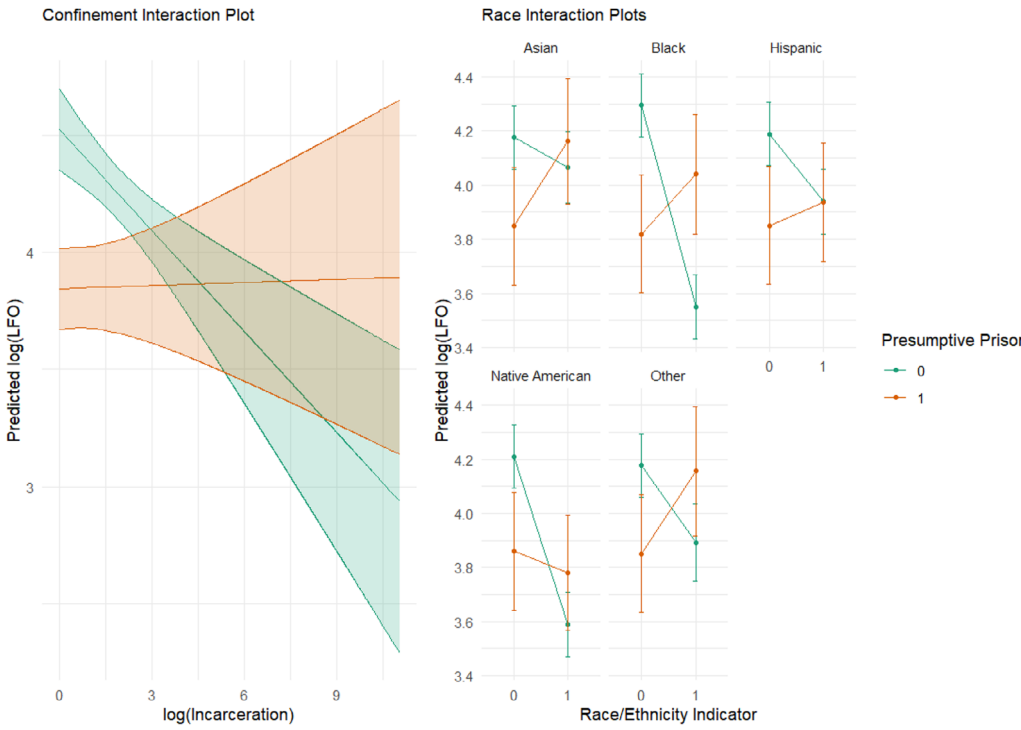

FIGURE 3 LFO interaction plots of incarceration and focal race effects by presumptive prison. [Color figure can be viewed at wileyonlinelibrary.com]

counter to those estimated in the multivariate approach depicted in Table 2. Specifically, Black and Native American defendants are associated with 7.4% and 66% longer probation lengths, respectively, whereas the Hispanic coefficient is also positive but not statistically significant. The effects for Black, Hispanic, and Native American defendants are substantially different from those in Table 2, suggesting that the negative bivariate relationships are driven by the substitution of incarceration and probation. When this packaging component is directly modeled, however, the coefficients indicate that probation, like LFOs, is racialized in its own right, above and beyond incarceration length disparities, but with Black and Native American defendants receiving a more onerous component of probation for similarly situated cases with comparable incarceration stays. In particular, this finding underscores the theoretical and empirical importance of accounting for the packaging mix of punishment in understanding race and sentencing. These overall models assume that packaging works similarly across different levels of sentencing.

However, if both the packaging interchangeabilities and racialization of the punishment axes are heterogeneous across sentencing contexts (i.e., subject to the sentencing guidelines or not), this stability assumption is untenable. Therefore, we estimate IV models that interact instrumented confinement and the focal race coefficients with the presumptive prison indicator in Models 3 and 4 of Table 3, with the interaction effects for LFOs and probation visualized in Figures 3 and 4 by predicted marginal effects, respectively. This approach allows the coefficients to vary across presumptive and non-presumptive cases and allows for an examination of potential effect heterogeneity in packaging and racialization across this crucial "tipping point" in Minnesota's sentencing grids at the felony level.

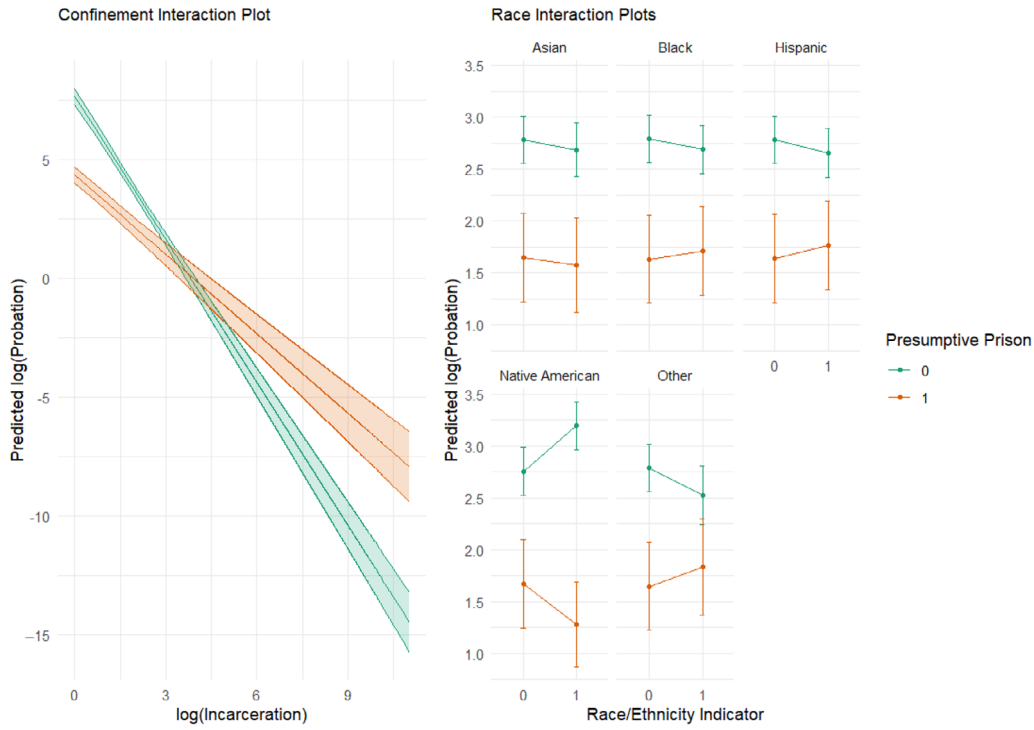

FIGURE 4 Probation interaction plots of incarceration and focal race effects by presumptive prison. [Color figure can be viewed at wileyonlinelibrary.com]

Among non-presumptive prison cases (third model in Table 3, visualized in Figure 3), the packaging effect of incarceration is estimated to be −0.14, indicative of a 15% substitution effect at the margin of incarceration. The interaction term is statistically significant and negative, indicating that the packaging effects among presumptive prison sentences are substantially constrained at higher levels of sentencing. Figure 3 displays this moderated relationship visually, with a significantly suppressed effect of incarceration lengths on LFOs for presumptive prison cases, indicating a severing of the relationship in this context (Wald F-test of effect of incarceration in presumptive cases: $F(192,106) = 0.071$, $p = 0.72$). The focal race coefficients also exhibit heterogeneity across sentencing contexts. All the main race estimates are statistically significant and negative, with positive, statistically significant interaction terms. These interaction effects are visually depicted in the second panel of Figure 3. At lower-level sentencing (i.e., non-presumptive cases), Black defendants are associated with significantly higher LFOs net of incarceration, whereas in presumptive prison cases, Black defendants receive significantly lower LFOs as compared to White defendants. For Hispanic, Native American, and Other Race defendants, the racialization of LFO amounts is restricted to non-presumptive cases, with each group receiving lower ordered LFO amounts as compared to White defendants net of incarceration. In sum, both the packaging exchangeability and racialization of LFOs are conditioned by the prescriptiveness of cases, with more pronounced packaging and racial impact in non-presumptive prison cases.

IV specifications of probation lengths (Model 4 in Table 3 and visualized in Figure 4) also exhibit similar heterogeneity in effect across presumptive and non-presumptive prison cases. The main effect of instrumented incarceration is statistically significant and negative for non-

presumptive prison cases, and the interaction term indicates a suppression of this effect among presumptive prison cases with a statistically significant and positive interaction term. In contrast to the LFO IV interaction model, the packaging relationship here is diminished as opposed to severed, with the combined effect at the presumptive level remaining statistically significant and negative (Wald F-test of effect of incarceration in presumptive cases: $F(192,106) = 2158.5$, $p < 0.001$), albeit significantly reduced in magnitude. The main race estimates parallel those in Table 2, with BIPOC defendants generally receiving significantly shorter probation lengths than White defendants at lower levels of sentencing (though Native Americans are an exception to this pattern, receiving longer probation sentences).

The focal race interaction terms are not statistically significant except for the Black and Native American terms. For Black defendants, they indicate less probation in non-presumptive cases, but significantly more probation at the presumptive level, suggesting that the higher probation burden identified in the overall model (Model 2 in Table 3) is driven by presumptive prison cases. However, a Wald linear hypothesis test indicates that the presumptive effect is not statistically different from zero (Wald F-test of effect of Black in presumptive cases: $F(192,106) = 1.05$, $p = 0.31$). The packaging effect heterogeneity likely explains the higher probation amounts in the overall model and that Black defendants experience slightly elevated or similar probation lengths to White defendants net of this packaging. In contrast, the Native American interaction effect is statistically significant and strongly negative as compared to the significant positive main effect. This shows that packaging for Native American defendants is dependent upon sentencing context, with a statistically significant increase in probation at the non-presumptive level (Wald F-test of effect of Native American in non-presumptive cases: $F(192,106) = 94.0$, $p < 1.001$) and a statistically significant decrease at the presumptive level (Wald F-test of effect of Native American in presumptive cases: $F(192,106) = 9.57$, $p = 0.002$). This highlights the bifurcated racialization of probation, net of the interchange of incarceration and probation, for Native American defendants in Minnesota.

Returning to our conceptual framework, we find evidence that the packaging of punishment is racialized in complex ways constrained by the prescriptiveness of the sentencing context (see Figure 2 for a comparison of the standard multivariate estimates with the IV estimates). Compared to White defendants, Black defendants receive punishment packages composed of comparatively longer incarceration terms, comparable probation terms, and lower LFOs for cases of similar incarceration length. However, Black defendants receive relatively less onerous probation lengths for lower-level cases. The punishment package for Hispanic defendants also includes comparatively longer incarceration terms and comparable probation terms, but smaller LFOs compared to White defendants, driven by lower levels of sentencing. Finally, for Native American defendants in Minnesota in our study period, the package of punishment is perhaps the most severe, consisting of much longer terms of incarceration and significantly longer probation terms at the non-presumptive level (with significantly shorter lengths at the presumptive level), albeit with lower LFOs for lower-level cases. The standard multivariate approach obscures much of this complex racialization net of incarceration and would suggest that the opposite is true for probation: that non-White defendants would be ordered to lower probation amounts on average overall. However, the IV estimates tell a similar racial story with most BIPOC groups receiving lower LFOs—a pattern we return to contextualize in the conclusion below. Further, our interaction models suggest that cases that do not carry a presumptive prison sentence have more *fluidity* between the punishment elements, indicating that decision makers are given more space to package and substitute punishment forms. This makes racialized packaging a more active social and punitive process at less-prescribed sentencing levels with widened discretion in this context. In sum, these

findings illustrate that the packaging of punishment a defendant receives is contingent in complex and bifurcating ways depending on the race of the defendant and the prescriptiveness of the sentencing context of the case under consideration.

5 | CONCLUSION

Incarceration has been a principal focus of criminological research on punishment, but it is only one of punishment's elemental axes. Punishment and sentencing scholars have devoted far less attention to the broader punishment *package*. In this paper, we have examined how the racialization of punishment depends upon the particular facet of the punishment matrix being analyzed. Specifically, we find that the "mix" or package of punishment that defendants receive is patterned by race and structured by institutional constraints. Further, both monetary sanctions and probation are *punitive terrains amenable to racialization in their own right* and are not mere downstream consequences of incarceration decisions. Our findings on incarceration are largely consistent with previous work in other social contexts that finds Black, Native American, and Hispanic defendants receive longer sentences. Standard estimates imply that Black, Hispanic, and Native American defendants generally receive lower fine and fee amounts and probation sentences descriptively.

Explicit modeling of the interrelationships between incarceration, probation, and monetary sanctions reveals a "trade-off" of punishment forms in which longer incarceration spells are offset by lower LFO and probation amounts overall. This is consistent with previous scholarship examining the combinatory character of these punishment forms (Martin, 2012; Waldfogel, 1995) and suggests that incarceration is a focal judicial decision point around which the other axes of punishment revolve. Further, the packaging action between punishment forms is most robust at lower levels of sentencing (i.e., non-presumptive prison cases), facilitated by the greater degree of judicial discretion in contexts in which judges are afforded more freedom to substitute one punishment form for one another.

However, when we statistically account for incarceration in the sentencing of both probation and monetary sanctions using an IV design, the racialization of noncustodial forms of punishment comes into sharper focus. Non-White defendants generally receive a reduced LFO burden on average as compared to White defendants, which is substantively consistent with standard regression approaches. However, the coefficient estimates in the standard approach are generally larger than the IV specifications, which suggest that these estimates could *overstate* racial advantage, as race covaries sharply with incarceration lengths. Further, this overall effect masks significant heterogeneity by the prescriptiveness of the sentencing context. The lower overall LFOs of BIPOC defendants are driven almost exclusively by less prescriptive sentencing contexts (i.e., non-presumptive prison cases), but all BIPOC groups receive more comparable LFOs to White defendants at the presumptive level.

This may be a reflection of judges taking defendants' ability to pay into consideration when making these assessments, particularly at lower levels of sentencing. As is true in many states, non-White residents are economically disadvantaged in Minnesota (Minnesota Department of Employment & Economic Development, 2024). The comparable monetary sanctions exhibited at the presumptive level could be a function of both limitations placed on judges as well as more expressive, symbolic displays of punitive practice. Under Minnesota's sentencing guidelines, judges enjoy broad discretion with respect to punitive fines, with broad statutory ranges. Judges at the presumptive prison level may be less likely to reduce monetary punishment to the extent they are able at lower levels of sentencing.

Alternatively (or in combination), presumptive prison cases often include incarceration sentences of substantial length, and LFOs may be a symbolic axis by which judges can express moral outrage in more serious cases by imposing monetary burdens incommensurate with defendant ability to pay in a show of cultural punitive condemnation (Garland, 1993; Harris et al., 2011). Thus, the more comparable LFOs for BIPOC defendants may be a terrain amenable to shows of symbolic-expressive punitiveness in the "degradation ceremonies" (Garfinkel, 1956) of the courtroom in more severe cases. BIPOC defendants may be further disadvantaged in such decisions if such considerations overtake or displace financial considerations regarding ability to pay.

The monetary sanction models demonstrate that the racialization of financial punishment is primarily concentrated in contexts of greater judicial discretion, and are racialized above and beyond extant disparities in incarceration at lower levels of sentencing. Importantly, while BIPOC defendants are associated with lower overall LFO *orders*, this does not necessarily imply lesser LFO *debt burden*. When extant levels of racial variation in defendant ability to pay are considered, the racial estimates in the model likely *understate* the differential impact monetary sanctions have on defendant experiences (e.g., debt loads) across racial groups. Our estimates here do not suggest an equal burden of monetary sanctions; in light of extant racial differences in ability to pay, the seeming "lightness" of LFO orders for BIPOC defendants could mask a racially disparate debt burden.

Our estimates also suggest that racial disparities in incarceration can have downstream consequences for disparities in probation. After adjusting for the exchanging interrelationship between incarceration and probation loads, the racialization of probation exhibits a different pattern than estimates from standard regression approaches would imply. For Hispanic and Asian defendants, the interaction models reveal overall comparable probation lengths to White defendants after adjustment for incarceration loads, whereas Black defendants exhibit a racial disadvantage with modestly higher probation lengths. This marks probation as an area of racial disadvantage for Black defendants, particularly among more severe cases. This suggests that the estimates from standard regression approaches also *overstate* the racial impacts of probation overall, as racial disparities in incarceration "bleed into" disparities in probation and account for the lower probation amounts for these groups descriptively. At the non-presumptive level, Hispanic and Asian defendants receive lower probation sentences, which may suggest advantage along this axis. However, Minnesota is a state with high average probation sentences (Stewart, 2019), and probation itself carries significant material and psychosocial burdens (Phelps & Ruhland, 2022), elevated risks of prison entry (Harding et al., 2017), as well as elevated burdens placed on family members (Comfort, 2008). Thus, while the shorter probation sentences experienced by BIPOC defendants at lower levels may be some form of reprieve, Minnesota's probation system still places a significant burden on defendants exposed to any level of probation.

Importantly, Native American defendants' pattern of racialization exhibits differences as compared to the other race/ethnicity groups: similar to Black defendants, Indigenous defendants receive *more* probation in the package on top of heftier incarceration sentences overall, resulting in a significantly heavy package of confinement and surveillance. Further, the negative estimate for Native Americans on probation holds only for the presumptive prison level; in those cases, any probation term greater than zero indicates a downward departure. Put another way, the negative association between Indigeneity and probation length signifies disadvantage rather than advantage, as these results indicate that Native American defendants are much less likely to receive downward departures than other groups. Indeed, not only is probation racialized in complex ways above and beyond the existing disparities in incarceration, but even the shorter probation sentences for Native American defendants with presumptive prison cases are a mechanism of

disadvantage. We interpret these patterns as reflecting a broader settler colonial project in which Indigenous Minnesotans are subject to stiffer punitive measures (Watters et al., 2024)—that is, American Indians account for 8.7% of the prison population—almost eight times their proportion in the state's population—and steeper LFOs and LFO debt, particularly in areas that overlap with Indian Country (Stewart et al., 2022).

Although the reasons for the remaining disparities in probation net of racial differences in incarceration are beyond the scope of our analysis, we can speculate about the reasons such disparities may exist. First, our analysis and the SCAO data do not account for the *qualities* of probation, only the quantity, and previous research suggests that non-White defendants in Minnesota receive more restrictive terms of probation as compared to White defendants (Kimchi, 2019), which also may help explain the shorter terms for BIPOC defendants. Second, previous research has indicated that Black and Indigenous defendants have higher rates of probation revocation (Wilson, 2023), which judges may consider when deliberating in the sentencing phase. Finally, probation often includes some degree of support services, such as drug treatment, mental health counseling, and job search assistance. If judges and/or courtroom cultures perceive non-White defendants as less deserving or that the support at lower levels of sentencing will be less efficacious for BIPOC groups, they may reduce probation lengths commensurate with their extant racial assumptions.

In total, our analysis provides further evidence of the racialized character of modern punishment in America. We extend this scholarship by identifying racialized packaging of punishment as a complex sociolegal process that is conditioned by institutional sentencing constraints. Specifically, the results here are consistent with much research on race and prison sentences (e.g., King & Light, 2019) but extend as well to monetary sanctions and probation, even net of incarceration sentences. Overall, the racialized packages identified in this analysis are more punitive with regard to incarceration for BIPOC defendants and are accompanied by lower LFO amounts and comparable or higher probation amounts. Further, these packages are contingent upon sentencing context and reveal a more complicated picture of racial advantage and disadvantage. This is further evidence that the criminal legal system continues to play a prominent role as a racialized social control institution (e.g., Wacquant, 2009), while placing a significant debt burden on disadvantaged communities of color (e.g., O'Neill et al., 2022). However, these racial impacts are not solely the purview of Blackness, as we document consistent racial effects for Hispanic and Native American defendants in incarceration, probation, and monetary sanctions. In sum, these results show how the "packaging" of punishment is racialized in complex and sometimes counterintuitive ways.

Additionally, the analysis here indicates that not only has crime become racially typified in American culture, but that punishment, at least in the Minnesota courts, also has a racially typified character. Even in an era of sentencing guidelines and more standardized sentences, race continues to color the shape, modality, and severity of a punishment package a defendant encounters. Our analysis also suggests that sentencing guidelines at the felony level play an important role in suppressing the packaging and racialization of punishment. Both the packaging exchangeability and racial impact are more muted at the presumptive prison level, and race estimates tend to be driven by sentencing among less severe cases on the whole. Nevertheless, continued racialization at lower levels has notable implications beyond sentencing, including both considerable collateral consequences and the reproduction of American racial stratification (Wakefield & Uggen, 2010). Further, these processes can be conceptualized within a broader framework in which the criminal legal system plays a role in the social construction of race (e.g., Ray, 2019) wherein racial ideologies, schemas, and assumptions become codified into formal and informal courtroom patterns in punishment packaging decisions at the sentencing phase.

Our analysis here is not meant to replace existing models of the racial effects on punishment, but rather to extend them and highlight the complexity of modeling these sociolegal processes. We have shown the interrelationships between varied modes of punishment and pointed toward an inclusive approach to capture a fuller *package* of punishment, while interrogating how this package is constructed along the social fault lines of race and ethnicity. Our analysis suggests that studies of race and punishment would benefit from considering multiple forms of punishment in tandem. For example, studies of monetary sanctions may be incomplete without considering incarceration. Similarly, studies of probation may overestimate or underestimate racial differences if they do not also incorporate incarceration as part of the model. We believe our analyses are more consistent with how punishment is experienced by defendants, who often leave court with a large and unwieldy package of monetary and nonmonetary sanctions.

Further, our analysis highlights not only the packaging of punishment by race but also how understanding of racial impacts can be misleading if the analyses do not account for other forms of punishment levied at sentencing. Controlling for the incarceration effect alters the magnitude, *and the direction*, of some of the race indicators. Therefore, considering one facet of punishment in isolation risks misidentifying how it fits into the overall package of punishment, yielding a biased or distorted picture of how race and other factors shape decision-making in relation to that punishment axis. In short, our analysis suggests that researchers should consider multiple facets of the punishment nexus to get a more complete, and arguably more nuanced, portrayal of punishment.

Our study is not without limitations. We have only considered the confinement, probation, and monetary components of punishment, rather than other court-ordered requirements such as conditions of release and probation-like mental health treatment or drug testing. While some of these facets are charge-type specific, they could be collinear with key variables and the components of punishment and therefore confound estimates of racial impact. Future research should expand the punishment outcomes to gain a more complete and resolute picture of the punishment package among different racial groups. Further, our models foreground the impact of incarceration on the other punishment forms but cannot neatly unpack the relationship between probation and LFOs. While incarceration is indeed a focal decision point as illustrated by the results here, future inquiry could identify IVs that more properly estimate the relationships between the other forms of punishment (e.g., punishment-specific sentencing policy changes, probation caseloads). IV estimators are classically low statistical power estimators, and therefore, our standard errors are slightly higher as compared to the ordinary least squares (OLS) multivariate estimator. This brings greater uncertainty as to our estimates of racial impact and represents a trade-off between identification and precision. Moreover, the SCAO data are limited at the defendant level, missing measures such as income and educational status. These measures represent further axes of stratification at the sentencing stage and are highly correlated with race. To the extent that our public defender variable picks up variation in defendant socioeconomic status, our focal race estimates are adjusted for this variation. However, unobserved heterogeneity by income, wealth, or other socioeconomic resources within the public defender-eligible pool could confound the focal race coefficient estimates. Minnesota also may be an outlier in terms of the use of LFOs and other forms of punishment, as other research shows Minnesota gives comparatively less fines and fees than many other states (Harris, 2016). The relative scarcity of research analyzing multiple components of punishment in tandem is likely due in part to the empirical modeling challenges discussed above, as well as access to complete and robust individual-level sentencing data.

Nonetheless, our analysis provides a foundation for future work to consider the nexus of race and the myriad forms of punishment, and these findings illustrate the benefits of modeling

punishment as it is experienced: as a *constitutive package* of different forms and intensities of punishment.

ACKNOWLEDGMENTS

This research was supported by a grant to the University of Washington from Arnold Ventures (Alexes Harris, PI) and a grant to the University of Minnesota from Arnold Ventures (Christopher Uggen and Kevin Reitz, Co-PIs). The authors thank Hannah Schwendeman, MayMay Seymour, and Caleigh Lueder for their feedback and assistance.

FUNDING INFORMATION

This work was supported by Arnold Ventures (Grant number: UWSC8574 BPO 27420 AMD 6) and the University of Minnesota through a Arnold Ventures grant (#23-09712; Christopher Uggen and Kevin Reitz, co-PIs), titled "Research Support for Minnesota Sentencing Guidelines Commission: Comprehensive Review of Minnesota's Sentencing System."

DATA AVAILABILITY STATEMENT

An online replication package at the Harvard Dataverse (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/PXQUB9&version=DRAFT) includes code files to replicate all analyses presented in the manuscript. Base SCAO data used in these analyses are not publicly available from the Minnesota Court Administrator's Office but can be requested at https://mncourts.gov/help-topics/court-statistics/contact (some SCAO data are publicly available via subscription, and some fields are not public [e.g., race] and therefore require an MN Supreme Court approval for use). MN Jail capacity data to construct the instrument are included in this repository.

REFERENCES

Alexander, M. (2012). *The new Jim Crow: Mass incarceration in the age of colorblindness*. The New Press.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444–455.

Baumer, E. P. (2013). Reassessing and redirecting research on race and sentencing. *Justice Quarterly*, *30*(2), 231–261.

Beck, A. J., & Blumstein, A. (2018). Racial disproportionality in US state prisons: Accounting for the effects of racial and ethnic differences in criminal involvement, arrests, sentencing, and time served. *Journal of Quantitative Criminology*, *34*, 853–883.

Beckett, K., & Murakawa, N. (2012). Mapping the shadow carceral state: Toward an institutionally capacious approach to punishment. *Theoretical Criminology*, *16*(2), 221–244.

Bobo, L. D., & Thompson, V. (2006). Unfair by design: The war on drugs, race, and the legitimacy of the criminal justice system. *Social Research: An International Quarterly*, *73*(2), 445–472.

Britt, C. L. (2000). Social context and racial disparities in punishment decisions. *Justice Quarterly*, *17*(4), 707–732.

Browne, S. (2015). *Dark matters: On the surveillance of blackness*. Duke University Press.

Buchner, D. (1979). Scale of sentence severity. *The Journal of Criminal Law and Criminology*, *70*(2), 182–187.

Campbell, M. C., & Schoenfeld, H. (2013). The transformation of America's penal order: A historicized political sociology of punishment. *American Journal of Sociology*, *118*(5), 1375–1423.

Chen, M. K., Christensen, K. L., John, E., Owens, E., & Zhuo, Y. (2025). Smartphone data reveal neighborhood-level racial disparities in police presence. *Review of Economics and Statistics*, *107*(6), 1734–1742.

Clair, M. (2020). *Privilege and punishment: How race and class matter in criminal court*. Princeton University Press.

Cleary, J. (2000). *Jail time for DWI offenders in Minnesota: A brief survey of practices*. Minnesota House of Representatives, House Research Department.

Comfort, M. (2008). *Doing time together: Love and family in the shadow of the prison*. University of Chicago Press.

Crawford, C., Chiricos, T., & Kleck, G. (1998). Race, racial threat, and sentencing of habitual offenders. *Criminology*, *36*(3), 481–512.

D'Alessio, S. J., & Stolzenberg, L. (1997). The effect of available capacity on jail incarceration: An empirical test of Parkinson's law. *Journal of Criminal Justice*, *25*(4), 279–288.

Demuth, S., & Steffensmeier, D. (2004). Ethnicity effects on sentence outcomes in large urban courts: Comparisons among White, Black, and Hispanic defendants. *Social Science Quarterly*, *85*(4), 994–1011.

Durante, K. A. (2021). County-level context and sentence lengths for Black, Latinx, and White individuals sentenced to prison: A multi-level assessment. *Criminal Justice Policy Review*, *32*(9), 915–937.

Duxbury, S. W. (2021). Who controls criminal law? Racial threat and the adoption of state sentencing law, 1975 to 2012. *American Sociological Review*, *86*(1), 123–153.

Forman, J. (2017). *Locking up our own: Crime and punishment in Black America*. Farrar, Straus and Giroux.

Franklin, T. W. (2013). Sentencing Native Americans in US federal courts: An examination of disparity. *Justice Quarterly*, *30*(2), 310–339.

Franklin, T. W., & Fearn, N. E. (2015). Sentencing Asian offenders in state courts: The influence of a prevalent stereotype. *Crime & Delinquency*, *61*(1), 96–120.

Franklin, T. W., & Henry, T. K. S. (2020). Racial disparities in federal sentencing outcomes: Clarifying the role of criminal history. *Crime & Delinquency*, *66*(1), 3–32.

Frase, R. S. (2009). What explains persistent racial disproportionality in Minnesota's prison and jail populations? *Crime and Justice*, *38*(1), 201–280.

Freiburger, T. L., & Hilinski, C. M. (2013). An examination of the interactions of race and gender on sentencing decisions using a trichotomous dependent variable. *Crime & Delinquency*, *59*(1), 59–86.

Freiburger, T. L., & Sheeran, A. M. (2020). The joint effects of race, ethnicity, gender, and age on the incarceration and sentence length decisions. *Race and Justice*, *10*(2), 203–222.

Friedman, B. (2020). Carceral immobility and financial capture: A framework for the consequences of racial capitalism penology and monetary sanctions. *UCLA Criminal Justice Law Review*, *4*(1), 177–184.

Garfinkel, H. (1956). Conditions of successful degradation ceremonies. *American Journal of sociology*, *61*(5), 420–424.

Garland, D. (1993). *Punishment and modern society: A study in social theory*. University of Chicago Press.

Garland, D. (2001). *The culture of control: Crime and social order in contemporary society*. Oxford university press.

Gregg, F., & Eder, D. (2022). Dedupe. https://github.com/dedupeio/dedupe

Gunderson, D. (2005, March 7). Minnesota jails bursting at the seams. *Minnesota Public Radio*. https://news.minnesota.publicradio.org/features/2005/03/07_gundersond_crowdedjails/

Harding, D. J., Morenoff, J. D., Nguyen, A. P., & Bushway, S. D. (2017). Short-and long-term effects of imprisonment on future felony convictions and prison admissions. *Proceedings of the National Academy of Sciences*, *114*(42), 11103–11108.

Harrington, M. P., & Spohn, C. (2007). Defining sentence type: Further evidence against use of the total incarceration variable. *Journal of Research in Crime and Delinquency*, *44*(1), 36–63.

Harris, A. (2016). *A pound of flesh: Monetary sanctions as punishment for the poor*. Russell Sage Foundation.

Harris, A., Evans, H., & Beckett, K. (2010). Drawing blood from stones: Legal debt and social inequality in the contemporary United States. *American Journal of Sociology*, *115*(6), 1753–1799.

Harris, A., Evans, H., & Beckett, K. (2011). Courtesy stigma and monetary sanctions: Toward a socio-cultural theory of punishment. *American Sociological Review*, *76*(2), 234–264.

Hattery, A. J., & Smith, E. (2021). *Policing Black bodies: How Black lives are surveilled and how to work for change*. Rowman & Littlefield Publishers.

Hickert, A., Bushway, S. D., Harding, D. J., & Morenoff, J. D. (2022). Prior punishments and cumulative disadvantage: How supervision status impacts prison sentences. *Criminology*, *60*(1), 27–59.

Hinton, E., & Cook, D. (2021). The mass criminalization of Black Americans: A historical overview. *Annual Review of Criminology*, *4*(1), 261–286.

Horowitz, V. L., Larson, R. P., Stewart, R., & Uggen, C. (2024). Fines, fees, and families: Monetary sanctions as stigmatized intergenerational exchange. *The Sociological Quarterly*, *65*(4), 469–488.

Jacobs, D., Carmichael, J. T., & Kent, S. L. (2005). Vigilantism, current racial threat, and death sentences. *American Sociological Review*, *70*(4), 656–677.

Johnson, B. D. (2003). Racial and ethnic disparities in sentencing departures across modes of conviction. *Criminology*, *41*(2), 449–490.
Johnson, B. D. (2006). The multilevel context of criminal sentencing: Integrating judge-and county-level influences. *Criminology*, *44*(2), 259–298.
Johnson, B. D., & Betsinger, S. (2009). Punishing the "model minority": Asian-American criminal sentencing outcomes in federal district courts. *Criminology*, *47*(4), 1045–1090.
Kaeble, D. (2025). Probation and parole in the United States, 2023 (NCJ 310118 Bulletin). U.S. Department of Justice, Bureau of Justice Statistics. https://bjs.ojp.gov/document/ppus23.pdf
Kimchi, A. (2017). Investigating the assignment of probation conditions: Heterogeneity and the role of race and ethnicity (Order No. 10621542). ProQuest Dissertations & Theses Global. https://www.proquest.com/dissertations-theses/investigating-assignment-probation-conditions/docview/2013199231/se-2
Kimchi, A. (2019). Investigating the assignment of probation conditions: Heterogeneity and the role of race and ethnicity. *Journal of Quantitative Criminology*, *35*, 715–745.
King, R. D., & Light, M. T. (2019). Have racial and ethnic disparities in sentencing declined? *Crime and Justice*, *48*(1), 365–437.
Kleiber, C., & Zeileis, A. (2008). *Applied econometrics with R*. Springer Verlag.
Kohler-Hausmann, I. (2013). Misdemeanor justice: Control without conviction. *American Journal of Sociology*, *119*(2), 351–393.
Kutateladze, B. L., Andiloro, N. R., Johnson, B. D., & Spohn, C. (2014). Cumulative disadvantage: Examining racial and ethnic disparity in prosecution and sentencing. *Criminology*, *52*(3), 514–551. https://doi.org/10.1111/1745-9125.12047
Larson, R., Shannon, S., Sojourner, A., & Uggen, C. (2022). Felon history and change in US employment rates. *Social Science Research*, *103*, 102649.
Liebler, C. A., Porter, S. R., Fernandez, L. E., Noon, J. M., & Ennis, S. R. (2017). America's churning races: Race and ethnicity response changes between census 2000 and the 2010 census. *Demography*, *54*(1), 259–284.
Light, M. T. (2022). The declining significance of race in criminal sentencing: Evidence from US federal courts. *Social Forces*, *100*(3), 1110–1141.
Light, M. T., Massoglia, M., & King, R. D. (2014). Citizenship and punishment: The salience of national membership in US criminal courts. *American Sociological Review*, *79*(5), 825–847.
Lizotte, A. J. (1978). Extra-legal factors in Chicago's criminal courts: Testing the conflict model of criminal justice. *Social Problems*, *25*(5), 564–580.
Lowder, E. M., Morrison, M. M., Kroner, D. G., & Desmarais, S. L. (2019). Racial bias and LSI- R assessments in probation sentencing and outcomes. *Criminal Justice and Behavior*, *46*(2), 210–233.
Martin, K. D. (2012). Monetary sanctions in federal criminal sentencing: Significance, prison, and policy (Doctoral dissertation). UC Berkeley.
Martin, K. D., Sykes, B. L., Shannon, S., Edwards, F., & Harris, A. (2018). Monetary sanctions: Legal financial obligations in US systems of justice. *Annual Review of Criminology*, *1*(1), 471–495.
Minnesota Department of Employment and Economic Development. (2024). Minnesota economic disparities by race and origin. https://mn.gov/deed/assets/24-02MN_tcm1045-435939.pdf
Minnesota Sentencing Guidelines Commission. (2015). Minnesota Sentencing Guidelines and Commentary: August 1, 2015. https://mn.gov/sentencing-guidelines/assets/2015%20Minnesota%20Sentencing%20Guidelines%20and%20Commentarytcm30-46700.pdf
Mueller, D., & Kluckow, R. (2025). Prisoners in 2023 – Statistical tables. U.S. Department of Justice. Bureau of Justice Statistics.
Murakawa, N., & Beckett, K. (2024). Reflections on the shadow carceral state. *Theoretical Criminology*, *28*(4), 424–436.
National Academies of Sciences, Engineering, and Medicine. (2023). *Reducing racial inequality in crime and justice: Science, practice, and policy*. The National Academies Press. https://doi.org/10.17226/26705
Omori, M., & Petersen, N. (2020). Institutionalizing inequality in the courts: Decomposing racial and ethnic disparities in detention, conviction, and sentencing. *Criminology*, *58*(4), 678–713.
O'Neill, K. K., Kennedy, I., & Harris, A. (2022). Debtors' blocks: How monetary sanctions make between-neighborhood racial and economic inequalities worse. *Sociology of Race and Ethnicity*, *8*(1), 43–61.
Page, J., & Soss, J. (2025). *Legal plunder: The predatory dimensions of criminal justice*. University of Chicago Press.

Pager, D. (2003). The mark of a criminal record. *American Journal of Sociology*, *108*(5), 937–975.
Pattillo, M., & Kirk, G. (2020). Pay unto Caesar: Breaches of justice in the monetary sanctions regime. *UCLA Criminal Justice Law Review*, *4*(1), 49–77.
Peterson, R. D., & Hagan, J. (1984). Changing conceptions of race: Towards an account of anomalous findings of sentencing research. *American Sociological Review*, *49*(1), 56–70.
Pettit, B., & Western, B. (2004). Mass imprisonment and the life course: Race and class inequality in US incarceration. *American Sociological Review*, *69*(2), 151–169.
Phelps, M. S. (2018). Discourses of mass probation: From managing risk to ending human warehousing in Michigan. *The British journal of criminology*, *58*(5), 1107–1126.
Phelps, M. S., & Ruhland, E. L. (2022). Governing marginality: Coercion and care in probation. *Social Problems*, *69*(3), 799–816.
Pleggenkuhle, B., Kras, K. R., & Huebner, B. M. (2021). Twice punished: Perceived procedural fairness and legitimacy of monetary sanctions. *Journal of Contemporary Criminal Justice*, *37*(1), 88–107.
Ray, V. (2019). A theory of racialized organizations. *American Sociological Review*, *84*(1), 26–53.
Roepke, D. (2004, November 28). Jail boom looming. Inforum. https://www.inforum.com/newsmd/jail-boom-looming
Ruzek, T. (2004, March 5). Austin edition—Board sums up jail issue. Post Bulletin. https://www.postbulletin.com/news/austin-edition-board-sums-up-jail-issue
Sampson, R. J., & Neil, R. (2024). The social foundations of racial inequalities in arrest over the life course and in changing times. *Criminology*, *62*(2), 177–204.
Schanzenbach, M., & Yager, M. (2006). Prison time, fines, and federal white-collar criminals: The anatomy of a racial disparity. *The Journal of Criminal Law and Criminology*, *96*(2), 757–793.
Sentencing Project. (2021). Criminal justice facts state-by-state data. https://www.sentencingproject.org/the-facts/#map?dataset-option=SIR
Sewell, A. A. (2016). The racism-race reification process: A mesolevel political economic framework for understanding racial health disparities. *Sociology of Race and Ethnicity*, *2*(4), 402–432.
Shannon, S. K., Uggen, C., Schnittker, J., Thompson, M., Wakefield, S., & Massoglia, M. (2017). The growth, scope, and spatial distribution of people with felony records in the United States, 1948–2010. *Demography*, *54*(5), 1795–1818.
Slavinski, I., & Pettit, B. (2022). Proliferation of punishment: The centrality of legal fines and fees in the landscape of contemporary penology. *Social Problems*, *69*(3), 717–742.
Soss, J., Fording, R. C., & Schram, S. (2011). *Disciplining the poor: Neoliberal paternalism and the persistent power of race*. University of Chicago Press.
Spohn, C. (2017). *Race and sentencing disparity*. In E. Luna (Ed.), *Reforming criminal justice: Vol. 4: A report of the Academy for Justice on bridging the gap between scholarship and reform* (pp. 169–186). Arizona State University.
Spohn, C., & Holleran, D. (2000). The imprisonment penalty paid by young, unemployed Black and Hispanic male offenders. *Criminology*, *38*(1), 281–306.
Stanoch, J. M., & Wieland, L. A. (2018). *System analysis of jail population drivers and trends*. Hennepin County. https://www.hennepin.us/cjcc/-/media/cjcc/other-docs/system-analysis-of-jail-population-drivers-and-trends-report-2018.pdf
Steffensmeier, D., & Demuth, S. (2000). Ethnicity and sentencing outcomes in U.S. federal courts: Who is punished more harshly? *American Sociological Review*, *65*(5), 705–729.
Stewart, R. (2019). *Probation sentences in Minnesota*. Minnesota Justice Research Center. https://653f73c4-99e7-4517-964b-f993fa08e67a.filesusr.com/ugd/d2a74f_fd57f4859d4e4ebbb4b943fb05ac4c5f.pdf
Stewart, R., Watters, B., Horowitz, V., Larson, R., Sargent, B., & Uggen, C. (2022). Native Americans and monetary sanctions. *The Russell Sage Foundation Journal of the Social Sciences*, *8*(2), 137–156.
Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* (80–108). Cambridge University Press.
Stolzenberg, L., D'Alessio, S. J., & Eitle, D. (2013). Race and cumulative discrimination in the prosecution of criminal defendants. *Race and Justice*, *3*(4), 275–299.
Strange, C. C., Cochran, J. C., Wooldredge, J., & Long, J. S. (2024). Jail utilization and court sentencing: Does jail overcrowding influence state court sentencing decisions? *Justice Quarterly*, *41*(5), 672–698.

Sutton, J. R. (2013). Structural bias in the sentencing of felony defendants. *Social Science Research*, *42*(5), 1207–1221.
Thomas, M., Herring, C., Horton, H. D., Semyonov, M., Henderson, L., & Mason, P. L. (2020). Race and the accumulation of wealth: Racial differences in net worth over the life course, 1989–2009. *Social Problems*, *67*(1), 20–39.
Tonry, M., & Melewski, M. (2008). The malign effects of drug and crime control policies on black Americans. *Crime and Justice*, *37*(1), 1–44.
Tonry, M. H. (2011). *Punishing race: A continuing American dilemma*. Oxford University Press.
Uggen, C., & Stewart, R. (2014). Piling on: Collateral consequences and community supervision. *Minnesota Law Review*, *99*, 1871–1910.
Ulmer, J., Painter-Davis, N., & Tinik, L. (2016). Disproportional imprisonment of Black and Hispanic males: Sentencing discretion, processing outcomes, and policy structures. *Justice Quarterly*, *33*(4), 642–681.
Ulmer, J. T., & Bradley, M. S. (2018). Punishment in Indian country: Ironies of federal punishment of Native Americans. *Justice Quarterly*, *35*(5), 751–781.
Ulmer, J. T., & Johnson, B. (2004). Sentencing in context: A multilevel analysis. *Criminology*, *42*(1), 137–178.
Ulmer, J. T., & Konefal, K. (2019). Sentencing the "other": Punishment of Latinx defendants. *UCLA Law Review*, *66*, 1716–1761.
Unnever, J. D. (1982). Direct and organizational discrimination in the sentencing of drug offenders. *Social Problems*, *30*(2), 212–225.
Vera Institute of Justice. (2020). Incarceration Trends dataset [Data set]. https://trends.vera.org/
Wacquant, L. (2001). Deadly symbiosis when ghetto and prison meet and mesh. *Punishment & Society*, *3*(1), 95–133.
Wacquant, L. J. D. (2009). *Punishing the poor: The neoliberal government of social insecurity*. Duke University Press.
Wakefield, S., & Uggen, C. (2010). Incarceration and stratification. *Annual Review of Sociology*, *36*(1), 387–406.
Waldfogel, J. (1995). Are fines and prison terms used efficiently? Evidence on federal fraud offenders. *The Journal of Law and Economics*, *38*(1), 107–139.
Wang, X., & Mears, D. P. (2010). A multilevel test of minority threat effects on sentencing. *Journal of Quantitative Criminology*, *26*(2), 191–215.
Watters, B. M., Stewart, R., & Statz, M. (2024). Retribution for tribal sovereignty: Settler colonial policing and civil justice impacts. *Punishment & Society*, *26*(4), 693–710.
Western, B. (2006). *Punishment and inequality in America*. Russell Sage Foundation.
Western, B., & Sirois, C. (2019). Racialized re-entry: Labor market inequality after incarceration. *Social Forces*, *97*(4), 1517–1542.
Wilson, A. (2023). Probation and parole revocations: Reviewing over two decades of revocation decisions in the United States. *Criminal Justice Review*, *48*(3), 392–415. https://doi.org/10.1177/07340168231170335
Wu, J. (2023). Within-race variations in sentencing outcomes: Nationality and punishment among Asians in United States federal courts. *Punishment & Society*, *25*(2), 449–470.
Zuberi, T. (2000). Deracializing social statistics: Problems in the quantification of race. *The ANNALS of the American Academy of Political and Social Science*, *568*(1), 172–185.

How to cite this article: Larson, R., Stewart, R., Horowitz, V., & Uggen, C. (2026). The racialized packaging of punishment: An instrumental variables approach to incarceration, probation, and monetary sanctions. *Criminology*, 1–40.
https://doi.org/10.1111/1745-9125.70045

APPENDIX A

In this appendix, we provide additional technical detail on our analytical approach. We first provide a more thorough discussion of the limits of the standard approach and our specification and follow with additional justification and explanation of our instrumental variable approach.

A.1 | Limits of the standard approach

The "standard" approach to the estimation of the effect of race on punishment measures is common in quantitative administrative designs assessing the extra-legal effects of defendant demographic variables (Baumer, 2013). This design has an appealing and intuitive logic: when statistically adjusting for characteristics such as charge severity, criminal history, and other legal, demographic, and contextual factors, this model attempts to isolate the impact of race on sentencing outcomes using "apples-to-apples" case comparisons, at least to the extent that the model adjusts for other relevant determinants of sentencing. This effect captures the "residual" rather than the total effect of race in sentencing.

This "residual" effect of race in the sentencing process is often conceptualized by analysts as racial discrimination, though it does not represent the totality of the impact of race. Race also patterns the distributions of criminal history (e.g., through the geography of police patrol), charging decisions, and other aspects of the criminal legal process (e.g., the presence of a public defender), some of which are essentially posttreatment controls in this model. This effect, however, is designed to capture the marginal change in sentence in response to a change in a defendant's racial group and the accompanying social implications, with all other observed variables held equal. Therefore, we join contemporary scholars (e.g., Sewell, 2016) in conceptualizing individual-level race estimates not simply as the impact of a simple demographic attribute in the sentencing phase. Rather, we view such estimates as the product of the broader constellation of constructed race meanings and structures within the criminal legal system, such as social inequalities, relationships to social institutions, and other social interactional processes that play out within social institutions. In other words, wrapped up in this ostensibly "simple" coefficient is not just a difference in racial identification but also a plethora of associated social inequities, relationships to social institutions, and social interactional processes that play out during and after sentencing.

Despite the attractive design of the standard approach, these models may suffer from omitted variable bias to the extent that defendant race is also related to forms of sentenced punishment left *unspecified* within the model. The previous literature, in large part, only examines the extra-legal effects of one form of punishment (e.g., monetary sanctions) or considers multiple axes of punishment in the study but specifies separate models for each punishment outcome. From a statistical perspective, models that examine only a single aspect of punishment may be misspecified if that component influences another aspect of the punishment defendants receive. In statistical parlance, each mode of punishment may be endogenous to the other forms, and therefore not explicitly modeling these interrelationships would result in model misspecification and omitted variable bias. To the extent that punishment outcomes influence one another, *and are also correlated with race*, omitting other forms of punishment from the model could bias the estimated race estimates, resulting in a risk of ill-informed conclusions about the role of defendant race in sentencing.

Previous scholarship suggests each axis of punishment considered here—incarceration, probation, and monetary sanctions—*are also raced* in and of themselves. Thus, a key innovation of our analysis is to account for the racial patterns in the other modes of punishment, to obtain not only a more complete, but also a more *nuanced*, estimation of the influence of race in the criminal sentencing process for each punishment mode. Analyses that do not account for these interdependencies risk having a biased estimation of racial impacts in the criminal sentencing process.

Moreover, estimating the effect of incarceration length on fines/fees and probation is complicated by two sources of bias. First, if judges adjust their confinement length on the basis of monetary sanctions or probation sentences (or vice-versa), it creates a circular causal chain in which it is difficult to pin down the direction of the causal arrow—commonly known as simultaneity bias— that can bias not only the effect of the endogenous regressor but also other effects of interest in the model. If monetary sanctions and/or probation lengths are determined in this way, it poses an endogenous simultaneity bias problem that interferes with the identification of the impact of incarceration length on the other punishment forms, as well as the identification of other effects of interest (e.g., race).

Even a specification that uses the standard approach while controlling for incarceration lengths is susceptible to bias as part of the effect of race on the other forms would be "controlled away" with the inclusion of incarceration, effectively inducing a form of posttreatment bias. Race-related factors that determine both incarceration and other forms would be purged from the race effects on the other forms. A second source of potential bias is that of collider bias, wherein conditioning on incarceration may induce a spurious correlation between race and the other forms if unobserved factors influence multiple punishment forms. A third source of bias is the traditional case of an omitted variable. In this case, an unobserved factor that causes changes in both the sentencing of incarceration length and the sentencing of another punishment component may lead to the wrong conclusion that the former caused the latter. One example of an omitted variable would be a judge's general level of punitiveness. Punitive judges may sentence more punitively on both components, thereby inducing a correlation, although a causal link between the two does not exist. Studies using the standard approach assume, often implicitly, that the omitted variables are uncorrelated with the control variables (e.g., other punishment axes), which is likely an untenable assumption given the dependent nature of punishment. Given the robust relationships exhibited between race and punishment in previous empirical work, having significant omitted variables bias in a key punishment control could bias the residual race estimates that remain.

A.2 | Justification for the instrumental variable approach

Considering the limitations of the standard approach, we introduce an instrumental variable estimator to appropriately model the endogenous nature of sentencing. This design allows us to model the effect of the incarceration component on both the monetary sanction and probation components while adjusting for multiple forms of bias. The IV approach alleviates both simultaneity and omitted variable bias that could result from key individual extralegal variables (i.e., defendant race) being collinear with incarceration sentences. Further, instrumenting incarceration isolates variation in the sentence that is determined by the instrument and observed predictors, effectively purging the mediator of unobserved heterogeneity. This allows for the inclusion of incarceration as a control while isolating the indirect racial pathway, thereby alleviating the posttreatment and collider bias concerns. The race effect, after adjusting for instrumented incarceration, can be interpreted as the direct effect of race on other punishment outcomes independent of exogenously determined incarceration differences and other observed covariates, effectively removing the indirect pathway from race to probation and monetary sanctions that operate through incarceration.

We use a time-varying measure of available county-level jail capacity as our instrument. To serve as a valid instrument and identify the causal effect of incarceration on the other punishment axes, our IV must meet three conditions or assumptions. First, the relevance condition requires that the instrument must induce a change in the endogenous independent variable, incarceration length, and that this change will translate into a sufficiently strong relationship between them. Second,

the exogeneity assumption requires that the change in capacity ratios should be uncorrelated with previous trends in the outcome variables. In other words, the values of the instrument should be independent of previous patterns in monetary sanction and probation sanctioning (i.e., the values should be distributed as though they are random), conditional on other variables in the model. Third, the exclusion restriction states that the instrument can only affect monetary sanctions and probation through its impact on confinement length.

We perform multiple robustness tests to evaluate instrument relevance and the validity of the exclusion restriction with our instrument. We show there is a strong relationship between jail capacity ratios and sentenced confinement length, providing direct evidence on the validity of the relevance condition. We follow the convention in the literature and consider the instrument to be sufficiently strong if the *F*-statistic from a Wald test on the instrumental variable is statistically significant in the first stage regression (Stock & Yogo, 2005). As Table 3 shows, the *F*-test on the instrument in the first-stage regressions yields a statistically significant *F*-statistic well above 10. Further, in Minnesota, the effect of jail capacity on incarceration length (i.e., the first-stage) exhibits significant heterogeneity (see below).

We also examine whether lagged monetary sanction and probation patterns explain changes in the jail capacity ratio, which would violate the exclusion restriction as the outcomes may lead to changes in jail capacity. We find little evidence that the changes in the jail capacity ratio are correlated with prior trends in monetary sanction sentencing or probation sentencing, with partial correlation coefficients of −0.05 and −0.06 for LFOs and probation, respectively, adjusting for lagged capacity ratios. Another key threat to the exogeneity of our instrument is that it is the consequence of prior incarceration sentencing decisions, thereby making it endogenous to incarceration. We help alleviate this concern by including sentence year and judicial district fixed effects in the first stage model, thereby adjusting the instrument for time-invariant differences in judicial punitive or sentencing philosophy as well as commonly shared shocks over time (e.g., state-wide changes in policy) that affect both capacity and sentencing decisions. We also include the full suite of covariates in the first stage specification, which will adjust for any observed differences and strengthen the assumption that the jail capacity ratio is conditionally exogenous. Finally, we argue that the instrument meets the exclusion restriction because monetary sanctions and probation lengths are unlikely to be directly impacted by jail capacities, apart from any perceptions or concerns with capacity that impact sentenced confinement amounts. If the instrumental variable assumptions are satisfied, it is possible to estimate the causal effect of confinement length on the other aspects of the punishment package.

If these conditions are met, this approach yields an estimate of the local average treatment effect (LATE) of confinement length. The estimated impact is local in that it is only estimated for the subset of cases for which the instrument—the available capacity ratio—induces a change in incarceration lengths (Angrist et al., 1996). As shown in Figure A1 and Table A1, the first-stage shows a heterogeneity in the effect of the capacity ratio on incarceration sentencing lengths in expected ways. At the felony level, a one unit increase in incarceration is associated with a 0.17 ($p < 0.001$) increase in logged incarceration lengths. At this level, we interpret this relationship to reflect overcrowding at local district levels as pushing cases to state prisons, to avoid filling local jails further that are already "bursting at the seams" as the capacity ratio increases. Conversely, among gross misdemeanor cases, the relationship between the county-level jail capacity ratio is large and negative, with a one unit increase in the jail capacity ratio associated with a −0.16 ($p < 0.001$) decrease in logged incarceration lengths. At the gross misdemeanor level, we interpret this relationship to be a more classic response to overcrowding or reducing sentence lengths to alleviate the burden on local jails.

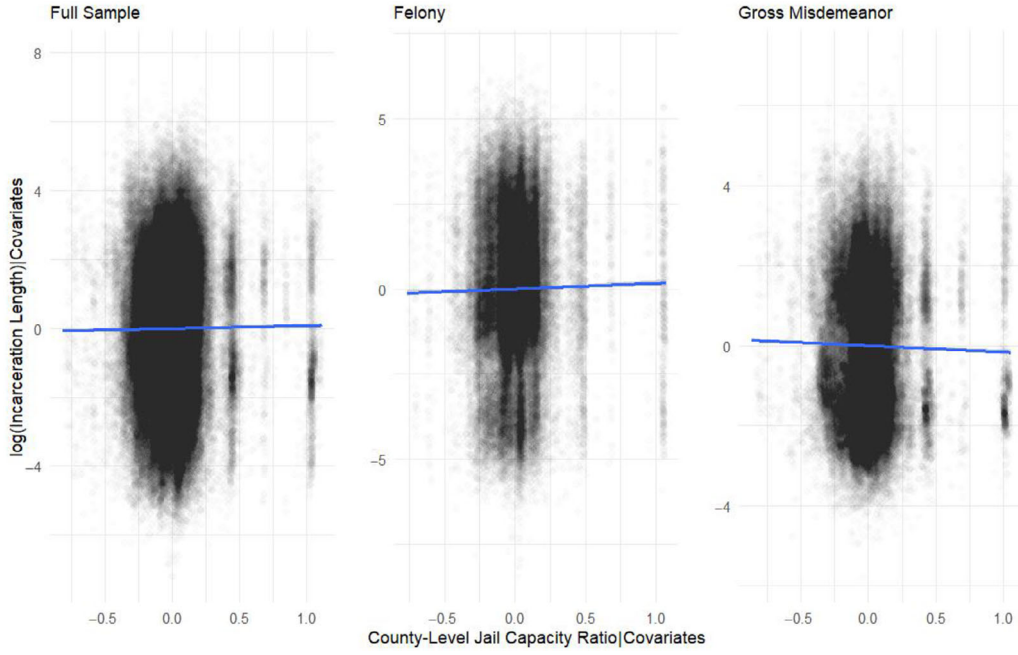

FIGURE A1 First stage partial relationships between jail capacity ratio and incarceration length. [Color figure can be viewed at wileyonlinelibrary.com]

Given this heterogeneity in the first-stage relationship of county-level jail capacity ratio and incarceration lengths, this violates the monotonicity assumption of IV estimates, as decision-makers respond differentially to the local capacity of their county jails. To estimate the LATE, treatment responses must be monotonic with the instrument, and in our analytical sample, response is not uniform across sentencing levels. This negates the LATE interpretation of the results - effectively estimating a weighted average of compliers and definers - when we treat the capacity ratio as having a uniform effect across all cases in the full sample.

In response to this heterogeneity, we formulate our first stage to include interaction terms between the county-level jail capacity ratio and indicators of felony and gross misdemeanor, which allows the effect of capacity ratio on incarceration rates to vary by sentencing-level context. In effect, this accounts for the moderating influence of sentencing level in the first stage by modeling a separate instrument effect for cases of different levels and retains the LATE interpretation respectively for case subgroups. Thus, at both levels of sentencing, we have statistically significant evidence of instrument relevance, but the direction of the effect exhibits directional heterogeneity. Felony-level cases are not unresponsive to localized capacity pressures at the district level but rather respond to capacity in a different form by pushing defendants toward longer incarceration sentences (i.e., state prison sentences), whereas gross misdemeanor cases respond to localized pressures of overcrowding by alleviating burden through reduced incarceration lengths on average.

TABLE A1 First-stage OLS models of incarceration length, Minnesota 2004–2017.

	Dependent variable: log(Incarceration)		
	Full sample	**Felony**	**Gross misdemeanor**
County-level jail capacity ratio	0.083** (0.026)	0.165*** (0.047)	−0.157*** (0.028)
Black	0.122*** (0.014)	0.456*** (0.025)	−0.142*** (0.016)
Hispanic	0.156*** (0.020)	0.499*** (0.039)	0.002 (0.022)
Asian	−0.079* (0.033)	0.021 (0.061)	−0.034 (0.037)
Native American	0.328*** (0.021)	0.459*** (0.036)	0.218*** (0.023)
Other race	−0.103* (0.041)	−0.030 (0.075)	−0.154*** (0.046)
Missing race	−0.314*** (0.018)	−0.128*** (0.038)	−0.370*** (0.018)
Male	0.700*** (0.011)	1.018*** (0.022)	0.413*** (0.012)
log(Age)	0.339*** (0.015)	0.594*** (0.028)	0.200*** (0.016)
Prior convictions	0.069*** (0.001)	0.088*** (0.002)	0.048*** (0.001)
Public defender	0.384*** (0.010)	0.404*** (0.021)	0.259*** (0.011)
Percent credit	−0.030*** (0.0002)	−0.035*** (0.0002)	−0.019*** (0.0002)
Percent stayed	−0.019*** (0.0001)	−0.035*** (0.0003)	−0.010*** (0.0002)
Trial	2.161*** (0.237)	1.884*** (0.295)	1.315** (0.432)
Felony	0.133*** (0.016)		1.101*** (0.019)
Gross misdemeanor	−1.049*** (0.016)	0.166*** (0.026)	
Violent	0.389*** (0.014)	0.721*** (0.021)	0.115*** (0.016)
Drug	−0.211*** (0.016)	−0.148*** (0.021)	0.054* (0.024)
Alcohol/DUI	0.954*** (0.012)	1.011*** (0.034)	0.959*** (0.012)
Constant	1.716*** (0.063)	0.647*** (0.132)	0.829*** (0.068)
District FE	Yes	Yes	Yes
Sentence year FE	Yes	Yes	Yes
Observations	192,155	67,289	118,713
F-statistic	2317.223*** (df = 41; 192,113)	1056.395*** (df = 40; 67,248)	717.908*** (df = 40; 118,672)

Note: All tests are two-tailed.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.