NBER WORKING PAPER SERIES

HETEROGENEOUS IMPACTS OF SENTENCING DECISIONS

Andrew Jordan
Ezra Karger
Derek Neal

Working Paper 31939
http://www.nber.org/papers/w31939

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2023

We thank Ilan Wolff, Maximiliano Gonzalez, Julien Bendelac, Ryan Lee, and Ora Halpern for excellent research assistance. We owe thanks to many people who gave generously of their time to help us better understand how the Illinois criminal justice system operates and records its activities: Jordan Boulger, Michael Cooney, Jennifer Dohm, Kendy Elberson, Ian Jantz, Alan Mills, Michael Moore, Max Schanzenbach, Jason Sweat, Gwyn Troyer, and Rob Warden. We owe special thanks to Stephen M Brandt, the Director of Legal Research for the Chief Judge, Judge Lawrence P. Fox, Sarah Staudt at the Prison Policy Initiative, and Forest Gregg of the Chicago Data Collaborative. We thank Magne Mogstad, Evan Rose, Azeem Shaikh, seminar participants at Brown Economics, the Opportunity & Inclusive Growth Institute at the Federal Reserve Bank of Minneapolis, the Institute for Research on Poverty's Summer Research Workshop, and Northwestern Law School's Law and Economics Workshop for useful feedback. We thank The Hymen Milgrom Supporting Organization's Successful Pathways from School to Work project for support. We thank Robert Goerge and Chapin Hall for supporting our work. We thank the Walton Family Foundation for research support. Any views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of Chicago, the Federal Reserve System, or the National Bureau of Economic Research.

Heterogeneous Impacts of Sentencing Decisions
Andrew Jordan, Ezra Karger, and Derek Neal
NBER Working Paper No. 31939
December 2023
JEL No. K0

ABSTRACT

We examine 70,581 felony court cases filed in Chicago, IL, from 1990–2007. We exploit case randomization to assess the impact of judge assignment and sentencing decisions on the arrival of new charges. We find that, in marginal cases, incarceration creates large and lasting reductions in recidivism among first offenders. Yet, among marginal repeat offenders, incarceration creates only short-run incapacitation effects and no lasting reductions in the incidence of new felony charges. These treatment-impact differences inform ongoing legal debates concerning the merits of sentencing rules that recommend leniency for first offenders while encouraging or mandating incarceration sentences for many repeat offenders. We show that methods that fail to estimate separate outcome equations for first versus repeat offenders or fail to model judge-specific sentencing tendencies separately for cases involving first versus repeat offenders produce misleading results for first offenders.

Andrew Jordan
Department of Economics
Washington University in St Louis
One Brookings Dr.
St Louis, MO 63105
awjordan@wustl.edu

Ezra Karger
Federal Reserve Bank of Chicago
230 South LaSalle Street
Chicago, IL 60604
ezra.karger@chi.frb.org

Derek Neal
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
n9na@uchicago.edu

Contents

Introduction

Between 1970 and the Great Recession in 2008, per-capita prison populations in the United States grew by roughly 400%.[1] Since then US incarceration rates have fallen, but in 2019, rates remained roughly 300% higher than their 1970 levels. Raphael and Stoll (2013), Neal and Rick (2016), and Neal and Rick (forthcoming) show that changes in sentencing policies, adopted by states during the 1980s and 1990s, drove this sharp increase in prison populations. Advocates of these reforms argue that incarceration sentences not only incapacitate potential criminals but also deter future offending. Critics of these reforms argue that, in marginal cases, sentencing offenders to incarceration rather than probation may increase future offending rates by disrupting ties to the community and increasing exposure to career criminals.

We exploit the random assignment of judges to felony court cases in Chicago, IL to evaluate the impacts of incarceration on future offending. Relative to most related studies, we have significantly larger samples, and in contrast to previous work, we are able to estimate separate models for first offenders versus repeat offenders. We compare treatment impacts among first versus repeat offenders for several reasons. To begin, throughout the US, sentencing guidelines often instruct judges to consider prior convictions as an aggravating factor, but judges exercise considerable discretion when implementing this guidance. So, the relative severity of a given judge when sentencing first offenders may not accurately predict her severity when sentencing repeat offenders. Further, since repeat offenders are the select group of former first offenders who have already recidivated, they may respond differently to incarceration than first offenders.

We show that, among first offenders who are marginal candidates for incarceration, prison sentences generate long and lasting reductions in recidivism, but this is not the case among repeat offenders. We also show that empirical models that do not include both separate outcome equations for first-offenders and measures of judge-severity that are specific to how judges sentence first offenders produce misleading results.

In the balance of the paper, we proceed as follows: we begin by briefly reviewing the related literature. We then explain our data sources, our empirical model, and how we construct key variables. Next, we present our main results and results from a series of robustness exercises and specification tests. In the final sections, we explore potential mechanisms that may drive our results, and we discuss the potential implications of our results for ongoing debates about sentencing policy.

1 Literature Review

We seek to understand how incarceration impacts future criminal justice outcomes, and we use the random assignment of cases to judges as a source of exogenous variation in the incidence of incarceration.[2] Three early studies that examine US data employ similar research designs. Green and Winik (2010), Nagin and Snodgrass (2013), and Loeffler (2013) exploit random case assignment in Washington, DC, Pennsylvania, and Cook County, IL to examine the impact of prison sentences on recidivism. None find that incarceration treatment has lasting impacts on future rates of recidivism.[3]

Norris et al. (2021) employ data from OH to examine the impacts of parental incarceration on children, but they also examine the direct impact of incarceration on adult recidivism. They find that, beyond three years after sentencing, incarceration has no impact on simple recidivism rates but does generate a lasting reduction in the total number of future charges. Dobbie et al. (2019) explores similar questions using data on random judge assignments in Sweden. They find that incarceration has little impact on criminal behavior six years after sentencing, but incarceration does lower future earnings and has negative impacts

[1] See Carson (2019).

[2] A recent review article, Loeffler and Nagin (2022), provides more detailed descriptions of most of the papers we discuss here. See Roodman (2017) for an earlier survey.

[3] Mueller-Smith (2015) exploits random judge assignment in Harris County, Texas, but he employs a different empirical methodology. He estimates a panel data model that employs several measures of the person's incarceration history and concludes that incarceration reduces crime in the short term but raises long-term re-offending rates.

on family structure. In contrast, Bhuller et al. (2020) exploit random judge assignment in Norway and find that incarceration reduces five-year recidivism rates by roughly 30 percentage points. They do not estimate separate models for both first versus repeat offenders but do present appendix results that document particularly large negative impacts of incarceration treatment among first offenders.[4]

A different literature does not rely on random judge assignment for identification but instead exploits discontinuities in rules that govern the length of mandatory sentences or early release from prison. These studies do not examine the impacts of assigning marginal defendants to at least some time in prison but rather the average impacts of increasing the length of time that defendants spend in prison. Rose and Shem-Tov (2021) estimate the impacts of receiving various doses of additional prison time by leveraging discontinuities in North Carolina's sentencing guidelines. They conclude that a one-year prison term reduces five-year arrest rates by 6.7 percentage points.[5] Kuziemko (2013) exploits discontinuities in Georgia Parole Board guidelines for early release to study the impacts of differences in time-served on recidivism. She finds that an extra month in prison reduces recidivism rates three years after release by 1.3 percentage points.

In the studies reviewed above, the vast majority of adult compliers assigned to incarceration or to additional incarceration are repeat offenders, and in Rose and Shem-Tov (2021) and Kuziemko (2013), many of these repeat offenders are not marginal candidates for incarceration.[6] Our paper is the first to examine how incarceration impacts the incidence of at least one new charge at different post-sentencing horizons among first-offenders and then compare these impacts to comparable results for repeat offenders. We find that, among repeat offenders, incarceration treatment has no long-term impact on recidivism rates, but among first offenders, incarceration treatment generates large and lasting reductions in recidivism rates.

2 Data

We employ two key data sources. We employ records from the Clerk of the Circuit Court of Cook County, IL that describe felony criminal proceedings held between 1984 and 2018, and we use data from the Illinois Department of Corrections (IDOC) for the years 1990 to 2014 that describe admissions to prison, exits from prison, and expected terms of Mandatory Supervised Release (MSR), which is Illinois' parole system. We combine these data to measure case outcomes and seven-year recidivism outcomes for cases filed in Cook County between 1990 and 2007. We include only male defendants in our estimation samples.

Appendix section 12 provides a detailed account of how we create our analysis samples and code key variables. Here, we briefly comment on several crucial steps.

We depart from most previous research by analyzing first and repeat offenders separately. We define first offenders as persons who have never been arraigned on a felony charge in adult court. To make sure that we have the opportunity to observe each defendant's entire criminal record in Cook County, we restrict our sample to persons born after 1966.[7] Our IDOC prison records allow us to see admissions and exits from the state prison system between 1990 and 2014. If a person exits IDOC before he faces his first felony charge recorded in our Cook County data, we classify all of his Cook County cases as repeat-offender cases.

[4]A smaller literature explores random assignment of criminal cases in juvenile court. Aizer and Doyle (2015) and Eren et al. (2018) report several results that are consistent with the hypothesis that incarcerating marginal juvenile offenders enhances the likelihood that a young offender will be involved in crime as an adult. Agan et al. (2021) consider decisions by prosecutors to drop charges against non-violent misdemeanor defendants and find that leniency for first-time offenders substantially reduces future recidivism, but these offenders are clearly not marginal candidates for incarceration.

[5]They also report that one year of incarceration has no impact on the arrival rate of recidivism events that occur more than 36 months after sentencing, and they interpret this as suggestive evidence that incapacitation effects drive the impacts of incarceration on five-year recidivism rates. Most of the offenders impacted by these discontinuities are repeat offenders.

[6]For example, the compliers in Kuziemko (2013) are prisoners who, on average, were scheduled to serve roughly two years in prison. Two years in prison is a substantial prison term in most states. Given the distribution of time-served in IL, we feel confident that many prisoners in the Kuziemko (2013) complier set were always takers and not marginal candidates for incarceration when they received their sentence.

[7]Our court data begin in 1984, and in Illinois, defendants ages 17 and over usually face criminal charges in adult courts.

We focus on cases that the Court randomly assigns to judges. Appendix 12 explains how we isolate cases that are eligible for random assignment. The Presiding Judge of the Criminal Division uses a computer program known as the *randomizer* to assign these cases to judges and then announces these assignments at each defendant's arraignment hearing. Although no state law requires that the Court assign any case randomly, the prosecutors, judges, and defense attorneys we interviewed all believe that the Presiding Judge assigns almost every eligible case randomly and only departs from random assignment in high-profile cases that involve serious violent crime. To minimize the chances that our sample includes these departures, we remove cases that involve the most serious violent crimes.[8]

Figure 1 describes how felony cases proceed following random assignment. To begin, just over four percent of these cases end because the SA drops the case or the judge dismisses the case after arraignment. In just under seven percent of cases, judges drop the felony case but require the defendant to plead guilty to a misdemeanor charge. These cases never result in prison sentences or in sentences to adult probation supervision. Instead, judges often assign fines, community service, or some form of community support and supervision.

Roughly six percent of cases proceed but end in acquittal. The remaining cases, 83 percent of the total, end in a felony conviction. Given a felony conviction, IL law requires the judge to sentence the defendant either to at least one year of adult probation supervision or to at least one year of incarceration.

Here, we note two details concerning the execution of incarceration sentences. First, while almost all incarceration sentences are sentences to prisons run by the Illinois Department of Corrections (IDOC), Cook County judges did assign some convicted felons to serve incarceration sentences in a small Bootcamp Program operated by the Cook County Department of Corrections (CCDOC) during part of our sample period.[9] Second, in less than one out of every forty cases where a judge sentences a defendant to IDOC, the judge awards the defendant so much credit for time-served in jail waiting for a verdict that the defendant serves no prison time.

Below, we code the outcome of a case as incarceration if the defendant serves four months in CCDOC Bootcamp or if he serves time in an IDOC prison. For all other cases, we code the case outcome as non-incarceration.

In our empirical models, we associate recidivism with arraignment on a new felony charge. However, we do not count technical violations of parole or probation as recidivism events. In addition, we do not treat revocations of parole or probation as evidence that a new crime has been committed unless we have clear evidence in either court or prison records that the offender in question is facing a new criminal charge.[10]

3 Empirical Model

Most of our empirical work involves 2SLS regression models that estimate the impact of sentencing decisions on future charges for defendants. The treatment variable in these models is an indicator variable that equals one if the defendant receives a sentence that requires him to serve an incarceration spell in either the CCDOC Bootcamp or an IDOC prison. Our first stage is

$$\tau_{j(i,t)} = z_{j(i,t)}\delta + x_{it}\gamma + e_{it} \tag{1}$$

where,

[8] We also remove traffic cases. See Bogira (2005) for more on the case randomization process. We estimate that, during our sample period, just over 80 percent of the cases arraigned in the George Leighton Criminal Courthouse were randomized. We use a subset of these cases that involve judges that we know handled at least 500 randomized cases.

[9] This one-year program involved four months of local incarceration and participation in special programs, followed by eight months of regular contact with persons working under the Sheriff. Four months is not a trivial incarceration spell. Over 12 percent of defendants who serve time in an IDOC prison serve less than four months.

[10] Staff in the Adult Probation Department informed us that violations of probation that are not accompanied by new criminal charges are not evidence of new criminal activity. See Appendix 12 for more details.

- $j(i,t)$ is a mapping that returns the judge j that the Court assigns to defendant i at time t.
- $\tau_{j(i,t)}$ is the treatment that judge $j(i,t)$ assigns to defendant i at time t.
- $z_{j(i,t)}$ is the severity of judge $j(i,t)$.
- x_{it} is a vector of characteristics that describe defendant i and the charges against him at t.
- e_{it} captures unobserved factors that influence sentencing for i at t.

Here, i does not index cases within a time period t. Rather, i is an index over all defendants in our data. We use the notation $j(i,t)$ to remind readers that the same defendant i may appear in many different cases that are randomly assigned to different judges at different points in time, t. Thus, when we present results for first offenders, we report HAC standard errors that reflect clustering at the judge level, but we use two-way clustering at the defendant and judge level when producing standard errors for our repeat offender results.[11]

Our second stage equation is

$$y_{its} = \tau_{j(i,t)}\theta_s + x_{it}\beta_s + v_{its} \tag{2}$$

y_{its} is an indicator that equals one if defendant i sentenced at time t is charged with a new crime before $t+s$. v_{its} captures unobserved factors that influence criminal justice outcomes between t and $t+s$.[12] We also present results from the following reduced form equation:

$$y_{its} = z_{j(i,t)}\alpha_s + x_{it}\pi_s + u_{its} \tag{3}$$

In all models, we employ the leave-out mean (LOM) of the treatment measure, $\tau_{j(i,t)}$, for judge $j(i,t)$ assigned to i at t, as our measure of judge severity, $z_{j(i,t)}$.[13]

$$z_{j(i,t)} = \frac{\sum\limits_{t'} \sum\limits_{\substack{i' \neq i \\ j(i',t')=j(i,t)}} \tau^*_{j(i',t')}}{\sum\limits_{t'} \sum\limits_{\substack{i' \neq i \\ j(i',t')=j(i,t)}} 1}$$

Here, $\tau^*_{j(i',t')}$ is the deviation of $\tau_{j(i',t')}$ from its expected value given the date the case is assigned and other defendant and case characteristics. We discuss how we create $\tau^*_{j(i',t')}$ in section 4.

3.1 Assumptions

We maintain the standard assumptions that define valid instruments in our setting.

[11] HAC standard errors, clustered at the judge level, are appropriate if we think of the asymptotic distribution of our 2SLS estimator as the limit achieved by letting the number of judges grow, while holding the cases that each judge handles fixed. In this case, our LOM measures of judge severity, $z_{j(i,t)}$, always share a common estimation error component within judge. If instead, we consider holding the number of judges fixed and letting the number of cases handled by each judge grow, there is no reason to cluster, given random case assignment. We have also produced Huber-White standard errors for our results as well as HAC standard errors that are clustered on the day of case assignment. These alternative methods produce similar results. See panels B and C of Appendix Table 10.3.

[12] We engage in a slight abuse of notation. t marks both the date of assignment and the date that the judge announces a verdict and, given a verdict of guilty, a sentence.

[13] We leave out the sentence assigned to i at t, and we leave out sentences assigned at t to any co-defendants of i. Among repeat offenders who appear in multiple cases, we leave out all cases that involve i.

Assumption 1 - Independence: $(e_{it}, v_{its}) \perp\!\!\!\perp z_{j(i,t)}, \quad \forall i, j, t, s$

Assumption 2 - Rank: $\delta \neq 0$

To illustrate the implications of heterogeneous treatment effects in this framework, we consider a special case. Let $v_{its} = v^0_{its} + \Delta_{its}\tau_{j(i,t)}$. Here, the error term in our recidivism equation takes on the value v^0_{its} if the defendant receives a probation sentence and $v^0_{its} + \Delta_{its}$ if the defendant receives an incarceration sentence. In this setting, independence requires $(e_{it}, v^0_{its}, \Delta_{its}) \perp\!\!\!\perp z_{j(i,t)}$.

Even if judges sentence offenders based on unmeasured defendant traits that are correlated with Δ_{its}, i.e. $E(\Delta_{its}|\tau_{j(i,t)}, x_{it}, z_{j(i,t)}) \neq 0$, our 2SLS estimator is a consistent estimator of the Local Average Treatment Effect (LATE) of incarceration. Here, this local average is a weighted average of the expected impacts of treatment among compliers, i.e. defendants who would not receive prison from the most lenient judge but would receive prison from at least one more severe judge.

The LATE interpretation of our 2SLS results requires that we impose an additional assumption. The relationship between true judge severity and sentencing outcomes must be monotonic.

Assumption 3 - Monotonicity: If judge j is more severe than j', then $\tau_{j(i,t)} \geq \tau_{j'(i,t)} \ \forall (i,t)$.

In section 6.3, we show that several common diagnostic tests produce no evidence that monotonicity fails within our samples of first and repeat offenders. We also present results from a test proposed by Frandsen et al. (2023) that provide support for the weaker assumption of average monotonicity.

Taken together, Assumptions 1-3 and the specification of equation 2 impose an important exclusion restriction. Judges impact recidivism through incarceration alone. No unmeasured decisions that judges make are correlated with $z_{j(i,t)}$ and impact y_{its} through channels other than incarceration. We return to this topic in section 6.4.

4 Descriptive Statistics and Balance

We have already discussed our measures of recidivism and incarceration treatment. Here, we discuss the remaining variables in our empirical models, present descriptive statistics, and present evidence that we have identified cases that the court assigned randomly to judges.

The vector x_{it} contains characteristics of the defendant and the case filed against him. It contains a full set of indicators for the year of case assignment, the offense class of the most serious charge against defendant i at time t, and a full set of interactions between year and class. It also contains indicators for crime category and interaction terms that capture whether the class designation is relatively less serious within the crime category. It also includes indicators for cases that involve multiple charges and cases that involve multiple defendants as well as an indicator for race and an indicator for residing in a high-crime neighborhood. In our repeat offender models, we also include an indicator for the presence of multiple prior felony charges.[14]

In our baseline specification, we create $z_{j(i,t)}$, the leave-out mean (LOM) of $\tau_{j(i,t)}$ by first running regressions of $\tau_{j(i,t)}$ on our full set of defendant and charge characteristics, x_{it}. We run separate regressions on our first and repeat offender samples and then capture the residuals from these regressions and average these residuals at the judge level, leaving out defendant i's case.

Figure 2 shows that many judges exhibit different relative severities when sentencing first versus repeat offenders. Panel A presents mean residuals by judge among cases that involve first offenders. Panel B presents mean residuals by judge among cases that involve repeat offenders. In both panels, we order judges from least severe to most severe, but we assign judge-id numbers based on each judge's relative

[14] Appendix section 12 provides more details concerning the construction of these conditioning variables.

severity when sentencing first offenders, i.e. judge number 1 is the least severe when sentencing first offenders. In both panels, we see significant variation among judges in their willingness to sentence offenders to incarceration. However, the judge-id numbers on the x-axis of Panel B show that some judges who are quite severe when sentencing first offenders are relatively lenient when sentencing repeat offenders and vice versa. To make this pattern more transparent, Panel C plots each judge's average judge severity when dealing with repeat offenders against the judge's severity rank when dealing with first offenders. Note that four of the nine most lenient judges for first offenders record positive mean residuals in cases involving repeat offenders.[15]

Taken as a whole, these figures support our decision to calculate separate LOM measures within first offenders and repeat offenders. We report in section 5.4 that, when we do not create separate LOM-severity measures for each of these groups, our empirical models produce quite different results.

Table 1 provides descriptive statistics for our two main analysis samples: first offenders and repeat offenders. Just under 48 percent of our total cases involve repeat offenders, and 41 percent of these cases began when the defendant was under MSR supervision. On average, repeat offenders are almost five years older than first offenders and have faced 2.64 prior felony charges.

Repeat offenders are more likely to be Black and more likely to live in high-crime areas. Repeat offenders are less likely to face charges in the lowest offense class, Class 4, and they are more likely to face drug charges. Repeat offenders are more than three times as likely to receive incarceration sentences, and conditional on receiving an incarceration sentence, repeat offenders are less likely to go to CCDOC Bootcamp and more likely to go to a state prison.

Our research design rests on the assertion that we have identified cases that the Court randomly assigned. Table 2 presents regression results that speak to the validity of this assertion. In each regression, we project our LOM measure of sentencing severity on a set of year dummies and one of the defendant or case characteristics. If a case involves a first offender, we assign a LOM measure calculated within the sample of first offenders, and if a case involves a repeat offender, we assign an LOM measure calculated within the sample of repeat offenders.

Table 2 presents balance tests for the combined sample, the first-offender sample, and the repeat-offender sample. The table contains 65 parameter estimates and associated p-values, and only one p-value is less than 0.1. Among first offenders, we reject the joint null that all case and defendant characteristics fail to predict judge severity. For the full sample and the repeat-offender sample, we fail to reject this joint null.

These results provide considerable support for our claim that we have constructed a sample of cases that the Court assigned to judges using the randomizer program. However, because our main LOM severity measures are created by summing residuals taken from projections of $\tau_{j(i,t)}$ on x_{it}, and x_{it} contains many of the case and defendant characteristics in Table 2, some readers may doubt the power of these tests. We have therefore conducted additional balance tests. We repeated these balance tests using LOM measures that are averages of residuals taken from regressions of sentencing outcomes on only a vector of dummies for year of case assignment. Appendix Table 10.2 presents the results, and they are quite similar to those in Table 2.

5 Impacts of Incarceration on Recidivism

Figure 3 presents our main results. Panel A presents results for first offenders, and Panel B presents results for repeat offenders. Both panels plot our 2SLS estimates of the impacts of incarceration on recidivism rates at horizons of $6, 12, 18, .., 84$ months. The bands around these estimates are 95% confidence intervals.[16]

Among first offenders, our results indicate that incarceration produces large and lasting reductions in recidivism. Among first offenders, almost 40 percent face a new felony charge within three years of

[15] The overall correlation between the two sets of judge effects is .38.

[16] See Appendix Table 10.3 for more information about the distributions of recidivism events, the distributions of time-served, and the sampling distributions of the estimators.

sentencing and just over 50 percent recidivate within seven years. However, our 2SLS results for this sample imply that incarceration reduces recidivism rates at horizons of four, five, and six years by roughly 30 percentage points and by 23 percentage points at seven years.

Among repeat offenders, our results are quite different. We see evidence that incarceration generates noteworthy reductions in recidivism rates during the first three years following sentencing, but beyond three years, we find no statistically significant impacts of incarceration on recidivism rates. Further, at horizons of 5-7 years, our point estimates indicate that incarceration has essentially no impact on recidivism among repeat offenders.

Our results for first and repeat offenders are both qualitatively and quantitatively different. Among first offenders, incarceration generates large absolute reductions in recidivism over horizons of 5-7 years. These 5-7 year impact estimates are highly significant statistically, and they are also statistically different from our estimates of the impacts of incarceration on recidivism among repeat offenders over the same horizons. Among repeat offenders, incarceration has no lasting impact on recidivism rates.

5.1 Robustness

The two panels of Figure 3 present our key results. We produced these results given specific choices concerning how we measure judge-stringency, the controls we include in our models, and the specific 2SLS estimator we employ. In this section, we examine the robustness of our key conclusions by discussing results from models that make different but reasonable choices on these dimensions.[17]

Appendix 10.4 presents results from these alternative models. We produce a second set of results by re-estimating our baseline model using only a set of indicators for year of case assignment as our conditioning set. We produce our third and fourth sets of results by re-estimating the first two models using LOM measures of severity that are created by summing residuals taken from first-stage equations that contain only controls for year of case assignment. Next, we drop the LOM severity measures and re-estimate the first two models using a vector of judge assignment indicators as our instrument set. Finally, we estimate both of these 2SLS models using the UJIVE estimator developed in Kolesar (2013). Thus, for both first and repeat offenders, we produce seven additional estimates of treatment effects plotted in the two panels of Figure 3.[18]

Our main results for first offenders imply that incarceration creates large and lasting reductions in recidivism. Even at the seven-year horizon, incarceration reduces recidivism rates by 23 percentage points among first offenders. Results from our seven alternative models imply that incarceration reduces seven-year recidivism rates among first offenders by 19 to 28 percentage points. The median of these implied reductions is just over 22 percentage points. The mean is just under 24 percentage points.

Our main results for repeat offenders imply that the long-term impacts of incarceration on recidivism are essentially zero. We find the same results in our seven alternative models. At horizons of five, six, or seven years, our 21 alternative estimates of the impacts of incarceration on recidivism among repeat offenders range from $-.001$ to .022, and none are close to being statistically significant.

Each set of results in Appendix 10.4 reproduce the patterns we highlight in our main results. Incarceration produces large and lasting reductions in recidivism among first offenders but generates no lasting impact on recidivism among repeat offenders.

5.2 Alternative Outcomes and Treatments

So far, we have discussed estimates of the impact of initial incarceration on the arrival of at least one new charge during observation windows that range from one to seven years. Appendix Table 10.5 presents

[17]Given the large number of robustness tests we conduct, these tables present results at annual horizons rather than six-month intervals.

[18]As another point of comparison, we also present results from standard 2SLS results that employ the UJIVE predicted treatment variables as instruments. The resulting point estimates are almost identical to the UJIVE results.

results that parallel those presented in Figure 3, but here the outcome is the total count of new charges. We see the same qualitative pattern that we see in Figure 3. Among first offenders, incarceration treatment reduces the total count of new charges at all horizons. Our LATE estimates for first offenders are highly statistically significant at each horizon, and at horizons beyond four years, the LATE of incarceration on the total count of new charges ranges from $-.61$ and $-.74$. Among repeat offenders, the parallel impact estimates range from $-.11$ to $-.17$, and none are close to statistically significant.

We have also re-estimated our main models using a different definition of incarceration treatment. Here, we define treatment, $\tau_{j(i,t)}$, not as any indicator for any incarceration but as the expected months of incarceration implied by the sentence judge j assigns to defendant i. For both first and repeat offenders, we first estimate these models using our original measures of each judge's tendency to assign any incarceration as our instrument. We then re-estimate these models letting $z_{j(i,t)}$ equal the LOM of the expected months of incarceration required by the sentences imposed by judge j.

Appendix Table 10.6 presents the results. Once again, we see the same stark contrast between first and repeat offenders. The mean of expected time-served among compliers sentenced to prison is roughly 14.5 months among both first and repeat offenders. When we use the LOM of any incarceration treatment as our instrument, our results imply that, relative to a non-incarceration sentence, a 14.5 month spell of incarceration reduces recidivism rates at horizons of five, six, and seven years by .30, .30, and .25 respectively among first offenders, but we find no impacts of expected time served on recidivism among repeat offenders. When we use the LOM of expected months incarcerated as our instrument, our results for first offenders are quite similar, but here we do find statistically significant long-term impacts of expected time-served on recidivism among repeat offenders.[19] Nonetheless, at longer horizons, our results for first and repeat offenders remain quite different. For example, at horizons of six and seven years, the implied reductions in recidivism generated an additional month of incarceration among repeat offenders are only one-fourth of the corresponding reductions we see among first offenders.

5.3 Heterogeneous Impacts Within First and Repeat Offenders

Appendix 10.7 presents results for five subsamples drawn from each of our first and repeat offenders samples. We present results for defendants charged with drug crimes, defendants who are not charged with drug crimes, defendants who reside in high-crime areas, defendants who live outside the high-crime areas of Chicago, and Black defendants. We do not have a large enough samples of non-Black defendants to estimate separate models for other race groups. We divide our samples according to the presence of felony drug charges and residence in high-crime areas because these case features may serve as signals that a defendant is more exposed to networks of criminal activity.

In each subsample, we see the same qualitative patterns that we observe in Figure 3. While our point estimates suggest that incarceration may generate particularly large, long-term reductions in recidivism among first offenders who are not charged with drug crimes or do not live in high-crime areas, we cannot reach firm conclusions given the precision of our estimated treatment impacts.

5.4 The Importance of Analyzing First Offenders Separately

From the outset, we have stressed the importance of analyzing first and repeat offenders separately and that our results for first and repeat offenders are quite different, and our results change dramatically if we do not estimate separate outcome equations for first and repeat offenders, or if we do not create measures of judge severity that are specific to first versus repeat-offender cases. Appendix Figure 10.1 presents three

[19] Repeat offenders are much more likely to face mandatory prison sentences. When we use judge propensities to assign incarceration treatment as our instrument, these defendants are not compliers but always takers. However, when we use the LOM of expected time-served as our instrument, the outcomes of any defendants who not only face a mandatory prison sentence but who also receive particularly long sentences as a result of their judge assignments contribute to our estimates of the impacts of serving an additional month in prison. In section 6.4, we show that judges' propensities to assign long sentences given the assignment of some incarceration are not correlated with judges' tendencies to assign any incarceration treatment.

sets of results. In Panel A, we present results from a model that employs a single LOM measure of judge severity that reflects each judge's overall severity, and we use the combined sample of first and repeat offenders to estimate a single outcome equation for each horizon. In Panels B and C, we present results from 2SLS models that are specific to first and repeat offenders respectively. However, in both cases, we employ the full-sample LOM measure of each judge's overall severity as our instrument for incarceration treatment.

In all three figures, we see no evidence that incarceration treatment generates lasting reductions in recidivism. Further, Panels B and C show that results for first and repeat offenders are quite similar when both models employ an overall LOM measure of severity, derived from the full sample of cases, as the instrument for incarceration. Any approach that does not model both judge severity and recidivism outcomes separately for first versus repeat offenders produces misleading results for first-offenders. This outcome, in part, reflects the fact that how severe judges are when sentencing repeat offenders is not a strong predictor of how likely they are to assign incarceration treatment to first-offenders.[20]

6 Maintained Assumptions

In section 3, we discuss the maintained assumptions that justify our 2SLS model. Here, we discuss evidence that speaks to the plausibility of these restrictions.

6.1 Independence

We assert that we have identified cases that the Court randomly assigned to judges, and we have presented balance tables, Table 2 and Appendix Table 10.2, that provide considerable support for this claim. In addition, Panels A and B of Appendix Table 10.4 show that the results from our various 2SLS estimators do not vary greatly when we change the conditioning set, x_{it}. We get similar results whether we condition on our full set of case and defendant characteristics or only a vector of indicators for year of case assignment. This pattern is expected if the cases in our analysis samples were randomly assigned to judges.

6.2 Rank

A significant literature addresses the concern that the partial correlation between $z_{j(i,t)}$ and $\tau_{j(i,t)}$ may be non-zero but also small enough that our 2SLS estimates $\hat{\theta}_s$ are asymptotically biased. The F-statistics for the null $\delta = 0$ are 284 for the first offender model and 782 for the repeat offender model. These values are well beyond the range of values that raise researcher concerns about weak instruments.

6.3 Monotonicity

Figure 2 provides evidence that some judges who are severe with first offenders are not severe with repeat offenders. Here, we explore whether monotonicity holds within our sample of first offenders and our sample of repeat offenders.

Within each sample, we create $\hat{\tau}_{j^*(i,t)}$, the likelihood that each offender i faces incarceration given his characteristics, x_{it}, and assignment to a reference judge, j^*. We then rank defendants in each sample by $\hat{\tau}_{j^*(i,t)}$ and divide both samples into quartiles. We run our first-stage regression within each of these eight

[20] Among first offenders, a .1 increase in the likelihood that a given judge assigns incarceration treatment to a given first offender, holding case characteristics constant, implies an increase of more than .08 in the likelihood that the defendant receives an incarceration sentence. However, a .1 increase in the likelihood that the same judge assigns incarceration treatment to repeat offenders implies less than a .02 increase in the likelihood that the judge will sentence a given first-offender to incarceration treatment.

quartile samples, and we always find that the conditional correlation between $\tau_{j(i,t)}$ and our first or repeat offender LOM severity measure is positive and highly significant. See Panel A of Appendix Table 10.8.[21]

These results provide evidence in favor of our monotonicity assumption, but none represent a formal test of monotonicity. Frandsen et al. (2023) suggest replacing monotonicity with a weaker assumption that they call average monotonicity. If for each first offender i, there is a positive correlation between our 44 measures of average judge severity when sentencing first offenders and the 44 counterfactual sentencing outcomes that would result from assigning i to each judge, then average monotonicity holds. A parallel condition applies to repeat offenders.

Frandsen et al. (2023) argue that one way to judge the plausibility of average monotonicity is to (a) assume that strict monotonicity holds within a group that shares an observed characteristic, (b) calculate a set of judge-specific severity measures using only within-group variation, and (c) compute the correlation between the set of full-sample judge severity measures and the group-specific judge severity measures.

For both first and repeat offenders, we create LOM severity measures that are specific to subsamples of five different types of defendants.[22] The last column of Panel B in Appendix 10.8 shows that correlations between these subsample-specific severity measures and the corresponding severity measure derived from all first or repeat offender cases range from .73 to .98, and the median is .85. These results suggest that, given our data, the weaker average monotonicity condition proposed by Frandsen et al. (2023) is a reasonable assumption.[23]

6.4 Exclusion

We assume that judge $j(i,t)$ impacts y_{its} only through the choice of whether to sentence defendant i to incarceration. However, judges make other decisions that may impact recidivism outcomes. For example, judges influence the amount of time that passes between arraignment and the resolution of a case. Thus, judges influence the length of court cases and therefore the likelihood that defendants serve significant time in jail before they ever receive a verdict. If judges who are relatively slow to resolve cases are also more or less prone to assign incarceration than the average judge, our judge-severity instruments are correlated with unobserved factors that may influence recidivism rates.

Both Kolesar et al. (2015) and Frandsen et al. (2023) discuss similar scenarios. Frandsen et al. (2023) recommend that researchers examine whether judges' tendencies that can be measured are correlated. We therefore construct judge-level measures of average case length and sentence length conditional on incarceration and correlate them with our judge-level measures of incarceration propensity.[24] We do this separately within our samples of first and repeat offenders.

Figure 4 shows the resulting four scatterplots. Each scatterplot includes a regression line and notes the p-value for the null hypothesis that the slope of the line is 0. The p-values range from .25 to .93. These results provide no evidence that our measures of judge-specific propensities to assign incarceration sentences are correlated with these other levers that allow judges to influence the amount of time that defendants interact with the criminal justice system.[25]

[21] Panel B of Appendix Table 10.8 repeats this exercise with subsamples of Black defendants and subsamples defined by the presence of drug charges in the case or whether the defendant lives in a high-crime neighborhood. In all 10 first-stage equations, we again find positive and highly significant conditional correlations between $\tau_{j(i,t)}$ and the corresponding LOM severity measures for either all first or all repeat offenders.

[22] Black offenders, those charged with drug crimes, those charged with non-drug crimes, those who reside in a high-crime neighborhood, and those who do not.

[23] Under this assumption, our estimates of the impacts of incarceration on recidivism are still local average treatment effects (LATE), but the weights placed on the treatment impacts for different defendants in our complier sets may differ from the weights under strict monotonicity. See page 16 in Frandsen et al. (2023).

[24] Frandsen et al. (2023) recommend these explorations as tools for exploring whether a condition they define as average exclusion is likely to hold in a given context. See page 274.

[25] We have also run our 2SLS models measuring recidivism from the date of the defendant's arraignment to the conclusion of his case. We find that predicted incarceration treatment, given judge assignment, is uncorrelated with this measure of pre-sentencing recidivism among both first and repeat offenders. The estimated treatment impacts are .05 and .03 for first and repeat offenders respectively and the corresponding standard errors are at least twice as large.

Other exclusion concerns arise from the fact that not all non-incarceration case outcomes are the same. If judges' tendencies to assign incarceration sentences are correlated with their tendencies to generate particular types of non-incarceration outcomes among those not sentenced to incarceration, these correlations may create violations of our exclusion conditions.

In Illinois, when judges sentence defendants convicted of felonies, they must sentence them either to prison or to adult probation supervision. Nonetheless, about ten percent of defendants in our data are not convicted of any charge and receive no punishment. Further, in about seven percent of our cases, judges allow all charges to be reduced from felonies to misdemeanors, and then judges may fine defendants, require them to perform community service, or require that they submit to certain forms of community supervision. Finally, in a small percentage of cases, judges nominally sentence convicted felons to prison, but they grant so much credit for time-already-served in jail that defendants are released the same week and never enter prison.

Appendix Table 10.9 presents results from 2SLS models of the impacts of case outcomes on recidivism at different horizons, but in contrast to our baseline models, these models do not use a single indicator for incarceration to summarize the outcome of each case. Here, we treat probation as a baseline (omitted) outcome and include indicators for incarceration, acquittal, and other outcomes. We instrument for these three endogenous treatment indicators using the judge-specific LOMs for each of these three outcomes, and we estimate these models separately for first and repeat offenders. For a moment, let us assume constant treatment effects for each treatment. Given this assumption, these results provide consistent estimates of the recidivism impacts of incarceration treatment and other treatments relative to probation sentences.

Our 2SLS estimates of the impacts of acquittal and other case outcomes are noisy. None of the results are statistically significant at any horizon among first or repeat offenders. In sum, we find no clear evidence that these outcomes produce different expected recidivism rates than probation sentences.

However, the results for incarceration treatment in Appendix Table 10.9 are not only statistically significant but also quite similar to the corresponding estimated impacts of incarceration we plot in Figure 3. In fact, the Appendix Table 10.9 estimates of the impacts of incarceration relative to probation among first offenders are almost identical to the Figure 3 Panel A estimates of the impacts of incarceration relative to any non-incarceration outcome among first offenders. Among repeat offenders, the differences between these two sets of results are a little greater, but they are still quite small. These results are consistent with our view that, in IL, judges are almost always choosing between incarceration and probation when they are sentencing defendants who are marginal candidates for incarceration.

When we relax the assumption of constant treatment effects, interpreting these results requires more work, but Bhuller and Sigstad (2023) derive conditions that guarantee that each 2SLS estimate of a given treatment impact, e.g. incarceration, is a positive-weighted average of the effects of judge assignment shifting offenders from probation to the treatment in question and also guarantee that none of these 2SLS estimates are contaminated by the effects of shifting offenders from probation to other treatment outcomes, and they propose a diagnostic test that researchers may use to uncover violations of these conditions.[26] Appendix Tables 10.10.1 through 10.10.5 contain a full description of the Bhuller and Sigstad (2023) test and results from applying it to 10 subsamples of our data. The results provide additional support for the hypothesis that estimates of the impacts of incarceration on recidivism are properly weighted averages of the impacts of shifting offenders between probation and incarceration and are not impacted by judges' tendencies to produce other types of case outcomes.

[26] Bhuller and Sigstad (2023) call these conditions: average conditional monotonicity and no cross effects. We implement 10 sets of diagnostic tests. For both first and repeat offenders, we use the same five subsamples that we used to conduct monotonicity tests.

7 Mechanisms

In a stationary environment, incarceration impacts recidivism through three channels.[27] Incarceration reduces exposure to recidivism risk by incapacitating offenders. Incarceration also shifts recidivism risk to older ages where most persons are less prone to engage in criminal activities. Finally, the experience of serving time in prison may directly impact recidivism rates following prison release.

Given this conceptual framework, consider a population of similar offenders who have just been convicted of the same crime and are now marginal candidates for incarceration. For this population, define $F(n|m,a)$. This function expresses the probability that a convicted offender sentenced today will face at least one new charge over the next n periods as a function of m, the number of periods of incarceration the defendant must serve, and a, the defendant's age at sentencing.

Next, consider a subpopulation of defendants who share a common age, a_0, and consider an experiment that randomly assigns half of them to probation and sentences the other half to serve exactly $m > 0$ periods of incarceration. Among those assigned to probation, the n-period recidivism rate is $F(n|0,a_0)$. Among those sentenced to m periods of incarceration, the n-period recidivism rate is $F(n|m,a_0)$. At any horizon n, the expected impact of m periods of incarceration on recidivism is the difference between these two probabilities, i.e. $\Delta(n,m) = F(n|m,a_0) - F(n|0,a_0)$. Since prison incapacitates offenders, we know that $F(n|m) = 0$ when $m > 0$ and $n \leq m$. This implies that, if the defendants randomly assigned to incarceration must serve m periods in prison, $\Delta(n,m) = -F(n|0,a_0)$ for all horizons $n \leq m$. In this section, we explain how specific assumptions about the channels through which incarceration impacts recidivism allow us to also form expressions for $F(n|m,a_0)$ and $\Delta(n,m)$ for horizons $n > m$.

We start by considering the possibility that neither aging while incarcerated nor the experience of serving time in prison influence recidivism rates after incarcerated offenders exit prison. Here, prison incapacitates, but after prisoners leave prison, they offend at the same rates they would have offended at age a_0 if they had been sentenced to probation. As a result, $F(n|m,a_0) = F(n-m|0,a_0)\ \forall n > m$. For example, with $(n = 5, m = 2)$, the expected 5-period recidivism rate for an incarcerated complier sentenced to serve 2 periods in prison equals the expected 3-period recidivism rate among non-incarcerated compliers. In this environment, prison does not change behavior. It only reduces exposure to recidivism risk. As a result, $\Delta(n,m) = -F(n|0,a_0)\ \forall n \leq m$, and $\Delta(n,m) = F(n-m|0,a_0) - F(n|0,a_0)\ \forall n > m$.

These equations show that, in a world where prison incapacitates but has no other impacts on the propensities of offenders to recidivate, we can estimate the expected n-period recidivism rate of any incarcerated complier if we know his release date from prison, m, and the expected $(n-m)$-period recidivism rates of non-incarcerated compliers, $F(n-m|0,a_0)$. This observation is key because, for both first and repeat offenders, Appendix 10.11.1 provides estimates of recidivism rates among non-incarcerated compliers at different horizons, and Appendix 10.11.2 provides estimates of the distributions of time-served among incarcerated compliers. The results in Appendix 10.11.1 allow us to form estimates of $\Delta(n,m)$ at any horizon n given any time served m, and the results in Appendix 10.11.2 tell us what fractions of incarcerated compliers serve sentences of different lengths, m.[28] Thus, we can form weighted averages of $\Delta(n,m)$ that serve as estimates of the impact of incarceration on recidivism at each horizon n, given a counterfactual setting where incarceration incapacitates but does not impact the behavior of offenders when they are not incarcerated. See Appendix 11 for more details.

These results create an interesting benchmark, but prisoners do age while incarcerated, and a large literature concludes that age directly impacts recidivism rates.[29] Therefore, we repeat these calculations and include adjustments for the impacts of aging during incarceration. These adjustments reduce our estimates of recidivism rates following release from prison by accounting for the fact that, when a prisoner

[27] Appendix 11 explains this decomposition argument. By imposing stationarity, we ignore the possibility that serving time in prison could shift recidivism risk to a future time when community conditions have changed in ways that make recidivism more or less likely. However, all of our models do contain controls for year of case assignment.

[28] Compliers are marginal candidates for incarceration. Appendix 10.11.1 shows that, among non-incarcerated compliers, first-offenders actually have higher recidivism rates than repeat offenders. This pattern may reflect the fact that sentencing guidelines instruct judges to set a higher threshold for assigning incarceration sentences to first-offenders.

[29] See Hirschi and Gottfredson (1983), Lussier et al. (2015), Sampson and Laub (2017), and Britt (2019).

leaves prison, he is older and less prone to recidivism than when he was originally sentenced. These adjustments are necessarily rough approximations because we do not have the data required to estimate separate recidivism rates at each horizon for each possible age at sentencing. Instead, we use the vector of estimated age effects in each of our main outcome equations to create our age-adjustment factors. Appendix 11 describes the construction of these adjustment factors.

7.1 Impacts of Incapacitation and Shifting Recidivism Risk to Older Ages

Figure 5 presents our results for first offenders. The figure contains three lines: our estimates of the impacts of incapacitation effects on recidivism, our estimates of the combined impacts of incapacitation and aging during incarceration on recidivism, and our Figure 3 estimates of the total impacts of incarceration on recidivism,

Even though Appendix Table 10.11.2 reports that over half of the incarcerated compliers in the first offender sample serve less than six months in prison, incapacitation still accounts for roughly 80% of the total impact of incarceration on one-year recidivism rates. However, at horizons of five years and beyond, incapacitation accounts for less than one-third of the total impact of incarceration on recidivism, and the gaps between our incapacitation effect estimates and our estimates of the total impacts of incarceration on recidivism are, on average, more than 20 percentage points.

The impacts of aging while incarcerated are trivial at short horizons because prisoners released after serving short prison terms aged little while in prison. These impacts grow as we move to longer horizons and incorporate released prisoners who served longer prison terms, but even at seven years, we estimate that, among first offenders, the total impacts of prisoners aging while incarcerated reduce recidivism rates by only an additional three percentage points.

The results in Figure 5 suggest that, among first offenders, prison does more than incapacitate offenders and shift recidivism opportunities to older and less risky ages. It appears that the experience of prison reduces recidivism rates following release from prison.

Figure 6 presents parallel results for repeat offenders. Here, at almost all horizons, our estimates of the absolute reductions in recidivism due to incapacitation effects are roughly the same or somewhat larger than the reductions implied by our estimates of the total impacts of incarceration recidivism. The impacts of aging while incarcerated among repeat offenders are small and similar to those we estimate among first offenders. While the incarcerated compliers in our repeat offender sample serve longer prison terms than their counterparts in our first-offender sample, repeat offenders are also older, and the impacts of aging while incarcerated decline with age at sentencing. Taken together, our results for repeat offenders suggest that the direct impacts of incapacitation account for most or all of the full impacts of incarceration on recidivism at all horizons.

The contrasts between our results for first and repeat offenders are striking. However, our approach does require strong assumptions. Since our motivating thought experiment involves random assignment, we clearly must assume that, within our complier sets, the distribution of unobserved traits that influence $F(n|0, a_0)$ is the same for those who receive incarceration versus non-incarceration sentences. Further, we must also assume that, among incarcerated compliers, the counterfactual distribution $F(n|0, a_0)$ is the same for prisoners who receive short, medium, or long sentences. In addition, because we do not have enough data to estimate our models separately by age, our estimates of the impacts of aging during incarceration rest on the assumption that the relationships between age and expected recidivism rates among our compliers match the relationships we estimate in our 2SLS models by including age indicators in x_{it}. Finally, since our 2SLS models contain numerous conditioning variables and true judge stringency takes on 44 different values in our setting, existing methods for calculating the characteristics of compliers may weigh complier observations quite differently than our 2SLS models. Here, we employ interpolation methods proposed by Dahl et al. (2014) to estimate complier characteristics. Appendix 11 presents parallel results using a different method proposed in Garin et al. (2023), and these results are remarkably similar to those presented in Figures 5 and 6.

8 Conclusions and Policy Implications

Over the past forty years or more, the state of Illinois and many other jurisdictions have adopted numerous sentencing policies that require or encourage judges to show relative leniency to convicted offenders who have no prior convictions.[30] However, numerous legal scholars have recently raised equity concerns about sentencing laws that assign punishment for a given offense, in part, based on the existence of previous offenses that have already been punished.[31]

We show that, among first offenders in Chicago who are marginal candidates for incarceration, incarceration sentences produce noteworthy and lasting reductions in recidivism, but among marginal repeat offenders, incarceration treatment has no lasting impact on recidivism rates. These results suggest that it may be possible to hold Illinois prison populations constant and reduce overall recidivism rates by sentencing marginally more first offenders to prison and marginally fewer repeat offenders to prison.

[30] Most states in the US, most English-speaking countries, Scandinavian countries, China, India, and South Korea have also adopted policies that make sentencing relatively more severe for repeat offenders.

[31] See Reitz (2014), Hester et al. (2018), and D'Alessio and Stolzenberg (2019).

References

Agan, Amanda, Doleac, Jennifer and Harvey, Anna. (2021), Misdemeanor Prosecution, Technical report, NBER WP 28600.

Aizer, Anna and Doyle, Jr. Joseph J. (2015). 'Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges', *Quarterly Journal of Economics* 130(2), 759–803.

Bhuller, Manudeep, Dahl, Gordon B., Loken, Katrine V. and Mogstad, Magne. (2020). 'Incarceration, Recidivism and Employment', *Journal of Political Economy* 128(4), 1269–1324.

Bhuller, Manudeep and Sigstad, Henrik. (2023), 2SLS With Multiple Treatments. arXiv2205.07836v3.

Bogira, Steve. (2005), *Courtroom 302: A Year Behind the Scenes in an American Criminal Courthouse*, Vintage Books.

Britt, Chester L. (2019). 'Age and crime', *The Oxford handbook of developmental and life-course criminology* pp. 13–33.

Carson, E. Ann. (2019), Prisoners in 2019, Technical report, Bureau of Justice Statistics.

Dahl, Gordon B., Kostol, Andreas Ravndal and Mogstad, Magne. (2014). 'Family welfare cultures', *Quarterly Journal of Economics* 129.

D'Alessio, Stewart J. and Stolzenberg, Lisa. (2019). 'Should Repeat Offenders Be Punished More Severely for Their Crimes?', *Criminal Justice Policy Review* 30, 731–47.

Dobbie, W., G., Hans, N., Susan, P., Martin and Piksk, M. (2019), The Intergenerational Effects of Parental Incarceration, Technical report, NBER WP 24186.

Eren, Ozkan, Lovenheim, Michael F. and Mocan, Naci H. (2018). 'The Effect of Grade Retention on Adult Crime: Evidence from a Test-Based Promotion Policy', *NBER WP 25384* .

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

Garin, Andrew, Koustas, Dmitri, McPherson, Carl, Norris, Samuel, Pecenco, Matthew, Rose, Evan, Shem-Tov, Yotam and Weaver, Jeffrey. (2023), The Impact of Incarceration on Employment, Earnings, and Tax Filing. working paper.

Green, Daniel and Winik, Daniel. (2010). 'Using Random Judge Assignments to Estimate the Effects of Incarceration and Probation on Recidivism among Drug Offenders', *Criminology* 48(2), 357–387.

Hester, Rhys, Frase, Richard S., Roberts, Julian V. and Mitchell, Kelly Lyn. (2018). 'Prior Record Enhancements at Sentencing: Unsettled Justifications and Unsettling Consequences', *Crime and Justice* 47(1), 209–254.

Hirschi, Travis and Gottfredson, Michael. (1983). 'Age and the Explanation of Crime', *American Journal of Sociology* 89(3), 552–584.

Kolesar, Michal. (2013), Estimation in an Instrumental Variables Model with Treatment Effect Heterogeneity. Working Paper.

Kolesar, Michal, Chetty, Raj, Friedman, John, Glaeser, Edward and Imbens, Guido W. (2015). 'Identification and Inference With Many Invalid Instruments', *Journal of Business and Economic Statistics* 33(4), 474–484.

Kuziemko, Ilyana. (2013). 'How Should Inmates be Released from Prison? An Assessment of Parole Versus Fixed-Sentence Regimes', *Quarterly Journal of Economics* 128(1), 371–424.

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

Loeffler, Charles E and Nagin, Daniel S. (2022). 'The impact of incarceration on recidivism', *Annual Review of Criminology* 5, 133–152.

Lussier, Patrick, McCuish, Evan and Corrado, Raymond R. (2015). 'The Adolescence-Adulthood Transition and Desistance from Crime: Examining the Underlying Structure of Desistance', *Journal of Developmental and Life Course Criminology* 1(2), 87–117.

Mueller-Smith, Michael. (2015), The Criminal and Labor Market Impacts of Incarceration. Working Paper.

Nagin, Daniel S. and Snodgrass, Matthew G. (2013). 'The Effect of Incarceration on Re-Offending: Evidence from a Natural Experiment in Pennsylvania', *Journal of Quantitative Criminology* 29(4), 601–642.

Neal, Derek and Rick, Armin. (2016). 'The Prison Boom and Sentencing Policy', *Journal of Legal Studies* 45.

Neal, Derek and Rick, Armin. (forthcoming), The Role of Policy in Prison Growth and Decline, *in* **Michael T. Light and Ryan D. King**., eds, 'Oxford Handbook of Sentencing', Oxford University Press.

Norris, Samuel, Pecenco, Matthew and Weaver, Jeffrey. (2021). 'The Effects of Parental and Sibiling Incarceration: Evidence from Ohio', *American Economic Review* 111(9), 2926–63.

Paral, Rob. (2003), Demographic and Epidemiological Profiles of Key Chicago Community Areas and Suburban/Downstate Places, Technical report, Rob Paral and Associates.

Raphael, Steven and Stoll, Michael A. (2013), *Why Are So Many Americans in Prison*, Russell Sage Foundation.

Reitz, Kevin. (2014), The Illusion of Proportionality: Desert and Repeat Offenders, *in* **Juila V Roberts and Andrew Von Hirsch**., eds, 'Previous Convictions at Sentencing: Theoretical and Applied Perspectives', Hart Publishing.

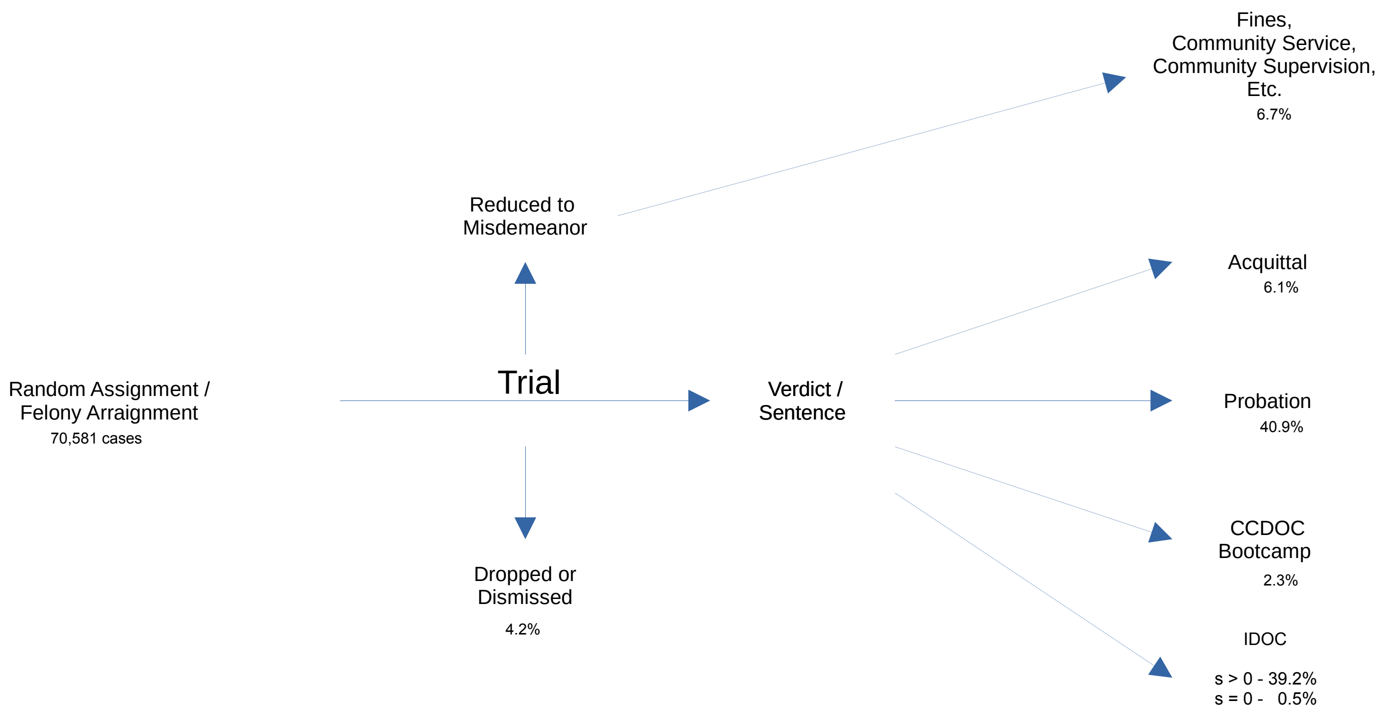
Roodman, David. (2017), The Impacts of Incarceration on Crime, Technical report, Open Philanthropy Project.

Rose, Evan K. and Shem-Tov, Yotam. (2021). 'How Does Incarceration Affect Reoffending? Estimating the Does-Response Function', *Journal of Political Economy* .

Sampson, Robert J and Laub, John H. (2017), Life-course desisters? Trajectories of crime among delinquent boys followed to age 70, *in* 'Developmental and life-course criminological theories', Routledge, pp. 37–74.

9 Figures and Tables

Figure 1
Felony Cases In Cook County

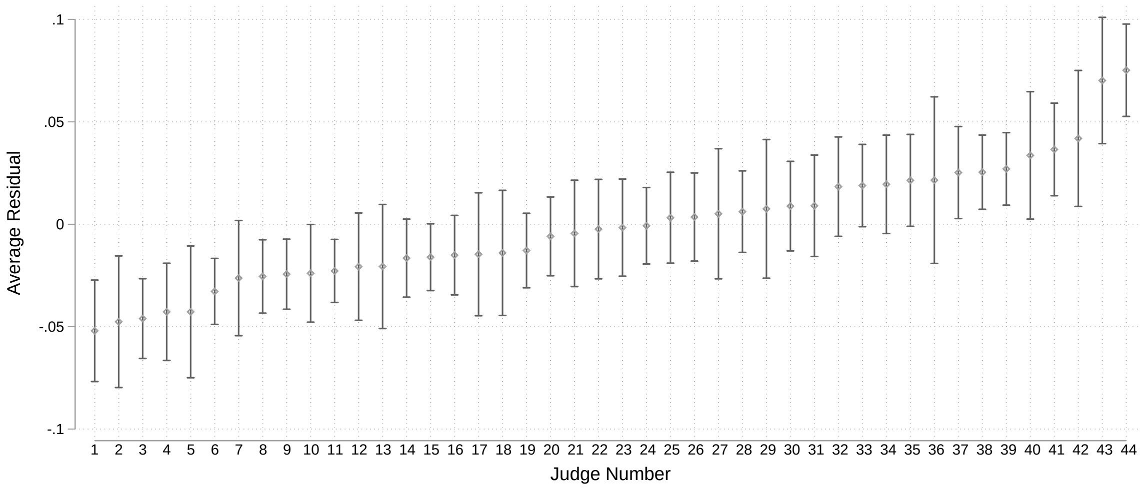


IDOC - Illinos Department of Corrections (State of IL)
CCDOC - Cook County Department of Correction (Sheriff)

Notes: This figure traces the flow of felony cases through the criminal courts in Cook County, IL. The figure starts with cases that have made it through the preliminary hearing stage and are eligible for random assignment. The Presiding Judge of the Criminal Division assigns cases randomly to calls run by specific judges, and defendants learn their judge assignments a few days later at their arraignment hearings. The vast majority of cases (83%) end in a felony conviction. However, some defendants have their cases dropped or dismissed. Some have all charges against them reduced to misdemeanors, and some are acquitted. Among those who receive a nominal sentence to IDOC, a small fraction receive so much credit for time-served in jail waiting a verdict that they do not have to serve time in prison. These are the $s = 0$ cases above. Cases that result in IDOC sentences that require defendants to serve prison time are $s > 0$.

Figure 2
Judge Severity Measures

Panel A: First Offenders



Panel B: Repeat Offenders

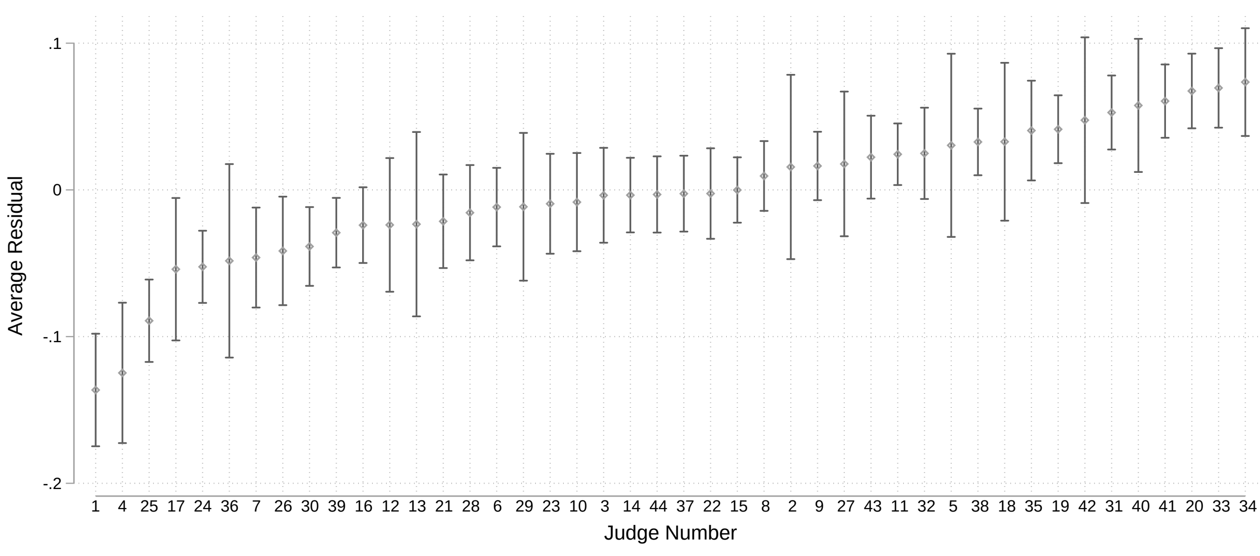
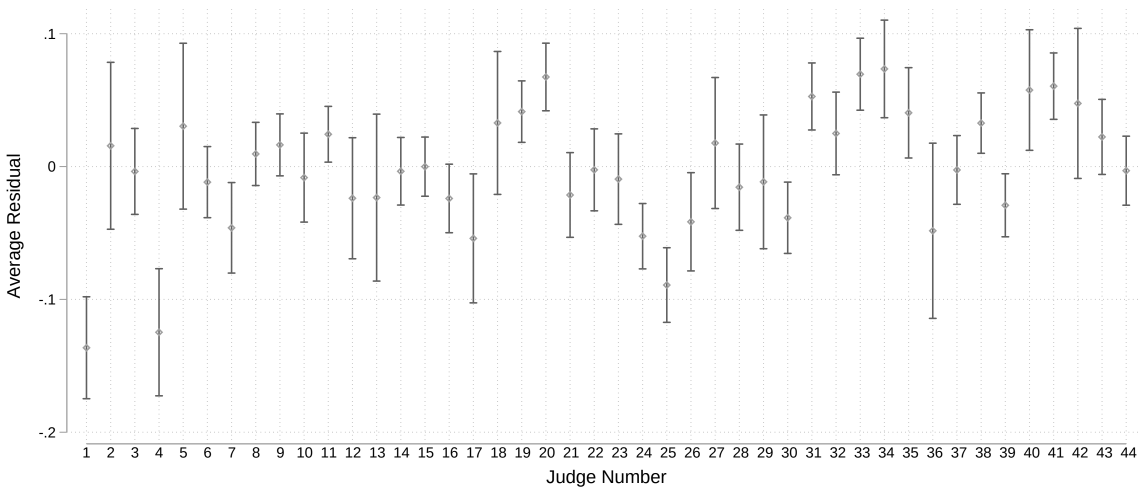


Figure 2
Judge Severity Measures

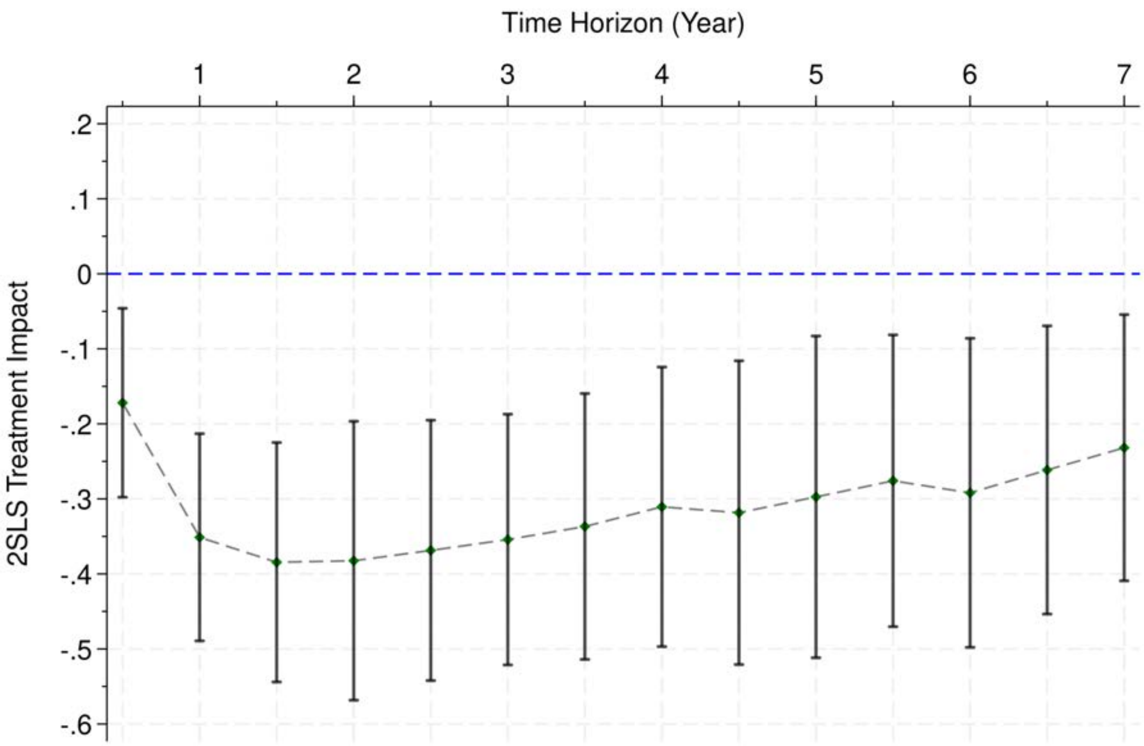
Panel C: Repeat Offenders -
Sorted by Judge Severity Among First Offenders



Notes: In all panels, we capture residuals from regressions of τ_{jit} on our full set of case and defendant characteristics listed in Appendix section 12.12. Each dot is the average sentencing residual for a judge, taken over the sample of first-offender cases or over the sample of repeat-offenders cases assigned to a given judge. In all three panels, we number judges according to their severity when dealing with first offenders. Judge 1 is the most lenient judge when dealing with first offenders. Judge 44 is most severe. The error bars are 95% confidence intervals. In Panel B, we label entries using these numbers, but we order judges on the X-axis by their severity when sentencing repeat offenders. For example, the third most lenient judge when dealing with repeat offenders is the 25th most severe when sentencing first offenders. Panel C presents the same information plotted in Panel B, but by sorting the X-axis by each judge's severity when sentencing first-offenders, this panel visually demonstrates the weakness of the correlation between our two measures of judge severity.

Figure 3
Impacts of Incarceration on Annual Recidivism Rates

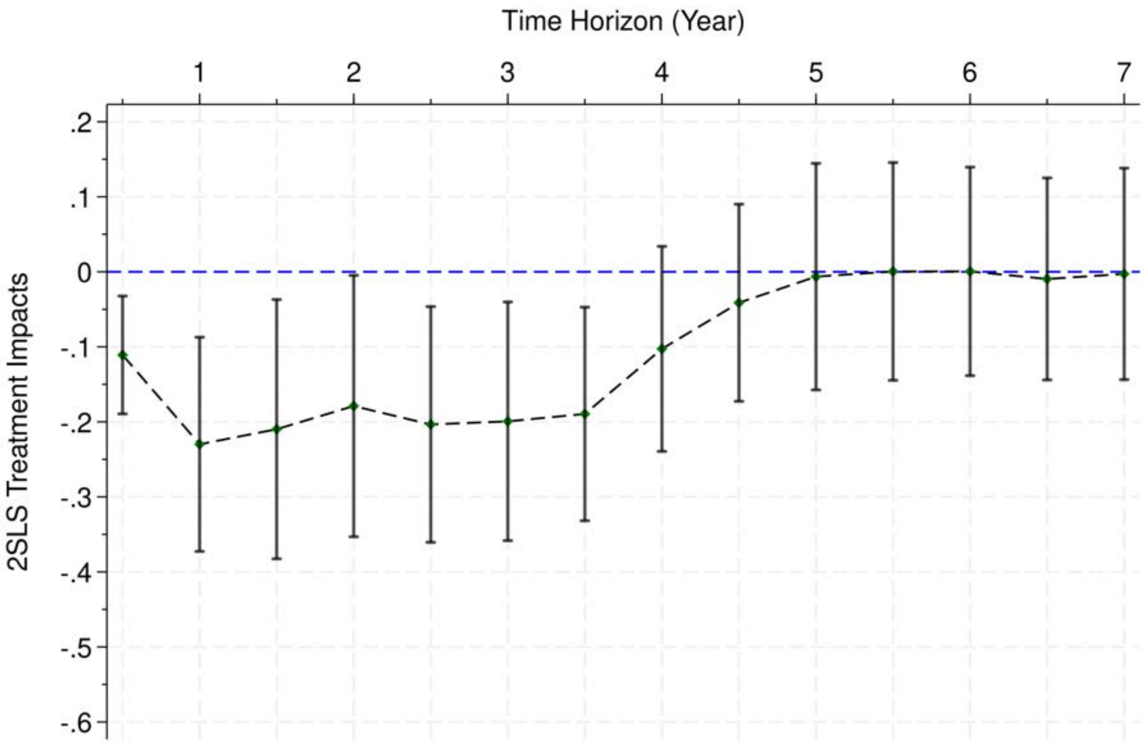
Panel A: First Offenders



Notes: This figure plots our estimates of the LATE of incarceration on recidivism rates among first offenders. Each dot is the estimated impact at a specific horizons. The horizons are $6, 12, 18, 24, ...84$ months The brackets around these estimates are 95% confidence intervals derived from HAC standard errors. We two-way cluster at the (judge-defendant) level. Given a five-percent significance level, the estimated impacts for horizons of 5-7 years are statistically different from the corresponding estimates plotted in Panel A above for first offenders. Appendix 11 presents these results in a table.

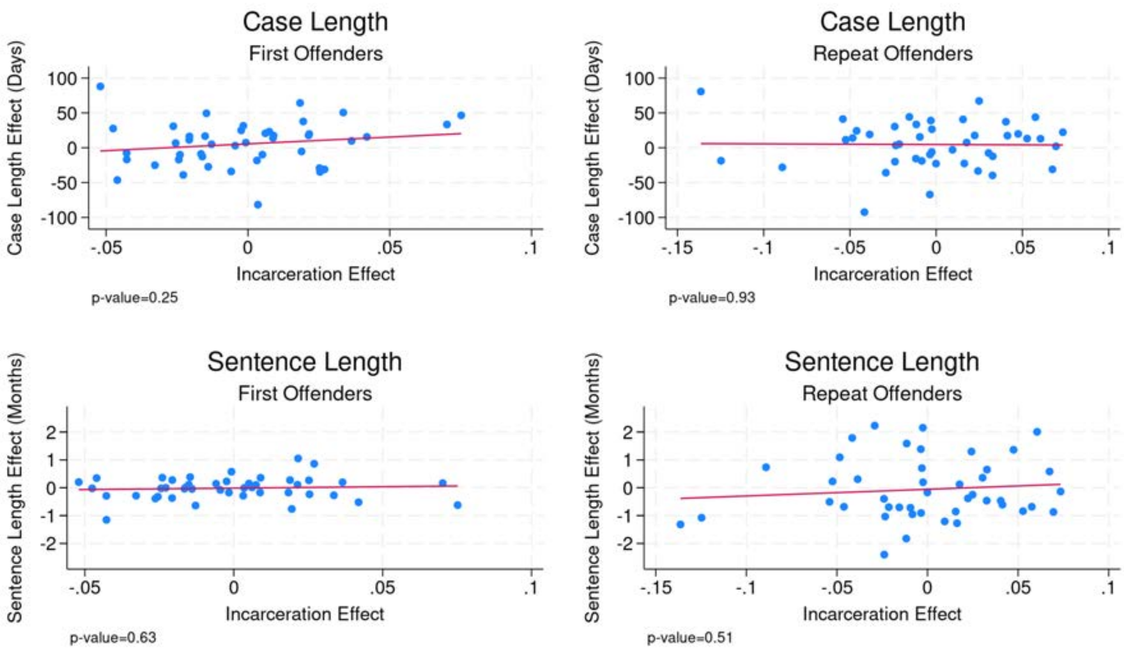
Figure 3
Impacts of Incarceration on Annual Recidivism Rates

Panel B: Repeat Offenders



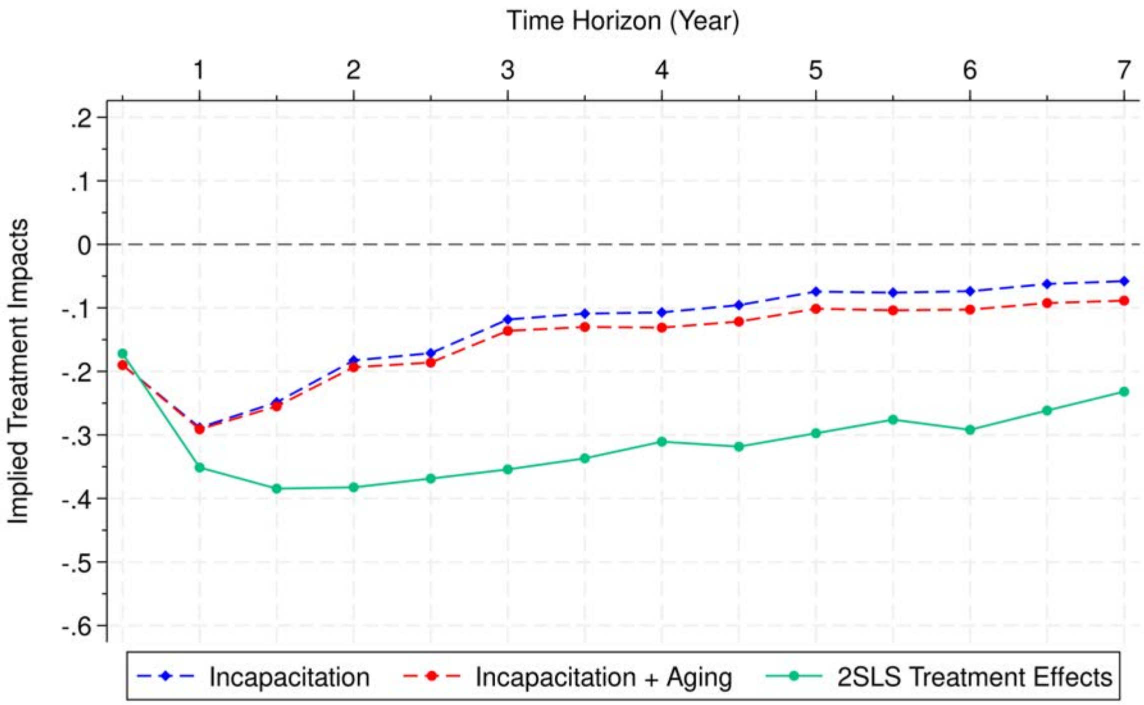
Notes: This figure plots our estimates of the LATE of incarceration on recidivism rates among repeat offenders. Each dot is the estimated impact at a specific horizons. The horizons are $6, 12, 18, 24, ...84$ months The brackets around these estimates are 95% confidence intervals derived from HAC standard errors. We two-way cluster at the (judge-defendant) level. Given a five-percent significance level, the estimated impacts for horizons of 5-7 years are statistically different from the corresponding estimates plotted in Panel A above for first offenders. Appendix 11 presents these results in a table.

Figure 4
Exclusion Plots



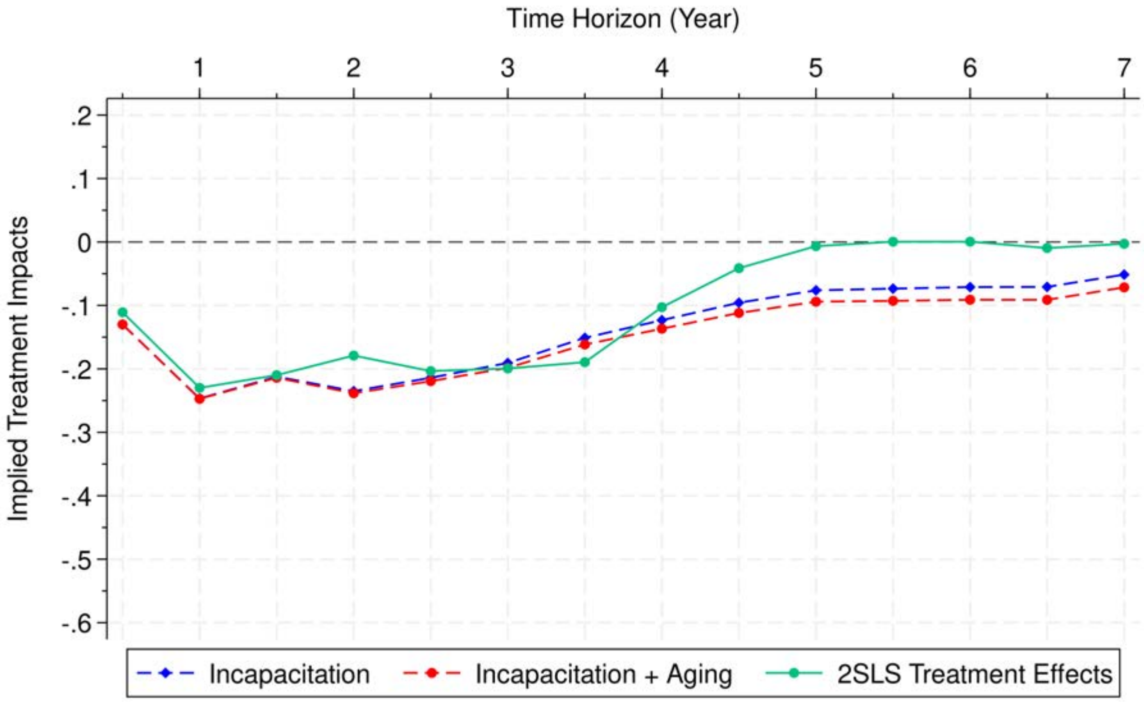
Notes: This figure reports the correlations between judge-level incarceration effects and judge-level case length and sentence length effects, separately for first and repeat offenders. Each judge-level effect is created by regressing one of these three outcome variables on our full control set. We run separate regression for first and repeat offenders, and we then form judge-level average residuals. Appendix Section 12 details our control set. The incarceration and case length regressions employ the first and repeat-offender samples that we use in our main analyses. The sentence length regressions employ the samples of defendants who receive incarceration sentences. Each judge-effects plot places the incarceration effect on the x-axis and either case-length effect or sentence-length effect on the y-axis. The linear fit is shown in red, and the p-values associated with the nulls that each slope is zero are printed below. We topcode the length of any given case at 1,000 days and the length of a sentence at 96 months. We measure sentence length as expected months of incarceration given the nominal sentence and credits for time-served in jail waiting a verdict.

Figure 5
First Offenders



Notes: For first offenders, we plot the total impacts of incarceration on recidivism that we present in Figure 3. We also plot our estimates of incapacitation effects and our estimates of the sum of incapacitation effects and the impacts of aging during incarceration. Let $R^0(n)$ be the recidivism rate among non-incarcerated compliers after n periods. Let $h(m)$ be the fraction of incarcerated compliers who exit prison during the m-th period after sentencing. If incapacitation is the only impact of incarceration, then the n-period recidivism rate among incarcerated compliers is $R^1(n) = \Sigma_{m=1}^{n} R^0(n-m)h(m)$. This formula applies the recidivism rates among non-incarcerated compliers to incarcerated compliers starting with the period they exit prison. Appendix 10.11.1, which presents estimates of recidivism rates among non-incarcerated compliers, gives $\hat{R}^0(n)$ for each $n > 0$. Appendix 10.11.2, which presents estimates of the time-served distribution among incarcerated compliers, provides the $\hat{h}(m)$ values. We also plot the combined impacts of incapacitation and aging while incarcerated among compliers sentenced to incarceration. Here, we adjust each value of, $\hat{R}^0(n-m)$, to account for aging during incarceration. The adjusted formula is: $\hat{R}^{0*}(n-m) = \hat{R}^0(n-m) - \Delta\hat{r}_{(n-m)}\Delta\hat{a}_m$, where $\Delta\hat{r}_{(n-m)}$ is our estimate of the absolute reduction in the $(n-m)$-year recidivism rates associated with aging one year, and $\Delta\hat{a}_m$ is the average years served prior to release among persons who served more than $m-1$ periods but less than m in prison. See Appendix 11 for more details.

Figure 6
Repeat Offenders



Notes: See note for Figure 5.

Table 1 - Descriptive Statistics

	First Offenders	Repeat Offenders
Age	21.12	25.91
Black	0.68	0.84
Prior Charges	.	2.64
Class X	0.16	0.15
Class 1	0.16	0.18
Class 2	0.30	0.33
Class 3	0.10	0.16
Class 4	0.28	0.18
High-Crime Area	0.56	0.71
Drug	0.42	0.46
Robbery	0.12	0.10
Burglary	0.11	0.09
Assault	0.05	0.05
Theft	0.11	0.10
Weapon	0.16	0.17
Guilty	0.90	0.89
Probation	0.71	0.22
Prison	0.17	0.65
CCDOC Bootcamp	0.03	0.02
On MSR	.	0.41
Case Length (Years)	0.48	0.53
Sample Size	37,055	33,526

Notes: These descriptive statistics describe our two analysis samples. The Appendix materials in section 12 detail the construction of these samples. The entries Guilty, Probation, Prison, and CCDOC Bootcamp describe sentencing outcomes. All other entries are characteristics of the defendant or the case against the defendant. Class X is the most serious offense class. Class 4 is the least serious.

Table 2 - Balance

	All	First Offenders	Repeat Offenders
Black	-0.000090 (p=0.959)	-0.000177 (p=0.704)	0.000610 (p=0.260)
Age	-0.000014 (p=0.968)	0.000037 (p=0.176)	0.000029 (p=0.585)
Height	0.000049 (p=0.451)	0.000010 (p=0.744)	0.000111 (p=0.121)
Weight	-0.000000 (p=0.996)	0.000001 (p=0.806)	0.000002 (p=0.701)
BMI	-0.000021 (p=0.815)	0.000006 (p=0.833)	-0.000027 (p=0.600)
Prior Cases	-0.000089 (p=0.955)	.	0.000178 (p=0.139)
Indictment	-0.000237 (p=0.550)	-0.000198 (p=0.553)	-0.000280 (p=0.619)
Multiple Defendant	-0.000339 (p=0.376)	-0.000253 (p=0.488)	-0.000475 (p=0.433)
Multiple Charge	0.000054 (p=0.902)	0.000207 (p=0.601)	-0.000237 (p=0.696)
Robbery	-0.000329 (p=0.593)	-0.000631 (p=0.315)	-0.000288 (p=0.753)
Assault	0.000579 (p=0.324)	0.000900 (p=0.208)	0.000207 (p=0.770)
Burglary	-0.000424 (p=0.312)	-0.000590 (p=0.112)	-0.000329 (p=0.661)
Theft	-0.000210 (p=0.609)	-0.000264 (p=0.623)	-0.000124 (p=0.829)
Other Non-Violent	0.000476 (p=0.582)	0.000440 (p=0.697)	0.000642 (p=0.644)
Drug	0.000214 (p=0.478)	0.000286 (p=0.405)	0.000270 (p=0.412)
Weapon	-0.000030 (p=0.950)	0.000112 (p=0.750)	-0.000194 (p=0.606)
High-Crime Area	-0.000174 (p=0.900)	-0.000410 (p=0.274)	0.000396 (p=0.400)
Class 0	0.000049 (p=0.913)	-0.000348 (p=0.485)	0.000586 (p=0.454)
Class 1	0.000006 (p=0.991)	-0.000138 (p=0.727)	0.000295 (p=0.586)
Class 2	-0.000262 (p=0.664)	-0.000499 (p=0.191)	0.000347 (p=0.435)
Class 3	0.000925 (p=0.350)	0.001447 (p<0.01)	-0.000183 (p=0.646)
Class 4	-0.000320 (p=0.826)	0.000146 (p=0.657)	-0.001152 (p=0.157)
All	. (p=0.226)	. (p=0.035)	. (p=0.387)

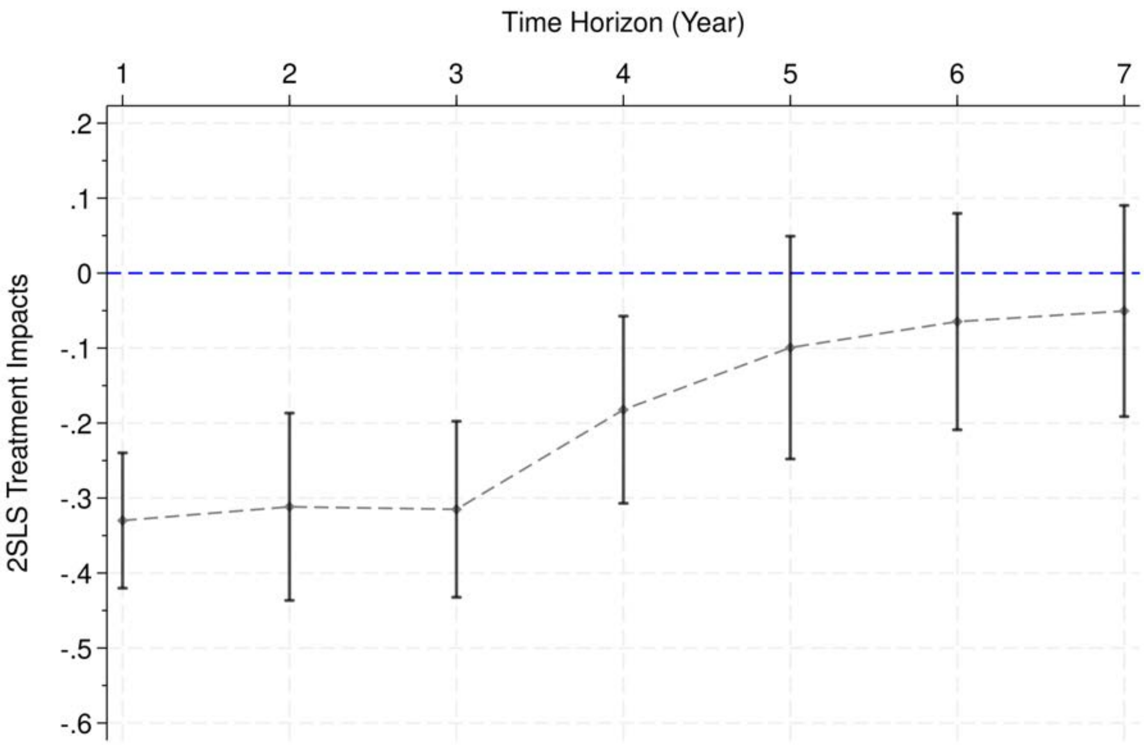
Notes: Each row reports three regression coefficients, e.g. the row *Age* reports the coefficients on age from three regressions of our LOM severity measure, $z_{j(i,t)}$, on age at arraignment and dummies for year of case assignment. The first regression pools all first and repeat offenders in one regression but employs $z_{j(i,t)}$ measures that are specific to the first vs repeat-offender status of defendant i. The other regressions restrict the sample to cases that involve either first or repeat offenders. We report p-values derived from HAC standard errors, and we cluster at the judge level. The final row presents results from regressing LOM severity measures on the entire set of characteristics. Here, the p-values are associated with the joint test that none of the characteristics predict judge severity.

10 Appendix Figures and Tables

10.1 Results from Pooled Models

Appendix Figure 10.1
Impacts of Incarceration on Annual Recidivism Rates: Models With Some Form of Pooling

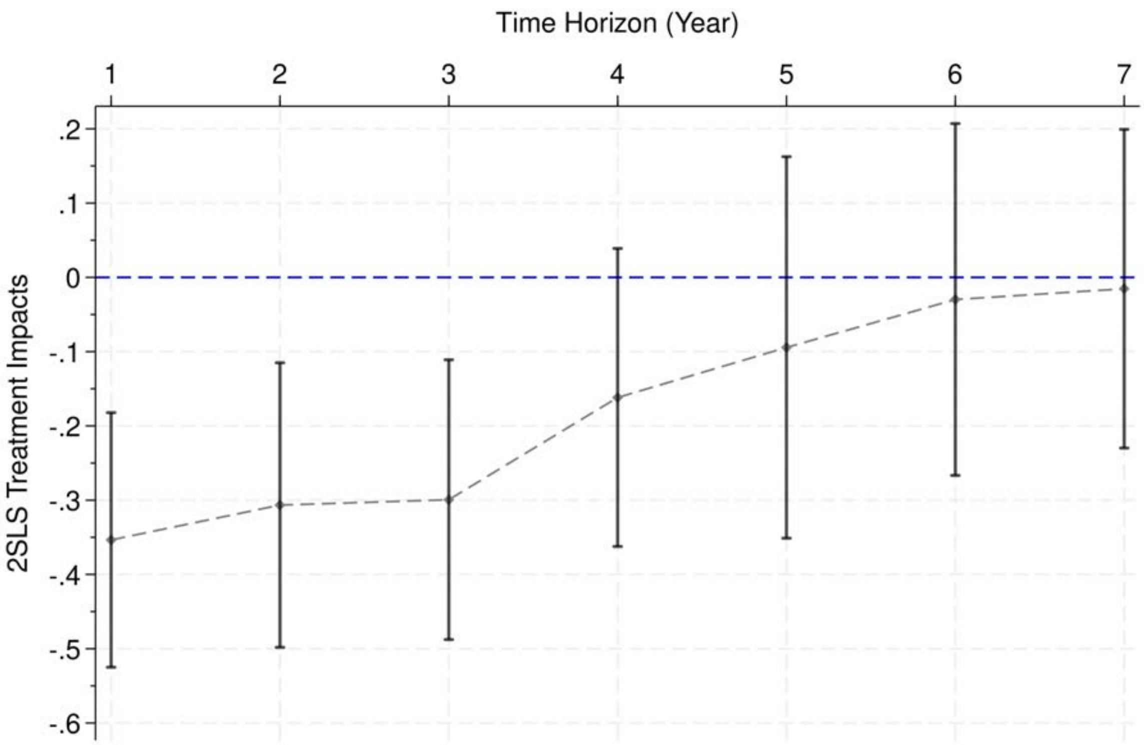
Panel A: Single Outcome Equation For Each Horizon -
w/ LOM Severity Measure Created Using Full-Sample Residuals



Notes: This figure plots our estimates of the LATE of incarceration on recidivism rates at different horizons when we estimate pooled 2SLS models that contain one outcome equation that is common to both first and repeat offenders. The first-stage equations also employ a single full-sample LOM measure of judge severity that captures average judge severity over all cases. Each dot is the estimated impact at a specific annual horizon of $1, 2, 3, .., 7$ years. The brackets around these estimates are 95% confidence intervals derived from HAC standard errors. We two-way cluster at the (judge-defendant) level.

Appendix Figure 10.1
Impacts of Incarceration on Annual Recidivism Rates: Models With Some Form of Pooling

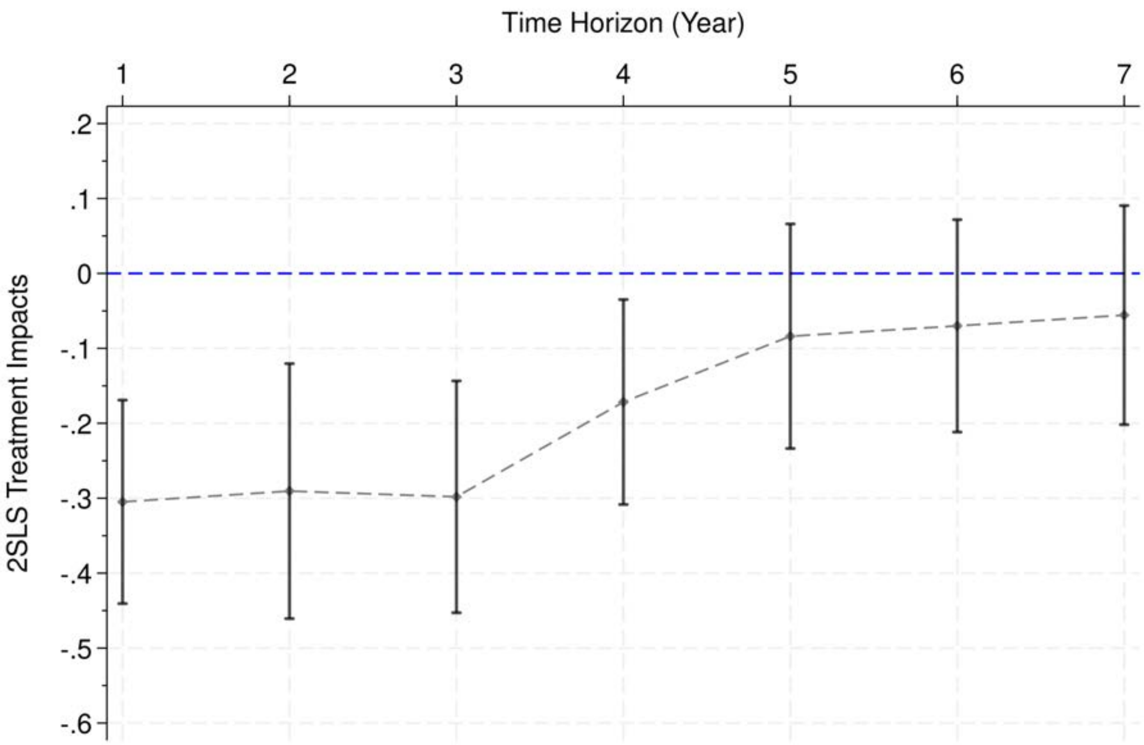
Panel B: Outcome Equations For First Offenders w/ LOM Severity Measure Created Using Full-Sample Residuals



Notes: This figure plots our estimates of the LATE of incarceration on recidivism rates among first offenders from 2SLS models that parallel the models that produce the results in Panel A of Figure 3, but here, we do not use an LOM severity measure derived solely from first-offender cases. Instead, we instrument for incarceration treatment in the first-stage using an LOM severity measure derived from all cases assigned to each judge. Again, each dot is the estimated impact at a specific annual horizon of $1, 2, 3, .., 7$ years. The brackets around these estimates are 95% confidence intervals derived from HAC standard errors. We cluster at the judge level.

Appendix Figure 10.1
Impacts of Incarceration on Annual Recidivism Rates: Models With Some Form of Pooling

Panel C: Outcome Equations For Repeat Offenders
w/ LOM Severity Measure Created Using Full-Sample Residuals



Notes: This figure plots our estimates of the LATE of incarceration on recidivism rates among repeat offenders from 2SLS models that parallel the models that produce the results in Panel B of Figure 3, but here, we do not use an LOM severity measure derived solely from repeat-offender cases. Instead, we instrument for incarceration treatment in the first-stage using an LOM severity measure derived from all cases assigned to each judge. Again, each dot is the estimated impact at a specific annual horizon of $1, 2, 3, .., 7$ years. The brackets around these estimates are 95% confidence intervals derived from HAC standard errors. We two-way cluster at the (judge-defendant) level.

10.2 Balance

Appendix Table 10.2
Alternative Balance Tests

LOM = average deviations from year-specific incarcerations rates

	All	First Offenders	Repeat Offenders
Black	-0.000158 (p=0.936)	-0.000527 (p=0.250)	0.000640 (p=0.245)
Age	0.000026 (p=0.951)	0.000044 (p=0.179)	0.000029 (p=0.601)
Height	0.000048 (p=0.523)	-0.000002 (p=0.946)	0.000107 (p=0.184)
Weight	0.000001 (p=0.955)	0.000001 (p=0.882)	0.000002 (p=0.811)
BMI	-0.000013 (p=0.897)	0.000006 (p=0.837)	-0.000031 (p=0.561)
Prior Cases	0.000042 (p=0.982)	.	0.000174 (p=0.182)
Indictment	-0.000098 (p=0.827)	0.000058 (p=0.881)	-0.000323 (p=0.594)
Multiple Defendant	-0.000275 (p=0.508)	-0.000176 (p=0.655)	-0.000462 (p=0.455)
Multiple Charge	0.000136 (p=0.771)	0.000444 (p=0.295)	-0.000253 (p=0.693)
Robbery	-0.000224 (p=0.744)	-0.000108 (p=0.873)	-0.000525 (p=0.591)
Assault	0.000609 (p=0.341)	0.001006 (p=0.190)	0.000124 (p=0.870)
Burglary	-0.000135 (p=0.759)	-0.000237 (p=0.583)	-0.000136 (p=0.870)
Theft	-0.000274 (p=0.548)	-0.000425 (p=0.421)	-0.000191 (p=0.754)
Other Non-Violent	0.000568 (p=0.504)	0.000334 (p=0.770)	0.000718 (p=0.617)
Drug	0.000197 (p=0.567)	0.000147 (p=0.676)	0.000257 (p=0.425)
Weapon	-0.000220 (p=0.683)	-0.000119 (p=0.761)	-0.000089 (p=0.828)
High-Crime Area	-0.000076 (p=0.961)	-0.000410 (p=0.256)	0.000393 (p=0.437)
Class 0	0.000128 (p=0.789)	-0.000119 (p=0.825)	0.000436 (p=0.601)
Class 1	0.000047 (p=0.937)	-0.000068 (p=0.870)	0.000159 (p=0.777)
Class 2	-0.000176 (p=0.798)	-0.000557 (p=0.153)	0.000375 (p=0.437)
Class 3	0.000825 (p=0.477)	0.001461 (p<0.01)	-0.000072 (p=0.866)
Class 4	-0.000447 (p=0.787)	0.000008 (p=0.980)	-0.001028 (p=0.224)
All	. (p=0.551)	. (p=0.080)	. (p=0.316)

Notes: Each row reports three regression coefficients, e.g. the row *Age* reports the coefficients on age from three regressions of alternative LOM severity measures, $z_{j(i,t)}$, on age at arraignment and dummies for year of case assignment. These LOM severity measures are sums of residuals taken from projections of $\tau_{j(i,t)}$ on only dummies for year of case assignments. The first regression pools all first and repeat offenders in one regression but employs $z_{j(i,t)}$ measures that are specific to the first vs repeat-offender status of defendant i. The other regressions restrict the sample to cases that involve either first or repeat offenders. We report p-values derived from HAC standard errors, and we cluster at the judge level. The final row presents results from regressing LOM severity measures on the entire set of characteristics. Here, the p-values are associated with the joint test that none of the characteristics predict judge severity.

10.3 Table of Main Results

Appendix Table 10.3

Expanded Tabular Presentation of Figure 3 Results

Panel A: Main Results
HAC Standard Errors Clustered by Judge

First Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.19	-0.146 (0.006) [p<0.01]	-0.290 (0.059) [p<0.01]	-0.351 (0.070) [p<0.01]
New Charge <24m	0.31	-0.124 (0.011) [p<0.01]	-0.316 (0.084) [p<0.01]	-0.382 (0.095) [p<0.01]
New Charge <36m	0.38	-0.091 (0.009) [p<0.01]	-0.292 (0.075) [p<0.01]	-0.354 (0.085) [p<0.01]
New Charge <48m	0.44	-0.064 (0.009) [p<0.01]	-0.256 (0.083) [p<0.01]	**-0.311 (0.095) [p<0.01]**
New Charge <60m	0.47	-0.048 (0.009) [p<0.01]	***-0.245 (0.094) [p=0.01]***	***-0.297 (0.109) [p<0.01]***
New Charge <72m	0.50	-0.036 (0.009) [p<0.01]	***-0.241 (0.090) [p=0.01]***	***-0.292 (0.105) [p<0.01]***
New Charge <84m	0.52	-0.027 (0.009) [p<0.01]	**-0.191 (0.077) [p=0.02]**	***-0.232 (0.091) [p=0.01]***

$\bar{\tau} = 0.19$, Standard Deviation of LOM: .028, F-Statistic: 284, N: 37,055
$f(l)$: 0 (81%), $(0, 12]$ (8%), $(12, 24]$ (4%), $(24, 36]$ (4%), $(36, 48]$ (2%), $(48, 60]$ (1%), $[60, \infty)$ (1%)

Repeat Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.14	-0.188 (0.007) [p<0.01]	-0.201 (0.067) [p<0.01]	-0.230 (0.073) [p<0.01]
New Charge <24m	0.31	-0.148 (0.008) [p<0.01]	-0.157 (0.080) [p=0.06]	-0.179 (0.089) [p=0.04]
New Charge <36m	0.44	-0.090 (0.008) [p<0.01]	-0.174 (0.073) [p=0.02]	-0.199 (0.081) [p=0.01]
New Charge <48m	0.52	-0.052 (0.009) [p<0.01]	-0.090 (0.062) [p=0.15]	**-0.103 (0.070) [p=0.14]**
New Charge <60m	0.58	-0.034 (0.009) [p<0.01]	***-0.006 (0.068) [p=0.93]***	***-0.006 (0.077) [p=0.93]***
New Charge <72m	0.62	-0.018 (0.008) [p=0.03]	***0.001 (0.063) [p=0.99]***	***0.001 (0.071) [p=0.99]***
New Charge <84m	0.65	-0.008 (0.008) [p=0.28]	**-0.002 (0.064) [p=0.97]**	***-0.003 (0.072) [p=0.97]***

$\bar{\tau} = 0.66$, Standard Deviation of LOM: .043, F-Statistic: 782, N: 33,526
$f(l)$: 0 (34%), $(0, 12]$ (35%), $(12, 24]$ (14%), $(24, 36]$ (10%), $(36, 48]$ (3%), $(48, 60]$ (2%), $[60, \infty)$ (2%)

Notes: Each panel reports results from seven OLS, RF, and 2SLS models. In the OLS and 2SLS models, each entry is the estimated coefficient on $\tau_{j(i,t)}$, which is an indicator that equals one if judge j assigns an incarceration sentence to defendant i at date t. In the RF column, each entry is the estimated coefficient on $z_{j(i,t)}$, the LOM severity measure associated with judge j. In each row, the outcome variable is an indicator for the presence of at least one new charge before a given horizon. The F-statistics are test statistics for the null that $z_{j(i,t)}$ does not predict $\tau_{j(i,t)}$ given our controls for case and defendant characteristics. We report HAC standard errors. For first offenders, we cluster at the judge level. For repeat offenders, we two-way cluster at the defendant*judge level. $\bar{\tau}$ gives the fraction of the sample that received an incarceration sentence. Entries in **bold** type are treatment impacts that are statistically different among first versus repeat offenders given $p = .1$, and entries in ***bold italics*** are different given $p = .05$. $f(l)$ is a discrete density that describes the distribution of expected incarceration time given the sentences assigned to defendants. Note that $f(0) = 1 - \bar{\tau}$ by definition.

Panel B: Main Results
w/ Huber White Standard Errors

First Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.19	-0.146 (0.005) [p<0.01]	-0.290 (0.071) [p<0.01]	-0.351 (0.087) [p<0.01]
New Charge <24m	0.31	-0.124 (0.007) [p<0.01]	-0.316 (0.082) [p<0.01]	**-0.382 (0.101) [p<0.01]**
New Charge <36m	0.38	-0.091 (0.007) [p<0.01]	-0.292 (0.086) [p<0.01]	-0.354 (0.106) [p<0.01]
New Charge <48m	0.44	-0.064 (0.007) [p<0.01]	-0.256 (0.087) [p<0.01]	-0.311 (0.107) [p<0.01]
New Charge <60m	0.47	-0.048 (0.007) [p<0.01]	***-0.245 (0.088) [p<0.01]***	***-0.297 (0.108) [p<0.01]***
New Charge <72m	0.50	-0.036 (0.007) [p<0.01]	***-0.241 (0.088) [p<0.01]***	***-0.292 (0.108) [p<0.01]***
New Charge <84m	0.52	-0.027 (0.007) [p<0.01]	**-0.191 (0.088) [p=0.03]**	**-0.232 (0.107) [p=0.03]**

$\bar{\tau} = 0.19$, Standard Deviation of LOM: .028, F-Statistic: 161, N: 37,055
$f(l) : 0\,(81\%),\ (0, 12]\,(8\%),\ (12, 24]\,(4\%),\ (24, 36]\,(4\%),\ (36, 48]\,(2\%),\ (48, 60]\,(1\%),\ [60, \infty)\,(1\%)$

Repeat Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.14	-0.188 (0.005) [p<0.01]	-0.201 (0.045) [p<0.01]	-0.230 (0.050) [p<0.01]
New Charge <24m	0.31	-0.148 (0.006) [p<0.01]	-0.157 (0.057) [p<0.01]	**-0.179 (0.065) [p<0.01]**
New Charge <36m	0.44	-0.090 (0.006) [p<0.01]	-0.174 (0.061) [p<0.01]	-0.199 (0.069) [p<0.01]
New Charge <48m	0.52	-0.052 (0.006) [p<0.01]	-0.090 (0.061) [p=0.14]	-0.103 (0.070) [p=0.14]
New Charge <60m	0.58	-0.034 (0.006) [p<0.01]	***-0.006 (0.060) [p=0.93]***	***-0.006 (0.069) [p=0.93]***
New Charge <72m	0.62	-0.018 (0.006) [p<0.01]	***0.001 (0.059) [p=0.99]***	***0.001 (0.068) [p=0.99]***
New Charge <84m	0.65	-0.008 (0.006) [p=0.15]	**-0.002 (0.058) [p=0.97]**	**-0.003 (0.067) [p=0.97]**

$\bar{\tau} = 0.66$, Standard Deviation of LOM: .043, F-Statistic: 247, N: 33,526
$f(l) : 0\,(34\%),\ (0, 12]\,(35\%),\ (12, 24]\,(14\%),\ (24, 36]\,(10\%),\ (36, 48]\,(3\%),\ (48, 60]\,(2\%),\ [60, \infty)\,(2\%)$

Panel C: Main Results
w/ HAC Standard Errors Clustered on Day of Assignment

First Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.19	-0.146 (0.005) [p<0.01]	-0.290 (0.072) [p<0.01]	-0.351 (0.088) [p<0.01]
New Charge <24m	0.31	-0.124 (0.007) [p<0.01]	-0.316 (0.085) [p<0.01]	**-0.382 (0.105) [p<0.01]**
New Charge <36m	0.38	-0.091 (0.007) [p<0.01]	-0.292 (0.089) [p<0.01]	-0.354 (0.110) [p<0.01]
New Charge <48m	0.44	-0.064 (0.007) [p<0.01]	-0.256 (0.089) [p<0.01]	-0.311 (0.110) [p<0.01]
New Charge <60m	0.47	-0.048 (0.007) [p<0.01]	***-0.245 (0.090) [p<0.01]***	***-0.297 (0.111) [p<0.01]***
New Charge <72m	0.50	-0.036 (0.008) [p<0.01]	***-0.241 (0.090) [p<0.01]***	***-0.292 (0.111) [p<0.01]***
New Charge <84m	0.52	-0.027 (0.008) [p<0.01]	**-0.191 (0.090) [p=0.03]**	**-0.232 (0.111) [p=0.04]**

$\bar{\tau} = 0.19$, Standard Deviation of LOM: 0.028, F-Statistic: 153, N: 37,055
$f(l) : 0$ (81%), $(0, 12]$ (8%), $(12, 24]$ (4%), $(24, 36]$ (4%), $(36, 48]$ (2%), $(48, 60]$ (1%), $[60, \infty)$ (1%)

Repeat Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.14	-0.188 (0.005) [p<0.01]	-0.201 (0.045) [p<0.01]	-0.230 (0.050) [p<0.01]
New Charge <24m	0.31	-0.148 (0.006) [p<0.01]	-0.157 (0.057) [p<0.01]	**-0.179 (0.065) [p<0.01]**
New Charge <36m	0.44	-0.090 (0.006) [p<0.01]	-0.174 (0.060) [p<0.01]	-0.199 (0.068) [p<0.01]
New Charge <48m	0.52	-0.052 (0.006) [p<0.01]	-0.090 (0.060) [p=0.13]	-0.103 (0.068) [p=0.13]
New Charge <60m	0.58	-0.034 (0.006) [p<0.01]	***-0.006 (0.060) [p=0.92]***	***-0.006 (0.068) [p=0.92]***
New Charge <72m	0.62	-0.018 (0.006) [p<0.01]	***0.001 (0.058) [p=0.99]***	***0.001 (0.067) [p=0.99]***
New Charge <84m	0.65	-0.008 (0.006) [p=0.16]	**-0.002 (0.058) [p=0.97]**	**-0.003 (0.066) [p=0.97]**

$\bar{\tau} = 0.66$, Standard Deviation of LOM: 0.043, F-Statistic: 238, N: 33,526
$f(l) : 0$ (34%), $(0, 12]$ (35%), $(12, 24]$ (14%), $(24, 36]$ (10%), $(36, 48]$ (3%), $(48, 60]$ (2%), $[60, \infty)$ (2%)

10.4 Alternative Models

Appendix Table 10.4
Alternative Model Specifications

Panel A: 2SLS Models with Alternative LOMs and Different Regression Controls

FIRST OFFENDERS

	Model (1)	Model (2)	Model (3)
$< 12m$	-0.385 (0.080) [p<0.01]	-0.341 (0.083) [p<0.01]	-0.300 (0.081) [p<0.01]
$< 24m$	**-0.421 (0.108) [p<0.01]**	-0.364 (0.114) [p<0.01]	-0.311 (0.113) [p<0.01]
$< 36m$	-0.398 (0.100) [p<0.01]	-0.349 (0.105) [p<0.01]	-0.288 (0.103) [p<0.01]
$< 48m$	***-0.353 (0.109) [p<0.01]***	**-0.331 (0.107) [p<0.01]**	-0.270 (0.106) [p=0.01]
$< 60m$	**-0.342 (0.122) [p<0.01]**	***-0.325 (0.118) [p<0.01]***	**-0.265 (0.117) [p=0.02]**
$< 72m$	***-0.340 (0.118) [p<0.01]***	***-0.342 (0.116) [p<0.01]***	***-0.280 (0.113) [p=0.01]***
$< 84m$	***-0.278 (0.103) [p<0.01]***	***-0.279 (0.100) [p<0.01]***	**-0.217 (0.096) [p=0.02]**

Model (1): LOM = residuals given full control set, x_{it}= indicators for year
Model (2): LOM = residuals given indicators for year, x_{it}= indicators for year
Model (3): LOM = residuals given indicators for year, x_{it}= full control set

REPEAT OFFENDERS

	Model (1)	Model (2)	Model (3)
$< 12m$	-0.228 (0.071) [p¡0.01]	-0.216 (0.076) [p¡0.01]	-0.218 (0.077) [p¡0.01]
$< 24m$	**-0.175 (0.091) [p=0.05]**	-0.158 (0.097) [p=0.10]	-0.164 (0.093) [p=0.08]
$< 36m$	-0.187 (0.086) [p=0.03]	-0.178 (0.092) [p=0.05]	-0.192 (0.085) [p=0.02]
$< 48m$	***-0.089 (0.076) [p=0.24]***	**-0.079 (0.079) [p=0.31]**	-0.094 (0.071) [p=0.18]
$< 60m$	***0.005 (0.084) [p=0.95]***	***0.012 (0.086) [p=0.89]***	**-0.001 (0.078) [p=0.99]**
$< 72m$	***0.012 (0.077) [p=0.87]***	***0.017 (0.078) [p=0.83]***	***0.004 (0.070) [p=0.96]***
$< 84m$	***0.009 (0.078) [p=0.91]***	***0.012 (0.078) [p=0.87]***	**-0.001 (0.071) [p=0.99]**

Notes: We present results from 3 alternative 2SLS models that employ LOM residual-severity measures as instruments for $\tau_{j(it,)}$. Model (1) employs the same LOM measures we use in Figure 3. Models (2) and (3) employ LOM severity measures that are sums of residuals taken from projections of $\tau_{j(i,t)}$ on only a set of indicators for year of case assignments. Models (1) and (2) employ only year of case assignment as controls in the first and second stage equations. Model (3) uses our full conditioning set. See Appendix section 12.12 for details. We report HAC standard errors. For first offenders, we cluster at the judge level. For repeat offenders, we two-way cluster at the defendant*judge level. Entries in **bold** type are treatment impacts that are statistically different among first versus repeat offenders given $p = .1$, and entries in ***bold italics*** are different given $p = .05$.

Panel B: 2SLS with Vector of Judge Indicators as Instruments

FIRST OFFENDERS

Baseline 2SLS with $z_{j(i,t)}$ = indicators for $j(i,t)$

	x_{it} = year indicators	x_{it} = all controls
$< 12m$	-0.296 (0.063) [p<0.01]	-0.319 (0.057) [p<0.01]
$< 24m$	-0.315 (0.088) [p<0.01]	-0.334 (0.081) [p<0.01]
$< 36m$	-0.296 (0.082) [p<0.01]	-0.306 (0.075) [p< 0.01]
$< 48m$	**-0.275 (0.084) [p<0.01]**	**-0.265 (0.083) [p<0.01]**
$< 60m$	***-0.264 (0.091) [p<0.01]***	***-0.248 (0.094) [p=0.01]***
$< 72m$	***-0.271 (0.089) [p<0.01]***	***-0.239 (0.090) [p=0.01]***
$< 84m$	***-0.221 (0.077) [p<0.01]***	**-0.191 (0.078) [p=0.02]**

UJIVE | **2SLS with $z_{j(i,t)} = \hat{p}_{i,ujive}$**

	UJIVE: x_{it} = year indicators	UJIVE: x_{it} = all controls	2SLS: x_{it} = year indicators	2SLS: x_{it} = all controls
$< 12m$	-0.331 (0.096) [p<0.01]	-0.352 (0.090) [p<0.01]	-0.331 (0.079) [p<0.01]	-0.352 (0.068) [p<0.01]
$< 24m$	-0.356 (0.122) [p<0.01]	-0.375 (0.122) [p<0.01]	-0.356 (0.109) [p<0.01]	-0.375 (0.093) [p<0.01]
$< 36m$	-0.342 (0.132) [p<0.01]	-0.348 (0.131) [p<0.01]	-0.342 (0.101) [p<0.01]	-0.348 (0.086) [p<0.01]
$< 48m$	-0.323 (0.154) [p=0.04]	-0.305 (0.164) [p=0.06]	**-0.323 (0.103) [p<0.01]**	**-0.305 (0.095) [p<0.01]**
$< 60m$	**-0.315 (0.161) [p=0.05]**	-0.287 (0.172) [p=0.10]	***-0.315 (0.113) [p<0.01]***	***-0.287 (0.109) [p<0.01]***
$< 72m$	**-0.329 (0.159) [p=0.04]**	-0.279 (0.173) [p=0.11]	***-0.329 (0.111) [p<0.01]***	***-0.279 (0.105) [p<0.01]***
$< 84m$	-0.268 (0.151) [p=0.08]	-0.223 (0.162) [p=0.17]	***-0.268 (0.095) [p<0.01]***	**-0.223 (0.091) [p=0.01]**

Notes: The two sets of results in the upper sub-panel come from running 2SLS models that employ the vector of judge assignment indicators as instruments for $\tau_{j(i,t)}$. The left column presents results given indicators for year of case assignment as the only control variables, x_{it}. The right column presents results given controls for our full set of case and defendant characteristics. The four sets of results in the lower sub-panel are biased-corrected versions of the 2SLS results in the upper sub-panel. The two columns on the left present UJIVE results. See Kolesar (2013). The two columns on the right employ the same $\hat{p}_{i,ujive}$ instrument employed in the UJIVE estimator, but instead of regressing recidivism on only a constant and a treatment indicator while using $\hat{p}_{i,ujive}$ as an instrument for treatment, these models add either year of assignment indicators or our full set of conditioning variables as controls in both the first and second stage equations. The absolute differences between the corresponding point estimates produced by the two procedures are all .001 or less. Kolesar (2013) does not propose an estimator for the variance of the UJIVE estimator. However, conventional standard errors for 2SLS are valid if we simply use Kolesar's $\hat{p}_{i,ujive}$ as our instrument for incarceration. We report HAC standard errors. For first offenders, we cluster at the judge level. For repeat offenders, we two-way cluster at the defendant*judge level. Entries in **bold** type are treatment impacts that are statistically different among first versus repeat offenders given $p = .1$, and entries in ***bold italics*** are different given $p = .05$.

Panel B continued: 2SLS w/ Vector of Judge Indicators as Instruments

REPEAT OFFENDERS

Baseline 2SLS with $z_{j(i,t)}$ = indicators for $j(i,t)$

	x_{it} = year indicators	x_{it} = all controls
$< 12m$	-0.210 (0.066) [p<0.01]	-0.224 (0.064) [p<0.01]
$< 24m$	-0.155 (0.084) [p=0.07]	-0.174 (0.078) [p=0.03]
$< 36m$	-0.161 (0.080) [p=0.05]	-0.181 (0.072) [p=0.01]
$< 48m$	**-0.068 (0.068) [p=0.32]**	**-0.091 (0.061) [p=0.14]**
$< 60m$	***0.013 (0.075) [p=0.86]***	***-0.005 (0.067) [p=0.94]***
$< 72m$	***0.020 (0.067) [p=0.77]***	***0.002 (0.062) [p=0.97]***
$< 84m$	***0.017 (0.068) [p=0.80]***	**0.000 (0.063) [p=1.00]**

UJIVE **2SLS with $z_{j(i,t)} = \hat{p}_{i,ujive}$**

	x_{it} = year indicators	x_{it} = all controls	x_{it} = year indicators	x_{it} = all controls
$< 12m$	-0.213 (0.087) [p=0.01]	-0.227 (0.086) [p<0.01]	-0.213 (0.075) [p<0.01]	-0.228 (0.072) [p< 0.01]
$< 24m$	-0.155 (0.112) [p=0.16]	-0.176 (0.110) [p=0.11]	-0.155 (0.096) [p=0.10]	-0.177 (0.088) [p=0.04]
$< 36m$	-0.173 (0.101) [p=0.09]	-0.194 (0.099) [p=0.05]	-0.173 (0.091) [p=0.06]	-0.194 (0.080) [p=0.02]
$< 48m$	-0.073 (0.086) [p=0.40]	-0.096 (0.087) [p=0.27]	**-0.073 (0.077) [p=0.34]**	**-0.096 (0.068) [p=0.16]**
$< 60m$	**0.017 (0.094) [p=0.85]**	-0.001 (0.095) [p=0.99]	***0.017 (0.085) [p=0.84]***	***-0.001 (0.076) [p=0.99]***
$< 72m$	**0.022 (0.085) [p=0.80]**	0.006 (0.088) [p=0.94]	***0.022 (0.077) [p=0.78]***	***0.006 (0.070) [p=0.93]***
$< 84m$	0.016 (0.087) [p=0.85]	0.001 (0.090) [p=0.99]	***0.016 (0.077) [p=0.83]***	**0.001 (0.071) [p=0.99]**

Notes: The two sets of results in the upper sub-panel come from running 2SLS models that employ the vector of judge assignment indicators as instruments for $\tau_{j(i,t)}$. The left column presents results given indicators for year of case assignment as the only control variables, x_{it}. The right column presents results given controls for our full set of case and defendant characteristics. The four sets of results in the lower sub-panel are biased-corrected versions of the 2SLS results in the upper sub-panel. The two columns on the left present UJIVE results. See Kolesar (2013). The two columns on the right employ the same $\hat{p}_{i,ujive}$ instrument employed in the UJIVE estimator, but instead of regressing recidivism on only a constant and a treatment indicator while using $\hat{p}_{i,ujive}$ as an instrument for treatment, these models add either year of assignment indicators or our full set of conditioning variables as controls in both the first and second stage equations. The absolute differences between the corresponding point estimates produced by the two procedures are all .001 or less. Kolesar (2013) does not propose an estimator for the variance of the UJIVE estimator. However, conventional standard errors for 2SLS are valid if we simply use Kolesar's $\hat{p}_{i,ujive}$ as our instrument for incarceration. We report HAC standard errors. For first offenders, we cluster at the judge level. For repeat offenders, we two-way cluster at the defendant*judge level. Entries in **bold** type are treatment impacts that are statistically different among first versus repeat offenders given $p = .1$, and entries in ***bold italics*** are different given $p = .05$.

10.5 Recidivism Outcome = Count of New Charges

Appendix Table 10.5

Impact of Incarceration on Total Future Charges

Panel A: First Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
Charge Count <12m	0.20	-0.154 (0.007) [p<0.01]	-0.360 (0.069) [p<0.01]	-0.436 (0.085) [p<0.01]
Charge Count <24m	0.36	-0.165 (0.012) [p<0.01]	**-0.507 (0.116) [p<0.01]**	***-0.615 (0.129) [p<0.01]***
Charge Count <36m	0.51	-0.142 (0.014) [p<0.01]	-0.538 (0.147) [p<0.01]	**-0.653 (0.163) [p<0.01]**
Charge Count <48m	0.64	-0.121 (0.016) [p<0.01]	**-0.597 (0.188) [p<0.01]**	***-0.724 (0.208) [p<0.01]***
Charge Count <60m	0.76	-0.116 (0.019) [p<0.01]	**-0.609 (0.209) [p<0.01]**	**-0.738 (0.231) [p<0.01]**
Charge Count <72m	0.86	-0.102 (0.022) [p<0.01]	-0.559 (0.206) [p<0.01]	**-0.677 (0.232) [p<0.01]**
Charge Count <84m	0.96	-0.084 (0.022) [p<0.01]	-0.503 (0.192) [p=0.01]	-0.609 (0.219) [p<0.01]

$\bar{\tau} = 0.19$, Standard Deviation of LOM: .028, F-Statistic: 284, N: 37,055

$f(l) : 0\,(81\%),\ (0, 12]\,(8\%),\ (12, 24]\,(4\%),\ (24, 36]\,(4\%),\ (36, 48]\,(2\%),\ (48, 60]\,(1\%),\ [60, \infty)\,(1\%)$

Panel B: Repeat Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
Charge Count <12m	0.15	-0.194 (0.007) [p<0.01]	-0.212 (0.075) [p<0.01]	-0.242 (0.083) [p<0.01]
Charge Count <24m	0.36	-0.188 (0.010) [p<0.01]	**-0.210 (0.103) [p=0.05]**	***-0.240 (0.116) [p=0.04]***
Charge Count <36m	0.56	-0.140 (0.013) [p<0.01]	-0.240 (0.125) [p=0.06]	**-0.274 (0.141) [p=0.05]**
Charge Count <48m	0.75	-0.108 (0.015) [p<0.01]	**-0.180 (0.137) [p=0.20]**	***-0.205 (0.155) [p=0.18]***
Charge Count <60m	0.92	-0.093 (0.016) [p<0.01]	**-0.134 (0.173) [p=0.44]**	**-0.153 (0.195) [p=0.43]**
Charge Count <72m	1.08	-0.060 (0.018) [p<0.01]	-0.092 (0.195) [p=0.64]	**-0.105 (0.219) [p=0.63]**
Charge Count <84m	1.22	-0.039 (0.020) [p=0.05]	-0.149 (0.227) [p=0.51]	-0.171 (0.256) [p=0.51]

$\bar{\tau} = 0.66$, Standard Deviation of LOM: .043, F-Statistic: 782, N: 33,526

$f(l) : 0\,(34\%),\ (0, 12]\,(35\%),\ (12, 24]\,(14\%),\ (24, 36]\,(10\%),\ (36, 48]\,(3\%),\ (48, 60]\,(2\%),\ [60, \infty)\,(2\%)$

Notes: These tables present results that parallel the main results we present in Figure 3. However, here the outcome is not any new charge but the count of new charges. Each panel reports results from seven OLS, RF, and 2SLS models. In the OLS and 2SLS models, each entry is the estimated coefficient on $\tau_{j(i,t)}$, which is an indicator that equals one if judge j assigns an incarceration sentence to defendant i at date t. In the RF column, each entry is the estimated coefficient on $z_{j(i,t)}$, the LOM severity measure associated with judge j. In each row, the outcome variable is the total number of new charges filed before a given horizon. The F-statistics are test statistics for the null that $z_{j(i,t)}$ does not predict $\tau_{j(i,t)}$ given our controls for case and defendant characteristics. We report HAC standard errors. For first offenders, we cluster at the judge level. For repeat offenders, we two-way cluster at the defendant*judge level. $\bar{\tau}$ gives the fraction of the sample that received an incarceration sentence. Entries in **bold** type are treatment impacts that are statistically different among first versus repeat offenders given $p = .1$, and entries in ***bold italics*** are different given $p = .05$. $f(l)$ is a discrete density that describes the distribution of expected incarceration time given the sentences assigned to defendants. Note that $f(0) = 1 - \bar{\tau}$ by definition.

10.6 Treatment = Expected Months Incarcerated

Appendix Table 10.6

Main Results Given
Treatment = Expected Months Incarcerated
Instrument = LOM of Incarceration Treatment

Panel A: First Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.19	-0.004 (0.000) [p<0.01]	-0.290 (0.059) [p<0.01]	-0.025 (0.007) [p<0.01]
New Charge <24m	0.31	-0.005 (0.000) [p<0.01]	-0.316 (0.084) [p<0.01]	-0.027 (0.010) [p<0.01]
New Charge <36m	0.38	-0.006 (0.000) [p<0.01]	-0.292 (0.075) [p<0.01]	-0.025 (0.008) [p<0.01]
New Charge <48m	0.44	-0.005 (0.000) [p<0.01]	-0.256 (0.083) [p<0.01]	-0.022 (0.009) [p=0.01]
New Charge <60m	0.47	-0.005 (0.000) [p<0.01]	***-0.245 (0.094) [p=0.01]***	**-0.021 (0.009) [p=0.02]**
New Charge <72m	0.50	-0.005 (0.000) [p<0.01]	***-0.241 (0.090) [p=0.01]***	***-0.021 (0.009) [p=0.02]***
New Charge <84m	0.52	-0.004 (0.000) [p<0.01]	**-0.191 (0.077) [p=0.02]**	**-0.017 (0.008) [p=0.03]**

$\bar{\tau} = 4.20$, Standard Deviation of LOM: .028, F-Statistic: 19, N: 37,055

$f(l) : 0\ (81\%), (0, 12]\ (8\%), (12, 24]\ (4\%), (24, 36]\ (4\%), (36, 48]\ (2\%), (48, 60]\ (1\%), [60, \infty)\ (1\%)$

Panel B: Repeat Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.14	-0.005 (0.000) [p<0.01]	-0.201 (0.067) [p<0.01]	-0.017 (0.006) [p<0.01]
New Charge <24m	0.31	-0.009 (0.000) [p<0.01]	-0.157 (0.080) [p=0.06]	-0.013 (0.006) [p=0.03]
New Charge <36m	0.44	-0.009 (0.000) [p<0.01]	-0.174 (0.073) [p=0.02]	-0.015 (0.006) [p<0.01]
New Charge <48m	0.52	-0.008 (0.000) [p<0.01]	-0.090 (0.062) [p=0.15]	-0.008 (0.005) [p=0.11]
New Charge <60m	0.58	-0.007 (0.000) [p<0.01]	***-0.006 (0.068) [p=0.93]***	**-0.000 (0.006) [p=0.93]**
New Charge <72m	0.62	-0.006 (0.000) [p<0.01]	***0.001 (0.063) [p=0.99]***	***0.000 (0.005) [p=0.99]***
New Charge <84m	0.65	-0.006 (0.000) [p<0.01]	**-0.002 (0.064) [p=0.97]**	**-0.000 (0.005) [p=0.97]**

$\bar{\tau} = 11.63$, Standard Deviation of LOM: .043, F-Statistic: 9, N: 33,526

$f(l) : 0\ (34\%), (0, 12]\ (35\%), (12, 24]\ (14\%), (24, 36]\ (10\%), (36, 48]\ (3\%), (48, 60]\ (2\%), [60, \infty)\ (2\%)$

Notes: These results parallel the main results we present in Figure 3. However, here the treatment, $\tau_{j(i)t}$, is not an indicator for incarceration but instead our estimate of how many months defendant i will spend incarcerated. These estimates are based on court records, IDOC admission dates, and IDOC release dates. $\bar{\tau}$ is the average expected months served. $\tau_{j(i)t} = 0$ for those not sentenced to incarceration. Each panel reports results from seven OLS, RF, and 2SLS models. In each row, the outcome variable is an indicator for the presence of at least one new charge before a given horizon. Among repeat offenders, the five, six, and seven-year treatment impacts round to zero because all are less than .0005 in absolute value. The F-statistics are test statistics for the null that $z_{j(i,t)}$ does not predict $\tau_{j(i,t)}$ given our controls for case and defendant characteristics. We report HAC standard errors. For first offenders, we cluster at the judge level. For repeat offenders, we two-way cluster at the defendant*judge level. Entries in **bold** type are treatment impacts that are statistically different among first versus repeat offenders given $p = .1$, and entries in ***bold italics*** are different given $p = .05$. $f(l)$ is a discrete density that describes the distribution of expected incarceration time given the sentences assigned to defendants.

Appendix Table 10.6

Main Results Given
Treatment = Expected Months Incarcerated
Instrument = LOM of Expected Months Incarcerated

Panel C: First Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.19	-0.004 (0.000) [p<0.01]	-0.011 (0.003) [p<0.01]	**-0.019 (0.005) [p<0.01]**
New Charge <24m	0.31	-0.005 (0.000) [p<0.01]	-0.012 (0.004) [p<0.01]	-0.021 (0.006) [p<0.01]
New Charge <36m	0.38	-0.006 (0.000) [p<0.01]	-0.011 (0.004) [p<0.01]	-0.019 (0.006) [p<0.01]
New Charge <48m	0.44	-0.005 (0.000) [p<0.01]	-0.011 (0.004) [p=0.02]	-0.019 (0.006) [p<0.01]
New Charge <60m	0.47	-0.005 (0.000) [p<0.01]	-0.011 (0.005) [p=0.02]	**-0.019 (0.007) [p<0.01]**
New Charge <72m	0.50	-0.005 (0.000) [p<0.01]	-0.011 (0.005) [p=0.02]	***-0.020 (0.007) [p<0.01]***
New Charge <84m	0.52	-0.004 (0.000) [p<0.01]	-0.009 (0.004) [p=0.02]	**-0.016 (0.006) [p<0.01]**

$\bar{\tau} = 4.20$, Standard Deviation of LOM: 0.576, F-Statistic: 35, N: 37,055

$f(l) : 0\,(81\%), (0, 12]\,(8\%), (12, 24]\,(4\%), (24, 36]\,(4\%), (36, 48]\,(2\%), (48, 60]\,(1\%), [60, \infty)\,(1\%)$

Panel D: Repeat Offenders

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.14	-0.005 (0.000) [p<0.01]	-0.009 (0.002) [p<0.01]	**-0.010 (0.002) [p<0.01]**
New Charge <24m	0.31	-0.009 (0.000) [p<0.01]	-0.013 (0.002) [p<0.01]	-0.015 (0.002) [p<0.01]
New Charge <36m	0.44	-0.009 (0.000) [p<0.01]	-0.012 (0.002) [p<0.01]	-0.014 (0.002) [p<0.01]
New Charge <48m	0.52	-0.008 (0.000) [p<0.01]	-0.009 (0.002) [p<0.01]	-0.011 (0.002) [p<0.01]
New Charge <60m	0.58	-0.007 (0.000) [p<0.01]	-0.006 (0.002) [p<0.01]	**-0.007 (0.002) [p<0.01]**
New Charge <72m	0.62	-0.006 (0.000) [p<0.01]	-0.004 (0.002) [p=0.04]	***-0.005 (0.002) [p=0.03]***
New Charge <84m	0.65	-0.006 (0.000) [p<0.01]	-0.004 (0.002) [p=0.07]	**-0.004 (0.002) [p=0.06]**

$\bar{\tau} = 11.63$, Standard Deviation of LOM: 1.265, F-Statistic: 647, N: 33,526

$f(l) : 0\,(34\%), (0, 12]\,(35\%), (12, 24]\,(14\%), (24, 36]\,(10\%), (36, 48]\,(3\%), (48, 60]\,(2\%), [60, \infty)\,(2\%)$

Notes: See notes for Panels A and B above. These results parallel those presented in Panels A and B above, but here $z_{j(i,t)}$ is not the LOM of incarceration treatment, $\tau_{j(i)t}$, but the LOM of expected months of incarceration.

10.7 Heterogeneity in Treatment Effects

Panel A: First Offenders: Blacks

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.22	-0.167 (0.008) [p<0.01]	-0.249 (0.071) [p<0.01]	-0.317 (0.086) [p<0.01]
New Charge <24m	0.35	-0.144 (0.012) [p<0.01]	-0.256 (0.082) [p<0.01]	-0.327 (0.094) [p<0.01]
New Charge <36m	0.44	-0.107 (0.010) [p<0.01]	-0.264 (0.074) [p<0.01]	-0.336 (0.088) [p<0.01]
New Charge <48m	0.50	-0.076 (0.010) [p<0.01]	**-0.262 (0.073) [p<0.01]**	***-0.333 (0.084) [p<0.01]***
New Charge <60m	0.54	-0.057 (0.011) [p<0.01]	**-0.227 (0.082) [p<0.01]**	***-0.290 (0.099) [p<0.01]***
New Charge <72m	0.57	-0.043 (0.011) [p<0.01]	***-0.204 (0.078) [p=0.01]***	***-0.259 (0.097) [p<0.01]***
New Charge <84m	0.59	-0.032 (0.010) [p<0.01]	-0.125 (0.077) [p=0.11]	-0.160 (0.094) [p=0.09]

$\bar{\tau} = 0.20$, Standard Deviation of LOM: .032, F-Statistic: 157, N: 25,223
$f(l)$: 0 (80%), $(0, 12]$ (8%), $(12, 24]$ (4%), $(24, 36]$ (4%), $(36, 48]$ (2%), $(48, 60]$ (1%), $[60, \infty)$ (1%)

Panel B: Repeat Offenders: Blacks

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.14	-0.192 (0.006) [p<0.01]	-0.189 (0.066) [p<0.01]	-0.219 (0.074) [p<0.01]
New Charge <24m	0.32	-0.154 (0.008) [p<0.01]	-0.134 (0.087) [p=0.13]	-0.156 (0.098) [p=0.11]
New Charge <36m	0.45	-0.094 (0.008) [p<0.01]	-0.163 (0.083) [p=0.06]	-0.189 (0.094) [p=0.04]
New Charge <48m	0.54	-0.056 (0.009) [p<0.01]	**-0.078 (0.069) [p=0.26]**	***-0.091 (0.079) [p=0.25]***
New Charge <60m	0.60	-0.039 (0.009) [p<0.01]	**-0.007 (0.077) [p=0.92]**	***-0.009 (0.088) [p=0.92]***
New Charge <72m	0.64	-0.023 (0.008) [p<0.01]	***0.011 (0.072) [p=0.88]***	***0.012 (0.083) [p=0.88]***
New Charge <84m	0.67	-0.013 (0.008) [p=0.10]	0.011 (0.075) [p=0.89]	0.013 (0.086) [p=0.88]

$\bar{\tau} = 0.67$, Standard Deviation of LOM: .045, F-Statistic: 564, N: 28,087
$f(l)$: 0 (33%), $(0, 12]$ (37%), $(12, 24]$ (14%), $(24, 36]$ (10%), $(36, 48]$ (3%), $(48, 60]$ (2%), $[60, \infty)$ (2%)

Notes: These panels present results that parallel our main results, see Appendix 10.3. However, here we restrict both our first and repeat offender samples to Black defendants. We employ LOM measures of judge severity that are specific to first and repeat offender samples that contain only Black defendants.

Panel C: First Offenders: Drug Crimes

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.24	-0.158 (0.009) [p<0.01]	-0.200 (0.089) [p=0.03]	-0.300 (0.117) [p=0.01]
New Charge <24m	0.37	-0.126 (0.012) [p<0.01]	-0.207 (0.127) [p=0.11]	-0.310 (0.172) [p=0.07]
New Charge <36m	0.45	-0.113 (0.011) [p<0.01]	-0.285 (0.117) [p=0.02]	-0.427 (0.157) [p<0.01]
New Charge <48m	0.50	-0.091 (0.012) [p<0.01]	-0.254 (0.132) [p=0.06]	-0.380 (0.184) [p=0.04]
New Charge <60m	0.53	-0.079 (0.012) [p<0.01]	-0.155 (0.153) [p=0.32]	-0.231 (0.217) [p=0.29]
New Charge <72m	0.56	-0.075 (0.011) [p<0.01]	-0.130 (0.135) [p=0.34]	-0.195 (0.194) [p=0.32]
New Charge <84m	0.58	-0.068 (0.011) [p<0.01]	-0.075 (0.134) [p=0.58]	-0.113 (0.194) [p=0.56]

$\bar{\tau} = 0.13$, Standard Deviation of LOM: .036, F-Statistic: 57, N: 15,542
$f(l)$: 0 (87%), $(0, 12]$ (6%), $(12, 24]$ (3%), $(24, 36]$ (2%), $(36, 48]$ (1%), $(48, 60]$ (0%), $[60, \infty)$ (1%)

Panel D: Repeat Offenders: Drug Crimes

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.16	-0.174 (0.008) [p<0.01]	-0.180 (0.073) [p=0.02]	-0.212 (0.081) [p<0.01]
New Charge <24m	0.35	-0.129 (0.011) [p<0.01]	-0.074 (0.085) [p=0.38]	-0.088 (0.096) [p=0.36]
New Charge <36m	0.48	-0.074 (0.011) [p<0.01]	-0.121 (0.071) [p=0.09]	-0.143 (0.080) [p=0.08]
New Charge <48m	0.56	-0.038 (0.011) [p<0.01]	-0.024 (0.071) [p=0.73]	-0.028 (0.081) [p=0.73]
New Charge <60m	0.61	-0.025 (0.011) [p=0.02]	0.028 (0.066) [p=0.68]	0.033 (0.077) [p=0.67]
New Charge <72m	0.65	-0.013 (0.009) [p=0.19]	0.043 (0.068) [p=0.53]	0.051 (0.080) [p=0.52]
New Charge <84m	0.68	-0.007 (0.009) [p=0.44]	0.049 (0.069) [p=0.48]	0.057 (0.081) [p=0.48]

$\bar{\tau} = 0.62$, Standard Deviation of LOM: .059, F-Statistic: 312, N: 15,557
$f(l)$: 0 (38%), $(0, 12]$ (39%), $(12, 24]$ (12%), $(24, 36]$ (8%), $(36, 48]$ (2%), $(48, 60]$ (1%), $[60, \infty)$ (1%)

Notes: These panels present results that parallel our main results, see Appendix 10.3. However, here we restrict both our first and repeat offender samples to defendants facing a lead charge that involves a drug crime. We employ LOM measures of judge severity that are specific to first and repeat offender samples that contain only drug cases.

Panel E: First Offenders: Non-Drug Crimes

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.16	-0.138 (0.008) [p<0.01]	-0.209 (0.059) [p<0.01]	-0.310 (0.095) [p<0.01]
New Charge <24m	0.26	-0.121 (0.013) [p<0.01]	-0.194 (0.083) [p=0.02]	-0.288 (0.131) [p=0.03]
New Charge <36m	0.34	-0.075 (0.012) [p<0.01]	-0.164 (0.086) [p=0.06]	-0.244 (0.135) [p=0.07]
New Charge <48m	0.39	-0.045 (0.011) [p<0.01]	-0.157 (0.085) [p=0.07]	-0.233 (0.132) [p=0.08]
New Charge <60m	0.43	-0.026 (0.012) [p=0.04]	-0.206 (0.078) [p=0.01]	-0.305 (0.126) [p=0.02]
New Charge <72m	0.45	-0.010 (0.012) [p=0.42]	-0.241 (0.082) [p<0.01]	-0.358 (0.133) [p<0.01]
New Charge <84m	0.47	-0.000 (0.012) [p=0.99]	-0.233 (0.070) [p<0.01]	-0.346 (0.117) [p<0.01]

$\bar{\tau} = 0.23$, Standard Deviation of LOM: .031, F-Statistic: 91, N: 21,513

$f(l) : 0\,(77\%),\ (0, 12]\,(9\%),\ (12, 24]\,(5\%),\ (24, 36]\,(5\%),\ (36, 48]\,(2\%),\ (48, 60]\,(1\%),\ [60, \infty)\,(1\%)$

Panel F: Repeat Offenders: Non-Drug Crimes

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.12	-0.203 (0.010) [p<0.01]	-0.174 (0.061) [p<0.01]	-0.229 (0.073) [p<0.01]
New Charge <24m	0.28	-0.169 (0.010) [p<0.01]	-0.268 (0.069) [p<0.01]	-0.353 (0.081) [p<0.01]
New Charge <36m	0.40	-0.108 (0.011) [p<0.01]	-0.265 (0.073) [p<0.01]	-0.349 (0.092) [p<0.01]
New Charge <48m	0.49	-0.069 (0.011) [p<0.01]	-0.213 (0.075) [p<0.01]	-0.281 (0.103) [p<0.01]
New Charge <60m	0.55	-0.045 (0.010) [p<0.01]	-0.064 (0.074) [p=0.39]	-0.084 (0.098) [p=0.39]
New Charge <72m	0.60	-0.025 (0.010) [p=0.01]	-0.083 (0.064) [p=0.20]	-0.109 (0.084) [p=0.19]
New Charge <84m	0.63	-0.012 (0.010) [p=0.24]	-0.093 (0.064) [p=0.15]	-0.122 (0.085) [p=0.15]

$\bar{\tau} = 0.70$, Standard Deviation of LOM: .042, F-Statistic: 132, N: 17,969

$f(l) : 0\,(30\%),\ (0, 12]\,(32\%),\ (12, 24]\,(15\%),\ (24, 36]\,(12\%),\ (36, 48]\,(5\%),\ (48, 60]\,(3\%),\ [60, \infty)\,(4\%)$

Notes: These panels present results that parallel our main results, see Appendix 10.3. However, here we restrict both our first and repeat offender samples to defendants who are not facing a lead charge that involves a drug crime. We employ LOM measures of judge severity that are specific to first and repeat offender samples that contain only non-drug cases.

Panel G: First Offenders: High-Crime Residence

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.23	-0.164 (0.009) [p<0.01]	-0.223 (0.064) [p<0.01]	-0.285 (0.079) [p<0.01]
New Charge <24m	0.37	-0.139 (0.015) [p<0.01]	-0.248 (0.086) [p<0.01]	-0.317 (0.099) [p<0.01]
New Charge <36m	0.46	-0.105 (0.013) [p<0.01]	-0.251 (0.079) [p<0.01]	-0.321 (0.091) [p<0.01]
New Charge <48m	0.51	-0.071 (0.012) [p<0.01]	-0.236 (0.076) [p<0.01]	-0.302 (0.086) [p<0.01]
New Charge <60m	0.55	-0.055 (0.012) [p<0.01]	-0.209 (0.087) [p=0.02]	-0.267 (0.103) [p<0.01]
New Charge <72m	0.58	-0.041 (0.011) [p<0.01]	-0.185 (0.083) [p=0.03]	-0.237 (0.100) [p=0.02]
New Charge <84m	0.60	-0.030 (0.010) [p<0.01]	-0.119 (0.084) [p=0.16]	-0.152 (0.101) [p=0.13]

$\bar{\tau} = 0.20$, Standard Deviation of LOM: .035, F-Statistic: 165, N: 20,605

$f(l): 0$ (80%), $(0, 12]$ (8%), $(12, 24]$ (4%), $(24, 36]$ (4%), $(36, 48]$ (2%), $(48, 60]$ (1%), $[60, \infty)$ (1%)

Panel H: Repeat Offenders: High-Crime Residence

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.14	-0.187 (0.007) [p<0.01]	-0.197 (0.067) [p<0.01]	-0.236 (0.078) [p<0.01]
New Charge <24m	0.32	-0.154 (0.009) [p<0.01]	-0.120 (0.096) [p=0.22]	-0.143 (0.113) [p=0.20]
New Charge <36m	0.45	-0.095 (0.008) [p<0.01]	-0.162 (0.090) [p=0.08]	-0.194 (0.106) [p=0.07]
New Charge <48m	0.54	-0.053 (0.008) [p<0.01]	-0.087 (0.075) [p=0.26]	-0.104 (0.090) [p=0.25]
New Charge <60m	0.60	-0.037 (0.008) [p<0.01]	-0.024 (0.081) [p=0.77]	-0.029 (0.096) [p=0.76]
New Charge <72m	0.64	-0.018 (0.008) [p=0.03]	0.020 (0.072) [p=0.78]	0.024 (0.085) [p=0.78]
New Charge <84m	0.67	-0.010 (0.008) [p=0.21]	0.024 (0.068) [p=0.72]	0.029 (0.080) [p=0.72]

$\bar{\tau} = 0.68$, Standard Deviation of LOM: .045, F-Statistic: 365, N: 23,833

$f(l): 0$ (32%), $(0, 12]$ (37%), $(12, 24]$ (14%), $(24, 36]$ (10%), $(36, 48]$ (3%), $(48, 60]$ (2%), $[60, \infty)$ (2%)

Notes: These panels present results that parallel our main results, see Appendix 10.3. However, here we restrict both our first and repeat offender samples to defendants who reside in one of 25 community areas that we designate as high-crime areas. See Appendix 12 for a discussion of these designations. We employ LOM measures of judge severity that are specific to first and repeat offender samples that contain only defendants who reside in these communities.

Panel I: First Offenders: Not High-Crime Residence

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.15	-0.121 (0.006) [p<0.01]	-0.297 (0.066) [p<0.01]	-0.558 (0.126) [p<0.01]
New Charge <24m	0.24	-0.103 (0.011) [p<0.01]	-0.333 (0.097) [p<0.01]	-0.625 (0.200) [p<0.01]
New Charge <36m	0.30	-0.072 (0.011) [p<0.01]	-0.245 (0.105) [p=0.02]	-0.459 (0.200) [p=0.02]
New Charge <48m	0.34	-0.053 (0.010) [p<0.01]	-0.201 (0.112) [p=0.08]	-0.377 (0.209) [p=0.07]
New Charge <60m	0.37	-0.037 (0.011) [p<0.01]	-0.256 (0.112) [p=0.03]	-0.481 (0.211) [p=0.02]
New Charge <72m	0.40	-0.027 (0.011) [p=0.02]	-0.303 (0.108) [p<0.01]	-0.569 (0.202) [p<0.01]
New Charge <84m	0.42	-0.021 (0.011) [p=0.07]	-0.312 (0.103) [p<0.01]	-0.586 (0.203) [p<0.01]

$\bar{\tau} = 0.19$, Standard Deviation of LOM: .027, F-Statistic: 33, N: 16,450

$f(l) : 0$ (81%), $(0, 12]$ (8%), $(12, 24]$ (4%), $(24, 36]$ (4%), $(36, 48]$ (2%), $(48, 60]$ (1%), $[60, \infty)$ (1%)

Panel J: Repeat Offenders: Not High-Crime Residence

Y_s	$\bar{Y}_s$	OLS	RF	2SLS
New Charge <12m	0.13	-0.188 (0.010) [p<0.01]	-0.153 (0.068) [p=0.03]	-0.215 (0.081) [p<0.01]
New Charge <24m	0.27	-0.132 (0.013) [p<0.01]	-0.145 (0.079) [p=0.07]	-0.204 (0.104) [p=0.05]
New Charge <36m	0.39	-0.076 (0.015) [p<0.01]	-0.142 (0.090) [p=0.12]	-0.200 (0.117) [p=0.09]
New Charge <48m	0.47	-0.048 (0.014) [p<0.01]	-0.065 (0.082) [p=0.43]	-0.092 (0.110) [p=0.41]
New Charge <60m	0.53	-0.026 (0.014) [p=0.08]	0.028 (0.073) [p=0.70]	0.039 (0.102) [p=0.70]
New Charge <72m	0.57	-0.015 (0.013) [p=0.26]	0.024 (0.073) [p=0.74]	0.034 (0.102) [p=0.74]
New Charge <84m	0.60	-0.003 (0.012) [p=0.84]	-0.005 (0.075) [p=0.95]	-0.007 (0.103) [p=0.95]

$\bar{\tau} = 0.66$, Standard Deviation of LOM: .054, F-Statistic: 118, N: 9,693

$f(l) : 0$ (34%), $(0, 12]$ (35%), $(12, 24]$ (14%), $(24, 36]$ (10%), $(36, 48]$ (3%), $(48, 60]$ (2%), $[60, \infty)$ (2%)

Notes: These panels present results that parallel our main results, see Appendix 10.3. However, here we restrict both our first and repeat offender samples to defendants who do not reside in one of 25 community areas that we designate as high-crime areas. See Appendix 12 for a discussion of these designations. We employ LOM measures of judge severity that are specific to first and repeat offender samples that contain only defendants who reside outside these 25 high-crime communities.

10.8 Monotonicity Tests

Appendix Table 10.8

Panel A: Sample Divided By Incarceration Propensity

Group	Quartile	FS Coeff	FS SE	FS p-value
First Offenders	1st	.264	(.065)	<0.001
	2nd	.678	(.098)	<0.001
	3rd	1.385	(.141)	<0.001
	4th	.946	(.171)	<0.001
Repeat Offenders	1st	.865	(.115)	<0.001
	2nd	1.01	(.122)	<0.001
	3rd	.947	(.106)	<0.001
	4th	.705	(.092)	<0.001

Panel B: Five Subpopulations

Group	Subsample	FS Coeff	FS SE	FS p-value	Corr. with Group LOM
First Offenders	Black	.911	(.077)	<0.001	.93
	Drug Charge	.881	(.092)	<0.001	.728
	High Crime Area	1.004	(.085)	<0.001	.917
	Not High Crime Area	.596	(.092)	<0.001	.804
	Nondrug Charge	.763	(.085)	<0.001	.821
Repeat Offenders	Black	.895	(.06)	<0.001	.976
	Drug Charge	1.09	(.082)	<0.001	.852
	High Crime Area	.864	(.065)	<0.001	.948
	Not High Crime Area	.904	(.102)	<0.001	.78
	Nondrug Charge	.691	(.073)	<0.001	.855

Notes: These tables examine how our LOM measure of judge severity varies among subsamples of of first and repeat offenders respectively. In Panel A, we first predict incarceration using our standard controls and assignment to a reference judge with $z_{j^*(i,t)} = 0 \ \forall i, t$. We form these predictions separately for first versus repeat offenders, and then within each sample, we divide defendants into quartiles based on their predicted incarceration rate. Next, within these quartiles, we estimate the first-stage relationship between the leave-out-mean and actual assignment to incarceration. The columns in Panel A describe the coefficient, standard error, and p-value from tests of significance specific to these subsample first-stages. Panel B repeats this exercise, but instead of dividing groups by quartiles of predicted incarceration, we create subsamples using demographic information. In Panel B, we also calculate the average LOM for each judge within each subpopulation. The final column presents the judge-level correlation between those subpopulation-specific LOMS and our overall LOMs. See Frandsen et al. (2023) for more on these tests.

10.9 2SLS: Multiple Treatment Impacts

Appendix Table 10.9
Three Estimated Treatment Impacts - Relative to Probation

Sample	Horizon	Incar 2SLS	SE	No Conv 2SLS	SE	Other 2SLS	SE
First Offenders	12	-.35	(.068)	-.053	(.119)	-.031	(.023)
First Offenders	24	-.386	(.092)	-.109	(.191)	-.028	(.038)
First Offenders	36	-.349	(.088)	.012	(.185)	-.028	(.034)
First Offenders	48	-.314	(.094)	-.075	(.191)	-.015	(.038)
First Offenders	60	-.304	(.106)	-.137	(.215)	-.023	(.039)
First Offenders	72	-.302	(.098)	-.193	(.209)	-.029	(.039)
First Offenders	84	-.243	(.078)	-.231	(.186)	-.041	(.033)
Repeat Offenders	12	-.218	(.074)	.135	(.119)	-.072	(.109)
Repeat Offenders	24	-.154	(.088)	.182	(.165)	-.018	(.125)
Repeat Offenders	36	-.167	(.076)	.209	(.151)	.014	(.12)
Repeat Offenders	48	-.089	(.07)	.108	(.145)	-.018	(.107)
Repeat Offenders	60	.008	(.08)	.138	(.131)	-.052	(.088)
Repeat Offenders	72	.015	(.075)	.126	(.125)	-.038	(.076)
Repeat Offenders	84	.008	(.077)	.109	(.119)	-.05	(.075)

Note: This table presents 2SLS estimates of the impacts of incarceration, acquittal, and other case outcomes on recidivism (relative to probation sentences). Other outcomes include convictions after charges are reduced to misdemeanors and prison sentences that require no time-served in prison given credits awarded for jail time prior to a verdict. Each row presents the impacts at a different recidivism horizon. We present separate results for first and repeat offenders. Each regression also includes, x_{it}, the full set of case and defendant characteristics that we include in our baseline models. We report HAC standard errors. For first offenders, we cluster at the judge level. For repeat offenders, we two-way cluster at the judge-defendant level.

10.10 Predicted Treatment Realization In Subsamples

10.10.1 Predicted Treatments: Black Subsample

Appendix Table 10.10.1
Treatments in Black Subsample on Full-Sample Predicted Treatments

	Incarceration	Acquittal	Other	Incarceration	Acquittal	Other
Pred-Incarceration	1.093***	-0.0169	-0.0491	1.019***	0.00622	-0.0180
	(0.0953)	(0.0666)	(0.0532)	(0.0506)	(0.0234)	(0.0210)
Pred-Acquittal	-0.0991	0.954***	0.0851	-0.0279	1.013***	-0.0163
	(0.104)	(0.0972)	(0.0646)	(0.0730)	(0.0610)	(0.0327)
Pred-Other Sentence	0.0198	-0.0210	1.030***	0.00444	0.0127	0.954***
	(0.0183)	(0.0131)	(0.0113)	(0.0482)	(0.0466)	(0.0310)
Observations	25223	25223	25223	28087	28087	28087
Adjusted R^2	0.290	0.040	0.090	0.161	0.040	0.034
Sample	First	First	First	Repeat	Repeat	Repeat
Treatment Mean	0.20	0.09	0.10	0.67	0.11	0.05

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents results from first and repeat-offender regressions that, within the subsamples of Black defendants, predict one of three case outcomes (incarceration, acquittal, and other outcome - probation is the excluded category) using the three predicted case outcomes calculated in the full samples. Other outcomes include convictions after charges are reduced to misdemeanors and prison sentences that require no time-served in prison given credits awarded for jail time prior to a verdict. Each column predicts a different one of these treatments, and we run these three regressions separately for first and repeat offenders. Each regression also includes, x_{it}, the full set of case and defendant characteristics that we include in our baseline models. We report HAC standard errors. For first offenders, we cluster at the judge level. For repeat offenders, we two-way cluster at the judge-defendant level. These results are the first of five sets of results. Bhuller and Sigstad (2023) argue that if, within homogeneous subsamples, full-sample predicted values of a given treatment predict the treatment in question but not other treatments, the results are consistent with the hypothesis that each of the 2SLS treatment effects produced by our three-treatment model is a proper weighted average of the recidivism impacts of assigning defendants to one specific treatment instead of probation. We implement these tests on the same five subsamples that we used to conduct monotonicity tests in the previous section: Black defendants, high-crime neighborhoods, not high-crime neighborhoods, drug cases, non-drug cases. See Bhuller and Sigstad (2023) for more on these diagnostic tests.

10.10.2 Predicted Treatments: Drug Subsample

Appendix Table 10.10.2
Treatments in Drug Subsample on Full-Sample Predicted Treatments

	Incarceration	Acquittal	Other	Incarceration	Acquittal	Other
Pred-Incarceration	1.037***	0.171*	-0.00734	1.222***	0.0295	0.0646
	(0.147)	(0.0655)	(0.109)	(0.166)	(0.0587)	(0.0482)
Pred-Acquittal	-0.259	1.167***	0.0458	0.0119	1.150***	-0.0975
	(0.207)	(0.128)	(0.207)	(0.165)	(0.115)	(0.0616)
Pred-Other Sentence	0.0248	-0.00717	1.167***	-0.237	0.0995	1.316***
	(0.0239)	(0.0173)	(0.0948)	(0.138)	(0.0737)	(0.110)
Observations	15542	15542	15542	15557	15557	15557
Adjusted R^2	0.123	0.050	0.116	0.185	0.035	0.043
Sample	First	First	First	Repeat	Repeat	Repeat
Treatment Mean	0.13	0.10	0.08	0.62	0.11	0.05

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents results from first and repeat-offender regressions that, within the subsamples of defendants charged with drug crimes, predict one of three case outcomes (incarceration, acquittal, and other outcome - probation is the excluded category) using the three predicted case outcomes calculated in the full samples. See note below Table 10.10.1 for more details.

10.10.3 Predicted Treatments: Non-Drug Subsample

Appendix Table 10.10.3
Treatments in Non-Drug Subsample on Full-Sample Predicted Treatments

	Incarceration	Acquittal	Other	Incarceration	Acquittal	Other
Pred-Incarceration	0.943***	-0.119*	0.0114	0.803***	-0.0285	-0.0531
	(0.105)	(0.0580)	(0.101)	(0.149)	(0.0534)	(0.0399)
Pred-Acquittal	0.209	0.858***	-0.00946	-0.0554	0.874***	0.104
	(0.195)	(0.0912)	(0.173)	(0.137)	(0.0909)	(0.0681)
Pred-Other Sentence	-0.0265	0.00965	0.876***	0.203	-0.0811	0.716***
	(0.0260)	(0.0208)	(0.0565)	(0.152)	(0.0628)	(0.0760)
Observations	21513	21513	21513	17969	17969	17969
Adjusted R^2	0.347	0.042	0.067	0.157	0.050	0.031
Sample	First	First	First	Repeat	Repeat	Repeat
Treatment Mean	0.23	0.10	0.11	0.70	0.10	0.05

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents results from first and repeat-offender regressions that, within the subsamples of defendants not charged with drug crimes, predict one of three case outcomes (incarceration, acquittal, and other sentence - probation is the excluded category) using the three predicted case outcomes calculated in the full samples. See note below Table 10.10.1 for more details.

10.10.4 Predicted Treatments: High-Crime Area Subsample

Appendix Table 10.10.4
Treatments in High-Crime Area Subsample
on Full-Sample Predicted Treatments

	Incarceration	Acquittal	Other	Incarceration	Acquittal	Other
Pred-Incarceration	1.212***	-0.0585	-0.0511	0.994***	0.0278	-0.0145
	(0.106)	(0.0711)	(0.0634)	(0.0635)	(0.0361)	(0.0209)
Pred-Acquittal	-0.0103	1.006***	0.186	0.0268	0.959***	0.000961
	(0.112)	(0.115)	(0.0954)	(0.110)	(0.0701)	(0.0482)
Pred-Other Sentence	0.0149	0.00694	1.021***	0.0177	0.0111	0.956***
	(0.0208)	(0.0136)	(0.0210)	(0.0546)	(0.0610)	(0.0442)
Observations	20605	20605	20605	23833	23833	23833
Adjusted R^2	0.289	0.034	0.088	0.157	0.037	0.035
Sample	First	First	First	Repeat	Repeat	Repeat
Treatment Mean	0.20	0.09	0.10	0.68	0.11	0.05

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents results from first and repeat-offenders regressions that, within the subsamples of defendants from high-crime areas, predict one of three case outcomes (incarceration, acquittal, and other sentence - probation is the excluded category) using the three predicted case outcomes calculated in the full samples. See note below Table 10.10.1 for more details.

10.10.5 Predicted Treatments: Not High-Crime Area Subsample

Appendix Table 10.10.5
Treatments in Not High-Crime Area Subsample
on Full-Sample Predicted Treatments

	Incarceration	Acquittal	Other	Incarceration	Acquittal	Other
Pred-Incarceration	0.727***	0.0660	0.0776	1.015***	-0.0631	0.0209
	(0.105)	(0.0928)	(0.0728)	(0.125)	(0.0671)	(0.0511)
Pred-Acquittal	0.00979	0.987***	-0.224*	-0.0702	1.119***	-0.00921
	(0.133)	(0.199)	(0.110)	(0.221)	(0.127)	(0.0890)
Pred-Other Sentence	-0.0234	-0.0115	0.972***	-0.0320	-0.0383	1.109***
	(0.0254)	(0.0212)	(0.0469)	(0.120)	(0.0894)	(0.0938)
Observations	16450	16450	16450	9693	9693	9693
Adjusted R^2	0.280	0.042	0.082	0.203	0.051	0.035
Sample	First	First	First	Repeat	Repeat	Repeat
Treatment Mean	0.18	0.12	0.09	0.63	0.11	0.05

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: This table presents results from first and repeat-offender regressions that, within the subsamples of defendants who are not from high-crime areas, predict one of three case outcomes (incarceration, acquittal, and other sentence - probation is the excluded category) using the three predicted case outcomes calculated in the full samples. See note below Table 10.10.1 for more details.

10.11 Characteristics of Compliers

10.11.1 Expected Recidivism Rates Given No Incarceration

Appendix Table 10.11.1
Expected Recidivism Rates given $\tau_{j(i,t)} = 0$
All Non-Incarcerated, Never Takers, and Compliers

Panel A: First Offenders

Horizon	E[Y(0)|τ=0]	E[Y(0)|NT=1]	E[Y(0)|C=1]
6 months	.13	.12	.19
12 months	.22	.21	.41
18 months	.29	.27	.53
24 months	.34	.32	.58
30 months	.38	.35	.64
36 months	.41	.39	.65
42 months	.43	.41	.67
48 months	.45	.43	.70
54 months	.47	.45	.72
60 months	.49	.46	.72
66 months	.50	.48	.73
72 months	.51	.49	.76
78 months	.52	.50	.74
84 months	.53	.51	.74

The three columns in each panel present the expected values of our recidivism indicators given a non-incarceration sentence for a specific subsample of first offenders. The thee subsamples are: all offenders not sentenced to incarceration, never takers, and compliers. We use the linear extrapolation method presented in Dahl et al. (2014) to create the estimates in the final two columns. Extrapolation is required because not all judges handle cases in all years. We extrapolate to estimate how many defendants the most and least severe judges would have sentenced to incarceration given a random sample of cases drawn from all years, and what the recidivism rates would be among those given non-incarceration sentences by the most and least severe judges. Among first offenders, $P(NT) = .736$, $P(C) = .124$, $P(AT) = .140$.

Appendix Table 10.11.1
Expected Recidivism Rates given $\tau_{j(i,t)} = 0$
All Non-Incarcerated, Never Takers, and Compliers

Panel B: Repeat Offenders

Horizon	E[Y(0)|τ=0]	E[Y(0)|NT=1]	E[Y(0)|C=1]
6 months	.14	.14	.13
12 months	.26	.25	.30
18 months	.34	.34	.36
24 months	.40	.39	.45
30 months	.45	.44	.50
36 months	.49	.48	.54
42 months	.52	.51	.56
48 months	.54	.54	.58
54 months	.57	.56	.59
60 months	.59	.58	.60
66 months	.60	.60	.62
72 months	.61	.61	.64
78 months	.63	.62	.66
84 months	.63	.63	.66

The three columns in each panel present the expected values of our recidivism indicators given a non-incarceration sentence for a specific subsample of repeat offenders. The thee subsamples are: all offenders not sentenced to incarceration, never takers, and compliers. We use the linear extrapolation method presented in Dahl et al. (2014) to create the estimates in the final two columns. Extrapolation is required because not all judges handle cases in all years. We extrapolate to estimate how many defendants the most and least severe judges would have sentenced to incarceration given a random sample of cases drawn from all years, and what the recidivism rates would be among those given non-incarceration sentences by the most and least severe judges. Among repeat offenders, $P(NT) = .261$, $P(C) = .221$, $P(AT) = .518$.

10.11.2 Distribution of Time-Served Given Incarceration

Appendix Table 10.11.2
Expected Future Rates of Incarceration given $\tau_{j(i,t)} = 1$
All Incarcerated, Always Takers, and Compliers

Panel A: First Offenders

Horizon	$E[I_s \mid \tau{=}1]$	$E[I_s \mid AT = 1, \tau = 1]$	$E[I_s \mid C = 1, \tau = 1]$
6 months	.65	.72	.36
12 months	.51	.57	.26
18 months	.41	.46	.17
24 months	.32	.36	.13
30 months	.22	.24	.13
36 months	.17	.18	.09
42 months	.13	.14	.08
48 months	.10	.11	.08
54 months	.08	.08	.08
60 months	.07	.06	.09
66 months	.06	.06	.07
72 months	.05	.04	.07
78 months	.04	.04	.07
84 months	.04	.03	.06

Notes: The three columns present fractions of first offenders sentenced to incarceration who remain incarcerated at different horizons. The three columns present results for three subsamples: all offenders sentenced to incarceration, always takers, and compliers. As in Table 10.11.1, we use the linear extrapolation method presented in Dahl et al. (2014) to create the estimates in the final two columns. Extrapolation is required for reasons that parallel those discussed in the notes to Table 10.11.1. $P(NT) = .736$, $P(C) = .124$, $P(AT) = .140$.

Appendix Table 10.11.2
Expected Future Rates of Incarceration given $\tau_{j(i,t)} = 1$
All Incarcerated, Always Takers, and Compliers

Panel B: Repeat Offenders

Horizon	E$[I_s \vert \tau{=}1]$	E$[I_s \vert AT = 1, \tau = 1]$	E$[I_s \vert C = 1, \tau = 1]$
6 months	.71	.74	.59
12 months	.48	.50	.40
18 months	.34	.34	.32
24 months	.25	.25	.25
30 months	.16	.16	.18
36 months	.12	.12	.10
42 months	.09	.09	.07
48 months	.07	.07	.05
54 months	.05	.06	.04
60 months	.04	.05	.03
66 months	.03	.04	.03
72 months	.03	.03	.03
78 months	.02	.02	.03
84 months	.02	.02	.02

Notes: The three columns present fractions of repeat offenders sentenced to incarceration who remain incarcerated at different horizons. The three columns present results for three subsamples: all offenders sentenced to incarceration, always takers, and compliers. As in Table 10.11.1, we use the linear extrapolation method presented in Dahl et al. (2014) to create the estimates in the final two columns. Extrapolation is required for reasons that parallel those discussed in the notes to Table 10.11.1. $P(NT) = .261$, $P(C) = .221$, $P(AT) = .518$.

11 Theory Appendix: Impacts of Incarceration

Consider two populations of offenders who are the same age, face the same charge, are demographically similar, and have comparable criminal histories. Further, assume that a court randomly assigns incarceration for $m > 0$ periods to one population while assigning no incarceration to the other group.

We use this thought experiment to show that three mechanisms determine the impacts of incarceration on recidivism. The first impact is incapacitation. Here, we assume that prison fully incapacitates offenders, so incarceration for $m > 0$ periods reduces exposure to recidivism risk by $m > 0$ periods over all horizons $n > m$. Second, since prisoners age while incarcerated, serving $m > 0$ periods implies that a defendant's first opportunity to recidivate occurs at age $a_0 + m$ instead of a_0, where a_0 is the defendant's age at sentencing. Finally, the experience of incarceration may have a direct impact on age-specific risks of recidivism following release from prison, i.e. during periods $n > m$.

The first two impacts are unavoidable consequences of incarceration. When the state imprisons someone, the state incapacitates them for a period of time, and then returns them to the community at an older age. This appendix explains how we create empirical measures of the first two impacts and also explains how the total impacts of incarceration should evolve over time if these are the only impacts of incarceration.

Notation

We model the time that elapses between sentencing and the arrival of the first new charge as a random failure time, τ. Time is discrete. We employ the following notation to describe the statistical process that governs recidivism risk:

- a_0 is the age at sentencing date.
- m is the number of periods of incarceration imposed by the sentence.
- $\tau \sim F(n|m, a_0)$ where $n \in \mathbb{Z}^+ \cup \{\infty\}$ and $F(0|m, a_0) = 0 \ \forall\, m, a_0$.
- $F(n|m, a_0) = 0 \quad \forall\, a_0, n \leq m \quad$ (full incapacitation)

Given our full-incapacitation assumption and the fact that $F(0|m, a_0) = 0 \ \forall\, m, a_0$, the impact of $m > 0$ periods of incarceration on the n-period recidivism rate is

$$\Delta(n, m) = -F(n|0, a_0) \quad \forall\, 0 \leq n \leq m$$

$$\Delta(n, m) = F(n|m, a_0) - F(n|0, a_0) \quad \forall\, n > m$$

For a moment, let us assume that the experience of incarceration has no impact on age-specific recidivism rates following release from prison. We can now write

$$F(n|m, a_0) = 0 + F(n - m|0, a_0 + m) \quad \forall\, (n, m) \text{ s.t. } n > m$$

This equation is key. If we can estimate $F(n - m|0, a_0 + m)$, we can calculate how incarceration impacts recidivism in a world where the experience of incarceration does not impact recidivism rates following release from prison. Here, we attack this estimation problem in two steps.

To start, let us make the additional assumption that age does not matter directly for recidivism rates. This assumption implies that $R(n) = F(n|0, a_0) = F(n|0, a_0 + m) \quad \forall\, n \geq 0, a_0 > 0, m > 0$. Here, $R(n)$ is the probability that a defendant, who is not sentenced to incarceration, faces a new charge within n periods of receiving his non-incarceration sentence. Likewise, for $n > m$, we have

$R(n-m) = F(n-m|0,a_0) = F(n-m|0,a_0+m)$ is the probability that a defendant sentenced to m periods of incarceration faces a new charge within the first $n-m$ periods following his release from prison. Given these objects, we can define the impact of incapacitation on recidivism in our thought experiment above:

$$\Delta(n,m|\text{Incapacitation}) = -R(n) \quad \forall\, 0 \leq n \leq m$$

$$\Delta(n,m|\text{Incapacitation}) = R(n-m) - R(n) \quad \forall\, n > m \tag{4}$$

Here, we have assumed that recidivism risk is not a function of age and that the experience of prison does not impact future recidivism risk. Thus, the only impact of incarceration on recidivism is incapacitation, i.e. incarceration reduces recidivism risk by reducing exposure to recidivism risk. Given this formulation, we can estimate $\Delta(n,m|\text{Incapacitation})\ \forall\, m > 0$ if we can estimate a single distribution, $F(n|0,a_0)$.

Recall that Appendix Table 10.11.1 presents, for both first and repeat offenders, estimates of annual recidivism rates among compliers who were not sentenced to prison, $m = 0$. If we let one period equal six months and maintain the assumptions that neither age nor the experience of incapacitation impacts recidivism rates, then these estimates of n-period recidivism rates among non-incarcerated compliers also serve as $\hat{R}(n) = \hat{F}(n|0,a_0)$, which by assumption also equals $\hat{F}(n|0,a_0+m)\ \forall\, m > 0$. Similar expressions define $\hat{R}(n-m)$, i.e. the $(n-m)$-period recidivism rate that applies both to the non-incarcerated in the $n-m$ periods following sentencing and to the incarcerated in the $n-m$ periods following their release from prison.

Still, we cannot use equation 4 directly to estimate incapacitation effects because not all incarcerated compliers serve prison terms of the same length. To estimate the impact of incapacitation on recidivism among the incarcerated, we form a weighted average of the expression in equation 4, where the weights reflect the fractions of sentenced compliers that serve prison terms of various lengths.

$$\hat{\Delta}(n|\text{Incapacitation}) = \left[\sum_{m=1}^{n} \hat{R}(n-m)\hat{h}(m)\right] - \hat{R}(n)$$

Here, $\hat{h}(m)$ is our estimate of the fraction of incarcerated compliers who leave prison after serving $m-1$ periods but without serving m periods. These estimates are derived from the results in Appendix Table 10.11.2, which describe the distribution of prison time served by incarcerated compliers.

In Figures 5 and 6, we plot $\hat{\Delta}(n|\text{Incapacitation})$ separately for first and repeat offenders. At horizons beyond five years, our estimates of incapacitation effects are quite modest. This result is expected because, among both first and repeat offenders, less than 25 percent of incarcerated compliers serve more than two years in prison, and among both sets of compliers, recidivism rates rise rapidly over the first three years of risk exposure.[32]

Finally, note that, given the timing conventions we adopt in these discrete time formulas, these impacts are upper bounds on the absolute size of incapacitation effects. Here, all persons who are released from prison after serving $m-1$ periods but without serving m periods do not face recidivism risk until the end of period m.

The incapacitation effects plotted in Figures 5 and 6 provide answers to interesting thought experiments. However, much evidence suggests that age does have a direct impact on recidivism rates. Therefore, it makes sense to go beyond pure incapacitation effects and estimate the combined effects of incapacitating offenders and shifting exposure to recidivism risk to older ages. We estimate these combined impacts as follows:

[32] Among first offenders, estimation errors create small departures from monotoncity in Appendix Tables 10.11.1 and 10.11.2. Since both distribution functions must be monotonic, we impose monotonicity by making small adjustments these Appendix results. However, these adjustments matter little for the results.

$$\hat{\Delta}(n|\text{Incapacitation, Aging}) = \left[\sum_{m=1}^{n}\left[\hat{R}(n-m) - \Delta\hat{r}_{(n-m)}\Delta\hat{a}_m\right]\hat{h}(m)\right] - \hat{R}(n)$$

Here, we do not use $\hat{R}(n-m)$ as the recidivism rate for incarcerated offenders who have just been released after serving m years in prison. Instead, we adjust each $n-m$ year recidivism rate to account for the fact that prisoners age while they are incarcerated. Here, $\Delta\hat{r}_{(n-m)}$ is our estimate of the absolute reduction in $n-m$ year recidivism rates associated with aging one year. $\Delta\hat{a}_m$ is the average years served prior to release among persons who served more that $m-1$ but less than m periods in prison. It is therefore an estimate of how much those released in period m aged before they were released.

The following thought experiment illustrates the need for the adjustment factor, $\Delta\hat{r}_{(n-m)}$. Let the expected 60-month recidivism rate for a first-offender sentenced to non-incarceration at age 23 equal X. Next, assume that an identical 23 year-old is instead sentenced to serve one year in prison and therefore is age 24 when he first faces the possibility of receiving a new criminal charge, and let the expected recidivism for this defendant during the first 60 months following his release prison equal Y. Then, $X-Y$ is the reduction in the expected 60-month recidivism rate associated with aging one year in prison, assuming that a defendant enters prison at age 23. Next, given this thought experiement, recall that we construct our Figure 3 results for first and repeat offenders by estimating 28 2SLS regressions that each contain a vector of indicator variables for the age of the defendant at sentencing. Thus, if we consider the model for first-offenders where the outcome, y_{its}, equals one if defendant i sentenced at date t faces a new charge within $s = 60$ months following sentencing, the difference between the estimated coefficients on the indicators for sentenced at age 23 and sentenced at age 24 provides a proxy for the difference $X-Y$ in our thought experiment above. Given this observation, for first and repeat offenders separately and for each post-release horizon $n-m$, we estimate $\Delta\hat{r}_{(n-m)}$ by forming weighted averages of the implied reductions in recidivism associated with aging one year. Here, we are weighting over different sentencing ages, and we use the distribution of ages at sentencing among those sentenced to incarceration to form these weights. We create separate weights for first and repeat offenders.[33]

Since prisoners age one month during every month they serve in prison, we use the distribution of release dates among the incarcerated to estimate $\Delta\hat{a}_m$ for $m = 12, 18, 24, 78$. The product $\Delta\hat{r}_{(n-m)} * \Delta\hat{a}_m$ gives the expected reduction in $n-m$-period recidivism rates associated with aging one year in prison multiplied by the average years of ex post prison time served among those sentenced to m periods of incarceration.[34]

Figures 5 and 6 show that these age-adjustment factors are trivial at short horizons because persons released after serving short prison terms aged little while in prison. Further, these aging-while-incarcerated adjustment factors are never greater than 3.1 percentage points at any horizon. Among first offenders, the potential size of these adjustments is limited by the fact that three-fourths of incarcerated compliers in the first-offender sample serve prison terms of less than a year. These adjustments are limited among repeat offenders by the fact that repeat offenders are significantly older when they are sentenced, and the impacts of aging on recidivism rates are smaller at older ages.

Figures 5 and 6 also plot our 2SLS estimates of the total impacts of incarceration on recidivism at each horizon. We first plot these results in figure 3 and also present them in table format at the end of this section. At long horizons, it is obvious that the impacts of incapacitation and aging while incarcerated cannot account for the full impacts of incarceration on recidivism among first offenders.

[33] When $n = m$, we have $\Delta\hat{r}_{(0)} = 0$ because we assume that that prisoners released during period m are not exposed to recidivism risk until the end of period m.

[34] These averages provide rough approximations for the impacts of aging while incarcerated on post-release recidivism rates. We cannot do better because we do not have the statistical power required to estimate our models within cells defined by the intersection of first-offender status and age at sentencing.

Alternative Estimators for Complier Characteristics

We created Figures 5 and 6 using the estimates of recidivism rates among non-incarcerated compliers presented in Appendix 10.11.1 and the estimates of the distribution of time-served among incarcerated compliers presented in Appendix 10.11.2. We produced both sets of estimates using the interpolation methods developed in Dahl et al. (2014). We also produced alternative estimated of complier characteristics using a method presented in Garin et al. (2023).[35] We use these alternative estimates of complier characteristics to create alternative versions of Figures 5 and 6. We present these results at the end of this section.

Readers may have a hard time telling the difference between these results for first offenders and the first-offender results in Figure 5. This is because the Garin et al. (2023) method produces estimates of first-offender complier characteristics that are remarkably close to those we produce using the Dahl et al. (2014) interpolation method. The above results for repeat offenders are also quite close to those in Figure 6, but here incapacitation effects are slightly smaller given the Garin et al. (2023) estimates of complier characteristics.

Decomposing the Total Impacts of Incarceration

In the analyses above, all impacts of incarceration on recidivism operate through the direct effects of incapacitation or the shifting of recidivism risk to older ages. We now show that, given our thought experiment, all other differences in recidivism rates between those who receive no incarceration, $m = 0$, and those who receive $m > 0$ must be attributed to the impacts of incarceration on age-specific offending rates following release from prison.

Note that among offenders who receive their sentences at the same age, a_0, we can define age at duration n to be $a_n = n + a_0$. Then, we can define the current-period risk of recidivism for an offender sentenced to $m \geq 0$ years of incarceration who has survived since sentencing without a new charge and is now age a_n:

$$r(a_n|m, a_0) = 0 \quad \forall\, a_n \leq m + a_0$$

$$r(a_n|m, a_0) = 1-P((\tau > (a_{n+1}-a_0)|\tau > (a_n-a_0))|m, a_0) = 1-\frac{1 - F((a_{n+1} - a_0 - m)|m, a_0 + m)}{1 - F((a_n - a_0 - m)|m, a_0 + m)} \quad \forall a_n > m+a_0$$

Note that for any a_n, m enters these equations in three ways. First, prison incapacitates, i.e. $r(a_n|m, a_0) = 0 \quad \forall\ m > 0, m + a_0 > a_n$, and the survivor functions that determine $r(a_n|m, a_0)$ for $n > m$ record survival age relative to the base age $a_0 + m$, i.e. the age at prison release and not the age at sentencing, a_0. Second, prison time shifts recidivism risk to older ages, and this impacts survivor functions through the conditioning term $a_0 + m$. Even if the experience of prison has no direct impact on recidivism behavior, $1 - F(n|0, a_0) \neq 1 - F(n|0, a_0 + m)$ for $m > 0$. Put differently, even if we assume that prisoners who exit prison face the same per-period recidivism risk upon release that the non-incarcerated would face if sentenced at the same age, $a_0 + m$, this rate is different, and almost always, lower than the rate that the non-incarcerated would face at age a_0.

The only other way that m enters these equations is through the direct impact of serving prison time on the survivor functions, i.e. m impacts $F(n|m, a_0 + m)$ directly as well as indirectly through $a_0 + m$. If we assume that these direct impacts do not exist, as we did when creating Figures 5 and 6, we have

$$r(a_n|m, a_0) = 0 \quad \forall\, a_n \leq m + a_0$$

[35] The note below these figures provides more detail about this alternative method.

$$r(a_n|m, a_0) = 1 - P((\tau > (a_{n+1} - a_0)|\tau > (a_n - a_0))|m, a_0) = 1 - \frac{1 - F((a_{n+1} - a_0 - m)|0, a_0 + m)}{1 - F((a_n - a_0 - m)|0, a_0 + m)} \quad \forall a_n > m + a_0$$

Comparing the two sets of equations, we realize that any impact of incarceration on recidivism that does not operate through incapacitation or shifting recidivism risk to older ages must arise because $F(n|m, a_0 + m) \neq F(n|0, a_0 + m)$ for $m > 0$, i.e. the experience of prison changes behavior following release from prison for reasons above and beyond the fact that inmates age while incarcerated. Our results for first offenders in Figure 5 suggest that the experience of prison lowers recidivism rates following release.

This decomposition result holds given two key assumptions that are implicit in our notation. First, there are no impacts of calendar time on recidivism. Prisoners who are released at at given age $a_0 + m$ in a given year behave just like prisoners released at age $a_0 + m$ in any other year. Second, our initial thought experiment involves random assignment, which guarantees that the distributions of risk types among compliers sentenced to incarceration of any length m are the same as the distribution of risk types among compliers who receive non-incarceration sentences. Thus, we are able to analyze these issues using a single survivor function, $F(n|m, a)$, without addressing unobserved risk types and the distribution of these types among those assigned to incarceration or non-incarceration sentences.

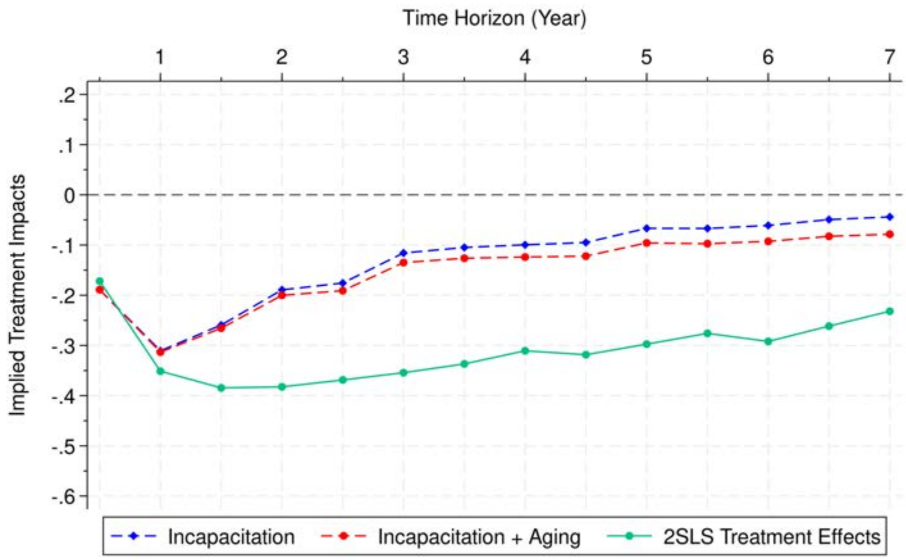
Treatment Impacts Plotted in Figure 3

Sample	Horizon	2SLS	SE
First Offenders	6	-.172	(.064)
First Offenders	12	-.351	(.07)
First Offenders	18	-.384	(.081)
First Offenders	24	-.382	(.095)
First Offenders	30	-.369	(.089)
First Offenders	36	-.354	(.085)
First Offenders	42	-.337	(.09)
First Offenders	48	-.311	(.095)
First Offenders	54	-.318	(.103)
First Offenders	60	-.297	(.109)
First Offenders	66	-.276	(.099)
First Offenders	72	-.292	(.105)
First Offenders	78	-.261	(.098)
First Offenders	84	-.232	(.091)
Repeat Offenders	6	-.111	(.04)
Repeat Offenders	12	-.23	(.073)
Repeat Offenders	18	-.21	(.088)
Repeat Offenders	24	-.179	(.089)
Repeat Offenders	30	-.203	(.08)
Repeat Offenders	36	-.199	(.081)
Repeat Offenders	42	-.189	(.073)
Repeat Offenders	48	-.103	(.07)
Repeat Offenders	54	-.041	(.067)
Repeat Offenders	60	-.006	(.077)
Repeat Offenders	66	.001	(.074)
Repeat Offenders	72	.001	(.071)
Repeat Offenders	78	-.01	(.069)
Repeat Offenders	84	-.003	(.072)

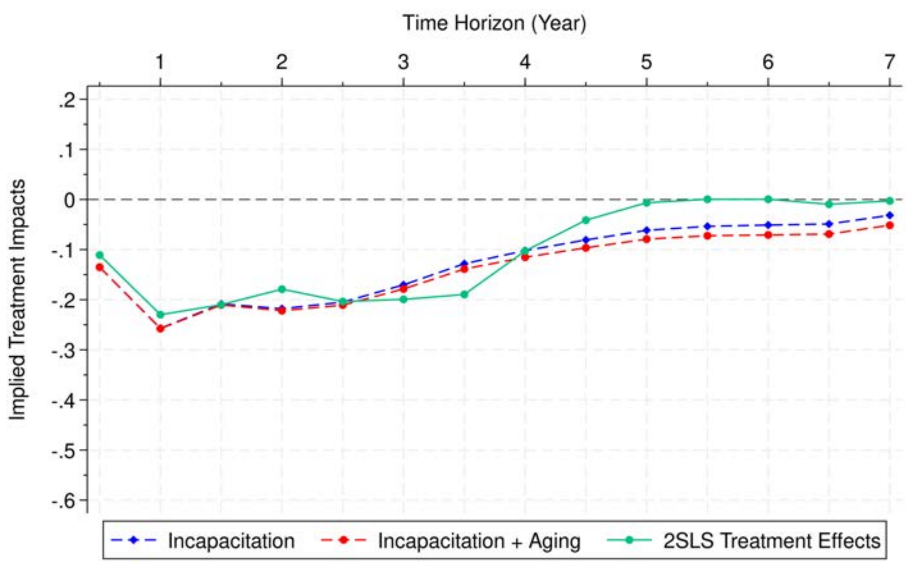
Notes: Each row presents the estimate LATE effect of incarceration on recidivism as a different horizons. These results are plotted in Figure 3. We report HAC standard errors. For first offenders, we cluster at the judge level. For repeat offenders, we two-way cluster at the defendant*judge level.

Alternative Versions of Figures 5 and 6

First Offenders



Repeat Offenders



Notes: These figures parallel Figures 5 and 6, but here we use a different method to estimate recidivism rates among non-incarcerated compliers and the distribution of time-served among incarcerated compliers. The alternative estimator presented in Garin et al. (2023) involves three steps. Consider the task of estimating recidivism rates among non-incarcerated compliers. First, form new outcome variables that are the product of our recidivism outcomes and a dummy for non-incarceration treatment. Next, define treatment as non-incarceration, and then separately for first and repeat offenders, run the resulting 2SLS models using our LOM severity measures as instruments. For each model, the estimated coefficient on the new treatment dummy is an estimate of the sample-specific recidivism rate among non-incarcerated compliers at a given horizon, e.g. the recidivism rate for non-incarcerated compliers in the first-offender sample at 60 months. We follow parallel steps to produce alternative estimates of the distribution of time-served among incarcerated compliers.

12 Data Appendix

Our raw data come from the Clerk of Court for Cook County, IL, and the Illinois Department of Corrections (IDOC). We begin with electronic records from the Clerk of Court that describe cases that were active in the Court between January, 1984 and December, 2019. The data contain 531,388 defendants who were involved in 1,273,605 felony cases.

We only use a subset of these records. A later section of this appendix describes all of the sample selection rules we impose. Here, we comment on four key selection rules.

First, we do not include female defendants. The sample of female defendants is too small to analyze separately, and we are not willing to assume that judge severity is invariant to defendant gender.

Second, since we perform separate analyses for first and repeat offenders, we eliminate defendants born before 1967. Since defendants age 17 or older are typically tried as adults and our court records begin in 1984, we cannot determine from court records whether those born before 1967 are facing their first felony charge in Cook County.

Third, we do not consider cases that the Court initiates before 1990 or after 2007. For cases before 1990, we are not able to use IDOC data to help identify cases that involve a nominal prison sentence but no time served in prison. For cases that begin after 2007, we do not have a full seven years of IDOC data following sentencing, and this means that we are not able to see all the recidivism events that we observe for cases that begin in 1990-2007 over a seven-year horizon. Whenever we observe an IDOC prison admission from a county in IL other than Cook County, we know that the admitted person committed a crime elsewhere in IL.

Finally, we only use cases that we feel confident are randomly assigned to judges who work in the main criminal court in Chicago. Below, we explain how we identify these judges.

Given these key sample restrictions and others motivated by missing data and measurement objectives, our final analysis sample consists of 55,285 defendants involved in 70,581 felony cases initiated between the 2nd of January, 1990 and the 17th of December, 2007.

The Clerk of Court of Cook County provides three types of data:

- the *root* data contain basic demographic information about the defendant and the case initiation date.
- the *charge* data describe each charge initiated by prosecutors.
- the *dispositions* file describes the 54 million dispositions filed during these felony cases.

We rely heavily on the root and charge data when creating variables that characterize defendants and the cases against them. All of these variables describe the state of the case prior to random assignment. For example, our charge measures capture the charges filed against defendants before their arraignment. We use the disposition data in concert with IDOC data to determine the effective sentencing decisions made by judges. We use the court files in concert with IDOC data to mark recidivism events.

12.1 Initial Cleaning

After receiving the raw court data, we interviewed a former judge, two former public defenders, a former prosecutor, employees of the office of the Clerk of the Circuit Court of Cook County, employees at the Adult Probation department, and representatives of nonprofits that specialize in the criminal justice system. Based on these conversations, we made the following edits:

1. The Clerk of Court occasionally mis-records credit for time served dispositions as probation sentences because the disposition codes are off by one digit. The disposition code for credit for time served is 521 and the disposition code for probation is 531. We identified these typos by checking whether the sentence length was denominated in days. Probation sentences are never denominated in days, so if a probation sentence length is denominated in days, it is a typo, and the disposition represents credit for time served. We correct these typos in the raw data. In total, we corrected 1,019 dispositions for this reason.
2. The clerk also occasionally mis-records probation sentences as Credit for Time Served dispositions. We correct 273 cases where we feel confident that this typo occurred.
3. The Court occasionally indicates a sentence to CCDOC when the individual is under CCDOC's authority but not actually held in CCDOC. We recode 1,905 dispositions to mark that the defendant was not incarcerated. In these cases, the "free description" (notes) section for each disposition reveals what really happened. We code these as "other" sentences. This category contains all defendants who are found guilty but not required to be supervised by the probation department, IDOC, or the Bootcamp program run in CCDOC by the Sheriff. As we explain further below, these sentences are assigned in cases where a defendant arraigned on felony charges resolves his case by pleading guilty to misdemeanor charges. The "free description" codes associated with "other" sentences are:
   - TASC – Treatment Alternatives for Safe Communities, see: https://www.tasc.org/tascweb/home.aspx

- ELECTRONIC – Electronic Monitoring, see:
  https://www.cookcountysheriff.org/cook-county-department-of-corrections/electronic-monitoring-program-placement/
- GATEWAY – Drug treatment and other services/programming to reduce recidivism, see:
  http://gatewaycorrections.org/locations/illinois/
- HRDI – Drug and alcohol treatment.
  See https://www.hrdi.org/
- SFFP – Sheriff's Female Furlough Program, see:
  https://www.cookcountysheriff.org/cook-county-department-of-corrections/sheriffs-female-furlough-program-sffp/
- HAYMARKET – Drug treatment program, see:
  http://www.hcenter.org/about-us
- WESTCARE – Primarily a drug treatment program, but they offer other interventions as well.
  See: https://www.westcare.com/page/what-we-do_01

4. If a CCDOC sentence free description included the substring "PROB", we recode it as a probation sentence. We recode 460 sentences for this reason.
5. If a CCDOC sentence free description included the substring "BOOT", we recode it as CCDOC Boot Camp. CCDOC Boot Camp is 4 months of incarceration in CCDOC and 8 months of probation. See: http://www.digibridge.net/bootcamp/facts.htm. We recode 68 dispositions as CCDOC Boot Camp.

12.2 Identifying Sentences

We use the raw disposition codes to identify and record the sentencing information for each case. We focus on the first four sentencing dates in each court case. While approximately 98% of the cases in the sample have two or fewer sentencing dates, a small subset of cases have 3 or more. 862 cases (less than 0.1% of the sample) have more than 4 sentencing dates. In those cases, we still limit our attention to the first four sentencing dates. If a defendant is not convicted, there is no sentence. Everyone who is convicted receives a sentence of some type. We use the sentencing disposition codes to place sentences into one of four categories:

1. Incarceration in IDOC
2. Incarceration in CCDOC Boot Camp
3. Probation
4. Other

According to state law, defendants convicted of felonies must receive either (1), (2), or (3), and defendants convicted of reduced, misdemeanor charges cannot receive (1), (2), or (3). The Other sentences in (4), which include the sentences described in section 12.1 above, are therefore misdemeanor sentences.

Occasionally, sentences of multiple types will be given on the same day. We record all of the sentence types given on that date. Within each sentence type (IDOC, CCDOC, Probation, and Other), we record the longest sentence length. For example, if an individual is given two IDOC sentences, one for 6 months and one for 12 months, we record the most severe sentence as 12 months. There is one exception to this rule. If the sentences are set to run consecutively (as noted by a disposition in the disposition file), we set the sentence length for each type to be equal to the sum of the sentences of that type on that day. This is rare. Most sentences given on or near the same day run concurrently.

Next, we identify credit for time served information for each sentencing date. In many cases, the Court records these credits in a separate disposition. We see some sentences marked as "Time Already Served." In these cases, the Court is typically noting that the requirements of a sentence that mandates additional jail time have been satisfied because the defendant spent a long time in jail waiting for a verdict. A variety of special disposition codes mark these sentences. If any of these codes appears on the sentencing date, we consider the sentence time already served.

The Court does not always record time-already-served sentences correctly. Based on conversations with staff from Adult Probation, we flagged jail sentences that are denominated in days but not equal to 364 days or multiples of 30 days. When such sentences are not accompanied by any dispositions marking credit for time served, we assume that these are actually time-already-served sentences. This decision affects 13,192 sentences. We classify all time-already-served sentences as "other."

In the end, a small fraction of sentences appear to require defendants to serve some time in jail but not participate in the CCDOC Boot Camp program. We do not code these sentences as incarceration sentences. If these sentences are paired with adult probation sentences, we treat them as felony probation sentences. In the rare cases where these sentences are stand alone events, we treatment them as misdemeanor sentences and classify them as "other."

Several factors drive our decision to code stand alone misdemeanor jail sentences as non-incarceration sentences. We know that, in a number of misdemeanor cases, credits for time-served are awarded by the judge but never recorded by the Clerk of Court, and we have no data that document the timing of post-sentencing exits from CCDOC jails.[36] As a result, in the rare cases where a defendant pleads guilty to a reduced misdemeanor charge and receives a sentence of additional jail time, we have no reliable way to confirm whether or for how long the defendant remained in jail. Finally, given our numerous interactions with lawyers, judges, and court staff, we know that court actors view IDOC sentences and CCDOC Bootcamp sentences as the two ways that judges sentence offenders to meaningful periods of incarceration.

12.3 Constructing a Case-Level Dataset

The Court assigns each case to a call. A call is a calendar of cases that a particular judge is responsible for handling. Other judges may work on cases in the call because vacations, sick leaves, and other factors make it impossible for one judge to handle all hearings for all cases assigned to a given call, but the Court organizes case assignment by calls. Calls have numbers, and in the electronic files produced by the Clerk, these numbers are labeled "Courtroom," but call numbers do not reveal physical locations in a particular Courthouse.

Case numbers identify both collections of charges and defendants. If a defendant is charged with multiple offenses, all of the offenses share the same case ID number. However, if a group of defendants are all charged with committing a crime together, the Clerk will record a separate case ID number for the charges against each defendant. We save case-level information from the disposition history by flagging various dispositions of interest. Our final case-level data set saves a single record for each case.

Many court participants told us that judges almost always award credit for time-served in jail when they sentence a defendant to IDOC. Therefore, when we see IDOC sentences, we assume that defendants receive credit for their jail time, even if we cannot find an explicit sentencing credit in the disposition history. In cases where the Court failed to record the exact amounts of these credits, we can often find records in the IDOC files that list them. In cases where we cannot find an explicit credit record in either the Court or IDOC data, we estimate the amount of time each convicted defendant spent in jail prior to sentencing using dispositions that indicate whether the defendant was in custody or on bond at each court appearance.

We determine whether a case was dropped by beginning with the data set containing all of the charges for each case. We mark a charge as dropped if any disposition code indicates that the prosecution dropped the charge. If every charge in the case was dropped, we consider the case dropped.

12.4 Tracking Individuals

To identify defendants, we rely on the fingerprint ID associated with each case. A fingerprint ID is a unique numerical identifier given by the Cook County Court system to each person upon intake. In some cases, the system assigns multiple IDs to the same individual, and we combine the two fingerprint IDs into a new unique individual identifier. We make these combinations on the basis of FBI numbers, IDOC numbers, and demographic information. In some cases, especially from the 1980s, a fingerprint ID is missing. In these cases, we use a defendant's name, race, sex, and exact birth date to try to find a different case he was involved in where a valid fingerprint ID exists. When there is no other case with a valid fingerprint ID for a defendant, we assign a synthetic fingerprint ID to defendants with unique names. We drop defendants who are missing both fingerprint IDs and valid demographic information.

12.5 Matching Court Records to Prison Records

To improve our measure of effective sentences and recidivism, we rely on both the Court's records and IDOC records. We match our case-level data from the Cook County Court system to IDOC records by creating a crosswalk between the unique individual identifiers in the court data and the unique individual identifiers in the prison data. The court and prison data both include demographic information as well as sentencing dates. We match individuals on the basis of shared demographic information and sentencing dates in the court and prison data. This allows us to mark recidivism events that take the form of criminal charges filed outside Cook County that result in new prison admissions.

To learn more about the time-served required by various sentences recorded in the court records, we locate court cases that resulted in admissions to an IDOC prison. We start with IDOC admission records that result from sentences announced in a Cook County court. Next, we identify the sentences in the Cook County Court data that could produce an IDOC admission record. Our IDOC data begin in 1990 and end in early 2015.

We now match each individual's eligible court records to his eligible IDOC admissions. The court and IDOC data both contain sentencing dates, sentence length, crime category and class number variables. We match IDOC spells with any court sentence that has the same sentencing date and either the same sentence length, or the same crime category and class number.

[36]Among those sentenced to IDOC, we may be able to infer CCDOC exit dates my matching the sentence to an IDOC admission record.

12.6 Combining Cases into Episodes

Sometimes the Court opens multiple cases against an individual simultaneously. We combine information from these simultaneous cases. We treat two cases as one case if the initiation date for the second case occurs before the terminal date for the first case. We define the terminal date as follows:

1. First sentencing date, if the case includes any sentences.
2. First not guilty disposition date if the case did not end in a sentence, and the case had a not guilty disposition.
3. First date of a disposition indicating the case was dismissed if the case did not end in a sentence, did not have a not guilty disposition, and did have a dismissed disposition.
4. The date the case was dropped if the case was dropped.
5. For all remaining cases, the final disposition date we have on record is the terminal date.

Combining cases that were tried simultaneously decreases our sample of felony cases with valid fingerprint IDs from 1,231,946 to 1,018,702. We refer to combined cases as episodes.

The Court occasionally initiates new cases against a defendant while the defendant is serving a prison spell in IDOC associated with a previous court sentence. These cases are not associated with crimes committed in prison. When inmates commit crimes in prison, the charges are filed in the County where the prison is located. There are no state prisons in Cook County. These cases appear to be the result of information gathered while investigating a previous case. We delete these cases from our data.

12.7 Treatment Variable Creation

This section explains how we define our key treatment variable, $\tau_{j(i,t)}$. This is an indicator for whether defendant i received a sentence, after being assigned at t to the call run by judge j, that required i to serve time in a state prison or the CCDOC Bootcamp program. We set this indicator to zero if the case against i at t:

1. Contains no sentence to IDOC or CCDOC Bootcamp
2. Contains a sentence that results in a match to IDOC admission records followed by an exit within two weeks. We have learned that, even in cases where the defendant is admitted to the IDOC system, receives an MSR (parole) agent assignment, and exits prison on the same day, the exit may be recorded with a lag. Also, inmates who stay less than two weeks in reception centers are never evaluated and assigned to a regular prison.
3. Contains a sentence that matches to an IDOC admission record but there is no corresponding exit record, and the sentencing and credit for time served information in the prison records implies that the sentence required less than two weeks of additional time served.
4. Contains a sentence to prison that does not match any IDOC admission record, and the implied additional time-served based on court records is less than three weeks after the initiation date.

Else, $\tau_{j(i,t)} = 1$

We based both the two and three week rules on observed relationships between the additional expected time-served implied by a common rule of thumb formula, i.e. .5(nominal sentence) - (credit for time already served), and the prevalence, among matched sentences, of admission and exit records that share a common date. When we see a prisoner enter and exit the IDOC system on the same day, we know that the prisoner did not owe any time and that the admission process served only as a means for assigning the offender to an MSR agent.

12.8 Artificial Records of Recidivism Events

If an individual commits a crime outside of Cook County, the offense is not recorded by the Clerk in Cook County. However, when these crimes result in IDOC admissions, we observe them in our IDOC data. We count these events as recidivism by creating artificial court records for them. We date these events by estimating initiation dates for the cases that created the admissions. Matched court and IDOC data allow us to build a model of the time between the date a charge is filed in court and the date a sentenced defendant enters the prison system.

We create artificial records for admission from courts outside Cook County or MSR violations associated with a new court charge outside Cook County. We do not count technical MSR violations as recidivism events.

12.9 Outcome Variables

Our key outcome variables are indicator variables for the presence of at least one new charge within 12, 24, 36, 48, 60, 72, or 84 months of the terminal date of a case. A new charge may be any of the following:

1. The initiation of a new court case in Cook County.
2. An imputed initiation date associated with a case outside of Cook County. Some of these cases may begin while the offender is on MSR.

We require that all recidivism events occur after the potential recidivism date for a case, which is the date when the offender is assumed to be at risk of recidivism. Charges associated with all crimes committed in an IDOC prison are filed in the county where the prison is located. There are no IDOC prisons in Cook County. Therefore, any charged filed in Cook County against a person who is currently an inmate in an IDOC prison must be linked to criminal activity that took place before the inmate entered prison.

We ignore new charges that occur before these dates:

1. The date of exit from prison - if the case resulted in an IDOC sentence and a matched prison spell with an exit record.
2. An estimated exit date from prison (based on information in IDOC records) - if the case resulted in an IDOC sentence and matched prison spell without an exit record.
3. An estimated exit date from prison (based on information in court records) - if the case resulted in an IDOC sentence and no matched prison spell in the IDOC admission records.

For cases where we expect defendants to serve more than one year, we multiply our best estimate of expected time served by .8 to create a conservative estimate of each incarcerated defendant's release date. We do this to avoid missing recidivism events among persons who may have earned early release due to good behavior in prison. We have also re-estimated our main empirical specifications using .5 instead of .8, and our results do not change. Thirteen of our 14 LATE estimates of the impact of incarceration for first and repeat offenders at different horizons change by less that .02 in absolute value, and the remaining one changed by .021.

For defendants incarcerated in the CCDOC Bootcamp, we count new charges filed during the required four month spell of incarceration as recidivism events. Any crime committed while CCDOC custody will be charged in Cook County and therefore in our data.

12.10 Geography

We create an indicator variable that marks offenders who likely grew up in a high-crime area. Cases in the Cook County Court data record the defendant's address at the initiation of the case. For each defendant in our analysis dataset, we use GIS software to geocode the first address associated with that defendant. We then project the resulting latitudes and longitudes onto a shapefile for Chicago's 77 community areas. A small number of addresses cannot be geocoded and are instead assigned to a community area by hand. If an address cannot be geocoded by hand or is located outside of Chicago, we treat the defendant as not coming from a high crime area.

A report by Rob Paral and Associates, Paral (2003), documents the average homicide rate in each Community Area over the five-year period 1994-1998. Twenty-five of the 77 areas had murder rates over 40 per 100,000 people during this period. We mark these 25 community as high-crime areas.

We explored several alternative methods. One designated defendants as having grown up in a high-crime community area based on per-capita charges in the court system. Another employed reports from the Chicago Police Department concerning index crime rates by community areas in some years and police districts in others. Both procedures involved a number of necessarily arbitrary choices concerning the weighting of various offenses, interpolation methods, and imputation rules. However, in both cases, the resulting indices provided support for the claim that the 25 high-crime community areas that we identify based on the 1994-1998 homicide rates are the high-crime community areas in Chicago during our sample period.

12.11 Waterfall of Data Restrictions

To give readers a sense of how we use the cleaning procedures discussed above to arrive at a final sample of cases, we describe how various sample selection rules impact our sample. The rules below describe how we identify a sample of randomized cases. It is important to note that we do employ the information contained in many of the cases we drop. For example, a case that is assigned to a special drug court during a preliminary hearing may count as a recidivism event for a particular offender even though it is not part of our sample because it was diverted to a special court and not randomly assigned to a call.

Our data has over 1 million cases. However, we only consider cases assigned at Leighton Criminal Court House, the main criminal court in Chicago, to calls that could have received randomized cases. We are not sure how cases are assigned in suburban courts, and we eliminate some calls in Leighton that did not receive random cases, e.g. Narcotics courts or Mental

Health courts. We have 306,804 cases with valid identifiers that could have been randomly assigned at Leighton. Starting with this sample, we make the following sample restrictions:

1. We drop cases that either were not resolved by the first of January 2008 or were not initiated by the first of January 1990: 306,804 → 219,651 (87,153 dropped)
2. We drop cases associated with any defendants who were older than 17 in 1984 when our Court data begin. This allows us to observe the full criminal histories in Cook County for each defendant in our sample: 219,651 → 130,433 (89,218 dropped)
3. We drop cases where we have explicit or implicit evidence the defendant was on probation, because these cases are not randomized: 130,433 → 109,139 (dropped 21,294)
4. We drop all cases that begin while an individual is in prison if the prison spell began because of a court case. 109,139 → 107,874 (1,265 dropped)
5. We drop cases whose most severe charge by class is in one of the following crime categories: murder, sex crime, armed violence, prison, court, traffic, inchoate: 107,874 → 93,212 (dropped 14,662)
6. We drop cases with female defendants: 93,212 → 83,800 (dropped 9,412)
7. We drop cases with more than 4 defendants: 83,800 → 82,127 (dropped 1,673)
8. We drop cases where not all cases within the episode are assigned to the same courtroom: 82,127 → 80,958 (dropped 1,169)
9. We drop cases with an IR number we believe may be a combination of multiple distinct individuals: 80,958 → 78,435 (dropped 2,523)
10. We drop cases that begin during a technical MSR prison spell if the case was initiated more than 30 days after the prison admission or if the preceding case was initiated within 30 days of the prison admission. 78,435 → 78,358 (dropped 77 cases)
11. We drop all cases where the felony had a class number which implied it was actually a misdemeanor: 78,358 → 78,263 (dropped 95)
12. We drop cases missing the defendant's age: 78,263 → 78,231 (dropped 32)
13. We drop cases missing the defendant's race: 78,231 → 78,042 (dropped 189)
14. We drop cases missing the defendant's gender: 78,042 → 78,041 (dropped 1)
15. We drop cases missing the class of the charge: 78,041 → 78,037 (dropped 4)
16. We drop cases missing the crime category of the charge: 78,037 → 78,008 (dropped 29)
17. We drop cases where the defendant was defrauding the state: 78,008 → 77,977 (dropped 31)
18. We drop cases where it was impossible to properly identify the marginal length on the defendant's sentence: 77,977 → 77,909 (dropped 68)
19. We drop cases where the defendant died or fled, the case is ongoing, or the case ended but we are unable to determine how it was resolved: 77,909 → 76,574 (dropped 1,335)
20. We drop cases where the judge was a "floater" (temporary) judge: 76,574 → 76,561 (dropped 13)
21. We drop cases assigned to judges who did not have at least 500 cases in the analysis sample: 76,561 → 70,581 (5,980 dropped)

Loeffler (2013) also used Cook County data, but he did not separate repeat offenders from first offenders, so he did not need to restrict his sample on birth year. Note that, in the second step above, we lost more than one-third of the sample by eliminating offenders born before 1967. Based on the observed relationship between age and first-offender status in later birth cohorts, we feel confident that the majority of these deleted cases involve charges against repeat offenders.

12.12 Leave-Out Mean Creation

To create the LOM instruments for our key regression models, we divide our analysis sample into first offenders and repeat offenders. We then regress $\tau_{j(i,t)}$ on the following variables

1. A vector of indicator variables for the case's initiation year
2. A vector of indicator variables for the class of the most severe charge in the case
3. A vector of indicator variables for interactions between class and year
4. A vector of indicator variables for the crime category of the most severe charge.
5. A vector of indicator variables for interactions between crime category and an indicator for a less severe class assignment within the given category. The distribution of class assignments varies greatly with category. A full set of class*category interactions is not feasible because there are many empty cells.

6. A vector of indicator variables for the number of prior charges on the defendants record
7. A vector of indicator variables for the defendant's age
8. An indicator variable for the presence of multiple defendants
9. An indicator variable for the presence of multiple charges
10. An indicator variable for Black
11. An indicator for initial residence in a high-crime area.

We capture the residuals from these regressions, and we form LOM averages at the assigned judge level within first offenders and within repeat offenders. We form additional LOM measures for some subsample analyses, e.g. first offenders facing drug charges, by summing these residuals within specific subsets of first or repeat offender cases.

When forming these LOM averages for $j(i,t)$, we "leave out" the case in question, (i,t), all cases against co-defendants that are bundled with the case in question at assignment, and all other cases involving defendant i.