

# Smart Sentencing Guidelines: The Effect of Marginal Policy Changes on Recidivism

Sarah M. Estelle  
Hope College

David C. Phillips  
University of Notre Dame<sup>1</sup>

May 2018

## Abstract

Public appeals regarding criminal justice have shifted somewhat from “tough on crime” to “smart justice” that is more lenient when tradeoffs merit it. Among other considerations, smart sentencing policy depends on how sentence severity affects recidivism. Using administrative data on two common non-violent felonies committed by adults in Michigan, we measure the effect of sentences on offenders’ future criminal activity. Discontinuities in the legislative guidelines that constrain sentences chosen by Michigan judges provide exogenous variation in sentence severity. Harsher sentences generated by sentencing guidelines significantly reduce recidivism by felony shoplifters but not repeat drunk drivers. Recidivism falls most for young, male offenders from Southeast Michigan and varies non-monotonically with prior criminal record. Because of such heterogeneity, any empirical strategy measures a local average treatment effect relevant to a particular population of offenders. Contrary to our main results, we find no evidence that harsher sentences induced by judge assignments reduce recidivism in our sample. When sentencing guidelines provide the primary practical policy lever, “smart justice” should incorporate directly relevant empirical evidence that accounts for offense- and offender-specific tradeoffs between public safety and the public budget.

**JEL Codes:** K42; K12

**Keywords:** sentencing guidelines; recidivism; mandatory minimum

---

<sup>1</sup> E-mail: [estelle@hope.edu](mailto:estelle@hope.edu) and [David.Phillips.184@nd.edu](mailto:David.Phillips.184@nd.edu). We have benefitted from comments by Steve Billings, Peter Brummund, Doug Kosinski, Camille Landais (the editor), Thomas Loughran, Robert Mohr, seminar participants at Notre Dame, and two anonymous reviewers. This project benefitted from excellent research assistance from Alex Belica, Tim Bloemendal, and Katlyn Hettinger. The Michigan Colleges Alliance provided financial support for this project. We are grateful for help from the Michigan Department of Corrections Office of Research and Planning, especially Jeff Anderson, for extensive help with the data. The views expressed are those of the authors and do not necessarily reflect the official policies of the Michigan Department of Corrections. All remaining errors and omissions are the responsibility of the authors.

“So our criminal justice system isn’t as smart as it should be. It’s not keeping us as safe as it should be. It is not as fair as it should be. Mass incarceration makes our country worse off, and we need to do something about it.” – President Barack Obama (D), address to NAACP Conference, July 14, 2015

"It's time to reinvent public safety in Michigan. ...By moving forward with a system of 'smart justice,' we will hold chronic offenders accountable for their actions, bring peace of mind to community residents, help to break the cycles that perpetuate crime, and unleash Michigan's economic growth." – Michigan Gov. Rick Snyder (R), Special Message on Public Safety, March 7, 2012

If the last half of the 20th century was an era when elected officials needed to prove they were “tough on crime,” the tide has shifted to include a new and widespread call for criminal justice reform sometimes referred to as “smart justice.” In a speech to the NAACP, President Obama admitted that the issue has “created some unlikely bedfellows” and praised the ACLU and attorneys general of his administration, but also Newt Gingrich, Rand Paul, Americans for Tax Reform, and Charles and David Koch. What seems to unify these otherwise dissimilar camps is an understanding that there are myriad unintended consequences to universally harsh criminal sentencing. This unusual political convergence matches recent high-quality empirical evidence that the extent of incarceration can be reduced while simultaneously reducing average recidivism. Aizer and Doyle (2015) and Mueller-Smith (2016), for example, find that those assigned more severe sentences because they were randomly assigned to more severe judges recidivate at higher rates. While exogenous variation in sentences driven by judge assignments provides direct evidence on particular public policies, such as replacing one judge with another or reducing judicial discretion, criminal justice reform more often engages other policy levers. The present research aims to extend this empirical evidence by examining the effects of marginal changes to a policy frequently under the authority of legislatures: sentencing guidelines.

Structured sentences include any constraint on or recommendation for a sentence that is determined by a legislative body as opposed to a presiding judge. They are currently utilized at the federal level and roughly half of states have sentencing guidelines and/or sentencing commissions.<sup>2</sup> The structure may be imposed on judges, parole boards, and/or juries with sentencing influence. Legislation can institute advisory sentencing guidelines that only recommend a range of possible sentences based on the severity of a criminal offense and the extent of an offender's criminal history, mandatory sentencing guidelines which judges cannot depart from for particular criminal charges, or presumptive sentences that allow some intermediate variation in uttered sentence or sentence served. If structured sentences increase sentences served above what they would be otherwise, e.g. through mandatory minimum sentences, they might be considered tough. However, if they provide or improve clarity and predictability; punishment proportional to the crime; and blindness in sentencing to race, gender and other factors, they might also be smart. Despite some states and the federal government reducing the influence of structured sentencing strategies, it is likely they will remain in place in some capacity.<sup>3</sup>

We exploit discontinuities in sentencing guidelines in Michigan to identify a causal path from sentencing guidelines to the sentence an offender receives and later criminal activity, or

---

<sup>2</sup>Characterizing the prevalence of sentencing guidelines across states and time is nontrivial. For a current list of sentencing commissions see the National Association of Sentencing Commissions (<http://thenasc.org/>). For the most recent unified discussion of state sentencing guidelines and a presentation of the nature of those guidelines ranging from voluntary to mandatory, see a 2008 report from the National Center for State Courts [http://www.ncsc.org/~media/Microsites/Files/CSI/State\\_Sentencing\\_Guidelines.ashx](http://www.ncsc.org/~media/Microsites/Files/CSI/State_Sentencing_Guidelines.ashx). Differences between these two lists can reflect changes in state policies over time, but also that not all states with sentencing guidelines utilize sentencing commissions. Michigan's guidelines, for example, were initially a recommendation of the State Supreme Court. In contrast, since 2014 a Criminal Justice Policy Commission, largely appointed by the governor (<http://council.legislature.mi.gov/CouncilAdministrator/cjpc>), has made proposals related to use of criminal justice resources including the effectiveness of misdemeanor and felony sentences.

<sup>3</sup> In fact, legislative constraints on judicial discretion pre-date even the tough on crime era. For example, the 1790 Crimes Act, which created the first federal crimes, imposed a mandatory death penalty for roughly one-third of those crimes. By the 1950s, federal mandatory minimums were extended to drug crimes (a policy that lost favor and legal support in the 1970s and regained relevance in the 1980s). (United States Sentencing Commission, 2011) Other modern examples of legal support for structured sentencing include "truth in sentencing" laws (1984) and "three strikes rules" (starting in Washington state in 1993).

recidivism. As in the extant literature, this research recognizes that both future criminal behavior and sentencing decisions likely reflect some unobservable characteristics of offenders, and we therefore exploit exogenous variation in sentencing. The State of Michigan assigns different sentencing guidelines to offenders convicted of the same crime but who have criminal history scores that place them on different sides of various but distinct thresholds. The discontinuities in sentencing guidelines allow this research to identify the causal connection between the uttered sentence and future outcomes.

The data focus on individuals convicted of two non-violent, mid-level felonies: Operating While Intoxicated-3<sup>rd</sup> Offense (hereafter “OWI” i.e. repeat drunk driving) and Retail Fraud 1<sup>st</sup> Degree (severe or repeated shoplifting). We are constrained to these two categories of crimes because the Michigan Department of Corrections had already constructed the data set for other administrative reporting needs. We are fortunate, however, that these crimes are quite common and provide two benefits beyond their large sample sizes. First, both are policy-relevant “straddle cell” crimes for which the sentence may or may not include any incarceration. Second, the nature of the crimes, judged in Michigan to be of similar egregiousness, are still sufficiently different as to provide insights on heterogeneous treatment effects by crime. On average, if one criminal activity indicates a substance abuse issue and another reflects calculated rational choice, then how much time one spends within the state corrections system may matter differently. Since crimes related to substance abuse and property compose the vast majority of non-violent felonies, our research provides two useful examples of how “smart” criminal justice reform may respond to dissimilar offenses.

Our results indicate that incarceration affects future criminal behavior differently depending on an offender’s initial charge, prior criminal record, age, sex, and location within the state.

Averaging across both offenses and three cutoffs in the guidelines, offenders assigned to harsher guidelines commit 20 percent fewer future crimes. Decreased recidivism among Retail Fraud offenders drives this change, while we find no evidence that future criminal behavior of those convicted of OWI responds to more severe penalties. The guidelines incorporate cutoffs for three groups of offenders with criminal records of different severity. Harsher sentencing reduces recidivism among the groups with the most and least severe criminal records but not those in the middle. Differences among the cutoffs in whether harsher guidelines move offenders from county jail to state prison may explain heterogeneity by criminal record. Finally, we show that the effect of harsher sentencing guidelines varies even among offenders with similar prior criminal record convicted of the same charge. Harsher guidelines can reduce recidivism more among offenders who are young, male, and convicted in the Southeast (Detroit Metro) region. In sum, how incarceration affects future crime depends on the characteristics of both the crime and the offender.

Consequently, appropriate public policy depends on knowing the diverse effects of sentencing guidelines. While the current political environment points toward reducing the severity of structured sentences, some form of sentencing guidelines will likely remain. The important next question regards how to re-design sentencing. As we demonstrate, harsher sentences generated by sentencing guidelines can reduce recidivism in some cases. However, in other cases harsher guidelines may not affect or even may increase recidivism while incurring a large financial cost for the state. Maximizing the public benefits of sentencing reform thus requires carefully identifying which offenses and offenders to target for reduced versus increased sentencing guidelines.

We also show that, in the presence of such heterogeneity, different empirical strategies can generate qualitatively different results because they estimate sentencing effects applicable to

different sets of offenders. As noted above, many high-quality studies use random judge assignment to identify the causal effect of sentence severity on recidivism. They identify a local average treatment effect (LATE) relevant for cases over which different judges would apply different sentences. Alternatively, our strategy examining discontinuities in sentencing guidelines puts weight on cases where sentencing guidelines bind and the offender's record is near the cutoff. Differences between the sets of compliers matter empirically. We implement a judge IV specification in the same sample and show that it provides qualitatively and quantitatively different estimates from our discontinuity approach. The results of different estimation strategies thus prove more informative for different criminal justice reform policies. Judge assignment IVs estimate a LATE that matches the effect of replacing judges or restricting one set of judges to follow the guidance of another set of judges. Discontinuities in sentencing guidelines, on the other hand, estimate the effect of adjusting eligibility for different sentencing guidelines at the margin. Of course, neither estimate is universally more informative because both estimates correspond to a particular policy and context. In the context of Michigan, the estimated effect of harsher sentencing on recidivism depends on the estimation strategy, and correct policy evaluation depends on attaching the appropriate estimate to the right policy.

More generally, our paper supports the idea that common empirical strategies estimate LATEs that correspond to particular groups of people and public policies. Imbens and Angrist (1994) document this point theoretically. A growing empirical literature clearly documents the link between estimation strategy and estimand. For prominent empirical questions, including how labor force participation responds to motherhood (Angrist and Evans, 1994; Lundborg et al., 2017)<sup>4</sup> and

---

<sup>4</sup>Angrist and Evans use twins or sex composition of existing children as instruments which estimates only the intensive margin, and Lundborg, et al. exploit the success of fertility treatment, thus estimating effects at the extensive margin.

how neighborhood quality shapes child outcomes (Ludwig et al. 2013; Jacob, Kapustin and Ludwig, 2015; Chyn, 2016; Aliprantis and Richter, 2016),<sup>5</sup> measured effects depend on the empirical strategy. Similar results have emerged within the economics of crime literature, though perhaps less frequently. Whether Mexican immigration to the U.S. (Chalfin, 2014; Chalfin and Deza, 2018) or letting teenagers out of school (Jacob and Lefgren, 2003; Billings and Phillips, 2017) affect crime rates depends on who the particular quasi-experiment under examination affects.<sup>6</sup> Thus, our study contributes to a growing awareness that the answer to an important empirical question can change when one considers a new source of variation.

## **1. Literature**

Our research contributes to the large literature examining the effects of incarceration on recidivism. In identification strategy, our research is closest to previous regression discontinuity (RD) papers but is most similar in scope to the literature exploiting other natural experiments, especially those using the random assignment of cases to judges as an instrumental variable (IV).<sup>7</sup>

### **1.1 RD Studies of Incarceration Effects**

A pair of papers studies how cutoffs in sentencing guidelines affect recidivism. Hjalmarsson (2009) exploits variation in juvenile sentences that results from a sentencing grid in Washington State. The grid assigns less severe local punishment or state incarceration according

---

<sup>5</sup> Both Ludwig et al. and Jacob, Kapustin, and Ludwig find housing voucher lotteries provide limited gains for children. But Chyn argues that these disappointing results are due to selection into the lottery, in particular that parents whose children would gain most do not sign up for the lotteries and are thus excluded from the estimation sample. Aliprantis and Richter, on the other hand, argue that housing vouchers in the Moving to Opportunity study actually change the treatment of interest, neighborhood quality, very little, and taking this fact into account reverses the finding that high quality neighborhoods do not affect adult economic outcomes.

<sup>6</sup> Whether US crime rates rise with the presence of Mexican immigrants depends on whether one uses rainfall shocks in Mexico (Chalfin) or the introduction of E-verify (Chalfin and Deza) to instrument for immigration. Having school in session can increase violent crime (Jacob and Lefgren), but this effect only occurs at schools with many high-risk peers and reverses at schools with few high-risk students (Billings and Phillips).

<sup>7</sup> For applications of difference in differences approaches to similar questions, see Abrams (2012), Kessler and Levitt (1999), Helland and Tabarrok (2007), and Owens (2009).

to criminal record categories formed from an underlying criminal record score. Her results indicate that incarceration reduces the probability that juvenile offenders reoffend. Our empirical approach is similar, but we focus on an adult population and can measure the sentence length. Most similar to our paper, a contemporaneous paper by Franco, et al. (2017) examines the same discontinuities in Michigan's adult felony sentencing guidelines that we study with similar data, though some choices in the empirical specification differ and they examine a shorter time frame for a broader set of offenses. Overall, they find that defendants just above a sentencing guidelines cutoff receive more severe sentencing regimes and recidivate less often. Franco, et al. (2017) differs from our analysis in that they focus on the average effect of prison and show that an original prison sentence can reduce crime by leading to re-imprisonment on technical violations. We focus new attention on how sentencing effects vary by context, offense, and offender. We show heterogeneity matters for interpreting the existing empirical literature and making public policy.<sup>8</sup>

Several papers on juvenile offenders in other states exploit sharp changes in sentence severity that occur at age 18 and find similar results. Guarin, Medina, and Tamayo (2013) and Loeffler and Grunwald (2015) conclude that more severe penalties reduce recidivism. Lee and McCrary (2017) also obtain negative estimates of sentence severity on recidivism, but with a magnitude so small as to lack much economic significance.

Hansen (2015), conversely, examines data on working-age adults and narrows his analysis to driving under the influence (DUI). Using the abrupt cutoff in blood alcohol content level (BAC) that distinguishes a standard DUI charge from an aggravated DUI in Washington State, Hansen finds that the harsher penalties imposed just above each of the BAC cutoffs reduce recidivism, qualitatively similar to the studies on younger populations. In short, extant RD studies suggest that

---

<sup>8</sup> Kuziemko (2013) also examines sharp changes in criminal justice guidelines, but she uses these to study the effect of parole decisions for inmates.



recidivism decreases with sentence severity for a variety of offenses among juvenile populations and for drunk driving among adults.

## **1.2 IV Studies of Incarceration Effects**

Another literature considers similar questions about the relationships between criminal punishment and future offense but utilizes random assignment of cases to judges. The random assignment of cases to judges allows researchers to gain traction from the varying proclivity of judges to assign harsher sentences.

In contrast to the related RD literature, the IV literature presents mixed results for the effect of incarceration. Aizer and Doyle (2015) consider the effects of juvenile incarceration with data from a large urban county in the United States where cases are randomly assigned to judges. They find that harsher penalties for juveniles lead to large increases in the likelihood of adult incarceration. Nagin and Snodgrass (2013) use a sample of all convicted criminal offenders from six Pennsylvania counties in one year, which results in a wider age range than Aizer and Doyle's. Green and Winik (2010) narrow their analysis to adults convicted of drug-related felonies in Washington, DC. Neither of these papers examining adult populations (which feature notably smaller sample sizes than Aizer and Doyle's and ours) finds evidence that sentence length impacts the likelihood of re-arrest. A more recent wave of research exploits the judge IV to focus on outcomes for the children of incarcerated parents (Dobbie et al. 2018; Bhuller et al. 2018) and the effects of pre-trial detention (Kleinberg, et al. 2017; Dobbie, Goldin and Yang, 2018). Most similar to the context of our analysis, Harding, et al. (2017) examine sentencing effects using random judge assignment in Michigan and find that prison time (compared to probation) increases future imprisonment within three years. Their contribution is largely to understand the role of technical

violations of probation as a factor in reincarceration, rather than the crime- and criminal-specific heterogeneity we explore.

The scope of this current research is perhaps best compared to Mueller-Smith (2016) which employs a large data set on multiple initial crimes, as we do. In his Harris County, Texas data, adult defendants are assigned randomly to courtrooms (staffed with judges and prosecutors). Mueller-Smith finds results for his adult population similar to those of the juvenile offenders examined by Aizer and Doyle. He finds specifically that increased sentence length elevates both frequency and severity of recidivism. We complement this literature by examining a similar question in a different context with an alternate source of variation in sentencing.

## **2. Theory and Implications for Estimation Strategies**

### **2.1. Sentences and Public Safety**

The term criminal justice stresses the role of law enforcement in assigning the proper penalty to an offender in order to satisfy society. Philosophies of punishment that emphasize justice, retribution, or restitution are by definition backward-looking, while the concern for public safety that currently dominates public discourse is forward-thinking. Similar to Becker (1968), we will consider a forward-looking policymaker whose goal is to maximize a wealth-neutral utilitarian welfare function  $W(\sum_i Y_i)$ , where  $Y_i$  represents outcomes for person  $i$  that affect overall welfare. For the exposition, suppose  $Y_i$  is a scalar that counts the number of future felony convictions. The severity of the sentence,  $S_i$ , handed down by courts may affect the number of future convictions in a linear manner such as

$$Y_i = \alpha + \beta_i S_i + f(X_i) + \omega_i + u_i \quad (1)$$

Outcomes may also depend on the defendant’s observable characteristics  $X_i$  (e.g., prior criminal record),  $\omega_i$  characteristics observed by the court but not the econometrician or the policymaker (defendant’s behavior in court), and unobservable characteristics  $u_i$  (motivation to rehabilitate).

Our research focuses on the coefficient  $\beta_i$ , the effect of increased sentence severity on public safety. Existing research suggests that the relationship between past sentencing and future crime is complex so that  $\beta_i$  will vary in both magnitude and direction from person to person. Longer terms of incarceration may result in more or fewer criminal acts due to three distinct mechanisms: incapacitation, general deterrence, and specific deterrence. Incapacitation may reduce future crimes because some crimes are more difficult to commit while in jail or prison. General deterrence refers to the ability of the mere threat of future expected incarceration to reduce crime. Specific deterrence encompasses a variety of factors that reflect the potentially lasting impact of incarceration on a former offender after release. For example, a former offender may update his assessment of the cost of incarceration; training and rehabilitation interventions in prisons might expand the former offender’s post-incarceration choice set or influence his utility function; incarceration may increase an individual’s criminal capital, e.g. by expanding his criminal network; or incarceration may reduce the set of labor market alternatives to crime. The extent to which these different mechanisms matter likely varies from person to person; hence, public policy needs to consider heterogeneity in the effects of sentencing on public safety.

## 2.2. A Simple Model of Sentence Assignment

The sentences that a court assigns to an offender depends on the preferences of the judge on the case, sentencing guidelines that constrain the judge, and the characteristics of the defendant. Suppose that each judge  $j$  has quadratic preferences regarding sentence length with the ideal sentence for defendant  $i$  defined as  $\tilde{S}_{ij} = g_j(X_i, \omega_i)$ . Judges may differ in their sentencing

behavior to the extent that  $g_j(\cdot)$  varies across judges; hence, the sentence depends on which judge is assigned to the case, which we denote by  $J_i$ .

State sentencing guidelines may limit judge discretion by prescribing maximum and minimum sentence severity,  $\bar{S}_i \geq \underline{S}_i$ , that may differ from person to person. For simplicity of exposition, suppose that sentencing guidelines are totally binding so that judges must assign a sentence in the prescribed range.<sup>9</sup> If the judge's preferred sentence falls in the relevant range, the sentence that judge  $j$  would assign is  $\tilde{S}_{ij}$ . If not, then the judge would assign the limit of the guideline because it is as close to her preferred sentence as possible. Stated differently, the full range of hypothetical sentences that different judges would assign, denoted  $S_{ij}$ , is

$$S_{ij} = \begin{cases} \underline{S}_i & \text{if } \tilde{S}_{ij} < \underline{S}_i \\ \tilde{S}_{ij} & \text{if } \underline{S}_i \leq \tilde{S}_{ij} \leq \bar{S}_i \\ \bar{S}_i & \text{if } \tilde{S}_{ij} > \bar{S}_i \end{cases} \quad (2)$$

Given judge assignment  $J_i$ , the actual sentence is assigned as  $S_i = S_{iJ_i}$  which may or may not be the same as the judge's preferred sentence,  $\tilde{S}_i = \tilde{S}_{iJ_i}$ .

### 2.3. Welfare Effects of a Change to Sentencing Policy

Suppose that policymakers want to know the potential welfare consequences of changing sentencing policy. Consider a policy  $\hat{X}$  that affects welfare only through changing the sentences received by defendants. The welfare effects of changing this policy are

$$\frac{\partial W}{\partial \hat{X}} = \frac{\partial W}{\partial \sum Y_i} \sum_i \beta_i \Delta S_i \quad (3)$$

---

<sup>9</sup> This assumption changes the exact effects of sentencing policy on individual sentences but not the overall conclusions regarding the parameters estimated by different empirical strategies.

How much this policy increases or decreases welfare depends on which offenders' sentences change and by how much. A policymaker who wishes to set optimal sentencing guidelines could treat this expression as a first order condition and change  $\hat{X}$  until (3) is zero.

Suppose that the current sentencing guidelines take a simple form; they divide offenders into groups based on prior criminal history,  $X_i$ , and assign different guidelines to offenders above and below a cutoff  $\hat{X}$  as follows:

$$\bar{S}_i = \begin{cases} \bar{S}^H & \text{if } X_i \geq \hat{X} \\ \bar{S}^L & \text{if } X_i < \hat{X} \end{cases} \quad \underline{S}_i = \begin{cases} \underline{S}^H & \text{if } X_i \geq \hat{X} \\ \underline{S}^L & \text{if } X_i < \hat{X} \end{cases} \quad (4)$$

where  $\bar{S}^H > \bar{S}^L$  and  $\underline{S}^H > \underline{S}^L$ . The cutoff between the groups,  $\hat{X}$ , provides a feasible policy lever that can be used to lower some sentences by moving some offenders into the lower guidelines group. Increasing the cutoff slightly will affect the sentences assigned to some defendants as follows:

$$\Delta S_i|_{X_i=\hat{X}} = \begin{cases} -(\underline{S}^H - \underline{S}^L) & \text{if } \tilde{S}_i \leq \underline{S}^L \\ -(\underline{S}^H - \tilde{S}_i) & \text{if } \underline{S}^L \leq \tilde{S}_i < \underline{S}^H \\ 0 & \text{if } \underline{S}^H \leq \tilde{S}_i < \bar{S}^L \\ -(\tilde{S}_i - \bar{S}^L) & \text{if } \bar{S}^L \leq \tilde{S}_i < \bar{S}^H \\ -(\bar{S}^H - \bar{S}^L) & \text{if } \tilde{S}_i \geq \bar{S}^H \end{cases} \quad \text{and} \quad \Delta S_i|_{X_i \neq \hat{X}} = 0 \quad (5)$$

For concreteness, the Michigan felony sentencing guidelines assign different guidelines to Class E (a mid-level category of offenses on a scale ranging from A to H, that is high to low) offenders with at least 25 prior record points versus those with less than 25 points. If a policy moved the boundary to 26, offenders with 25 points would move to the lower group, and the applicable sentencing guidelines for those offenders would decrease from a range of 5-23 months to a range of 0-9 months. The changes in sentences would be

$$\Delta S_i|_{X_i=25} = \begin{cases} -5 & \text{if } \tilde{S}_i = 0 \\ -(5 - \tilde{S}_i) & \text{if } 0 \leq \tilde{S}_i < 5 \\ 0 & \text{if } 5 \leq \tilde{S}_i < 9 \\ -(\tilde{S}_i - 9) & \text{if } 9 \leq \tilde{S}_i < 23 \\ -14 & \text{if } \tilde{S}_i \geq 23 \end{cases} \quad \Delta S_i|_{X_i \neq 25} = 0$$

Whether and how much an offender's sentence changes in response to the policy depends on the offender's prior criminal history; the change in categorization only affects offenders with a criminal history just at the 25-point cutoff. Additionally, the relationship of the judge's preferred sentence to the guidelines matters. The policy change most affects sentencing for extreme cases. If the unconstrained judge would assign a sentence of either 0 months or more than 23 months, then the guidelines move sentences by the full difference between old and new guidelines. For an intermediate group of defendants, the sentence still falls but not by as much. If the judge wishes to issue a sentence between 0 and 5 months, these are now viable options and the sentence falls from 5 months to the judge's preferred sentence. Finally, in some cases, the guidelines overlap, and the policy change does not matter. A judge wishing to issue a sentence between 5 and 9 months can do so with either set of guidelines.

#### 2.4. Using Discontinuities in Sentencing Guidelines to Estimate Policy-Relevant Effects

A regression discontinuity design naturally estimates the effects of changing the sentencing guidelines cutoff.

$$Y_i = \alpha + \beta^{RD} I\{X_i \geq \hat{X}\} + \rho_1(X_i - \hat{X}) + I\{X_i \geq \hat{X}\} * \rho_2(X_i - \hat{X}) + \epsilon_i \quad (6)$$

$I(\cdot)$  is an indicator function;  $\rho_1(\cdot)$  and  $\rho_2(\cdot)$  are polynomials;  $\epsilon_i$  is a composite of the determinants of sentencing and recidivism not considered by the sentencing guidelines; and all other variables are as defined above. As is well known, a regression discontinuity design estimates causal effects in the immediate neighborhood of the cutoff. In a sample, regression discontinuity estimates the limit of the difference in means just above and just below the cutoff.

$$\hat{\beta}^{RD} = \lim_{X \downarrow \hat{X}} \bar{Y} - \lim_{X \uparrow \hat{X}} \bar{Y} = \frac{1}{N} \sum_i \beta_i \left( \lim_{X \downarrow \hat{X}} S_i - \lim_{X \uparrow \hat{X}} S_i \right) = -\frac{1}{N} \sum_i \beta_i \Delta S_i \quad (7)$$

The second equality follows from equation (1) and the final equality defines  $\Delta S_i$  as above. Subject to a simple re-scaling, the regression discontinuity design estimates the quantity needed for welfare calculations in equation (3), provided the policymaker is prepared to make normative claims about the contribution of the outcome to welfare,  $\frac{\partial W}{\partial \Sigma Y_i}$ .

The regression discontinuity design can also be used to identify an instrumental variables estimate of the effect of sentences on outcomes.

$$S_i = \alpha_0 + \beta^{RD} I\{X_i \geq \hat{X}\} + \rho_1(X_i - \hat{X}) + I\{X_i \geq \hat{X}\} * \rho_2(X_i - \hat{X}) + \epsilon_i \quad (8)$$

$$Y_i = \alpha_1 + \beta^{IV} S_i + \rho_3(X_i - \hat{X}) + I\{X_i \geq \hat{X}\} * \rho_4(X_i - \hat{X}) + v_i \quad (9)$$

This specification will measure the effect of sentence severity on future crime, averaging over defendants for whom a change in sentencing guidelines actually leads to different sentences. In standard instrumental variables language (Imbens and Angrist, 1994), we estimate a local average treatment effect (LATE) for the set of compliers who receive different sentences in response to the guidelines, as defined by equation (5). Importantly, these compliers are exactly the same people who would be affected by a small change in sentencing policy. A regression discontinuity design measures the marginal effect of a feasible policy change.

## 2.5. Comparison to Empirical Strategies Based on Random Judge Assignment

An empirical strategy based on random judge assignment measures the effect of harsher sentences for a different population. A simple judge-IV estimates the following specification:

$$S_i = \alpha_1 + \psi \bar{S}_{(i)} + Z_i \theta_1 + \epsilon_i \quad (10)$$

$$Y_i = \alpha_2 + \beta^{IVJ} S_i + Z_i \theta_2 + v_i \quad (11)$$

$\bar{S}_{(i)}$  is the average sentence assigned by  $i$ 's judge to other cases not involving  $i$ , and judges are assigned randomly to courtrooms conditional on observable variables  $Z_i$ . The coefficient  $\beta^{IVJ}$  also measures a LATE of harsher sentences on the outcome.

The population of compliers for a random judge assignment IV differs significantly from compliers with a sentencing guidelines discontinuity instrument. Most simply and significantly, defendants with prior criminal history away from the cutoff,  $X_i \neq \hat{X}$ , may be compliers for a judge IV. Even among defendants with prior criminal history near the cutoff, random judge assignment and discontinuity instruments will emphasize different populations of compliers. For simplicity, suppose that the data include only two judges,  $M$  and  $N$ , such that judge  $M$  always wishes to assign a harsher sentence,  $\tilde{S}_{iM} > \tilde{S}_{iN} \forall i$ , and recall the example from above with guidelines of 5 to 23 months. Suppose that judges  $M$  and  $N$  wish to assign a defendant to 36 months and 24 months, respectively. Because of the sentencing laws, both judges actually assign the defendant to 23 months. Hence, the LATE from a judge IV puts no weight on this defendant. However, if the defendant was sentenced under the more lenient guidelines of 0 to 9 months, both judges would lower their sentences to 9 months. Sentencing guidelines matter for this defendant but judge assignment does not. Conversely, consider a middle-of-the-road defendant for whom judges  $M$  and  $N$  wish to select sentences of 9 and 5 months, respectively. The sentence for this defendant varies with judge assignment but not sentencing regime.

In general, the judge IV computes a LATE for cases where judges both differ in their preferred sentences and have freedom to choose. It ignores offenders for whom judges would have applied the same sentence, including those for whom the guidelines bind regardless of the judge. On the other hand, a discontinuity-based LATE puts weight on cases near the cutoff where the guidelines bind.



These different LATEs correspond to different policy changes. The judge-assignment LATE measures the effect of replacing all severe judges with lenient judges (or vice versa) or, equivalently, totally removing judicial discretion by collapsing sentencing guidelines down to the sentence applied by the strict judge. Note, though, that the judge assignment LATE does not measure the effect of expanding judicial discretion, since widening the sentencing guideline range would affect people for whom the current guidelines bind regardless of the judge. On the other hand, the discontinuity-based estimates measure the effect of marginally expanding or contracting who falls under harsher sentencing guidelines. Of course, both estimates depend on context because judge preferences and the details of sentencing guidelines vary. To the extent that these two estimates differ within the same context, the latter will more directly inform legislators who can set the rules governing who receives which sentencing regime but cannot dictate individual sentences.

### **3. Data and Institutional Background**

#### **3.1. Michigan Department of Corrections Sentencing Data**

This paper uses administrative felony sentencing data from the Michigan Department of Corrections (MDOC). MDOC maintains a database named OMNI that includes sentencing information for all felony convictions in Michigan. The data record limited demographics (e.g., current age), extensive prior criminal history, characteristics of the sentence assigned, and information on state sentencing guidelines applying to the particular case. Most of these variables derive from a pre-sentencing report. After a felony case has been disposed with a guilty (or similar) verdict, an employee of MDOC composes this report to inform the court on the prior criminal record of the offender and other facts relevant to the state's sentencing guidelines. Because an employee of the state government rather than the local district court completes this report, the

OMNI database covers all felony cases in the state; however, this database includes only felonies because misdemeanors do not fall under the state's sentencing guidelines and hence do not require a pre-sentencing report.

As mentioned above, from this database, we analyze a pre-existing extract of data (selected prior to the inception of our research by MDOC officials for administrative purposes) which covers two common mid-level offenses: Retail Fraud – 1<sup>st</sup> Degree and Operating While Intoxicated, 3<sup>rd</sup> offense or greater. Retail Fraud entails shoplifting, i.e. stealing from a store when it is open, and similar activities. A Retail Fraud case becomes classified as first degree, rather than either of two lesser categories, if the amount stolen is greater than \$1,000 or if the offender has prior Retail Fraud convictions. OWI – 3<sup>rd</sup> is Michigan's most common felony drunk driving charge. An individual's first and second OWI convictions result in misdemeanor charges, but any subsequent OWI offense results in a felony charge for OWI – 3<sup>rd</sup>. Michigan law classifies both Retail Fraud – 1<sup>st</sup> Degree and OWI-3<sup>rd</sup> as mid-level class E offenses (on a scale of decreasing severity from A to H). We use an extract from the OMNI database provided by MDOC that starts with a set of base cases sentenced between January 2003 and July 2013 for which Retail Fraud – 1<sup>st</sup> Degree and OWI – 3<sup>rd</sup> are the sole charge. The extract then includes all other felony convictions for those same people, both those sentenced to incarceration and those not, in the same time period, i.e. cases matched by State Identification Number (SID) to a base case. We limit our sample to base cases sentenced by July 2008 so that we can measure 5-year recidivism. This results in a full sample of 14,804 individual sentencing observations covering 10,915 unique people. The full sample breaks into 6,662 OWI – 3<sup>rd</sup> sentences, 2,435 Retail Fraud – 1<sup>st</sup> Degree cases, and 5,707 other cases.

While this paper is constrained by the availability of data, focusing on Retail Fraud and OWI cases is still conducive to our research task in that they are representative of common, non-

violent felonies. They are the 23<sup>rd</sup> and 11<sup>th</sup> most common felony charges, respectively, in Michigan out of about 800 different charges, and – like 7 of the 10 most common felonies in Michigan – they focus on offenses related to substance abuse or property crime.<sup>10</sup> They are also distinct felony charges treated similarly by state sentencing guidelines. Drunk driving is a public safety charge potentially related to addictive behavior. Retail Fraud is a property crime perhaps more likely to respond to rational cost-benefit calculations. The people committing these crimes are also quite different. Table 1 shows summary statistics for our sample. The first two columns show means and standard deviations of relevant variables for Retail Fraud and OWI cases. Compared to those convicted of OWI, people convicted of Retail Fraud are almost twice as likely to be sentenced in the Southeast (Detroit Metro) region of the state and have much more extensive criminal records. We can compute prior criminal record in the data in two ways. First, the data directly records MDOC’s summary measures of number of prior misdemeanors and felonies. Second, we can calculate prior felonies within the range of the data extract we have available (i.e., since 2003). Both of these measures indicate that Retail Fraud offenders have more extensive criminal histories, on average, than repeat OWI offenders. Nonetheless, state law classifies both of these offenses as mid-level Class E offenses subject to the same sentencing guidelines. Thus, Retail Fraud and OWI provide a useful context to test how changes in sentencing regimes affect criminal activity in potentially different ways for different crimes.

We compute recidivism in the data similarly by linking individuals across time. MDOC uses SIDs administratively to track people over time, e.g., identifying people with prior convictions. Hence, they provide a high-quality way for us to match people over time in the data.

---

<sup>10</sup> The other three are violent and/or weapons charges. The top ten are, in order, possession of narcotics, delivery of narcotics, assault with a dangerous weapon, delivery of marijuana, felony firearms, uttering & publishing, concealed carry, larceny in a building, breaking and entering a building with intent, and home invasion 2<sup>nd</sup> degree.

Given a particular case with a particular SID, we count the number of future felony sentences within 5 years after the original sentencing date for the same SID. We use this as our measure of recidivism. Table 1 reports averages for our sample. In addition to having a more extensive criminal history, Retail Fraud offenders also commit more future crimes, 1.43 future sentences compared to 0.51 for OWI offenders.

The data also provide information on crimes other than Retail Fraud and OWI committed by the same offenders. While these offenses do not provide a fully representative sample of all offenses committed in Michigan, they provide a broader group of crimes to which we can compare Retail Fraud and OWI offenses. The third and fourth columns of Table 1 report summary statistics for other class E and non-class E offenses, respectively. For other class E offenses, offenders in our sample receive an average sentence of 213 days incarceration and 743 days of probation, which is between Retail Fraud and OWI offenses. The non-class E offenses have an average sentence of 268 days incarceration and 1004 days probation. In addition to including these other crimes in our measures of past and future offending, we will occasionally compare results with these samples to the Retail Fraud and OWI samples to give a sense of the broader context in which these two crimes are sentenced.

### **3.2. Felony Sentencing and Sentencing Guidelines in Michigan**

The state of Michigan uses a felony sentencing structure that combines some elements similar to other US states with features unique to Michigan.<sup>11</sup> As in many states, the sentencing process is indeterminate. A court sets the *uttered sentence* which defines the minimum time of incarceration, while state statutes define the maximum sentence. The length of the uttered sentence also affects the type of incarceration with terms longer than one year served in the more severe,

---

<sup>11</sup> The following description relies on extensive communication with Michigan Department of Corrections staff as well as two outside reports: Council of State Governments (2009) and Council of State Governments (2014).

standardized environment of state prisons while offenders serve terms less than one year in county jails that vary more by jurisdiction. The court also sets a probation period and any fees. Table 1 shows these variables for our sample. Retail Fraud and OWI receive uttered sentences of 217 and 175 days incarceration, on average, respectively. OWI offenses do receive significantly longer probation periods, though, on average. The statutory maximum time in prison for both crimes is 60 months, which is reflected in our data. Thus, the typical offender we examine will spend between 6 and 60 months incarcerated. A parole board then has discretion within that range regarding when to release the prisoner, though prisoners are typically released close to the minimum sentence.

We do not observe actual time served in our data, which makes some distinction between uttered jail and prison sentences important. As noted above, sentences longer than one year lead to the qualitatively different experience of state prison. Because of “truth in sentencing” laws, uttered prison sentences clearly indicate the time before which almost no offenders will be released. For the crimes we study, actual time served in prison tends to closely reflect the minimum sentence. On average, Michigan prisoners serve 127 percent of the minimum sentence (Council of State Governments, 2009). For county jails, though, the relationship between uttered sentences and time served varies more significantly due to differences in discretion across counties and conditions allowing for early release (e.g., crowding, “good time”), and offenders assigned to jail frequently serve less time than the uttered sentence. We will use the minimum sentence which we observe as an approximation of the length of incarceration, and this approximation will be most accurate for prison sentences.

Sentencing guidelines constrain the range of the uttered sentence that can be selected by the court for a given offender. The guidelines set a lower and upper limit for the uttered minimum

sentence. During our sample<sup>12</sup>, these sentencing guidelines are “presumptive,” meaning that offenders are presumed to be sentenced within that range. Judges can elect a “departure” sentence outside the range, though this is not the norm and has some cost (e.g., makes the case more likely to be subject to appeal). Table 1 shows average sentencing constraints for the cases we consider. On average, the sentencing guidelines prescribe a sentence between 5 and 23 months for Retail Fraud offenders and 2 and 14 months for OWI offenders. OWI offenders have less stringent recommendations because of their, on average, more limited criminal history (see below). As expected, the average sentence actually applied by the court, 6 to 7 months, falls within the range of the guidelines.

The sentence limits prescribed by the guidelines depend mechanically on attributes of the offense and offender. The lower limit is a function of the class of the offense, an offense variable total capturing the severity of the offense, and a prior record total capturing the offender’s criminal history. The prior record total will be particularly important for our analysis. The MDOC employee conducting the pre-sentence report scores seven variables based on the offender’s prior convictions. For example, one variable assigns 0 points to someone with no prior class E through H felony convictions. It then assigns 5 points, 10 points, 20 points, and 30 points for 1, 2, 3, or 4+ such prior E through H felony convictions, respectively. The prior record total then sums these seven underlying variables. Table 1 lists the prior record total for our data. Matching the previous data on prior crimes committed, the sentencing guidelines assign Retail Fraud offenders 33 prior record points but OWI offenders only 15 points. The offense variable total similarly sums 20

---

<sup>12</sup> The existing guidelines system was ruled to be unconstitutional in 2015 by the Michigan Supreme Court in *People v Lockridge* and became non-binding. This event occurred after our sample period, creating a difference between the practice of sentencing during our sample period versus the present. Sentencing guidelines are still relevant in Michigan to the extent that advisory structured sentences influence sentencing as previous research has suggested in other states. The results of our study will also be useful for contexts where sentencing guidelines remain presumptive.

underlying variables, now related to the nature of the offense. A grid, described in more detail below, then assigns the lower limit of the sentencing guidelines based on the class of the offense, the prior record total, and the offense variable total. The upper limit of the sentencing guideline range depends on the prior record and offense totals in a similar manner, but it also considers a habitual offender indicator scored with four levels. The MDOC administrative data includes the exact values of each of these determinants of the sentencing guidelines which will be important for the empirical strategy we develop below.

#### **4. Empirical Strategy**

##### **4.1. Adapting a Regression Discontinuity Framework to Michigan's Guidelines**

Simply comparing outcomes for offenders under harsher versus laxer sentences will not generally measure the causal effect of changing sentencing guidelines of an individual offender's sentence. The law explicitly links prior offending to more severe sentencing guideline ranges, which will generate a positive correlation between recommended sentences, actual sentences, and recidivism. Furthermore, the tendency of courts to assign harsher sentences to those prone to re-offending for reasons unobservable to the econometrician will likewise generate spurious positive correlation between sentences and recidivism. It follows that measuring the causal effect of sentencing guidelines and sentences on re-offending requires a more careful econometric framework.

The statutory sentencing guidelines map criminal history and offense variables into ranges of possible sentence in a discontinuous manner. Figure 1 displays the sentencing grid for class E offenses, which includes both OWI – 3<sup>rd</sup> Offense and Retail Fraud – 1<sup>st</sup> Degree. For example, an offender with 30 prior record points, 15 offense points, and no habitual offender tag receives a prior record level of D and an offense level of II, which maps to sentencing range of 7 months to

23 months. Thus, the sentencing guidelines truncate continuous offense and prior record totals into a discrete number of sentencing cells with different sentencing guidelines governing different cells. For the prior record total, cells of 0, 1-9, 10-24, 25-49, 50-74, and 75+ points receive potentially different guidelines. The vast majority of cases in our sample reside in lowest offense levels (I, II, or III). For these cases, mandatory minimums only change at the 25-, 50-, and 75-point cutoffs. Hence, we focus on four categories: 0-24, 25-49, 50-74, and 75+.

The empirical patterns in Figure 2 clearly match the *de jure* sentencing guidelines from Figure 1. Figure 2.A. uses sentencing data to display the discontinuity in the guidelines at 25, 50, and 75 prior record points for Retail Fraud offenders. The left and right images show these jumps for the lower and upper limits of the sentencing range, respectively. Within a sentencing guideline grid cell, the guidelines remain constant.<sup>13</sup> However, when the prior record total crosses the boundary between two cells, the minimum sentence jumps. For example, the guidelines do not require most offenders with prior record total between 20 and 24 to receive any time in prison or jail, but the average person with a prior record total of 25 to 29 must be incarcerated at least 6 months. Smaller discrete jumps in the mandatory minimum sentence appear at the 50- and 75-point cutoffs. The maximum end of the sentencing guidelines also jumps at the 25-point cutoff, though not at the other two cutoffs. So, crossing the 25-point cutoff shifts the entire range upward, but crossing the 50- and 75-point cutoffs only increases the lower bound. Figure 2.B. shows a similar pattern of discontinuities for OWI cases.

We will focus on the discontinuities in sentencing at 25, 50, and 75 prior record total points. Michigan's sentencing guidelines feature a number of other discontinuities. We ignore the discontinuities at 1 and 10 prior record points because, as shown in Figures 1 and 2, the actual

---

<sup>13</sup> In our data, the sentences mandated by the guidelines increase slightly even within cells because offenders with higher prior record total are more likely to also have higher offense variable totals



guidelines do not move much at these points. The guidelines also employ offense levels determined by discretizing an underlying offense total, but the offense variable total is reportedly more subjective in its construction and does not generate large discontinuities in practice. Thus, we will exploit the variation in Michigan’s felony sentencing guidelines around 25, 50, and 75 prior record points.

This institutional context naturally maps into the regression discontinuity design described above. We use the prior record total as the running variable  $X_i$  with cutoffs at 25, 50, and 75. Equation (6) can be adapted to examine multiple cutoffs:

$$Y_i = \alpha + \rho_0(X_i) + \sum_{k \in \{25, 50, 75\}} [\beta_k^{RD} I\{X_i \geq k\} + I\{X_i \geq k\} * \rho_k(X_i - k)] + \epsilon_i \quad (6')$$

The coefficients of interest,  $\beta_k^{RD}$ , measure the jump in outcomes at each cutoff  $k$ . The set of polynomials allow outcomes to respond flexibly to the prior record total within each sentencing range. In our main specifications, we control for quadratic polynomials in the running variable and limit the sample to offenders with prior record totals no more than 100 points. This specification can be interpreted as a set of local quadratic regressions with a bandwidth of 25.<sup>14</sup> We eliminate cases far from the discontinuities since our goal is to measure the jump at the discontinuity and because polynomial regressions can be sensitive to outliers. For robustness, all of the figures also display results from a specification that allows for cutoffs at 1, 10, 25, 50, and 75 and linear control for the running variable within each window. As shown in Figure 2, this alternate specification

---

<sup>14</sup> We also modify the prior record total slightly, removing the first of the 7 prior record variables from the total. This variable counts points for prior severe felonies in increments of exactly 25 points. Since these increments exactly match the cutoff, we remove this variation to avoid placing a disproportionate number of offenders with severe felonies just above the cutoffs. Because most cutoffs in the guidelines are multiples of 25, removing 25, 50, or 75 points from prior high severity felony convictions simply moves offenders from a higher cutoff. For example, two offenders with actual prior record totals of 49 and 50 reside near the 50 cutoff. If each receives 25 points from prior severe felonies, we will omit those 25 points, move these individuals to 24 and 25 points, which still leaves them on opposing sides of a sentencing grid cutoff.

generates similar discontinuities in the sentencing guidelines. Finally, we can average results across multiple cutoffs. Consider a regression of outcomes on the number of cutoffs exceeded:

$$Y_i = \alpha + \rho_0(X_i) + \beta^{RD} \sum_{k \in \{25, 50, 75\}} I\{X_i \geq k\} + \sum_{k \in \{25, 50, 75\}} [I\{X_i \geq k\} * \rho_k(X_i - k)] + \epsilon_i \quad (6'')$$

This regression is identical to equation (6') except it forces the treatment effect coefficients,  $\beta_k^{RD}$ , to be equal.

Table 2 shows the results of this specification inserting the sentencing guidelines minimum and maximum as outcomes. The upper pane shows results averaging across all three cutoffs as in equation (6''). Columns (1) and (5) examine together those originally charged with Retail Fraud or OWI. Offenders with prior record totals just above the cutoff face lower and upper limits of the sentencing guideline range that are, respectively, 4.70 and 7.95 months larger than similar offenders just below the cutoff. The bottom pane splits out these results across the three cutoffs, as in equation (6'). The guidelines are most sensitive at the lowest cutoff. At 25 points, the lower limit of the sentencing guideline range jumps by 6 months, compared to 2 months at the 50- and 75-point cutoffs. The upper limit also increases by 11 months at the 25-point cutoff but does not obviously increase at the 50- and 75-point cutoffs. As shown in the other columns of Table 2, results are qualitatively similar when we split the sample into offenders convicted of Retail Fraud, OWI, or some other offense.<sup>15</sup> Overall, we precisely measure a large jump in the sentencing guidelines applying to cases around the cutoffs in prior record total.

Finally, we can use these effects on sentencing as the first stage of an instrumental variables specification that adapts equations (8) and (9) to have multiple cutoffs:

---

<sup>15</sup> We have data on other offenses only for offenders who also commit OWI or Retail Fraud at some point in our time period. Since being in the sample depends on OWI or Retail Fraud offending, we can examine differences in the sentencing guidelines but not recidivism for these other offenses.

$$S_i = \alpha_1 + \phi_0(X_i) + \sum_{k \in \{25, 50, 75\}} [\gamma_k I\{X_i \geq k\} + I\{X_i \geq k\} * \phi_k(X_i - k)] + \epsilon_i \quad (8')$$

$$Y_i = \alpha_2 + \beta^{IV} S_i + \rho_0(X_i) + \sum_{k \in \{25, 50, 75\}} [I\{X_i \geq k\} * \rho_k(X_i - k)] + v_i \quad (9')$$

The excluded discontinuities,  $I\{X_i \geq k\}$ , serve as instruments for sentence severity  $S_i$ .

## 4.2. Validity of the Regression Discontinuity Strategy

Our empirical strategy requires that offenders on either side of sentencing guideline discontinuities would experience similar outcomes in the absence of the effect of the guidelines. As noted above, the guidelines assign offenders to one side of the discontinuity or another based on the prior record and offense totals. By the nature of the guidelines, these variables are most often assigned in multiples of 5. Figure 3 demonstrates this heaping on multiples of 5. Such lumpiness in the running variable prevents us from using local linear regression with optimal bandwidth selection, and the specification defined in equations (6) through (9) may thus be best interpreted as a parametric regression discontinuity framework. We report recidivism results for different functional forms and bandwidths in Appendix Table 1. The robustness of our results both graphically and in response to different functional forms justifies our approach.

The state exogenously imposes sentencing guideline cutoffs, but offenders, prosecutors, and judges could sort around the cutoffs through selective recidivism, charging, and plea bargaining. Because of the discrete running variable, we cannot apply standard tests for the continuity of the distribution of the running variable (e.g., McCrary, 2008). Visual inspection of the distribution of the running variable in Figure 3 suggests that any manipulation is small, though we still wish to ensure that sorting across sentencing boundaries does not contaminate our results. Thus, we check for balance in offender characteristics around the cutoff. Figure 4 provides reassurance that any sorting across boundaries does not introduce bias into our identification

strategy. These figures test for discontinuities in four pre-determined variables: age at offense, number of prior felonies and number of prior misdemeanors as measured by MDOC, and the lagged value of our recidivism measure, that is, the number of sentences for the individual in our dataset prior to the present sentence.<sup>16</sup> Figure 4.A shows the results for Retail Fraud and 4.B for OWI. For both groups of offenders, baseline variables show no clear tendency to jump at the sentencing guideline cutoff at 25 points. Some gaps may appear at the 50- and 75-point cutoffs, though the graphs are quite noisy due to small samples. More formally, we can test for baseline balance by implementing regression (6'') using various predetermined baseline characteristics. Table 3 quantifies these results. As shown in Column (1), this test registers no statistically significant differences in prior felonies (as measured by the court), prior misdemeanors, or age at offense for the combined sample. The number of prior felonies (as measured in our sample) is statistically different. This would be concerning except, as shown in Appendix Table 2, there is no imbalance for the discontinuity at 25 points, which drives our main results. For Retail Fraud alone, there is some imbalance in age at offense which comes from the 25-point cutoff; however, Figure 4.A suggests that this estimate is an artefact of the quadratic functional form. Any sorting around sentencing guideline boundaries does not appear to be related to recidivism for Retail Fraud and OWI offenders. Discontinuities in the sentencing grid appear to be a valid way to measure the causal effect of sentencing guidelines on uttered sentences and future criminal behavior.

## **5. Main Results**

### **5.1. Sentences**

---

<sup>16</sup> Prior felonies do not equal prior sentences in our data because we have a shorter dataset than that available to MDOC.

We first test for whether more severe sentencing guidelines lead to more severe sentences.<sup>17</sup> In Michigan, uttered sentences reflect the mandatory minimum time served as imposed by the court. Sentencing guideline regimes could have no effect on uttered sentences if the law draws sentencing ranges so widely that they do not constrain judges or if departing from the guidelines is low cost and common. We thus test for whether Michigan's sentencing guidelines actually constrain sentences. Figure 5 provides a first piece of graphical evidence that the guidelines do matter for sentencing. In Figure 5.A, each histogram shows the distribution of uttered sentences relative to the sentencing guidelines' limits for a different offense. For OWI and Retail Fraud, the vast majority of sentences assigned by courts fall within the limits set by the sentencing guidelines, and sentences tend to heap on the upper and lower limits of the sentencing range. Similar features can be observed for other class E felonies and felonies in other classes. Figure 5.B examines the guidelines cutoff at 25 prior record points. When an offender barely qualifies for harsher guidelines, judges more commonly issue departures for sentences below the guidelines' range. However, sentences still heap on the lower limit, suggesting that the sharp change in guidelines with respect to prior criminal record binds for some offenders. These facts suggest that

---

<sup>17</sup> The question of whether structured sentences (e.g., mandatory minimums, presumptive sentences, guidelines) bind is of particular importance to our research in light of our intended contribution, shedding light on a practical policy lever. The existing literature on the effects of sentencing regimes on sentences provides mixed results. For example, a handful of papers consider the influence of federal mandatory minimums and report different rates of departure. Bjerk (2014) finds that between October 2010 and September 2012 less than half of drug trafficking offenders nationwide receive sentences consistent with mandated minimums. Lacasse and Payne (1999) examine an earlier time period (1981-1995) and a broader set of eligible crimes in only two districts in New York and find that after the imposition of mandatory minimum policies in 1987, the variance in sentences increased for eligible crimes relative to ineligible crime types. That is, mandatory minimums seem to matter, but not in the expected way. Considering a similar time period and the same policy shift, Anderson, Kling, and Stith (1999) use national data and narrow their focus to the portion of sentence variation that can be attributed to judges. The legal strength behind sentencing guidelines and particularly the costs judges face should they depart from them can vary by jurisdiction. While fewer papers consider state-level sentencing guidelines, Bushway, Owens, and Piehl (2012) employ circuit court data from Maryland to conclude that the state's guidelines do exert influence on a judge's uttered sentence but not time served. However, because less than 50 percent of sentences fall within the recommended range and Maryland's sentencing guidelines have never imposed legal requirements upon judges who wish to depart from them, these guidelines can be considered advisory in nature, which differs from our sample.

the guidelines constrain the sentencing process, pushing judges to assign sentences within the guideline range.

We can also test whether discontinuities in the sentencing guidelines with respect to prior criminal record translate through to uttered sentences. Figure 6.A. displays the results of applying the empirical framework in equation (6) to uttered sentences for Retail Fraud offenders. The left image demonstrates that the number of incarceration days in the uttered sentences jumps upward at the cutoffs in the same manner as the sentencing guidelines. Figure 6.B. shows similar, though noisier, results for OWI sentences. Column (1) of Table 4 quantifies these effects for the two crime types combined. The first row averages over all three cutoffs. Offenders facing harsher sentencing guidelines receive uttered sentences with 42 more days of incarceration, on average, than similar offenders facing laxer sentencing guidelines. The bottom pane shows that increased incarceration appears at all three cutoffs, though the 75-point cutoff is noisy due to a small sample. As shown in columns (2) and (3), the guidelines generate discontinuities in sentencing for both Retail Fraud and OWI, though the effects for OWI are somewhat smaller and less statistically significant. The final three columns of Table 4 and the righthand images in Figure 6 suggest that courts may offset longer incarceration with shorter probation. As shown in Column (4), for example, offenders sentenced to harsher guidelines receive 46 fewer probation days. This point estimate is very noisy, but we cannot reject the hypothesis that sentencing guidelines leave the combined sentence length unchanged. Overall, we find that being subject to more severe sentencing guidelines changes the mix of the average sentence, increasing incarceration.

The increase in incarceration caused by the sentencing guidelines includes changes in the type of incarceration. In Michigan, offenders serve sentences less than one year in county jail, and actual time served may vary considerably depending on differences across counties, crowding

concerns, etc. Sentences longer than one year direct the offender to the more severe and standardized environment of state prison where time served tends to follow the uttered sentence. Table 5 demonstrates how sentencing guidelines affect the type of incarceration in addition to its length. Column (1) repeats the results of Table 4, showing that harsher sentencing guidelines increase incarceration. Focusing on the top row, Columns (2) and (3) demonstrate that the increase in incarceration reflects a combination of state prison and county jail time, with jail increasing by 17 days and prison increasing by 26 days, on average. As shown in columns (4) and (5), prison increases on the extensive margin as 3.8 percentage points of sentences shift from either jail or no incarceration to prison. Column (6) shows that jail increases on the intensive margin. For those assigned to jail, duration increases by 33 days on average, despite selection effects that move the most severe offenders in jail to prison. We cannot say definitely, though, whether prison time increases on the intensive margin because the new arrivals in prison are among the least severe cases. The positive but statistically insignificant coefficient in column (7) combines any increase in prison sentence lengths with this dilution in the severity of offenders sent to prison. Overall, increased incarceration includes both longer jail terms and a shift from jail to prison.

How the sentencing guidelines affect incarceration differs qualitatively across the three cutoffs. The bottom pane of Table 5 demonstrates this fact. Column (1) shows that incarceration increases 38, 46, and 85 days respectively at the 25-, 50- and 75-point cutoffs; however, columns (2) through (7) show that the composition of this increase varies. Crossing the 25-point cutoff both intensifies jail and moves offenders from jail to prison, but the other cutoffs only do one or the other. Crossing the 50-point cutoff only intensifies jail, but the 75-point threshold solely moves people from jail to prison. These empirical facts match the legal guidelines in Figure 1. Recall that prison occurs for sentences greater than 12 months. For the most common group in our data (OV

level II, not habitual) the 25-point cutoff both institutes a minimum jail sentence and makes prison possible. The 50-point cutoff only pushes jail sentences of 5-9 months up to 10 months. The 75-point cutoff makes a 12 month prison sentence the minimum. Overall, the sentencing guidelines increase incarceration at each of the three cutoffs, but they do so in different ways. The 50-point cutoff primarily increases jail, the 75-point cutoff primarily moves people from jail to prison, and the 25-point cutoff does both.

## **5.2. Recidivism**

The effect of harsher sentencing guidelines and uttered sentences on recidivism depends on the original offense. We measure recidivism as the number of future felony sentences observed in our data. Figures 7.A. and 7.B. display the graphical results for Retail Fraud and OWI, respectively. As expected, offenders with more severe prior criminal records receive more future felony sentences, leading to the upward sloping graph in both cases. For Retail Fraud cases, though, recidivism jumps downward when prior record totals cross the 25-point threshold that leads to more severe sentencing guidelines and uttered sentences. Because of the more severe penalties, Retail Fraud offenders with prior record totals just above the cutoff recidivate at a lower rate than those just below the cutoff. This effect is large. Offenders with a prior record total between 25 and 29 have committed 2 to 3 more prior felonies or 5 to 6 more prior misdemeanors than offenders with prior record totals between 10 and 14, but because of more severe sentences applied to those with greater criminal records they recidivate at similar rates. We observe a similar, large decrease in recidivism at the 75-point cutoff for Retail Fraud offenders. On the other hand, Retail Fraud offenders near the 50-point cutoff and OWI offenders at all three cutoffs show no response to harsher penalties. The similar increase in incarceration experienced by these offenders has no effect on future criminal behavior.



Table 6 quantifies the effect of harsher sentences on recidivism more precisely. Combining Retail Fraud and OWI offenders in the first column, harsher sentencing guidelines cause the number of future sentences incurred by an offender to drop by 0.15 on average. Columns (2) and (3) demonstrate that Retail Fraud, not OWI, drives this effect, and the bottom pane shows that particularly Retail Fraud offenders near the 25- and 75-point cutoffs respond to harsher sentencing guidelines.<sup>18</sup> The effects for Retail Fraud offenders are meaningfully large. The 0.31 decrease in future sentences represents a 22 percent decrease in recidivism, relative to a mean of 1.43, in response to a 23 percent increase in length of incarceration. On the other hand, the point estimates for OWI are small for policymaking purposes. Those sentenced under harsher guidelines commit 0.004, or 0.9 percent, fewer offenses in response to an 18 percent increase in length of incarceration.

### **5.3. Mechanisms**

We find some suggestive evidence that moving offenders from jail to prison drives the effects we observe. As discussed above, harsher guidelines both increase uttered jail sentences and shift offenders from jail to prison. We cannot cleanly separate these two changes. In Appendix Table 5, we split the sample between offenders who receive any prison time and those who do not. The effect of crossing the cutoffs on recidivism is stronger for those who go to prison. This result could indicate that moving offenders from jail to prison drives effects among Retail Fraud offenders. However, the apparent effectiveness of prison could also result from selection effects. Among offenders sentenced to prison, those sentenced under harsher guidelines commit fewer future crimes, but harsher guidelines also induce courts to move marginal offenders into to prison

---

<sup>18</sup> A reduction in similar property crimes (including Retail Fraud) drive the reduction in crime, indicating that Retail Fraud offenders respond to incarceration by avoiding conviction of the original offense that placed them behind bars. Results on recidivism by type of crime are available upon request.

who become the lowest risk among those in prison. Hence, harsher guidelines may reduce recidivism among those sentenced to prison by simply shifting composition. Appendix Table 6 shows that selection into prison matters. Those who receive prison are older and have more prior felony convictions than those not receiving prison (conditional on prior record total). On the other hand, a central role for the distinction between prison and jail would match how recidivism effects vary across the three cutoffs. Recall from Table 5 that crossing the 25-point and 75-point cutoffs moves many offenders from jail to prison, while crossing the 50-point cutoff only increases jail time. These changes line up with our recidivism effects, which concentrate at the 25- and 75-point cutoffs. Of course, the prior criminal record of offenders near the three cutoffs also varies considerably, and treatment effects may depend on criminal history. However, the simplest models of heterogeneous effects would predict a monotonic relationship. While not definitive, the available evidence suggests that moving offenders from jail to prison matters more than other components of sentence severity.

Both incapacitation and specific deterrence may matter in our sample. Incarcerating offenders may *incapacitate* them, making it physically difficult to commit crimes while locked up. On the other hand, a person's experience of incarceration may affect recidivism after release by providing information on incarceration, access to rehabilitation programs, peers that build criminal capital, etc. Criminologists refer to these effects as *specific deterrence*.<sup>19</sup> Our ability to identify these two effects differs across the cutoffs. The left pane of Figure 8 shows results for the 25-point cutoff. As the histogram shows, nearly all offenders with prior records near this cutoff receive no more than one year of incarceration, so any effect on recidivism in the first year after sentencing will combine incapacitation and specific deterrent effects, while effects after the first year can be

---

<sup>19</sup> This effect differs from *general deterrence*, i.e. when the general population avoids criminal behavior because they know it may result in a penalty.

attributed to specific deterrence. The solid line shows the effect of harsher sentencing guidelines on recidivism, over time. Crime reduction accrues entirely after one year, so we conclude that specific deterrence matters for this group with moderate criminal histories. Firm conclusions about incapacitation effects are more elusive because the first year combines incapacitation with specific deterrence for offenders released early in the first year. The right pane of Figure 8 shows offenders with severe records near the 75-point cutoff. Decreased recidivism among this group accumulates entirely within two years, which overlaps with the period during which some would be incarcerated. Hence, it is difficult to conclude whether specific deterrence or incapacitation drives decreased recidivism for the most severe offenders.<sup>20</sup>

#### **5.4. Robustness of Recidivism Results**

Exercises in the appendix demonstrate the robustness of the results on recidivism. First, to verify that we measure true discontinuities rather than statistical noise, we conduct placebo tests considering each score between 1 and 75 as a cutoff. Statistically significant drops in recidivism appear near the actual cutoffs in the guidelines and nowhere else. (See Appendix Figure 1.) Statistically significant decreases also occur at placebo cutoffs immediately adjacent to the actual cutoffs, e.g. 24 and 74, which is not surprising. Perhaps less obvious, we also detect increased recidivism away from the actual cutoffs near 10 and 65. However, the actual discontinuities mechanically create false discontinuities with an opposing sign far from the actual cutoffs. Appendix Figure 2 shows that we find similar results using sentencing guidelines variables, which have known discontinuities only at 25, 50 and 75 (not 10, 40 and 65), as the outcome. Appendix

---

<sup>20</sup> We could instead split the outcome period by the person's release date to separate incapacitation and specific deterrence. However, individual release dates are endogenous to sentencing. Offenders sentenced under harsher guidelines will be both older and in a different economic/criminal environment at release. Also, our data ends at a fixed date, and those assigned longer sentences have less follow-up data after release. To measure recidivism from release date, we would disproportionately drop severe offenders above the cutoff.

Figure 3 shows an example of the geometry of how actual increases in recommended sentences at the 25- and 50-point cutoffs create a false negative discontinuity at 40. Overall, this placebo test finds evidence consistent with cutoffs only at 25, 50, and 75 points.

Second, in our preferred specifications we limit the sample to people we can follow for five years in the data. We make this choice because the data end in 2013 for all individuals, and cases after 2008 have less follow-up data. We could instead include observations of varying time lengths to use all of the available information in the sample, but OLS gives equal weight to observations with longer and shorter follow-up periods. As we show in Appendix Table 7, a least squares specification using the full dataset but weighted by the length of the follow-up period generates results similar to our main specification. Different follow-up periods and censoring could also suggest modeling the data with a survival model. In Appendix Table 7, we estimate a Cox-proportional hazard survival model to test whether formally modeling differential follow-up and censoring affects the results. The hazard rate of recidivism falls by 26 percent at the 25-point cutoff for Retail Fraud offenders but not for OWI, which is very similar to our main results. Altogether, our results are robust to different methods of modeling recidivism.

## **5.5. Heterogeneity**

The recidivism response of Retail Fraud offenders to more severe sentences varies with observable, individual characteristics. As noted above, recidivism effects differ by offense and for offenders with prior record totals of 25, 50, and 75. Treatment effects also vary within a group of offenders with similar prior records. The first column of Table 7 examines how the effect of harsher sentences varies with age. We split the sample at the median age. The statistically significant coefficients of -0.68 and -2.79 indicate that recidivism decreases for younger retail fraud offenders near the 25- and 75-point cutoffs, respectively, when they are sentenced under

harsher guidelines. The statistically significant coefficients of 0.85 and 2.66 on the interactions between the older age group and the prior record cutoffs indicate that these benefits disappear for older offenders. Consistent with much of the existing literature, the coefficient of -0.43 on the older offender dummy indicates that older offenders in our sample re-offend at a lower rate. As above, we do not detect an effect at the 50-point cutoff. Results on sex in column (2) are mixed. Recidivism falls more for males than females at the 75-point cutoff but not at other cutoffs. Column (3) indicates that, near the 25-point cutoff, more severe sentences prevent future crime only in the southeast portion of Michigan (in and near Detroit). In sum, 5 of 9 interaction coefficients are statistically significant at the 10 percent level. We conclude that the effect of a harsher sentencing regime on recidivism varies considerably with observable characteristics.

## **6. Comparing Estimates Based on Sentencing Cutoffs vs. Judge Assignment**

We can directly compare our main estimates to those from a specification using judge assignment instruments. In theory, these two estimates may differ to the extent that they estimate local average treatment effects for different populations of compliers. The results above suggest that treatment effects do in fact vary with observable characteristics such as charge, criminal history, age, sex, and location. The added value of alternative identification strategies will depend on whether the two groups of compliers also experience different treatment effects. Comparing estimates provides direct evidence on such heterogeneity.

We implement a straightforward judge-assignment instrument in the same sample. Equations (10) and (11) provide the basic outline of this empirical strategy. We use the average sentence assigned by the same judge for the same offense to other defendants as an instrument for the sentence in the present case. We calculate average sentences by the same judge across years to increase power. We include information from cases during 2009 through 2013 in these judge

averages, though we do not have recidivism follow-up data for these cases. Due to the limitations of our data, we only pool sentences for guilty defendants and hence instrument with a measure of the judge's harshness for cases sentenced. To the extent that judges differ in their tendency to give any sentence at all, we will miss this effect.

Implementing a judge-assignment instrument in our context provides some practical obstacles. Existing studies focus on individual counties with uniform case assignment rules, allowing the econometrician to identify and control for all possible non-random judge assignment. This strategy relies heavily on identifying proper covariates and sample restrictions. For example, Mueller-Smith (2016) drops cases with an open probation sentence from the sample because these cases go to the original judge rather than a randomly assigned judge. Michigan circuit courts use conditional random assignment similar to other courts, making a similar empirical strategy possible in principle. In practice, though, we use data from dozens of courts spanning an entire state, and some of these courts condition random assignment on variables that cannot be observed in our statewide administrative data. We do observe the circuit court in which each judge presides (within which random assignment occurs) and can control for a full vector of circuit court-year fixed effects. Assignment of judges within a court-year should be mostly random, but may be subject to some bias due to selection on variables we cannot observe.

We provide some evidence on the validity of judge instruments by testing the first stage and baseline balance. Columns (1) through (3) of Table 8 compare the first stage for OLS, a judge assignment IV, and our main discontinuity specification implemented as the IV described in equations (8) and (9). As shown in the first column, an OLS regression of days incarcerated on itself trivially gives a coefficient of one. The second column shows that a judge's sentencing tendency in other cases carries over to the present case. For every additional day of incarceration

that the judge assigns to other Retail Fraud cases, the sentence for the present offender increases by one-third of a day. Finally, the third column replicates our prior result that sentences discontinuously increase by 46, 58, and 40 days, respectively, at the prior record thresholds. Columns (4) to (6) confirm similar results for OWI cases.

Table 9 tests baseline balance, showing whether the variation in sentences used by each method correlates with observable characteristics. As shown in the first row, a simple regression of baseline characteristics on days of incarceration while controlling for circuit-year fixed effects performs poorly. For both Retail Fraud and OWI offenders, those sentenced to longer stays in jail or prison have more prior felonies, more prior misdemeanors, and are older. The second row demonstrates that using sentences for other cases assigned to the same judge improves balance considerably. Both Retail Fraud and OWI offenders assigned to harsh versus lenient judges balance on age and prior criminal record. The third row confirms that characteristics of offenders on either side of the prior record discontinuity also balance. As discussed above, the marginally significant imbalance in age for Retail Fraud offenders appears to be only an artifact of a quadratic functional form. Overall, the judge assignment and discontinuity IVs both provide plausibly exogenous variation in sentences.

We compare the effects of sentencing across different specifications and find that the choice of empirical strategy matters considerably. Table 10 reports estimated treatment effects for different specifications. While sentences are multi-dimensional, we focus on the number of days incarcerated as a summary of the sentence. The first column replicates our main results for Retail Fraud in an instrumental variables specification. Increasing the sentence of Retail Fraud offenders by 100 days prevents 0.55 future felonies, which is a large reduction equal to 38 percent of the mean. The point estimate in the judge assignment specification in column (2) implies very different

effects from harsher sentences. Focusing on defendants whose sentences depend on the judge, adding 100 days to sentences causes 24 percent *more* future crimes. Though not different from zero, the judge IV point estimate is outside the 95 percent confidence interval of the discontinuity-based estimate. On the other hand, we cannot reject that the estimates for OWI cases shown in columns (4) and (5) are statistically different.

Two possible explanations exist for the difference in results for Retail Fraud. First, bias in the judge IV could matter. However, the baseline balance tests suggest that the judge-assignment IV performs reasonably well and should at least perform better than OLS. Column (3) of Table 10 shows the OLS results. If anything, instrumenting with judge assignment moves the treatment effects upward, away from the discontinuity-based estimates. We cannot completely rule out that judge-assignment IV estimates drawn from our data include bias, but this explanation seems less likely. Second, different instruments produce different sets of compliers and thus different local average treatment effects. In this interpretation, judge-assignment instruments measure the effect of sentencing on offenders to whom judges would give different sentences. As indicated by existing studies and the present data, harsher sentences for these offenders can be counterproductive. On the other hand, estimates based on discontinuities in sentencing guidelines measure treatment effects for offenders whose sentences would change if the state tinkered with sentencing guideline rules. We find that harsher sentences can reduce future crime committed by this group of offenders.

## **7. Conclusion**

This study finds that Retail Fraud and repeat drunk driving offenders respond differently to sentencing guidelines requiring more incarceration. Offenders for whom Michigan state sentencing guidelines discontinuously recommend longer incarceration are, in fact, assigned



longer sentences by the courts when convicted, 51 additional days for Retail Fraud – 1<sup>st</sup> Degree and 32 additional days for OWI – 3<sup>rd</sup> offense. We find no evidence that OWI offenders commit fewer offenses when receiving harsher sentences, while Retail Fraud offenders receive 22 percent fewer felony convictions in the future when sentenced under harsher guidelines. The effects for Retail Fraud vary with age, sex, geography, and prior criminal record. Because of this heterogeneity, estimates based on different empirical strategies should be understood to correspond to the effects of different policies. As an example, we show that our main results differ from those obtained from a standard instrumental variables model based on random judge assignment. Our main results directly apply to a natural policy lever in the state of Michigan: moving offenders at the margin from more lenient to harsher sentencing guidelines. Conversely, estimates from a judge assignment model identify the effects of replacing judges or restricting one set of judges to mimic another set. We conclude that policies that lead to harsher sentences for different groups of people can affect recidivism in quite different ways.

Our conclusion should be qualified in a few key ways. First, results for both empirical strategies – those based on random assignment of judges and ours exploiting discontinuities in sentencing guidelines – depend on context. States or time periods where the sentencing structure differs from Michigan between 2003 and 2013 may see different results. Second, due to available data, we narrow our focus to two crimes of interest. While Retail Fraud and OWI represent very common, mid-level, non-violent felonies, the data prevent us from directly testing whether their results are representative for other similar offenses. Third, because we examine only two sources of variation that correspond to distinct policy changes, we cannot estimate the effect of other policies, such as universally granting judges more discretion, with these sources of variation.

We can, however, consider the implications of our results for the design of Michigan's sentencing guidelines in the context of cost-benefit analysis and subsequently explore just how "smart" smart justice might have to be. Consider a much-simplified analysis of Retail Fraud offenders near the sentencing guidelines' cutoffs. The benefit of increased incarceration for Retail Fraud offenses is only partially reflected in the 0.31 fewer future sentences that follow, most of which result from preventing future retail fraud. Our data only include crimes that result in a felony sentence. Using available national and state data to account for unreported crimes, unresolved criminal cases, and non-conviction, we estimate the reduction in felony sentences corresponds to 13 fewer crimes committed.<sup>21</sup> If all avoided crimes are Retail Fraud and the costs imposed by each are \$400 (including victim costs and community expenses like policing),<sup>22</sup> harsher sentencing provides \$5,200 of value to society. Incarceration in Michigan costs the state approximately \$100 per day (Maxwell, 2015), or \$5,100 for the 51 additional days generated by harsher sentencing guidelines. Thus, longer incarceration for a Retail Fraud offender at the margin of the sentencing guidelines passes a simple cost-benefit test comparing the benefits of crime reduction to the costs of incarceration. Since longer incarceration for OWI offenders provides no evidence of crime reduction benefits, it fails a similar, simple cost-benefit analysis. Cost-benefit analysis would also differ for some sub-samples or with a fuller analysis of crime reduction benefits. For instance, consider the larger 0.68 sentence reduction in future recidivism for young Retail Fraud offenders near the 25-point cutoff. It implies \$11,657 of public benefits from crime

---

<sup>21</sup> National surveys indicate that 40 percent of property crimes are reported (Rand, 2009), and 12 percent of reported property crime incidents are cleared in Michigan, typically with an arrest (MICR, 2008). NIBRS-type data from Michigan indicate 18,652 Retail Fraud (of all degrees) cleared in Michigan in 2008, and Michigan's state Judicial Warehouse data show a final guilty decision for 8,966 Retail Fraud cases (of all degrees) in 2008, or 48 percent of cleared incidents. Thus, approximately 1 in 43 crimes receives a sentence and enters the OMNI database.

<sup>22</sup> This value matches the mean value of restitution paid in our data, which is \$362. The literature commonly uses crime costs from Miller, Cohen, and Wiersma (1996); however, they do not include retail fraud. If we assign MCW costs to crimes in our data and then impute the cost of retail fraud according to its group and class, the cost of retail fraud is \$473.

reduction compared to an incarceration cost of \$2,800.<sup>23</sup> More generally, the incapacitation and specific deterrence effects we measure in this research do not include any general equilibrium effects of harsher sentencing policies including any enhanced general deterrence within the whole population. A more broadly understood and quantified measure of the social benefits of reduced crime could tip the balance in either direction.

In this manner, optimal sentencing guidelines will likely depend on the characteristics of the offender and the crime. Net benefits of incarceration for a wider group of crimes, including violent crimes, will likely hinge on the direction and magnitude of the response of recidivism to sentencing for the particular crime and offender. We find that such responses to harsher sentences are heterogeneous, leading to lower recidivism for some offenders and no effect on others. Thus, making sentencing guidelines smarter will require careful attention to tailoring sentences to different crime categories and offenders.

---

<sup>23</sup> We estimate changes to sentencing as in column (1) of Table 5 but limiting the sample to younger offenders. In this specification, the jump in the guidelines increases incarceration in the uttered sentence by 28 days.

## References

- Abrams, D. S., 2012. "Estimating the Deterrent Effect of Incarceration Using Sentencing Enhancements." *American Economic Journal: Applied Economics*, 4(4): 32-56.
- Aizer, A. and Doyle, J. J., 2015. "Juvenile Incarceration, Human Capital, and Future Crime: Evidence from Randomly Assigned Judges." *The Quarterly Journal of Economics*, Oxford University Press, 130(2); 759-803.
- Aliprantis, D. and Richter, F.G.C., 2016. "Evidence of Neighborhood Effects from Moving to Opportunity: LATEs of Neighborhood Quality." Unpublished working paper.
- Anderson, J. M., Kling, J. R., and Stith, K., 1999. "Measuring Interjudge Sentencing Disparity: Before and After the Federal Sentencing Guidelines." *Journal of Law and Economics*, 42(1): 271-307.
- Angrist, J. D. and Evans, W. N., 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review*, 88(3), pp.450-477.
- Becker, G. S., 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy*, 76(2): 169-217.
- Bhuller, M., Dahl, G. B., Loken, K. V., Mogstad, M., 2018. "Intergenerational Effects of Incarceration." (No. w24227). National Bureau of Economic Research.
- Billings, S. B. and Phillips, D. C., 2017. "Why do kids get into trouble on school days?" *Regional Science and Urban Economics*, 65, pp.16-24.
- Bjerk, D., 2014. "Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion Under Mandatory Minimum Sentencing." *Journal of Law and Economics*, 48: 591-625.
- Bushway, S. D., Owens, E. G., and Piehl A. M., 2012. "Sentencing Guidelines and Judicial Discretion: Quasi - Experimental Evidence from Human Calculation Errors." *Journal of Empirical Legal Studies*, 9(2): 291 - 319.
- Chalfin, A., 2013. "What is the contribution of Mexican immigration to US crime rates? Evidence from rainfall shocks in Mexico." *American Law and Economics Review*, 16(1), 220-268.
- Chalfin, A. and Deza, M., 2018. "New Evidence on Mexican Immigration and Crime in the United States: Evidence from a Natural Experiment in Immigration Enforcement." Unpublished working paper.
- Chyn, E. 2016. "Moved to Opportunity: The Long-Run Effect of Public Housing Demolition on Labor Market Outcomes of Children." Unpublished working paper.

Council of State Governments (2009) "Justice Reinvestment in Michigan: Analyses of Crime, Community Corrections, and Sentencing Policies."

Council of State of Governments (2014) "Applying a Justice Reinvestment Approach to Improve Michigan's Sentencing System: Summary Report of Analyses and Policy Options."

Dobbie, W., Goldin, J., Yang, C. S. (2018) "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges." *American Economic Review*, 108(2): 201-240.

Dobbie, W., Grönqvist, H., Niknami, S., Palme, M. and Priks, M., 2018. "The Intergenerational Effects of Parental Incarceration" (No. w24186). National Bureau of Economic Research.

Ehrlich, I., 1973. "Participation in Illegitimate Activities: A Theoretical and Empirical Investigation." *Journal of Political Economy*, 81(3): 521-565.

Franco, C., Harding, D. J., Morenoff, J. and Bushway, S. D. 2017 "Estimating the Effect of Imprisonment on Recidivism: Evidence from a Regression Discontinuity Design." Unpublished working paper. Accessed 1/16/2018 from <https://sites.google.com/site/cfrancoecon/research>.

Green, D. P. and Winik, D., 2010. "Using Random Judge Assignments to Estimate the Effects of Incarceration and Probation on Recidivism Among Drug Offenders." *Criminology*, 48: 357-387.

Guarín, A., Medina, C., and Tamayo, J. A., 2013. "The Effects of Punishment of Crime in Colombia on Deterrence, Incapacitation, and Human Capital Formation." *Borradores de Economía* 774, Banco de la Republica de Colombia.

Hansen, B., 2015. "Punishment and Deterrence: Evidence from Drunk Driving." *American Economic Review*, 105(4): 1581-1617.

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

Helland, E. and Tabarrok, A., 2007. "Does Three Strikes Deter?: A Nonparametric Estimation." *Journal of Human Resources*, 42:309-330.

Hjalmarsson, R., 2009. "Juvenile Jails: A Path to the Straight and Narrow or to Hardened Criminality?" *The Journal of Law and Economics*, 52(4): 779-809.

Imbens, G. W., and Angrist, J. D., 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2): 467-475.

Jacob, B. A., Kapustin, M. and Ludwig, J., 2014. "The impact of housing assistance on child outcomes: Evidence from a randomized housing lottery." *Quarterly Journal of Economics*, 130(1): 465-506.

Jacob, B. A. and Lefgren, L., 2003. "Are idle hands the devil's workshop? Incapacitation, concentration, and juvenile crime." *American Economic Review*, 93(5): 1560-1577.

Kauder, N. B. and Ostrom, B. J., 2008. "State Sentencing Guidelines Profiles and Continuum." National Center for State Courts.

Kessler, D. P. and Levitt, S. D., 1999. "Using Sentence Enhancements to Distinguish Between Deterrence and Incapacitation." *Journal of Law and Economics*, 42(1): 343-63.

Kleinberg, J. Lakkaraju, H., Leskovec, J., Ludwig, J., Mullainathan, S, 2018. "Human Decisions and Machine Predictions" (No. w24186). National Bureau of Economic Research.

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

LaCasse, C. and Payne, A. A., 1999. "Federal Sentencing Guidelines and Mandatory Minimum Sentences: Do Defendants Bargain in the Shadow of the Judge?" *Journal of Law and Economics*, 42(1): 245-69.

Lee, D. and McCrary, J., 2017. "The Deterrence Effect of Prison: Dynamic Theory and Evidence," in Cattaneo, M. and Escanciano, J.C. (eds.) *Regression Discontinuity Designs (Advances in Econometrics, Volume 38)*: 73-146.

Loeffler, C. E. and Grunwald, B., 2015. "Decriminalizing Delinquency: The Effect of Raising the Age of Majority on Juvenile Recidivism." *The Journal of Legal Studies*, 44(2): 361 - 388.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R. and Sanbonmatsu, L., 2013. "Long-term neighborhood effects on low-income families: Evidence from Moving to Opportunity." *American Economic Review: Papers & Proceedings*, 103(3): 226-231.

Lundborg, P., Plug, E. and Rasmussen, A.W., 2017. "Can Women Have Children and a Career? IV Evidence from IVF Treatments." *American Economic Review*, 107(6): 1611-1637.

Maxwell, J., 2015. "Marginal Cost of Corrections: Michigan's Experience." Papers Examining Critical Issues Facing Michigan paper series, Senate Fiscal Agency.

McCrary, J., 2008. "Manipulation of the running variable in the regression discontinuity design: A density test," *Journal of Econometrics*, 142(2): 698-714.

Michigan Judicial Institute, 2012. "State of Michigan Sentencing Guidelines Manual."

Michigan Supreme Court, 2008. "Annual Report of the Michigan Supreme Court 2008."

- MICR, 2008. "Michigan Incident Crime Reporting." Accessed from [http://www.michigan.gov/documents/msp/Aa-Introduction\\_GroupA\\_B\\_305552\\_7.pdf](http://www.michigan.gov/documents/msp/Aa-Introduction_GroupA_B_305552_7.pdf)
- Mueller-Smith, M. G., 2016. "The Criminal and Labor Market Impacts of Incarceration." mimeo.
- Nagin, D. S. and Snodgrass, G. M., 2013. "The Effect of Incarceration on Re-Offending: Evidence from a Natural Experiment in Pennsylvania." *Journal of Quantitative Criminology*, 29(4): 601-642.
- Owens, E. G., 2009. "More Time, Less Crime? Estimating the Incapacitative Effect of Sentence Enhancements." *Journal of Law and Economics*, 52(3): 551-579.
- Obama, Barack, July 14, 2015, Remarks by the President at the NAACP Conference. Philadelphia, Pennsylvania.
- Rand, M.R., 2009. "Criminal Victimization, 2008." *Bureau of Justice Statistics Bulletin*. U.S. Department of Justice.
- Snyder, Rick, March 7, 2012. "Special Message to the Legislature on Public Safety." Flint, Michigan.
- United States Sentencing Commission, 2011. "2011 Report to the Congress: Mandatory Minimum Penalties in the Federal criminal Justice System."

**Figure 1: Michigan Felony Sentencing Guidelines, Class E Grid**

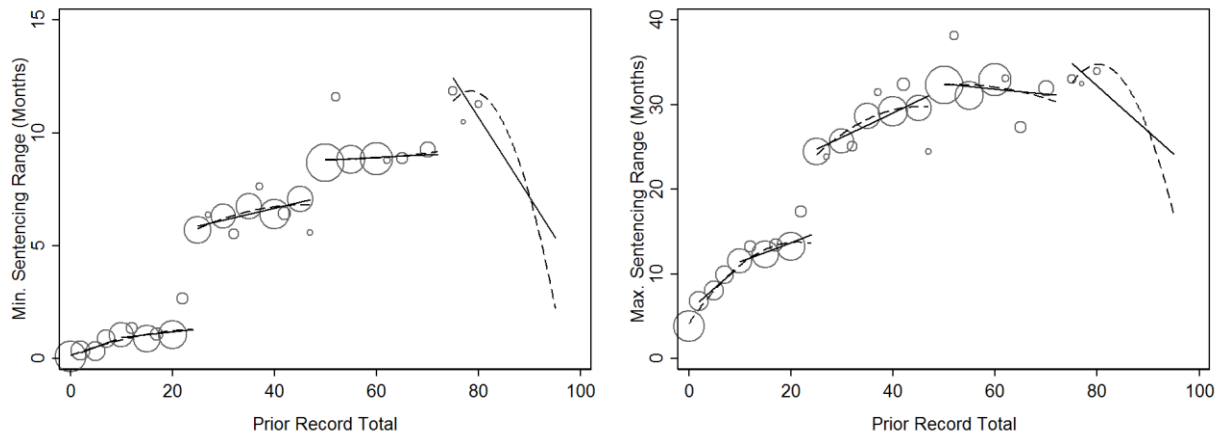
OV Level	PRV Level											Offender Status	
	A 0 Points		B 1-9 Points		C 10-24 Points		D 25-49 Points		E 50-74 Points		F 75+ Points		
I 0-9 Points	0	3*	0	6*	0	9*	5	23	7	23	9	23	
		3*		7*		11*		28		28		28	HO2
		4*		9*		13*		34		34		34	HO3
		6*		12*		18*		46		46		46	HO4
II 10-24 Points	0	6*	0	9*	0	11*	7	23	10	23	12	24	
		7*		11*		13*		28		28		30	HO2
		9*		13*		16*		34		34		36	HO3
		12*		18*		22		46		46		48	HO4
III 25-34 Points	0	9*	0	11*	0	17*	10	23	12	24	14	29	
		11*		13*		21		28		30		36	HO2
		13*		16*		25		34		36		43	HO3
		18*		22		34		46		48		58	HO4
IV 35-49 Points	0	11*	0	17*	5	23	12	24	14	29	19	38	
		13*		21		28		30		36		47	HO2
		16*		25		34		36		43		57	HO3
		22		34		46		48		58		76	HO4
V 50-74 Points	0	14*	5	23	7	23	14	29	19	38	22	38	
		17*		28		28		36		47		47	HO2
		21		34		34		43		57		57	HO3
		28		46		46		58		76		76	HO4
VI 75+ Points	0	17*	7	23	12	24	19	38	22	38	24	38	
		21		28		30		47		47		47	HO2
		25		34		36		57		57		57	HO3
		34		46		48		76		76		76	HO4

Source: State of Michigan Sentencing Guidelines Manual, 2016. Asterisks denote intermediate sanction cells, while straddle cells are shaded and prison cells are unmarked.

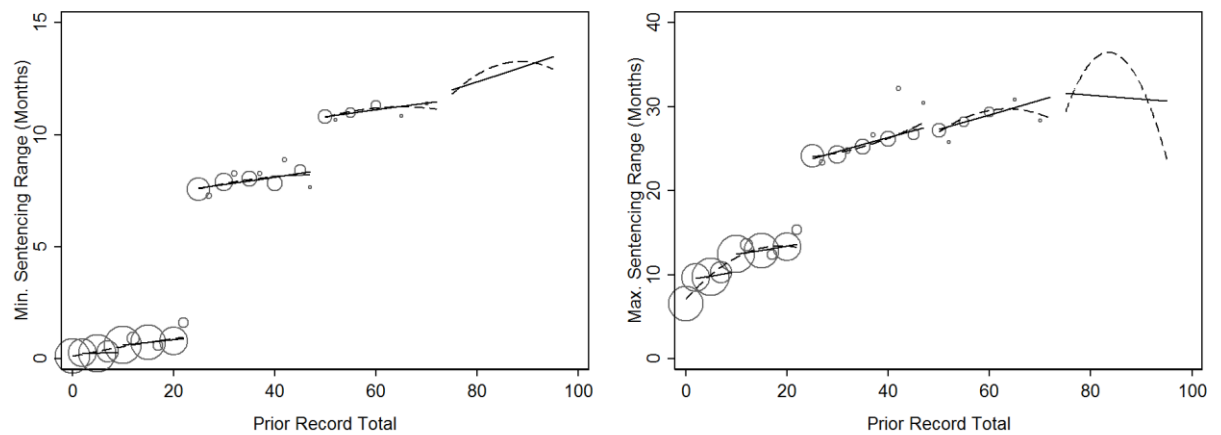


**Figure 2: Discontinuities in Sentencing Guidelines, Months of Incarceration**

**A. Retail Fraud – 1<sup>st</sup> Degree**

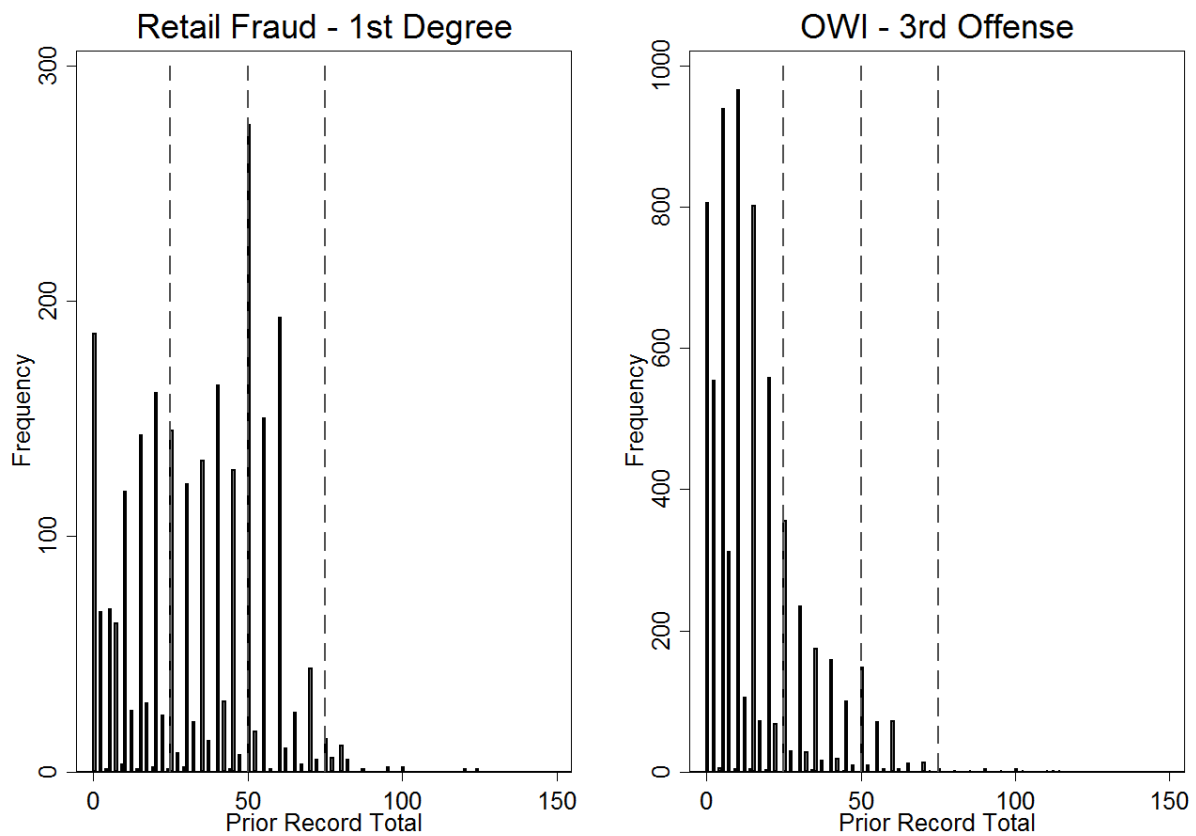


**B. Operating While Intoxicated – 3<sup>rd</sup> Offense**



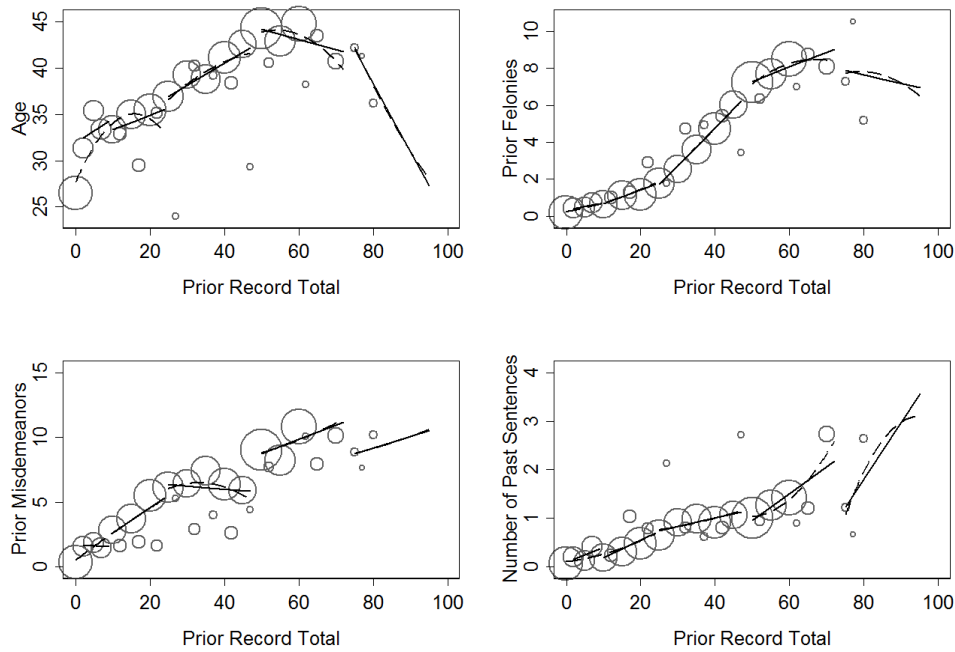
Plotted circles show sample means for charges with the same prior record total. The size of the circle indicates the number of charges with that total. The dashed curve displays quadratic polynomials fitted separately above and below cutoffs at 25, 50 and 75. The solid curve fits the data linearly within bins defined by all potential sentencing guideline categories: 0, 1-9, 10-24, 25-49, 50-74, and 75-100.

**Figure 3. Histogram of Prior Record Total, Retail Fraud and OWI-3<sup>rd</sup>**

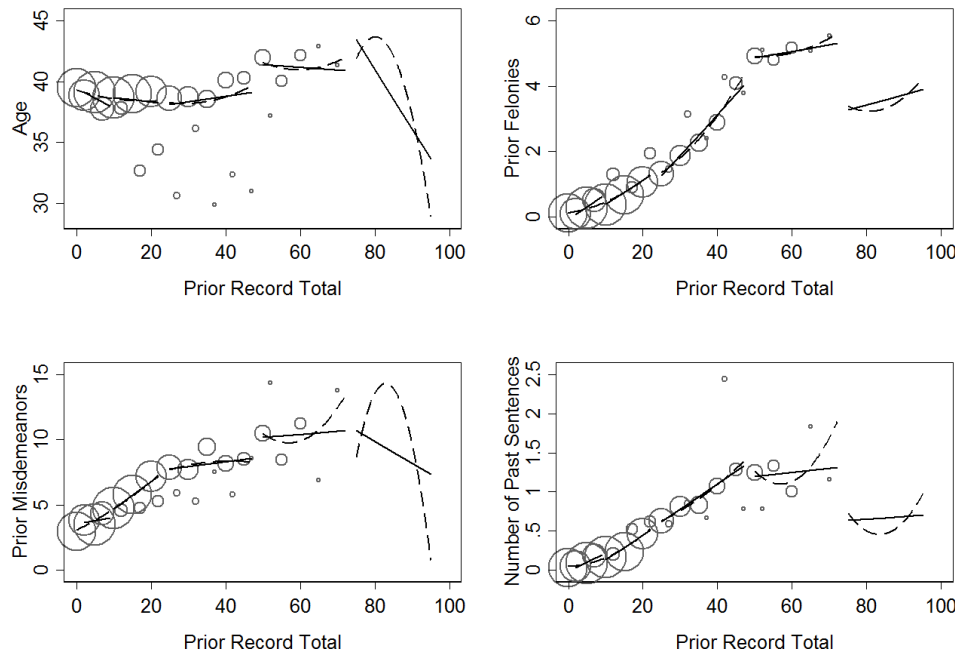


Sample includes all sentences for OWI – 3<sup>rd</sup> and Retail Fraud – 1<sup>st</sup> Degree in the MDOC sentencing data. The vertical lines indicate the 25-, 50-, and 75-point cutoffs dividing cells of the sentencing guidelines.

**Figure 4. Baseline Balance**  
**A. Retail Fraud – 1<sup>st</sup> Degree**



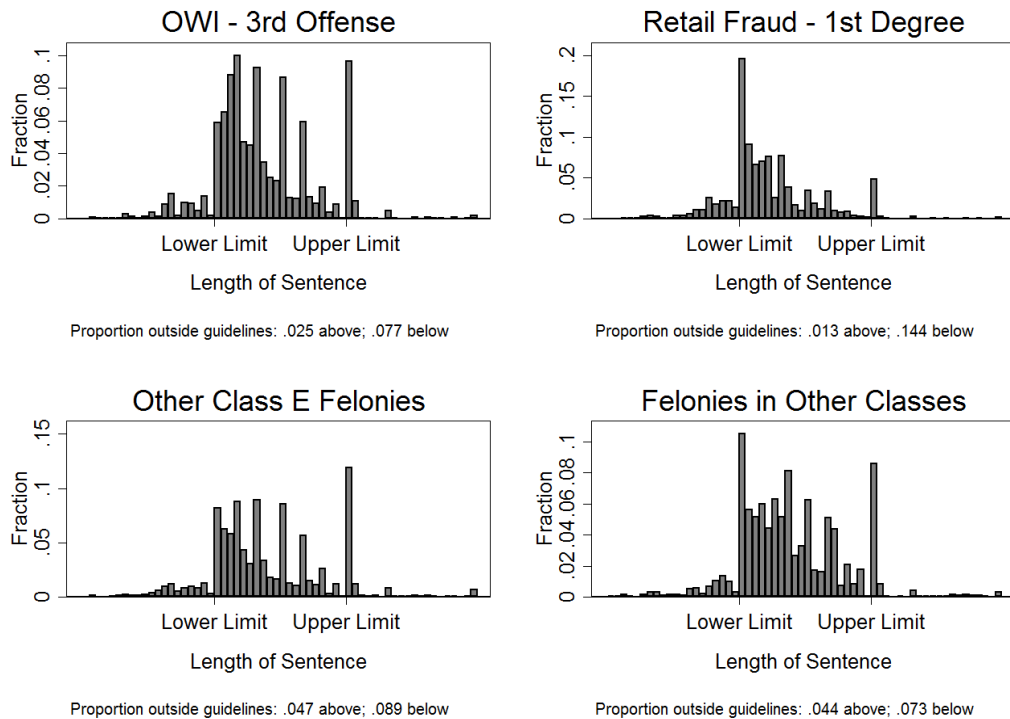
**B. Operating While Intoxicated – 3<sup>rd</sup> Offense**



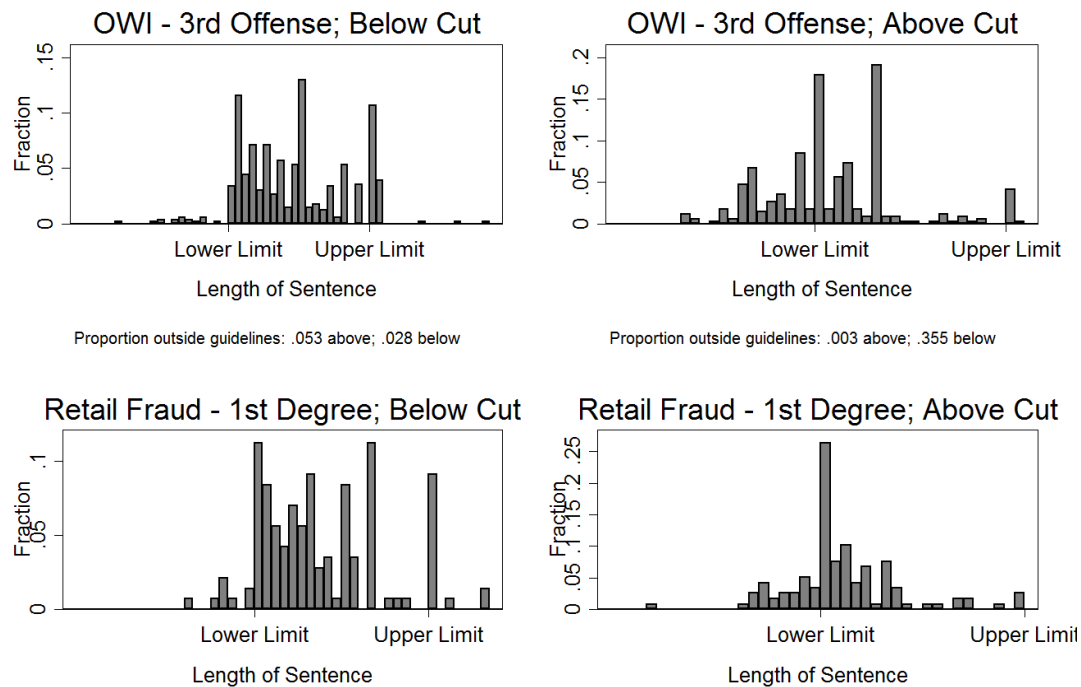
Plotted circles show sample means for charges with the same prior record total. The size of the circle indicates the number of charges with that total. The dashed curve displays quadratic polynomials fitted separately above and below cutoffs at 25, 50 and 75. The solid curve fits the data linearly within bins defined by all potential sentencing guideline categories: 0, 1-9, 10-24, 25-49, 50-74, and 75-100.

**Figure 5. Histograms of Uttered Sentences Relative to Sentencing Guidelines**

**A. By Type of Charge**



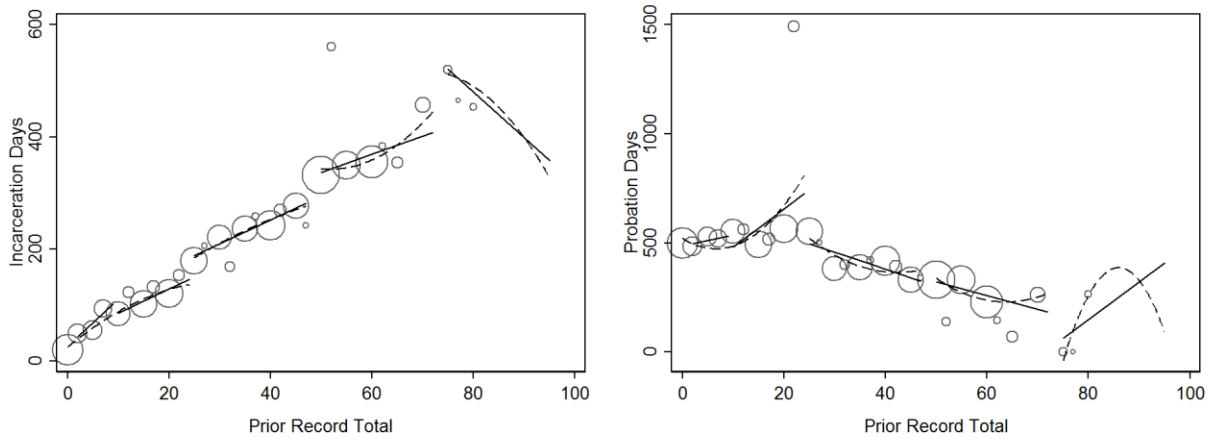
**B. Cases Just Above and Just Below 25-point Prior Record Cutoff**



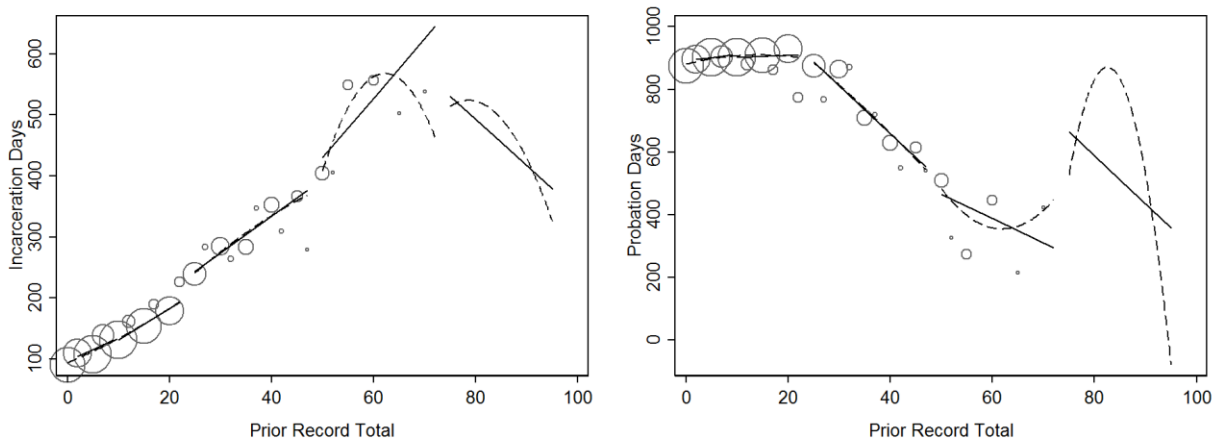
Data show the number of days incarcerated for the uttered sentence. Sentences are normalized to be a percentage of the difference between the upper and lower limits of the sentencing guidelines for that case. “Above Cut” limits the sample to prior record totals between 20 and 24; “Below Cut” limits the sample to prior record totals between 25 and 29.

**Figure 6. Discontinuities in Uttered Sentences**

**A. Retail Fraud – 1<sup>st</sup> Degree**



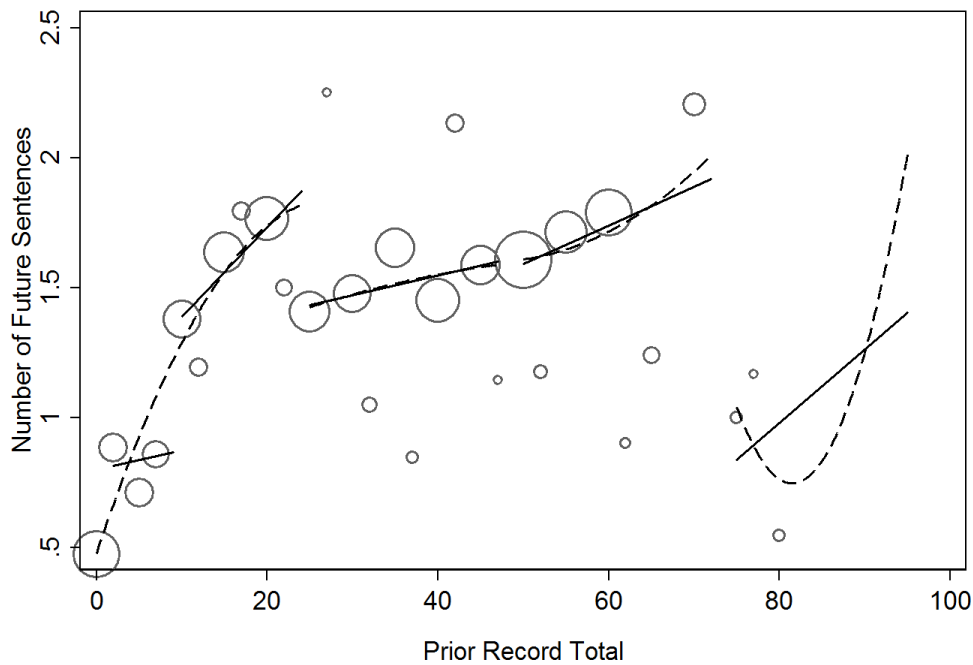
**B. Operating While Intoxicated – 3<sup>rd</sup> Offense**



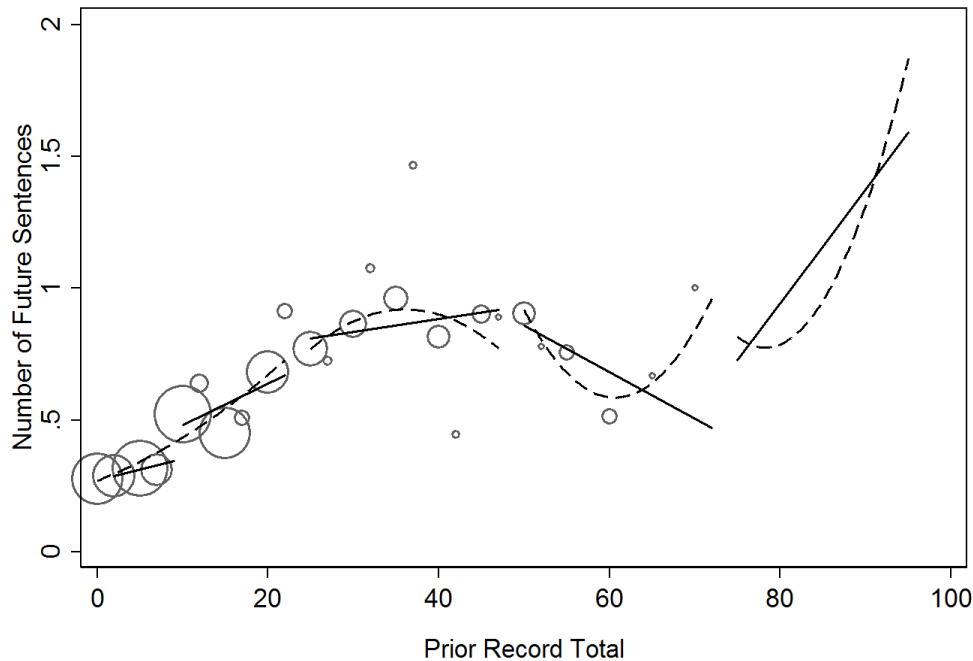
Plotted circles show sample means for charges with the same prior record total. The size of the circle indicates the number of charges with that total. The dashed curve displays quadratic polynomials fitted separately above and below cutoffs at 25, 50 and 75. The solid curve fits the data linearly within bins defined by all potential sentencing guideline categories: 0, 1-9, 10-24, 25-49, 50-74, and 75-100.

**Figure 7. Discontinuities in Felony Recidivism**

**A. Retail Fraud – 1<sup>st</sup> Degree**

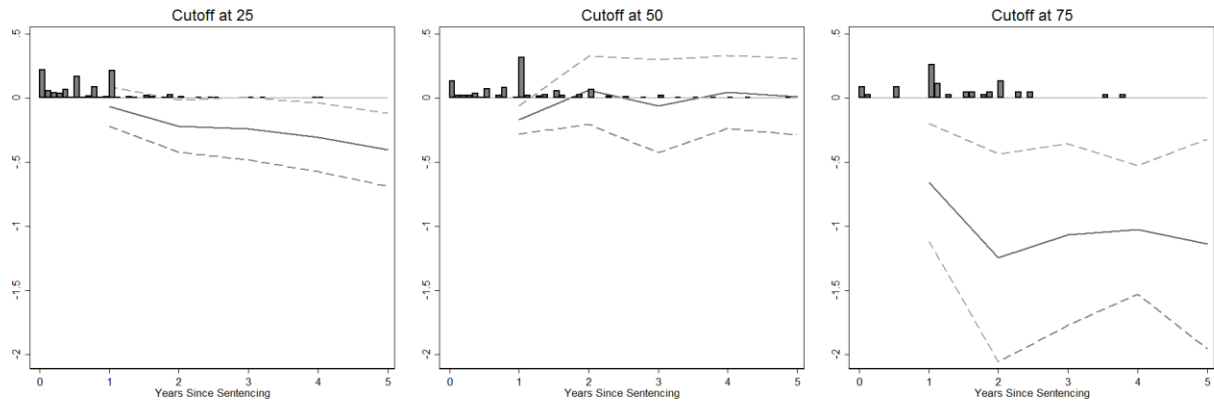


**B. Operating While Intoxicated – 3<sup>rd</sup> Offense**



Plotted circles show sample means for charges with the same prior record total. The size of the circle indicates the number of charges with that total. The dashed curve displays quadratic polynomials fitted separately above and below cutoffs at 25, 50 and 75. The solid curve fits the data linearly within bins defined by all potential sentencing guideline categories: 0, 1-9, 10-24, 25-49, 50-74, and 75-100.

**Figure 8. Histograms of Sentence Length and Recidivism Effects by Years Since Sentencing, Retail Fraud**



The bars display a histogram for years of incarceration for cases with prior record total within 5 points of the cutoff. Each point on the solid line shows the coefficient on a dummy for passing the listed sentencing cutoff in a regression where the dependent variable is the number of future felony sentences in the Xth year since the original sentencing. All other aspects of the specifications are identical to those in the bottom pane of column (2) in Table 6. The dashed lines show pointwise 95-percent confidence intervals.

**Table 1. Summary Statistics**

Sample:	Retail Fraud - 1st Deg.	OWI - 3rd Offense	Other Class E	Other Not Class E	All
Age	38.1 (11.2)	38.8 (10.1)	36.6 (10.0)	34.8 (10.1)	37.5 (10.4)
Southeast Region	0.64 (0.48)	0.36 (0.48)	0.32 (0.47)	0.45 (0.50)	0.42 (0.49)
Prior Misdemeanors	5.83 (5.93)	5.43 (4.38)	5.27 (4.64)	6.07 (5.87)	5.58 (5.02)
Prior Felonies	3.94 (4.31)	0.98 (1.78)	1.50 (2.42)	2.60 (3.14)	1.87 (2.94)
Prior Felonies in Sentencing Data	0.81 (1.44)	0.34 (0.92)	0.43 (1.20)	1.00 (1.53)	0.56 (1.23)
Future Felony Sentences	1.43 (1.89)	0.51 (1.08)	1.03 (1.69)	1.88 (2.15)	1.02 (1.68)
--Controlled Substance	0.26 (0.74)	0.044 (0.30)	0.13 (0.54)	0.39 (0.91)	0.16 (0.60)
--Personal	0.19 (0.59)	0.075 (0.38)	0.13 (0.50)	0.27 (0.69)	0.14 (0.52)
--Property	1.21 (1.71)	0.041 (0.30)	0.53 (1.42)	1.07 (1.88)	0.52 (1.35)
--Public Order	0.027 (0.20)	0.018 (0.15)	0.039 (0.25)	0.059 (0.34)	0.031 (0.22)
--Public Safety	0.090 (0.35)	0.39 (0.75)	0.57 (0.87)	0.68 (0.88)	0.43 (0.78)
--Public Trust	0.017 (0.16)	0.0090 (0.11)	0.011 (0.10)	0.036 (0.24)	0.016 (0.15)
Uttered Sentence - Min. Incarceration Days	217.4 (229.3)	174.9 (180.8)	213.2 (246.8)	268.2 (501.1)	206.9 (288.5)
Uttered Sentence - Probation Days	420.7 (675.3)	853.7 (522.4)	742.6 (680.8)	1003.7 (13907.6)	788.1 (6012.6)
Statutory Maximum Sentence (Months)	57.8 (7.83)	60.0 (0.90)	64.8 (25.6)	792.2 (8504.0)	196.3 (3671.0)
Min. Sentencing Range (Months)	5.14 (4.35)	2.26 (3.91)	3.18 (4.49)	4.65 (13.6)	3.36 (7.05)
Max. Sentencing Range (Months)	22.5 (13.9)	14.1 (7.99)	16.0 (9.87)	22.0 (30.0)	17.4 (16.1)
Prior Record Total	32.9 (21.1)	15.1 (14.8)	20.3 (18.2)	30.1 (19.6)	21.9 (19.1)

Notes: Means for all observations in the sentencing data. Standard deviations in parentheses.



**Table 2. Discontinuities in Sentencing Guidelines, Months Incarceration**

Outcome: Sample:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Minimum Guideline (months)				Maximum Guideline (months)			
	RF/OWI Combined	Retail Fraud	OWI	Other	RF/OWI Combined	Retail Fraud	OWI	Other
Number of Cutoffs Exceeded	4.70*** (0.64)	3.26*** (0.57)	5.63*** (0.55)	2.95*** (0.42)	7.95*** (1.52)	6.68*** (1.46)	8.41*** (1.53)	7.06*** (1.09)
R-Squared	0.75	0.66	0.77	0.13	0.64	0.54	0.67	0.18
Prior Record Total Exceeds 25	5.93*** (0.3)	4.46*** (0.55)	6.54*** (0.24)	3.50*** (0.33)	11.3*** (0.87)	10.6*** (1.45)	11.3*** (0.72)	8.57*** (0.9)
Prior Record Total Exceeds 50	2.14*** (0.4)	2.04*** (0.52)	2.55*** (0.32)	3.38*** (0.49)	0.59 (1.08)	2.88** (1.1)	-2.08 (1.39)	7.57*** (1.6)
Prior Record Total Exceeds 75	2.12*** (0.49)	2.60*** (0.48)	0.81* (0.42)	-3.79* (1.97)	3.97* (2.15)	3.55 (2.67)	3.11 (3.86)	-8.34* (4.62)
R-Squared	0.75	0.66	0.78	0.13	0.65	0.54	0.68	0.18
Prior Record Total Quadratic × Prior Record Bins	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	3.03	5.13	2.26	3.88	16.4	22.5	14.1	18.9
N	9091	2433	6658	5705	9.91	2433	6658	5705

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total and its interaction with the three prior record cutoffs (at least 25, at least 50, and at least 75). The top and bottom panes estimate separate models. The top pane reports the coefficient on a variable counting how many of the three cutoffs the person's record exceeds. The bottom pane reports coefficients from a model including separate dummies for exceeding each cutoff. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variables are the lower and upper limits of the sentencing guideline range, in months.

**Table 3. Baseline Balance**

Sample:	(1) RF/OWI Combined	(2) Retail Fraud	(3) OWI	(4) Other
Outcome:				
Prior Felonies	-0.14 (0.21)	0.01 (0.35)	-0.22 (0.17)	-0.78*** (0.22)
Prior Misdemeanors	0.81 (0.86)	1.69 (1.13)	0.15 (0.71)	0.47 (1.02)
Prior Felonies (Sentencing Data)	-0.17** (0.076)	-0.28 (0.2)	-0.11* (0.05)	-0.16 (0.1)
Age at Offense	1.49 (1.32)	3.34** (1.5)	0.11 (1.38)	0.5 (1.78)

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total and its interaction with the three prior record cutoffs (at least 25, at least 50, and at least 75). Each cell reports the coefficient on a variable counting how many of the three cutoffs the person's record exceeds. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variables are characteristics determined prior to sentencing. "Prior felonies" records all felonies previously committed by the same person while "prior felonies (sentencing data)" counts only those prior felonies we can observe in the timeframe of our sentencing data.

**Table 4. Discontinuities in Uttered Sentences**

Sentencing Outcome:  Sample:	(1)	(2)	(3)	(4)	(5)	(6)
	Incarceration Days			Probation Days		
	RF/OWI Combined	Retail Fraud	OWI	RF/OWI Combined	Retail Fraud	OWI
Number of Cutoffs Exceeded	42.4*** (9.87)	50.8*** (12.4)	31.5*** (10.8)	-45.9 (27.6)	-216.9 (148.8)	-6.35 (41.9)
R-Squared	0.27	0.28	0.3	0.097	0.036	0.051
Prior Record Total at Least 25	38.2*** (12.1)	47.0*** (17.4)	30.4** (12.1)	-39.7 (38.1)	-328.5 (290)	-8.5 (48.2)
Prior Record Total at Least 50	45.6** (20.9)	58.2*** (20)	27.5 (27.3)	-21.4 (22.2)	-56.1 (70.6)	-4.45 (78.1)
Prior Record Total at Least 75	85 (58.7)	39.9 (52.8)	151 (160.5)	-289.5*** (80.4)	-333.4*** (101.7)	84.7 (205.3)
R-Squared	0.27	0.28	0.3	0.097	0.037	0.051
Prior Record Total Quadratic X Prior Record Bins	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	186.2	217.4	174.8	738.1	421.1	854
N	9091	2433	6658	9091	2433	6658

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total and its interaction with the three prior record cutoffs (at least 25, at least 50, and at least 75). The top and bottom panes estimate separate models. The top pane reports the coefficient on a variable counting how many of the three cutoffs the person's record exceeds. The bottom pane reports coefficients from a model including separate dummies for exceeding each cutoff. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variables are the number of days of incarceration (prison plus jail) and probation in the uttered sentence.

**Table 5. Decomposing Changes in Sentences**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Incarceration Days	Jail Days	Prison Days	Any Jail	Any Prison	Jail Days (if > 0)	Prison Days (if > 0)
Number of Cutoffs Exceeded	42.4*** (9.87)	16.9*** (5.2)	25.5*** (7.96)	-0.014 (0.02)	0.038*** (0.01)	33.3*** (7.32)	5.66 (31.7)
R-Squared	0.27	0.043	0.23	0.12	0.25	0.21	0.024
Prior Record Total at Least 25	38.2*** (12.1)	15.1** (6.58)	23.1** (9.24)	-0.021 (0.02)	0.033* (0.02)	30.1*** (8.57)	69.9 (47.4)
Prior Record Total at Least 50	45.6** (20.9)	30.7*** (9)	14.9 (18)	0.014 (0.03)	0.029 (0.02)	51.3*** (17)	-37.8 (35.2)
Prior Record Total at Least 75	85 (58.7)	-42.0* (21.3)	127.0* (70.2)	-0.089* (0.05)	0.16*** (0.05)	-49.6 (75.7)	42.9 (94.7)
R-Squared	0.27	0.043	0.23	0.12	0.26	0.21	0.024
Prior Record Total Quadratic X Prior Record Bins	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	186.2	125.3	60.8	0.77	0.11	163.6	576
N	9091	9091	9091	9091	9091	6965	960

Combined results for Retail Fraud and OWI. Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total and its interaction with the three prior record cutoffs (at least 25, at least 50, and at least 75). The top and bottom panes estimate separate models. The top pane reports the coefficient on a variable counting how many of the three cutoffs the person's record exceeds. The bottom pane reports coefficients from a model including separate dummies for exceeding each cutoff. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variables are characteristics of the "uttered sentence" assigned by the court, which defines minimum time served.

**Table 6. Discontinuities in Recidivism**

Sample:	(1)	(2)	(3)
	Number of Future Offenses RF/OWI Combined	Retail Fraud	OWI
Number of Cutoffs Exceeded	-0.15** (0.07)	-0.31** (0.14)	0.0044 (0.07)
R-Squared	0.07	0.037	0.038
Prior Record Total at Least 25	-0.21*** (0.07)	-0.40*** (0.15)	-0.05 (0.08)
Prior Record Total at Least 50	0.15 (0.1)	0.0084 (0.15)	0.24* (0.13)
Prior Record at Least 75	-1.07*** (0.39)	-1.14*** (0.42)	-0.46 (0.42)
R-Squared	0.071	0.038	0.038
Prior Record Total Quadratic X Prior Record Bins	Yes	Yes	Yes
Mean of Dep. Var.	0.76	1.43	0.51
N	9091	2433	6658

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total and its interaction with the three prior record cutoffs (at least 25, at least 50, and at least 75). The top and bottom panes estimate separate models. The top pane reports the coefficient on a variable counting how many of the three cutoffs the person's record exceeds. The bottom pane reports coefficients from a model including separate dummies for exceeding each cutoff. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variable is the number of future felony sentences observed in the data.

**Table 7. Heterogeneity in Recidivism Effects**

Sample:	(1)	(2)	(3)
Characteristic:	Age > Median	Retail Fraud Female	Southeast Region
Prior Record Total at Least 25	-0.68*** (0.19)	-0.47 (0.49)	0.76*** (0.23)
Prior Record Total at Least 50	0.63 (0.41)	-0.059 (0.26)	-0.62** (0.31)
Prior Record Total at Least 75	-2.79*** (0.50)	-1.44*** (0.46)	-0.80*** (0.16)
At Least 25 × Characteristic	0.85* (0.43)	0.12 (0.79)	-1.92*** (0.32)
At Least 50 × Characteristic	-0.84 (0.50)	0.067 (0.57)	0.97** (0.43)
At Least 75 × Characteristic	2.66*** (0.28)	1.75*** (0.49)	-0.66 (0.59)
Characteristic Dummy	-0.43* (0.25)	-0.42*** (0.13)	0.094 (0.061)
Quadratic in PRT X PRT Bins	Yes	Yes	Yes
Quadratics in PRT X Bins X Characteristic	Yes	Yes	Yes
Mean of Dep. Var.	1.43	1.43	1.43
R-Squared	0.063	0.049	0.055
N	2433	2433	2433

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total and its interaction with the three prior record cutoffs (at least 25, at least 50, and at least 75) and the listed characteristic. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variable is the number of future felony sentences observed in the data. Results restricted to cases with an original charge of Retail Fraud.

**Table 8. First Stage for Various Empirical Strategies, Regressions of Days of Incarceration on Instruments**

	(1)	(2)	(3)	(4)	(5)	(6)
Sample:		Retail Fraud			OWI	
Strategy:	OLS	Judge IV	Discontinuity IV	OLS	Judge IV	Discontinuity IV
Days of Incarceration	1*** (0)			1*** (0)		
Judge Mean Days Incarc.		0.33*** (0.11)			0.56*** (0.1)	
Prior Record Total At Least 25			45.5** (17.6)			30.2** (12.2)
Prior Record Total At Least 50			58.2*** (20)			27.5 (27.3)
Prior Record Total At Least 75			39.9 (52.8)			151 (160.4)
Circuit X Year FE	Yes	Yes	No	Yes	Yes	No
Prior Record Total Quadratic	No	No	Yes	No	No	Yes
Mean of Dep. Var.	224.9	224.7	217.3	180.1	180	174.7
R-Squared	1	0.18	0.28	1	0.21	0.3
N	2752	2726	2435	6948	6939	6661

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. The columns indicate different estimation strategies and samples, and the rows display the instrument(s) corresponding to each strategy. The first row corresponds to simple OLS, the second to a judge IV, and the next three rows to each of the cutoffs in our discontinuity-based IV. OLS and Judge IV estimates control for circuit court × year fixed effects. The Judge IV instrument is the average sentence in other cases assigned to the same judge with different defendants. The discontinuity IV instruments are three indicators for prior record totals above the relevant cutoffs, and that specification controls for a quadratic polynomial in the prior record total and a quadratic polynomial interacted with the three cutoff dummies. OLS and the Judge IV cluster standard errors by judge, while the discontinuity IV clusters standard errors by prior record total.

**Table 9. Baseline Balance for Various Empirical Strategies, Coefficient on Days of Incarceration**

Sample:	(1)	(2)	(3)	(4)	(5)	(6)
		Retail Fraud			OWI	
Characteristic:	Prior Felonies	Prior Misdemeanors	Age at Current Offense	Prior Felonies	Prior Misdemeanors	Age at Current Offense
OLS	0.0087*** (0.0005)	0.0075*** (0.0009)	0.01*** (0.00)	0.0045*** (0.0003)	0.0066*** (0.0007)	0.002*** (0.001)
Judge IV	-0.0037 (0.0047)	-0.0054 (0.0050)	-0.01 (0.02)	0.0018 (0.0012)	-0.0028 (0.0027)	0.005 (0.005)
Discontinuity IV	0.0064 (0.0072)	0.0260 (0.0200)	0.0450* (0.025)	0.0080 (0.0057)	0.0150 (0.0140)	0.0190 (0.0180)

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Each cell shows the coefficient on days of incarceration with standard errors in parentheses. The columns indicate different dependent variables, namely characteristics that are determined prior to sentencing. The rows indicate different estimation strategies with different instruments. The first row conducts simple OLS. The second row instruments days of incarceration with average days incarcerated for other cases assigned to the same judge with different defendants. Both of the first two rows control for circuit court  $\times$  year fixed effects. In the third row, a set of three dummies for the sentencing guideline cutoffs at 25, 50, and 75 are the instruments, and we control for a quadratic polynomial in the prior record total and its interaction with the three prior record dummies. OLS and the Judge IV cluster standard errors by judge, while the discontinuity IV clusters standard errors by prior record total.



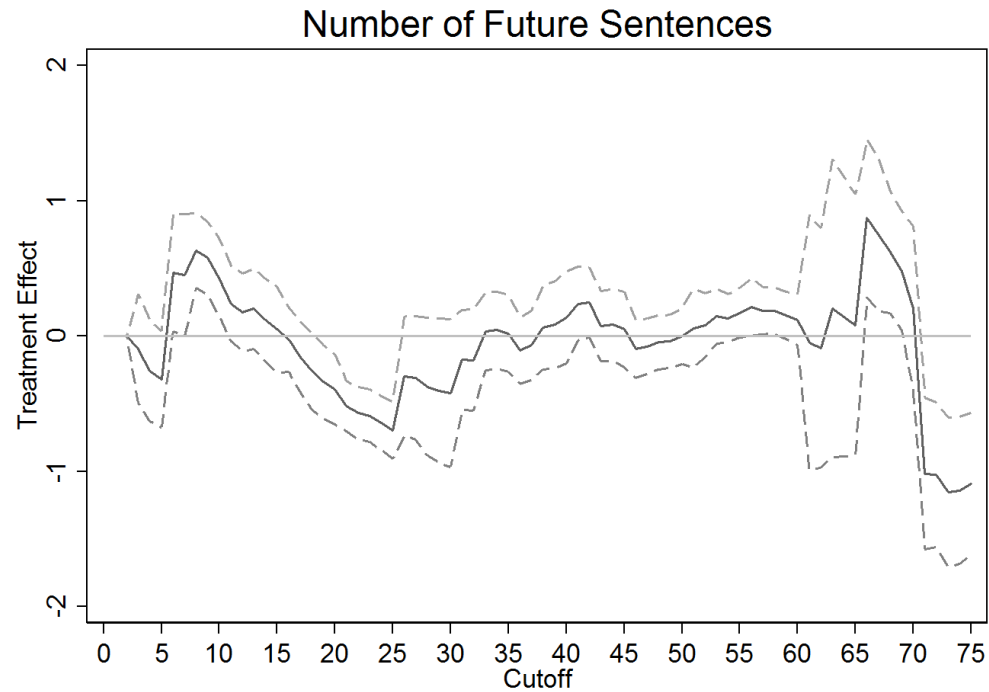
**Table 10. Recidivism Effects for Different Estimation Strategies**

Sample:	(1)	(2)	(3)	(4)	(5)	(6)
	Retail Fraud			OWI		
Estimation Method:	Discontinuity	Judge IV	OLS	Discontinuity	Judge IV	OLS
	IV			IV		
Days of Incarceration	-0.0055* (0.0029)	0.0034 (0.0028)	0.0004** (0.0002)	-0.0005 (0.0020)	-0.0012* (0.0007)	0.0003*** (0.0001)
Circuit × Year FE	No	Yes	Yes	No	Yes	Yes
Prior Record Total						
Quadratic	Yes	No	No	Yes	No	No
Mean of Dep. Var.	1.43	1.40	1.39	0.51	0.53	0.53
N	2,435	2,726	2,752	6,661	6,939	6,948

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors are in parentheses. The outcome is the number of future felony convictions. The columns indicate different estimation strategies. The judge IV instruments for days of incarceration with average sentences for other cases assigned to the same judge but with different defendants. The discontinuity IV instruments for days of incarceration with dummies reflecting prior record totals above the relevant cutoffs (25, 50, and 75). The Judge IV clusters standard errors by judge and controls for circuit court × year fixed effects. The discontinuity IV clusters standard errors by prior record total and controls for a quadratic polynomial in the prior record total and its interaction with the three prior record dummies.

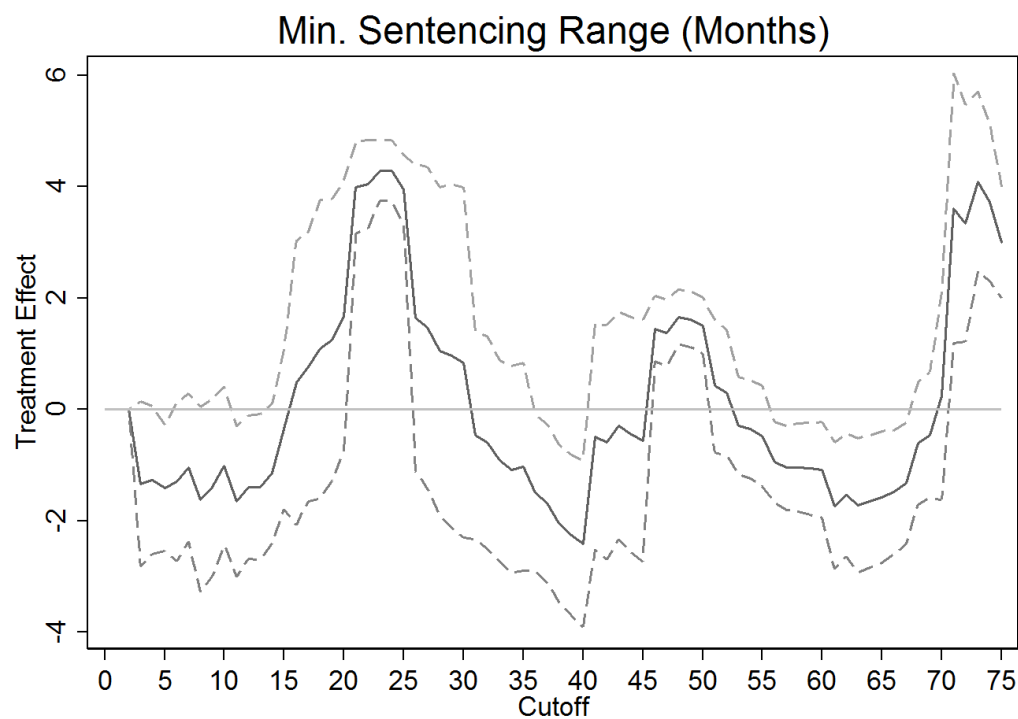
## Appendix Tables and Figures

**Appendix Figure 1. Placebo Effects at All Possible Cutoffs, 1 to 75, Effect on Recidivism**



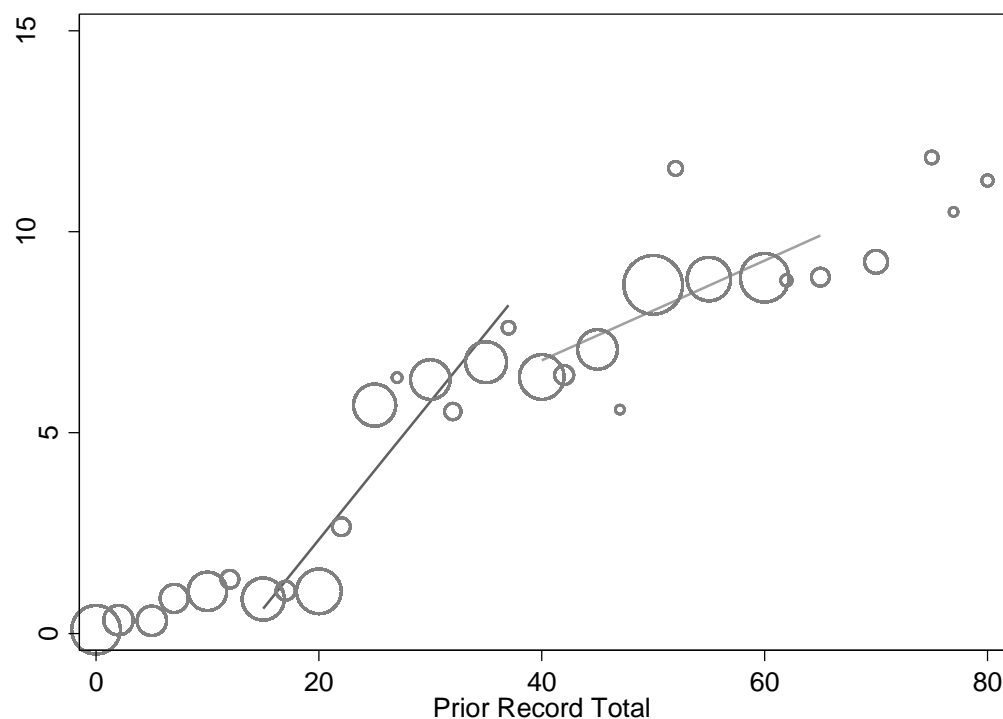
Each point on the solid line shows the measured discontinuity in recidivism assuming that a cutoff exists at the value on the horizontal axis. The dashed line shows a pointwise 95 percent confidence interval. See Appendix Figure 3 for details of the specification.

**Appendix Figure 2. Placebo Effects at All Possible Cutoffs, 1 to 75, Effect on Minimum of Sentencing Guidelines Range**



Each point on the solid line shows the measured discontinuity in the low end of the sentencing guidelines range assuming that a cutoff exists at the value on the horizontal axis. The dashed line shows a pointwise 95 percent confidence interval. See Appendix Figure 3 for details of the specification.

**Appendix Figure 3. An Example of Placebo Effects on the Minimum of the Sentencing Guidelines Range, Cutoff at 40**



Plotted circles show sample means for charges with the same prior record total. The size of the circle indicates the number of charges with that total. The solid curves display linear fits separately within 25 points above and below the assumed cutoff of 40. We use a linear fit because a quadratic fit exacerbates the spurious negative jump. The dependent variable is months of incarceration for the minimum of the sentencing guidelines range.

**Appendix Table 1. Robustness of Recidivism Effects at 25-Point Cutoff to Choice of Bandwidth and Functional Form**

	(1) Quadratic (bw = 25)	(2) Quadratic (bw = 30)	(3) Quadratic (bw = 20)	(4) Piecewise Linear
Retail Fraud	-0.40** (0.15)	-0.39** (0.15)	-0.087 (0.14)	-0.47*** (0.092)
OWI	-0.037 (0.086)	-0.039 (0.085)	-0.031 (0.13)	0.091 (0.089)
PRT Quadratic x 2 Bins	Yes	Yes	Yes	No
PRT Linear x Many Bins	No	No	No	Yes

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. The sample includes only sentences with prior record total less than 50. Each cell of the table shows the effect of moving across the 25-point cutoff in the guidelines, i.e. the coefficient on an “over 25” dummy. The first column replicates the main results with a bandwidth of the prior record total set to 25. The next two columns widen and narrow the bandwidth as listed. The final column estimates models that split the sample according to the many prior record total groups in Figure 1 and include dummies for these groups as well as their interaction with a linear control for the prior record total. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variable is the number of future felony sentences observed in the data.

**Appendix Table 2. Baseline Balance, 25-Point Cutoff**

Outcome:	Sample:	(1) RF/OWI Combined	(2) Retail Fraud	(3) OWI	(4) Other
Prior Felonies		-0.24 (0.28)	-0.16 (0.57)	-0.23 (0.21)	-0.53** (0.26)
Prior Misdemeanors		0.086 (0.99)	0.60 (1.72)	-0.34 (0.78)	-0.45 (1.27)
Prior Felonies (Sentencing Data)		-0.029 (0.057)	0.036 (0.17)	-0.040 (0.038)	-0.21 (0.17)
Age at Offense		1.34 (1.71)	4.11** (1.98)	-0.25 (1.64)	-0.036 (2.69)

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total, dummies for exceeding three prior record cutoffs (25, 50, and 75), and the interaction of the quadratic with the record categories. Each cell reports the coefficient on a dummy for prior record total of at least 25. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variables are characteristics determined prior to sentencing. “Prior felonies” records all felonies previously committed by the same person while “prior felonies (sentencing data)” counts only those prior felonies we can observe in the timeframe of our sentencing data.

**Appendix Table 3. Baseline Balance, 50-point Cutoff**

Outcome:	Sample:	(1) RF/OWI Combined	(2) Retail Fraud	(3) OWI	(4) Other
Prior Felonies		0.17 (0.42)	0.32 (0.55)	-0.045 (0.28)	-1.38*** (0.31)
Prior Misdemeanors		3.33*** (1.08)	4.24*** (1.13)	2.34** (1.03)	2.66*** (0.93)
Prior Felonies (Sentencing Data)		-0.24** (0.11)	-0.20 (0.25)	-0.27 (0.19)	-0.090 (0.082)
Age at Offense		1.56 (2.42)	2.04 (2.67)	1.40 (2.36)	1.79 (2.00)

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total, dummies for exceeding three prior record cutoffs (25, 50, and 75), and the interaction of the quadratic with the record categories. Each cell reports the coefficient on a dummy for prior record total of at least 50. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variables are characteristics determined prior to sentencing. “Prior felonies” records all felonies previously committed by the same person while “prior felonies (sentencing data)” counts only those prior felonies we can observe in the timeframe of our sentencing data.

**Appendix Table 4. Baseline Balance, 75-point Cutoff**

Outcome:	Sample:	(1) RF/OWI Combined	(2) Retail Fraud	(3) OWI	(4) Other
Prior Felonies		-0.43 (0.99)	-0.40 (1.01)	-2.49*** (0.83)	-0.14 (0.37)
Prior Misdemeanors		-3.92* (1.97)	-3.25 (2.39)	-5.28 (6.19)	-1.67 (1.32)
Prior Felonies (Sentencing Data)		-1.78*** (0.35)	-2.04*** (0.68)	-1.57 (1.12)	-0.069 (0.23)
Age at Offense		3.36* (1.90)	4.83** (1.84)	0.50 (3.05)	-0.85 (1.78)

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total, dummies for exceeding three prior record cutoffs (25, 50, and 75), and the interaction of the quadratic with the record categories. Each cell reports the coefficient on a dummy for prior record total of at least 75. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variables are characteristics determined prior to sentencing. “Prior felonies” records all felonies previously committed by the same person while “prior felonies (sentencing data)” counts only those prior felonies we can observe in the timeframe of our sentencing data.



**Appendix Table 5. Heterogeneous Effects by Sentence Type**

Sample:	(1) Retail Fraud	(2) Retail Fraud	(3) OWI	(4) OWI
	No Prison	Any Prison	No Prison	Any Prison
Prior Record Total at Least 25	-0.39** (0.15)	-1.24** (0.50)	-0.063 (0.087)	0.14 (0.26)
Prior Record Total at Least 50	0.20 (0.20)	-0.55*** (0.20)	0.36 (0.25)	0.0046 (0.28)
Prior Record Total at Least 75	-0.14 (0.26)	-2.10*** (0.76)	-0.74 (1.00)	-0.14 (0.30)
Prior Record Total Quadratic X PRT Bins	Yes	Yes	Yes	Yes
Mean of Dep. Var.	1.49	1.18	0.51	0.56
R-Squared	0.054	0.049	0.047	0.0089
N	1964	469	6167	491

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total interacted with the prior record categories. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variable is the number of future felony sentences observed in the data. The odd columns display results for the sub-sample not receiving state prison (i.e. incarceration  $\leq$  1 year) as part of their sentence and the even columns show results for those receiving state prison. These results should be interpreted carefully because sentencing to prison responds endogenously to the sentencing guidelines.

**Appendix Table 6. Differences in Baseline Characteristics for Prison vs. Jail**

Outcome:	Sample:	(1) RF/OWI Combined	(2) Retail Fraud	(3) OWI	(4) Other
Prior Felonies		0.34** (0.16)	0.93*** (0.26)	0.22* (0.13)	0.34 (0.27)
Prior Misdemeanors		-0.38** (0.18)	-0.51 (0.32)	-0.55* (0.29)	-1.16*** (0.37)
Prior Felonies (Sentencing Data)		0.17*** (0.06)	0.12** (0.055)	0.22** (0.11)	0.18* (0.11)
Age at Offense		0.70** (0.34)	0.31 (0.47)	1.42*** (0.45)	-0.56 (0.34)

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. All models are estimated by ordinary least squares including a quadratic polynomial in the prior record total and its interaction with the three prior record cutoffs (at least 25, at least 50, and at least 75). Each cell reports the coefficient on a variable indicating whether the sentence includes time in prison. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The dependent variables are characteristics determined prior to sentencing. “Prior felonies” records all felonies previously committed by the same person while “prior felonies (sentencing data)” counts only those prior felonies we can observe in the timeframe of our sentencing data.

**Appendix Table 7. Robustness of Recidivism Effects at the 25-Point Cutoff to Treatment of Survival Data**

Sample:	(1)	(2)	(3)	(5)	(6)	(7)
Specification:	Retail Fraud	Retail Fraud	Retail Fraud	OWI	OWI	OWI
	5-Year	Weight by	Cox-PH	5-Year	Weight by	Cox-PH
	Follow-Up	Obs. Years	Survival Model	Follow-Up	Obs. Years	Survival Model
Prior Record Total at Least 25	-0.40** (0.15)	-0.37** (0.15)	0.74*** (0.07)	-0.050 (0.084)	0.027 (0.058)	0.89 (0.07)
Prior Record Total Quadratic	Yes	Yes	No	Yes	Yes	No
Prior Record Total Quadratic $\times$ Above 25	Yes	Yes	No	Yes	Yes	No
Prior Record Total Exponential	No	No	Yes	No	No	Yes
Prior Record Total Exponential $\times$ Above 25	No	No	Yes	No	No	Yes

Statistical significance at the 1, 5, and 10 percent levels is denoted by \*\*\*, \*\*, and \* respectively. Standard errors clustered by the running variable are in parentheses. The first two models are estimated by least squares including a quadratic polynomial in the prior record total and a quadratic polynomial interacted with a prior record total of at least 25. The prior record total has been modified to exclude the prior record variable related to high severity felonies. The sample includes only sentences with prior record total less than 50. The dependent variable is the number of future felony sentences observed in the data. The first model uses ordinary least squares on a sample of observations with 5 years of follow-up data, and the second model uses all observations regardless of follow-up length and weights according to the time between the sentencing date for the original crime and the end of our data on August 1, 2013. The third model fits a Cox-proportional hazard survival model for time until first recidivism to all observations regardless of follow-up length accounting for censoring at August 1, 2013. All covariates in the proportional hazard model enter in an exponential functional form. The reported coefficient for the Cox-PH model is a hazard ratio; values less than 1 indicate a negative effect, and asterisks indicate statistical difference from 1.