NBER WORKING PAPER SERIES

CONVICTION, INCARCERATION, AND RECIDIVISM: UNDERSTANDING THE REVOLVING DOOR

John Eric Humphries
Aurelie Ouss
Kamelia Stavreva
Megan T. Stevenson
Winnie van Dijk

Working Paper 32894
http://www.nber.org/papers/w32894

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2024, Revised December 2024

Thanks to Meredith Farrar-Owens and others at the Virginia Criminal Sentencing Commission for providing data and answering questions, and to Ben Schoenfeld for web scraping Virginia criminal court records and making them publicly available. We are grateful to Alex Albright, Steve Berry, Jordan Cammarota, Jiafeng (Kevin) Chen, Will Dobbie, Deniz Dutz, Brigham Frandsen, Anjelica Hendricks, Felipe Goncalves, Hans Gr¨onqvist, Phil Haile, Randi Hjalmarsson, Rucker Johnson, Larry Katz, Emily Leslie, Charles Loeffler, Jens Ludwig, Alex Mas, Magne Mogstad, Jack Mountjoy, Derek Neal, Arnaud Philippe, Vitor Possebom, Steve Raphael, Yotam Shem-Tov, Elie Tamer, Pietro Tebaldi, Alex Torgovitsky, Crystal Yang, Ed Vytlacil, Chris Walker, and seminar participants for helpful comments. We thank Magdalena Dominguez, Jeff Grogger, Vishal Kamat, and Mike Mueller-Smith for serving as discussants. We thank Cecile Macaire, Naomi Shimberg, Joost Sijthoff, Iliana Cabral, and the UVA Law Librarians for excellent research assistance. We also thank Arnold Ventures and the Tobin Center for Economic Research for financial support. Any remaining errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2024 by John Eric Humphries, Aurelie Ouss, Kamelia Stavreva, Megan T. Stevenson, and Winnie van Dijk. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Conviction, Incarceration, and Recidivism: Understanding the Revolving Door
John Eric Humphries, Aurelie Ouss, Kamelia Stavreva, Megan T. Stevenson, and Winnie van Dijk
NBER Working Paper No. 32894
August 2024, Revised December 2024
JEL No. J0, K4

ABSTRACT

Noncarceral conviction is a common outcome of criminal court cases: for every individual incarcerated, there are approximately three who are recently convicted but not sentenced to prison or jail. We develop an empirical framework for studying the consequences of noncarceral conviction by extending the binary-treatment judge IV framework to settings with multiple treatments. We outline assumptions under which widely-used 2SLS regressions recover margin-specific treatment effects, relate these assumptions to models of judge decision-making, and derive an expression that provides intuition about the direction and magnitude of asymptotic bias when they are not met. Under the identifying assumptions, we find that noncarceral conviction (relative to dismissal) leads to a large and long-lasting increase in recidivism for felony defendants in Virginia. In contrast, incarceration relative to noncarceral conviction leads to a short-run reduction in recidivism, consistent with incapacitation. While the identifying assumptions include a strong restriction on judge decision-making, we argue that any bias resulting from its failure is unlikely to change our qualitative conclusions. Lastly, we introduce an alternative empirical strategy, and find that it yields similar estimates. Collectively, these results suggest that noncarceral felony conviction is an important and potentially overlooked driver of recidivism.

John Eric Humphries
Department of Economics
Yale University
87 Trumbull Street
New Haven, CT 06511
and NBER
johneric.humphries@yale.edu

Aurelie Ouss
Department of Criminology
University of Pennsylvania
571 McNeil Building
3718 Locust Walk
Philadelphia, PA 19104
and NBER
aouss@upenn.edu

Kamelia Stavreva
Columbia University
kes2220@columbia.edu

Megan T. Stevenson
University of Virginia
mstevenson@law.virginia.edu

Winnie van Dijk
Department of Economics
Yale University
87 Trumbull Street, B228
New Haven, CT 06520
and NBER
winnie.vandijk@yale.edu

1 Introduction

The U.S. criminal justice system is commonly referred to as a "revolving door" due to the high rate of recidivism among those who come into contact with it.[1] A key question for policy makers is whether the criminal justice system itself contributes to these patterns or whether they are driven by external factors such as addiction, mental health, neighborhood disadvantage, or limited labor market opportunities. Much of the available quantitative research has focused on how *incarceration* affects recidivism. However, *noncarceral conviction* (a conviction that does not result in incarceration) is a frequent outcome in the criminal court system.[2] For instance, in 2010, 2.7 individuals were on probation for every person who was incarcerated (Phelps, 2013). A noncarceral conviction could directly affect recidivism through several channels. It may induce crime by reducing its opportunity cost. For example, a conviction record could make it harder to find employment, making crime relatively more attractive. A conviction could also increase future criminal justice contact even if it has no impact on criminal behavior. For example, prosecutors may be more likely to pursue charges against someone with a recent conviction on their record, and judges may sentence them more harshly. Conversely, a conviction could act as a deterrent if it increases the expected penalties for future crime.

In this paper, we provide new evidence on how both felony noncarceral conviction and incarceration affect future criminal justice involvement. Our main approach follows existing research by using quasi-random assignment of cases to judges as a source of exogenous variation, but our discussion formalizes an extension of this research design from two to three treatments. Our goal is to learn about *margin-specific* treatment effects: causal impacts of noncarceral conviction relative to dismissal of all charges, and causal impacts of incarceration relative to noncarceral conviction. These quantities allow us to isolate the impact of mechanisms that come into play when someone is convicted without a carceral sentence (such as having a felony conviction record or increased supervision) from the impact of mechanisms that matter for incarceration (such as incapacitation).

We study a newly-constructed panel of felony cases in Virginia, spanning approximately two decades. Our outcomes are new felony charges, new convictions, and new carceral sentences. Following the literature, we refer to these outcomes as "recidivism." Our results point to noncarceral conviction as an important, long-lasting driver of recidivism, consistent with a criminogenic effect of a felony conviction record. By contrast, we only find evidence of a short-term decrease in recidivism due to incarcer-

[1] According to the Bureau of Justice Statistics, 44% of people released from prison in the U.S. in 2005 were rearrested within one year. Nine years later, 83% had been rearrested at least once (Alper et al., 2018).

[2] We will at times refer to "noncarceral conviction" as "conviction" for brevity.

ation, which is likely due to incapacitation.

Our discussion proceeds in three parts. First, we develop an empirical framework to discuss the interpretation of judge-stringency 2SLS estimands in a multiple-treatment setting with full treatment effect heterogeneity. Prior applied work using 2SLS with multiple treatments has often used instruments that are reasonably thought of as varying the net payoff to taking up a "focal" treatment (e.g., Kline and Walters, 2016; Kirkeboen et al., 2016; Mountjoy, 2022). For such instruments, it may be justifiable to assume that they are *treatment-specific*, i.e., they either encourage or discourage take-up of the focal treatment and do not cause any switches between other "non-focal" treatments. This property, combined with the usual IV assumptions, ensures that 2SLS regressions can be used to identify causal effects of the focal treatment, relative to one or potentially a mix of alternatives.[3] However, judge stringency instruments do not generally vary the net payoff to take-up. Instead, they represent the *shares* of cases a judge allocates to specific court outcomes.

We argue that this property of judge stringency instruments has a benefit and a drawback. On the one hand, we show that treatment specificity is sufficient for 2SLS with judge stringency instruments to identify *margin-specific* causal effects, unlike in the previously-cited applications. On the other hand, requiring these instruments to be treatment-specific could be considered a strong restriction on judge behavior. We provide intuition for the restrictiveness of this assumption by examining how it constrains models of judge decision-making. We consider three commonly-used discrete-choice models, applied to judge decision-making over three court outcomes: dismissal, non-carceral conviction, and incarceration. The models we consider are ordered, sequential, and multinomial choice models. Among these, the only model in which both of the judge-stringency instruments are treatment-specific is the ordered model. For the sequential and unordered models, which are arguably more realistic in our setting, at least one of the instruments is not treatment-specific. However, all satisfy a weaker assumption which we label conditional pairwise monotonicity (CPM).[4]

We then derive an expression for the asymptotic bias in the 2SLS estimand under CPM. The bias term is additive and easy to interpret. It provides intuition about the direction and magnitude of asymptotic bias when CPM holds, but treatment-specificity does not. Moreover, it clarifies how assumptions on treatment effect heterogeneity, or on the relative effects for compliers on different margins, can sign or eliminate the bias in lieu of assuming a more restrictive model of judge behavior.

In the second part of the paper, we turn to our main empirical contributions:

[3]Here, we follow the literature in referring to an estimand as "causal" if it is a non-negatively weighted average of local average treatment effects (LATEs).

[4]CPM is related to the "no defiers" assumption from the binary case in that it assumes that an instrument induces flows in only one direction across each margin.

estimating the impacts of noncarceral conviction and incarceration on future criminal justice involvement. We use 2SLS with the conviction propensity of judges as an instrument for conviction, while controlling for their incarceration propensity.[5] Analogously, we use judges' incarceration propensity as an instrument for incarceration and control for their dismissal propensity. Under the assumptions described in the first part of our discussion, our estimates imply that noncarceral conviction relative to dismissal leads to large and long-lasting increases in future justice involvement, while incarceration relative to noncarceral conviction decreases recidivism in the first year, likely due to incapacitation.[6]

Our results on noncarceral conviction are consistent with both increased criminal behavior and an escalation in subsequent criminal justice responses. We examine how impacts differ by prior records, types of offenses, and measures of recidivism, but do not find evidence that supports one hypothesis over another. Both channels imply that a felony conviction can lead individuals to cycle back into the criminal justice system, leading to increased charges, convictions, and future incarceration. Overall, when given a causal interpretation, our results underscore the significant role of conviction even in the absence of imprisonment.

To probe whether it is reasonable to interpret our estimates as causal and margin-specific effects, we conduct an empirical test of whether the instruments are treatment-specific. The test also lets us adjudicate between different models of judge decision-making. Our findings suggest that treatment specificity does not hold in our setting, meaning that neither stringency instrument moves people across only a single margin. This implies that we can empirically reject the ordered and sequential models of judge decision-making. We use our expression for asymptotic bias, along with theory and empirical evidence, to argue that the failure of treatment specificity is unlikely to overturn our qualitative conclusion regarding the effect of noncarceral conviction. The bias for the long-run effect of conviction is likely zero or small, owing to likely null effects of incarceration on recidivism post-incapacitation (e.g., Norris et al., 2021; Rose and Shem-Tov, 2021; Garin et al., 2023, and the regression-discontinuity estimates from our setting). The sign of the bias term for the short run impact of conviction is likely small and negative, based on the likely composition of compliers.

To assuage any remaining concerns about bias in the 2SLS estimates, in the third part of the paper, we provide an alternative approach for identifying and estimating the impacts of conviction and incarceration. We develop a novel approach that builds on Mountjoy (2022) to identify margin-specific treatment effects in a multiple-treatment

[5]This approach mirrors a strategy used in the literature studying the impact of incarceration on recidivism. See Loeffler and Nagin (2022) and Doleac (2023) for recent reviews of this literature.

[6]We additionally examine the effects of incarceration using a regression discontinuity design based on sentencing guidelines, yielding estimates that are consistent with our main findings.

context. This approach requires treatment-specific instruments, which we have argued judge stringencies generally are not. Following methods from the discrete choice literature, we impose additional structure on the choice problem to construct treatment-specific instruments from judge stringencies. We then use these newly-constructed instruments to obtain estimates of margin-specific treatment effects. The results are similar to our 2SLS estimates, although they are somewhat smaller and sometimes less precise.

This research contributes to both substantive and methodological literatures. First, our work is related to a small set of recent studies that explore the impact of criminal convictions. Two of these studies show that felony diversion causes large and sustained reductions in future criminal justice contact (Mueller-Smith and Schnepel, 2021; Augustine et al., 2022). Felony diversion helps avoid conviction, but can also affect recidivism through other channels. For instance, there may be enhanced deterrence, since rearrest leads to reinstated charges. Nonetheless, the authors present compelling evidence that felony conviction plays a substantial role in the documented effect. In the context of misdemeanors, Agan et al. (2023b) show that the decision to file charges increases future contact with the criminal justice system. However, only 26% of those charged receive a misdemeanor conviction, and the authors argue that the mark of a conviction is not the main channel explaining this effect. In related work, Kamat et al. (2024) adopt a partial-identification approach and find that misdemeanor conviction increases the number of future charges, but they cannot rule out large effects of felony conviction in either direction. Additionally, there is a deep socio-legal literature providing theoretical arguments, as well qualitative and descriptive evidence about the adverse effects of both felony and misdemeanor convictions (e.g., Chiricos et al., 2007; Natapoff, 2011; Phelps, 2017; Irankunda et al., 2020). We contribute to the existing literature by disentangling conviction from other aspects of the criminal process and by assessing the relative importance of felony conviction and incarceration in explaining future criminal justice involvement within the same setting.

Second, this paper contributes to the large body of work investigating the consequences of incarceration for recidivism.[7] A recent review shows that post-conviction incarceration generally is not found to have long-term effects on recidivism, while pretrial detention often is found to increase recidivism after the incapacitation period (Loeffler and Nagin, 2022). Our study suggests one way to reconcile these findings: since pretrial detention increases the likelihood of conviction, adverse effects of pretrial detention may be operating through conviction rather than the experience of

[7]E.g., Kling (2006); Hjalmarsson (2009); Kuziemko (2013); Loeffler (2013); Aizer and Doyle (2015); Mueller-Smith (2015); Gupta et al. (2016); Leslie and Pope (2017); Estelle and Phillips (2018); Harding et al. (2018); Dobbie et al. (2018); Bhuller et al. (2020); Norris et al. (2021); Rose and Shem-Tov (2021); Arteaga (2021); Franco et al. (2022); Jordan et al. (2023); Garin et al. (2023).

incarceration itself. Studies that identify the impacts of post-conviction incarceration, meanwhile, are often comparing incarceration to noncarceral conviction, with both the treatment and control groups being convicted.

We build on a methodological literature about the identification and estimation of treatment effects in the presence of multiple treatment alternatives. The prior and contemporaneous literature has outlined many of the challenges associated with multiple treatments (e.g., Heckman et al., 2006; Heckman and Vytlacil, 2007; Heckman et al., 2008; Kline and Walters, 2016; Kirkeboen et al., 2016; Heckman and Pinto, 2018; Lee and Salanié, 2018; Mountjoy, 2022; Heinesen et al., 2022; Bhuller and Sigstad, 2024; Kamat et al., 2024). However, not all of the insights developed in the prior literature apply to the judge IV setting, given the special nature of judge stringency instruments as shares.[8] Identification issues specific to judge IV in a multiple treatment setting have received sustained consideration in two prior papers studying the impacts of incarceration. Mueller-Smith (2015) provides one of the first in-depth discussions of the challenges inherent in this design and proposes controlling for judge stringency along "non-focal" dimensions (such as fine amount or probation length). Arteaga (2021) discusses multiple-treatment identification issues and shows how to identify causal effects along the incarceration margin within a sequential model.

Our paper contributes to the methodological literature in several ways. First, we lay out identifying assumptions sufficient for judge IV to yield a causal and margin-specific estimand when there are multiple treatments. In contemporaneous work, Bhuller and Sigstad (2024) present an alternative set of identifying conditions for 2SLS with multiple treatments. Their regression model is different: it instruments for all treatments simultaneously, and thus requires stronger functional form assumptions than our approach. The monotonicity conditions they propose are weaker than ours, but ours have the advantage of having straightforward and tractable relationships with economic models of judge behavior. Indeed, one of our contributions is to show how our econometric assumptions relate to three commonly used discrete choice models. This helps illuminate the econometric implications associated with different ways of modeling the court system. We also derive an expression for asymptotic bias under a weaker set of monotonicity assumptions which all of the choice models we consider satisfy. We suggest an empirical test for instruments' treatment-specificity, and we demonstrate how to reason about the sign and magnitude of the bias term if the assumption is rejected. Finally, we show how to derive treatment-specific instruments from judge stringency instruments, thus allowing the researcher to apply the identifi-

[8]For instance, we show that the treatment-specific instruments assumption is sufficient to yield causal and margin-specific treatment effects in our setting, while Mountjoy (2022) shows that it generally is not. Similarly, the differencing technique presented in Mountjoy (2022) is not possible with judge stringencies as instruments, although we show how to adapt the method to this setting.

cation approach presented in Mountjoy (2022), or other approaches that require such instruments (e.g., Lee and Salanié, 2018).

Lastly, our paper is related to a broad literature of applied work that uses judge instruments. We offer a practical guide for research designs using such instruments when judges choose between more than two options.[9] Researchers can use their institutional knowledge to reason about which choice model fits best, apply the tests that we suggest across models to see if the data is consistent with their institutional knowledge and, if necessary, reason about the likely sign and magnitude of the bias. Our paper suggests that if both institutional expertise and the tests support an ordered model, 2SLS is a good choice, assuming the other identifying assumptions are met. If either institutional knowledge or the empirical test reject the ordered model, then 2SLS estimates may have an additional bias term for at least one of the margin-specific contrasts. In that case, theory and estimates from prior literature can help the researcher to reason about the sign and magnitude of the bias, as we demonstrate in our setting. Lastly, our alternative approach to identification can be used if institutional knowledge and empirical tests support an unordered model. It can also be used as a robustness check to IV specifications.

The paper proceeds as follows. Section 2 describes the institutional setting and our data. Section 3 extends the random judge design to multiple treatments and presents a set of sufficient conditions for 2SLS to recover causal and margin-specific treatment effects. We show how the treatment-specific instruments assumption rules out some commonly used models of discrete choice, and then derive an expression for the asymptotic bias if this assumption is not met. Section 4 presents the empirical evidence based on 2SLS estimates and introduces an empirical test for treatment-specific instruments. Section 5 describes an alternative approach to identification and estimation, as well as corresponding empirical results. Section 6 summarizes results.

[9] Judge stringency instruments have been used in the criminal justice setting (e.g., Mueller-Smith, 2015; Bhuller et al., 2020; Norris et al., 2021; Arteaga, 2021; Huttunen et al., 2020), but also in other settings, such as foster care (Doyle, 2008; Gross and Baron, 2022; Baron and Gross, 2022), disability claims (Maestas et al., 2013; French and Song, 2014), bankruptcy (Dobbie and Song, 2015; Dobbie et al., 2017), eviction (Collinson et al., 2024), or patent decisions (Sampat and Williams, 2019; Feng and Jaravel, 2020; Gavrilova and Juranek, 2021). Many of these settings can be thought of as having multiple alternatives. For example, in the context of pretrial detention decisions, one could be interested in the effect on failure to appear of pretrial detention, vs electronic monitoring, vs release as in Rivera (2023). Outside of the criminal justice context, one might consider the effect on mortality of opioid prescription, vs other pain medication, vs no prescription, or the effect on homeowners' financial situation of foreclosure, vs loan modification, vs no court action.

2 Institutional details and data

2.1 Felony case processing in Virginia

This section describes felony criminal case processing in Virginia, with a focus on adjudication within the Circuit Court, which is the primary data source for this paper.

Between arrest and Circuit Court. After a person is arrested, they are brought to the local police station, booked, and held for their bail hearing. Bail is set by a magistrate, a member of the judiciary who will not preside over further hearings on the case. Charges are first filed in District Court, where the preliminary hearing will be held.[10] At this hearing, the prosecutor must convince the judge that there is probable cause that the defendant committed a felony. This hearing is also the first stage in which plea negotiations might occur. Felony charges might be negotiated down to misdemeanors, or the charges might be dropped or dismissed entirely. If the judge finds probable cause for a felony, the case will then proceed to a grand jury hearing in which a panel of citizens conducts an additional review of the evidence to ensure that probable cause has been met. If the grand jury finds probable cause that the defendant committed a felony, charges will be filed in Circuit Court, where the remainder of the criminal proceedings will occur.[11] Our analyses include only cases that make it to Circuit Court (roughly 90% of felony charges).

Assignment of cases to judges. Once charges have been filed in Circuit Court, the case will be assigned to a judge. The exact assignment procedure varies across jurisdictions.[12] A few examples include: (1) the clerk drawing colored stickers out of a can to assign judges; (2) a rotating schedule where a judge will see all cases scheduled for that court during that rotation; (3) assignment of judges to cases based on availability; and (4) cases assigned to judges based on whether the case number is odd or even. Appendix E shows that our results are robust to which jurisdictions we include.

Adjudication within Circuit Court. Once a judge has been assigned, the defendant must decide whether she wants to plead guilty or take the case to trial. Since the decision about how to plead depends partly on her expectations of success at trial,

[10] District Court is a court of limited jurisdiction, meaning that one cannot be convicted of a felony there. District Court adjudicates misdemeanors and provides initial screenings for felonies.

[11] There are some potential variations of this process. For instance, defendants can waive their right to a preliminary hearing or a grand jury hearing, and prosecutors can bypass the preliminary hearing and directly indict the case with the grand jury.

[12] We conducted phone interviews with court clerks to determine how cases were assigned to judges.

we describe the trial process first. Trials in Virginia can be either in front of a judge, which is called a bench trial, or a jury. Approximately 15% of felony convictions in our sample come from trials, almost all of which are bench trials. The remainder come from guilty pleas.[13] In a bench trial, the judge decides whether to convict and, if so, what sentence to give.[14] Judges also exert substantial indirect influence on adjudication and sentencing through various motions. For instance, judges decide what evidence is admissible, what charges can proceed, what must be struck from the record, and what instructions the jury receives. Many of these decisions are made prior to trial. Since they influence the expected outcome of a trial case, they also influence the willingness to offer or accept a plea deal. The more motions are resolved in favor of the defense, the stronger her bargaining position will be. Plea negotiations may result in a stipulated sentence and/or an agreement that the prosecutor will request a particular sentence. Virginia uses a sentence guidelines system, but the judge makes the final decision about the sentence: they have the latitude to reject any negotiated plea deal and to deviate from the sentence guidelines if they provide a written explanation.

These features show that judges influence both conviction and incarceration decisions in many ways, even if they do not fully control them. This is important for our research design since we use judge stringencies as instruments in our main analyses.[15]

Virginia's criminal justice system compared to other states. Appendix A compares aggregate statistics of Virginia's criminal justice system to both national averages and statistics for states considered in other recent studies of the impacts of incarceration. Virginia is similar in terms of incarceration and probation rates, and has similar racial and ethnic composition of its incarcerated population. However, it has lower than average parole rates. This is because Virginia adopted "truth in sentencing" for felony convictions starting in 1995, which requires people with felony convictions to serve at least 85% of their prison term. As a result, the initial sentence is much more closely linked to time spent incarcerated than in other places.

2.2 How noncarceral conviction and incarceration may affect recidivism

Noncarceral conviction. Receiving a felony conviction instead of a dismissal could increase or decrease recidivism through a number of channels. It could decrease

[13] Plea resolutions are somewhat less frequent in Virginia than in other states. For example, in 2009, nationally, 93% of felony convictions occurred through a guilty plea (Reaves, 2013).

[14] In a jury trial, the jury decides both guilt and sentencing, although the judge can reduce the sentence.

[15] We provide more institutional details related to the relevance of judge stringency for case outcomes as well as empirical evidence in Appendix D.

recidivism via deterrence. For example, a person who is convicted but not incarcerated is often placed on probation, which entails additional surveillance and scrutiny, thus increasing the probability of apprehension. It could also raise sentences conditional on conviction, since prior convictions are used to determine recommended sentences. Both of these channels suggest that noncarceral conviction increases the expected punishment for future offenses, thereby raising the costs of crime and potentially dampening recidivism (Drago et al., 2009; Philippe, 2020).

Alternatively, felony convictions may increase recidivism due to the stigma and destabilization associated with such records.[16] Employers or landlords conducting background checks may be dissuaded from hiring or renting to someone with a felony conviction, raising the cost of finding work in the formal sector, depressing future wages, and driving those with felony conviction to move into neighborhoods with higher overall crime rates (Pager, 2003; Holzer et al., 2006, 2007; Agan and Starr, 2018; Doleac and Hansen, 2020; Craigie, 2020; Rose, 2021a; Agan et al., 2023a).[17]

A prior conviction may also increase our measures of recidivism by changing the outcomes of future criminal justice interactions, even with no changes to future criminal behavior. Our recidivism measures are based on new felony charges, convictions, and carceral sentences, all of which involve discretionary decisions by various criminal justice actors. A prior conviction may influence these decisions, leading to a "ratcheting up" of penal responses, where each subsequent interaction with the criminal justice system results in more severe consequences. Criminal justice actors have access to the full criminal record at nearly all stages of decision-making, and prior convictions can impact, for example, the likelihood that someone will be detained pretrial, or the prosecutor's willingness to offer diversion or bargain the charges down to a misdemeanor.[18,19]

[16]We note that our paper focuses on felony charges, and not on misdemeanors. While misdemeanor charges are more common (Mayson and Stevenson, 2020), they generally carry fewer legal and extra-legal consequences (Agan et al., 2023a).

[17]Both arrests and convictions are visible on background checks and both may influence employers' and landlords' decisions. However, convictions are likely considered more serious than arrests that do not lead to conviction, since convictions have met a higher burden of proof. Agan et al. (2024) find evidence in support of such differential consideration of arrests and convictions in a survey of hiring professionals. Furthermore, those with a felony conviction are prohibited by law from certain types of employment and from receiving certain public benefits. In contrast, arrests that do not lead to conviction generally do not trigger automatic exclusion rules. In fact, exclusion rules based on arrests that do not lead to conviction are potentially unconstitutional (`https://www.eeoc.gov/arrestandconviction`). Employment background checks submitted to the Virginia criminal records database do not show arrests that did not lead to a conviction (see VA Code §19.2-389).

[18]Prior arrests that do not lead to a conviction also influence these decisions (Kohler-Hausmann, 2018). But convictions are generally thought of as more serious indicators of prior crime.

[19]Two other channels by which noncarceral conviction could affect recidivism (relative to dismissal) are fines and probation conditions. However, the existing evidence suggests that these are not the primary drivers of recidivism. A small but growing literature shows that court fines and fees do not affect recidivism

Incarceration. Incarceration could affect recidivism through a variety of channels. It could reduce future criminal justice contact through incapacitation (Avi-Itzhak and Shinnar, 1973).[20] Incarceration could also decrease recidivism through specific deterrence (Zimring et al., 1973; Drago et al., 2009; Jordan et al., 2023). Under this theory, the negative experience of incarceration discourages future criminal behavior. Alternatively, incarceration could increase recidivism because the trauma, disruption, and loss of human capital involved with time behind bars erode a person's capacity to make a living on the legal labor market (Sykes, 1958; Blevins et al., 2010). Crime becomes more attractive as the outside option becomes less lucrative or less accessible. Prison might also expand the criminal network, thus making illicit activity more profitable (Hagan, 1993; Bayer et al., 2009; Stevenson, 2017).

2.3 Data sources, sample construction, and summary statistics

This subsection provides a brief overview of our data as well as sample and variable construction. A much more detailed description can be found in Appendix B. This subsection also presents summary statistics.

Data. Our primary data source for the judge IV analysis in Section 4 comes from Virginia's Circuit Courts. The data was scraped from a publicly accessible website. The Circuit Court data are available from 2000-2020 and cover all of Virginia except Alexandria and Fairfax counties. This data contains information on charges (type and date), on the defendant (gender, race, and FIPS code of residence), and on court proceedings for these cases (type, outcome, and judge). We also use it to construct defendants' recidivism outcomes. We then supplement this data with information on prior felony convictions from the Virginia Criminal Sentencing Commission (VCSC), which covers everyone convicted of a felony in Virginia during the period 1996-2020.

Sample and variable construction. We drop courts where cases are assigned to judges based on judge specialization or some other non-random schema. We also drop courts where there is substantial missing data as well as those with only one judge. Observations are at the case level. We say that a person is "incarcerated" if at least one charge resulted in a carceral sentence. We define a person to be "convicted" if at least one charge led to a sentence, but none resulted in a carceral sentence (i.e., noncarceral

(Pager et al., 2022; Finlay et al., 2023; Lieberman et al., 2023). Similarly, several large-scale RCTs have shown that probation and parole conditions do not affect recidivism (for a recent review, see Doleac, 2023).

[20] This doesn't mean that incarceration prevents crime, since crime is common in jails and prisons (Wolff et al., 2007). However, most within-prison crime is either not reported or is punished using an internal disciplinary system. Generally, only very serious crimes result in new charges.

conviction). Lastly, we say that a person was "dismissed" if all of their charges led to a dismissal (either by prosecution or judge) or an acquittal. Our main measure of recidivism is whether a person has a new felony charge in Circuit Court for an offense that allegedly occurred after the focal disposition date. Our main recidivism measure does not include probation revocations unless these are also accompanied by a new felony charge for a new crime. We calculate recidivism in the first year, years two to four, years five to seven, and the first seven years after a person's initial conviction. We also consider two alternative measures of recidivism: a new conviction resulting from felony Circuit Court charges, or a new carceral sentence resulting from felony Circuit Court charges.

Summary statistics. Table 1 provides summary statistics for those dismissed, with a noncarceral conviction, or incarcerated, respectively. Slightly more than half of the defendants in our sample received a carceral sentence. Among the non-incarcerated cases, about 66% are convicted. The dismissed, convicted, and incarcerated groups are similar in terms of zip code-level poverty but differ demographically. Cases ending in a noncarceral conviction are more likely to have female and non-Black defendants. Cases ending in incarceration are more likely to have defendants with prior felony convictions (22%) compared to the noncarceral conviction and dismissed samples (10% and 14%, respectively). Drug charges are the most common charges for all groups, followed by larceny, assault, and fraud.[21] Appendix Figure E.1 presents disposition types for four common offenses: drugs, larceny, assault, and fraud. While there is variation in the breakdown, all three disposition types exist within offense type.

3 Extending binary-treatment judge IV to multiple treatments

In this section, we discuss an extension of the "random judge" framework from the binary-treatment case to the case with three possible court outcomes. We outline assumptions under which widely-used 2SLS regressions recover margin-specific treatment effects, provide intuition for their restrictiveness by relating them to models of judge decision-making, and derive an expression that can be used to reason about the likely sign and direction of bias when some of the assumptions are not met.

[21]Fraud includes offenses like forgery, credit card fraud, or issuance of false checks.

3.1 Notation and common regression specifications

We consider a setting where cases can end in one of three mutually exclusive and collectively exhaustive alternatives: dismissal (d), noncarceral conviction (c), or incarceration (i). We denote treatment by $T \in \{d, c, i\}$. To simplify the discussion below, we further define $T_k = \mathbb{1}\{T = k\}$ as an indicator for the outcome of the case being $k \in \{d, c, i\}$ and $T_{\setminus d} = \mathbb{1}\{T \in \{c, i\}\}$ as an indicator that is equal to one if an individual is convicted or incarcerated (i.e., their case is not dismissed). Finally, we let Y be a measure of recidivism.

Both T_c and T_i are likely to be affected by unobserved factors that also influence recidivism, such as the strength of the evidence or the details of the offense or criminal record. Therefore, in a regression of Y on these court outcomes, there is concern about selection bias in the estimates of their respective coefficients. To account for this, a common approach is to use judge propensities for specific case outcomes as instruments. Let J denote the identity of the judge randomly assigned to a case. Define incarceration stringency $Z_i = E[T_i|J]$ and let $z_i^j = E[T_i|J = j]$, where $j \in \{1, ..., \mathcal{J}\}$ indexes the judges. Similarly define Z_k and z_k^j for $k \in \{c, d\}$.

Using the notation above and abstracting away from covariates, the following specification is commonly used to study the impacts of incarceration (see, for example, Mueller-Smith, 2015; Bhuller et al., 2020; Arteaga, 2021; Norris et al., 2021):

$$T_i = \alpha_0 + \alpha_1 Z_i + \alpha_2 Z_d + U \tag{1}$$

$$Y = \beta_0 + \beta_1 T_i + \beta_2 Z_d + V. \tag{2}$$

This 2SLS regression instruments incarceration with the assigned judge's incarceration stringency, and controls for dismissal stringency Z_d to prevent exclusion violations stemming from the judge's likelihood of conviction.[22]

Analogously, one approach to learning about the impacts of a noncarceral conviction is to run the following specification, in which we instrument for conviction but control for incarceration stringency:

$$T_c = \gamma_0 + \gamma_1 Z_c + \gamma_2 Z_i + U \tag{3}$$

$$Y = \delta_0 + \delta_1 T_c + \delta_2 Z_i + V. \tag{4}$$

[22]Another common specification uses a second stage with two endogenous treatments, instrumented with both stringencies. Under A1-A4, this specification produces the same estimand as (1)-(2). However, it builds in linearity assumptions that can be relaxed in our approach. See Appendix F.1. Alternatively, researchers may instrument a binary treatment indicator (e.g., for incarceration) with judge stringency in that same dimension, omitting controls for other dimensions of sentencing. Under the standard Imbens and Angrist (1994) LATE assumptions, this approach does not recover a well-defined causal effect of incarceration relative to a mix of counterfactuals when there are multiple treatments and the relevant stringencies are correlated, which is likely given that $Z_i = 1 - (Z_c + Z_d)$ – see Appendix F.2.

In the next subsection, we discuss conditions under which δ_1 has a causal and margin-specific interpretation – i.e., whether it can be interpreted as the impact of noncarceral conviction relative to dismissal for some well-defined subgroup of the population (and, analogously, under which β_1 has such an interpretation for the effect of incarceration vs conviction).

3.2 Judge IV assumptions in the multiple-treatment case

For simplicity, our discussion in this section is organized around the interpretation of δ_1 in specification (3)-(4), but an analogous discussion holds for the interpretation of β_1 in specification (1)-(2).

We define, for each individual, the potential case outcomes $T(z_c, z_i) \in \{d, c, i\}$, and the potential recidivism outcomes $Y(t, z_i, z_c),\ t \in \{d, c, i\}$. In analogy to our notation from the previous section, we further define $T_k(z_c, z_i) = \mathbb{1}\{T(z_c, z_i) = k\}$, for $k \in \{d, c, i\}$. Using this notation, we can state the standard IV assumptions of exclusion, random assignment, and relevance for the multiple-treatment case:

A1. Exclusion: $Y(t, z_i, z_c) = Y(t)\ \forall\ t, z_i, z_c$.

A2. Random assignment: $Y(t),\ T(z_c, z_i) \perp\!\!\!\perp Z_i, Z_c\ \forall\ t, z_i, z_c$.

A3. Relevance: $\gamma_1 \neq 0$ in equation (3).

We additionally make an assumption on the way that Z_i enters the regression (following Blandhol et al., 2022).

A4. Rich covariates: The linear projection of Z_c on Z_i is equal to $E[Z_c|Z_i]$.

We instrument for conviction using Z_c while controlling for Z_i rather than instrumenting for conviction and incarceration jointly in the same 2SLS regression. An advantage of our approach is that Z_i is a *control* and Assumption A4 can be relaxed by controlling for Z_i more flexibly (see Appendix C).

Throughout the paper, unless specified otherwise, we assume A1-A4 are satisfied. A1-A3 represent straightforward analogs to the standard Imbens and Angrist (1994) assumptions. Extending the monotonicity assumption to the multiple-treatment setting is less straightforward. In other applications, researchers have assumed that instruments induce compliers to take up a specific treatment, without inducing anyone to switch into other "non-focal" treatments. For example, Kline and Walters (2016) study the impact of enrolling in Head Start in a setting with two outside options, using randomly-assigned offers of enrollment as an instrument. The Head Start offer is assumed to not induce switches between the outside options. Similarly, Kirkeboen et al. (2016) study the returns to college majors and use offers of admission to specific majors as instruments. Their irrelevance condition states that access to a major does not induce switches between other choices (e.g., increased access to an economics major won't induce students to switch between history and mathematics). In a similar vein, Mountjoy (2022) assumes that reducing the distance to a two-year college (while holding distance to four-year college fixed) lowers its relative costs, while it does not induce switches between four-year college and not enrolling.

The Unordered Partial Monotonicity (UPM) assumption in Mountjoy (2022) formalizes the treatment-specific instruments assumption. In our notation, this assumption may be stated as:

A5. Unordered Partial Monotonicity (UPM($Z_c|Z_i$)):

For all z_c, z'_c, z_i with $z'_c > z_c$ and holding z_i fixed:

i $T_c(z'_c, z_i) \geq T_c(z_c, z_i)$

ii $T_i(z'_c, z_i) \leq T_i(z_c, z_i)$

iii $T_d(z'_c, z_i) \leq T_d(z_c, z_i)$.

Treatment specificity of an instrument for conviction, as formalized by UPM, imposes three restrictions on substitution patterns when Z_c increases and Z_i is held fixed. First, it guarantees that individuals only move into (and not out of) noncarceral conviction. Second, it guarantees that individuals only (weakly) move in one direction across any margin. Third, it rules out flows between dismissal and incarceration.[23] The UPM assumption thus incorporates a property similar to the "no defiers" assumption in the binary setting (Imbens and Angrist, 1994), but additionally rules out switches between incarceration and dismissal.

When using judge stringencies as instruments, the UPM assumption imposes stronger restrictions on substitution patterns than the restrictions in the three studies discussed above. In those examples, the instruments reduce costs or increase access to specific choices. In contrast, judge stringency instruments are the judge-specific probabilities of a case ending with a particular outcome. Indeed, the stringency instruments will add up to one ($z_d^j + z_c^j + z_i^j = 1$) since our case outcomes are mutually exclusive. As such, judge stringency instruments vary the net probabilities of taking up particular

[23]Note that UPM can hold when varying one instrument and holding the other fixed, while it does not hold when switching the roles of the instruments. We therefore use the notation UPM($Z_c|Z_i$) for the definition above and UPM($Z_i|Z_d$) when incarceration is the focal treatment.

treatments. By the same logic, if we condition on the judge stringency for one particular treatment, we do not fix the average costs of that treatment but its net probability of take-up.

This feature of judge instruments is important for understanding judge IV with multiple treatments. If we increase conviction stringency Z_c while holding Z_i fixed, we increase the net probability of conviction while holding the net probability of incarceration constant. Thus, if increasing Z_c results in a shift from $i \to c$, there must also be a compensating same-sized shift from $d \to i$ in order to keep the net probability of incarceration constant. However, UPM$(Z_c|Z_i)$ rules out flows from dismissal to incarceration. This implies there can be no flows from $i \to c$ because the net probability of incarceration Z_i is held fixed. Therefore, UPM implies that judge stringency instruments are not only *treatment-specific*, as in the examples described above, but also *margin-specific*: they induce complier flows across only one margin, e.g., dismissal to noncarceral conviction. In the multi-treatment judge IV setting, UPM therefore helps recover margin-specific treatment effects, but it is also a less plausible assumption than in many other multiple-treatment IV settings. In section 3.3 we illustrate the latter point by examining how UPM restricts models of judge decision-making.

Given that UPM may be a particularly strong assumption with judge stringency instruments, we next introduce a weaker monotonicity assumption, which we call conditional pairwise monotonicity (CPM).[24]

A6. Conditional pairwise monotonicity (CPM($Z_c|Z_i$)):

For case outcomes c, i, and d, for all z_c, z_c', z_i with $z_c' > z_c$ and holding z_i fixed:

i $T_c(z_c', z_i) \geq T_c(z_c, z_i)$ for all individuals

ii if $T_i(z_c', z_i) = T_d(z_c, z_i) = 1$ for any individual, then $T_i(z_c, z_i) = 1$ implies $T_d(z_c', z_i) = 0$ for all individuals

iii if $T_d(z_c', z_i) = T_i(z_c, z_i) = 1$ for any individual, then $T_d(z_c, z_i) = 1$ implies $T_i(z_c', z_i) = 0$ for all individuals.

CPM imposes two of the three restrictions imposed by UPM. It guarantees that, in response to increasing Z_c while holding Z_i fixed, individuals only move into (and not out of) $T = c$ and that individuals only (weakly) move in one direction across any

[24]Another way to relax the UPM assumption would be to extend the concept of average monotonicity (Frandsen et al., 2023) to the multiple-treatment setting. We present a definition of "average UPM" in Appendix C.4 and discuss intuition. Bhuller and Sigstad (2024) provide a more general way to extend average monotonicity with an arbitrary number of treatments. They provide conditions that are both sufficient and necessary for an estimand to have "proper weights."

margin.[25]

CPM does not rule out flows across margins that are not adjacent to noncarceral conviction. For example, an increase in Z_c holding Z_i constant can induce flow from $d \to c$ and $i \to c$, but also flows from $d \to i$. Throughout this paper, we assume CPM holds, and we discuss the implications when CPM holds but UPM does not.[26]

3.3 Connecting assumptions to models of judge decision-making

In this subsection, we provide economic intuition for the assumptions in the previous subsection, by discussing how they restrict models of judge decision-making. We consider three index-crossing models of judge decision-making based on canonical models of multinomial discrete choice – an ordered choice model, a sequential choice model, and an unordered choice model – and discuss how they relate to the legal and institutional practices of criminal proceedings.[27] All three models satisfy the CPM assumption, but only the ordered choice model satisfies the UPM assumption for both instruments. The sequential model illustrates that UPM may be satisfied for one of the instruments but not the other.

3.3.1 Ordered choice

First, we consider a straightforward extension to a trinary model from the binary threshold-crossing model. This extension is an ordered choice model with a single dimension of case-specific unobserved heterogeneity W. Each judge has their own thresholds for the values of W that would result in dismissal, noncarceral conviction, and incarceration:

$$\begin{aligned} T_d &= \mathbb{1}\{W < \pi_c(Z_d)\}, \\ T_c &= \mathbb{1}\{\pi_c(Z_d) \leq W < \pi_i(Z_i)\}, \\ T_i &= \mathbb{1}\{W \geq \pi(Z_i)\}, \end{aligned} \tag{5}$$

[25]Note that conditions (ii) and (iii) in A6 can be replaced with $T_d(z_c', z_i) \leq T_d(z_c, z_i)$ within our setting with stringency instruments, which makes CPM equivalent to (i) and (iii) from the UPM definition.

[26]While CPM is weaker than UPM, it is worth noting that it still implies restrictions on judge behavior that may not hold. For example, suppose a judge with a high incarceration propensity overall is more lenient on drug cases. Switching to this judge would increase incarceration for most people, but decrease it for drug offenders, thus violating CPM. Violations of "no defier" assumptions have received considerable attention in the literature (de Chaisemartin, 2017; Chan et al., 2022; Frandsen et al., 2023; Sigstad, 2024). Given that these issues are already well understood, we set them aside and focus on the novel issues that arise with judge stringency instruments and multiple treatments.

[27]Throughout this subsection we use "models of judge decision-making" as a shorthand; in practice, court outcomes reflect a combination of decisions by multiple actors, as we discussed in Section 2.

where the judge's conviction threshold $\pi_c(Z_d)$ is less than their incarceration threshold $\pi_i(Z_i)$ for all Z_d and Z_i. Panel (a) in Figure 1 visualizes, for two different judges, the regions of W under which each judge dismisses, convicts, and incarcerates. In this example, judge 1 has higher thresholds for both noncarceral conviction and for incarceration than judge 2.

In an ordered choice model, we can estimate margin-specific treatment effects for both the $T = c$ vs $T = d$ margin and the $T = i$ vs $T = c$ margin. To illustrate this, consider panel (b) of Figure 1, in which both judges have the same incarceration threshold, but judge 2 has a lower noncarceral conviction threshold, meaning that they convict more and dismiss less than judge 1. This figure demonstrates a key point: fixing Z_i and increasing Z_c will result in holding $\pi_i(Z_i)$ fixed and decreasing $\pi_c(Z_d)$. The only people who will switch treatment status are those who move from $d \to c$. When conditioning, the instruments are treatment-specific, since fixing Z_i and increasing Z_c will induce flows into only one choice ($T = c$) and not into any other treatment. Moreover, the instruments only move individuals across a single margin (from $d \to c$). Similarly, we can learn about the effect of incarceration vs noncarceral conviction using variation in Z_i and fixing Z_d. Thus, this choice model satisfies the unordered partial monotonicity assumption for both margins (i.e., UPM($Z_c|Z_i$) and UPM($Z_i|Z_d$) hold).

This model would be appropriate if all judges considered a single dimension of unobserved heterogeneity in their decision, and they agreed on how cases are ranked according to this dimension. The only ways in which judges can differ in their decision-making is by setting different thresholds for assigning cases to each of the outcomes. In practice, however, judges may take into account more than one measure of unobserved heterogeneity. In the remainder of this section, we consider models that allow for multiple dimensions of unobserved differences between defendants.

3.3.2 Sequential choice

Next we consider a sequential choice model in which the court process consists of two decisions: (1) a dismissal decision and, if not dismissed, (2) an incarceration decision. This reflects the two-step process of criminal cases: a trial to adjudicate guilt or innocence, followed by a sentencing hearing if the person is found guilty. The model allows judges to consider different, though potentially correlated, unobserved factors in each decision. For example, conviction decisions may depend on the strength of the evidence, which is not observed in our data, while incarceration decisions may depend on other aspects, such as the propensity to re-offend or severity of the crime, which are also not observed in our data.

We can write this as a threshold-crossing model:

$$
\begin{aligned}
T_d &= \mathbb{1}\{U_c < \pi_c(Z_d)\} \\
T_c &= \mathbb{1}\{U_c \geq \pi_c(Z_d), U_i < \pi_i(Z_i, Z_d)\} \\
T_i &= \mathbb{1}\{U_c \geq \pi_c(Z_d), U_i \geq \pi_i(Z_i, Z_d)\}.
\end{aligned}
$$

In this model, the first choice is between $T \in \backslash d$ (not dismissed) and $T = d$ and depends on the value of case-specific unobservable U_c relative to judge-specific threshold π_c. For cases that switch from dismissed to "not dismissed," there is then a second choice: noncarceral conviction or incarceration. This choice depends on the value of case-specific unobservable U_i, which can be correlated with U_c, relative to judge-specific π_i.[28] This model is consistent with only a subset of the information available to the judge being used in each of the two steps. It is also consistent with new information arriving at the incarceration stage, such as letters of support for the person convicted of the crime or victim impact statements.

Under the assumptions of the sequential model, it is possible to use 2SLS and the stringency instruments to recover margin-specific treatment effects between $T = i$ and $T = c$, but not between $T = c$ and $T = d$ or $T \neq d$ and $T = d$. Figure 2 illustrates this point. Panel (a) visualizes one judge's decision regions based on U_c and U_i. Panel (b) then compares two judges who have the same probability of dismissal, but where the second judge has a higher probability of incarceration. Here, variation in Z_i holding Z_d fixed induces only changes in court outcomes from $c \to i$ for a set of compliers.

In contrast, panel (c) compares two judges who have the same probability of incarceration (Z_i), but where judge 2 has a lower probability of dismissal (Z_d). Recall that Z_i is the *proportion* of cases that a judge incarcerates. In this figure, Z_i is represented by the fraction of people in the top-right section. For two judges to have the same incarceration stringency, both π_i and π_c must differ across these judges. This comparison then induces three sets of compliers, those moving from $d \to c$, those moving from $i \to c$, and those moving from $d \to i$. This example satisfies CPM since there is only a one-way flow across any given margin and no flows out of treatment. But the flows from $T = d$ to $T = i$ mean that the instrument is not treatment-specific, and UPM$(Z_c|Z_i)$ is not satisfied.

While the sequential model captures the two-step nature of the criminal proceeding, it may not be a good model if case outcomes are determined by a *joint consideration* of the two dimensions, as may be the case when plea bargaining occurs. We thus also consider a multinomial choice model, which similarly has two dimensions of unob-

[28] See Heckman et al. (2016) for details on identifying treatment effects in this type of sequential choice model, and Arteaga (2021) for a criminal court application studying the impacts of incarceration using a model similar to the sequential model described above.

served heterogeneity but allows for both unobservables to affect both conviction and incarceration.

3.3.3 Unordered multinomial choice

We now consider an unordered multinomial choice model, where outcomes can be thought of as being determined by maximizing over their "returns": $R_c \equiv V_c - \pi_c(Z_c, Z_i)$, $R_i \equiv V_i - \pi_i(Z_c, Z_i)$, and $R_d \equiv 0$.[29] Treatment depends on the judge's threshold for noncarceral conviction ($\pi_c(Z_c, Z_i)$), the judge's threshold for incarceration ($\pi_i(Z_c, Z_i)$), and two case-specific unobserved characteristics (V_c and V_i). Thus, case outcomes are modeled as being determined by a joint consideration across the two unobserved dimensions, which may better capture the intertwined decisions that are common in Virginia and other US jurisdictions due to plea bargaining. In a plea deal, a defendant typically agrees to plead guilty in exchange for a lower sentence, making conviction and sentencing determinations closely connected; unobserved determinants of the sentencing decision may affect the decision to plead guilty.

The unordered multinomial choice model can also be written as a threshold-crossing model:

$$\begin{aligned} T_d &= \mathbb{1}\{V_c < \pi_c(Z_c, Z_i), V_i < \pi_i(Z_c, Z_i)\} \\ T_c &= \mathbb{1}\{V_c \geq \pi_c(Z_c, Z_i), V_c - V_i \geq \pi_c(Z_c, Z_i) - \pi_i(Z_c, Z_i)\} \\ T_i &= \mathbb{1}\{V_i \geq \pi_i(Z_c, Z_i), V_i - V_c \geq \pi_i(Z_c, Z_i) - \pi_c(Z_c, Z_i)\}. \end{aligned} \tag{6}$$

In this model, the instruments are not treatment-specific. For example, the propensity of a judge to convict depends on both π_i and π_c, neither of which are directly observed. Panel (a) of Figure 3 visualizes the court outcomes and how they depend on judge thresholds and the two unobservables.

Under this model, 2SLS with stringency instruments does not recover margin-specific or treatment-specific treatment effects without further assumptions. To see this, consider panel (b) of Figure 3, which shows how treatment assignment changes when holding Z_i fixed and increasing Z_c. In this case, individuals shift from incarcerated to convicted and from dismissed to convicted but, in order to hold the probability of incarceration (Z_i) constant, individuals also need to shift from dismissed to incarcerated. This flow from dismissal to incarceration violates UPM and demonstrates that instruments neither move individuals into a single treatment nor across a single margin. Results are similar when holding Z_c (or Z_d) fixed and varying Z_i.

These observations illustrate how judge stringency instruments differ from those

[29] See, e.g., Heckman et al. (2006) for a discussion of treatment effects in a unordered multinomial choice model and Mountjoy (2022) for an application in the context of college choice.

in Kirkeboen et al. (2016), Kline and Walters (2016), and Mountjoy (2022). The difference stems from the fact that stringency instruments are generally not treatment-specific. The judge stringency for conviction, for example, does not correspond to π_c; it corresponds to the fraction of court cases in the conviction section of the graph. If we could directly shift π_c, then decreasing π_c holding π_i constant would result in flows into conviction from the other two treatments and no flows between incarceration and dismissal, as shown in panel (c) of Figure 3. Given that π_c and π_i are not observed, we instead can only shift or condition on Z_c and Z_i, resulting in variation that violates UPM and does not solely shift people into or out of a particular choice.

3.4 Asymptotic bias under different monotonocity assumptions

The prior subsection showed how UPM rules out some reasonable models of judge behavior, while the weaker CPM condition is not sufficient for 2SLS to recover margin-specific or treatment-specific effects. Here we derive what the Wald estimand recovers under CPM, which is satisfied by all three models. As in the prior section, we will consider the impacts of conviction vs dismissal and study the case where Z_c takes on two values and Z_i is fixed. Analogous results for the incarceration-conviction margin can be obtained by rearranging subscripts.

To begin, consider increasing conviction stringency from z_c to z'_c while holding incarceration stringency fixed at z_i. Let $\omega_{i\to c}$ represent the proportion of cases switching from $i \to c$ in response to the instrument shift. Similarly, allow $\omega_{d\to c}$ and $\omega_{c\to i}$ to represent the proportions of cases responding by switching across the other margins. Next, let $\Delta^{Y_c-Y_i}_{i\to c}$ represent the local average $Y_c - Y_i$ treatment effect for those who switch from $i \to c$ when the instrument shifts from z_c to z'_c, holding Z_i fixed. More generally, $\Delta^{Y_m-Y_n}_{k\to l}$ denotes the treatment effect of $T = m$ vs $T = n$ for cases induced to move from $k \to l$.[30]

Proposition 1 *Under A1-A4 and CPM, the Wald estimand of increasing conviction stringency Z_c from z_c to z'_c, while holding incarceration stringency fixed at $Z_i = z_i$, is*

[30] For simplicity, we suppress notation indicating instrument values; for example, we write $\omega_{d\to c}$ rather than $\omega_{d\to c}(z'_c, z_c|z_i)$ and $\Delta^{Y_m-Y_n}_{j\to k}$ rather than $\Delta(z'_c, z_c|z_i)^{Y_m-Y_n}_{j\to k}$.

given by:

$$\frac{E[Y(T(z_c', z_i)) - Y(T(z_c, z_i))]}{E[T_c(z_c', z_i) - T_c(z_c, z_i)]} =$$

$$\underbrace{\frac{\omega_{d\to c}\Delta_{d\to c}^{Y_c - Y_d} + \omega_{i\to c}\Delta_{d\to i}^{Y_c - Y_d}}{\omega_{d\to c} + \omega_{i\to c}}}_{\textit{Positively-weighted avg. of } Y_c - Y_d \textit{ treatment effects}} + \underbrace{\frac{\omega_{i\to c}}{\omega_{d\to c} + \omega_{i\to c}}\left[\Delta_{d\to i}^{Y_i - Y_c} - \Delta_{i\to c}^{Y_i - Y_c}\right]}_{\textit{Bias term}}. \quad (7)$$

Proof: See Appendix C.1.

Proposition 1 states that the Wald estimand can be decomposed into two terms. The first term is a weighted average of two LATEs for noncarceral conviction vs dismissal, corresponding to two different groups of compliers. The second term represents asymptotic bias relative to this weighted average. The bias term is the difference between the LATE for incarceration vs conviction for two equally-sized groups of compliers, weighted by the share of compliers moving from incarceration to noncarceral conviction. A direct consequence of Proposition 1 is that, when we replace the CPM assumption with the UPM assumption, the bias term in equation (7) is eliminated.

Corollary 1 *Under A1-A4 and UPM, the Wald estimand of increasing conviction stringency Z_c from z_c to z_c', while holding incarceration stringency fixed at $Z_i = z_i$, is given by:*

$$\begin{aligned}\frac{E[Y(T(z_c', z_i)) - Y(T(z_c, z_i))]}{E[T_c(z_c', z_i) - T_c(z_c, z_i)]} &= E[Y(T(z_c', z_i)) - Y(T(z_c, z_i)) \mid T_c(z_c', z_i) = T_d(z_c, z_i) = 1] \\ &= \Delta_{d\to c}^{Y_c - Y_d} \quad (8)\end{aligned}$$

To see this, note that the bias term is zero if $\omega_{i\to c}$ equals zero, which is the case if no compliers shift from incarceration to conviction. As discussed in Section 3.2, this is what UPM imposes when combined with judge stringency instruments. Thus, under UPM, the Wald estimand will be $\Delta_{d\to c}^{Y_c - Y_d}$, which is the LATE for noncarceral conviction vs dismissal for those shifted across that margin by the instrument.

Proposition 1 and Corollary 1 allow us to reason about conditions under which asymptotic bias will be quantitatively important for our 2SLS estimands. Under A1-A4 and UPM, the 2SLS specification in equations (3)-(4) yields a positively-weighted sum of unbiased Wald estimands.[31] If CPM holds but UPM does not, then the 2SLS estimates will recover a positively-weighted sum of the biased Wald estimands from equation (7) unless we impose additional assumptions. One possibility is to restrict

[31] Note that assumptions A1-A5 imply the assumptions needed in Blandhol et al. (2022) for 2SLS to recover causal estimands. In particular, A5 implies their "Ordered strong monotonicity" (OSM). See Appendix C.3 for details, and see Appendix C.5 for how to interpret the 2SLS estimand when there are additional control variables.

treatment effect heterogeneity. However, it is not necessary to assume treatment effect homogeneity for both margins, or even for all cases.

Treatment effect homogeneity assumptions under which the bias term is zero. As can be seen from equation (7), the bias term will be zero if $\Delta_{d\to i}^{Y_i-Y_c} - \Delta_{i\to c}^{Y_i-Y_c} = 0$. Thus, if the average treatment effects of incarceration vs conviction are the same for the $d \to i$ compliers and $i \to c$ compliers, the bias will be zero. For this result, we do not need the stronger assumption that treatment effects are homogeneous across all cases. Nor do we need to assume treatment effect homogeneity across the conviction-dismissal margin.[32] A special case is the one where the impact of incarceration vs conviction is zero for these two groups. This case is of specific interest in our context, because prior studies find long run null effects across this margin (see, e.g., Loeffler and Nagin, 2022; Garin et al., 2023). We return to this point in our discussion of our empirical results in Section 4.5.2.

Reasoning about sign and magnitude of the bias. Equation (7) also allows us to reason about the likely sign and magnitude of the bias when we are unwilling to make the homogeneity assumptions discussed above. We know that the bias is less than and proportional to $\Delta_{d\to i}^{Y_i-Y_c} - \Delta_{i\to c}^{Y_i-Y_c}$, i.e., the difference in the impact of incarceration (relative to noncarceral conviction) between those shifted from $d \to i$ and those shifted from $i \to c$. Thus the sign and the magnitude of the bias depend on the differential impact between these two groups. In Section 4.5, we discuss this point in more detail within the context of our institutional setting and draw on results from the existing literature.

4 Conviction, incarceration, and recidivism: 2SLS estimates

4.1 Regression specifications for estimation

Using leave-one-out estimates of judge stringency as our instruments, we consider the following 2SLS regression, which is common in the literature (stated here for noncarceral conviction; the specification for incarceration is analogous):

[32] Also note that homogeneous treatment effects still allow for selection on level (e.g., individuals more prone to recidivism can be more likely to be incarcerated), but not selection on gains.

$$T_c = \delta_0 + \delta_1 Z_c + \delta_2 Z_i + {\delta_3}' X + U \tag{9}$$

$$Y = \gamma_0 + \gamma_1 T_c + \gamma_2 Z_i + {\gamma_3}' X + V, \tag{10}$$

where Y is one of the measures of recidivism described in Section 2.3. The vector X includes court-by-year, court-by-month-of-year, and day-of-the-week fixed effects, as well as controls for offense type, race, gender, and a flag for prior felony convictions. For our main measure of judge stringency, we use the tri-yearly leave-one-out conviction and incarceration rates for the judge handling the case.[33] We run these 2SLS regressions on the sample described in Section 2.3.[34]

In Appendix D, we discuss at length how assumptions A1-A3 are supported by features of the institutional environment and provide empirical evidence, based on a standard battery of tests, that these assumptions likely hold in our setting. For both the conviction and incarceration regressions, we have a strong first stage with F-statistics of 165 and 288 respectively (Table 2), suggesting that relevance holds in our setting. Panels A and B of Figure 4 plot the variation in residualized judge conviction and incarceration stringency, showing that there is substantial variation in each. Panel C of Figure 4 provides a scatter plot of residualized conviction and incarceration stringency and shows that there is also substantial variation in Z_c conditional on Z_i, and vice versa. For balance, Table 3 shows that, while case characteristics are strong predictors of conviction and incarceration, they largely do not predict judge stringencies. For the few covariates with statistically significant loadings, the predicted difference in stringency tends to be very small (0.016 to 0.036 standard deviations of the residualized stringency measure, see Appendix Table D.1). In addition, Appendix Tables D.2 and D.3 show that our main results are broadly similar when systematically dropping certain case types, such as assault. For the exclusion restriction, we discuss potential violations and provide tests suggesting that these would not have qualitative impacts on our results. For instance, we show in Figures E.3-E.6 that estimates remain largely unchanged when including sentence-length stringencies as additional controls. Finally, we provide a test of the "no defiers" assumption that is part of both CPM and UPM, with Tables D.5 and D.6 reporting split-sample monotonicity tests and finding the same sign for the first stage across various splits of the data. We postpone the discussion and implementation of an additional test of the UPM assumption to Section 4.5.

[33] We choose a tri-yearly specification to allow for a large number of cases per judge, without requiring that judges behave identically for their entire tenure. We exclude cases assigned to judges who see fewer than 100 cases in the 3-year period.

[34] As discussed in Section 3, under A1-A5, these regression estimates can be interpreted as causal and margin-specific. See Appendix C for additional discussion of what 2SLS identifies when including controls based on Blandhol et al. (2022), and details on the assumption of sufficiently rich controls.

4.2 Noncarceral conviction

Table 4 presents 2SLS estimates of the model in equations (9)-(10). When given a causal and margin-specific interpretation, these estimates represent the impact of noncarceral conviction on recidivism relative to dismissal for those near the margin.

We consider three measures of future criminal justice contact: new felony charges in Circuit Court, a new conviction resulting from felony Circuit Court charges, or a new carceral sentence resulting from felony Circuit Court charges. We use various time windows to measure recidivism, all measured from the time of disposition: year 1, years 2-4, years 5-7, and cumulatively for the first 7 years. For each of these outcomes, we present OLS and 2SLS regressions.[35]

As discussed in Section 2.2, noncarceral conviction (instead of a dismissal) could increase or decrease recidivism through a number of channels, and the sign of the net effect is not clear a priori. If given a causal and margin-specific interpretation, our 2SLS estimates suggest that noncarceral conviction increases future criminal justice contact relative to dismissal. The estimates for future charges within the first year after conviction are large: around 10.5 percentage points (95% CI, 0.02 to 0.20), which is a 66% increase relative to the control complier mean. The impact on cumulative recidivism 1-7 years later is also statistically significant, with an estimate of 23 percentage points (95% CI, 0.04 to 0.42), a 47% increase relative to the control complier mean. The effects for years 1-7 are approximately twice as large as the effects in year 1, with positive but statistically insignificant effects in years 2-4 and 5-7. Results are similar for the other measures of recidivism we consider.

These point estimates are similar in magnitude to estimates found in the related literature. For instance, Mueller-Smith and Schnepel (2021) find that diversion cuts reoffending rates in half, and Agan et al. (2023b) find that nonprosecution reduces the likelihood of a new criminal complaint by 53%. Mueller-Smith et al. (2023) find that adult conviction increases the total number of future felony charges by roughly 75%. While our point estimates could be considered fairly large, the confidence intervals leave room for a wide range of values, as is typical for judge IV research designs.

Our 2SLS estimates are similarly signed but substantially larger than the OLS estimates. However, the OLS estimates likely suffer from omitted variable bias. One important omitted variable is the strength of the evidence, which often consists primarily of witness testimony. Graef et al. (2023) show that witness appearance in court is by far the most predictive factor in whether the defendant will be convicted. Thus, the sign of the bias in the OLS estimates depends in part on the relationship

[35] Appendix Table E.1 presents reduced-form estimates. The OLS estimate is from a regression of recidivism on a conviction indicator that is one if the individual is convicted or convicted and incarcerated, and controls for an incarceration indicator.

between witness appearance and the defendant's risk of recidivism. These could be positively correlated if, e.g., witnesses are more invested in securing punishment for high-recidivism defendants. Or they could be negatively correlated if, e.g., witnesses are scared of testifying against high-recidivism defendants. The fact that victims and bystander witnesses often come from the same socioeconomic groups as defendants also suggests a negative correlation. The same factors that give someone a high-recidivism potential – for example, poverty or social marginalization – may also make it harder for the witnesses to take time off work for a court date, or make them less willing to cooperate with a system they distrust. If so, OLS estimates will be downward biased.[36] Alternatively, IV compliers may be more impacted by conviction than the average defendant. In Appendix Table E.2, we show that the racial composition of the complier group is similar to the overall sample, but that on average this group is less likely to be in court for violent offenses and is less likely to have a prior conviction. Our OLS estimates for noncarceral conviction are somewhat larger when reweighting with complier weights, while the estimates for incarceration do not notably change (see Appendix Table E.3).

We next explore whether our results are coming from an increase in criminal behavior or an escalation in subsequent criminal justice responses ("ratcheting up") – mechanisms we discussed in Section 2.2. While we cannot answer this question definitively, we consider two tests to help provide suggestive evidence.

First, if conviction makes it harder to find employment due to the mark of a felony record, we might expect to see a more pronounced increase in income-generating crime. We test for this in Appendix Table E.4 and find similar point estimates across income-generating and non-income-generating crime; the confidence intervals are too large to draw a firm conclusion.[37]

Second, if the ratcheting up effect is operative, conviction may have a larger effect on the more downstream measures of future criminal justice contact, such as future conviction or incarceration. The logic here is that if a felony conviction increases the likelihood of a negative outcome at each discretionary stage, the negative impact of a conviction will accumulate. Downstream outcomes, like incarceration, will be impacted more than upstream outcomes, like the charging decision. Comparing the three measures of recidivism in Table 4, the point estimates are larger relative to the control complier means for outcomes with more discretionary decisions.

[36]Witness cooperation is only one potential omitted variable. There are many others that could also bias the OLS upwards or downwards, depending on the correlation structure. For instance, if people with a skillful lawyer are both less likely to be convicted and less likely to recidivate, our OLS estimates would be upward biased; if people with substance abuse or untreated mental health concerns are less likely to be convicted and more likely to recidivate, the OLS estimates would be downward biased.

[37]Likewise, there are no consistent differential patterns for drug vs. non-drug crimes, as shown in Appendix Table E.5.

While we cannot conclusively say whether increased recidivism is driven primarily by increased criminal behavior or a ratcheting up effect, both mechanisms imply that felony conviction can trap a person in the revolving door of criminal justice, increasing not just future charges and convictions, but also future incarceration.

4.3 Incarceration

Table 5 presents 2SLS estimates of the model in equations (9)-(10), but instrumenting for incarceration with incarceration stringency and controlling for dismissal stringency. If given a causal and margin-specific interpretation, these estimates represent the impacts of incarceration relative to noncarceral conviction for those near the margin.

We find that incarceration causes a decline in recidivism in the first year after sentencing. Our 2SLS estimates suggest a 10 percentage point reduction in future charges in the first year (95% CI, -0.15 to -0.04). This reduction is likely due, at least partially, to incapacitation. While people are incarcerated, new crimes are usually addressed with internal sanctions and are unlikely to result in new felony charges. However, we find no evidence that incarceration affects future criminal justice interactions beyond the impact in the first year. The 2-4 year and 5-7 year estimates are small and statistically insignificant. The cumulative estimate across all seven years implies a seven percentage point reduction in new felony charges (95% CI, -0.19 to 0.05). We can reject increases in recidivism larger than 2.7 percentage points at the .05 level. Results are similar for future convictions and future incarceration.

Our qualitative conclusions are further strengthened by the fact that we find very similar results using another research design within the same institutional setting. We leverage the fact that judges' sentencing decisions are influenced by sentence guidelines. The guidelines-recommended sentence is calculated using a scoring system in which various characteristics of the offense and criminal record are assigned points which are then summed to create the sentence guidelines score. Exploiting two different discontinuities in the sentence guidelines recommendations within a regression discontinuity design framework, we estimate the effects of incarceration on the intensive margin (sentence length) and on the extensive margin (short jail sentences vs probation). As when exploiting quasi-random assignment of cases to judges, we find that incarceration leads to short-term decreases in criminal justice contact. We find no evidence of longer-term impacts of exposure to incarceration. We refer the reader to Appendix H for details on the empirical approach and findings.

We acknowledge some limitations to our analysis of incarceration. First, incarceration may affect other dimensions of well-being besides recidivism, or affect outcomes among subgroups that we are underpowered to detect (Aizer and Doyle, 2015; Mueller-Smith, 2015; Jordan et al., 2023). Second, our research design does not allow us to isolate the effects of long carceral sentences (e.g., five or ten years) vs noncarceral conviction. A higher "dosage" of incarceration may have more impact. Third, some people with noncarceral convictions could have been incarcerated pretrial and thus may have already experienced some negative effects of incarceration, reducing the difference between these groups in terms of their carceral exposure.

Similarly, some people who receive noncarceral conviction become incarcerated in the future, both because of new criminal convictions, as we showed in Section 4.2, or because of technical violations. This will further reduce the differences in carceral exposure between the incarcerated group and those with noncarceral conviction. However, our evidence suggests that there remains a substantial difference in exposure to incarceration across these two groups. Appendix Figure E.2 shows how much "incarceration catch-up" occurs for those who receive noncarceral sentences compared to those who receive carceral sentences, both for new crimes and for technical violations resulting in probation revocation. These results suggest that while there is some catch-up, more than 50% of those receiving a noncarceral sentence are never incarcerated over the next seven years.

Overall, the results from Sections 4.2 and 4.3 imply that incarceration's influence on the revolving door is limited, and noncarceral conviction may hold greater importance. Our findings on the effects of incarceration align with the conclusion drawn in a recent literature review that most of the papers that find incarceration to be criminogenic are looking at pretrial detention, rather than post-sentencing incarceration (Loeffler and Nagin, 2022). Since pretrial detention also increases the probability of conviction (Gupta et al., 2016; Leslie and Pope, 2017; Dobbie et al., 2018), these papers are effectively estimating the joint effect of conviction and incarceration. In contrast, most papers evaluating the impact of post-conviction incarceration do not find evidence of effects lasting beyond the incapacitation period. Incarceration may be a traumatic experience, but, in line with our findings, most studies find no evidence that it is an important contributor to the revolving door.

4.4 Robustness and subgroup analyses

In this subsection, we provide a brief overview of robustness checks that are discussed in detail in Appendix E.1. Our results are robust to our choice of sample restrictions and controls, as shown in Appendix Figures E.3-E.6. In particular, our results are similar when we drop specific crime types, for example drug cases, for which diversion is more likely to happen than for other offenses. Appendix Figures E.3-E.6 also show that our estimates and standard errors remain similar when we more flexibly control for non-focal stringency.[38] Appendix Table E.9 shows that our results are robust to varying our definition of recidivism, and considering counts of new offenses and charges. Appendix F.3 shows that our results are robust to correcting for measurement error in stringency using Empirical Bayes methods. Additionally, Appendix Figure E.7 demonstrates no differential mobility out of Virginia based on incarceration outcomes.[39]

To examine effect heterogeneity, we first break out our results based on whether a person has a prior felony conviction or not (Appendix Table E.6), since avoiding a first felony conviction might play an especially pivotal role in people's future trajectories. We find that people without a recent felony conviction have large and sustained increases in recidivism as a result of a felony conviction. Yet, we cannot reject that these estimates are equal to estimates for those with a recent felony conviction, for whom estimates are imprecise—likely because they make up only 20% of the sample. Sample size limitations again preclude clear inference about heterogeneity in the impacts of incarceration across those with and without a recent felony, although point estimates are similar for the two groups.[40]

We additionally explore heterogeneity across race and zip code income level. These results are also described in more detail in Appendix E.1. We find qualitatively similar patterns across Black and non-Black defendants. We find suggestive evidence that the impacts of noncarceral conviction are larger for people living in zip codes with above-median poverty rates. This might be due to felony convictions having greater consequences for poorer individuals, perhaps because such convictions block access to housing or other social services.

4.5 Testing for and characterizing bias in the 2SLS results

In Section 3.4, we showed that the 2SLS estimates may be asymptotically biased if the UPM assumption doesn't hold. In this subsection, we describe and implement an empirical test for this assumption. We then use theory and external evidence to discuss the likely magnitude and direction of the bias in our context.

4.5.1 Testing the UPM assumption

The UPM assumption has testable implications. If instrumental variation is only causing flows between two treatments, there should be no movement in or out of the third

[38] See also Table C.2, which provides further robustness to the choice of controls.

[39] We are unable to study differential mobility out of Virginia due to conviction, as less information about defendants is collected for cases ending in dismissal, prohibiting linkage to data on out-of-state moves.

[40] We define our prior felony indicator as a prior felony in the last five years. When considering the heterogeneous effects of incarceration, Jordan et al. (2023) are able to better isolate first felony convictions as they observe age for everyone in their sample, which allows them to construct first felonies using age restrictions. Our data does not include age.

treatment. In our setting, this implies:

(1) Under UPM($Z_c|Z_i$), the observable characteristics of those with $T = i$ should not change when holding Z_i constant and varying Z_c.

(2) Under UPM($Z_i|Z_d$), the observable characteristics of those with $T = d$ should not change when holding Z_d constant and varying Z_i.

To build intuition for the first testable implication, consider those incarcerated in the ordered model, which we discussed in Section 3.3. When holding incarceration stringency fixed, varying conviction stringency will move people between dismissal and conviction, but will not move people in or out of incarceration. This implies that the observed characteristics of incarcerated individuals should not change, and motivates the first testable implication above. If the characteristics of incarcerated individuals do change, then there must be flows in and out of incarceration, which implies that the instrument is moving people across more than one margin. More generally, this would imply that UPM($Z_c|Z_i$) is violated, as the UPM assumption plus stringency instruments (and the other IV assumptions) ensures compliers move across only one margin. A similar argument holds for our proposed testable implication of whether UPM($Z_i|Z_d$) holds.

Importantly, these conditions allow us to test across models of judge decision-making introduced in Section 3.3. In particular, (1) and (2) above must hold for the ordered model, and (2) must hold for the sequential model.

We implement our test using predicted recidivism: an index constructed by regressing recidivism on individual and case characteristics.[41] We test implication (1) by regressing predicted recidivism on our noncarceral conviction instrument, restricting the sample to those incarcerated and controlling for the incarceration instrument and court-by-time fixed effects. Similarly, we test implication (2) by regressing predicted recidivism on the incarceration instrument, restricting to the dismissed sample and controlling for the dismissal instrument and court-by-time fixed effects. Results are shown in Table 6, where Panel A presents tests for (1) and Panel B tests for (2).[42] Appendix Table E.10 shows results for both tests using a variety of defendant characteristics (criminal record, offense and demographics) instead of predicted recidivism.

Using the predicted recidivism index, we reject UPM($Z_c|Z_i$) and UPM($Z_i|Z_d$), which also means we reject both the ordered and sequential models. For (1), we

[41]Predicted recidivism variables are created by regressing recidivism post-release if incarcerated, or post-conviction/dismissal otherwise, on offense type, socio-demographic controls, and month, court, and day-of-the-week fixed effects. Using these regressions, we construct measures of predicted recidivism within one year, two to four years, five to seven years, and within seven years after case disposition.

[42]When implementing this test, we are maintaining other assumptions we make throughout the paper, such as the assumption that judge stringencies do not idiosyncratically depend on defendant characteristics and CPM. Results are similar when including flexible controls for the other stringency measure.

find that predicted recidivism for the incarcerated group increases with the judge's conviction propensity, holding incarceration propensity constant. For (2) we find that the predicted recidivism for the dismissed group decreases with the judge's incarceration propensity, holding fixed the dismissal propensity. These results suggest the UPM assumption does not hold exactly in our setting, and so our 2SLS estimates are potentially biased.

4.5.2 Sign and magnitude of asymptotic bias

Proposition 1 implies that when UPM does not hold (but A1-A4 and CPM do) 2SLS estimands will be positively-weighted averages of the Wald estimands in equation (7). In this section, we demonstrate how the expression in equation (7) can be combined with theory and external evidence to reason about the direction and quantitative importance of bias in 2SLS estimands. We consider each margin of interest separately. Throughout this discussion, we will assume that CPM holds, as it does in each of the three judge decision-making models we considered. We also assume A1-A4 from Section 3 hold.

Impact of noncarceral conviction vs dismissal. For simplicity, we discuss the bias term in the context of the special case where two judges have the same incarceration rate but differing rates of noncarceral conviction – as we also did when deriving equation (7).

Equation (7) shows that the bias term in the Wald estimand is less than but proportional to $\Delta_{d\to i}^{Y_i - Y_c} - \Delta_{i\to c}^{Y_i - Y_c}$, which is the difference in the impact of incarceration vs noncarceral conviction between those near the incarceration-dismissal margin and those near the incarceration-conviction margin. Hence, we can reason about the likely sign and magnitude of the bias based on conjectures and evidence that inform how incarceration vs conviction may differentially impact recidivism for these two groups. We separately consider the long- and short-run effects – where "long run" corresponds loosely to the post-incapacitation period.

Table 6 shows that the average predicted recidivism rate of the incarcerated group increases in response to increasing Z_c while controlling for Z_i (i.e., holding the net probability of incarceration constant). This implies that those shifting into incarceration from dismissal have a higher predicted recidivism rate than those shifting out of it into conviction.[43] It's reasonable to think that, in the short run, incarceration affects recidivism primarily through incapacitation (for both groups). If so, shifting

[43]This empirical finding is consistent with a scenario where the individuals on the incarceration-dismissal margin are those whose evidence is borderline but the case is serious enough to guarantee incarceration upon conviction, while those on the incarceration-conviction margin have sufficient evidence against them but marginal case severity.

prison beds towards those at a higher risk of recidivism will reduce recidivism, and $\Delta_{d\to i}^{Y_i-Y_c} < \Delta_{i\to c}^{Y_i-Y_c}$. If this is the case, the bias term in equation 5 would be negative and our short-run estimates would underestimate the increase in recidivism caused by conviction. However, the magnitude of the composition change shown in Table 6 is relatively small: a ten percentage point increase in noncarceral conviction stringency increases one-year predicted recidivism among the incarcerated group by 0.1 percentage points. This suggests that either the proportion of $i \to c$ compliers is small, or the two groups have similar observable characteristics and therefore potentially similar treatment effects. Both imply that the magnitude of the bias is likely small.

Turning to the longer run, if incarceration *only* has incapacitation effects, we would expect the impact of incarceration vs conviction to be zero after the incapacitation period. However, incarceration could affect recidivism through channels other than incapacitation, which could produce upward bias. For example, prison may be a stronger deterrent after release for people with fewer priors, as in Jordan et al. (2023). Since those with fewer priors typically have lower predicted recidivism, they are over-represented in the group at the incarceration-conviction margin, relative to those at the incarceration-dismissal margin.[44] Then, $\Delta_{d\to i}^{Y_i-Y_c} > \Delta_{i\to c}^{Y_i-Y_c}$ and the bias term would be positive. However, we think this type of upward bias is unlikely in our setting for two reasons. First, we find no evidence of differential treatment effects of incarceration by prior conviction status (see Panel B of Appendix Table E.6), though these estimates are imprecise. Second, multiple pieces of evidence suggest that longer-term effects of incarceration vs conviction on recidivism are negligible. In our setting and using the same data set but a different research design, the RD evidence we present in Appendix H shows that incarceration reduces recidivism only in the short run (likely due to incapacitation) for those on the margin of conviction and incarceration. In other settings, the majority of studies on the impact of incarceration finds similarly that the impact of incarceration on recidivism is negligible (Loeffler and Nagin, 2022).

Overall, the arguments above suggest that a violation of UPM would lead our 2SLS estimand of the effects of noncarceral conviction to have a small negative bias in the short run and negligible bias in the long run. Hence, it is unlikely that our qualitative conclusions about the impact of noncarceral conviction vs dismissal would be overturned as a result of a violation of the UPM assumption.

Impact of incarceration vs noncarceral conviction. Here, we discuss the bias in the context of the simple case where two judges have the same noncarceral conviction rate but differing rates of incarceration. Using a similar derivation as in the proof

[44] Indeed, when we run test (1) using prior convictions instead of predicted recidivism, we see that those shifting from $c \to i$ have a lower prior conviction rate than those shifting from $d \to i$ (Appendix Table E.10).

of Proposition 1, we know that the bias term for the impact of incarceration will be smaller than but proportional to $\Delta_{d\to i}^{Y_c-Y_d} - \Delta_{c\to d}^{Y_c-Y_d}$, and zero if there are no $d \to i$ compliers. As previously discussed, we expect cases on incarceration-dismissal margin to be high-severity, meaning that the charges are serious and/or the criminal record is long. Meanwhile, those on the conviction-dismissal margin are expected to have lower case severity, with less serious charges and a limited criminal record.[45]

Thus, in order to evaluate the bias on incarceration vs conviction, we need to know whether the mark of a felony conviction (vs dismissal) will affect recidivism more for high-severity cases than for low-severity cases. We may expect a felony conviction to increase recidivism more for low-severity cases through two channels. First, low-severity cases are less likely to already have a felony on their criminal record, and the impact of the first felony conviction is likely to be greater than future ones. Second, people with low-severity cases might have greater labor market attachment prior to conviction, and thus more to lose. If either channel is present, we would expect the bias term to be negatively signed. As a result, our 2SLS estimates would then underestimate incarceration's impact on recidivism. However, if our prediction is correct – that the marginal impact of conviction is greater for low-severity cases – the difference is not large enough for us to detect. We find no discernible difference in the impact of conviction vs dismissal across crime types or priors (Appendix Tables E.4 - E.6). In addition, the compositional changes shown in Table 6 and Appendix Table E.10 are relatively small. If the compositional shifts are minimal, then either the proportion of $d \to i$ compliers is small, or the $c \to d$ and $d \to i$ compliers have similar observable characteristics and, therefore, potentially similar treatment effects. The bias term on the incarceration effect is therefore likely to be small as well.

While we argue that the 2SLS bias is likely to be small in our setting, we present an alternative identification approach in the section below, which yields similar results. Beyond our context, this approach may also be useful in other applications where bias may be larger.

5 An alternative approach to identification and estimation of margin-specific treatment effects

In the prior section, we found that our empirical test rejects the UPM assumption, ruling out the ordered and sequential models and implying that our 2SLS estimates

[45]This is consistent with the results of our empirical test in Section 4.5.1, which suggests that those who flow into dismissal from noncarceral conviction have lower predicted recidivism than those who flow out of dismissal into incarceration.

will be asymptotically biased. Although we believe the bias resulting from this violation of UPM is likely small in our setting, it is worth considering alternative approaches based on assumptions that are not rejected by our test. In this section we therefore present a method for estimating margin-specific treatment effects under the unordered multinomial model, which we discussed in Section 3.3.3.

The method builds on Mountjoy's 2022 approach for identifying margin-specific treatment effects in unordered choice settings. Because this approach requires treatment-specific instruments, we begin by constructing such instruments from the panel of judge decisions in our data.

5.1 Recovering treatment-specific instruments from judge stringencies

Mountjoy (2022) studies enrollment in two-year and four-year college, modelling this decision using the unordered multinomial choice model discussed in Section 3.3.3, and using distances to the nearest two-year and four-year colleges as instruments. These distance instruments are plausibly treatment-specific and shift the cost associated with attending either two-year or four-year college. Varying one distance instrument while holding the other fixed is equivalent to exogenously shifting one of the latent thresholds in the decision model while holding the other fixed. In our notation, treatment-specificity of instruments $\tilde{Z}_c$ and $\tilde{Z}_i$ would imply that π_c (the latent threshold for noncarceral conviction) is a function of only $\tilde{Z}_c$, and not $\tilde{Z}_i$, and that π_i (the latent threshold for incarceration) is a function of only $\tilde{Z}_i$.

Even with such treatment-specific instruments, 2SLS estimands are difficult to interpret, as they are weighted averages of treatment effects that correspond to different margins, as visualized in Panel (c) of Figure 3. For example, in our context, shifting from $\tilde{z}_c$ to $\tilde{z}'_c$ while holding $\tilde{Z}_i$ fixed would yield a weighted average of the LATE for those switching from $i \to c$ and the LATE for those switching from $d \to c$. The central objective of Mountjoy (2022) is to decompose the 2SLS estimand, obtained using a treatment-specific instrument, into two margin-specific effects.

To apply Mountjoy's (2022) method, we first conduct an intermediate step of inverting the choice shares (judge stringencies), which we observe for each judge, to recover thresholds (π_c and π_i). These thresholds are treatment-specific instruments.

For each judge, we observe the shares of cases ending in $T = d$, $T = c$, and $T = i$, where individual cases are randomly assigned to each judge. Using the shares, we aim to recover the unkown judge-specific thresholds. Rewriting equation (6), we have:

$$\begin{aligned} R_c &= V_c - \pi_c(Z_c, Z_i) \\ R_i &= V_i - \pi_i(Z_c, Z_i) \\ R_d &= 0, \end{aligned} \tag{11}$$

where we have normalized the return of $T = d$ to zero. This setup has similarities to models in industrial organization where shares are observed for different markets.[46] We leverage results from the IO literature and adapt them to our context of judge decision-making. Berry, Gandhi and Haile (2013) show that the inversion between shares and thresholds exists under weak assumptions,[47] and Berry and Haile (2022) show that judge-specific thresholds can be identified without invoking identification at infinity arguments.[48]

While these papers show that the π's are identified under relatively weak conditions, we make additional assumptions for tractability in estimation and show that results are broadly similar under a few different assumptions. Our main specification assumes the shocks (η and ϵ in the equation below) follow a standard logistic distribution plus a random effect with a correlated multivariate normal distribution. We can then write the returns as

$$\begin{aligned} R_{ncj} =& \beta_c - \pi_c^j + \gamma_c' X_n + \eta_{nc} + \epsilon_{nc}, \\ R_{nij} =& \beta_i - \pi_i^j + \gamma_i' X_n + \eta_{ni} + \epsilon_{ni}, \end{aligned}$$

where n represents the case, c and i indicate conviction or incarceration, j the judge, X_n are characteristics about the defendant or case, and R_{ncj} and R_{nij} represent the returns

[46]Unlike most applications in the industrial organization literature, our setting has quasi-random assignment of cases to judges, implying that the joint distribution of (V_c, V_i) is not judge-specific, and therefore $\pi_c(Z_c, Z_i)$ and $\pi_i(Z_c, Z_i)$ are independent of V_c and V_i.

[47]They assume the structural choice probability function can be written with a nonparametric index where judges' latent preferences enter linearly into the index. Then the key assumption is that a "connected substitutes" condition holds. In a multinomial choice setting, this condition implies that the probability of choosing j is strictly increasing in the index, which is an input into the structural choice probability function. In a linear-in-parameters unordered choice model, this is satisfied if the support of the additive errors (i.e., the Vs) is $\mathbb{R}^K$, where K is the number of choices.

[48]This proof assumes an index structure on the structural choice probability function where judges' latent preferences enter linearly into the index. Using this setup, the paper shows how the latent judge preferences π^j can be identified using a combination of variation in latent preferences across judges and variation in case characteristics within each judge. In particular, identification requires three continuous covariates whose loadings do not vary across judges. The proof does not assume the distribution of error terms is independent or identically distributed. Similarly, beyond the assumption on the index function, linearity is not required. Kamat et al. (2024) provide an alternative approach that uses the sequential model and does not require covariates, but recovers bounds rather than point estimates.

to a specific outcome for a specific case assigned to judge j.[49] Here we assume $f(\epsilon_{nc}, \epsilon_{ni})$ has a standard logistic distribution and $g(\eta_{nc}, \eta_{ni}) \sim N(0, \Sigma)$. We estimate the model by judicial circuit and 3-year bin, which further allows the model parameters to differ across circuits and over time. Importantly, the random effects allow for correlation between V_c and V_i and for V_c and V_i to have different variances.[50]

5.2 Margin-specific effects in the unordered model

We refer to these newly constructed instruments—the estimated judge-specific thresholds—as $\tilde{Z}_c$ and $\tilde{Z}_i$, to distinguish them from the stringency instruments Z_c and Z_i. With these treatment-specific instruments in hand, we closely follow Mountjoy (2022) for estimating the impacts on the two margins discussed above. This method relies on assumptions A1-A4, defined for $\tilde{z}_c$ and $\tilde{z}_i$, plus one additional assumption: "comparable compliers" (CC). This assumption requires that the $i \to c$ compliers from decreasing $\tilde{z}_i$ have the same potential outcome when convicted as $i \to c$ compliers from increasing $\tilde{z}_c$ at their limits (see Appendix G for a formal definition). Under this set of assumptions, Mountjoy (2022) shows how to identify and estimate $E[Y(c) - Y(d) \mid d \to c$ complier w.r.t $(\tilde{z}_c, \tilde{z}_i) \to (\tilde{z}'_c, \tilde{z}_i)]$ and $E[Y(i) - Y(c) \mid i \to c$ complier w.r.t $(\tilde{z}_c, \tilde{z}_i) \to (\tilde{z}'_c, \tilde{z}_i)]$. We follow Mountjoy (2022) in our approach to estimation and provide additional details in Appendix G.

While we do not invoke the UPM assumption in this section, we introduce additional assumptions in both the construction of treatment-specific instruments and in applying Mountjoy (2022).[51] Therefore, the assumptions we consider in this section are not necessarily weaker or stronger than those supporting a causal interpretation of the 2SLS estimates.

5.3 Results

Table 7 presents the results of this alternative approach. These results assume a mixed-logit structure with a multivariate normal random effect whose variance and correlation

[49]Note that, while we make (flexible) parametric assumptions regarding the joint distribution of V_c and V_i for estimation, we do not make assumptions regarding the relationship between the errors in the choice model and the outcome equations. An alternative approach would be to directly model the joint distribution of error terms in the choice equation and outcomes, e.g., using a latent factor structure (Heckman et al., 2018).

[50]In Appendix G, we include additional results under two alternative assumptions: (1) that V_c and V_i follow standard logistic distributions and (2) that Σ is a diagonal matrix. Both are less flexible but easier to implement. For (1), the thresholds are simply $\pi_c(z_c, z_i) = log(z_c) - log(1 - z_c - z_i)$ and $\pi_i(z_c, z_i) = log(z_i) - log(1 - z_c - z_i)$.

[51]For identification, we assume the unordered model, "comparable compliers," and the existence of additive covariates whose loadings do not vary across judges. For estimation, we additionally make distributional assumptions about the error terms.

are allowed to vary by judicial circuit and year.[52] Panel A reports estimates for the noncarceral conviction vs dismissal margin. The point estimates are qualitatively similar to the 2SLS estimates reported in Section 4. Compared to the 2SLS estimates, the new estimates for noncarceral conviction are somewhat smaller. For example, the 2SLS estimate for a future felony charge within the first seven years is 0.23 (95% CI: 0.04-0.42), while the estimate from this alternative approach is 0.19 (95% CI: 0.03,0.42). However, its 95% confidence interval contains nearly the entire confidence interval of the 2SLS estimate. Panel B reports the incarceration vs noncarceral conviction (I vs C) effect. Again, results are qualitatively similar to the 2SLS effects of incarceration on recidivism.

Overall, the results from applying this method tell a similar story to that of the 2SLS estimates: noncarceral conviction increases future criminal justice contact in the long run, and incarceration only has short-term incapacitation effects. This similarity suggests that any bias in the 2SLS estimates coming from the failure of UPM is likely small and therefore unlikely to change our qualitative conclusions.

6 Conclusion

In this paper, we study the impacts of noncarceral conviction on future criminal justice contact and draw a comparison to the impacts of incarceration. Across different analyses, we find that noncarceral conviction increases future criminal justice contact. In contrast, our analysis of the impact of incarceration only finds evidence for a shorter-term decrease in recidivism, which approximately coincides with the typical period of incapacitation. Thus, we find evidence for a "revolving door" effect of criminal justice contact, but this effect primarily operates through noncarceral conviction rather than through incarceration.

In addition to these substantive findings, this paper discusses the challenges stemming from multiple treatment alternatives in the commonly-used random judge research design. We develop an empirical framework to aid the interpretation of 2SLS estimands using judge stringency instruments when treatment effects can be heterogeneous. Within this framework, we provide assumptions that allow the estimands to be interpreted as causal and margin-specific. In particular, we show that requiring judge instruments to be treatment-specific is sufficient (in addition to straightforward extensions of exclusion, random assignment, relevance, and rich controls). We discuss which models of judge decision-making are consistent with treatment specificity, and propose an approach for testing this assumption empirically. We also derive the asymptotic

[52]Tables G.1 and G.2 in Appendix G.2 report results for alternative specifications that assume a standard logit structure and assume the correlation of the random effect is zero, respectively.

bias when it does not hold. Using this expression, it is possible to reason about the likely sign and magnitude of bias using features of the institutional setting. Finally, we propose and implement an empirical approach that better accommodates the fact that judge stringency instruments are not treatment-specific.

A number of papers have looked at how to reduce the number of felony convictions or their impact. Felony convictions could be reduced by increasing felony diversion (Mueller-Smith and Schnepel, 2021; Augustine et al., 2022), decriminalizing certain offenses, or downgrading the charge of conviction to a misdemeanor. Alternatively, the *impact* of felony convictions may be reduced by limiting the accessibility or permissible uses of criminal records. For instance, limiting how long criminal records are publicly available could mitigate employment effects of having a criminal record, potentially reducing recidivism by increasing formal employment options (Cullen et al., 2023). Likewise, reducing feedback loops within the penal system, such as automatic charge upgrades or sentence increases for those with a felony conviction, could mitigate the impact of a criminal record (Rose, 2021b).

Our findings suggest that these policies could contribute to reducing the penal system's revolving door problem. Of course, various other considerations may play a role. For example, there can be valid reasons for using felony conviction records in the hiring decision or to ratchet up punishment. However, given the prevalence of felony convictions in the U.S. – with 8% of adults and 33% of Black adult men estimated to have a felony conviction record (Shannon et al., 2017) – the impact of felony conviction on future criminal justice contact is an important part of this discussion.

References

Agan, Amanda Y. and Sonja Starr, "Ban the box, criminal records, and racial discrimination: A field experiment," *The Quarterly Journal of Economics*, 2018, *133* (1), 191–235.

Agan, Amanda Y, Andrew Garin, Dmitri K Koustas, Alexandre Mas, and Crystal Yang, "Can you Erase the Mark of a Criminal Record? Labor Market Impacts of Criminal Record Remediation," Technical Report, National Bureau of Economic Research 2024.

Agan, Amanda Y., Andrew Garin, Dmitri Koustas, Alex Mas, and Crystal S. Yang, "The Impact of Criminal Records on Employment, Earnings, and Tax Outcomes," 2023.

—, Jennifer L. Doleac, and Anna Harvey, "Misdemeanor prosecution," *The Quarterly Journal of Economics*, 2023, *138* (3), 1453–1505.

—, Matthew Freedman, and Emily Owens, "Is your lawyer a lemon? Incentives and selection in the public provision of criminal defense," *Review of Economics and Statistics*, 2021, *103* (2), 294–309.

Aizer, Anna and Joseph J. Jr. Doyle, "Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges," *Quarterly Journal of Economics*, May 2015, *130* (2), 759–803. MAG ID: 2118051501.

Alper, Mariel, Matthew R. Durose, and Joshua Markman, *2018 update on prisoner recidivism: a 9-year follow-up period (2005-2014)*, US Department of Justice, Office of Justice Programs, Bureau of Justice, 2018.

Angrist, Joshua D. and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton university press, 2009.

Arnold, David, Will Dobbie, and Peter Hull, "Measuring Racial Discrimination in Bail Decisions," *American Economic Review*, September 2022, *112* (9), 2992–3038.

Arteaga, Carolina, "Parental Incarceration and Children's Educational Attainment," *The Review of Economics and Statistics*, 10 2021, pp. 1–45.

Augustine, Elsa, Johanna Lacoe, Steven Raphael, and Alissa Skog, "The impact of felony diversion in San Francisco," *Journal of Policy Analysis and Management*, 2022, *41* (3), 683–709.

Avi-Itzhak, Benjamin and Reuel Shinnar, "Quantitative models in crime control," *Journal of Criminal Justice*, 1973, *1* (3), 185–217.

Baron, Jason E. and Max Gross, "Is There a Foster Care-To-Prison Pipeline? Evidence from Quasi-Randomly Assigned Investigators," Technical Report, National Bureau of Economic Research 2022.

Bayer, Patrick, Randi Hjalmarsson, and David Pozen, "Building Criminal Capital behind Bars: Peer Effects in Juvenile Corrections*," *The Quarterly Journal of Economics*, 02 2009, *124* (1), 105–147.

Berry, Steven T., Amit Gandhi, and Philip Haile, "Connected Substitutes and Invertibility of Demand," *Econometrica*, 2013, *81* (5), 2087–2111.

_ and Philip A. Haile, "Nonparametric Identification of Differentiated Products Demand Using Micro Data," 2022.

Bhuller, Manudeep and Henrik Sigstad, "2SLS with multiple treatments," *Journal of Econometrics*, 2024, *242* (1), 105785.

_ , Gordon B Dahl, Katrine V. Løken, and Magne Mogstad, "Incarceration, recidivism, and employment," *Journal of Political Economy*, 2020, *128* (4), 1269–1324.

Blandhol, Christine, John Bonney, Magne Mogstad, and Alexander Torgovitsky, "When is tsls actually late?," Technical Report, National Bureau of Economic Research 2022.

Blevins, Kristie R., Shelley Johnson Listwan, Francis T. Cullen, and Cheryl Lero Jonson, "A general strain theory of prison violence and misconduct: An integrated model of inmate behavior," *Journal of Contemporary Criminal Justice*, 2010, *26* (2), 148–166.

Chan, David C., Matthew Gentzkow, and Chuan Yu, "Selection with Variation in Diagnostic Skill: Evidence from Radiologists," *Quarterly Journal of Economics*, 2022, *137* (2), 729–83.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff, "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood," *American Economic Review*, September 2014, *104* (9), 2633–79.

Chiricos, T., K. Barrick, W. Bales, and S. Bontrager, "The labeling of convicted felons and its consequences for recidivism.," *Criminology: An Interdisciplinary Journal*, 2007, *45* (3), 547–581.

Collinson, Robert, John Eric Humphries, Nicholas Mader, Davin Reed, Daniel Tannenbaum, and Winnie van Dijk, "Eviction and poverty in American cities: Evidence from Chicago and New York," *Quarterly Journal of Economics*, 2024.

Craigie, Terry-Ann, "Ban the box, convictions, and public employment," *Economic Inquiry*, 2020, *58* (1), 425–445.

Cullen, Zoë, Will Dobbie, and Mitchell Hoffman, "Increasing the Demand for Workers with a Criminal Record*," *The Quarterly Journal of Economics*, February 2023, *138* (1), 103–150.

Dahl, Gordon B., Andreas Ravndal Kostøl, and Magne Mogstad, " Family Welfare Cultures *," *The Quarterly Journal of Economics*, 08 2014, *129* (4), 1711–1752.

de Chaisemartin, Clement, "Tolerating Defiance? Local Average Treatment Effects without Monotonicity," *Quantitative Economics*, 2017, *8* (2), 367–96.

Deshpande, Manasi and Michael Mueller-Smith, "Does welfare prevent crime? The criminal justice outcomes of youth removed from SSI," *The Quarterly Journal of Economics*, 2022, *137* (4), 2263–2307.

Dobbie, Will S. and Jae Song, "Debt relief and debtor outcomes: Measuring the effects of consumer bankruptcy protection," *American economic review*, 2015, *105* (3), 1272–1311.

_ , Jacob Goldin, and Crystal S. Yang, "The Effects of Pre-Trial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges," *American Economic Review*, 2018, *108* (2), 201–240.

_ , Paul Goldsmith-Pinkham, and Crystal S. Yang, "Consumer bankruptcy and financial health," *Review of Economics and Statistics*, 2017, *99* (5), 853–869.

Doleac, Jennifer L., "Encouraging Desistance from Crime," *Journal of Economic Literature*, 2023, *61* (2), 383–427.

_ and Benjamin Hansen, "The Unintended Consequences of "Ban the Box": Statistical Discrimination and Employment Outcomes When Criminal Histories Are Hidden," *Journal of Labor Economics*, 2020, *38* (2).

Doyle, Joseph J. Jr., "Child protection and adult crime: Using investigator assignment to estimate causal effects of foster care," *Journal of political Economy*, 2008, *116* (4), 746–770.

Drago, Francesco, Roberto Galbiati, and Pietro Vertova, "The deterrent effects of prison: Evidence from a natural experiment," *Journal of political Economy*, 2009, *117* (2), 257–280.

Enamorado, Ted, Benjamin Fifield, and Kosuke Imai, "Using a probabilistic model to assist merging of large-scale administrative records," *American Political Science Review*, 2019, *113* (2), 353–371.

Estelle, Sarah M. and David C. Phillips, "Smart sentencing guidelines: The effect of marginal policy changes on recidivism," *Journal of public economics*, 2018, *164*, 270–293.

Farrar-Owens, Meredith, "The evolution of sentencing guidelines in Virginia: An example of the importance of standardized and automated felony sentencing data," *Federal Sentencing Reporter*, 2013, *25* (3), 168–170.

Feng, Josh and Xavier Jaravel, "Crafting intellectual property rights: Implications for patent assertion entities, litigation, and innovation," *American Economic Journal: Applied Economics*, 2020, *12* (1), 140–81.

Finlay, Keith, Matthew Gross, Carl Lieberman, Elizabeth Luh, and Michael G Mueller-Smith, "The Impact of Criminal Financial Sanctions: A Multi-State Analysis of Survey and Administrative Data," Working Paper 31581, National Bureau of Economic Research August 2023.

Franco, Catalina, David Harding, Jeffrey Morenoff, and Shawn Bushway, "Failing to Follow the Rules: Can Imprisonment Lead to More Imprisonment Without More Actual Crime," Working Paper, RAND 2022.

Frandsen, Brigham, Lars Lefgren, and Emily Leslie, "Judging Judge Fixed Effects," *American Economic Review*, January 2023, *113* (1), 253–277.

French, Eric and Jae Song, "The effect of disability insurance receipt on labor supply," *American economic Journal: economic policy*, 2014, *6* (2), 291–337.

Garin, Andrew, Dmitri Koustas, Carl McPherson, Samuel Norris, Matthew Pecenco, Evan K Rose, Yotam Shem-Tov, and Jeffrey Weaver, "The Impact of Incarceration on Employment, Earnings, and Tax Filing," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2023, (2023-108).

Gavrilova, Evelina and Steffen Juranek, "Female Inventors: The Drivers of the Gender Patenting Gap," *Available at SSRN 3828216*, 2021.

Goldsmith-Pinkham, Paul, Maxim Pinkovskiy, and Jacob Wallace, "The great equalizer: Medicare and the geography of consumer financial strain," Technical Report, National Bureau of Economic Research 2023.

Graef, Lindsay, Sandra G Mayson, Aurelie Ouss, and Megan T Stevenson, "Systemic Failure to Appear in Court," *U. Pa. L. Rev.*, 2023, *172*, 1.

Gross, Max and E Jason Baron, "Temporary stays and persistent gains: The causal effects of foster care," *American Economic Journal: Applied Economics*, 2022, *14* (2), 170–99.

Gupta, Arpit, Christopher Hansman, and Ethan Frenchman, "The Heavy Costs of High Bail: Evidence from Judge Randomization," *The Journal of Legal Studies*, 2016, *45* (2), 471–505.

Hagan, John, "The social embeddedness of crime and unemployment," *Criminology*, 1993, *31* (4), 465–491.

Harding, David J, Jeffrey D Morenoff, Anh P Nguyen, and Shawn D Bushway, "Imprisonment and labor market outcomes: Evidence from a natural experiment," *American Journal of Sociology*, 2018, *124* (1), 49–110.

Heckman, James J. and Edward J Vytlacil, "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments," *Handbook of econometrics*, 2007, *6*, 4875–5143.

— **and Rodrigo Pinto**, "Unordered Monotonicity," *Econometrica*, 2018, *86* (1), 1–35.

— **, John Eric Humphries, and Gregory Veramendi**, "Dynamic treatment effects," *Journal of Econometrics*, 2016, *191* (2), 276–292. Innovations in Measurement in Economics and Econometrics.

Heckman, James J, John Eric Humphries, and Gregory Veramendi, "Returns to education: The causal effects of education on earnings, health, and smoking," *Journal of Political Economy*, 2018, *126* (S1), S197–S246.

Heckman, James J., Sergio Urzua, and Edward Vytlacil, "Understanding Instrumental Variables in Models with Essential Heterogeneity," *The Review of Economics and Statistics*, 08 2006, *88* (3), 389–432.

— **,** — **, and** — , "Instrumental Variables in Models with Multiple Outcomes: the General Unordered Case," *Annals of Economics and Statistics*, 2008, (91/92), 151–174.

Heinesen, Eskil, Christian Hvid, Lars Johannessen Kirkebøen, Edwin Leuven, and Magne Mogstad, "Instrumental Variables with Unordered Treatments: Theory and Evidence from Returns to Fields of Study," NBER Working Paper 30574, National Bureau of Economic Research, Cambridge, MA October 2022.

Hjalmarsson, Randi, "Juvenile jails: A path to the straight and narrow or to hardened criminality?," *The Journal of Law and Economics*, 2009, *52* (4), 779–809.

Holzer, Harry J, Steven Raphael, and Michael A Stoll, "Perceived criminality, criminal background checks, and the racial hiring practices of employers," *The Journal of Law and Economics*, 2006, *49* (2), 451–480.

— **,** — **, and** — , "The effect of an applicant's criminal history on employer hiring decisions and screening practices: Evidence from Los Angeles," *Barriers to reentry*, 2007, *4* (15), 117–150.

Huttunen, Kristiina, Martti Kaila, and Emily Nix, "The Punishment Ladder: Estimating the Impact of Different Punishments on Defendant Outcomes," 2020.

Imbens, Guido and Stefan Wager, "Optimized regression discontinuity designs," *Review of Economics and Statistics*, 2019, *101* (2), 264–278.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475.

Irankunda, Armel, Gregory N. Price, Norense E. Uzamere, and Miesha J. Williams, "Ex-Incarceree/Convict Status: Beneficial for Self-Employment and Entrepreneurship?," *The American Economist*, 2020, *65* (1), 144–162.

Jordan, Andrew, Ezra Karger, and Derek Neal, "Heterogeneous Impacts of Sentencing Decisions," Working Paper 31939, National Bureau of Economic Research December 2023.

Kamat, Vishal, Samuel Norris, and Matthew Pecenco, "Conviction, Incarceration, and Policy Effects in the Criminal Justice System," *Available at SSRN*, 2024.

Kane, Thomas J and Douglas O Staiger, "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation," Working Paper 14607, National Bureau of Economic Research December 2008.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad, "Field of Study, Earnings, and Self-Selection," *The Quarterly Journal of Economics*, 2016, *131* (3), 1057–1112.

Kline, Patrick and Christopher R. Walters, "Evaluating Public Programs with Close Substitutes: The Case of Head Start*," *The Quarterly Journal of Economics*, 07 2016, *131* (4), 1795–1848.

Kling, Jeffrey R., "Incarceration Length, Employment, and Earnings," *American Economic Review*, June 2006, *96* (3), 863–876.

Kohler-Hausmann, Issa, *Misdemeanorland: Criminal courts and social control in an age of broken windows policing.*, Princeton University Press, 2018.

Kolesár, Michal and Christoph Rothe, "Inference in regression discontinuity designs with a discrete running variable," *American Economic Review*, 2018, *108* (8), 2277–2304.

Kuziemko, Ilyana, "How should inmates be released from prison? An assessment of parole versus fixed-sentence regimes," *The Quarterly Journal of Economics*, 2013, *128* (1), 371–424.

LaCasse, Chantale and A Abigail Payne, "Federal sentencing guidelines and mandatory minimum sentences: Do defendants bargain in the shadow of the judge?," *The Journal of Law and Economics*, 1999, *42* (S1), 245–270.

Lee, Sokbae and Bernard Salanié, "Identifying Effects of Multivalued Treatments," *Econometrica*, 2018, *86* (6), 1939–1963.

Leslie, Emily and Nolan G. Pope, "The Unintended Impact of Pretrial Detention on Case Outcomes: Evidence from New York City Arraignments," *The Journal of Law and Economics*, 2017, *60* (3), 529–557.

Lieberman, Carl, Elizabeth Luh, Michael Mueller-Smith, and US CensusBureau UniversityofMichigan UniversityofMichigan, *Criminal court fees, earnings, and expenditures: A multi-state RD analysis of survey and administrative data*, US Census Bureau, Center for Economic Studies, 2023.

Loeffler, Charles E., "Does Imprisonment Alter the Life Course? Evidence on Crime and Employment From a Natural Experiment," *Criminology*, 2013, *51* (1), 137–166.

— **and Daniel S. Nagin**, "The impact of incarceration on recidivism," *Annual Review of Criminology*, 2022, *5*, 133–152.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand, "Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt," *American economic review*, 2013, *103* (5), 1797–1829.

Mayson, Sandra Gabriel and Megan Stevenson, "Misdemeanors by the Numbers," *Boston College Law Review*, March 2020, *61* (3).

Morris, Carl N., "Parametric Empirical Bayes Inference: Theory and Applications," *Journal of the American Statistical Association*, 1983, *78* (381), 47–55.

Mountjoy, Jack, "Community colleges and upward mobility," *American Economic Review*, 2022, *112* (8), 2580–2630.

Mueller-Smith, Michael, "The Criminal and Labor Market Impacts of Incarceration: Identifying Mechanisms and Estimating Household Spillovers," Working Paper, University of Michigan 2015.

— **and Kevin T. Schnepel**, "Diversion in the criminal justice system," *The Review of Economic Studies*, 2021, *88* (2), 883–936.

Mueller-Smith, Michael G, Benjamin Pyle, and Caroline Walker, "Estimating the Impact of the Age of Criminal Majority: Decomposing Multiple Treatments in a Regression Discontinuity Framework," Working Paper 31523, National Bureau of Economic Research August 2023.

Natapoff, Alexandra, "Misdemeanors," *S. Cal. L. Rev.*, 2011, *85*, 1313.

Norris, Samuel, "Examiner inconsistency: Evidence from refugee appeals," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2019, (2018-75).

_ , Matthew Pecenco, and Jeffrey Weaver, "The Effects of Parental and Sibling Incarceration: Evidence from Ohio," *American Economic Review*, September 2021, *111* (9), 2926–63.

Pager, Devah, "The mark of a criminal record," *American journal of sociology*, 2003, *108* (5), 937–975.

_ , Rebecca Goldstein, Helen Ho, and Bruce Western, "Criminalizing Poverty: The Consequences of Court Fees in a Randomized Experiment," *American Sociological Review*, 2022, *87* (3), 529–553.

PASC, "County Time Served and Revocations: 2013 report," Technical Report, Pennsylvania Commission on Sentencing 2013.

Phelps, Michelle S., "The paradox of probation: Community supervision in the age of mass incarceration," *Law & policy*, 2013, *35* (1-2), 51–80.

_ , "Mass probation: Toward a more robust theory of state variation in punishment," *Punishment & society*, 2017, *19* (1), 53–73.

Philippe, Arnaud, "Learning by doing. How do criminals learn about criminal law?," *University of Bristol, Working Paper*, 2020.

Reaves, Brian A, "Felony defendants in large urban counties, 2009-statistical tables," *Washington, DC: US Department of Justice*, 2013.

Rivera, R, "Release, detain or surveil? The effects of electronic monitoring on defendant outcomes," *Unpublished manuscript, Columbia University*, 2023.

Rose, Evan, "Does Banning the Box Help Ex-Offenders Get Jobs? Evaluating the Effects of a Prominent Example," *Journal of Labor Economics*, 2021, *39* (1).

_ , "Who gets a second chance? Effectiveness and equity in supervision of criminal offenders," *The Quarterly Journal of Economics*, 2021, *136* (2), 1199–1253.

_ **and Yotam Shem-Tov**, "How does incarceration affect reoffending? estimating the dose-response function," *Journal of Political Economy*, 2021, *129* (12), 3302–3356.

Sampat, Bhaven and Heidi L. Williams, "How do patents affect follow-on innovation? Evidence from the human genome," *American Economic Review*, 2019, *109* (1), 203–36.

Shannon, Sarah K. S., Christopher Uggen, Jason Schnittker, Melissa Thompson, Sara Wakefield, and Michael Massoglia, "The Growth, Scope, and Spatial Distribution of People With Felony Records in the United States, 1948-2010," *Demography*, 2017, *54* (5), 1795–1818.

Sigstad, Henrik, "Monotonicity among Judges: Evidence from Judicial Panels and Consequences for Judge IV Designs," *Working Paper*, 2024.

Stevenson, Megan T., "Breaking Bad: Mechanisms of Social Influence and the Path to Criminality in Juvenile Jails," *The Review of Economics and Statistics*, 2017, *99* (5), 824–838.

Sykes, Gresham, *The Society of Captives*, Princeton University Press, 1958.

Wolff, Nancy, Cynthia L. Blitz, Jing Shi, Jane Siegel, and Ronet Bachman, "Physical violence inside prisons: Rates of victimization," *Criminal justice and behavior*, 2007, *34* (5), 588–599.

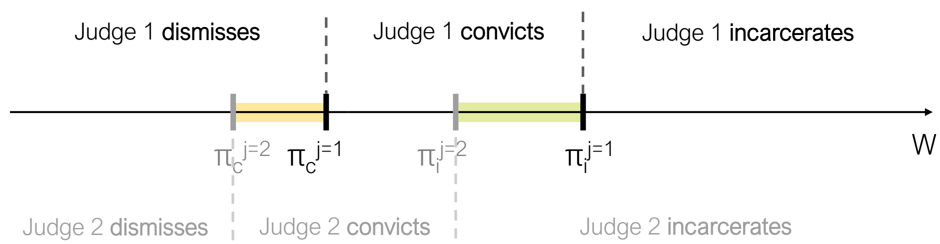
Zimring, Franklin E, Gordon Hawkins, and James Vorenberg, *Deterrence: The legal threat in crime control*, University of Chicago Press Chicago, 1973.

7 Figures and tables

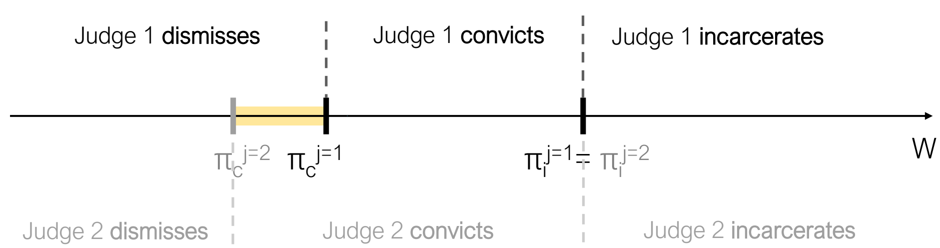
7.1 Figures

Figure 1: Ordered choice model

(a)

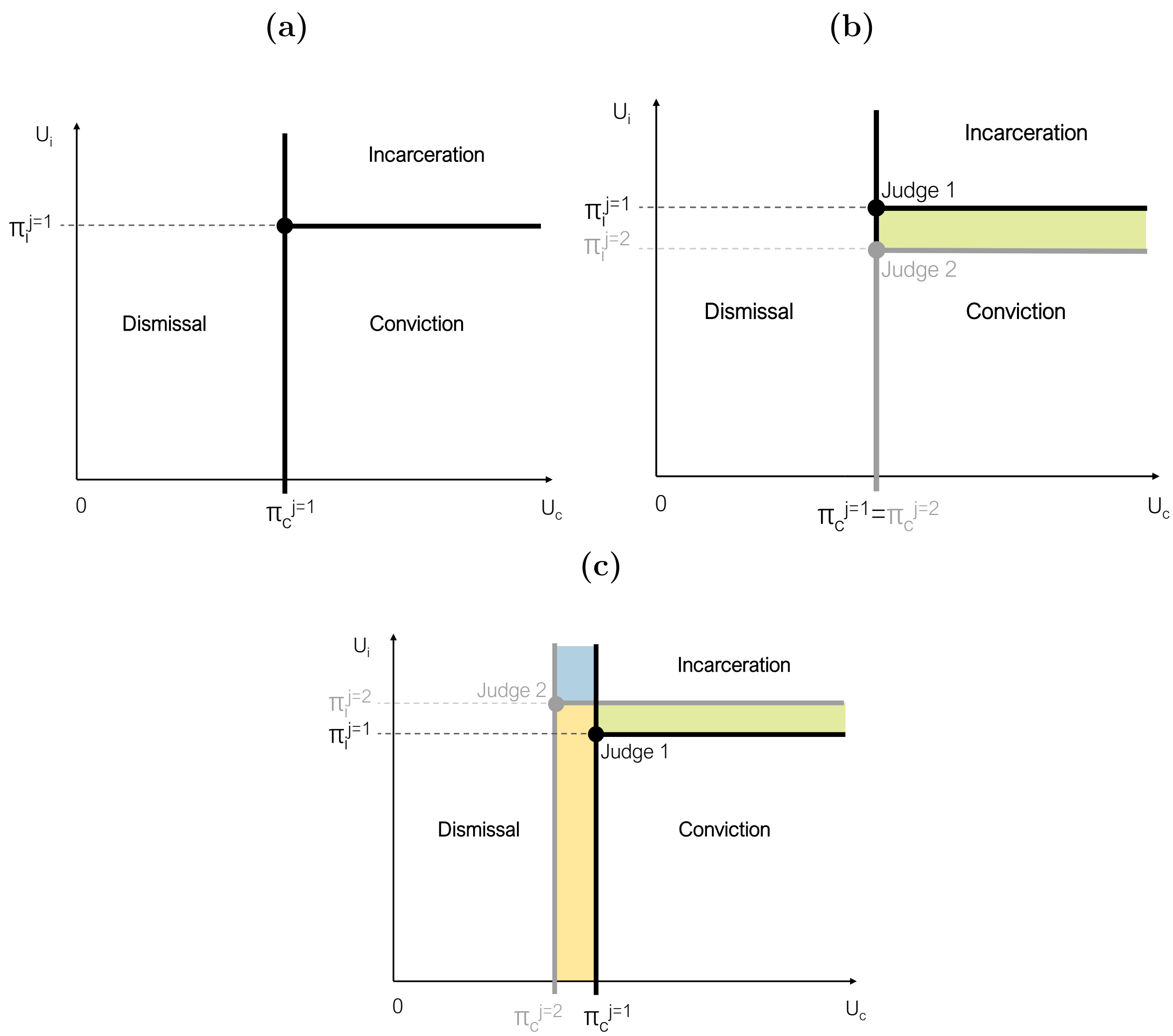


(b)



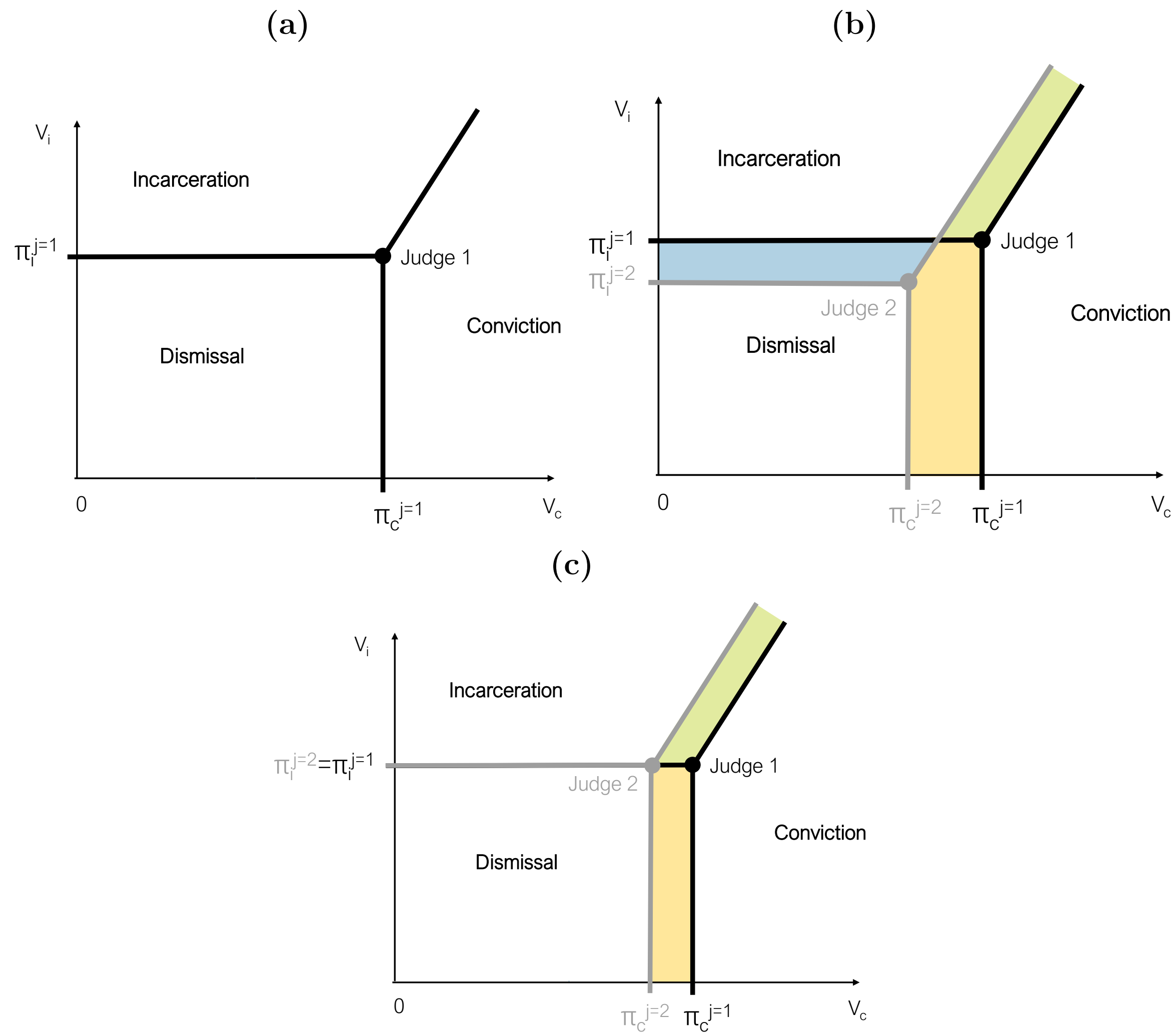
Note: This figure visualizes how, under the ordered choice model discussed in Section 3.3.1, judges classify individuals into incarceration, conviction, and dismissal depending on the cases' unobservable W. Panel (a) visualizes this for two arbitrary judges, and Panel (b) does so for two judges with the same incarceration stringency but different conviction stringencies.

Figure 2: Sequential choice model



Note: This figure visualizes how, under the sequential choice model discussed in Section 3.3.2, judges classify individuals into incarceration, conviction, and dismissal depending on the cases' unobservable U_i and U_c. Panel (a) visualizes this for an arbitrary judge, Panel (b) does so for two judges with the same dismissal stringency and different conviction stringencies, and Panel (c) for two judges with the same incarceration stringency but where judge 2 has a higher conviction stringency.

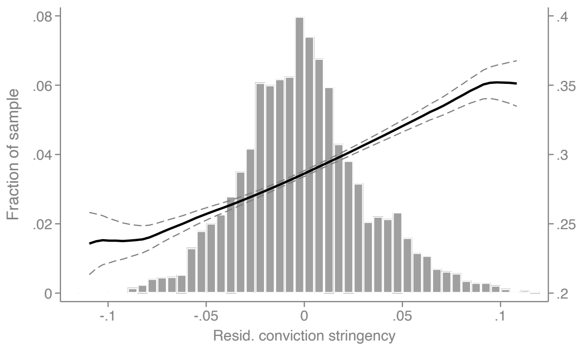
Figure 3: Unordered multinomial choice model



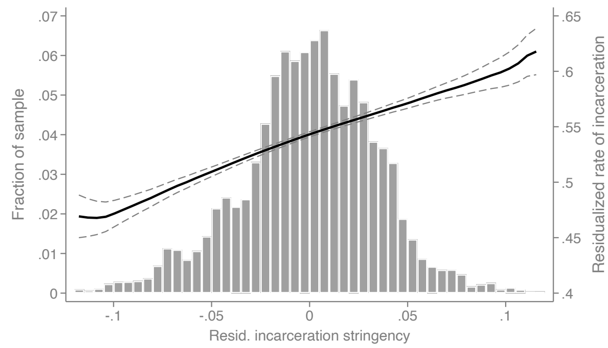
Note: This figure visualizes how, under the unordered multinomial choice model discussed in Section 3.3.2, judges classify individuals into incarceration, conviction, and dismissal depending on the cases' unobservable V_i and V_c. Panel (a) visualizes this for an arbitrary judge, Panel (b) does so for two judges with the same incarceration stringency but where judge 2 has higher conviction stringency, and Panel (c) for two judges with the same threshold for incarceration but where judge 2 has a higher conviction stringency.

Figure 4: Distribution of the stringency instruments

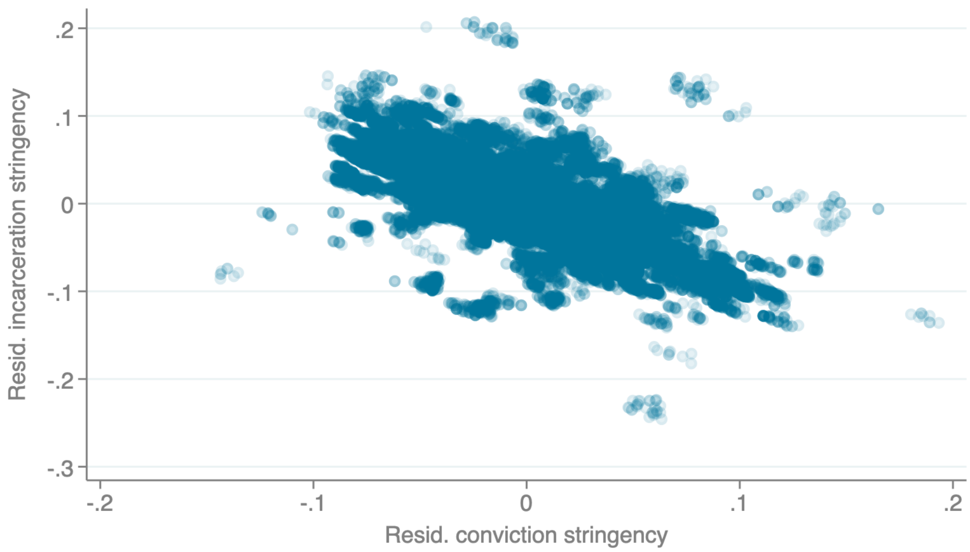
(a) Noncarceral conviction



(b) Incarceration



(c) Scatter plot of residualized instruments



Note: This figure presents our first stages in graphical format for noncaceral conviction (Panel A), where the outcome is an indicator for the case ending in conviction without incarceration, or incarceration (Panel B), where the outcome is an indicator for the case ending with a carceral sentence. The histograms plot the density of our residualized measures of conviction or incarceration stringency, and the line plots estimates of the first stage regression with conviction (Panel A) and incarceration (Panel B) as the dependent variable. Panel C is a scatter plot of the residualized incarceration and conviction instruments. In all three panels, the corresponding instrument is residualized against day-of-the-week, and court-by-year fixed effects. In all figures the sample is restricted to all cases for which outcomes can be observed for at least seven years.

7.2 Tables

Table 1: Summary statistics: 2SLS sample

	Dismissed	Convicted	Incarcerated
	(1)	(2)	(3)
Offenses			
Drugs	0.35	0.33	0.29
Larceny	0.17	0.29	0.25
Assault	0.19	0.08	0.18
Fraud	0.09	0.16	0.10
Traffic	0.04	0.05	0.13
Burglary	0.06	0.07	0.08
Robbery	0.05	0.02	0.06
Sexual assault	0.03	0.02	0.03
Kidnapping	0.03	0.01	0.02
Murder	0.01	0.00	0.01
Defendant characteristics			
Black	0.57	0.51	0.60
Female	0.22	0.32	0.16
% of ppl in zip earning <25K	0.46	0.44	0.46
Incarceration			
Has misdemeanor	0.06	0.09	0.08
Prior conviction within 5 years	0.14	0.10	0.22
Incarceration length	0.00	0.00	27.45
Probation length	0.00	31.50	39.34
Post-release			
Any charge within 1 year	0.09	0.09	0.07
Median incar. leng.	0	0	12
Median prob. leng.	0	12	12
Percent of sample	16	30	55
Observations	28,589	54,640	100,152

Note: This table shows means and select medians of relevant variables for the data used in the 2SLS analysis split into the three subsamples. The first column shows estimates for those whose cases were dismissed or who were found not guilty. The second column shows estimates for those whose cases ended with a conviction but without incarceration. The final column shows results for those whose cases ended with incarceration. The summary statistics are for cases adjudicated in 2012 or earlier, representing our seven year estimates. The incarceration and probation length medians and means are in months. Probation length is top-coded at 20 years.

Table 2: Relevance: first stage coefficients for the 2SLS analysis

	Non-carceral conviction			Incarceration		
	(1)	(2)	(3)	(4)	(5)	(6)
Conviction stringency	0.63***	0.60***	0.59***			
	(0.033)	(0.032)	(0.046)			
Incarceration stringency			-0.010	0.62***	0.59***	0.60***
			(0.041)	(0.033)	(0.032)	(0.035)
Dismissal stringency						0.033
						(0.051)
Controls	No	Yes	Yes	No	Yes	Yes
Mean dep. var.	0.298	0.298	0.298	0.546	0.546	0.546
F-stat	360.3	339.7	165.4	346.7	350.5	287.8
N	183,381	183,381	183,381	183,381	183,381	183,381

Note: This table reports the coefficient on the instruments from the first stage of the 2SLS regressions. Columns (1)-(3) report these coefficients for the conviction analysis, where the outcome is an indicator for the case ending in conviction (without incarceration). The first column includes only the instrument, the second column adds controls for the individual and case, and the third column controls for the leave-one-out judge incarceration stringency. Columns (4)-(6) repeat this analysis, but for the case ending in incarceration, and the final row controlling for judge dismissal stringency. Regression includes court-by-year fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. The first stage analysis in this table is on those cases adjudicated in 2012 or earlier. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

Table 3: Balance

	Convicted	**Conv. stringency**	**Incarceration**	**Incar. stringency**
	(1)	(2)	(3)	(4)
Any prior conv.	-0.1370***	-0.0000	0.1691***	0.0003
	(0.0029)	(0.0002)	(0.0032)	(0.0002)
Female	0.1207***	-0.0003*	-0.1242***	0.0002
	(0.0032)	(0.0002)	(0.0031)	(0.0002)
Black	-0.0416***	0.0002	0.0460***	-0.0002
	(0.0025)	(0.0002)	(0.0026)	(0.0002)
Has misdemeanor	0.0436***	0.0001	-0.0150***	0.0003
	(0.0047)	(0.0003)	(0.0050)	(0.0003)
Drugs	-0.0283***	0.0003	0.0706***	-0.0000
	(0.0037)	(0.0002)	(0.0041)	(0.0003)
Larceny	-0.0095***	0.0003	0.0996***	0.0003
	(0.0035)	(0.0002)	(0.0037)	(0.0002)
Assault	-0.1542***	-0.0011***	0.1576***	0.0012***
	(0.0035)	(0.0002)	(0.0043)	(0.0003)
Fraud	0.0251***	0.0004	0.0515***	0.0006*
	(0.0040)	(0.0003)	(0.0042)	(0.0003)
Traffic	-0.1860***	-0.0003	0.3309***	0.0006*
	(0.0042)	(0.0003)	(0.0048)	(0.0004)
Burglary	-0.0406***	-0.0001	0.0780***	0.0005
	(0.0043)	(0.0003)	(0.0047)	(0.0003)
Robbery	-0.0948***	-0.0002	0.1645***	0.0004
	(0.0048)	(0.0004)	(0.0059)	(0.0004)
Sexual assault	-0.1680***	-0.0007	0.2069***	0.0012**
	(0.0062)	(0.0005)	(0.0074)	(0.0006)
Kidnapping	-0.0631***	-0.0005	-0.0023	0.0006
	(0.0066)	(0.0006)	(0.0085)	(0.0006)
Murder	-0.1538***	-0.0012	0.1763***	0.0010
	(0.0076)	(0.0008)	(0.0119)	(0.0010)
F-stat joint F-test	568.532	3.757	803.043	2.666
P-value joint F-test	0.000	0.000	0.000	0.001
Observations	183,381	183,381	183,381	183,381

Note: This table shows estimates from regressions of either case outcomes (noncarceral conviction or incarceration indicators) or judge stringency measures on case characteristics. Regressions include court-by-year fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. The offenses are ordered by their prevalence in the data. The balance outcomes shown are for those cases adjudicated in 2012 or earlier. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01. To see the balance table in standard deviation units, see Appendix Table D.1

Table 4: Noncarceral conviction and recidivism

	Year 1		Year 2-4		Year 5-7		Year 1-7	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Fut. charge	-0.002	0.105**	0.004	0.085	0.006**	0.077	0.011***	0.233**
	(0.002)	(0.046)	(0.003)	(0.075)	(0.002)	(0.075)	(0.004)	(0.097)
Fut. conviction	0.001	0.136***	0.007***	0.114	0.007***	0.054	0.014***	0.298***
	(0.002)	(0.043)	(0.003)	(0.072)	(0.002)	(0.071)	(0.004)	(0.095)
Fut. incarceration	0.001	0.113***	0.006**	0.059	0.005**	-0.025	0.012***	0.214**
	(0.002)	(0.037)	(0.002)	(0.063)	(0.002)	(0.057)	(0.003)	(0.083)
Ctrl. comp. mean: fut. chrg.	0.158	0.158	0.302	0.302	0.237	0.237	0.494	0.494
Ctrl. mean: fut. chrg.	0.089	0.089	0.170	0.170	0.129	0.129	0.297	0.297
Ctrl. comp. mean: fut. conv.	0.138	0.138	0.264	0.264	0.225	0.225	0.460	0.460
Ctrl. mean: fut. conv.	0.076	0.076	0.148	0.148	0.114	0.114	0.268	0.268
Ctrl. comp. mean: fut. incar.	0.135	0.135	0.288	0.288	0.276	0.276	0.523	0.523
Ctrl. mean: fut. incar.	0.054	0.054	0.109	0.109	0.083	0.083	0.204	0.204
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381

Note: This table shows regression estimates of the impact of conviction on future recidivism. The four columns report results for four time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). For each panel, we report ordinary least squares (OLS) and instrumental variable (IV) estimates. The sample includes cases adjudicated in 2012 and earlier. All regressions control for race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. The first three rows report the estimated impact of conviction on different measures of recidivism. The first row is for any future felony charge, the second row is for any future felony conviction, and the third row is for any future felony incarceration. All IV regressions control for judge incarceration stringency. For the OLS estimates, we regress recidivism on having a conviction (regardless of incarceration status), controlling for incarceration. The estimates presented are the coefficient on the conviction variable. The middle portion of the table reports the control complier mean and control mean for each of the three outcomes we consider. Control means are calculated for cases that end in dismissal. See Appendix F.4 for details on the calculation of control complier means. Standard errors are clustered at the judge-year level. Stars denote * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Incarceration and recidivism

	Year 1		Year 2-4		Year 5-7		Year 1-7	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Fut. charge	-0.022***	-0.097***	0.013***	-0.015	0.025***	0.004	0.023***	-0.070
	(0.002)	(0.029)	(0.002)	(0.047)	(0.002)	(0.040)	(0.003)	(0.059)
Fut. conviction	-0.018***	-0.112***	0.014***	-0.037	0.023***	0.021	0.022***	-0.106*
	(0.001)	(0.029)	(0.002)	(0.047)	(0.002)	(0.039)	(0.003)	(0.058)
Fut. incarceration	-0.010***	-0.072***	0.017***	0.008	0.021***	0.053	0.027***	-0.030
	(0.001)	(0.024)	(0.002)	(0.041)	(0.002)	(0.032)	(0.003)	(0.051)
Ctrl. comp. mean: fut. chrg.	0.122	0.122	0.199	0.199	0.147	0.147	0.370	0.370
Ctrl. mean: fut. chrg.	0.088	0.088	0.175	0.175	0.132	0.132	0.306	0.306
Ctrl. comp. mean: fut. conv.	0.084	0.084	0.168	0.168	0.113	0.113	0.310	0.310
Ctrl. mean: fut. conv.	0.077	0.077	0.159	0.159	0.120	0.120	0.283	0.283
Ctrl. comp. mean: fut. incar.	0.043	0.043	0.071	0.071	0.051	0.051	0.166	0.166
Ctrl. mean: fut. incar.	0.055	0.055	0.115	0.115	0.084	0.084	0.212	0.212
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381

Note: This table shows regression estimates of the impact of incarceration on future recidivism. The four columns report results for four time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). For each panel, we report ordinary least squares (OLS) and instrumental variable (IV) estimates. Each time period restricts the sample to cases adjudicated in 2012 or earlier. All regressions control for race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. The first three rows report the estimated impact of incarceration on different measures of recidivism. The first row is for any future felony charge, the second row is for any future felony conviction, and the third row is for any future felony incarceration. All IV regressions control for the leave-one-out judge dismissal stringency. For the OLS estimates, we regress our measures of recidivism on incarceration, controlling for having a conviction (regardless of incarceration status). The middle portion of the table reports the control complier mean and control mean for each of the three outcomes we consider. Control means are calculated for cases that end in noncarceral conviction. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

Table 6: Testing the models with predicted recidivism

	Pred. recid. 1 year	Pred. recid. 2-4 years	Pred. recid. 5-7 years	Pred. recid. 1-7 years
Panel A: Ordered				
Conviction stringency (Z_c)	0.013***	0.030***	0.023***	0.048***
	(0.0039)	(0.0092)	(0.0072)	(0.014)
Mean dep. var.	0.093	0.202	0.153	0.346
N	100152	100152	100152	100152
Panel B: Sequential and ordered				
Incarceration stringency (Z_i)	-0.012***	-0.027**	-0.020**	-0.042**
	(0.0045)	(0.010)	(0.0083)	(0.017)
Mean dep. var.	0.090	0.183	0.138	0.321
N	28589	28589	28589	28589

Note: Predicted recidivism variables are created by taking the fitted values from a regression of recidivism after release on controls for demographics, charge, criminal record, and month, year-by-court, court-by-month-of-year, and day-of-week FE. For Panel A, we restrict to the incarcerated sample and regress predicted recidivism on conviction stringency, controlling for incarceration stringency and court-by-time fixed effects. For Panel B, we restrict to the dismissed sample and regress predicted recidivism on incarceration stringency, controlling for dismissal stringency and court-by-time fixed effects. The sample is restricted to cases adjudicated in 2012 or earlier. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

Table 7: Margin-specific treatment effects: an alternative approach

	Mixed logit with correlated normal random effects			
	Year 1	Year 2-4	Year 5-7	Year 1-7
Panel A: Noncarceral conviction vs dismissal (C vs D)				
Felony charge:	0.092**	0.188***	0.098	0.193**
	[0.010,0.182]	[0.057,0.348]	[-0.023,0.222]	[0.028,0.418]
	{0.061}	{0.138}	{0.125}	{0.311}
Felony conviction:	0.100***	0.206***	0.085	0.245***
	[0.026,0.194]	[0.095,0.347]	[-0.018,0.211]	[0.063,0.450]
	{0.053}	{0.113}	{0.121}	{0.256}
Felony incarceration:	0.063**	0.142***	0.043	0.155*
	[0.006,0.123]	[0.027,0.264]	[-0.063,0.144]	[-0.000,0.334]
	{0.053}	{0.097}	{0.100}	{0.248}
Panel B: Incarceration vs noncarceral conviction (I vs C)				
Felony charge:	-0.048***	0.005	-0.032	-0.077*
	[-0.081,-0.012]	[-0.045,0.062]	[-0.101,0.029]	[-0.166,0.006]
	{0.086}	{0.185}	{0.161}	{0.358}
Felony conviction:	-0.039***	0.002	-0.020	-0.075
	[-0.068,-0.010]	[-0.051,0.056]	[-0.076,0.044]	[-0.164,0.011]
	{0.075}	{0.171}	{0.142}	{0.334}
Felony incarceration:	-0.015	0.023	-0.016	-0.075*
	[-0.042,0.013]	[-0.027,0.077]	[-0.062,0.031]	[-0.146,0.003]
	{0.055}	{0.114}	{0.111}	{0.261}
Controls	Yes	Yes	Yes	Yes

This table reports margin-specific estimates of the impact of noncarceral conviction vs dismissal (Panel A) and incarceration vs noncarceral conviction (Panel B) using an unordered multinomial model based on the methodology developed in Mountjoy (2022). The treatment-specific instruments are recovered as described in Section 5.1 using a mixed-logit specification for the choice model where the intercept includes a correlated multivariate normal random effect and controls for female and Black indicators, an indicator for whether any charges are for violent crimes, an indicator for whether any charges are for property crimes, and indicator for whether any charges are for drug crimes, the number of charges, the time since last offense, and the number of misdemeanor charges associated with the case. The choice model is fit by district and 3-year bin. The estimates then use the recovered treatment-specific instruments in the method developed by Mountjoy (2022), where we include the same controls plus district and year fixed effects. The curly brackets report control-group complier means. In the top panel, this is the mean outcome for compliers whose cases were dismissed, while for the bottom panel, it is for those convicted but not incarcerated. 95% confidence intervals are reported in brackets and are based on 500 bootstrap samples. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01 based on the 90%, 95%, and 99% confidence intervals.

A Comparing Virginia's criminal justice system to other states

This appendix section shows how Virginia's criminal justice system compares to the U.S. overall, as well as to several states considered in recent related studies: Georgia, Michigan, North Carolina, Ohio, and Texas. First, we re-create figures from Norris et al. (2021) with an additional label for Virginia. Following Norris et al. (2021), we use 2004 data from the Pew Center on three-year recidivism rates, 2004 data on incarceration rates from the Bureau of Justice Statistics, and 2004 data on violent and property crime rates from the FBI Uniform Crime Reporting Program.[53]

Panel (a) of Appendix Figure A.1 shows that while Virginia has similar incarceration rates to the US average and other states, it has slightly lower recidivism (around 28% 3-year recidivism rates). Panel (b) shows that Virginia's property and violent crime rates are lower than the selection of states highlighted, but it is not an outlier in comparison to the rest of the states in the sample.

Appendix Figure A.2 shows prison and jail incarceration rates for the U.S., Virginia, and the five comparison states.[54] Virginia's prison incarceration rate, shown in Panel (a), is 447 per 100,000 people. This rate is somewhat higher, but comparable to the national rate, and roughly equal to the median among the five comparison states. The rate at which people are jailed in Virginia – 273 per 100,000 – is on the higher end compared to the national average and the five comparison states. Although it is not an obvious outlier relative to either the national average or the five comparison states, when interpreting our results, it is helpful to keep in mind that Virginia tends to rely more on jails than prisons and that conditions may vary across these two settings.

We next consider the racial and ethnic make-up of the prison population in Virginia. Figure A.3 displays the relative ratio of incarceration rates for Black vs White and Hispanic vs White residents.[55] The ratio for Black:White residents in Virginia is 4.3, just below the national average of 4.8 and roughly equal to the average of 4.4 of the other five comparison states. As in others states, Black residents are over-represented in the carceral population. The ratio for Hispanic:White residents is 0.5 for Virginia, lower than national average of 1.3 and most comparison states.

Lastly, we compare probation and parole rates (Figure A.4). Virginia's probation rate is close to the national average, as are most comparison states, with the exception of Georgia. However, the parole rate in Virginia – 22 per 100,000 residents – is much lower than the benchmarks. This difference is because discretionary parole was virtually abolished in Virginia for felonies committed after 1995, with inmates being required to serve at least 85% of their sentences, with the possibility to earn good-time credits toward early release. This also means that the initial sentence is more closely linked to time spent incarcerated than in other places.

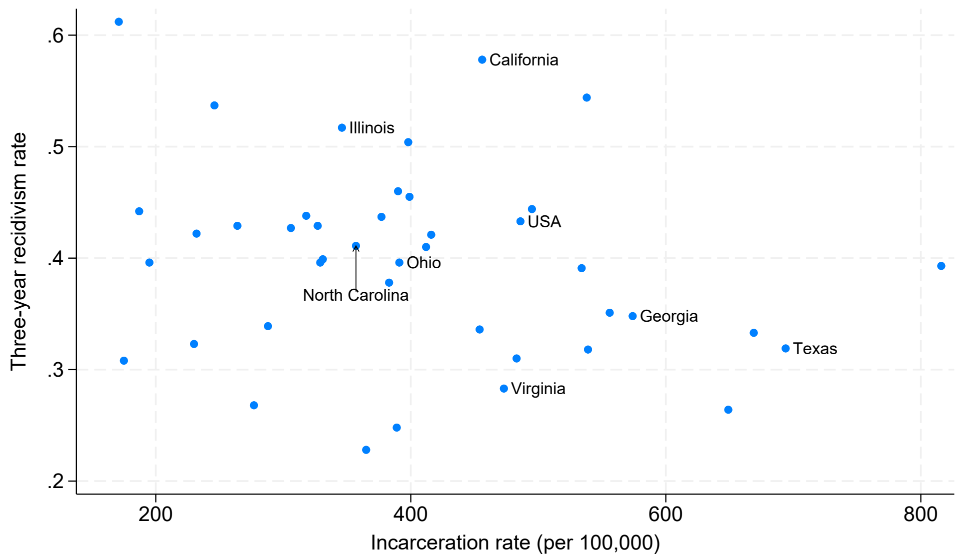
[53]This data can be found at https://www.pewtrusts.org/-/media/legacy/uploadedfiles/pcs_assets/2011/pewstateofrecidivismpdf.pdf, https://bjs.ojp.gov/content/pub/pdf/p04.pdf, and https://www2.fbi.gov/ucr/cius_04/.

[54]We use data from the Prison Policy Initiative. This data can be downloaded from https://www.prisonpolicy.org/reports/correctionalcontrol2018.html.

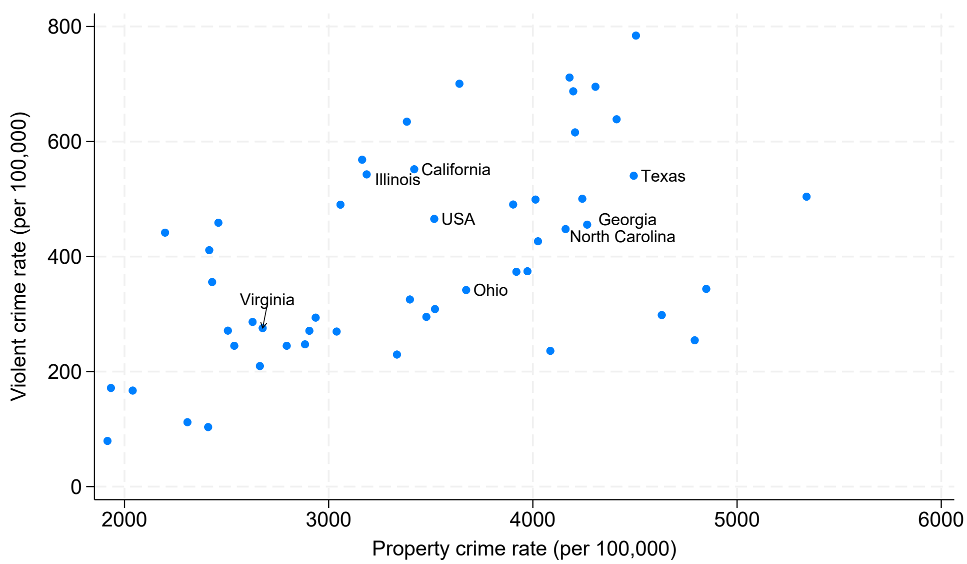
[55]These ratios read as follows: If out of every 100,000 Hispanic residents 200 are incarcerated, and out of every 100,000 White residents 400 are incarcerated, the Hispanic:White ratio is 0.5.

Figure A.1: State-level comparisons of recidivism, incarceration, and crime

(a) 3yr-recidivism rates vs incarceration rates



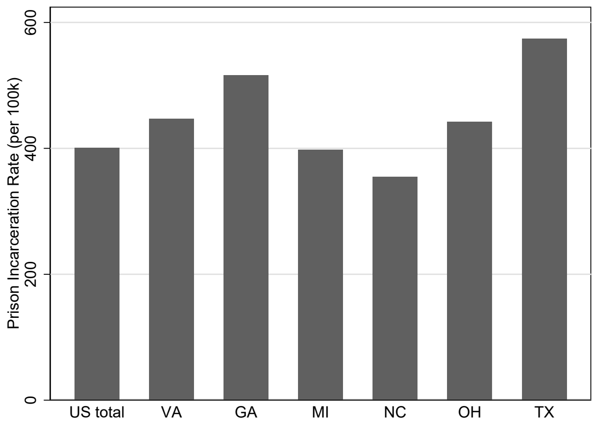
(b) Violent crime rates vs property crime rates



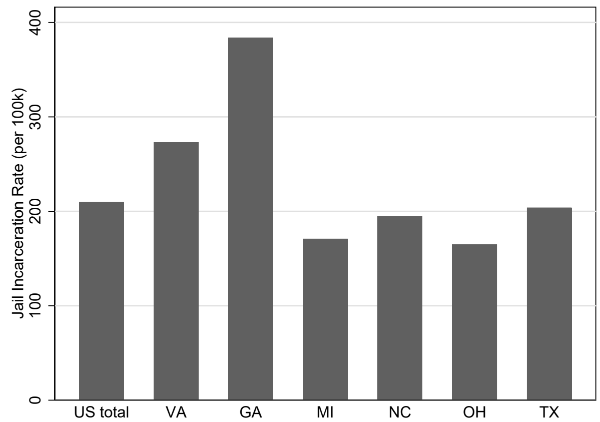
Note: Scatterplots of 2004 incarceration rates, 2004 three-year recidivism rates, and 2004 crime rates. Data gathered from the Pew Center, Bureau of Justice Statistics, and the FBI Uniform Crime Reporting Program.

Figure A.2: Incarceration rates

(a) Prison incarceration rates

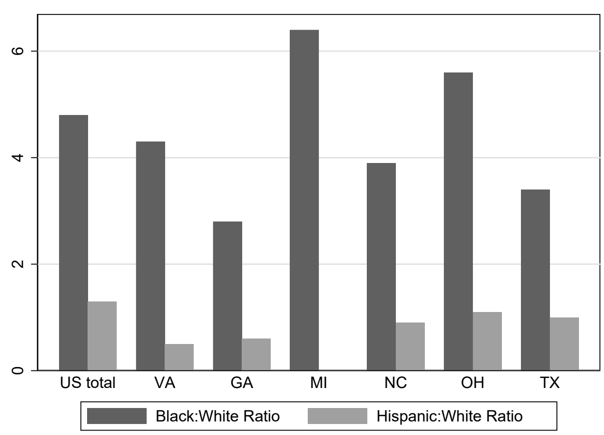


(b) Jail incarceration rates



Note: This figure shows the prison (Panel A) and jail (Panel B) incarceration rates, respectively, per 100,000 residents for Virginia, the U.S. overall, Georgia, Michigan, North Carolina, Ohio, and Texas. Based on 2017 and 2014 data respectively from the Prison Policy Initiative (December 2018 press release).

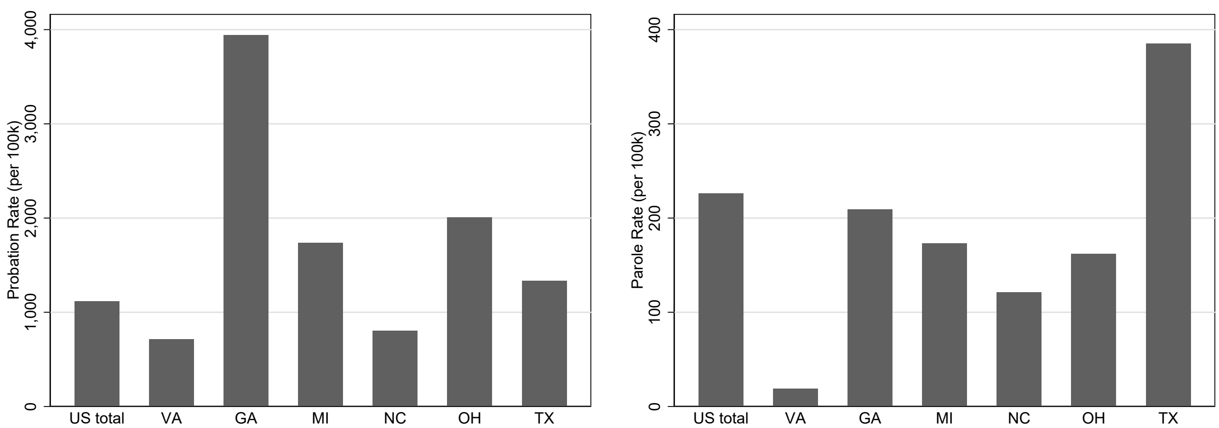
Figure A.3: Racial and ethnic composition of the imprisoned population



Note: This figure plots the ratio of incarceration rates for Black vs White residents (darker bars) and Hispanic vs White residents (lighter bars), for Virginia, the U.S. overall, Georgia, Michigan, North Carolina, Ohio, and Texas in 2019. Data from sentencingproject.org, used to calculate incarceration by ethnicity, is not available for Michigan.

Figure A.4: Virginia supervision rates comparison

(a) Probation rates **(b) Parole rates**



Note: Panel (a) shows the probation rate in Virginia per 100,000 people and Panel (b) shows the parole rate in Virginia per 100,000 people, both compared to the rates for the U.S. total, Georgia, Michigan, North Carolina, Ohio, and Texas. Based on 2016 data from the Prison Policy Initiative (December 2018 press release).

B Additional details on data construction

B.1 Main data sources

Virginia Circuit Courts (VCC) data. The Virginia Court system keeps all Circuit Court case records publicly available for anyone to search. We obtained this data from Ben Schoenfeld who web-scraped records from the courts and made the corresponding data available on http://virginiacourtdata.org/ for public download. This data covers criminal cases in which at least one charge is a felony. It contains information on charges (type and date), on the defendant (gender, race, partial birth date, and FIPS code of residence), and on Circuit Court proceedings for these cases (type, outcome, and judges on the proceedings) and is available for the period 2000-2019. All of Virginia is covered except for Alexandria and Fairfax counties. This is the primary data source for our 2SLS analysis with judge stringencies.

Virginia Criminal Sentencing Commission (VCSC) data. The VCSC provided a dataset that contains information on individuals in Virginia sentenced for a felony. This is used as supplementary data for our 2SLS analysis (to construct our measure of prior convictions) and as the main source for the RD analysis. The data provided to us by the VCSC includes records on all people convicted of a felony in Virginia from 1996 to 2020. This data includes information on the charge(s) of conviction, date of sentencing, sentence imposed for this conviction, guidelines-recommended sentence, points accrued on each item in a worksheet, and total worksheet scores. This data does not contain information on demographics or prior and future charges, so we match it to data from Virginia's Circuit Courts as described below.

B.2 Supplementary data sources

Virginia District Courts (VDC) data. The Virginia Court system also keeps all District Court case records publicly available for anyone to search. As with the Circuit Court data, we obtained this data from Ben Schoenfeld's web-scraped records (http://virginiacourtdata.org/). This data covers all dockets filed in District Court, including felonies and misdemeanors. The District Court is a court of limited jurisdiction; felony charges that are filed there cannot be adjudicated there. We use this data to obtain information about pretrial detention, as used in the RD specification that subsets to those never previously incarcerated.

Virginia residency data. We obtain information on residency status from a private vendor, matched to the VCSC data with name, social security number and partial birth date. We use the residency data to look at differential mobility in the RD sample. The vendor provided us with information as to which state the matched individual resides in post-sentencing. We receive snapshots of information from them 1, 3, 5, 7 years post-sentencing date, and we construct a variable indicating if an individual is in the state of Virginia 5 and 7 years post-sentencing. 7.7% of observations are missing residency.

IRS zip code income data. This is publicly available data produced by the IRS of average zip code earnings. We use the 2005 vintage and match in by zip onto our samples. This is supplementary data to our IV and RD analysis.

B.3 Data construction

This section details the data construction and cleaning process as well as the matching procedure implemented between the various raw datasets described above.

IV data. We begin with the sample of 3.4 million dockets from the VCC data between 2000 and 2019.

- In addition to dockets with felony charges – the focus of our analysis – the data also includes many dockets pertaining to technical issues (failures to appear in court, revocations, bond hearings, etc.) as well as some pertaining to misdemeanors. We only keep dockets pertaining to new felony charges (roughly 50% of all dockets), leaving roughly 1.6 million felony dockets remaining. We also drop dockets that are missing disposition date or initiation date, as well as cases where the disposition is on a weekend. This represents roughly 77,000 dockets, or less than 5% of the remaining sample.
- Sometimes prosecutors file separate dockets for different charges against the same defendant. This could happen if, for instance, the defendant was arrested for multiple burglaries or drug selling occasions. These nonetheless get processed together as one effective case. For our analyses, we define a "case" – our main unit of analysis – as composing all dockets with the same defendant and either the same or consecutive case numbers. Consecutive case numbers means that they were all filed at the same time. Docket level descriptors are aggregated to the case level (i.e., a case is considered "convicted" if at least one charge was adjudicated guilty). The 1.6 million dockets correspond to 773,553 cases.
- Some courts do not regularly fill out judge information. We drop all courts where less than 80% of judge names are filled out. These courts cover 171,718 cases or 22.2% of cases resulting in 601,835 remaining cases.
- Each case can have multiple hearings. Judge information is provided at the hearing level. We have hearing-level data for 502,732 cases, or 84% of cases.
- We then drop cases entirely missing judge information (37,191 cases dropped or 7.4% of cases resulting in 465,541 cases left).
- We limit ourselves to larger courts with multiple judges overseeing felony cases. In our main sample, we drop judges who see less than 100 cases over 3 years, and all observations in a court-by-year with only one judge. In our main specification, we require that we have at least 3 years per court where multiple judges are present, to avoid including courts and years in which judges simply overlapped because of turnover. In total, these sample restrictions lead us to drop 18,777 cases (4% of the sample), leaving us with 446,764 cases.
- We called clerks in the remaining courts to understand how cases were assigned. In our main specification, we dropped courts where the clerks described a case assignment mechanism that clearly wasn't quasi-random; for instance, ones in which cases are assigned based on judge specialization. We also drop one court after 2010 due to decreased data availability. This represents 121,931 cases, (27% of remaining cases). This leaves us with 324,799 cases.
- Lastly, we use the VCC data to calculate recidivism, defined as a new felony charge in Circuit Court within X years for various values of x. The VCC data

goes through 2019. The sample we use for most of the analysis, which is cases that have 7 years of data, is 183,381 observations. In a robustness check, we expand our analysis sample to include all available years. Our main sample includes only cases disposed prior to 2012 in order to have seven years post-disposition. In this robustness check, we expand the sample up to 2015 for outcomes in years 2-4 and up to 2018 for outcomes in year 1.

RD data. We begin by using the VCSC felony data as our universe of cases for each individual convicted of a felony in Virginia. We start with 458,164 observations between 2000 and 2018 (years for which we also have CC data, used to measure recidivism). From there we create two main samples for the RD analyses, as well as a supplementary sample that we use for robustness checks.

Incarceration-length RD data. The first sample leverages the discontinuity in the incarceration-length score as calculated in Worksheet A. This is the sample that we use to measure the effect of longer prison stays vs. shorter jail stays. For this set of analyses, we impose four restrictions on the sample.

- First, we drop offense categories in which the seriousness of the offense mandates a recommended prison sentence, since we do not have variation at the margins for these cases. The omitted offense categories include murder and voluntary manslaughter, rape, aggravated DWI, some more serious drug offenses, more serious types of assault, burglary, robbery, and other miscellaneous offenses. These constitute roughly 26% of the sample, or 118,364 cases.
- Second, we drop certain offense categories because the distribution of the sentence guidelines scores is not smooth, potentially due to the scoring of worksheets for those categories. Since the RD method requires a smooth evolution of potential outcomes across the running variable, these could be problematic for our design, even if this is mechanically due to the way in which points are accrued. The offense categories dropped are fraud, traffic, and weapons; these constitute 20% of the remaining data, or 72,026 cases. Our main results are robust to including these offense categories.
- Third, we drop individuals who are recorded as having no points in the incarceration-length score: 0.2% of the sample, or 758 cases. We infer that these are likely data errors, since about 10% of these individuals are recommended for prison despite being far below the cutoff at which a prison recommendation is warranted.
- We then match the VCSC sentencing data to the VCC data. VCC data allows us to construct our primary measure of new criminal justice contact (new felony charges in circuit court) as well as race, gender, arrest date, and prior charges. We drop cases from Fairfax and Alexandria, which are not in the CC data. We use the fuzzy matching method developed by Enamorado et al. (2019) and match on first name, last name, middle initial, FIPS code, birth month, and sentence date. For the years and counties in which a match is feasible, our match rate is 92%. Our final sample has 230,357 observations.

Probation/jail RD data. The second sample leverages the discontinuity in the probation/jail score found in Worksheet B. For this set of analyses, we impose similar sample restrictions as described previously.

- First, we drop anyone whose primary offense makes them ineligible for probation, as well as those convicted of violent offenses, since almost none of these are probation-eligible. This represents 59% of the data, or 269,437 cases.
- As previously, we drop individuals who are recorded as having no points in the probation/jail score (0.8%, or 1,576 cases) due to suspected data entry errors. We also drop offense categories for which there are only 2 points between our focal cutoff (probation/jail) and the secondary cutoff (short jail/long jail sentence), which represents 6.8% (or 12,765 cases) of the Worksheet B sample. The remaining offense categories either only have one cutoff (about half of cases) or have 3 points between the focal and secondary cutoff.
- For this data we also restrict to a sample where the Circuit Court match is feasible, using the same procedure as that described for the incarceration-length RD data. Our final sample has 130,692 cases.

Supplementary RD data. Finally, we create a supplementary sample that matches the Worksheet B sample to information on pretrial detention from the VDC data. This reduces our sample significantly since the VDC data is only available from 2010-2019. Since we use three years of follow up, the sample includes those convicted of a felony between 2010-2016: 49,246 cases.

Comparison between IV and RD data. While the data for the RD and the IV analyses come from the same general sources and have significant overlap, there are some key differences.

- The group of cases in the RD data is a subset of those in the 2SLS data, since the RD sample just covers those whose felony charges led to a conviction. For both sets of analyses, we have approximately 80% of Virginia's population since the VCC data misses Alexandria and Fairfax counties.
- In addition, as described above, we further subset the RD sample to include offense types that could, in theory, have led to defendants being on either side of the different RD thresholds.
- Tables 1 and H.1 present summary statistics for each sample.

B.4 Variable construction and definitions

Variable definitions.

- *Incarceration.* We define a person to be incarcerated if at least one of the charges resulted in a positive (greater than zero) carceral sentence.
- *Noncarceral conviction.* We define a person to be convicted if at least one charge led to a sentence, but no charge resulted in a carceral sentence.
- *Dismissal.* We define a case as dismissed if all charges were dismissed or withdrawn by prosecution (nolle prosequi); or if the defendant was acquitted of all charges.

- *Recidivism.* Our main measure of recidivism is whether a person has a new felony charge in Circuit Court for an offense that allegedly happened after the focal charge date. This measure does not include revocations unless these are also accompanied by a new felony charge for a new crime. We create these variables for recidivism in year 1, years 2-4, years 5-7, and years 1-7 cumulative. For the RD analyses, since we have more years of data, we also include measures for years 8-10 and years 1-10.
- *Recidivism-new conviction.* This is similar to our main recidivism measure, but here the indicator refers to a new conviction on a Circuit Court felony charge for a crime committed within the relevant time periods.
- *Recidivism-new incarceration.* Again, this outcome is similar to the previous variable, except the indicator means there is a new carceral sentence resulting from a Circuit Court felony charge for a crime committed within the relevant time period.
- *Prior conviction flag.* We define someone as having a prior felony conviction if they have a case in the VCSC data in the 5 years prior to the first offense date of their current case. We use VCSC data to build our prior conviction flag because our data goes back to 1996. This gives us at least 5 years of information on prior felony convictions for all cases in the 2SLS sample.
- *Judge on the case.* We define the judge on the case in the following way. Our main measure is the judge that appears when the "pleading" or the "remarks" variable in the hearings data is marked as "sentencing", "judgement", "dismissal", "conviction", or "final order". If this does not appear on a case, we fill in with the judge present on the disposition date. Finally, if the judge is still missing, for any remaining listings where there is an available judge, we use the maxmode to determine the presiding judge. In our sample, roughly 80% of hearings are in front of the judge whom we define as the judge for the case.[56]
- *Black.* Race of the defendant as defined in the VCC data. Almost all of the people for which race information is available are labeled either "Black" or "White." Ethnicity is not available.
- *Female.* Gender of the defendant as defined in the VCC data.
- *Incarceration Length.* This variable indicates how long in months an individual is imprisoned (if they have a carceral sentence). It will be 0 otherwise.
- *Income generating.* This is a variable that is used to determine whether the individual has new felony charges for an income-generating type of crime. We consider the following charges to be income-generating: burglary, drug charges (excluding drug possession), fraud, larceny, robbery, or prostitution.
- *Has misdemeanor.* An indicator if the current case has a misdemeanor charge as recorded in the Circuit Court data.
- *% of people in zip earning <25K.* Share of people earning less than 25K in a zip code, using matched IRS average zip code level earnings data.

[56]The other hearings could be seen by another judge because the primary judge is absent that day (sick or on vacation) or if the case was reassigned.

C Additional details on bias in 2SLS estimands

C.1 Proof of Proposition 1

When CPM holds but UPM does not, a shift from z_c to z_c' holding z_i fixed induces three types of flows: $d \to c$, $d \to i$, and $i \to c$. The reduced form effect is thus given by

$$E[Y(T(z_c', z_i)) - Y(T(z_c, z_i))] = \omega_{d\to c}\Delta_{d\to c}^{Y_c - Y_d} + \omega_{i\to c}\Delta_{i\to c}^{Y_c - Y_i} + \omega_{d\to i}\Delta_{d\to i}^{Y_i - Y_d}. \tag{1}$$

Since the overall probability of incarceration is fixed at z_i, the share of cases flowing into and out of incarceration must be equal in size (i.e., it must be that $\omega_{d\to i} = \omega_{i\to c}$). Hence, we can rewrite equation (1) as

$$E[Y(T(z_c', z_i)) - Y(T(z_c, z_i))] = \omega_{d\to c}\Delta_{d\to c}^{Y_c - Y_d} + \omega_{i\to c}\left[\Delta_{i\to c}^{Y_c - Y_i} + \Delta_{d\to i}^{Y_i - Y_d}\right]. \tag{2}$$

Next, observe that

$$\Delta_{d\to i}^{Y_i - Y_d} = \Delta_{d\to i}^{Y_i - Y_c} + \Delta_{d\to i}^{Y_c - Y_d}.$$

Hence, equation (2) can be rewritten as:

$$\begin{aligned}
E[Y(T(z_c', z_i)) - Y(T(z_c, z_i))] &= \omega_{d\to c}\Delta_{d\to c}^{Y_c - Y_d} + \omega_{i\to c}\left[\Delta_{i\to c}^{Y_c - Y_i} + \Delta_{d\to i}^{Y_i - Y_c} + \Delta_{d\to i}^{Y_c - Y_d}\right] \\
&= \omega_{d\to c}\Delta_{d\to c}^{Y_c - Y_d} + \omega_{i\to c}\Delta_{d\to i}^{Y_c - Y_d} + \omega_{i\to c}\left[\Delta_{d\to i}^{Y_i - Y_c} - \Delta_{i\to c}^{Y_i - Y_c}\right].
\end{aligned}$$

For the denominator of the Wald estimand, we have

$$E[T_c(z_c', z_i) - T_c(z_c, z_i)] = \omega_{d\to c} + \omega_{i\to c}.$$

Constructing the Wald estimand, we obtain equation (7):

$$\frac{E[Y(T(z_c', z_i)) - Y(T(z_c, z_i))]}{E[T_c(z_c', z_i) - T_c(z_c, z_i)]} =$$

$$\underbrace{\frac{\omega_{d\to c}\Delta_{d\to c}^{Y_c - Y_d} + \omega_{i\to c}\Delta_{d\to i}^{Y_c - Y_d}}{\omega_{d\to c} + \omega_{i\to c}}}_{\text{Positively-weighted avg. of } Y_c - Y_d \text{ treatment effects}} + \underbrace{\frac{\omega_{i\to c}}{\omega_{d\to c} + \omega_{i\to c}}\left[\Delta_{d\to i}^{Y_i - Y_c} - \Delta_{i\to c}^{Y_i - Y_c}\right]}_{\text{Bias term}}.$$

Moving from CPM to the stronger UPM assumption simplifies equation (7). First, recall that UPM($Z_c|Z_i$) implies that there can only be flows into $T = c$ when increasing Z_c from z_c to z_c'. Second, recall that fixing judge stringency $Z_i = z_i$ implies that the net probability of incarceration must remain constant. This second point implies that any flows from $T = i$ to $T = c$ would need to be compensated by flows from $T = d$ to $T = i$. Since UPM($Z_c|Z_i$) rules out flow from $T = d$ to $T = i$, there can be no flows from $T = i$ to $T = c$ since Z_i is fixed. This implies that $\omega_{i\to c} = 0$, which simplifies equation (7) to

$$\frac{E[Y(T(z_c', z_i)) - Y(T(z_c, z_i))]}{E[T_c(z_c', z_i) - T_c(z_c, z_i)]} = \Delta_{d\to c}^{Y_c - Y_d}.$$

C.2 Bias with four treatments

Here, we calculate the bias from a 2SLS estimate in a simple setting with four mutually exclusive treatments. For example, these could be dismissed; convicted without incarceration; convicted with a short carceral sentence; or convicted with a long carceral sentence: $T \in \{d, c, s, l\}$. The mutually-exclusive stringencies would then be: Z_d, Z_c, Z_s, Z_l. We assume CPM and the other assumptions, except for UPM (see Section 3.1 for details).

In the example below, we characterize bias when using differential stringencies to determine the causal effect of conviction vs dismissal. Let's consider two judges who have the same z_s and z_l, but different z_c. Following the notation from Appendix C.1, ωs represent shares of switchers. For example, $\omega_{d\to c}$ represents the proportion of people switching from $T = d$ to $T = c$ when shifting conviction stringency from z_c to z_c', holding z_s and z_l fixed.

The set of potential movers when changing z_c (holding fixed z_s and z_l) under CPM are: (1) $d \to c$, (2) $s \to c$, (3) $l \to c$, (4) $d \to s$, (5) $d \to l$, and (6) $l \to s$. Note that this is just one possible direction of switches that would be compatible with CPM. For instance, for (6), we could have reversed flows and allowed for $s \to l$ instead $l \to s$; but under CPM we can only have one, not the other. The same applies for (5).

As with 3 treatments, holding z_s fixed means that flows in and out of $T = s$ have to be equal, and holding z_l fixed means flows in and out of $T = l$ have to be equal. This means that $\omega_{s\to c} = \omega_{l\to s} + \omega_{d\to s}$ and $\omega_{d\to l} = \omega_{l\to s} + \omega_{l\to c}$.

The reduced form effect is thus given by:

$$\begin{aligned} &E[Y(T(z_c', z_s, z_l)) - Y(T(z_c, z_s, z_l))] = \\ &\Delta^{Y_c-Y_d}_{d\to c}\omega_{d\to c} + [\Delta^{Y_s-Y_d}_{d\to s}\omega_{d\to s} + \Delta^{Y_c-Y_s}_{s\to c}\omega_{s\to c}] + [\Delta^{Y_l-Y_d}_{d\to l}\omega_{d\to l} + \Delta^{Y_c-Y_l}_{l\to c}\omega_{l\to c}] + \Delta^{Y_s-Y_l}_{l\to s}\omega_{l\to s}, \end{aligned} \tag{3}$$

where brackets have been placed around two sets of terms to simplify the explanation of the next steps below.

For any difference in two potential outcomes, we can always rewrite it as $Y_k - Y_j = (Y_k - Y_m) - (Y_j - Y_m)$. Using this, the first term in the square brackets in equation (3) can be rewritten as follows:

$$\begin{aligned} [\Delta^{Y_s-Y_d}_{d\to s}\omega_{d\to s} + \Delta^{Y_c-Y_s}_{s\to c}\omega_{s\to c}] &= [\Delta^{Y_s-Y_d}_{d\to s}\omega_{d\to s} + \Delta^{Y_c-Y_s}_{s\to c}(\omega_{d\to s} + \omega_{l\to s})] \\ &= [\Delta^{Y_s-Y_d}_{d\to s}\omega_{d\to s} + (\Delta^{Y_c-Y_d}_{s\to c} - \Delta^{Y_s-Y_d}_{s\to c})\omega_{d\to s} + \Delta^{Y_c-Y_s}_{s\to c}\omega_{l\to s}] \\ &= [\Delta^{Y_c-Y_d}_{s\to c}\omega_{d\to s} + (\Delta^{Y_s-Y_d}_{d\to s} - \Delta^{Y_s-Y_d}_{s\to c})\omega_{d\to s} + \Delta^{Y_c-Y_s}_{s\to c}\omega_{l\to s}]. \end{aligned} \tag{4}$$

Similarly, the second term in the square brackets from equation (3) can be rewritten:

$$\begin{aligned} [\Delta^{Y_l-Y_d}_{d\to l}\omega_{d\to l} + \Delta^{Y_c-Y_l}_{l\to c}\omega_{l\to c}] &= [\Delta^{Y_l-Y_d}_{d\to l}(\omega_{l\to s} + \omega_{l\to c}) + \Delta^{Y_c-Y_l}_{l\to c}\omega_{l\to c}] \\ &= [\Delta^{Y_l-Y_d}_{d\to l}\omega_{l\to s} + \Delta^{Y_l-Y_d}_{d\to l}\omega_{l\to c} + (\Delta^{Y_c-Y_d}_{l\to c} - \Delta^{Y_l-Y_d}_{l\to c})\omega_{l\to c}] \\ &= [\Delta^{Y_l-Y_d}_{d\to l}\omega_{l\to s} + \Delta^{Y_c-Y_d}_{l\to c}\omega_{l\to c} + (\Delta^{Y_l-Y_d}_{d\to l} - \Delta^{Y_l-Y_d}_{l\to c})\omega_{l\to c})]. \end{aligned} \tag{5}$$

So, equation (3) can be written as:

$$
\begin{aligned}
E[Y(T(z_c', z_s, z_l)) - Y(T(z_c, z_s, z_l))] = & \\
& \Delta_{d\to c}^{Y_c - Y_d}\omega_{d\to c} + \Delta_{s\to c}^{Y_c - Y_d}\omega_{d\to s} + \Delta_{l\to c}^{Y_c - Y_d}\omega_{l\to c} \\
& + (\Delta_{d\to s}^{Y_s - Y_d} - \Delta_{s\to c}^{Y_s - Y_d})\omega_{d\to s} \\
& + (\Delta_{d\to l}^{Y_l - Y_d} - \Delta_{l\to c}^{Y_l - Y_d})\omega_{l\to c} \\
& + \Delta_{l\to s}^{Y_s - Y_l}\omega_{l\to s} + \Delta_{s\to c}^{Y_c - Y_s}\omega_{l\to s} + \Delta_{d\to l}^{Y_l - Y_d}\omega_{l\to s}.
\end{aligned} \tag{6}
$$

Next, the last row of equation (6) can be rewritten as:

$$
\begin{aligned}
& \Delta_{l\to s}^{Y_s - Y_l}\omega_{l\to s} + \Delta_{s\to c}^{Y_c - Y_s}\omega_{l\to s} + \Delta_{d\to l}^{Y_l - Y_d}\omega_{l\to s} \\
& = \Delta_{l\to s}^{Y_s - Y_l}\omega_{l\to s} + (\Delta_{s\to c}^{Y_c - Y_d}\omega_{l\to s} - \Delta_{s\to c}^{Y_s - Y_d}\omega_{l\to s}) + \Delta_{d\to l}^{Y_l - Y_d}\omega_{l\to s} \\
& = \Delta_{l\to s}^{Y_s - Y_l}\omega_{l\to s} + \Delta_{s\to c}^{Y_c - Y_d}\omega_{l\to s} - (\Delta_{s\to c}^{Y_s - Y_l}\omega_{l\to s} + \Delta_{s\to c}^{Y_l - Y_d}\omega_{l\to s}) + \Delta_{d\to l}^{Y_l - Y_d}\omega_{l\to s} \\
& = \Delta_{s\to c}^{Y_c - Y_d}\omega_{l\to s} + (\Delta_{l\to s}^{Y_s - Y_l} - \Delta_{s\to c}^{Y_s - Y_l})\omega_{l\to s} + (\Delta_{d\to l}^{Y_l - Y_d} - \Delta_{s\to c}^{Y_l - Y_d})\omega_{l\to s}.
\end{aligned} \tag{7}
$$

Rewriting equation (6), we get:

$$
\begin{aligned}
E[Y(T(z_c', z_s, z_l)) - Y(T(z_c, z_s, z_l))] = & \\
& \Delta_{d\to c}^{Y_c - Y_d}\omega_{d\to c} + \Delta_{s\to c}^{Y_c - Y_d}\omega_{d\to s} + \Delta_{l\to c}^{Y_c - Y_d}\omega_{l\to c} + \Delta_{s\to c}^{Y_c - Y_d}\omega_{l\to s} \\
& + (\Delta_{d\to s}^{Y_s - Y_d} - \Delta_{s\to c}^{Y_s - Y_d})\omega_{d\to s} \\
& + (\Delta_{d\to l}^{Y_l - Y_d} - \Delta_{l\to c}^{Y_l - Y_d})\omega_{l\to c} \\
& + (\Delta_{l\to s}^{Y_s - Y_l} - \Delta_{s\to c}^{Y_s - Y_l})\omega_{l\to s} + (\Delta_{d\to l}^{Y_l - Y_d} - \Delta_{s\to c}^{Y_l - Y_d})\omega_{l\to s}.
\end{aligned} \tag{8}
$$

And the first row of equation (8) can be rewritten as:

$$
\begin{aligned}
& \Delta_{d\to c}^{Y_c - Y_d}\omega_{d\to c} + \Delta_{s\to c}^{Y_c - Y_d}\omega_{d\to s} + \Delta_{l\to c}^{Y_c - Y_d}\omega_{l\to c} + \Delta_{s\to c}^{Y_c - Y_d}\omega_{l\to s} \\
= & \Delta_{d\to c}^{Y_c - Y_d}\omega_{d\to c} + \Delta_{s\to c}^{Y_c - Y_d}(\omega_{d\to s} + \omega_{l\to s}) + \Delta_{l\to c}^{Y_c - Y_d}\omega_{l\to c} \\
= & \Delta_{d\to c}^{Y_c - Y_d}\omega_{d\to c} + \Delta_{s\to c}^{Y_c - Y_d}\omega_{s\to c} + \Delta_{l\to c}^{Y_c - Y_d}\omega_{l\to c}.
\end{aligned} \tag{9}
$$

Equation (3) can thus be expressed in terms of $d \to c$ treatment effects (first line of equation (10)) and differences in the same treatment effects between different subgroups (remaining lines of equation (10)):

$$
\begin{aligned}
& E[Y(T(z_c', z_s, z_l)) - Y(T(z_c, z_s, z_l))] = \\
& \underbrace{\Delta_{d\to c}^{Y_c - Y_d}\omega_{d\to c} + \Delta_{s\to c}^{Y_c - Y_d}\omega_{s\to c} + \Delta_{l\to c}^{Y_c - Y_d}\omega_{l\to c}}_{\text{Weighted } d \to c \text{ treatment effects}} \\
& \underbrace{\begin{aligned}
& + (\Delta_{d\to s}^{Y_s - Y_d} - \Delta_{s\to c}^{Y_s - Y_d})\omega_{d\to s} \\
& + (\Delta_{d\to l}^{Y_l - Y_d} - \Delta_{l\to c}^{Y_l - Y_d})\omega_{l\to c} \\
& + (\Delta_{l\to s}^{Y_s - Y_l} - \Delta_{s\to c}^{Y_s - Y_l})\omega_{l\to s} \\
& + (\Delta_{d\to l}^{Y_l - Y_d} - \Delta_{s\to c}^{Y_l - Y_d})\omega_{l\to s}
\end{aligned}}_{\text{Differences in subgroup treatment effects}}
\end{aligned} \tag{10}
$$

Next, the denominator of the Wald estimator will be given by:

$$
E[T_C(z_c', z_s, z_l) - T_C(z_c, z_s, z_l)] = \omega_{d\to c} + \omega_{s\to c} + \omega_{l\to c}. \tag{11}
$$

Finally, dividing equation (10) by equation (11), we end up with two terms. The first term is a weighted average of margin-specific treatment effects of moving from $T = d$ to $T = c$ for three groups of compliers. The weights here are all weakly positive and sum to one. The second term is a weighted average of the four bias terms, where each term is the difference in the treatment effect of a given margin for two different sets of compliers, and the weights are weakly positive.[57] This implies that the bias will depend on the heterogeneity of treatment effects. For example, under a constant effects assumption, the bias terms are all zero.

Note that this expression parallels the expression derived in Appendix C.1 where we have a proper weighted average of the margin-specific effects of interest and an additive weighted bias term, where the size of the bias depends on how heterogeneous the margin-specific treatment effects are.

C.3 Interpreting conditional 2SLS estimates

In the main paper, we consider the comparison of two judges that have the same stringency on one margin, but different stringencies on another margin. For example, for the Wald estimands, we consider two judges that have the same incarceration stringency $Z_i = z_i$, but different conviction stringencies Z_c. Here, we consider what the IV estimand identifies when exclusion, random assignment, relevance, and the conditional pairwise monotonicity (CPM) assumptions hold, and what changes when swapping out CPM for the unordered partial monotonicity assumption (UPM). Specifically, we consider the case where we first condition on a set of judges who have the same incarceration stringency $Z_i = z_i$ but potentially differ in their conviction stringency. We assume Z_c can take on values $\{z_c^0, ..., z_c^K\}$ where the set is ordered such that $z_c^k \leq z_c^{k'}$ if $k \leq k'$.

In Appendix C.1, we derive the Wald estimand when comparing two judges with the same incarceration stringency but different conviction stringencies. This gives us:

$$\text{Wald}(z_c', z_c|z_i) =$$
$$\frac{E[Y(T(z_c', z_i)) - Y(T(z_c, z_i))]}{E[T_c(z_c', z_i) - T_c(z_c, z_i)]} = \frac{E[Y|Z_c = z_c', Z_i = z_i] - E[Y|Z_c = z_c, Z_i = z_i]}{E[T_c|Z_c = z_c', Z_i = z_i] - E[T_c|Z_c = z_c, Z_i = z_i]} =$$
$$\underbrace{\frac{\omega_{d\to c}\Delta_{d\to c}^{Y_c - Y_d} + \omega_{i\to c}\Delta_{d\to i}^{Y_c - Y_d}}{\omega_{d\to c} + \omega_{i\to c}}}_{\text{Weighted avg. of } Y_c - Y_d \text{ treatment effects}} + \underbrace{\frac{\omega_{i\to c}}{\omega_{d\to c} + \omega_{i\to c}}\left[\Delta_{d\to i}^{Y_i - Y_c} - \Delta_{i\to c}^{Y_i - Y_c}\right]}_{\text{Bias term}}.$$

Now, we derive what is identified in this setting by IV when using judges with varying conviction stringency but the same incarceration stringency. For notational simplicity, we leave the conditioning on $Z_i = z_i$ implicit throughout this derivation. The IV estimand is given by: $\alpha^{IV} = \frac{E[Y(Z_c - E[Z_c])}{E[T_c(Z_c - E[Z_c]]} = \frac{cov(Y, Z_c)}{cov(T_c, Z_c)}$ Following Imbens and

[57] As discussed above, $\omega_{s\to c} = \omega_{l\to s} + \omega_{d\to s}$ and $\omega_{d\to l} = \omega_{l\to s} + \omega_{l\to c}$. With these two identities, it is straightforward algebra to show that $\omega_{d\to c} + \omega_{s\to c} + \omega_{l\to c} = \omega_{d\to s} + \omega_{l\to c} + \omega_{l\to s} + \omega_{l\to s}$, making the second term a weighted average of the four bias terms.

Angrist (1994) closely, first consider the numerator:

$$
\begin{aligned}
E[Y\cdot(Z_c-E[Z_c])] &= \sum_{l=0}^{K}\lambda_l E[Y|Z_c=z_c^l](z_c^l-E[Z_c]) \\
&= \sum_{l=0}^{K}\lambda_l E[Y|Z_c=z_c^0](z_c^l-E[Z_c]) \\
&+ \sum_{l=1}^{K}\lambda_l\sum_{k=1}^{l}\Big(E[Y|Z_c=z_c^k]-E[Y|Z_c=z_c^{k-1}]\Big)\,(z_c^l-E[Z_c]) \\
&= \sum_{k=1}^{K}\left(\Big(E[Y|Z_c=z_c^k]-E[Y|Z_c=z_c^{k-1}]\Big)\sum_{l=k}^{K}\lambda_l(z_c^l-E[Z_c])\right) \\
&= \sum_{k=1}^{K}\text{Wald}(z_c^k,z_c^{k-1}|z_i)\left(\Big(E[T_c|Z_c=z_c^k]-E[T_c|Z_c=z_c^{k-1}]\Big)\sum_{l=k}^{K}\lambda_l(z_c^l-E[Z_c])\right)
\end{aligned}
$$

Next, the denominator using a similar set of steps can be written as:

$$
\begin{aligned}
E[T_c(Z_c-E[Z_c])] &= \sum_{l=0}^{K}\lambda_l E[T_c|Z_c=z_c^l](z_c^l-E[Z_c]) \\
&= \sum_{k=1}^{K}\left(\Big(E[T_c|Z_c=Z_c^k]-E[T_c|Z_c=Z_c^{k-1}]\Big)\sum_{l=k}^{K}\lambda_l(z_c^l-E[Z_c])\right)
\end{aligned}
$$

Putting these together, we get:

$$
\alpha^{IV}=\sum_{k=1}^{K}\theta_{k,k-1}\text{Wald}(z_c^k,z_c^{k-1}|z_i)
$$

where

$$
\theta_{k,k-1}=\frac{\big(E[T_c|Z_c=z_c^k]-E[T_c|Z_c=z_c^{k-1}]\big)\sum_{l=k}^{K}\lambda_l(z_c^l-E[Z_c]))}{\sum_{k=1}^{K}\Big(E[T_c|Z_c=z_c^k]-E[T_c|Z_c=z_c^{k-1}]\Big)\sum_{l=k}^{K}\lambda_l(z_c^l-E[Z_c])}.
$$

Other than the implicit conditioning on $Z_i=z_i$, this formula is the same as the formula derived in Imbens and Angrist (1994), but the Wald estimand may not always be a pairwise LATE as in Imbens and Angrist (1994). Under the CPM assumption and other standard IV assumptions, the Wald estimand recovers the term given in equation (7) in Section 3. Thus, rather than a weighted average of pairwise local-average treatment effects, we recover a weighted average of the potentially biased margin-specific local average treatment effects. Under the stronger UPM assumption, or under a constant-effects assumption, equation (7) reduces down to a standard margin-specific LATE as in Imbens and Angrist (1994) and the conditional 2SLS estimand can be interpreted as a positively-weighted average of LATEs where the weights sum to one.

Based on these results, a natural path forward would be to estimate separate 2SLS regressions, conditional on each value of Z_i. Angrist and Pischke (2009) propose doing this in a single 2SLS regression where the instrument Z_c is interacted with all possible values of Z_i. They refer to this as the "saturate and weight" approach. However, in finite samples, this approach can result in many weak instruments and the problems that arise in such setting (Angrist and Pischke, 2009; Blandhol et al., 2022).

Table C.1 shows estimates where the treatment and instrument have been interacted with the other judge stringency. Some caution should be taken in interpreting these estimates, as splitting our sample into thirds quickly leads to large standard errors and small first-stage F-statistics. We report four specifications that include increasingly rich sets of controls which are described in the table notes. Across all specifications for the impacts of conviction, the majority of estimates are positive, nearly all estimates are positive when including richer controls, and all negative estimates are statistically insignificant with very large standard errors. Across estimates we see very similar trends with large impacts of conviction in the first year that accumulate over time.

Table C.1: The impacts of conviction and incarceration on recidivism: interacting treatment and instruments with stringency bins

	Impacts of conviction with incarceration stringency bins				Impacts of incarceration with dismissal stringency bins			
	(1) **Year 1**	(2) **Year 2-4**	(3) **Year 5-7**	(4) **Year 1-7**	(5) **Year 1**	(6) **Year 2-4**	(7) **Year 5-7**	(8) **Year 1-7**
Specification 1								
Convict x bottom 3rd	0.121	-0.048	0.118	0.161	-0.016	-0.108	0.039	-0.114
	(0.075)	(0.140)	(0.108)	(0.165)	(0.067)	(0.088)	(0.089)	(0.117)
Convict x middle 3rd	0.263**	0.184	0.290	0.574**	-0.187**	0.023	-0.050	-0.081
	(0.125)	(0.281)	(0.180)	(0.289)	(0.073)	(0.096)	(0.091)	(0.134)
Convict x top 3rd	0.345	-1.060	0.463	-0.106	-0.025	0.018	0.053	0.091
	(0.440)	(0.966)	(0.643)	(0.949)	(0.054)	(0.093)	(0.075)	(0.115)
Specification 2								
Convict x bottom 3rd	0.112	-0.042	0.157	0.206	-0.017	-0.105	0.014	-0.137
	(0.089)	(0.148)	(0.129)	(0.195)	(0.071)	(0.093)	(0.094)	(0.124)
Convict x middle 3rd	0.275*	0.114	0.372	0.628*	-0.194**	0.006	-0.039	-0.099
	(0.154)	(0.306)	(0.229)	(0.357)	(0.078)	(0.105)	(0.096)	(0.147)
Convict x top 3rd	0.224	-0.981	0.460	-0.053	-0.026	0.057	0.065	0.126
	(0.393)	(0.752)	(0.600)	(0.865)	(0.058)	(0.096)	(0.079)	(0.120)
Specification 3								
Convict x bottom 3rd	0.073	0.111	0.131	0.275**	-0.073	-0.098	0.026	-0.146
	(0.055)	(0.086)	(0.088)	(0.121)	(0.076)	(0.081)	(0.080)	(0.110)
Convict x middle 3rd	0.109	0.122	0.218*	0.391**	-0.271	-0.007	-0.044	-0.158
	(0.075)	(0.122)	(0.117)	(0.165)	(0.218)	(0.208)	(0.204)	(0.295)
Convict x top 3rd	0.112	-0.067	0.236	0.266	-0.041	0.063	-0.034	0.055
	(0.101)	(0.148)	(0.156)	(0.227)	(0.141)	(0.177)	(0.155)	(0.234)
Specification 4								
Convict x bottom 3rd	0.032	-0.007	0.008	0.030	-0.103***	-0.069*	-0.058*	-0.141***
	(0.028)	(0.043)	(0.038)	(0.054)	(0.023)	(0.036)	(0.032)	(0.044)
Convict x middle 3rd	0.021	0.023	0.008	0.054	-0.065**	-0.042	-0.045	-0.094*
	(0.034)	(0.051)	(0.048)	(0.068)	(0.027)	(0.041)	(0.036)	(0.052)
Convict x top 3rd	0.057*	0.032	0.085**	0.097	-0.004	0.079**	0.068*	0.138***
	(0.030)	(0.049)	(0.042)	(0.061)	(0.025)	(0.040)	(0.036)	(0.052)
Observations	183381	183381	183381	183381	183381	183381	183381	183381

Note: This table shows 2SLS estimates of the impact of conviction and incarceration on future charges. For conviction, each specification interacts conviction and conviction stringency with residualized incarceration stringency terciles. For incarceration, each specification interacts incarceration and incarceration stringency with residualized dismissal stringency terciles. Specification 1 includes our standard set of fixed effects: court-by-year, court-by-month of year, and day-of-week dummies. Specification 2 replaces court-by-year and court-by-month of year dummies with court-by-year-by-month of year dummies. As the tercile interactions only condition on three bins of incarceration or dismissal stringency, Specification 3 further adds dummies for deciles of residualized judge incarceration or dismissal stringency. The final specification replaces the conviction instrument or incarceration instrument interacted with residualized incarceration or dismissal stringency terciles with judge dummies. Standard errors are clustered at the judge-year level. The sample is restricted to cases observed for 7 years. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

C.4 Average UPM

UPM($Z_c|Z_i$) represents a form of "strict" monotonicity, in that it is defined over every z_c shift, holding z_i constant. Yet, similar to what has been shown in the binary context, such a strict assumption is not necessary to yield a causal estimand. Frandsen et al.

(2023) propose a condition called "average monotonicity," which requires a positive correlation between each individual's *potential* treatment status and judge stringency across all judges. They show that average monotonicity is sufficient (along with other standard IV assumptions) to yield a causal estimand in the binary treatment context.

Here we propose an extension of Frandsen et al.'s (2023) average monotonicity condition into the three treatment setting and refer to this as "average UPM($Z_c|Z_i$)." We focus on the condition that is relevant to the specification where we are instrumenting for conviction and controlling for the incarceration stringency; average UPM($Z_i|Z_d$) is defined similarly.

We first introduce an additional piece of notation. Let G be a group variable where $g \in G$ maps (Z_c, Z_i) onto potential treatment $T_c(Z_c, Z_i)$. G is the collective and mutually exclusive set of groups g. In the binary treatment, binary instrument context, G consists of compliers, defiers, always takers, and never takers.

A5b: Average UPM($Z_c|Z_i$): For all (g, z_i) in the support of (G, Z_i) the following conditions must hold:

$$Cov(T_c(Z_c, Z_i), Z_c|Z_i = z_i, G = g) \geq 0 \tag{12}$$

$$Cov(T_i(Z_c, Z_i), Z_c|Z_i = z_i, G = g) = 0 \tag{13}$$

To illustrate a difference between UPM($Z_c|Z_i$) and average UPM($Z_c|Z_i$), consider a shift from z_c to $z_c' > z_c$, holding z_i constant. If there exists a group g for whom this instrument shift would induce them from conviction to dismissal, UPM($Z_c|Z_i$) would be violated but average UPM($Z_c|Z_i$) might not be. As long as the probability of conviction for each group is positively correlated with the overall conviction propensity of judges, average UPM($Z_c|Z_i$) is satisfied.

Average UPM($Z_c|Z_i$), along with A1-A4, is sufficient for equations (3) and (4) to yield margin-specific and causal estimands. We build off of Blandhol et al. (2022) for the proof. First, note that the second line of A5b, combined with A2 and A3 (random assignment and exclusion) assure that the exogeneity condition outlined in Blandhol et al. (2022) is met. In our setting, this exogeneity condition means that $G, Y(T = c) \perp Z_c|Z_i$. G is orthogonal to Z_c (conditional on Z_i) due to the random assignment assumption. $Y(T = c)$ is orthogonal to Z_c because, if you hold Z_i fixed, Z_c will not be correlated with the probability of incarceration for any group.

With exogeneity in hand, the remainder of the proof is provided by Blandhol et al. (2022). Blandhol et al. (2022) focus on a condition they call "monotonicity-correct," which they show is sufficient for the 2SLS estimator with covariates to be weakly causal (i.e., the weights on all group-specific treatment effects are weakly positive and the estimate does not depend on the levels of the dependent variable). In the appendix, they derive the monotonicity condition that is both sufficient and necessary for weakly causal estimates, which is the condition in line one of A5b, when written in our notation and in the terms relevant to our setting.[58] They do not focus on this condition in the

[58]The necessary and sufficient condition for weakly causal estimates is presented in the paragraph between equation (28) and equation (29) in the appendix proof for Proposition 9 (page 50) of the version from August 9, 2022. Our Z_c would be written $\dot{Z}$ in their notation, our Z_i would be their X, and our $T_c^g(Z_c)$ would be

main text because "such fortuitous averaging would be difficult to defend." In the judge IV context, however, this "fortuitous averaging" could naturally occur. For instance, a judge who punishes harshly overall may be relatively lenient on certain types of offenders. This would violate both the monotonicity-correct condition as well as UPM. But as long as relatively harsh judges increase punishment *on average* for all groups, an occasional judge who bucks the trend for certain groups is not a problem.

C.5 Interpreting 2SLS estimates with controls

Appendix section C.3 derived the 2SLS estimand when conditioning on a specific value of Z_i. The estimation results reported in Section 4 control for Z_i rather than condition. This section discusses how to interpret these 2SLS estimates. In particular, following Blandhol et al. (2022), 2SLS specifications that control for Z_i (and potentially other covariates) can still be interpreted as a positively-weighted sum of the Wald estimates we derived in Section 3, as long as one additional assumption is met.

Blandhol et al. (2022) considers what 2SLS recovers when covariates are included as controls, but are not fully saturated as in the "saturate and weight" approach. They show that covariates can introduce substantial bias and result in estimands that are not what they call "weakly causal." They define an estimand as weakly causal when it (i) does not depend on the levels of the potential outcomes when holding treatment effects (differences) constant and (ii) it does not apply negative weights to any subgroup. Blandhol et al. (2022) goes on to discuss what assumptions are necessary and sufficient for 2SLS with controls to recover weakly causal parameters. For our setting, with a scalar multi-valued instrument, one additional assumption needs to hold:[59]

A4b. Rich covariates: The linear projection of Z on X is equal to the conditional expectation of Z given X. That is $L[Z|X] = X'E[XX']^{-1}E[XZ] = E[Z|X]$.

Assumption A4b implies that we need to include a rich set of controls. Note that assumption A4b differs from assumption A4 as Section 3.2 abstracted away from covariates. Here we provide the more general version of the assumption, which allows for other covariates. When the only covariate is Z_i, we need rich controls for Z_i. When instruments are only randomly assigned conditional on a vector of covariates $\mathbf{X}$, then we must include a sufficiently rich set of controls for the full vector of covariates, including Z_i.

Blandhol et al.'s (2022) Proposition 11 provides an expression for what the 2SLS estimand recovers. A small rearrangement of that expression allows it to be written as a positively-weighted average of Wald estimands. Under assumptions A1-A5 or A1-A4 and A6, these Wald estimands are equivalent to those we derive in Section 3.4. Thus, under assumptions A1-A3, A4b and A6, 2SLS recovers a positively-weighted average of terms that are margin-specific causal effects plus additive bias terms. Under assumptions A1-A3, A4b, and A5, 2SLS recovers a positively-weighted average of

$\mathbb{1}(Z \in \mathbb{Z}_j(g))$.

[59]Note that assumptions A1-A3, and A5 satisfy the other needed assumptions in Blandhol et al. (2022). In particular, A5 implies their "Ordered strong monotonicity" (OSM). Assumption A6 also satisfies the OSM, but violates their definition of exclusion, which can result in biased Wald estimates, similar to those we derive under CPM.

margin-specific treatment effects.

Table C.2 shows that our estimates are not sensitive to the richness of our control variables. Each specification adds increasingly detailed sets of dummies for place, time, and the other judge stringency as described in the table notes. All specifications are similar to the estimates we report in the main paper, and trend towards larger estimates when including richer set of controls. Like our main estimates, we find large increases in recidivism from conviction that accumulate over time, while incarceration has a negative effect in the first year, which remains relatively constant when looking at one year, one to four year, or one to seven years.

Table C.2: The impacts of conviction and incarceration on recidivism: robustness to richness of controls

	Impacts of conviction				Impacts of incarceration			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Year 1	**Year 2-4**	**Year 5-7**	**Year 1-7**	**Year 1**	**Year 2-4**	**Year 5-7**	**Year 1-7**
Specification 1								
Fut. charge	0.100**	0.123	0.104	0.290***	-0.100***	-0.048	-0.011	-0.099*
	(0.051)	(0.083)	(0.080)	(0.109)	(0.029)	(0.046)	(0.041)	(0.060)
Fut. conviction	0.134***	0.159**	0.074	0.354***	-0.113***	-0.067	0.005	-0.133**
	(0.048)	(0.079)	(0.076)	(0.106)	(0.028)	(0.046)	(0.039)	(0.058)
Fut. incarceration	0.101**	0.084	-0.005	0.251***	-0.076***	-0.021	0.035	-0.059
	(0.042)	(0.069)	(0.061)	(0.093)	(0.024)	(0.039)	(0.032)	(0.050)
Specification 2								
Fut. charge	0.089	0.151	0.154	0.350***	-0.103***	-0.041	-0.017	-0.105*
	(0.060)	(0.094)	(0.095)	(0.128)	(0.031)	(0.050)	(0.044)	(0.064)
Fut. conviction	0.129**	0.188**	0.123	0.425***	-0.118***	-0.062	0.002	-0.139**
	(0.057)	(0.090)	(0.091)	(0.126)	(0.030)	(0.049)	(0.042)	(0.062)
Fut. incarceration	0.106**	0.107	0.035	0.327***	-0.078***	-0.025	0.031	-0.073
	(0.048)	(0.078)	(0.074)	(0.110)	(0.026)	(0.042)	(0.034)	(0.054)
Specification 3								
Fut. charge	0.116**	0.153*	0.097	0.320***	-0.100***	-0.038	-0.015	-0.094
	(0.054)	(0.089)	(0.081)	(0.115)	(0.031)	(0.047)	(0.042)	(0.062)
Fut. conviction	0.153***	0.192**	0.071	0.403***	-0.116***	-0.058	-0.008	-0.131**
	(0.053)	(0.085)	(0.077)	(0.113)	(0.030)	(0.046)	(0.040)	(0.060)
Fut. incarceration	0.119***	0.120	-0.021	0.292***	-0.075***	-0.018	0.023	-0.068
	(0.044)	(0.074)	(0.063)	(0.099)	(0.025)	(0.039)	(0.033)	(0.053)
Specification 4								
Fut. charge	0.108	0.191*	0.148	0.391***	-0.103***	-0.043	-0.035	-0.119*
	(0.067)	(0.103)	(0.099)	(0.140)	(0.034)	(0.052)	(0.045)	(0.066)
Fut. conviction	0.155**	0.225**	0.119	0.487***	-0.121***	-0.065	-0.025	-0.158**
	(0.065)	(0.099)	(0.095)	(0.140)	(0.033)	(0.050)	(0.043)	(0.064)
Fut. incarceration	0.130**	0.146*	0.013	0.378***	-0.078***	-0.036	0.009	-0.102*
	(0.054)	(0.086)	(0.077)	(0.122)	(0.028)	(0.043)	(0.036)	(0.057)
Observations	183371	183371	183371	183371	183371	183371	183371	183371

Note: This table reports estimates of the impact of conviction and incarceration on our three measures of recidivism. Each specification adds richer controls. Specification 1 includes the fixed effects included in the paper: court-by-year, court-by-month of year, and day of week dummies, plus percentile dummies for residualized judge incarceration or dismissal stringency. Specification 2 matches specification 1 but swaps out court-by-year and court-by-month of year fixed effects with court-by-year-by-month of year fixed effects. Specification 3 includes the main place and location fixed effects plus year-by-decile of residualized incarceration or dismissal stringency dummies. Specification 4 is the same as specification three, but swaps out year-by-court and year-by-month of year dummies with court-by-year-by-month of year dummies. Standard errors are clustered at the judge-year level. The sample is restricted to cases observed for 7 years. Stars denote * $p< 0.10$, ** $p< 0.05$, *** $p< 0.01$.

D Validating assumptions A1-A4

In this section, we discuss whether Assumptions A1-A4 from Section 3 are supported by features of the institutional environment and provide empirical evidence, based on a standard battery of tests, to help assess their validity.

Relevance. Here, we explain the various ways through which judges can influence both conviction and incarceration outcomes, expanding on Section 2.2. We also present empirical evidence of the relevance of judges' influence on these decisions.

Judges influence conviction in several ways. In all cases, they have the latitude to dismiss charges if they find the evidence insufficient. They are directly responsible for adjudicating guilt during bench trials (that is, trials by judge, without lay jurors). They also exert indirect influence on the likelihood of conviction through multiple channels. First, they make the determination on various pretrial motions, which can have a large impact on the likelihood of conviction. For example, they can refuse to grant a continuance if a key witness does not show up to court on a given day. They rule on the admissibility of evidence, including critical pieces like confessions, possession of contraband, or expert testimony. Finally, they can affect jury composition by ruling on motions to strike and by formulating jury instructions.

Judges also influence sentences in several ways. In the case of a bench trial, they directly choose the sentence. In the case of guilty pleas, they can reject the negotiated plea agreement. Moreover, their reputation as a tough or lenient judge might shape what offers prosecutors and defense attorneys are willing to put forward (LaCasse and Payne, 1999). For example, if the judge has a reputation for choosing short sentences, the prosecutor may adjust and offer shorter sentences as part of the plea deal.

Empirically, we find persistent differences in case outcomes across judges. Panels A and B of Figure 4 in the main paper shows the histogram of judge noncarceral conviction stringency (Panel A) and judge incarceration stringency (Panel B). Each panel plots the residualized leave-one-out judge propensity for that case outcome. In both panels there is substantial variation in the instrument.[60] Both panels also plot the local linear regression of the residualized court outcome on the instrument.

Panel C of Figure 4 plots the residualized noncarceral conviction and incarceration stringencies against each other. The two instruments are negatively correlated, which is expected, since the probability of all three case outcomes adds up to one. Importantly for our research design, there is substantial variation in Z_c across most of the support of Z_i and vice versa.

Table 2 in the main paper presents our first-stage estimates, and confirms that judge stringency has a large and statistically significant effect on conviction and incarceration. The first three columns show the results for the first stage on noncarceral conviction. The first column shows the loading on conviction stringency when only including interacted court and time fixed effects as controls. The second column adds detailed case-level controls. The third column additionally controls for incarceration stringency. Across all three specifications, the conviction stringency remains large, with partial F-statistics between 165 and 360. Columns 4 through 6 perform similar first-stage regressions on incarceration stringency, with the sixth column controlling

[60] Conviction stringency was constructed by residualizing an indicator for noncarceral conviction against county-by-year, county-by-month-of-year, and day-of-week fixed effects, then constructing leave-one-out averages at the judge-by-three-year level. Incarceration stringency is similarly constructed.

for dismissal stringency. Again, the loading on incarceration stringency is large and statistically significant, with partial F-statistics between 288 and 351.

Random assignment. As discussed in Section 2.1, within our sample, cases are quasi-randomly assigned to judges within court. There is either actual randomization, or case assignment is done based on scheduling or judge availability.[61] In addition, we confirm empirically that judge stringency is largely not predicted by case characteristics. In Table 3 of the main paper, we show that case characteristics are strong predictors of being convicted and of being incarcerated (columns 1 and 3). We then show that case characteristics largely do not predict with judge conviction stringency (column 2) or incarceration stringency (column 4). For the few instances where covariates have statistically significant loadings, the predicted difference in stringency tends to be very small. Table D.1 replicates columns (2) and (4) from Table 3 but using standardized stringency measures. The odd columns regress non-carceral conviction stringency and incarceration stringency on case characteristics where the stringency measure has been standardized to have a mean of zero and a standard deviation of one. We see that the largest loading is on an indicator for assault cases which predicts assault cases are associated with a 0.015 standard deviation change in stringency. The odd columns do not account for variation in stringency caused by variation over time or across courts. The even columns replicate the regressions from the odd columns but first residualize the stringency instruments for the set of district and time fixed effects used in our analysis before standardizing. Here we find the largest coefficient in absolute value to be 0.036. Overall, this suggests that while there are a few instances where covariates have statistically significant loadings, these loadings imply small predicted differences in stringency.

We additionally provide robustness showing that our results are not sensitive to fully excluding certain types of cases from our analysis, that is, both from the construction of the stringency instruments and from the 2SLS regressions. Table D.2 provides our main OLS and IV estimates for noncarceral conviction for four different subsets of cases. In Panel A we drop all cases involving assault charges when constructing the instrument and running the analyses, as this is the offense type that is most predictive of both noncarceral conviction and incarceration stringency in our balance tables. These results are broadly similar to our main estimates in Table 4 with the same sign and magnitude, with the two main differences being that point estimates are moderately smaller, and standard errors are somewhat larger (likely due to the 15% reduction in sample size).

Panel B and C repeat the prior exercise, but throw out cases with drug offenses and cases with violent offenses, respectively. We focus on drug offenses and violent offenses since these are offense types where we believe judges may be most likely to differ in opinion on appropriate case outcome. We again find that dropping these offense types lead to broadly similar results, with similar point estimates and somewhat larger standard errors. Finally, Panel D drops cases with assault, sexual assault, fraud, or traffic charges (all offense types where there is any evidence of imbalance in Table 3). Estimates again are broadly similar. For this specification, we lose statistical significance on several coefficients that are significant in our main table. This may

[61] In Appendix E, we show that IV estimates are similar when we remove courts where assignment is by judge availability.

be in part due to moderately smaller (though similar in magnitude) estimates in Year 1-7, but is largely driven by larger standard errors, likely because of the 39% reduction in sample size. Table D.3 replicates the analysis in Table D.2, but for incarceration. Similar to results for conviction, results are broadly similar. Overall, Tables D.2 and D.3 suggest that our results are not driven by potential exclusion violations.

In our general robustness analysis in Appendix Section E, we compare how our estimates vary under several different assumptions. There we additionally include results in Figures E.3-E.6 where we use the full sample, but allow judge stringency to differ by (1) if the case has an assault charge or not and (2) if the case has a drug charge or not. These alternative constructions of our instrument are more demanding on our data, but also find statistically significant increases in recidivism from non-carceral conviction 1-7 years after the case, and statistically significant decreases from incarceration only in the first year after the case.

Finally, as additional evidence of exogeneity, first-stage estimates barely change when we add controls to our first-stage regression, as seen by comparing columns 2 and 3 and columns 5 and 6 of Table 2 in the main paper.

Exclusion. Our identification strategy relies on the assumption that the conviction stringency instrument affects recidivism outcomes only through its effects on conviction once we control for judges' incarceration stringency, and vice versa. Here we argue that the risk of potential exclusion violations is low. We consider sentence length to be the most important potential violation. For example, if a high-conviction judge also tends to give longer sentences (holding incarceration probability fixed) it would violate exclusion. We test for this by regressing sentence length on our measure of conviction stringency, controlling for incarceration stringency. As shown in Appendix Table D.4, we find no evidence of a violation of the exclusion restriction for conviction. In addition, when we re-estimate the main IV regressions with an additional control for sentence length stringency or probability of sentence length shorter than 6 months and longer than 1 year and 4 years, we find that the main conclusions are unchanged (see Appendix Figures E.3-E.6).[62]

A judge may influence other aspects of the case, such as probation and parole terms, or fines and fees. While we do not rule these channels out, we do not expect them to be as important. There are a number of large-scale RCTs that have shown probation and parole conditions do not affect recidivism (for a recent review, see Doleac, 2023). There is also a small but growing literature showing that court fines and fees do not affect recidivism (Pager et al., 2022; Finlay et al., 2023; Lieberman et al., 2023). The findings in this literature add confidence that even if judge stringency in conviction and incarceration were correlated with these other factors, they would not bias the results.

We do not expect decisions made at the beginning of the case, such as bail or pretrial detention, to lead to an exclusion violation. These decisions are made by bail magistrates that have no later influence over the case. Furthermore, there is often a month between the date of the arrest and when the defendant arrives at circuit court and the judge is assigned. It follows that the Circuit Court judge has no influence over these early aspects of the defendant's criminal justice experience.

[62] We define sentence length stringency as the tri-yearly leave-one-out average sentence for the judge handling the case, setting sentences to 0 if a person has no carceral sentence and to the sentence length in months if a person is sentenced to a carceral sentence.

Although we are comfortable arguing that conviction and incarceration are likely the most important channels by which criminal justice involvement can affect recidivism, we see expanding beyond a trinary model to include these alternatives as an important area of future research. Given the tradeoffs, we have chosen tractability over complexity.

Lastly, in Appendix Table E.1, we present reduced-form estimates, which regress outcomes on our instruments, and do not require the exclusion assumption to hold.

Monotonicity. As discussed previously, one consequence of CPM (and the stronger condition, UPM) is that there will only be one-way flows across any margin. Here we present some empirical evidence in support of this assumption. Following common practice for binary treatments (see, for example, Bhuller et al., 2020 or Norris et al., 2021), we conduct split-sample regressions where the data is bifurcated using observed characteristics such as race and gender. Judge stringency is then estimated on each subsample, and the first stage regression is then run on its complement, controlling for stringency along the other margin. If the "no defiers" condition holds, we would expect positive coefficients for each sub-sample. Appendix Tables D.5 and D.6 report the coefficient on the instrument from split-sample first-stage regressions. Each row presents a particular case characteristic. For example, the first row breaks our sample into whether a person has a drug charge or does not. The "Zero" column for that row calculates the stringency on the individuals without a drug charge and then estimates the first stage on those with a drug charge, reporting the coefficient on that instrument. The "One" column does the converse of that – calculates the stringency on the individuals with a drug charge and then estimates the first stage on those without a drug charge, reporting the coefficient on that instrument. For both conviction and incarceration, we find positive coefficients on the instrument for all split-sample estimates. Also see Section 4.5 of the paper where we present a test of the UPM assumption.

Table D.1: Balance: outcomes in standard deviations

	Conv. string.	Resid. conv. string.	Incar. string.	Resid. incar. string.
	(1)	(2)	(3)	(4)
Any prior conv.	-0.0004	-0.0011	0.0031	0.0073
	(0.0025)	(0.0060)	(0.0027)	(0.0062)
Female	-0.0043*	-0.0102*	0.0025	0.0058
	(0.0023)	(0.0056)	(0.0024)	(0.0057)
Black	0.0031	0.0075	-0.0028	-0.0065
	(0.0022)	(0.0054)	(0.0022)	(0.0053)
Has misdemeanor	0.0013	0.0031	0.0041	0.0097
	(0.0037)	(0.0089)	(0.0039)	(0.0090)
Drugs	0.0045	0.0108	-0.0003	-0.0006
	(0.0032)	(0.0077)	(0.0035)	(0.0081)
Larceny	0.0035	0.0085	0.0041	0.0096
	(0.0027)	(0.0066)	(0.0029)	(0.0068)
Assault	-0.0148***	-0.0355***	0.0142***	0.0332***
	(0.0031)	(0.0075)	(0.0031)	(0.0072)
Fraud	0.0047	0.0114	0.0068*	0.0160*
	(0.0034)	(0.0082)	(0.0039)	(0.0090)
Traffic	-0.0036	-0.0088	0.0076*	0.0177*
	(0.0043)	(0.0103)	(0.0045)	(0.0104)
Burglary	-0.0016	-0.0039	0.0056	0.0132
	(0.0039)	(0.0093)	(0.0042)	(0.0098)
Robbery	-0.0026	-0.0062	0.0043	0.0101
	(0.0052)	(0.0124)	(0.0055)	(0.0128)
Sexual assault	-0.0085	-0.0205	0.0143**	0.0335**
	(0.0067)	(0.0161)	(0.0070)	(0.0163)
Kidnapping	-0.0063	-0.0151	0.0070	0.0164
	(0.0076)	(0.0182)	(0.0075)	(0.0176)
Murder	-0.0149	-0.0357	0.0118	0.0275
	(0.0108)	(0.0259)	(0.0117)	(0.0273)
F-stat joint F-test	3.757	3.757	2.666	2.666
P-value joint F-test	0.000	0.000	0.001	0.001
Observations	183,381	183,381	183,381	183,381

Note: This table replicates Table 3, but where the left-hand-side variable in the regression (i.e., either noncarceral conviction, incarceration, noncarceral conviction stringency, or incarceration stringency) has been standardized in the sample to have a mean of zero and a standard deviation of one. For each outcome, we regress the standardized outcome on case characteristics. Regression includes court-by-year fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. The offenses are ordered by their prevalence in the data. The balance outcomes shown are for those cases adjudicated in 2012 or earlier, representing our seven-year sample. Star denote * $p< 0.10$, ** $p< 0.05$, *** $p< 0.01$.

Table D.2: Noncarceral conviction and recidivism– robustness to unbalanced offenses

	Year 1		Year 2-4		Year 5-7		Year 1-7	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Panel A: No assault offenses								
Fut. charge	0.004*	0.099*	0.015***	0.062	0.011***	0.045	0.026***	0.180*
	(0.002)	(0.053)	(0.003)	(0.080)	(0.003)	(0.082)	(0.004)	(0.105)
Fut. conviction	0.006***	0.122**	0.017***	0.106	0.012***	0.038	0.028***	0.254**
	(0.002)	(0.050)	(0.003)	(0.079)	(0.003)	(0.078)	(0.004)	(0.105)
Fut. incarceration	0.005***	0.105**	0.014***	0.018	0.009***	-0.035	0.025***	0.169*
	(0.002)	(0.043)	(0.003)	(0.067)	(0.002)	(0.064)	(0.003)	(0.090)
Observations	155100	155100	155100	155100	155100	155100	155100	155100
Panel B: No drug offenses								
Fut. charge	-0.013***	0.147**	-0.010***	0.188	-0.003	0.084	-0.015***	0.334**
	(0.003)	(0.073)	(0.004)	(0.127)	(0.003)	(0.109)	(0.005)	(0.151)
Fut. conviction	-0.010***	0.207***	-0.006	0.238*	-0.001	0.085	-0.011**	0.430***
	(0.003)	(0.069)	(0.003)	(0.125)	(0.003)	(0.107)	(0.004)	(0.157)
Fut. incarceration	-0.007***	0.164***	-0.006**	0.175	-0.002	-0.044	-0.011***	0.323**
	(0.002)	(0.063)	(0.003)	(0.111)	(0.003)	(0.087)	(0.004)	(0.136)
Observations	125602	125602	125602	125602	125602	125602	125602	125602
Panel C: No violent offenses								
Fut. charge	0.004*	0.096*	0.016***	0.072	0.012***	0.045	0.028***	0.193*
	(0.002)	(0.057)	(0.003)	(0.086)	(0.003)	(0.086)	(0.004)	(0.111)
Fut. conviction	0.006***	0.126**	0.018***	0.107	0.013***	0.031	0.031***	0.263**
	(0.002)	(0.054)	(0.003)	(0.084)	(0.003)	(0.082)	(0.004)	(0.112)
Fut. incarceration	0.005***	0.110**	0.016***	0.024	0.010***	-0.046	0.028***	0.176*
	(0.002)	(0.047)	(0.003)	(0.071)	(0.002)	(0.068)	(0.003)	(0.097)
Observations	149473	149473	149473	149473	149473	149473	149473	149473
Panel D: No assault, sexual assault, fraud, or traffic offenses								
Fut. charge	0.005*	0.106	0.018***	0.011	0.009***	-0.025	0.028***	0.152
	(0.003)	(0.073)	(0.004)	(0.102)	(0.003)	(0.101)	(0.004)	(0.135)
Fut. conviction	0.006***	0.157**	0.020***	0.035	0.012***	-0.070	0.031***	0.223*
	(0.002)	(0.070)	(0.003)	(0.099)	(0.003)	(0.096)	(0.004)	(0.133)
Fut. incarceration	0.006***	0.141**	0.017***	0.000	0.009***	-0.109	0.030***	0.154
	(0.002)	(0.059)	(0.003)	(0.088)	(0.002)	(0.081)	(0.004)	(0.117)
Observations	112135	112135	112135	112135	112135	112135	112135	112135

Note: Panel A of this table shows 2SLS estimates of the impact of conviction vs dismissal on future charges, convictions, and incarcerations. Here we recalculate the instrument by assault cases (assault and weapons) and then drop the assault cases from the sample. Panel B is similar except it recalculates the stringency splitting drug and non-drug cases and drops drug cases. Panel C recalculates the instruments using all violent offenses (assault, sexual assault, and murder). Panel D includes all unbalanced offenses which includes assault, sexual assault, fraud, and traffic. The columns report results for four recidivism time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). Each time period restricts the sample to cases observed for all 7 years. All regressions control for stringency on the other margin (i.e., z_i for the conviction specification), race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

Table D.3: Incarceration and recidivism– robustness to unbalanced offenses

	Year 1		Year 2-4		Year 5-7		Year 1-7	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Panel A: No assault offenses								
Fut. charge	-0.020***	-0.090***	0.018***	-0.001	0.028***	-0.004	0.030***	-0.053
	(0.002)	(0.031)	(0.003)	(0.048)	(0.002)	(0.042)	(0.003)	(0.059)
Fut. conviction	-0.017***	-0.101***	0.019***	-0.026	0.026***	0.012	0.029***	-0.094
	(0.002)	(0.030)	(0.002)	(0.048)	(0.002)	(0.040)	(0.003)	(0.058)
Fut. incarceration	-0.009***	-0.066***	0.021***	0.008	0.024***	0.036	0.034***	-0.048
	(0.001)	(0.025)	(0.002)	(0.042)	(0.002)	(0.033)	(0.003)	(0.052)
Observations	155100	155100	155100	155100	155100	155100	155100	155100
Panel B: No drug offenses								
Fut. charge	-0.017***	-0.117***	0.015***	0.001	0.026***	0.036	0.028***	-0.038
	(0.002)	(0.040)	(0.003)	(0.064)	(0.002)	(0.053)	(0.003)	(0.080)
Fut. conviction	-0.013***	-0.138***	0.016***	-0.032	0.024***	0.036	0.028***	-0.091
	(0.002)	(0.038)	(0.003)	(0.062)	(0.002)	(0.051)	(0.003)	(0.077)
Fut. incarceration	-0.006***	-0.085***	0.019***	0.002	0.022***	0.085**	0.034***	-0.014
	(0.002)	(0.032)	(0.002)	(0.055)	(0.002)	(0.043)	(0.003)	(0.069)
Observations	125602	125602	125602	125602	125602	125602	125602	125602
Panel C: No violent offenses								
Fut. charge	-0.020***	-0.085***	0.020***	0.011	0.030***	0.003	0.033***	-0.034
	(0.002)	(0.032)	(0.003)	(0.049)	(0.002)	(0.043)	(0.003)	(0.060)
Fut. conviction	-0.017***	-0.098***	0.020***	-0.008	0.027***	0.020	0.032***	-0.075
	(0.002)	(0.031)	(0.003)	(0.049)	(0.002)	(0.041)	(0.003)	(0.058)
Fut. incarceration	-0.009***	-0.065**	0.022***	0.024	0.025***	0.042	0.035***	-0.033
	(0.001)	(0.026)	(0.002)	(0.043)	(0.002)	(0.034)	(0.003)	(0.053)
Observations	149473	149473	149473	149473	149473	149473	149473	149473
Panel D: No assault, sexual assault, fraud, or traffic offenses								
Fut. charge	-0.021***	-0.075*	0.022***	0.037	0.033***	0.037	0.036***	0.001
	(0.002)	(0.041)	(0.003)	(0.059)	(0.003)	(0.055)	(0.003)	(0.076)
Fut. conviction	-0.017***	-0.095**	0.022***	0.028	0.030***	0.080	0.035***	-0.025
	(0.002)	(0.040)	(0.003)	(0.058)	(0.002)	(0.052)	(0.003)	(0.073)
Fut. incarceration	-0.009***	-0.062*	0.024***	0.043	0.026***	0.054	0.037***	0.001
	(0.002)	(0.035)	(0.003)	(0.055)	(0.002)	(0.043)	(0.003)	(0.067)
Observations	112135	112135	112135	112135	112135	112135	112135	112135

Note: Panel A of this table shows 2SLS estimates of the impact of incarceration vs conviction on future charges, convictions, and incarcerations. Here we recalculate the instrument by assault cases (assault and weapons) and then drop the assault cases from the sample. Panel B is similar except it recalculates the stringency splitting drug and non drug cases and drops drug cases. Panel C recalculates the instruments using all violent offenses (assault, sexual assault, and murder). Panel D uses all unbalanced offenses which includes assault, sexual assault, fraud, and traffic. The columns report results for four recidivism time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). Each time period restricts the sample to cases observed for all 7 years. All regressions control for stringency on the other margin (i.e., z_i for the conviction specification), race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

Table D.4: 2SLS regressions of sentence length on conviction stringency

	Conviction reg controling for incarceration									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Sent length	Any incar	6mo	1y	2y	3y	4y	5y	6y	7y
Pr. convict	7.68	-0.033	0.088*	-0.032	-0.043	-0.027	0.0023	-0.011	-0.0071	-0.00018
	(64.2)	(0.051)	(0.047)	(0.043)	(0.033)	(0.029)	(0.025)	(0.021)	(0.017)	(0.017)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	322.018	0.546	0.374	0.203	0.113	0.078	0.061	0.042	0.035	0.030
N	183381	183381	183381	183381	183381	183381	183381	183381	183381	183381

Note: This table shows a regression of various sentence length variables on z_c. The first column uses sentence length as the outcome, the second any incarceration, third to tenth any incarceration greater than 6 months, 1 year, 2 years, 3 years, 4 years, 5 years, 6 years, and 7 years respectively. All regressions control for z_i, race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. The sample is restricted to cases observed for 7 years. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

Table D.5: Split sample monotonicity test: conviction

	Zero	One
Any drug charges	0.548	0.199
Any property charges	0.463	0.238
Any violent charges	0.430	0.099
Black	0.308	0.393
Female	0.875	0.168
Prior conviction	0.274	0.148

Note: This table shows first-stage estimates for the conviction (without incarceration) instrument where, for each regression, the stringency measure is calculated on a specific subpopulation, and the regression is then run on its complement. For example, the "Zero" column of the "Any drug charges" row calculates judge stringency on those without drug charges, then estimates the first stage on those *with* drug charges, and reports the coefficient on the instrument. Regression includes court-by-year fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. The regression also controls for the leave-one-out propensity of the judge to have cases that end in incarceration. The sample is restricted to cases observed for 7 years.

Table D.6: Split sample monotonicity test: incarceration

	Zero	One
Any drug charges	0.538	0.366
Any property charges	0.677	0.342
Any violent charges	0.319	0.191
Black	0.460	0.592
Female	0.658	0.269
Prior conviction	0.750	0.337

Note: This table shows first-stage estimates for the incarceration instrument where, for each regression, the stringency measure is calculated on a specific subpopulation, and the regression is then run on its complement. For example, the "Zero" column of the "Any drug charges" row calculates judge stringency on those without drug charges, then estimates the first stage on those *with* drug charges, and reports the coefficient on the instrument. Regression includes court-by-year fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. Regression also controls for the leave-one-out propensity of the judge to dismiss cases. The sample is restricted to cases observed for 7 years.

E Additional figures and tables: IV analyses

In this appendix, we present a series of additional analyses and robustness tests for our main IV analyses.

E.1 Overview of analyses

E.1.1 Disposition types

Disposition type by offense. Figure E.1 shows the breakdown of disposition types for four common offenses: drugs, fraud, larceny, and assault. These offense categories differ in seriousness and, while the exact breakdown varies, all disposition types are present in each offense type considered.

Future exposure to incarceration Appendix Figure E.2 illustrates the extent of "incarceration catch-up" for individuals given noncarceral sentences compared to those given carceral sentences, considering both new crimes and technical violations leading to probation revocation. These results indicate that although some catch-up occurs, over 50% of those receiving noncarceral sentences avoid incarceration over the next seven years.

E.1.2 Reduced-form estimates

Appendix Table Panel A E.1 presents reduced-form estimates, showing the relationship between our outcome variables and the conviction instrument controlling for race, gender, prior conviction, offense type dummies, and year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects as well as the leave-one-out judge incarceration stringency. We find that the instrument positively and significantly affects the year 1 and the year 1-7 outcomes. Appendix Table E.1 Panel B shows comparable reduced-form estimates for the incarceration results.

E.1.3 Compliers

Characterizing compliers. Appendix Table E.2 compares compliers for the conviction and incarceration margins to the full sample. The distribution of offenses is mostly similar for compliers to both instruments, with a few exceptions. Compliers to the conviction instrument are more likely to be female (27% vs 22%) and are more likely to have a property crime charge (42% vs 38%). Moreover, they are less likely to have a prior conviction (10% vs 17%), less likely to have a violent charge than the general sample (8% vs 19%), and less likely to have charges that fall into the other category (6% vs 16%). Compliers to the incarceration instrument are slightly more similar to the full sample, but exhibit some of the same notable differences. First, prior conviction rates and share of women are more similar, 19% vs 17% and 21% vs 22%, respectively. For property charges and violent charges we continue to see disparities with 46% vs 38% having a property charge and 8% vs 19% having a violent charge.

Complier weighted OLS. In Appendix Table E.3 we reweight the OLS for incarceration and conviction margin compliers. The OLS estimates do not change much when re-weighting for compliers. The reweighted estimates for noncarceral conviction are somewhat larger, while the estimates for incarceration are nearly identical.

E.1.4 Heterogeneity

Increased criminal behavior or "ratcheting up"? We take two strategies to provide suggestive evidence on whether the recidivism effects come from increased criminal behavior or the "ratcheting up" effect. First, we look at differences across different stages of the criminal justice process. If each discretionary decision is influenced by the criminal record, then the influence of the conviction will accumulate as someone advances through the criminal proceedings. If the ratcheting up effect is operative, it may have a larger effect on the more downstream measures of future criminal justice contact, like incarceration, than on the more upstream measures, like new charges. Consistent with this mechanism, we note that in all of our estimates presented in Table 4, the percent changes are larger for more downstream measures of future criminal justice contact.[63]

Second, we consider recidivism across crime types. Following Deshpande and Mueller-Smith (2022), we break out new crimes into income generating crimes or other crimes.[64] If our results are driven by increases in income-generating crime, this would be more consistent with the destabilization channel. Appendix Table E.4 shows that our point estimates are similar for both crime types. The impacts are larger in percent change terms for more downstream measures of future criminal justice contact. Results are similar if we break out drug crimes from non-drug crimes (Appendix Table E.5). These analyses are far from definitive, but they provide some suggestive evidence in favor of the "ratcheting up" channel.

2SLS estimates for other subgroups. In Appendix Tables E.6 - E.8, we present 2SLS estimates conditional on various offense categories and sociodemographic characteristics. Appendix Table E.6 separately considers people with or without prior convictions in the last 5 years. We find large effects of conviction for those with no prior felony conviction. Our sample of those with a prior felony conviction is quite small and standard errors are too large to inform us about differences in effect sizes across groups.

For incarceration, we find that both groups have similar patterns: short-term incapacitation effects, but no long-term effects, for either group. This result differs from findings in Jordan et al. (2023). This could partially be caused by two limitations in our data. First, we can only observe prior felony convictions if they appear in our data set. Given this, our indicator for prior felony conviction is "prior felony conviction within the last 5 years" (and would miss all felony convictions outside of the state). Presumably, some subset of our sample with no felony conviction within the last five years have older felony convictions we cannot observe. Jordan et al. (2023) solve this

[63]The fact that conviction increases the probability of future incarceration also indicates that there are direct future financial costs within the criminal justice of these marginal convictions.

[64]Income generating crimes are cases with at least one burglary, drug (excluding drug possession), fraud, larceny, robbery, or prostitution charge.

issue by restricting their analysis to individuals who are younger than 18 at the start of their sample. We are not able to include a similar restriction as we do not know the age or date of birth for many people in our sample. It is possible that we would find different results for incarceration if our data allowed us to fully restrict the sample to first-time offenders.

We find no substantial differences between Black and White defendants (Appendix Table E.7). We do find some evidence that impacts are larger for people living in zip codes with above median poverty rates (Appendix Table E.8). This could be because felony convictions have more consequences in terms of access to relevant social services or housing, or in terms of future criminal justice scrutiny, for poorer people.

E.1.5 Robustness checks

Robustness to sample choice and specification. In Appendix Figures E.3-E.6, we examine how our main 2SLS estimates for conviction and incarceration change when we alter our sample or specifications, for our 1 year, 2-4 year, 5-7 year, and 1-7 year estimates. We consider the following variations:[65]

- Changing the required number of cases seen by a judge in our 3-year window (50 or 150 instead of 100);
- Varying which courts are included. We conducted phone interviews in 2021 with court clerks in all courts in Virginia for which we had data. We asked the clerks how cases were allocated. Our main sample includes courts where cases are quasi-randomly allocated (see Section 2.1 for more details.) We vary which courts we include:
  - Keep all courts, even if there appears to be selection in the kinds of cases that judges handle. This can happen for example if there are specialized courts, in particular drug courts.
  - Drop courts where the clerks said that cases were assigned based on judge availability, which may be more subject to discretion in what cases to work on.
- Clustering our standard errors at the month court level or at the defendant level;
- Changing what offenses are included:
  - Dropping drug cases. Although diversion is rare for felonies in Virginia, it is more likely in drug cases. Thus, dropping drug cases means eliminating the cases where diversion is most probable.
  - Dropping offenses types that are not balanced across judges (see Table 3).
- Varying how we control for non-focal stringency. In our main specification, we control for incarceration stringency, defined as the fraction of cases that end in carceral sentences. Here, we consider including controls for sentence length stringency, probability of sentence length shorter than 6 months or longer than 1 year or 4 years, flexibly controlling for deciles of the non-focal stringency, or no controls at all.
- Reconstructing the judge stringency instrument by crime type (assault or not and drug or not).

[65] The sample and specification changes are detailed in the footnotes of the figures.

- Including all years for which we can construct recidivism. We expand the sample up to 2015 for outcomes in years 2-4 and up to 2018 for outcomes in year 1.

Generally, our estimates are very close to our main specification (colored in green and denoted by the red dotted line). Although we occasionally lose statistical significance, estimates from the majority of the specifications remain significantly different from zero at the 95% level when our main estimate is also significant. Our main estimates also tend to fall towards the middle of the range of point estimates.

Robustness to different definitions of recidivism. In Appendix Table E.9, we show that our results are robust to defining recidivism in a variety of ways. In panel A we count recidivism as the total number of future charges (i.e., if you have 3 future charges in a case 1 year later, we count that as 3.) In panel B we count the total number of future charge events. Meaning that if you have a case in year 1 and another separate case in year 2 we count that as 2. Finally in panels C, D, and E we look at recidivism where there is one charge, two to three charges, and four or more charges respectively. This tests our results using slightly different definitions of recidivism. While overall we see the same general patterns, our estimates occasionally fall in and out of significance. Furthermore, much of our results seem to be coming from recidivism with more than one charge as evidenced from panels D and E.

Empirical Bayes Shrinkage. We correct for potential measurement error in judge stringency instruments using Empirical Bayes methods. We implement an Empirical Bayes procedure where we assume that judge stringencies are drawn from a Beta distribution, and the individual stringencies follow a Bernoulli distribution. We consider two specifications: in the first, we assume that judge stringencies are drawn from a single Beta distribution, while the second assumes that the Beta distribution varies by district-year. We provide detailed descriptions of our methodology and results in Appendix F.3. Overall, our results are not sensitive to using shrunken leniency estimates, which is consistent with the fact that judges in our sample see many cases per year.

Differential mobility. Our results could be confounded if conviction or incarceration influence the likelihood of moving outside of Virginia, and therefore change the likelihood that we would capture their recidivism in our data. Due to data limitations, we cannot test for this in the IV setting. However, for our RD analyses, we can test to see if there is any discontinuity in the likelihood of living in Virginia for those right above/below the cutoff in the incarceration length score and the probation/jail score. We build an indicator for Virginia residency that is equal to one if the person is marked as being in the state of VA in year 5 post-sentencing and year 7 post-sentencing. Missing observations are excluded.[66] As we can see in Appendix Figure E.7, there is no discontinuity at our cutoff score. Notably, in the incarceration-length sample, the share of people remaining in Virginia 5-7 years after the sentencing date ranges from 79-83% at every score. This consistency suggests that neither conviction nor incarceration affect migration from Virginia.

[66] If we instead include missings as 0s the results are very similar. Around 7.7% of the sample is missing this information.

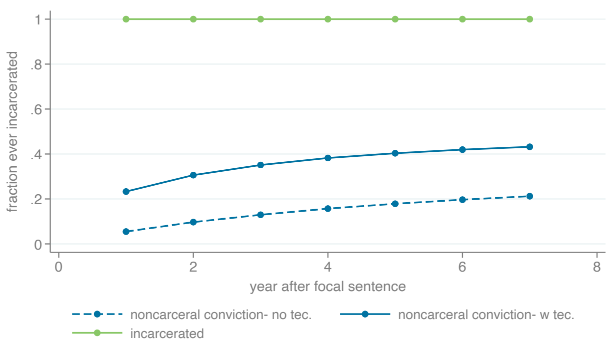
E.2 Appendix figures: 2SLS analyses

Figure E.1: Dismissed, convicted, and incarcerated percentages by offenses

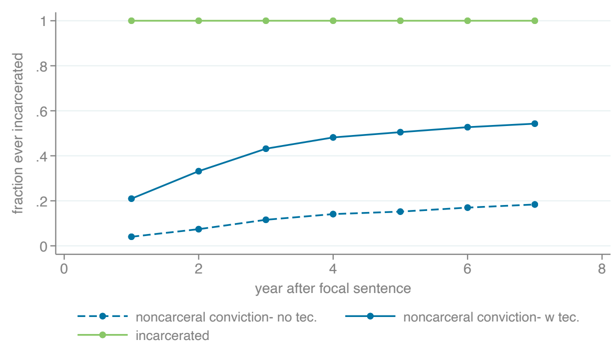
Note: This figure shows the variation in dismissal, conviction, and incarceration by four offense categories. The top left depicts fraud cases, the top right larceny, the bottom left assault, and the bottom right drugs. There is variation in the percent of cases dismissed, convicted, and incarcerated within each offense. The sample is restricted to cases observed for 7 years.

Figure E.2: Dynamics of incarceration

(a) Ever incarcerated with and without technical revocations



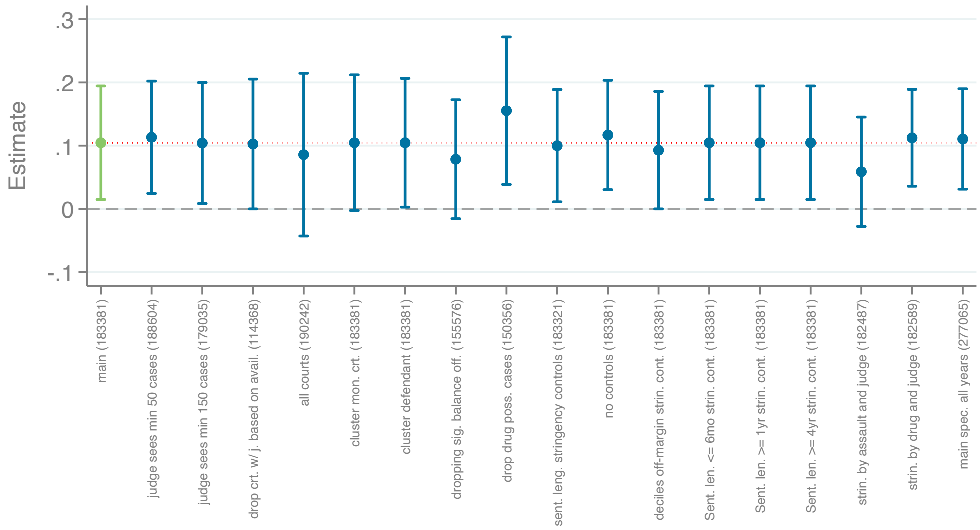
(b) Ever incarcerated with and without technical revocations– complier reweighted



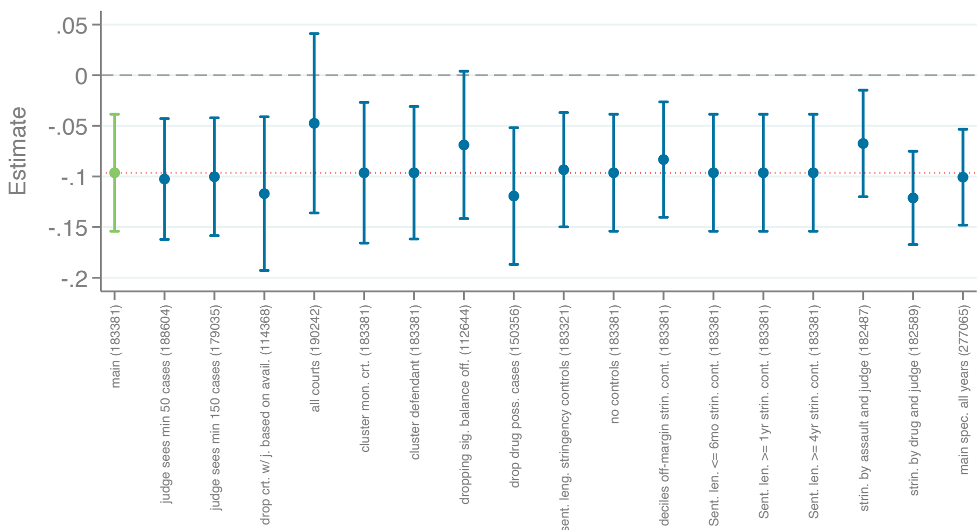
Note: Panel (a) shows the fraction of individuals that ever experience an incarceration in each year post the focal sentence date. This takes into account the current incarceration, prior incarcerations, and any future incarcerations. It is a cumulative measure, so naturally for our incarceration group the estimate is 100%. We have split out including incarceration for technical probation violations in the solid line from incarceration only due to a new crime in the dotted line. The blue lines are for people who initially got a noncarceral conviction sentence while the green line is for those who initially got a carceral sentence. Panel (b) is the same figure but complier weighted, meaning that it is reweighted to match the compliers for each margin.

Figure E.3: Robustness for 2SLS results: recidivism in year 1

(a) Noncarceral conviction



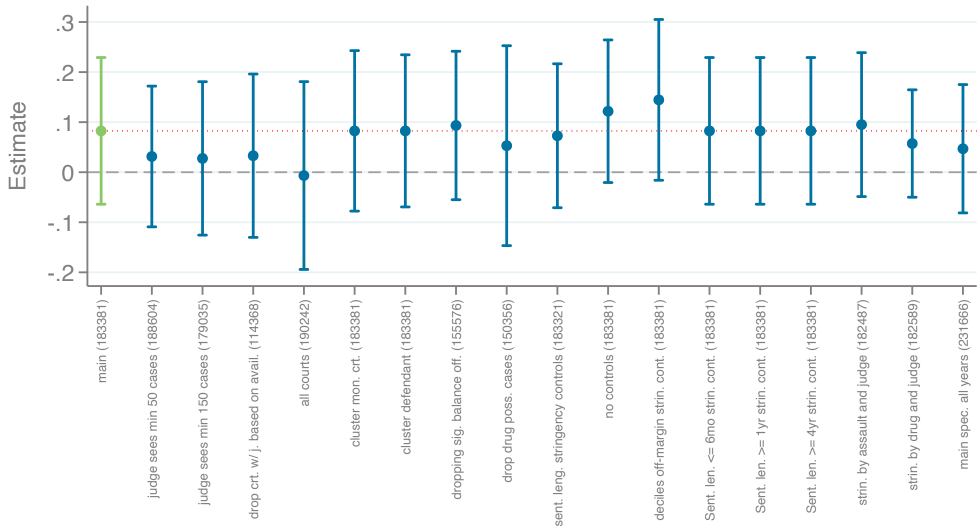
(b) Incarceration



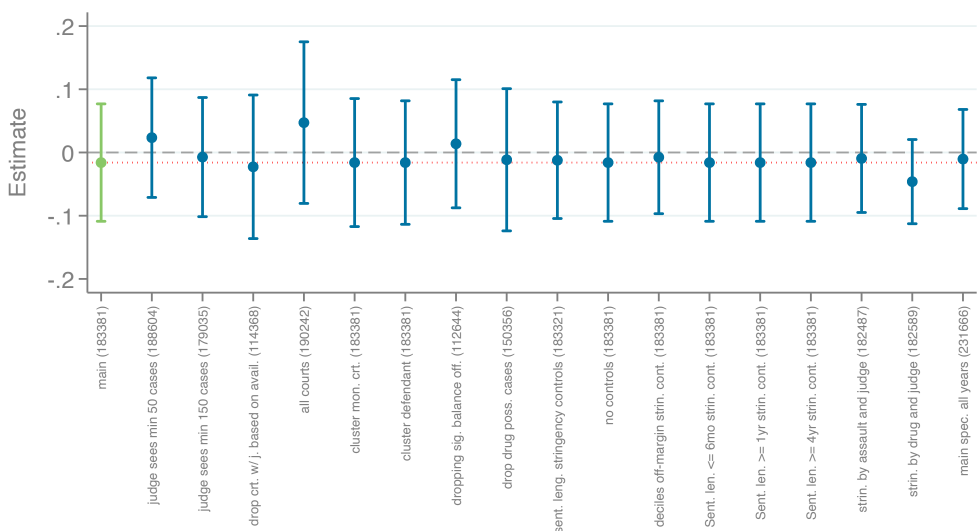
Note: This figure shows various estimates of the impact of conviction (panel a) and incarceration (panel b) on recidivism within the first year after sentencing. The main sample is restricted to cases observed for 7 years. Sample size is noted in parentheses and the main estimate is highlighted in green; 95% confidence intervals are shown. The red dotted line is located at the height of the main estimate and the dashed gray line is located at 0. The sample restrictions on the estimates are the following: (1) Our main specification. (2) Requiring that a judge sees a minimum of 50 cases in our 3 year window. (3) Requiring that a judge sees a minimum of 150 cases in our 3 year window. (4) Dropping courts that use judges based on availability. (5) Keeping courts where clerks described an assignment process that seemed non-random. (6) Clustering standard errors at the court month level. (7) Clustering standard errors at the defendant level. (8) Dropping any offenses that are significant in our balance tests. (9) Dropping any cases that relate to drug possession. (10) Including a sentence length stringency instrument control. (11) Main specification without any of our controls. (12) Including decile bins of our off-margin stringency as controls. (13) Controlling for "probability of sentence length less than 6 months" stringency. (14) Controlling for "probability of sentence length greater than 1 year" stringency. (15) Controlling for "probability of sentence length greater than 4 years" stringency. (16) Controlling for judge stringency instruments recalculated by assault only/non-assault charges. (17) Controlling for judge stringency instruments recalculated by drug/non-drug cases. (18) Using the full sample of available years for the estimate.

Figure E.4: Robustness for 2SLS results: recidivism in years 2-4

(a) Noncarceral conviction



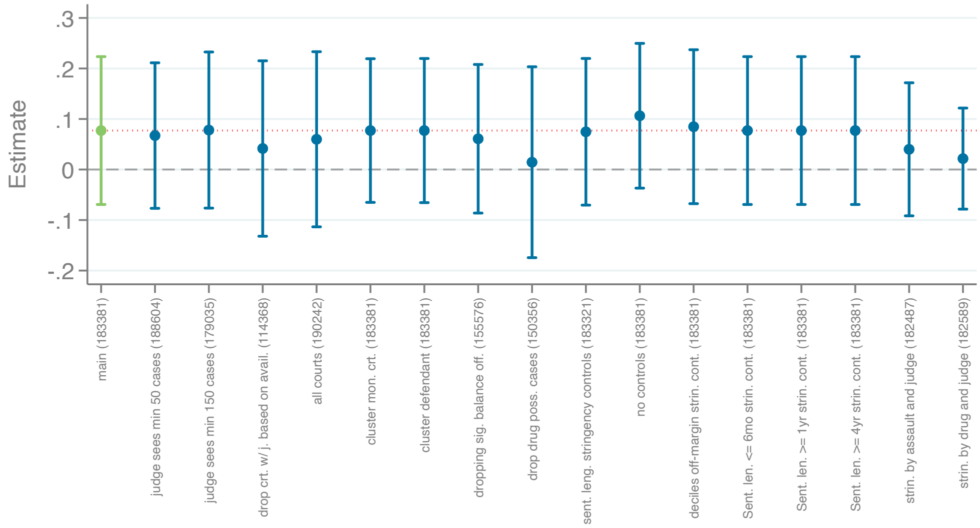
(b) Incarceration



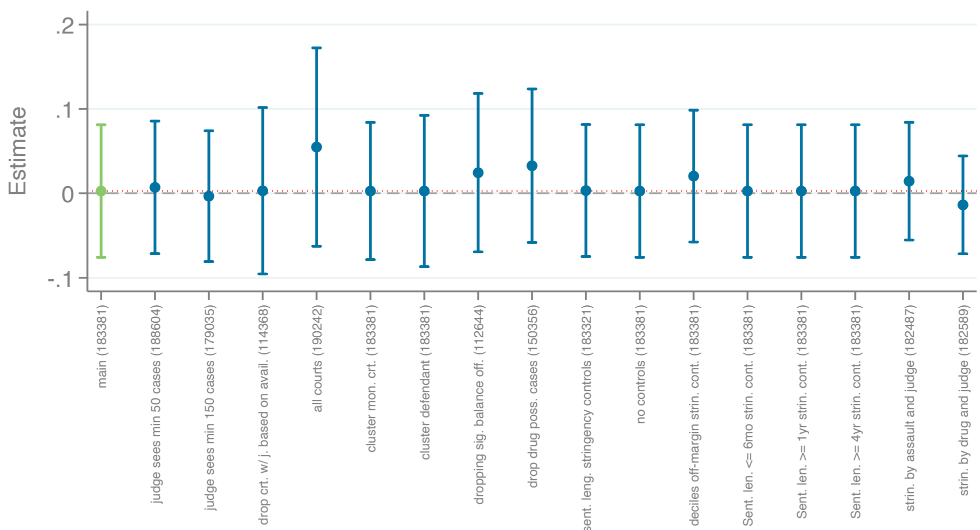
Note: This figure shows various estimates of the impact of conviction (panel a) and incarceration (panel b) on recidivism 2-4 years after sentencing. The main sample is restricted to cases observed for 7 years. Sample size is noted in parentheses and the main estimate is highlighted in green; 95% confidence intervals are shown. The red dotted line is located at the height of the main estimate and the dashed gray line is located at 0. The sample restrictions on the estimates are the following: (1) Our main specification. (2) Requiring that a judge sees a minimum of 50 cases in our 3 year window. (3) Requiring that a judge sees a minimum of 150 cases in our 3 year window. (4) Dropping courts that use judges based on availability. (5) Keeping courts where clerks described an assignment process that seemed non-random. (6) Clustering standard errors at the court month level. (7) Clustering standard errors at the defendant level. (8) Dropping any offenses that are significant in our balance tests. (9) Dropping any cases that relate to drug possession. (10) Including a sentence length stringency instrument control. (11) Main specification without any of our controls. (12) Including decile bins of our off-margin stringency as controls. (13) Controlling for "probability of sentence length less than 6 months" stringency. (14) Controlling for "probability of sentence length greater than 1 year" stringency. (15) Controlling for "probability of sentence length greater than 4 years" stringency. (16) Controlling for judge stringency instruments recalculated by assault only/non-assault charges. (17) Controlling for judge stringency instruments recalculated by drug/non-drug cases. (18) Using the full sample of available years for the estimate.

Figure E.5: Robustness for 2SLS results: recidivism in years 5-7

(a) Noncarceral conviction



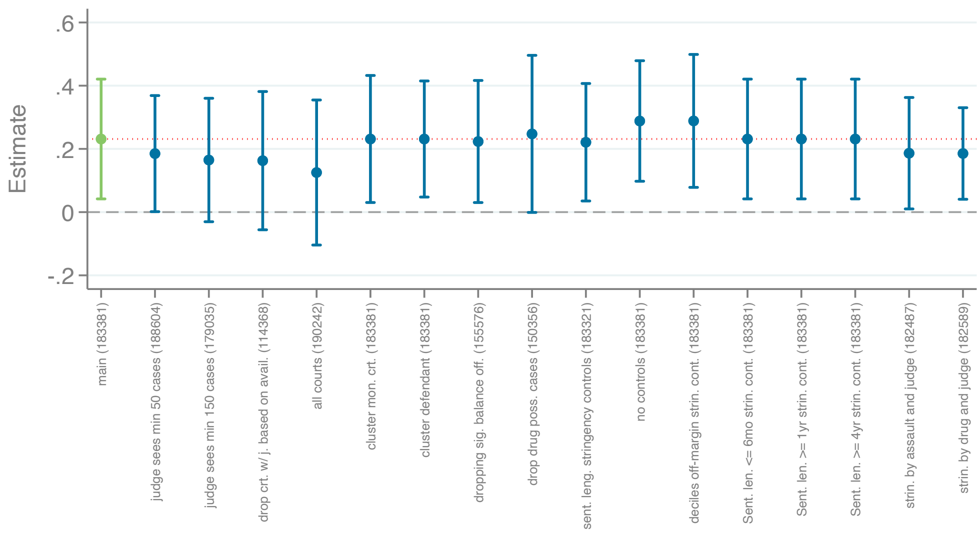
(b) Incarceration



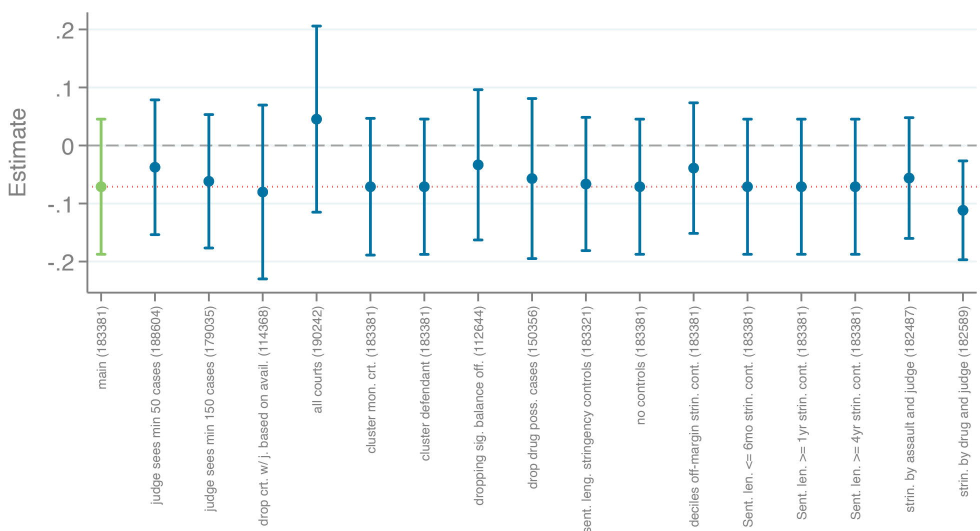
Note: This figure shows various estimates of the impact of conviction (panel a) and incarceration (panel b) on recidivism 5-7 years after sentencing. The main sample is restricted to cases observed for 7 years. Sample size is noted in parentheses and the main estimate is highlighted in green; 95% confidence intervals are shown. The red dotted line is located at the height of the main estimate and the dashed gray line is located at 0. The sample restrictions on the estimates are the following: (1) Our main specification. (2) Requiring that a judge sees a minimum of 50 cases in our 3 year window. (3) Requiring that a judge sees a minimum of 150 cases in our 3 year window. (4) Dropping courts that use judges based on availability. (5) Keeping courts where clerks described an assignment process that seemed non-random. (6) Clustering standard errors at the court month level. (7) Clustering standard errors at the defendant level. (8) Dropping any offenses that are significant in our balance tests. (9) Dropping any cases that relate to drug possession. (10) Including a sentence length stringency instrument control. (11) Main specification without any of our controls. (12) Including decile bins of our off-margin stringency as controls. (13) Controlling for "probability of sentence length less than 6 months" stringency. (14) Controlling for "probability of sentence length greater than 1 year" stringency. (15) Controlling for "probability of sentence length greater than 4 years" stringency. (16) Controlling for judge stringency instruments recalculated by assault only/non-assault charges. (17) Controlling for judge stringency instruments recalculated by drug/non-drug cases.

Figure E.6: Robustness for 2SLS results: recidivism in years 1-7

(a) Noncarceral conviction



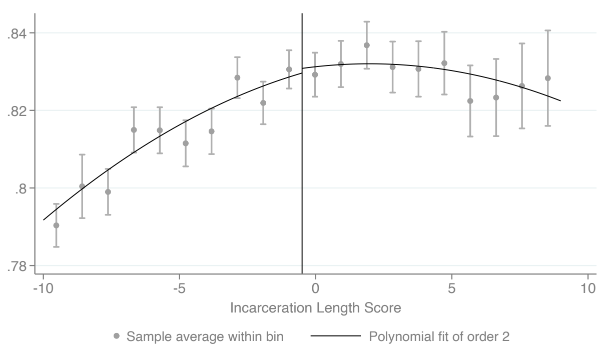
(b) Incarceration



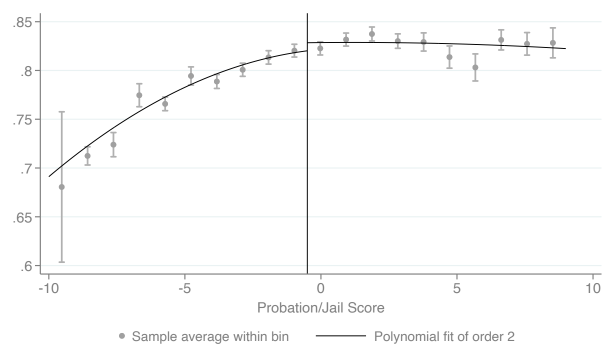
Note: This figure shows various estimates of the impact of conviction (panel a) and incarceration (panel b) on recidivism within the first seven years after sentencing. The main sample is restricted to cases observed for 7 years. Sample size is noted in parentheses and the main estimate is highlighted in green; 95% confidence intervals are shown. The red dotted line is located at the height of the main estimate and the dashed gray line is located at 0. The sample restrictions on the estimates are the following: (1) Our main specification. (2) Requiring that a judge sees a minimum of 50 cases in our 3 year window. (3) Requiring that a judge sees a minimum of 150 cases in our 3 year window. (4) Dropping courts that use judges based on availability. (5) Keeping courts where clerks described an assignment process that seemed non-random. (6) Clustering standard errors at the court month level. (7) Clustering standard errors at the defendant level. (8) Dropping any offenses that are significant in our balance tests. (9) Dropping any cases that relate to drug possession. (10) Including a sentence length stringency instrument control. (11) Main specification without any of our controls. (12) Including decile bins of our off-margin stringency as controls. (13) Controlling for "probability of sentence length less than 6 months" stringency. (14) Controlling for "probability of sentence length greater than 1 year" stringency. (15) Controlling for "probability of sentence length greater than 4 years" stringency. (16) Controlling for judge stringency instruments recalculated by assault only/non-assault charges. (17) Controlling for judge stringency instruments recalculated by drug/non-drug cases.

Figure E.7: Testing for discontinuities in Virginia residency

(a) Incarceration-length sample



(b) Probation/jail sample



Note: The outcome variable here is a flag indicating that the person is still residing in Virginia 5-7 years after their sentencing date, based on data obtained from a private vendor. Panel (a) is restricted to the RD incarceration-length sample; panel (b) is restricted to the RD probation/jail sample. There is no discontinuity across either threshold. People whose residency information is missing (7.7% of the sample) were excluded from the analysis.

E.3 Appendix tables: 2SLS analyses

Table E.1: Reduced form estimates

	Year 1	Year 2-4	Year 5-4	Year 1-7
	RF	RF	RF	RF
Panel A: Conviction				
Fut. charge	0.062**	0.049	0.046	0.137**
	(0.027)	(0.044)	(0.044)	(0.055)
Fut. conviction	0.080***	0.065	0.032	0.174***
	(0.025)	(0.042)	(0.042)	(0.053)
Fut. incarceration	0.066***	0.032	-0.015	0.124***
	(0.022)	(0.037)	(0.034)	(0.048)
Observations	183381	183381	183381	183381
Panel B: Incarceration				
Fut. charge	-0.058***	-0.010	0.002	-0.043
	(0.018)	(0.029)	(0.024)	(0.036)
Fut. conviction	-0.067***	-0.022	0.012	-0.064*
	(0.017)	(0.028)	(0.023)	(0.035)
Fut. incarceration	-0.042***	0.006	0.031	-0.017
	(0.014)	(0.025)	(0.019)	(0.031)
Observations	183381	183381	183381	183381

Note: This table shows estimates from reduced form regressions of recidivism on z_c in Panel A and regressions of recidivism on z_i in Panel B. The four columns report results for four recidivism time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). The sample is restricted to cases observed for 7 years. All regressions control for z_i in the first Panel and z_d in the second as well as, race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. The table reports the estimated impact of conviction and incarceration. The first row is for any future felony charge, the second row is for any future conviction, and the third row is for any future incarceration. Standard errors are clustered at the judge-year level. Stars denote * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E.2: Complier characteristics (noncarceral conviction)

	Pr(X=x)	Pr(X=x—complier)	$\frac{Pr(X=x\|complier)}{Pr(X=x)}$
Panel A: Conviction			
Prior conviction	0.172	0.101	0.586
	(0.003)	(0.032)	(0.187)
Female	0.218	0.273	1.248
	(0.003)	(0.041)	(0.186)
Black	0.568	0.557	0.981
	(0.015)	(0.049)	(0.085)
Has misdemeanor	0.078	0.080	1.024
	(0.004)	(0.020)	(0.254)
Drugs	0.313	0.316	1.011
	(0.007)	(0.034)	(0.105)
Property	0.377	0.417	1.104
	(0.008)	(0.045)	(0.115)
Violent	0.194	0.084	0.433
	(0.004)	(0.031)	(0.158)
Other	0.160	0.064	0.397
	(0.002)	(0.027)	(0.170)
Panel B: Incarceration			
Prior conviction	0.172	0.186	1.084
	(0.003)	(0.021)	(0.119)
Female	0.218	0.209	0.956
	(0.003)	(0.031)	(0.140)
Black	0.568	0.549	0.967
	(0.015)	(0.029)	(0.047)
Has misdemeanor	0.078	0.061	0.787
	(0.004)	(0.018)	(0.222)
Drugs	0.313	0.264	0.845
	(0.007)	(0.028)	(0.088)
Property	0.377	0.460	1.220
	(0.008)	(0.034)	(0.091)
Violent	0.194	0.084	0.431
	(0.004)	(0.028)	(0.139)
Other	0.160	0.150	0.937
	(0.002)	(0.023)	(0.144)

Note: This table shows the characteristics of compliers for our 2SLS conviction analysis in Panel A and incarceration analysis in Panel B. The first column reports average characteristics for the full 2SLS sample. The second column reports the estimated average coefficients for compliers. The third column reports the ratio of column 2 to column 1. The sample is restricted to cases observed for 7 years. Standard errors are calculated via bootstrap using 500 bootstrap samples.

Table E.3: Complier weighted OLS

	Year 1		Year 2-4		Year 5-7		Year 1-7	
	OLS	OLS weighted	OLS	OLS weighted	OLS	OLS weighted	OLS	OLS weighted
Panel A: Conviction								
Fut. charge	-0.002	0.004*	0.004	0.010***	0.006**	0.010***	0.011***	0.021***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.004)
Fut. conviction	0.001	0.006***	0.008***	0.012***	0.007***	0.011***	0.014***	0.023***
	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)
Fut. incarceration	0.001	0.006***	0.006**	0.010***	0.005**	0.008***	0.012***	0.021***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Ctrl. mean: fut. chrg.	0.081	0.089	0.154	0.170	0.115	0.129	0.270	0.297
Ctrl. mean: fut. conv.	0.069	0.076	0.135	0.148	0.102	0.114	0.243	0.268
Ctrl. mean: fut. incar.	0.048	0.054	0.097	0.109	0.073	0.083	0.182	0.204
Panel B: Incarceration								
Fut. charge	-0.022***	-0.022***	0.013***	0.013***	0.025***	0.025***	0.023***	0.024***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Fut. conviction	-0.018***	-0.018***	0.013***	0.014***	0.023***	0.023***	0.022***	0.023***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Fut. incarceration	-0.010***	-0.010***	0.017***	0.017***	0.021***	0.021***	0.027***	0.028***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Ctrl. mean: fut. chrg.	0.088	0.088	0.177	0.175	0.133	0.132	0.308	0.306
Ctrl. mean: fut. conv.	0.078	0.077	0.160	0.159	0.121	0.120	0.285	0.283
Ctrl. mean: fut. incar.	0.055	0.055	0.116	0.115	0.085	0.084	0.214	0.212
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381

Note: This table shows regression estimates of the impact of conviction on future recidivism in Panel A and the impact of incarceration on future recidivism in Panel B, showing our ordinary least squares (OLS) regressions and complier weighted OLS etimates. The four columns report results for four time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). Each time period restricts the sample to cases observed for 7 years. All regressions control for race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. The first three rows of each panel report the estimated impact of conviction or incarceration on different measures of recidivism. The first row is for any future charge, the second row is for any future conviction, and the third row is for any future incarceration. For the OLS estimates in Panel A, we regress our measures of recidivism on having a conviction (regardless of incarceration status) controlling for incarceration. For for the OLS weighted estimates, we use the same regression but weighted by IV compliers. For the OLS estimates in Panel B, we regress our measures of recidivism on incarceration, controlling for dismissal. For the OLS weighted estimates we use the same regression but weighted by incarceration IV compliers. The estimates presented are the coefficients on the conviction variable. Standard errors are clustered at the judge-year level. Stars denote * $p< 0.10$, ** $p< 0.05$, *** $p< 0.01$.

Table E.4: Income-generating vs non-income-generating recidivism

	Income generating recidivsim				Non-income generating recidivsim			
	Year 1	Year 2-4	Year 5-7	Year 1-7	Year 1	Year 2-4	Year 5-7	Year 1-7
Panel A: Conviction								
Fut. charge	0.061*	0.031	0.042	0.131*	0.092**	0.035	0.006	0.128
	(0.034)	(0.060)	(0.054)	(0.077)	(0.039)	(0.063)	(0.062)	(0.082)
Fut. conviction	0.070**	0.095	0.042	0.196**	0.102***	0.032	-0.020	0.118
	(0.032)	(0.059)	(0.052)	(0.077)	(0.036)	(0.058)	(0.060)	(0.081)
Fut. incarceration	0.046*	0.045	-0.016	0.096	0.091***	0.040	-0.031	0.131**
	(0.027)	(0.050)	(0.044)	(0.068)	(0.030)	(0.049)	(0.045)	(0.066)
Ctrl. comp. mean: fut. chrg.	0.023	0.117	0.076	0.181	0.002	0.121	0.100	0.181
Ctrl. mean: fut. chrg.	0.054	0.106	0.079	0.196	0.053	0.108	0.081	0.204
Ctrl. comp. mean: fut. conv.	0.017	0.087	0.073	0.144	-0.013	0.105	0.098	0.159
Ctrl. mean: fut. conv.	0.047	0.092	0.070	0.175	0.044	0.091	0.070	0.178
Ctrl. comp. mean: fut. incar.	0.015	0.057	0.077	0.119	-0.015	0.082	0.076	0.113
Ctrl. mean: fut. incar.	0.034	0.069	0.053	0.135	0.030	0.064	0.048	0.128
Panel B: Incarceration								
Fut. charge	-0.043*	0.004	-0.020	-0.054	-0.066***	-0.022	0.029	-0.032
	(0.024)	(0.036)	(0.031)	(0.048)	(0.024)	(0.041)	(0.032)	(0.051)
Fut. conviction	-0.053**	-0.013	-0.027	-0.073	-0.072***	-0.018	0.054*	-0.016
	(0.023)	(0.036)	(0.030)	(0.048)	(0.023)	(0.039)	(0.031)	(0.049)
Fut. incarceration	-0.033*	-0.002	0.025	-0.019	-0.044**	0.009	0.059**	0.025
	(0.019)	(0.031)	(0.026)	(0.042)	(0.019)	(0.032)	(0.024)	(0.040)
Ctrl. comp. mean: fut. chrg.	0.089	0.117	0.124	0.259	0.066	0.131	0.094	0.242
Ctrl. mean: fut. chrg.	0.056	0.112	0.080	0.204	0.047	0.102	0.080	0.195
Ctrl. comp. mean: fut. conv.	0.092	0.127	0.117	0.259	0.074	0.117	0.072	0.227
Ctrl. mean: fut. conv.	0.049	0.102	0.073	0.188	0.041	0.090	0.071	0.174
Ctrl. comp. mean: fut. incar.	0.070	0.077	0.063	0.175	0.048	0.076	0.038	0.140
Ctrl. mean: fut. incar.	0.035	0.074	0.053	0.141	0.029	0.062	0.047	0.123
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381

Note: The first panel of this table shows 2SLS estimates of the impact of conviction vs dismissal on future recidivism. In the first four columns, recidivism is defined in reference to new income-generating felony charges, in the last four columns recidivism is defined in reference to new non-income generating charges. The second panel is similar except it shows 2SLS estimates of the impact of incarceration vs conviction. The columns report results for four recidivism time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). The sample is restricted to cases observed for 7 years. All regressions control for stringency on the other margin (i.e., z_i for the conviction specification), race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. Stars denote * $p< 0.10$, ** $p< 0.05$, *** $p< 0.01$.

Table E.5: Drug vs non-drug recidivism

	Drug charges				Non-drug charges			
	Year 1	Year 2-4	Year 5-7	Year 1-7	Year 1	Year 2-4	Year 5-7	Year 1-7
Panel A: Conviction								
Fut. charge	0.146**	0.001	-0.015	0.161	0.077	0.148	0.141	0.298**
	(0.070)	(0.105)	(0.110)	(0.140)	(0.057)	(0.097)	(0.086)	(0.121)
Fut. conviction	0.115*	0.026	-0.052	0.204	0.144***	0.173*	0.124	0.369***
	(0.062)	(0.098)	(0.104)	(0.134)	(0.054)	(0.093)	(0.083)	(0.118)
Fut. incarceration	0.114**	-0.027	-0.084	0.130	0.110**	0.110	0.019	0.272***
	(0.055)	(0.091)	(0.087)	(0.128)	(0.048)	(0.083)	(0.068)	(0.103)
Ctrl. comp. mean: fut. chrg.	0.149	0.356	0.249	0.554	0.164	0.273	0.229	0.460
Ctrl. mean: fut. chrg.	0.079	0.159	0.123	0.282	0.094	0.176	0.132	0.306
Ctrl. comp. mean: fut. conv.	0.134	0.328	0.232	0.523	0.142	0.231	0.221	0.425
Ctrl. mean: fut. conv.	0.067	0.137	0.108	0.252	0.080	0.154	0.117	0.277
Ctrl. comp. mean: fut. incar.	0.136	0.334	0.281	0.572	0.137	0.268	0.279	0.504
Ctrl. mean: fut. incar.	0.047	0.097	0.076	0.184	0.058	0.116	0.087	0.216
Panel B: Incarceration								
Fut. charge	-0.060	0.000	-0.012	-0.046	-0.114***	-0.026	0.004	-0.090
	(0.061)	(0.089)	(0.081)	(0.109)	(0.034)	(0.055)	(0.046)	(0.069)
Fut. conviction	-0.073	0.001	0.030	-0.082	-0.130***	-0.056	0.014	-0.123*
	(0.055)	(0.089)	(0.078)	(0.109)	(0.033)	(0.053)	(0.045)	(0.066)
Fut. incarceration	-0.062	0.082	0.010	-0.017	-0.079***	-0.021	0.062*	-0.042
	(0.049)	(0.079)	(0.067)	(0.102)	(0.028)	(0.047)	(0.037)	(0.058)
Ctrl. comp. mean: fut. chrg.	0.123	0.177	0.105	0.340	0.119	0.205	0.162	0.377
Ctrl. mean: fut. chrg.	0.100	0.189	0.143	0.336	0.082	0.168	0.126	0.291
Ctrl. comp. mean: fut. conv.	0.084	0.132	0.064	0.260	0.083	0.182	0.133	0.328
Ctrl. mean: fut. conv.	0.088	0.171	0.130	0.311	0.072	0.153	0.114	0.269
Ctrl. comp. mean: fut. incar.	0.043	0.048	0.026	0.136	0.043	0.083	0.064	0.180
Ctrl. mean: fut. incar.	0.063	0.125	0.090	0.234	0.051	0.110	0.081	0.201
Observations	57,249	57,249	57,249	57,249	126,134	126,134	126,134	126,134

Note: The first panel of this table shows 2SLS estimates of the impact of conviction vs dismissal on future recidivism. In the first four columns, recidivism is defined in reference to new drug charges; in the last four columns recidivism is defined in reference to new non-drug charges. The second panel is similar except it shows 2SLS estimates of the impact of incarceration. The columns report results for four recidivism time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). The sample is restricted to cases observed for 7 years. All regressions control for stringency on the opposite margin (i.e., z_i in the conviction specification) race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. Stars denote * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E.6: 2SLS estimates for those with/without prior felony convictions

	Priors				No priors			
	Year 1	Year 2-4	Year 5-7	Year 1-7	Year 1	Year 2-4	Year 5-7	Year 1-7
Panel A: Conviction								
Fut. charge	0.092	0.306	-0.204	0.319	0.102**	0.073	0.113	0.229**
	(0.211)	(0.365)	(0.398)	(0.443)	(0.046)	(0.073)	(0.069)	(0.093)
Fut. conviction	0.205	0.239	-0.216	0.354	0.125***	0.109	0.085	0.293***
	(0.200)	(0.346)	(0.379)	(0.447)	(0.044)	(0.071)	(0.065)	(0.090)
Fut. incarceration	0.152	0.356	-0.340	0.444	0.107***	0.034	0.005	0.190**
	(0.178)	(0.315)	(0.327)	(0.430)	(0.036)	(0.060)	(0.052)	(0.078)
Ctrl. comp. mean: fut. chrg.	0.342	0.506	0.544	0.955	0.139	0.272	0.190	0.427
Ctrl. mean: fut. chrg.	0.147	0.294	0.239	0.503	0.080	0.151	0.112	0.265
Ctrl. comp. mean: fut. conv.	0.273	0.422	0.506	0.847	0.124	0.240	0.182	0.402
Ctrl. mean: fut. conv.	0.129	0.264	0.217	0.471	0.067	0.130	0.098	0.236
Ctrl. comp. mean: fut. incar.	0.291	0.503	0.705	1.071	0.120	0.262	0.223	0.455
Ctrl. mean: fut. incar.	0.098	0.209	0.173	0.386	0.047	0.093	0.069	0.176
Panel B: Incarceration								
Fut. charge	-0.101	-0.039	0.099	-0.021	-0.090***	-0.010	-0.013	-0.073
	(0.071)	(0.112)	(0.111)	(0.131)	(0.032)	(0.050)	(0.042)	(0.062)
Fut. conviction	-0.148**	-0.092	0.102	-0.106	-0.099***	-0.025	0.007	-0.098
	(0.069)	(0.107)	(0.107)	(0.126)	(0.031)	(0.049)	(0.040)	(0.060)
Fut. incarceration	-0.091	-0.062	0.152	-0.022	-0.064**	0.024	0.036	-0.023
	(0.064)	(0.097)	(0.094)	(0.120)	(0.025)	(0.042)	(0.034)	(0.053)
Ctrl. comp. mean: fut. chrg.	0.097	0.209	0.239	0.451	0.127	0.201	0.129	0.357
Ctrl. mean: fut. chrg.	0.119	0.302	0.234	0.496	0.084	0.161	0.120	0.285
Ctrl. comp. mean: fut. conv.	0.087	0.194	0.195	0.402	0.085	0.167	0.099	0.297
Ctrl. mean: fut. conv.	0.106	0.280	0.217	0.471	0.074	0.145	0.109	0.261
Ctrl. comp. mean: fut. incar.	0.052	0.130	0.115	0.264	0.043	0.065	0.043	0.154
Ctrl. mean: fut. incar.	0.076	0.216	0.161	0.375	0.052	0.103	0.075	0.194
Observations	31,505	31,505	31,505	31,505	151,878	151,878	151,878	151,878

Note: The first panel of this table shows 2SLS estimates of the impact of conviction on future recidivism for those with/without a prior felony conviction within 5 years. The second panel is similar except it shows 2SLS estimates of the impact of incarceration. The columns report results for four recidivism time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). The sample is restricted to cases observed for 7 years. All regressions control for stringency on the opposite margin (i.e., z_i in the conviction specification) race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

Table E.7: 2SLS estimates for Black and non-Black defendants

	Black				Non-Black			
	Year 1	Year 2-4	Year 5-7	Year 1-7	Year 1	Year 2-4	Year 5-7	Year 1-7
Panel A: Conviction								
Fut. charge	0.102	0.073	0.112	0.238	0.103	0.115	0.042	0.241**
	(0.069)	(0.117)	(0.109)	(0.147)	(0.063)	(0.090)	(0.094)	(0.114)
Fut. conviction	0.125*	0.112	0.100	0.335**	0.142**	0.128	0.011	0.274**
	(0.067)	(0.110)	(0.102)	(0.146)	(0.058)	(0.088)	(0.092)	(0.111)
Fut. incarceration	0.164***	0.008	-0.012	0.259**	0.058	0.115	-0.027	0.181*
	(0.061)	(0.088)	(0.080)	(0.126)	(0.049)	(0.082)	(0.077)	(0.103)
Ctrl. comp. mean: fut. chrg.	0.150	0.303	0.222	0.493	0.155	0.260	0.216	0.433
Ctrl. mean: fut. chrg.	0.104	0.196	0.148	0.339	0.070	0.135	0.104	0.241
Ctrl. comp. mean: fut. conv.	0.136	0.260	0.212	0.460	0.133	0.235	0.210	0.407
Ctrl. mean: fut. conv.	0.088	0.169	0.129	0.305	0.059	0.120	0.093	0.218
Ctrl. comp. mean: fut. incar.	0.137	0.324	0.289	0.567	0.128	0.233	0.249	0.447
Ctrl. mean: fut. incar.	0.064	0.126	0.094	0.235	0.040	0.087	0.069	0.164
Panel B: Incarceration								
Fut. charge	-0.131***	-0.009	-0.061	-0.113	-0.057	-0.022	0.073	-0.027
	(0.042)	(0.070)	(0.060)	(0.086)	(0.040)	(0.065)	(0.057)	(0.082)
Fut. conviction	-0.125***	-0.025	-0.044	-0.140	-0.099**	-0.048	0.092	-0.072
	(0.041)	(0.068)	(0.057)	(0.085)	(0.040)	(0.064)	(0.056)	(0.079)
Fut. incarceration	-0.105***	0.013	0.016	-0.070	-0.032	0.004	0.089*	0.014
	(0.035)	(0.056)	(0.047)	(0.074)	(0.032)	(0.057)	(0.049)	(0.072)
Ctrl. comp. mean: fut. chrg.	0.164	0.226	0.173	0.435	0.097	0.208	0.147	0.361
Ctrl. mean: fut. chrg.	0.094	0.193	0.144	0.332	0.081	0.157	0.118	0.279
Ctrl. comp. mean: fut. conv.	0.122	0.194	0.136	0.371	0.062	0.171	0.113	0.300
Ctrl. mean: fut. conv.	0.082	0.173	0.130	0.305	0.073	0.144	0.109	0.259
Ctrl. comp. mean: fut. incar.	0.066	0.070	0.062	0.193	0.034	0.098	0.059	0.185
Ctrl. mean: fut. incar.	0.059	0.127	0.091	0.232	0.050	0.102	0.076	0.192
Observations	104,225	104,225	104,225	104,225	79,158	79,158	79,158	79,158

Note: The first panel of this table shows 2SLS estimates of the impact of conviction on future recidivism for Black and non-Black defendants. The second panel is similar except it shows 2SLS estimates of the impact of incarceration. The columns report results for four recidivism time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). The sample is restricted to cases observed for 7 years. All regressions control for stringency on the opposite margin (i.e., z_i in the conviction specification) race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

Table E.8: 2SLS estimates for those from zip codes above and below median poverty level

	Above median poverty zip				Below median poverty zip			
	Year 1	Year 2-4	Year 5-7	Year 1-7	Year 1	Year 2-4	Year 5-7	Year 1-7
Panel A: Conviction								
Fut. charge	0.173*	0.186	0.048	0.351**	0.019	0.002	0.041	0.055
	(0.096)	(0.138)	(0.130)	(0.171)	(0.055)	(0.100)	(0.109)	(0.123)
Fut. conviction	0.166*	0.245*	0.000	0.383**	0.070	-0.009	0.043	0.122
	(0.086)	(0.131)	(0.123)	(0.157)	(0.051)	(0.095)	(0.104)	(0.121)
Fut. incarceration	0.111	0.181	-0.105	0.296**	0.062	-0.064	0.004	0.041
	(0.068)	(0.119)	(0.104)	(0.148)	(0.048)	(0.080)	(0.086)	(0.102)
Ctrl. comp. mean: fut. chrg.	0.164	0.319	0.188	0.471	0.129	0.253	0.210	0.442
Ctrl. mean .	0.110	0.204	0.151	0.350	0.078	0.150	0.115	0.268
Ctrl. comp. mean: fut. conv.	0.144	0.284	0.182	0.438	0.117	0.231	0.203	0.419
Ctrl. mean: fut. conv.	0.091	0.177	0.133	0.316	0.067	0.133	0.102	0.242
Ctrl. comp. mean: fut. incar.	0.145	0.301	0.251	0.515	0.108	0.245	0.245	0.462
Ctrl. mean: fut. incar.	0.065	0.131	0.096	0.241	0.046	0.097	0.075	0.183
Panel B: Incarceration								
Fut. charge	-0.100**	0.006	0.079	0.010	-0.071*	0.041	-0.011	-0.018
	(0.046)	(0.073)	(0.066)	(0.086)	(0.042)	(0.067)	(0.061)	(0.080)
Fut. conviction	-0.092**	-0.023	0.102	-0.001	-0.097**	0.024	-0.004	-0.063
	(0.045)	(0.071)	(0.062)	(0.081)	(0.040)	(0.067)	(0.059)	(0.080)
Fut. incarceration	-0.055	-0.020	0.150***	0.041	-0.054	0.079	0.005	0.002
	(0.036)	(0.062)	(0.056)	(0.076)	(0.035)	(0.059)	(0.052)	(0.072)
Ctrl. comp. mean: fut. chrg.	0.145	0.194	0.137	0.364	0.094	0.204	0.170	0.374
Ctrl. mean: fut. chrg.	0.101	0.202	0.153	0.352	0.084	0.169	0.125	0.295
Ctrl. comp. mean: fut. conv.	0.099	0.155	0.094	0.279	0.066	0.181	0.139	0.329
Ctrl. mean: fut. conv.	0.088	0.181	0.138	0.324	0.075	0.155	0.114	0.274
Ctrl. comp. mean: fut. incar.	0.056	0.083	0.056	0.180	0.031	0.067	0.051	0.157
Ctrl. mean: fut. incar.	0.063	0.133	0.099	0.248	0.053	0.110	0.079	0.202
Observations	73,473	73,473	73,473	73,473	73,533	73,533	73,533	73,533

Note: The first panel of this table shows 2SLS estimates of the impact of conviction on future recidivism for those who live in zip codes where the percent earning under 25K (percent in poverty) is above/below median. The second panel is similar except it shows 2SLS estimates of the impact of incarceration. The columns report results for four recidivism time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). The sample is restricted to cases observed for 7 years. All regressions control for stringency on the opposite margin (i.e., z_i in the conviction specification) race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. Stars denote * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E.9: New charges and conviction

	Noncarceral conviction sample				Incarceration sample			
	Year 1	**Year 2-4**	**Year 5-7**	**Year 1-7**	**Year 1**	**Year 2-4**	**Year 5-7**	**Year 1-7**
Panel A: Total number of future charges								
Fut. charge	0.200	0.287	0.050	0.537	-0.166	-0.040	0.121	-0.084
	(0.201)	(0.414)	(0.300)	(0.575)	(0.114)	(0.215)	(0.167)	(0.312)
Fut. conviction	0.285	0.863**	0.306	1.454***	-0.169	-0.241	0.154	-0.256
	(0.182)	(0.356)	(0.254)	(0.494)	(0.103)	(0.202)	(0.152)	(0.283)
Ctrl Mean: fut. charge	0.222	0.498	0.362	1.083	0.213	0.483	0.353	1.050
Ctrl Mean: fut. conv.	0.000	0.000	0.000	0.000	0.209	0.473	0.344	1.025
Panel B: Number of future charge events								
Fut. charge	0.095	0.114	-0.092	0.117	-0.128***	-0.064	0.073	-0.120
	(0.060)	(0.128)	(0.121)	(0.212)	(0.038)	(0.078)	(0.062)	(0.117)
Fut. conviction	0.192***	0.371***	0.078	0.641***	-0.127***	-0.089	0.085	-0.132
	(0.056)	(0.115)	(0.101)	(0.193)	(0.036)	(0.075)	(0.061)	(0.113)
Ctrl Mean: fut. charge	0.115	0.251	0.183	0.549	0.108	0.247	0.185	0.540
Ctrl Mean: fut. conv.	0.000	0.000	0.000	0.000	0.106	0.241	0.179	0.526
Panel C: Future recidivism with 1 charge								
Fut. charge	0.044	-0.055	-0.004	0.044	-0.042*	0.058	0.020	0.019
	(0.038)	(0.055)	(0.062)	(0.076)	(0.025)	(0.036)	(0.032)	(0.045)
Fut. conviction	0.075**	0.066	0.064	0.182***	-0.032	0.050	0.029	0.024
	(0.034)	(0.052)	(0.051)	(0.070)	(0.022)	(0.035)	(0.031)	(0.045)
Ctrl Mean: fut. charge	0.063	0.122	0.092	0.213	0.060	0.119	0.089	0.208
Ctrl Mean: fut. conv.	0.000	0.000	0.000	0.000	0.059	0.116	0.086	0.203
Panel D: Future recidivism with 2 to 3 charges								
Fut. charge	0.051**	0.067*	0.050	0.126***	-0.035**	-0.063***	-0.008	-0.078**
	(0.022)	(0.037)	(0.034)	(0.045)	(0.015)	(0.024)	(0.022)	(0.031)
Fut. conviction	0.069***	0.115***	0.070**	0.193***	-0.037**	-0.062***	-0.009	-0.087***
	(0.020)	(0.035)	(0.032)	(0.043)	(0.014)	(0.022)	(0.021)	(0.029)
Ctrl Mean: fut. charge	0.019	0.037	0.028	0.064	0.021	0.042	0.032	0.074
Ctrl Mean: fut. conv.	0.000	0.000	0.000	0.000	0.020	0.041	0.031	0.072
Panel E: Future recidivism with 4 or more charges								
Fut. charge	0.003	0.069***	0.027	0.065**	-0.014*	-0.018	-0.006	-0.015
	(0.013)	(0.020)	(0.020)	(0.028)	(0.008)	(0.013)	(0.012)	(0.017)
Fut. conviction	0.016	0.088***	0.035*	0.100***	-0.013*	-0.020	-0.003	-0.016
	(0.013)	(0.019)	(0.019)	(0.026)	(0.007)	(0.013)	(0.011)	(0.016)
Ctrl Mean: fut. charge	0.005	0.011	0.008	0.019	0.006	0.013	0.010	0.023
Ctrl Mean: fut. conv.	0.000	0.000	0.000	0.000	0.006	0.013	0.010	0.022
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381

Note: The first panel of this table shows 2SLS estimates of the impact of conviction vs dismissal on the number of future charges and convictions. The second panel is similar except it shows 2SLS estimates of the impact of incarceration vs conviction on the number of future charges. This table shows the results using both the noncarceral conviction sample (columns 1-4) and the incarceration sample (columns 5-8). The columns report results for four recidivism time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). Each time period restricts the sample to cases observed for all 7 years. All regressions control for stringency on the other margin (i.e., z_i for the conviction specification), race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

Table E.10: Testing the models with descriptive characteristics

	Prior Conviction	Female	Black	Misdemeanors	Assualt	Burglary	Drugs	Fraud	Kidnapping	Larceny	Misc	Murder	Robbery	Sexual Assault
Panel A: Ordered														
Conviction stringency (Z_c)	0.13***	-0.10**	0.0015	0.084**	-0.087*	0.036	0.026	0.055	-0.0020	0.047	-0.032**	-0.054*	0.022	0.0070
	(0.046)	(0.046)	(0.055)	(0.036)	(0.049)	(0.032)	(0.058)	(0.034)	(0.017)	(0.052)	(0.015)	(0.030)	(0.022)	(0.040)
Mean dep. var.	0.228	0.176	0.583	0.097	0.185	0.075	0.299	0.097	0.020	0.260	0.014	0.059	0.033	0.112
N	153692	153692	153692	153692	153692	153692	153692	153692	153692	153692	153692	153692	153692	153692
Panel B: Sequential and ordered														
Incarceration stringency (Z_i)	-0.051	0.11	-0.059	0.0081	0.19**	-0.0060	0.14	-0.13**	0.016	-0.039	0.039*	0.028	0.048	-0.036
	(0.064)	(0.074)	(0.085)	(0.039)	(0.073)	(0.041)	(0.091)	(0.055)	(0.033)	(0.071)	(0.021)	(0.043)	(0.033)	(0.032)
Mean dep. var.	0.136	0.220	0.570	0.065	0.192	0.057	0.352	0.093	0.027	0.175	0.011	0.045	0.034	0.037
N	28589	28589	28589	28589	28589	28589	28589	28589	28589	28589	28589	28589	28589	28589

Note: This table replicates the test of the UPM assumption conducted in Table 6, but using individual covariates as the dependent variables rather than predicted recidivism. For Panel A, we restrict to the incarcerated sample and regress case characteristics on conviction stringency, controlling for incarceration stringency and court-by-time fixed effects. For Panel B, we restrict to the dismissed sample and regress case characteristics on incarceration stringency, controlling for dismissal stringency and court-by-time fixed effects. The sample is restricted to cases adjudicated in 2012 or earlier. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

F Additional derivations and results

F.1 2SLS with two endogenous variables

Here we briefly discuss why our specification – which instruments for a binary treatment indicator (such as T_c) with one stringency (such as Z_c) while controlling for another stringency (such as Z_i) – should have the same estimand as running a single 2SLS regression with two endogenous treatment variables and both stringencies. In the main paper, we consider the following population regression:

$$\begin{aligned} T_c &= \delta_0 + \delta_1 Z_c + \delta_2 Z_i + U \\ Y &= \gamma_0 + \gamma_1 T_c + \gamma_2 Z_i + V \end{aligned}$$

In the population, we should have $\delta_0 = 0$, $\delta_1 = 1$, and $\delta_2 = 0$. Thus, γ_1 should be equal to γ_1' in the following regression:

$$Y = \gamma_0' + \gamma_1' Z_c + \gamma_2' Z_i + V'$$

Consider now a specification in which both endogenous variables, T_c and T_i, are instrumented for in the same second-stage regression:

$$\begin{aligned} T_c &= \delta_0 + \delta_1 Z_c + \delta_2 Z_i + U \\ T_i &= \omega_0 + \omega_1 Z_c + \omega_2 Z_i + U \\ Y &= \gamma_0'' + \gamma_1'' T_c + \gamma_2'' T_i + V'' \end{aligned}$$

By similar logic, $\omega_0 = 0$, $\omega_1 = 0$, and $\omega_2 = 1$. Thus, $\gamma_1 = \gamma_1' = \gamma_1''$ and $\gamma_2 = \gamma_2' = \gamma_2''$.

In our sample, the first-stage coefficients are not precisely zero or one, as is common in the applied literature. Yet, these two approaches produce similar estimates. Table F.1 shows that, when running 2SLS with two instruments and two endogenous variables, our estimates are similar to those in the main paper and we reach similar conclusions. Note that in these 2SLS and OLS regressions we replace T_c with $T_{\setminus d}$ (i.e., the conviction instrument dummy that remains equal to one for those incarcerated) so that the loading on T_i can be interpreted as the change relative to $T = C$ rather than $T = D$.

Table F.1: Two instruments and two endogenous variables

	Year 1		Year 2-4		Year 5-7		Year 1-7	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Convict: fut. charge	-0.002	0.105**	0.004	0.086	0.006**	0.082	0.011***	0.241**
	(0.002)	(0.048)	(0.003)	(0.078)	(0.002)	(0.078)	(0.004)	(0.100)
Incar: fut. charge	-0.022***	-0.097***	0.013***	-0.016	0.025***	0.003	0.023***	-0.071
	(0.002)	(0.029)	(0.002)	(0.047)	(0.002)	(0.040)	(0.003)	(0.059)
Convict: fut. conv.	0.001	0.136***	0.008***	0.115	0.007***	0.059	0.014***	0.306***
	(0.002)	(0.044)	(0.003)	(0.075)	(0.002)	(0.074)	(0.004)	(0.098)
Incar: fut. conv.	-0.018***	-0.112***	0.013***	-0.037	0.023***	0.020	0.022***	-0.106*
	(0.001)	(0.029)	(0.002)	(0.047)	(0.002)	(0.039)	(0.003)	(0.058)
Convict: fut. incar.	0.001	0.114***	0.006**	0.058	0.005**	-0.023	0.012***	0.220**
	(0.002)	(0.039)	(0.002)	(0.066)	(0.002)	(0.059)	(0.003)	(0.086)
Incar: fut. incar.	-0.010***	-0.071***	0.017***	0.009	0.021***	0.052	0.027***	-0.028
	(0.001)	(0.024)	(0.002)	(0.041)	(0.002)	(0.032)	(0.003)	(0.051)
Ctrl Mean: fut. charge	0.088	0.088	0.175	0.175	0.132	0.132	0.306	0.306
Ctrl Mean: fut. conv.	0.077	0.077	0.159	0.159	0.120	0.120	0.283	0.283
Ctrl Mean: fut. incar.	0.055	0.055	0.115	0.115	0.084	0.084	0.212	0.212
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381

Note: This table shows regression estimates of the impacts of conviction and incarceration on future recidivism. The four columns report results for four time ranges (1 year, 2-4 years, 5-7 years, and 1-7 years). For each panel we report ordinary least squares (OLS) and instrumental variable (IV) estimates with two instruments and two endogenous variables. Each time period restricts the sample to cases observed for seven years. All regressions control for race, gender, prior conviction, offense type dummies, year-by-court fixed effects, court-by-month-of-year fixed effects, and day-of-week fixed effects. The first six rows report the estimated impact of conviction or incarceration on different measures of recidivism. The first two rows are for any future charge, the second two rows are for any future conviction, and the third two rows are for any future incarceration. For the OLS estimates, we regress our measures of recidivism on having a conviction (regardless of incarceration status) controlling for incarceration. The estimates presented are the coefficient on the conviction variable. For the IV estimates, this provides an estimate of the impacts of conviction compared to dismissal for the set of compliers at that margin and incarceration compared to conviction for the set of compliers at the other margin. Standard errors are clustered at the judge-year level. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01.

F.2 Binary treatment

Consider an attempt to estimate the impacts of incarceration vs non-incarceration using the following 2SLS specification:

$$T_i = \delta_0 + \delta_1 Z_i + U$$
$$Y = \gamma_0 + \gamma_1 T_i + V$$

This specification is similar to equations (1) and (2) from the main text, but does not include judge dismissal stringency as a control. Under the standard LATE assumptions, γ_1 will not yield a weighted average of LATEs of incarceration vs non-incarceration, since an increase in Z_i could generate flows between dismissal and conviction in the non-incarcerated group if Z_i and Z_c are correlated, which is likely given that $Z_i = 1 - (Z_c + Z_d)$ by construction.

F.3 2SLS estimates with Empirical Bayes Shrinkage

We estimate judge stringency using leave-one-out means. To help ensure stringency measures are not too noisy, we restrict our analysis to judges who see at least 100 cases over the three-year windows we use to calculate stringency. We can further correct for potential measurement error using Empirical Bayes methods. Empirical Bayes was developed in the context of the teacher valued added literature (Chetty et al., 2014; Kane and Staiger, 2008), where the population distribution of teacher value added is typically assumed to be normally distributed, but measured with noise, also typically assumed to be normally distributed. This approach has also been applied to judge stringency measures in some papers (Arnold et al., 2022; Norris, 2019), using standard Empirical Bayes shrinkage procedures (Morris, 1983).

Here we similarly perform parametric Empirical Bayes, but we assume that the population distribution judge stringencies are drawn from a Beta distribution, and the individual stringencies follow a Bernoulli distribution. We believe these parametric assumptions are better than assuming normality since judge stringencies are probabilities.

We take two approaches. The first assumes judge stringencies are drawn from a single Beta distribution, while the second assumes the Beta distribution varies by district and year.

Empirical Bayes with a single Beta prior: First, we assume that judge stringencies are drawn from a $Beta(\alpha, \beta)$ distribution, and we estimate $\hat{\alpha}$ and $\hat{\beta}$ via maximum likelihood based on our sample of judge stringencies, which are calculated in 3-year bins by judge, restricting to judges who handle at least 100 cases.[67] Let's consider noncarceral conviction stringency (the same derivations apply for incarceration or dismissal stringencies). Let C_j be the number of cases ending in a noncarceral conviction for judge j and N_j be the total number of cases they handle. Based on the estimated Beta prior, the posterior conviction stringency is given by:

$$\frac{C_j + \alpha}{N_j + \alpha + \beta}.$$

We then adjust construct the leave-one-out posterior stringency as:

$$\frac{C_j - C_{j,i} + \alpha}{N_j - 1 + \alpha + \beta}.$$

where i represents that particular case.

Figure F.1 plots our main stringency measures (x-axis) against the estimates Empirical Bayes estimates (y-axis). The measures are similar; they largely fall close to the 45 degree line.

Panel (a) of table F.2 reports our main first stage estimates; panel (b) reports the first-stage estimates using the shrunk stringency estimates. The results are very similar. The first-stage coefficients and F-statistics are slightly larger when we use the Empirical Bayes estimates. Panel (a) of table F.3 reproduces our main estimates, and panel (b) reports our 2SLS estimates for noncarceral conviviction using the Empirical

[67] To simplify, we use "judge" and j subscripts though, as in the rest of the paper, these are three-year rolling averages.

Bayes stringencies. Results are nearly identical. Table F.4 produces a similar table for incarceration, with similar conclusions.

Empirical Bayes with priors that vary by district-year: So far, we have used the same Beta prior for all judges. Here, we estimate priors that vary parametrically by district-year. We can express $\alpha = \gamma/\sigma$ and $\beta = (1-\gamma)/\sigma$, where γ is the average stringency and σ is the spread. We then estimate $\gamma_j = \gamma_0 + \gamma_{d,y}$ where $\gamma_{d,y}$ shifts the average stringency by district-year. We estimate this regression using a Bayesian Beta-Binomial regression, then with estimates of γ_0 and $\gamma_{d,y}$, we construct α_j and β_j for each judge-district-year. We construct the leave-one-out posterior stringency as:

$$\frac{C_j - C_{j,i} + \alpha_j}{N_j - 1 + \alpha_j + \beta_j}.$$

This is very similar to the previous approach; the difference is that we now shrink judge stringencies towards the average stringency within district-year, rather than the overall average in our sample. This approach is appealing, as it allows the prior distribution to vary by district-year, but requires estimating many more parameters to recover our empirical priors.

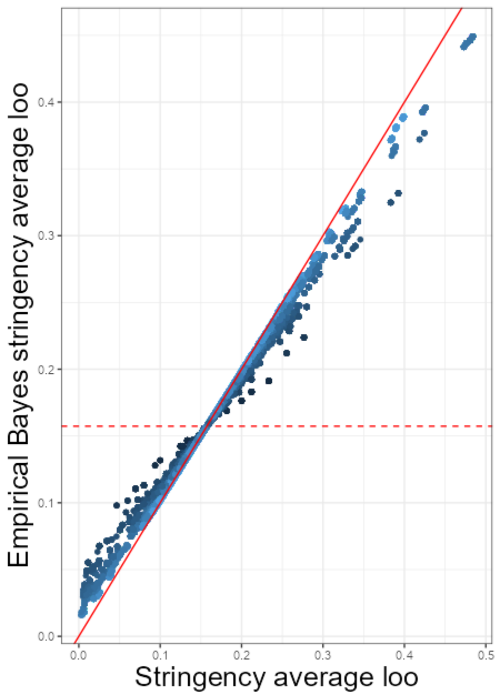
Analogous to panel (b), panel (c) of Table F.2 presents first-stage estimates using Empirical Bayes stringency with district-year priors. Here, we obtain first-stage coefficients that are closer to one, and larger F-statistics. A plausible interpretation is that this approach more effectively addresses measurement error in stringency measures.

Panel (c) of Table F.3 reports our main 2SLS estimates for noncarceral conviction using Empirical Bayes stringency with district-year priors. The results are very similar to our main specification, though estimates are somewhat smaller, particularly for the Year 1-7 time window where estimates are 12.5% to 23.3% smaller and the estimate on future charge is statistically significant at the 0.1 rather than 0.05 level. Table F.4 produces similar results for the 2SLS estimates of incarceration with similar conclusions.

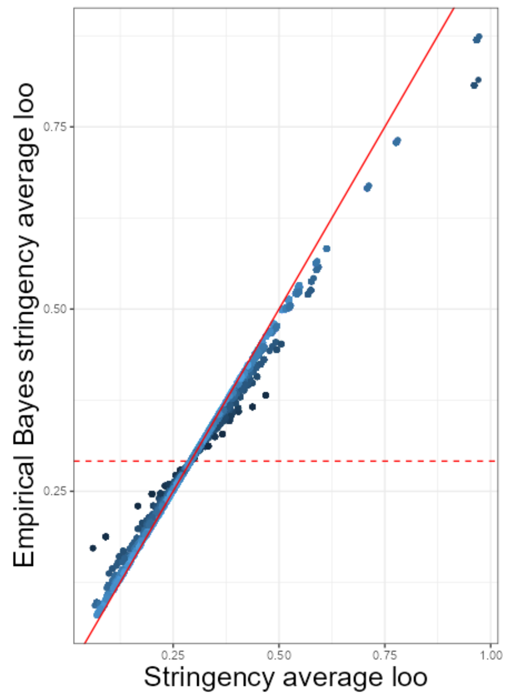
Overall, these results show that accounting for measurement error with either of the methods above does not qualitatively change our conclusions and does not lead to large quantitative differences.

Figure F.1: Leave-one-out raw stringency vs. leave-one-out empirical Bayes stringency

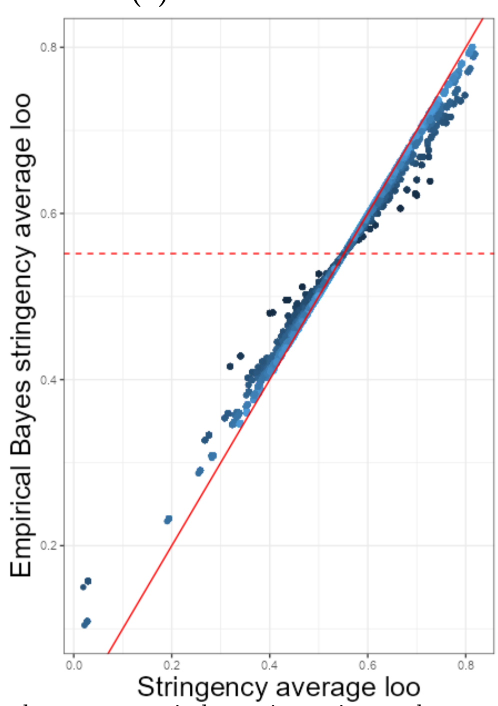
(a) Dismissal



(b) Conviction



(c) Incarceration



Notes: This figure compares the leave-one-out judge stringencies used on our main analysis to leave-one-out stringencies calculated via empirical Bayes with a single Beta prior. The brighter the blue points, the higher the total number of cases for judge j.

Table F.2: Relevance: first stage coefficients for the 2SLS analysis (Empirical Bayes shrinkage)

	Conviction			Incarceration		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: no shrinkage						
Conviction stringency	0.63***	0.60***	0.59***			
	(0.033)	(0.032)	(0.046)			
Incarceration stringency			-0.010	0.62***	0.59***	0.60***
			(0.041)	(0.033)	(0.032)	(0.035)
Dismissal stringency						0.033
						(0.051)
Controls	No	Yes	Yes	No	Yes	Yes
Mean dep. var.	0.298	0.298	0.298	0.546	0.546	0.546
F-stat	360.3	339.7	165.3	346.7	350.4	287.7
N	183,381	183,381	183,381	183,381	183,381	183,381
	(1)	(2)	(3)	(4)	(5)	(6)
Panel B: single Beta-prior (EB loo)						
Conviction stringency	0.69***	0.66***	0.65***			
	(0.035)	(0.034)	(0.048)			
Incarceration stringency			-0.011	0.68***	0.65***	0.66***
			(0.044)	(0.035)	(0.033)	(0.037)
Dismissal stringency						0.030
						(0.055)
Controls	No	Yes	Yes	No	Yes	Yes
Mean dep. var.	0.298	0.298	0.298	0.546	0.546	0.546
F-stat	397.0	373.7	177.9	369.0	372.9	308.4
N	183,381	183,381	183,381	183,381	183,381	183,381
	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: priors varying by district-year (BB loo)						
Conviction stringency	1.02***	0.97***	0.94***			
	(0.048)	(0.047)	(0.073)			
Incarceration stringency			-0.033	0.97***	0.94***	0.95***
			(0.067)	(0.050)	(0.047)	(0.051)
Dismissal stringency						0.044
						(0.094)
Controls	No	Yes	Yes	No	Yes	Yes
Mean dep. var.	0.298	0.298	0.298	0.546	0.546	0.546
F-stat	444.5	424.9	165.5	379.3	393.5	348.8
N	183,381	183,381	183,381	183,381	183,381	183,381

Note: This table compares the coefficients on the instruments from the first stage of the 2SLS regressions in our main analysis (Panel A) with the coefficients derived using Empirical Bayes with a single Beta prior (Panel B) and with the coefficients derived using Empirical Bayes with priors that vary by district-year (Panel C).

Table F.3: Noncarceral conviction and recidivism (Empirical Bayes shrinkage)

	Year 1		Year 2-4		Year 5-7		Year 1-7	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Panel A: no shrinkage								
Fut. charge	-0.002	0.105**	0.004	0.083	0.006**	0.077	0.011***	0.231**
	(0.002)	(0.046)	(0.003)	(0.075)	(0.002)	(0.075)	(0.004)	(0.097)
Fut. conviction	0.001	0.135***	0.008***	0.110	0.007***	0.055	0.014***	0.295***
	(0.002)	(0.043)	(0.003)	(0.072)	(0.002)	(0.071)	(0.004)	(0.095)
Fut. incarceration	0.001	0.111***	0.006**	0.055	0.005**	-0.025	0.012***	0.209**
	(0.002)	(0.037)	(0.002)	(0.063)	(0.002)	(0.057)	(0.003)	(0.083)
Ctrl comp. mean: fut. chrg.	0.158	0.158	0.302	0.302	0.237	0.237	0.494	0.494
Ctrl mean: fut. chrg.	0.089	0.089	0.170	0.170	0.129	0.129	0.297	0.297
Ctrl comp. mean: fut. conv.	0.138	0.138	0.264	0.264	0.225	0.225	0.460	0.460
Ctrl mean: fut. conv.	0.076	0.076	0.148	0.148	0.114	0.114	0.268	0.268
Ctrl comp. mean: fut. incar.	0.135	0.135	0.288	0.288	0.276	0.276	0.523	0.523
Ctrl mean: fut. incar.	0.054	0.054	0.109	0.109	0.083	0.083	0.204	0.204
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381
Panel B: single Beta prior (EB loo)								
Fut. charge	-0.002	0.109**	0.004	0.084	0.006**	0.075	0.011***	0.231**
	(0.002)	(0.046)	(0.003)	(0.072)	(0.002)	(0.073)	(0.004)	(0.094)
Fut. conviction	0.001	0.137***	0.008***	0.113	0.007***	0.056	0.014***	0.295***
	(0.002)	(0.043)	(0.003)	(0.070)	(0.002)	(0.069)	(0.004)	(0.092)
Fut. incarceration	0.001	0.108***	0.006**	0.055	0.005**	-0.028	0.012***	0.203**
	(0.002)	(0.037)	(0.002)	(0.061)	(0.002)	(0.055)	(0.003)	(0.080)
Ctrl comp. mean: fut. chrg.	0.156	0.156	0.299	0.299	0.236	0.236	0.492	0.492
Ctrl mean: fut. chrg.	0.089	0.089	0.170	0.170	0.129	0.129	0.297	0.297
Ctrl comp. mean: fut. conv.	0.136	0.136	0.263	0.263	0.224	0.224	0.458	0.458
Ctrl mean: fut. conv.	0.076	0.076	0.148	0.148	0.114	0.114	0.268	0.268
Ctrl comp. mean: fut. incar.	0.133	0.133	0.285	0.285	0.273	0.273	0.518	0.518
Ctrl mean: fut. incar.	0.054	0.054	0.109	0.109	0.083	0.083	0.204	0.204
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381
Panel C: priors varying by district-year (BB loo)								
Fut. charge	-0.002	0.102**	0.004	0.039	0.006**	0.088	0.011***	0.186*
	(0.002)	(0.047)	(0.003)	(0.075)	(0.002)	(0.075)	(0.004)	(0.097)
Fut. conviction	0.001	0.132***	0.008***	0.076	0.007***	0.067	0.014***	0.259***
	(0.002)	(0.045)	(0.003)	(0.072)	(0.002)	(0.072)	(0.004)	(0.097)
Fut. incarceration	0.001	0.112***	0.006**	0.019	0.005**	-0.026	0.012***	0.172**
	(0.002)	(0.039)	(0.002)	(0.063)	(0.002)	(0.058)	(0.003)	(0.084)
Ctrl comp. mean: fut. chrg.	0.131	0.131	0.269	0.269	0.213	0.213	0.448	0.448
Ctrl mean: fut. chrg.	0.089	0.089	0.170	0.170	0.129	0.129	0.297	0.297
Ctrl comp. mean: fut. conv.	0.115	0.115	0.238	0.238	0.200	0.200	0.417	0.417
Ctrl mean: fut. conv.	0.076	0.076	0.148	0.148	0.114	0.114	0.268	0.268
Ctrl comp. mean: fut. incar.	0.108	0.108	0.240	0.240	0.222	0.222	0.437	0.437
Ctrl mean: fut. incar.	0.054	0.054	0.109	0.109	0.083	0.083	0.204	0.204
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381

Note: This table compares the OLS and 2SLS regression estimates depicting the impact of noncarceral conviction on future recidivism in our main analysis (Panel A) with the estimates obtained using Empirical Bayes with a single Beta prior (Panel B) and with the coefficients derived using Empirical Bayes with priors that vary by district-year (Panel C).

Table F.4: Incarceration and recidivism (Empirical Bayes shrinkage)

	Year 1		Year 2-4		Year 5-7		Year 1-7	
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Panel A: no shrinkage								
Fut. charge	-0.022***	-0.096***	0.013***	-0.016	0.025***	0.003	0.023***	-0.071
	(0.002)	(0.029)	(0.002)	(0.047)	(0.002)	(0.040)	(0.003)	(0.059)
Fut. conviction	-0.018***	-0.111***	0.013***	-0.037	0.023***	0.020	0.022***	-0.106*
	(0.001)	(0.029)	(0.002)	(0.047)	(0.002)	(0.039)	(0.003)	(0.058)
Fut. incarceration	-0.010***	-0.071***	0.017***	0.009	0.021***	0.052	0.027***	-0.028
	(0.001)	(0.024)	(0.002)	(0.041)	(0.002)	(0.032)	(0.003)	(0.051)
Ctrl comp. mean: fut. chrg.	0.122	0.122	0.199	0.199	0.147	0.147	0.370	0.370
Ctrl mean: fut. chrg.	0.088	0.088	0.175	0.175	0.132	0.132	0.306	0.306
Ctrl comp. mean: fut. conv.	0.084	0.084	0.168	0.168	0.113	0.113	0.310	0.310
Ctrl mean: fut. conv.	0.077	0.077	0.159	0.159	0.120	0.120	0.283	0.283
Ctrl comp. mean: fut. incar.	0.043	0.043	0.071	0.071	0.051	0.051	0.166	0.166
Ctrl mean: fut. incar.	0.055	0.055	0.115	0.115	0.084	0.084	0.212	0.212
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381
Panel B: single Beta prior (EB loo)								
Fut. charge	-0.022***	-0.100***	0.013***	-0.016	0.025***	0.003	0.023***	-0.066
	(0.002)	(0.029)	(0.002)	(0.046)	(0.002)	(0.039)	(0.003)	(0.058)
Fut. conviction	-0.018***	-0.115***	0.013***	-0.037	0.023***	0.019	0.022***	-0.104*
	(0.001)	(0.028)	(0.002)	(0.046)	(0.002)	(0.038)	(0.003)	(0.056)
Fut. incarceration	-0.010***	-0.069***	0.017***	0.008	0.021***	0.049	0.027***	-0.029
	(0.001)	(0.024)	(0.002)	(0.040)	(0.002)	(0.032)	(0.003)	(0.050)
Ctrl comp. mean: fut. chrg.	0.124	0.124	0.200	0.200	0.147	0.147	0.373	0.373
Ctrl mean: fut. chrg.	0.088	0.088	0.175	0.175	0.132	0.132	0.306	0.306
Ctrl comp. mean: fut. conv.	0.087	0.087	0.170	0.170	0.113	0.113	0.314	0.314
Ctrl mean: fut. conv.	0.077	0.077	0.159	0.159	0.120	0.120	0.283	0.283
Ctrl comp. mean: fut. incar.	0.045	0.045	0.072	0.072	0.051	0.051	0.169	0.169
Ctrl mean: fut. incar.	0.055	0.055	0.115	0.115	0.084	0.084	0.212	0.212
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381
Panel C: priors varying by district-year (BB loo)								
Fut. charge	-0.022***	-0.106***	0.013***	0.001	0.025***	-0.004	0.023***	-0.055
	(0.002)	(0.028)	(0.002)	(0.045)	(0.002)	(0.038)	(0.003)	(0.056)
Fut. conviction	-0.018***	-0.116***	0.013***	-0.022	0.023***	0.012	0.022***	-0.094*
	(0.001)	(0.028)	(0.002)	(0.045)	(0.002)	(0.037)	(0.003)	(0.054)
Fut. incarceration	-0.010***	-0.073***	0.017***	0.016	0.021***	0.044	0.027***	-0.025
	(0.001)	(0.023)	(0.002)	(0.038)	(0.002)	(0.031)	(0.003)	(0.048)
Ctrl comp. mean: fut. chrg.	0.112	0.112	0.193	0.193	0.142	0.142	0.353	0.353
Ctrl mean: fut. chrg.	0.088	0.088	0.175	0.175	0.132	0.132	0.306	0.306
Ctrl comp. mean: fut. conv.	0.081	0.081	0.165	0.165	0.115	0.115	0.302	0.302
Ctrl mean: fut. conv.	0.077	0.077	0.159	0.159	0.120	0.120	0.283	0.283
Ctrl comp. mean: fut. incar.	0.045	0.045	0.085	0.085	0.061	0.061	0.180	0.180
Ctrl mean: fut. incar.	0.055	0.055	0.115	0.115	0.084	0.084	0.212	0.212
Observations	183,381	183,381	183,381	183,381	183,381	183,381	183,381	183,381

Note: This table compares the OLS and 2SLS regression estimates depicting the impact of incarceration on future recidivism in our main analysis (Panel A) with the estimates obtained using Empirical Bayes with a single Beta prior (Panel B) and with the coefficients derived using Empirical Bayes with priors that vary by district-year (Panel C).

F.4 Calculating control means for compliers

To calculate control-group complier means, we follow Dahl et al. (2014) and Agan and Starr (2018). First we show how to derive control-group complier means for the simple case of binary treatment and a binary instrument. We then expand this to our setting.

In the simple case where $Z \in 0, 1$ and $D \in 0, 1$, we aim to calculate $E[Y(0) \mid D(1) > D(0)]$. Here $Y(0)$ is the potential outcome when $D = 0$, $D(1)$ is the potential treatment when $Z = 1$, and $D(0)$ is the potential treatment when $Z = 0$. Note that

$$
\begin{aligned}
\underbrace{E[Y|D=0,Z=0]}_{\text{data}} =& \frac{\pi_c}{\pi_c+\pi_n}\underbrace{E[Y(0)|D(1)>D(0)]}_{\text{unknown}} + \frac{\pi_n}{\pi_c+\pi_n}E[Y(0)|D(1)=D(0)=0] \\
\text{where} & \\
\pi_n =& \underbrace{Pr(D=0|Z=1)}_{\text{data}} \\
\pi_a =& \underbrace{Pr(D=1|Z=0)}_{\text{data}} \\
\pi_c =& 1-\pi_n-\pi_a.
\end{aligned}
$$

In the expression above, the terms with "data" below them can be estimated directly from the data. The term $E[Y(0)|D(1) = D(0) = 1] = E[Y|D = 0, Z = 1]$, where the right-hand term can also be estimated directly from the data. This leaves only one unknown term: $E[Y(0)|D(1) > D(0)]$, which is the term of interest. Re-arranging the equations and plugging in, we get:

$$
E[Y(0)|D(1) > D(0)] = \frac{\pi_c + \pi_n}{\pi_c} E[Y|D = 0, Z = 0] - \frac{\pi_n}{\pi_c} E[Y|D = 0, Z = 1],
$$

where all the terms on the right side of the equality can be estimated from the data.

Our setting differs from this setting above as we have a continuous instrument, and D can take on 3 values. We follow Dahl et al. (2014) and Agan and Starr (2018) in adapting the math above to the case with continuous instruments. We use code from the replication file of Agan and Starr (2018), which is adapted from Dahl et al. (2014). This adaptation involves calculating the minimum and maximum values of the instrument (z_{min} and z_{max}). Following the papers above, we can then adapt the equations to be:

$$
\begin{aligned}
\underbrace{E[Y|D=0,Z=z_{min}]}_{\text{data}} =& \frac{\pi_c}{\pi_c+\pi_n}\underbrace{E[Y(0)|D(z_{max})>D(z_{min})]}_{\text{unknown}} \\
&+ \frac{\pi_n}{\pi_c+\pi_n}E[Y(0)|D(z_{max})=D(z_{min})=0] \\
\text{where} & \\
\pi_n =& \underbrace{Pr(D=0|Z=z_{max})}_{\text{data}} \\
\pi_a =& \underbrace{Pr(D=1|Z=z_{min}])}_{\text{data}} \\
\pi_c =& \beta * (z_{max} - z_{min})
\end{aligned}
$$

where β is from the regression of D on the instrument. Similar to the binary case, we have $E[Y(0)|D(z_{min}) = D(z_{max}) = 1] = E[Y|D = 0, Z = z_{max}]$. We use the first and 99th percentiles of the residualized instrument for z_{min} and z_{max}, respectively.

To address the fact that we consider multiple treatments, we include non-focal judge stringency as an additional control. For example, if D is the indicator for conviction, we use judge conviction stringency as the instrument, controlling for judge incarceration stringency. Under UPM and the other IV assumptions laid out in the main paper, the only compliers will be those shifting from $T = d$ to $T = c$ and, therefore, capture the margin-specific compliers of interest.

G Additional details for multinomial model with heterogeneous effects

This appendix discusses how we apply Mountjoy (2022) in our setting. First, we describe the identification and estimation of margin-specific treatment effects. Then, we report additional empirical results.

G.1 Additional details on identification and estimation

This subsection summarizes how we adapt the identification and estimation strategies from Mountjoy (2022) to obtain the results in Section 5. To begin, we state the "comparable compliers" assumption of Mountjoy (2022) in our notation:

A7. Comparable Compliers (CC)

For all $\tilde{z}_c$ and $\tilde{z}_i$,

$$\begin{aligned}&\lim_{\tilde{z}_c' \uparrow \tilde{z}_c} E[Y(c)|T(\tilde{z}_c', \tilde{z}_i) = c, T(\tilde{z}_c, \tilde{z}_i) = i]\\ =&\lim_{\tilde{z}_i' \downarrow \tilde{z}_i} E[Y(c)|T(\tilde{z}_c, \tilde{z}_i') = c, T(\tilde{z}_c, \tilde{z}_i) = i].\end{aligned}$$

This assumption says that $i \to c$ compliers from decreasing $\tilde{z}_i$ have the same potential outcome when convicted as $i \to c$ compliers from increasing $\tilde{z}_c$ at their limits, where $\tilde{z}_i$ and $\tilde{z}_c$ are the treatment-specific instruments.

Given a treatment-specific instrument for conviction, it is possible to identify a weighted average of two LATEs that are specific to two different margins as captured by the following expression:

$$LATE_c = \omega LATE_{d \to c} + (1 - \omega) LATE_{i \to c}.$$

This decomposition is visualized in Panel (c) of Figure 3, which shows that such variation induces two sets of compliers, those moving from $T = d$ to $T = c$ (in yellow) and those moving from $T = i$ to $T = c$ (in green).

Mountjoy (2022) shows that it is possible to recover the two margin-specific LATEs, as well as ω, by using variation in two treatment-specific instruments to construct the relevant expected potential outcomes for the two groups. His identification results directly apply once we have recovered choice-specific instruments.

We also follow Mountjoy (2022) in estimation. For example, we similarly assume the relevant conditional expectations are well approximated by a local linear regression centered around the chosen evaluation point of the instruments. These regressions include additive controls as specified in the notes of Table 7. We use an Epanechnikov kernel with a bandwidth of 3 and report estimates evaluated at the mean value of the instruments. This approach produces similar estimates when using smaller or larger bandwidths. Inference is based on 500 bootstrap samples. We report 95% confidence intervals based on the bootstraps and significance stars based on the 90%, 95%, and 99% two-sided confidence intervals.

We refer the reader to Mountjoy (2022) for a full discussion of identification and estimation.

G.2 Additional results

Tables G.1 and G.2 provide additional results under alternative assumptions used to construct the treatment-specific instruments. The first set of results comes from assuming a standard multinomial logistic model. While restrictive, this allows for a simple closed-form solution for constructing thresholds from shares, as explained in the main paper. The second mirrors the mixed model reported in Table 7, but assumes the random effects follow an independent multivariate normal distribution. Confidence intervals for all three approaches are calculated using 500 bootstrap samples.

Overall, the results in Tables G.1 and G.2 are similar in magnitude to Table 7, although the estimates are somewhat larger and tend to be closer to the 2SLS estimates reported in the paper.

Table G.1: Margin-specific treatment effects: alternative approach (robustness, multinomial logit)

	simple log-ratio			
	Year 1	Year 2-4	Year 5-7	Year 1-7
Panel A: Noncarceral conviction vs dismissal (C vs D)				
Felony charge:	0.067	0.172**	0.187**	0.264**
	[-0.024,0.183]	[0.015,0.353]	[0.047,0.390]	[0.035,0.519]
	{0.077}	{0.173}	{0.136}	{0.344}
Felony conviction:	0.090*	0.216**	0.156**	0.342***
	[-0.017,0.205]	[0.064,0.403]	[0.011,0.327]	[0.133,0.636]
	{0.068}	{0.132}	{0.146}	{0.282}
Felony incarceration:	0.056	0.149**	0.074	0.212**
	[-0.024,0.150]	[0.009,0.305]	[-0.038,0.236]	[0.018,0.400]
	{0.070}	{0.107}	{0.123}	{0.298}
Panel B: Incarceration vs noncarceral conviction (I vs C)				
Felony charge:	-0.043**	0.036	0.002	-0.037
	[-0.074,-0.007]	[-0.024,0.113]	[-0.061,0.072]	[-0.139,0.061]
	{0.084}	{0.178}	{0.138}	{0.334}
Felony conviction:	-0.035**	0.031	0.019	-0.027
	[-0.071,-0.003]	[-0.034,0.101]	[-0.050,0.092]	[-0.130,0.075]
	{0.074}	{0.163}	{0.120}	{0.306}
Felony incarceration:	-0.012	0.045	0.014	-0.037
	[-0.044,0.020]	[-0.020,0.106]	[-0.041,0.070]	[-0.141,0.077]
	{0.054}	{0.109}	{0.099}	{0.241}
Controls	Yes	Yes	Yes	Yes

Note: This table reports margin-specific estimates of the impact of noncarceral conviction vs dismissal (Panel A) and incarceration vs noncarceral conviction (Panel B) using an unordered multinomial model based on the methodology developed in Mountjoy (2022). The methodology is described in the notes of Table 7, except that here judge-specific latent preferences are calculated under the stronger assumption that case outcomes are determined by a multinomial logit. The curly brackets report control-group complier means. 95% confidence intervals are reported in brackets and are based on 500 bootstrap samples. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01 based on the 90%, 95%, and 99% confidence intervals.

Table G.2: Margin-specific treatment effects: alternative approach (robustness, independent mixed logit)

	mixed logit with independent normal random effects			
	Year 1	Year 2-4	Year 5-7	Year 1-7
Panel A: Noncarceral conviction vs dismissal (C vs D)				
Felony charge:	0.077*	0.185***	0.124**	0.209***
	[-0.007,0.159]	[0.069,0.309]	[0.004,0.234]	[0.060,0.368]
	{0.059}	{0.140}	{0.120}	{0.297}
Felony conviction:	0.086**	0.198***	0.110**	0.260***
	[0.011,0.151]	[0.092,0.319]	[0.007,0.220]	[0.106,0.447]
	{0.049}	{0.117}	{0.116}	{0.247}
Felony incarceration:	0.059*	0.137**	0.061	0.175**
	[-0.006,0.121]	[0.035,0.248]	[-0.035,0.153]	[0.006,0.342]
	{0.048}	{0.098}	{0.095}	{0.234}
Panel B: Incarceration vs noncarceral conviction (I vs C)				
Felony charge:	-0.054***	0.008	-0.021	-0.082*
	[-0.089,-0.020]	[-0.043,0.068]	[-0.076,0.036]	[-0.176,0.010]
	{0.090}	{0.181}	{0.150}	{0.354}
Felony conviction:	-0.044***	0.004	-0.012	-0.079*
	[-0.076,-0.008]	[-0.055,0.061]	[-0.065,0.042]	[-0.165,0.012]
	{0.079}	{0.168}	{0.133}	{0.330}
Felony incarceration:	-0.019	0.024	-0.006	-0.079**
	[-0.047,0.008]	[-0.026,0.076]	[-0.056,0.046]	[-0.158,-0.001]
	{0.057}	{0.111}	{0.103}	{0.257}
Controls	Yes	Yes	Yes	Yes

Note: This table reports margin-specific estimates of the impact of noncarceral conviction vs dismissal (Panel A) and incarceration vs noncarceral conviction (Panel B) using an unordered multinomial model based on the methodology developed in Mountjoy (2022). The methodology is described in the notes of Table 7, except that here judge-specific latent preferences are calculated under the stronger assumption that the intercepts include a random effect that is an uncorrelated multivariate normal. The curly brackets report control-group complier means. 95% confidence intervals are reported in brackets and are based on 500 bootstrap samples. Stars denote * p< 0.10, ** p< 0.05, *** p< 0.01 based on the 90%, 95%, and 99% confidence intervals.

H Impacts of incarceration: additional evidence from sentencing guidelines

In this Appendix, we provide supporting evidence on the effects of incarceration, exploiting an independent source of variation: discontinuous changes in recommended sentences in the Virginia sentencing guidelines. Although judges have the final say over sentencing in Virginia, each person convicted of a felony gets a guidelines-recommended sentence which is calculated using a series of worksheets. Sentence recommendations change discontinuously at some scores. Exploiting two different discontinuities, we estimate the effects of incarceration on the intensive margin (sentence length) and on the extensive margin (short jail sentences vs probation). We are also able to provide evidence on the extensive margin for those who had never previously been incarcerated.

H.1 Empirical setup

Sample and data. For these analyses, we focus on people who were convicted of a felony in Virginia and use data from the Virginia Criminal Sentencing Commission (VCSC). See Appendix B for more details on the data and sample construction, and Table H.1 for summary statistics on our sample.

Calculating the sentencing score. The Virginia sentence guidelines were developed in the 1980s to harmonize practices across judges and reduce disparities across similar defendants (Farrar-Owens, 2013). Information on the sentence guidelines is available to all parties during negotiations.

The diagram in Figure H.1 describes the order in which the different sentencing worksheets are filled out. The first worksheet determines whether a person convicted of a felony is recommended for prison (more than one year of incarceration). This worksheet, called "Worksheet A", consists of a series of questions pertaining to the offense and criminal history. Each question has a number of points associated with it; the sum of these points is the "incarceration-length score." Those who score above a cutoff are recommended for prison. Those who score below the cutoff are recommended for probation or jail, where recommended jail sentences are under a year in length.

Based on the cutoff in Worksheet A, either Worksheet B (for those below the cutoff) or Worksheet C (for those above the cutoff) is used to calculate the final guidelines-recommended sentence. Worksheet B also has a discontinuity that is useful for our analysis. Defendants who score above a particular cutoff on the "probation/jail score" are recommended for short jail sentences, while defendants who score below that cutoff are recommended for probation.

Offenses are sorted into 16 offense categories, and each category has a slightly different worksheet. The worksheets are filled out by a probation officer or a prosecutor and then given to a judge during sentencing. The worksheet package contains a cover sheet, which has a summary of information related to the case. The guidelines-recommended sentence and range is displayed prominently on the cover sheet. An example of Worksheet A can be found in Appendix H.7; the other worksheets follow a similar organization.

Empirical approach. To conduct this analysis, we compare people who score just below and just above our worksheet thresholds. The main assumption for this to yield causal estimates of the effects of tougher sentences is that potential outcomes are smooth across the cutoff. This might not hold if, for example, legal actors are able to manipulate the scores. Three institutional details in our setting help mitigate this concern. First, the sentence guidelines are discretionary, not binding. Thus it is not necessary for legal actors to manipulate the score to achieve a certain sentence. Second, legal actors may pay more attention to the final recommended sentence as calculated on Worksheet B or Worksheet C, rather than the intermediary score calculated on Worksheet A. Therefore, concerns of manipulation on the incarceration-length score (derived from Worksheet A) might not be as strong, simply because it's less salient. Third, from the legislator's standpoint, the goal of these worksheets was to reduce unjustified disparities. Therefore, it seems unlikely that the sharp sentencing discontinuities observed at the cutoff in the incarceration-length score were created on purpose. In Section H.4 below, we provide evidence that there is no change in characteristics at the cutoff, along with tests for bunching in the running variable on either side of the cutoff.

An additional challenge in our setting is that the running variable is discrete, generating difficulties in estimating accurate confidence intervals. To address this, we adopt the technique developed by Kolesár and Rothe (2018) – "K&R" henceforth – designed specifically for regression discontinuity with a discrete running variable. As in other RD settings, we want to estimate a function of the form:

$$y_{i,s} = \beta * \mathbb{1}(s \geq 0) + f(s) * \mathbb{1}(s \geq 0) + g(s) * \mathbb{1}(s < 0) + \epsilon \tag{14}$$

where $y_{i,s}$ is the outcome of the person in case i having obtained a sentencing score of s.[68] Our main coefficient of interest is β. The challenge is to estimate the form of $f(.)$ and $g(.)$, especially close to the cutoff.

Typical approaches in RD consist of fitting specifications on either side of the cutoff. However, these approaches assume that bias can be minimized by reducing the bandwidth. In the discrete setting, the bandwidth cannot asymptotically go to zero, because there are no observations in between each discrete bin. The scarcity of points close to the cutoff could lead to misspecification error: in the absence of additional assumptions, it is unclear what the behavior of the functions of interest would be close to the cutoff, resulting in misspecified confidence intervals.

K&R offer an approach to determine confidence intervals, by estimating plausible behaviors of the potential outcome function close to the cutoff based on its behavior at other points. By fitting a linear regression through points at the left and right of the cutoff, we might be missing non-linearities closer to the cutoff. We cannot use observations "very close" to the cutoff to estimate this, since the discrete nature of the score hinders the credibility of limit arguments. K&R determine credible bounds for the second derivatives of $f(.)$ and $g(.)$ close to the cutoff, based on its behavior further from the cutoff, to estimate the magnitude of plausible deviations from the linear estimation. We need to choose a parameter K which is the upper bound of the absolute value of the second derivative of the conditional expectation function. This tells us how quickly the functions $f(.)$ and $g(.)$ can change. Using K, we can

[68] As a reminder, the sentencing score is either the incarceration-length score or the probation/jail score.

construct confidence intervals that reflect how far away from the linear approximation the true conditional expectation function might be based on its expected behavior at other points.

To choose K, we follow the approach developed by Imbens and Wager (2019) and implemented by Goldsmith-Pinkham et al. (2023). We take a large window of nine points to the left of the cutoff and fit a quadratic function of the sentencing score to the data.[69] We take the coefficient on the quadratic term, take the absolute value, and multiply it by four. Intuitively, this means that we allow the rate of change (2nd derivative) of $f(.)$ at the cutoff to be two times that of the estimated rate of change between -9 and -1 from a second order polynomial. When we estimate the optimal bandwidth, we obtain an optimal choice of equal to or close to 5 for many of our main outcomes. In order to keep bandwidths constant across outcomes and time periods, we use a bandwidth of 5 in all specifications.

H.2 Intensive margin: effects of longer carceral sentences.

As expected from the way worksheets are designed, we find that small differences in the incarceration-length score translate into large changes in people's sentences. Columns 1 and 2 of Table H.2 show the regression discontinuity results and Appendix Figure H.4 presents graphic evidence. Scoring above the threshold generates large (42 ppt) changes in the probability of having a sentence greater than one year, and sentences are on average eight months longer, compared to the control-group mean of 4 months.[70]

By comparing people on either side of the threshold, we can estimate the causal effect on new criminal justice contact of going from a sentence of approximately four months to approximately one year. Columns 3-9 of Table H.2 present outcomes in various time periods, from year 1 to year 8-10 after a person's sentencing date.

Our results are consistent with those estimated using quasi-random assignment of cases to judges. In the first year after sentencing, people above the cutoff are less likely to recidivate. This is likely due to an incapacitation effect: those right below the cutoff have an average sentence of four months, while those right above have an average sentence of 12 months. However, in the longer run, this effect disappears, with no significant difference in recidivism. In our ten-year cumulative measure, we can reject anything larger than a 1.2 percentage point increase in new felony charges over a control group mean of 46%.

H.3 Extensive margin: effects of exposure to incarceration

We found no evidence that tripling the sentence length (from approximately four to 12 months) affected future criminal justice contact. This may be because the impacts of incarceration accrue rapidly in the first several months. For example, a few months in jail might lead a person to lose their job, or to experience ruptures in their family lives (Dobbie et al., 2018). We can test the impact of initial exposure by looking at variation in outcomes for people who score just above or just below the cutoff in the

[69] We focus on the left of the cutoff, since we have more observations there.

[70] Control-group means are calculated for people whose score is below the relevant cutoff, and whose score is within the bandwidth used in that RD estimate.

probation/jail score. The first two columns of Panel A of Table H.3 show that scoring above the threshold translates into a 43 ppt increase in the likelihood of receiving a carceral sentence, and the average sentence length increases by 0.73 months (Figure H.5 presents graphic evidence for this extensive margin). Estimates from the probation/jail sample therefore capture the effect of a short jail sentence relative to probation only.[71] Columns 3-5 of Panel A of Table H.3 present results for recidivism. Given that sentences around the cutoff are so short in the sample, we look at short-term results using the six months after sentencing, and longer-term results looking 2-3 years after sentencing. Here, we find no evidence of a short-term incapacitation effect—likely because the difference in sentences is only about a month.[72] As previously, we find no evidence of longer-term effects. In our 1-3 year cumulative measure we can reject anything larger than a 0.007 percentage point increase over a control mean of 20%.

It is also possible that a person's very first incarceration spell may be particularly destabilizing or traumatic. To get at that question, we re-run our analysis on the portion of the probation/jail sample who had not been incarcerated previously, and who had not been detained pretrial.[73] This lowers our sample size substantially, particularly since data on pretrial detention is only available after 2010. As seen in Panel B of Table H.3, there is still a strong discontinuity in the likelihood of receiving a carceral sentence for those right above the cutoff, but no evidence of a change in outcomes once the original carceral sentence is complete. However, the estimates are noisy and we can't reject moderate changes in either direction.

These results are very similar to those obtained exploiting quasi-random assignment of cases to judges: we find short-term decreases in criminal justice contact, likely due to incapacitation, but we do not identify any longer-term impacts of exposure to incarceration. Table E.2 Panel B and Table H.6 present complier characteristics for the IV analyses, and characteristics of defendants who score just above or just below the relevant cutoffs. There are similarities across these groups, but also some small differences. For example, marginal defendants in the RD analysis are more likely to have been convicted with a drug crime compared to the IV compliers—especially for the extensive margin analyses.

H.4 Balance and marginal cases

Balance tests. Figure H.2 (H.3) and Table H.4 (H.5) present balance tests for the intensive margin experiment based on Worksheet A (extensive margin experiment, Worksheet B). We first perform analyses of defendant characteristics, such as demographics or criminal history, and find no notable discontinuities. We then turn to legal actor decisions. Since inputs to the worksheets and how they translate into sentences is common knowledge, it is possible that some savvy legal actors might try to manipulate inputs. For example, a better defense attorney might push harder to drop certain charges if their client has a score close to the cutoff, in order to push them just below the cutoff and avoid longer recommended sentences. If defense attorneys were trying

[71] Short sentences such as those experienced right above the cutoff are not atypical. For example, in Pennsylvania, the average amount of time spent in jail post-sentencing upon release is 2.4 months (PASC, 2013).

[72] We do find short-term incapacitation effects when looking at quarterly data.

[73] Our data is limited to Virginia; it is possible that they had experienced incarceration in another state.

to push their clients to the left of the cutoff, one way this could manifest is by having more charges dropped just before the cutoff. That is because some of the points are linked to number of offenses for which a person is convicted. This does not seem to be happening. We also look at measures of defendant poverty, which can affect quality of representation (Agan et al., 2021).[74] We do not find evidence of a discontinuity at the cutoff, suggesting that quality of representation does not change at this point.

We do find one difference: defendants in the incarceration-length sample are about 2.8 percentage points more likely to have their case resolved by plea just before than just after the cutoff (Panel B of table H.4). This result could be because the longer sentences offered to those just above the threshold make people more willing to "risk it" in court. Since taking the case to trial increases the likelihood of dismissal by 10 percentage points, a 2.8 percentage point increase in the trial rate would lead to losing 0.28% of the sample right above the threshold. Given how small the differences in conviction is at the threshold, and the fact that we see no detectable differences in observable characteristics, we think that this is unlikely to affect our research design too much. We also note that we do not find this discontinuity for the probation/jail sample, so these concerns do not apply to that set of analyses.

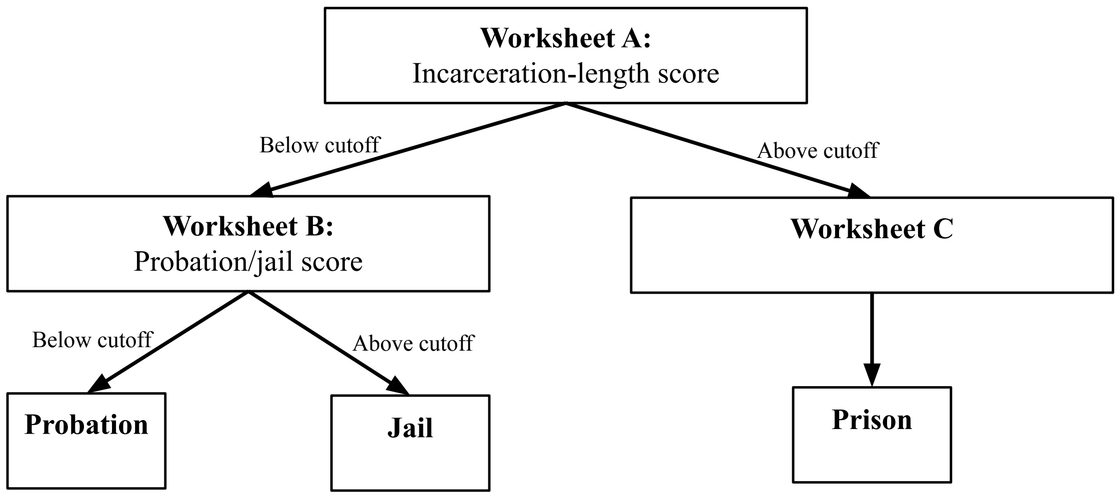
Lastly, we examine the distribution of the running variables to evaluate whether there is excess mass right above or below the cutoff. Such excess mass would be consistent with strategic manipulation of the scores to nudge defendants above or below the discontinuity in guidelines-recommended sentence. These analyses are shown in Figures H.2 (a) and H.3 (a) for the incarceration-length score and the probation/jail score, respectively. Visual inspection reveals possible excess mass below the cutoff for the incarceration-length score. Though, the distribution is not smooth, making it hard to infer whether this bunching is just a natural byproduct of a lumpy distribution or the result of strategic manipulation. There is no visible bunching around the cutoff for the probation/jail score.

Marginal case. Appendix Table H.6 compares the characteristics of marginal cases to those of the full sample in the relevant experiment, where marginal cases are defined as those scoring right below or right above the cutoff. The biggest difference between marginal cases and the full sample for Worksheet A is that marginal cases are much more likely to have prior incarceration: 87% had been incarcerated in the past, compared to 65% for the sample overall. This set aside, marginal cases are similar across offenses, but tend to be slightly younger. For worksheet B, there are differences across offense types: people convicted of a drug offense are more likely to be moved by the policy, while people convicted with property crimes are less so. Marginal cases are also more likely to have been incarcerated in the past (65% compared to 54%). Note that the marginal cases in the RD and IV experiments are different (as an example, 21% of the IV incarceration marginal cases had a prior felony conviction in the last 5 years, compared to 85% of the RD marginal cases). Yet, our results are similar across both experiments, suggesting that the differences in composition are not yielding different findings.

[74]We proxy poverty by whether a defendant comes from zip codes where the percent of people reporting less than $25,000 (less than $10,000) per year to the IRS was above the median within our sample.

H.5 Appendix figures: RD analyses

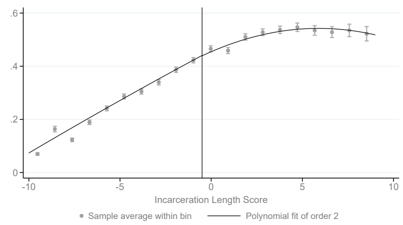
Figure H.1: Flowchart of felony sentencing determination in Virginia



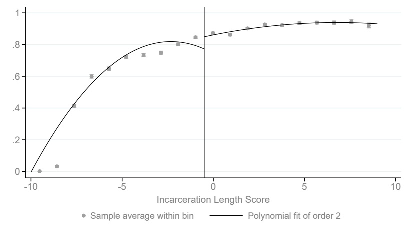
Note: This figure presents a flowchart describing the sentencing process in Virginia after a felony conviction, and how and when different Worksheets are used.

Figure H.2: Balance tests – incarceration-length sample

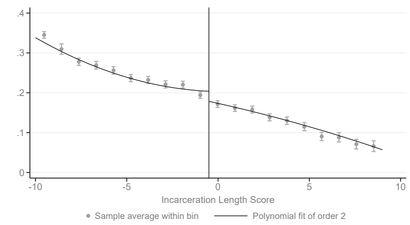
(a) Prior conviction



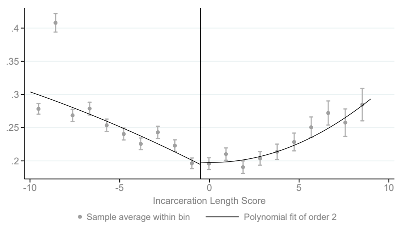
(b) Prior incarceration



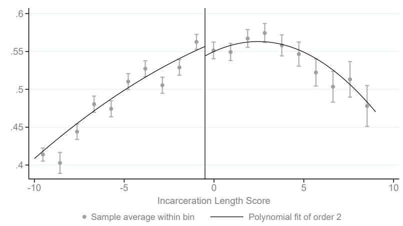
(c) Female



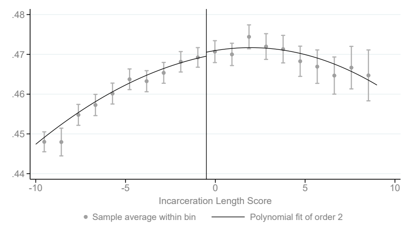
(d) Young



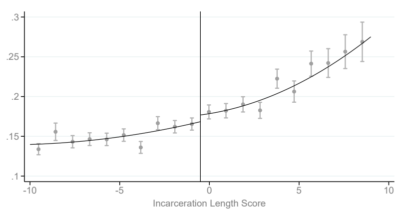
(e) Black



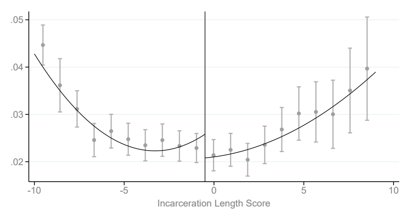
(f) Percent earning < $25K in zip code



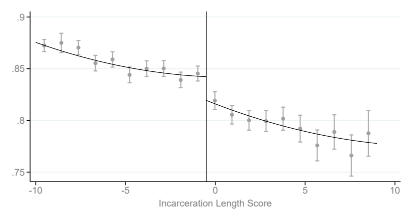
(g) Any charges dropped



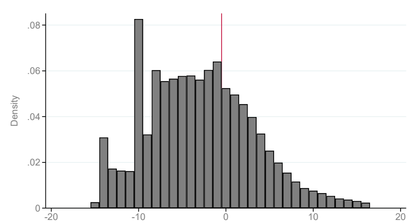
(h) Any up-charging



(i) Guilty plea



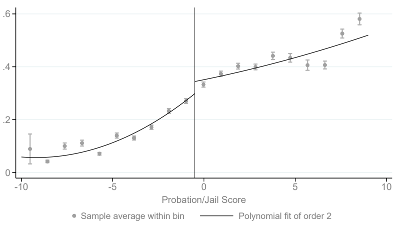
(j) Histogram



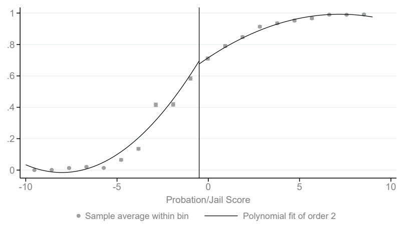
Note: Panels (a) - (i) show RD plots for various demographic variables and case characteristics. Panel (j) shows the distribution of incarceration-length scores around the cutoff. The incarceration-length score is normalized so that the cutoff is at zero.

Figure H.3: Balance tests – probation/jail sample

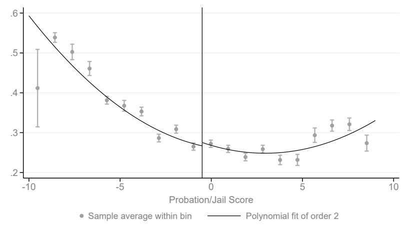
(a) Prior conviction



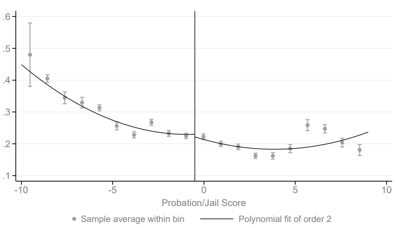
(b) Prior incarceration



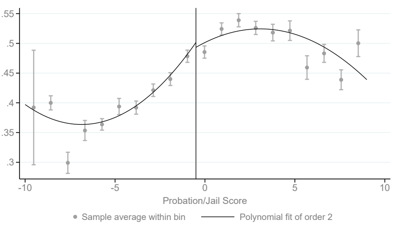
(c) Female



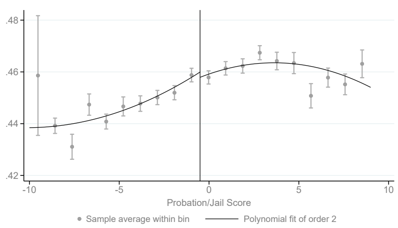
(d) Young



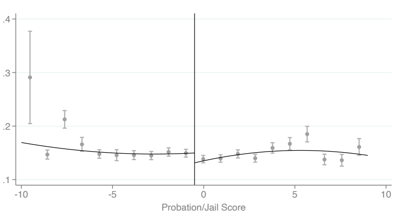
(e) Black



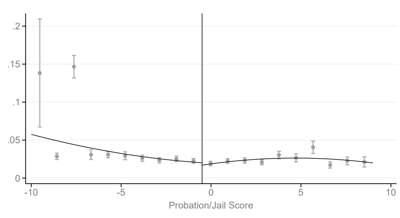
(f) Percent earning < $25K in zip code



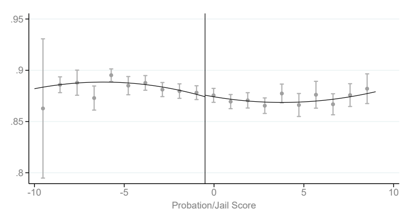
(g) Any charges dropped



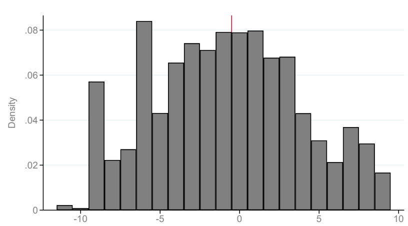
(h) Any up-charging



(i) Guilty plea



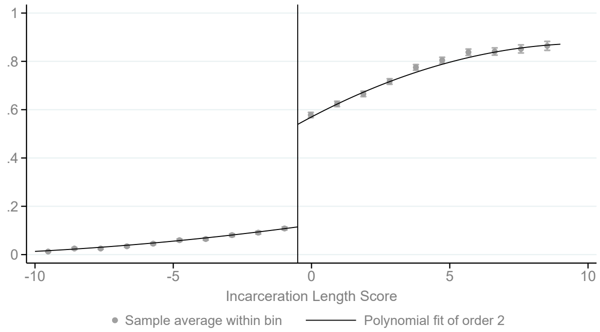
(j) Histogram



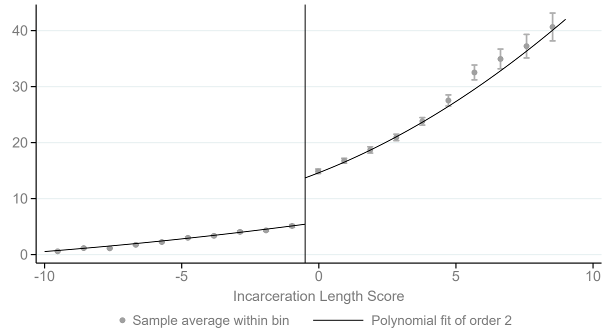
Note: Panels (a) - (i) show RD plots for various demographic variables and case characteristics. Panel (j) shows the distribution of probation/jail scores around the cutoff. The probation/jail score is normalized so that the cutoff is at zero.

Figure H.4: RD first stage and outcome graphs – incarceration-length sample

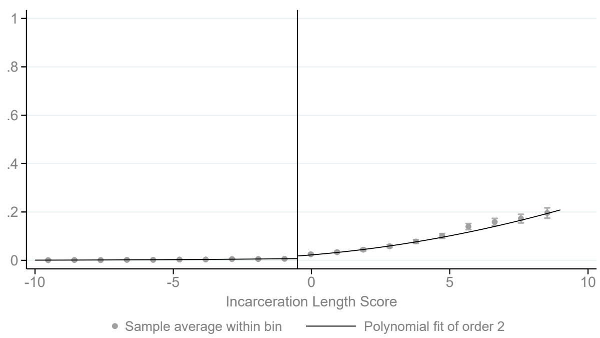
(a) Incarcerated for at least 1 year



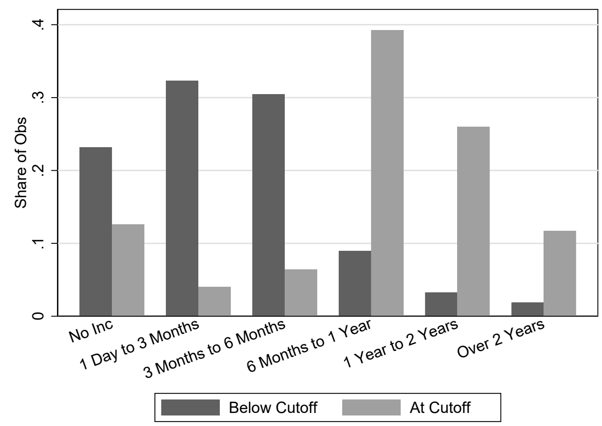
(b) Sentence in months



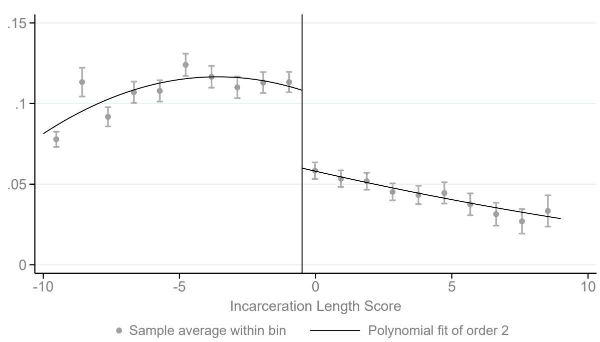
(c) Incarcerated for at least five years



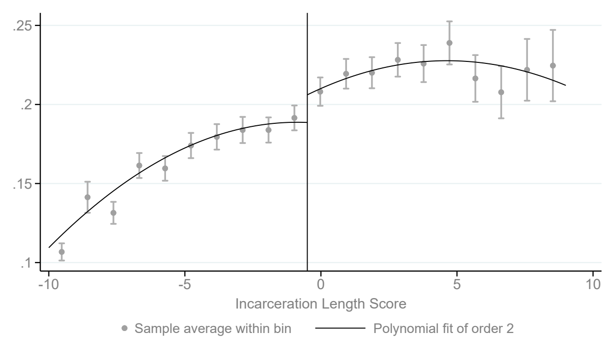
(d) Sentencing distributions for those right above/below the cutoff



(e) New charge within 1 year of sentencing



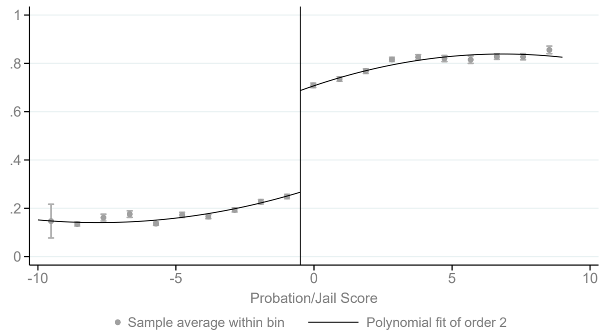
(f) New charge in years 5-7 after sentencing



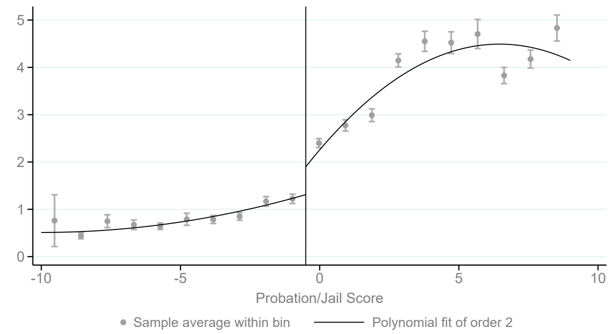
Note: Panel (a) shows the RD plot for being incarcerated for at least one year around the discontinuity in the incarceration-length score. Panel (b) shows the same plot for months sentenced and panel (c) shows the same plot for being sentenced to at least five years. Panel (d) shows the distribution of sentence lengths for those just above and just below the cutoff. Panel (e) shows RD plots for recidivism-defined as a binary variable for having at least one new charge one year post-sentencing and panel (f) shows recidivism within 5-7 years post-sentencing.

Figure H.5: RD First stage and outcome graphs – probation/jail score

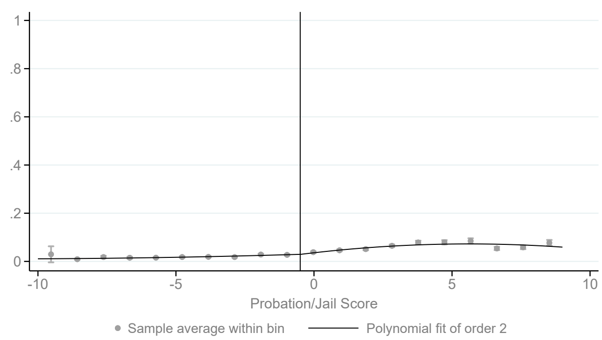
(a) Incarcerated at all



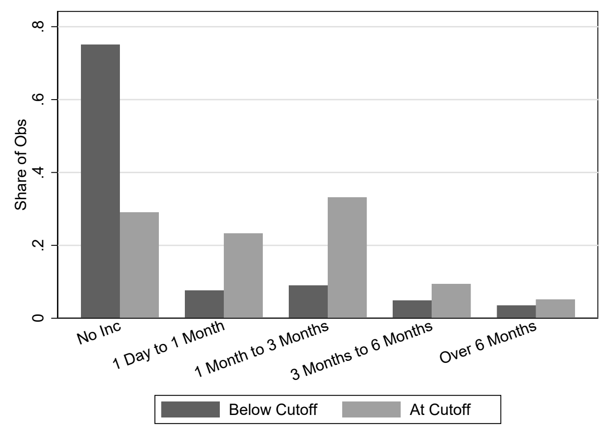
(b) Sentence in months



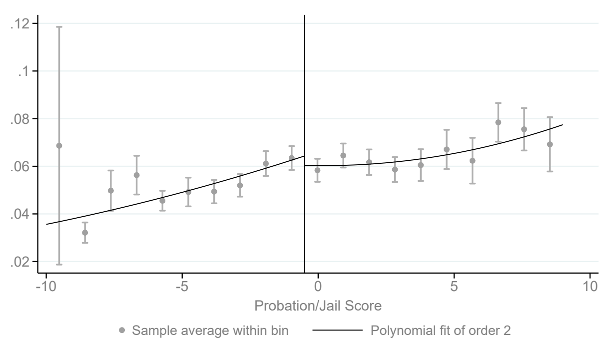
(c) Incarcerated for at least one year



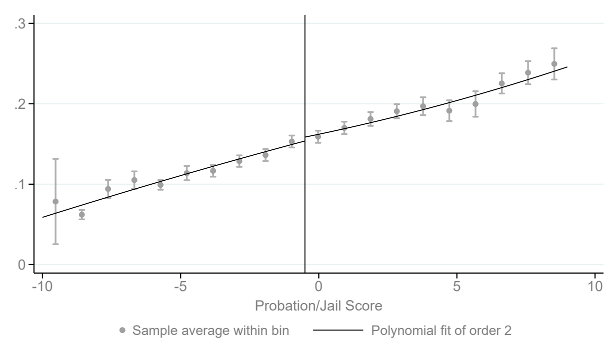
(d) Sentencing distributions for those right above/below the cutoff



(e) New charge within 6 months of sentencing



(f) New charge in years 2-3 after sentencing



Note: Panel (a) shows the RD plot for being incarcerated at all. Panel (b) shows the same plot for months sentenced and panel (c) shows the same plot for being sentenced to at least one year. Panel (d) shows the distribution of sentence lengths for those just above and just below the cutoff. Panel (e) shows RD plots for recidivism-defined as a binary variable for having at least one new charge six months post-sentencing and panel (f) shows recidivism within 2-3 years post-sentencing.

H.6 Appendix tables: RD analyses

Table H.1: Summary statistics: RD sample

	Incarceration length worksheet	Probation/jail worksheet
	mean	mean
Offenses		
Assault	0.05	0.00
Burglary	0.11	0.00
Drug	0.41	0.57
Larceny	0.35	0.42
Miscellaneous	0.02	0.01
Robbery	0.02	0.00
Sexual assault	0.03	0.00
Defendant characteristics		
Black	0.50	0.45
Female	0.23	0.32
Under 23	0.26	0.24
% of ppl in zip earning <25K	0.46	0.45
Incarceration		
Recommended for prison	0.34	0.00
Prior incarceration	0.63	0.54
Prior circuit crt. felony convic.	0.33	0.27
Carceral sentence	0.61	0.47
Jail sentence	0.34	0.45
Prison sentence	0.28	0.04
Sentence >= 5 years	0.04	0.00
Months of sentence	10.50	2.15
Post-release		
New felony charge within 1 year	0.08	0.11
Observations	151,778	115,300

Note: This table shows the means of relevant variables for the incarceration-length sample from Worksheet A and the probation/jail sample from Worksheet B.

Table H.2: Incarceration and recidivism: RD estimates for the intensive margin

	Sentence		Recidivism			
	(1) **Incar > 1 yr**	(2) **Months**	(3) **1 year**	(4) **2-4 years**	(5) **5-7 years**	(6) **1-7 years**
Treatment	0.440	8.451	-0.052	-0.009	0.015	-0.023
	[0.422,0.459]	[7.873,9.029]	[-0.065,-0.038]	[-0.028,0.011]	[-0.005,0.034]	[-0.048,0.002]
N	81,439	81,439	81,439	81,439	81,439	81,439
Control mean	0.08	4.00	0.12	0.23	0.18	0.41

Note: This table first shows the RD estimates of how the cutoff affects sentences (probability of getting a carceral sentence greater than 1 year and sentence length (columns 1-2) and recidivism (columns 3-6). We measure recidivism as the likelihood of receiving a new charge for various time windows: the first post-sentencing year, in which incapacitation is most likely, years 2-4, in which some incapacitation may still be present, as well as years 5-7, during which incarceration rates across treatment and control are equal. It also shows cumulative time windows of 1-7 years to compare to our IV estimates. Below the estimates, we present in brackets confidence intervals obtained following Kolesár and Rothe (2018). Our estimations are for a bandwidth of 5 above and below the cutoff. See Appendix H for a discussion of parameter choices.

Table H.3: Incarceration and recidivism: RD estimates for the extensive margin

	Sentence		Recidivism		
	(1) **Any incar**	(2) **Months**	(3) **6 months**	(4) **1-3 years**	(5) **2-3 years**
Panel A: probation/jail sample					
Treatment:	0.428	0.754	-0.006	-0.005	-0.003
	[0.391,0.465]	[0.523,0.986]	[-0.014,0.001]	[-0.019,0.008]	[-0.015,0.009]
N	80,304	80,304	80,304	80,304	80,304
Control mean	0.21	0.99	0.06	0.21	0.13
Panel B: no prior incar. probation/jail sample					
Treatment:	0.422	0.887	0.015	0.022	-0.002
	[0.340,0.504]	[0.254,1.520]	[-0.013,0.043]	[-0.036,0.080]	[-0.045,0.041]
N	7,851	7,851	7,851	7,851	7,851
Control mean	0.18	0.80	0.05	0.20	0.13

Note: This table first shows the RD estimates of how the cutoff affects sentences (probability of getting a carceral sentence and sentence length (columns 1-2) and recidivism (columns 3-5). We measure recidivism as the likelihood of receiving a new charge for various time windows: the first is 6 months post-sentencing year, in which incapacitation is most likely. It also shows cumulative 1-3 year estimates to compare more closely to our IV results. The third is years 2-3, during which incarceration rates across treatment and control are equal. The first panel is our probation/jail score sample while our second panel is for those in our probation/jail sample without prior incarceration post-2010. Below the estimates, we present in brackets confidence intervals obtained following Kolesár and Rothe (2018). Our estimations are for a bandwidth of 5 above and below the cutoff. See Appendix H for a discussion of parameter choices.

Table H.4: Balance: RD estimates for incarceration-length sample

	(1) In Virginia 5-7yrs	(2) Any prior chrg.	(3) Prior incarc.	(4) Female	(5) Young	(6) Black
Panel A: demographic balance						
RD estimate:	-0.004	0.006	-0.004	-0.016	0.002	-0.016
	[-0.020,0.012]	[-0.011,0.023]	[-0.173,0.165]	[-0.033,0.001]	[-0.038,0.042]	[-0.043,0.011]
N	81,439	81,268	81,439	81,439	81,022	81,439
Control mean	0.80	0.35	0.77	0.22	0.22	0.53
	Plea	Dropped chrg.	Upgrade chrg.	Zip <10K	Zip <25K	
Panel B: income & legal actor balance						
RD estimate:	-0.023	0.007	-0.002	-0.000	-0.000	
	[-0.041,-0.005]	[-0.013,0.028]	[-0.009,0.006]	[-0.003,0.002]	[-0.006,0.006]	
N	81,439	76,399	76,399	58,899	58,899	
Control mean	0.85	0.16	0.02	0.19	0.47	

Note: Panel A shows RD estimates of discontinuities in various predetermined characteristics across the cutoff in the incarceration-length score. Panel B tests for discontinuities at the cutoff in outcomes of the criminal proceedings, such as whether the case resolved in a guilty plea, whether there were any dropped charges, whether there were any charges that were upgraded from misdemeanor to felony, and various measures of indigency within the defendant's zip Code. Below the estimates, we present in brackets confidence intervals obtained following Kolesár and Rothe (2018). Our estimations are for a bandwidth of 5 above and below the cutoff.

Table H.5: Balance: RD estimates for the probation/jail sample

	(1) In Virginia 5-7yrs	(2) Any prior chrg.	(3) Prior incarc.	(4) Female	(5) Young	(6) Black
Panel A: demographic balance						
RD estimate:	-0.006	0.033	0.040	0.017	0.007	-0.000
	[-0.033,0.021]	[-0.014,0.079]	[-0.104,0.184]	[-0.025,0.058]	[-0.033,0.046]	[-0.050,0.050]
N	80,304	80,099	80,304	80,304	79,968	80,302
Control mean	0.79	0.20	0.36	0.31	0.24	0.43
	Plea	Dropped chrg.	Upgrade chrg.	Zip <10K	Zip <25K	
Panel B: zip income & legal actor balance						
RD estimate:	-0.001	-0.012	-0.003	-0.001	-0.003	
	[-0.016,0.013]	[-0.025,0.002]	[-0.009,0.003]	[-0.003,0.001]	[-0.009,0.003]	
N	80,304	72,503	72,503	57,005	57,005	
Control mean	0.88	0.14	0.02	0.18	0.45	

Note: Panel A shows RD estimates of discontinuities in various predetermined characteristics at the cutoff in the probation/jail score. Panel B tests for discontinuities across the cutoff in outcomes of the criminal proceedings, such as whether the case resolved in a guilty plea, whether there were any dropped charges, whether there were any charges that were upgraded from misdemeanor to felony, and various measures of indigency within the defendant's zip Code. Below the estimates, we present in brackets confidence intervals obtained following Kolesár and Rothe (2018). Our estimations are for a bandwidth of 5 above and below the cutoff.

Table H.6: Marginal cases in the RD study

	Incarceration length worksheet		Probation/jail worksheet	
	P(X=x)	P(X=x| Marginal)	P(X=x)	P(X=x| Marginal)
Prior conviction	0.636	0.852	0.521	0.565
	(0.481)	(0.355)	(0.500)	(0.496)
Female	0.245	0.204	0.320	0.277
	(0.430)	(0.403)	(0.466)	(0.447)
Black	0.458	0.507	0.438	0.459
	(0.498)	(0.500)	(0.496)	(0.498)
Prior incarceration	0.651	0.871	0.535	0.651
	(0.477)	(0.335)	(0.499)	(0.477)
Drugs	0.412	0.393	0.576	0.815
	(0.492)	(0.488)	(0.494)	(0.388)
Property	0.496	0.491	0.413	0.173
	(0.500)	(0.500)	(0.492)	(0.378)
Violent	0.073	0.098	0.000	0.000
	(0.260)	(0.298)	(0.000)	(0.000)
Other	0.040	0.047	0.011	0.012
	(0.197)	(0.212)	(0.105)	(0.111)
Observations	230357	27556	152663	20609

Note: This table compares socio-demographic characteristics of compliers to that of the full sample for the RD sample.

H.7 Example of sentencing worksheet

Drug/Schedule I/II ❖ Section A

Offender Name: ____________________

◆ **Primary Offense**

A. Possess Schedule I or II drug

	1 count	1
	2 counts	3
	3 counts	8

B. Sell, Distribute, Possession with Intent Schedule I or II drug

	1 count	12
	2 counts	13
	3 counts	14
	4 counts	15

C. Sell, etc. Schedule I, II drug to minor (1 count) 11

D. Accommodation - Sell, Distribute, Possession with Intent Schedule I or II drug

	1 count	5
	2 counts	7

E. Sell, etc. imitation Schedule I or II drug (1 count) 4

Score

◆ **Primary Offense Additional Counts** Total the maximum penalties for counts of the primary not scored above → 0

Years:				
	5 - 10	1	31 - 42	4
	11 - 21	2	43 or more	5
	22 - 30	3		

◆ **Additional Offenses** Total the maximum penalties for additional offenses, including counts → 0

Years:				
	Less than 4	0	22 - 30	3
	4 - 10	1	31 - 42	4
	11 - 21	2	43 or more	5

◆ **Knife or Firearm in Possession at Time of Offense** — **If YES, add 2** → 0

◆ **Conviction in Current Event Requiring Mandatory Minimum Term (6 mos or more)** **If YES, add 9** → 0

◆ **Mandatory Firearm Conviction for Current Event** — **If YES, add 7** → 0

◆ **Prior Convictions/Adjudications** Total the maximum penalties for the 5 most recent and serious prior record events → 0

Years:		
	Less than 7	0
	7 - 26	1
	27 - 48	2
	49 or more	3

◆ **Prior Incarcerations/Commitments** — **If YES, add 2** → 0

◆ **Prior Felony Drug Convictions/Adjudications** → 0

Number:		
	1 - 2	1
	3 - 4	2
	5	3
	6 or more	4

◆ **Prior Juvenile Record** — **If YES, add 1** → 0

◆ **Legally Restrained at Time of Offense** → 0

None	0
Other than parole/post-release, supervised probation or CCCA	1
Parole/post-release, supervised probation or CCCA	4

SCORE THE FOLLOWING FACTOR **ONLY** IF PRIMARY OFFENSE IS **POSSESSION OF SCHEDULE I/II DRUG** (§ 18.2-250(A,a))

◆ **Two or More Prior Felony Convictions/Adjudications** — **If YES, add 2** → 0
For Possession, Possession with Intent, Distribution, Manufacture or Sale of Schedule I or II Drug

Total Score →

If total is 10 or less, go to **Section B**. If total is 11 or more, go to **Section C**.

Drug Schedule I or II/ Section A Eff. 7-1-09