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

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, Jiafeng (Kevin) Chen, Will Dobbie, Deniz Dutz, Brigham Frandsen, Anjelica Hendricks, Felipe Goncalves, Hans Grönqvist, 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.

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
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

A data appendix is available at http://www.nber.org/data-appendix/w32894

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, 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 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 treatments.

[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.

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 $T_c(z_c', z_i) - T_i(z_c', z_i) \geq T_c(z_c, z_i) - T_i(z_c, z_i)$,

ii $T_c(z_c', z_i) - T_d(z_c', z_i) \geq T_c(z_c, z_i) - T_d(z_c, z_i)$,

iii $T_i(z_c', z_i) - T_d(z_c', z_i) \geq T_i(z_c, z_i) - T_d(z_c, z_i)$, or $T_i(z_c', z_i) - T_d(z_c', z_i) \leq T_i(z_c, z_i) - T_d(z_c, z_i)$.

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 (weakly) move in one direction across any margin and that individuals only move into (and not out of) $T = c$. 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

[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."

$d \to c$ 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.[25]

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.[26] 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:

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

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

[25] 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.

[26] 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.

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.[27] 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 unobserved heterogeneity but allows for both unobservables to affect both conviction and incarceration.

[27] 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.

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 $R_d \equiv 0$.[28] 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 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;

[28] 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.

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 $\Delta_{i\to c}^{Y_c-Y_i}$ 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, $\Delta_{k\to l}^{Y_m-Y_n}$ denotes the treatment effect of $T = m$ vs $T = n$ for cases induced to move from $k \to l$.[29]

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 given by:*

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

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

Proof: See Appendix C.1.

[29]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 $\Delta_{j\to k}^{Y_m-Y_n}$ rather than $\Delta(z_c', z_c|z_i)_{j\to k}^{Y_m-Y_n}$.

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:*

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

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.[30] 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 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-

[30] 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.

dismissal margin.[31] 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 $\Delta^{Y_i - Y_c}_{d \to i} - \Delta^{Y_i - Y_c}_{i \to c}$, 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):

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

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

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.[32] We run these 2SLS regressions

[31]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.

[32]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.

on the sample described in Section 2.3.[33]

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.

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

[33]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.

present OLS and 2SLS regressions.[34]

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 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.[35]

[34] 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.

[35] Witness cooperation is only one potential omitted variable. There are many others that could also

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.[36]

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.

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 im-

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.

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

pacts 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.[37] 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.[38]

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

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

[38]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.

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.[39]

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 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

[39] 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.

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.[40] 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).[41] 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 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.

[40]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.

[41]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.

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 $\Delta^{Y_i-Y_c}_{d\to i} - \Delta^{Y_i-Y_c}_{i\to c}$, 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.[42] It's reasonable to think that, in the short run, incarceration affects recidivism primarily through incapacitation (for both groups). If so, shifting prison beds towards those at a higher risk of recidivism will reduce recidivism, and $\Delta^{Y_i-Y_c}_{d\to i} < \Delta^{Y_i-Y_c}_{i\to c}$. 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

[42] 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.

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.[43] 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 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.[44]

[43]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).

[44]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

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 under-estimate 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 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-

dismissal into incarceration.

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.[45]

[45] 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.

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,[46] and Berry and Haile (2022) show that judge-specific thresholds can be identified without invoking identification at infinity arguments.[47]

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 to a specific outcome for a specific case assigned to judge j.[48] 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.[49]

[46] 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.

[47] 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.

[48] 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).

[49] 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 $\pi_i(z_c, z_i) = log(z_i) - log(1 - z_c - z_i)$.

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).[50] 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 are allowed to vary by judicial circuit and year.[51] 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

[50] 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.

[51] 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.

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 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 9% 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 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 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.

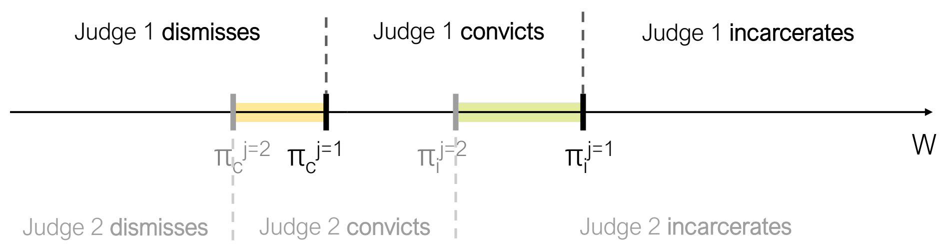
Zimring, Franklin E, Gordon Hawkins, and James Vorenberg, *Deterrence: The legal threat in crime control*, University of Chicago Press Chicago, 1973.

7 Figures and tables

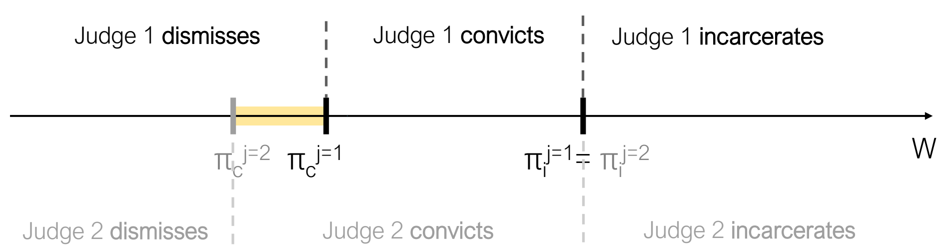
7.1 Figures

Figure 1: Ordered choice model

(a)

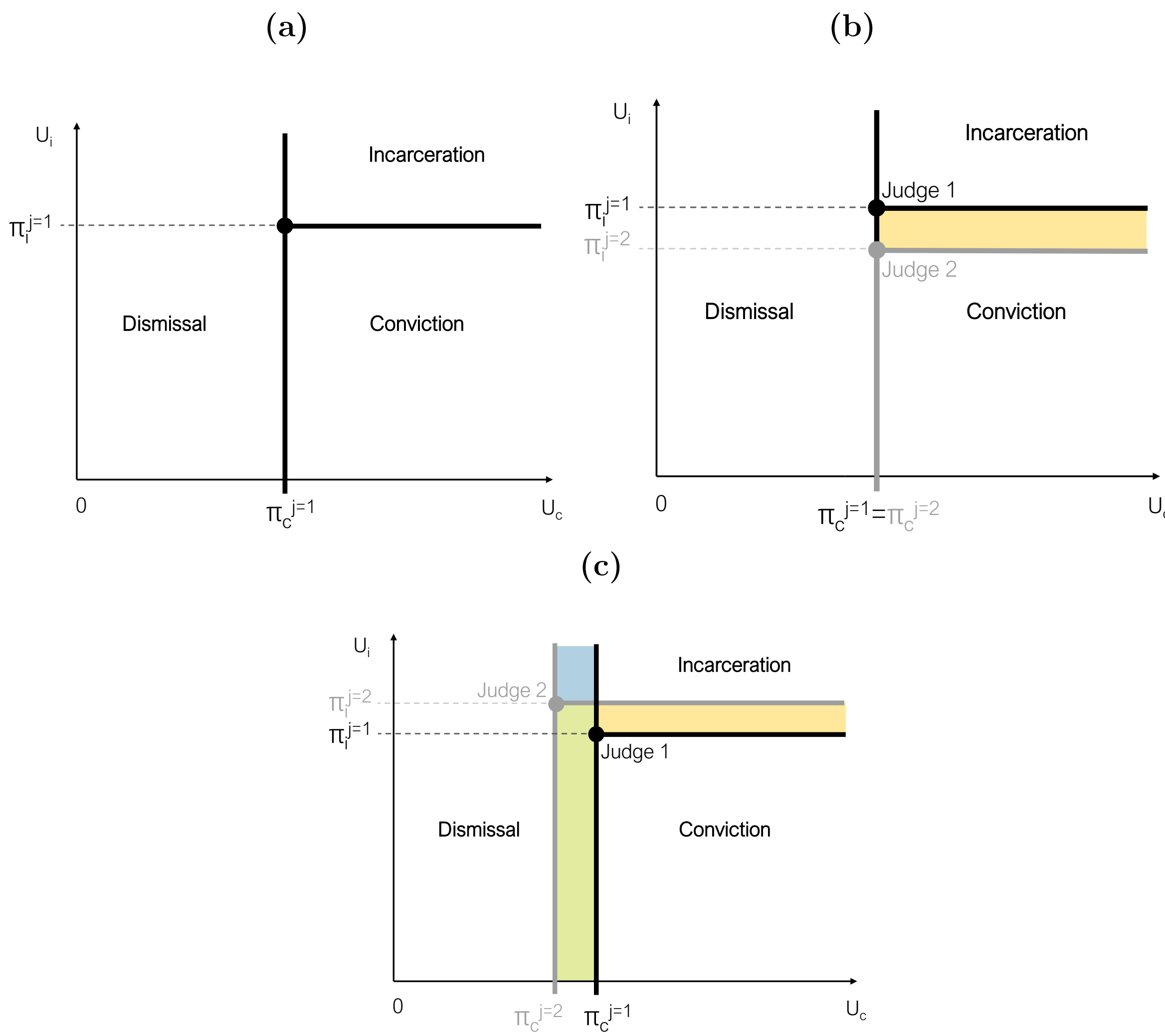


(b)



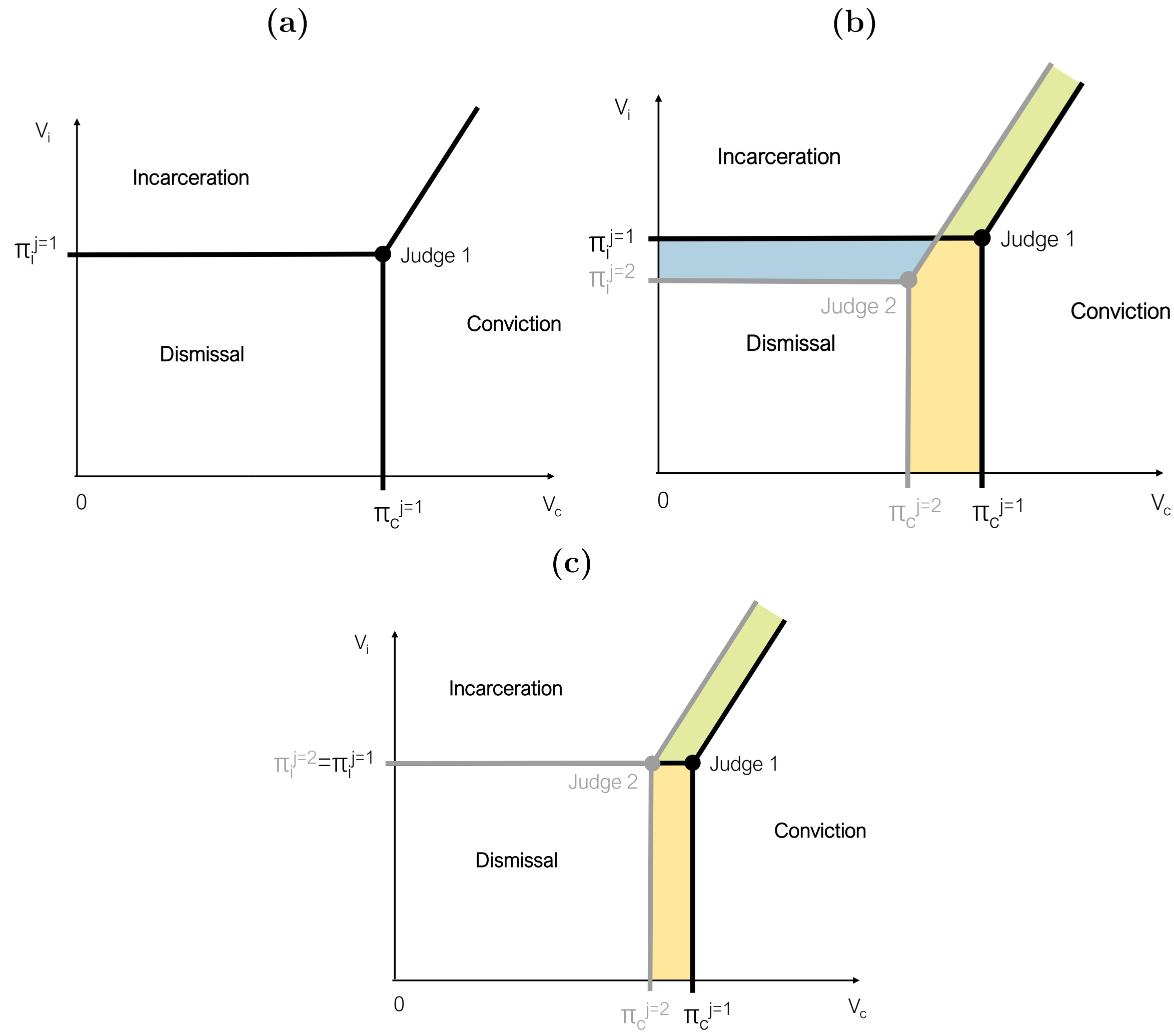
Note: This figure visualizes how, under the ordered choice model discussed in Section 3.3.1, judges classify individuals into incarceration, conviction, and dismissal depending on the cases' unobservable W. Panel (a) visualizes this for two arbitrary judges, and Panel (b) does so for two judges with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

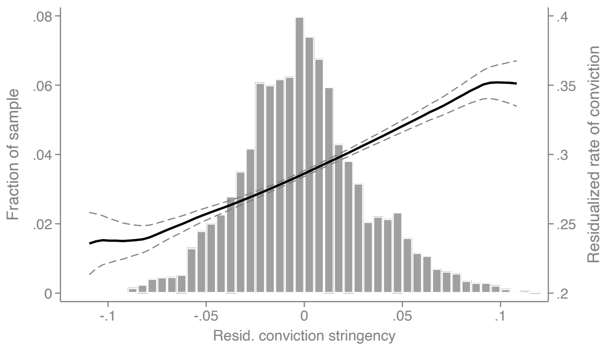
Figure 3: Unordered multinomial choice model



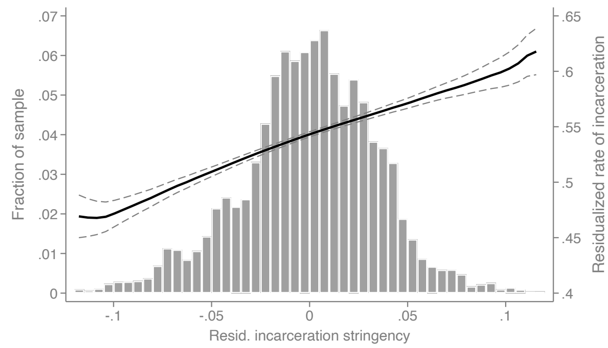
Note: This figure visualizes how, under the unordered multinomial choice model discussed in Section 3.3.2, judges classify individuals into incarceration, conviction, and dismissal depending on the cases' unobservable V_i and V_c. Panel (a) visualizes this for an arbitrary judge, Panel (b) does so for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

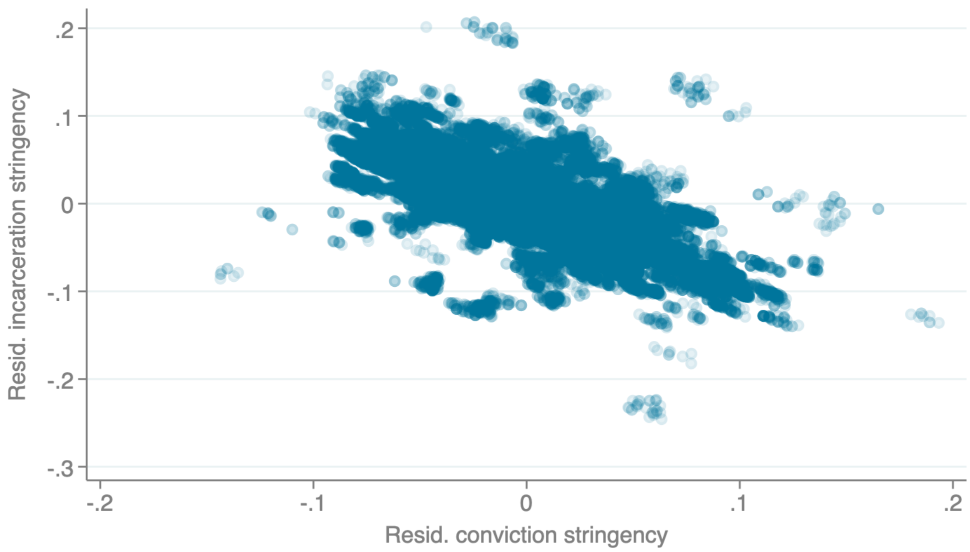
(a) Noncarceral conviction



(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.