To Plea or Not to Plea: Evidence from North Carolina

David Abrams[*] *and Ryan Fackler*[†]

Abstract

The decision to accept a plea bargain is one of the highest stakes decisions under uncertainty that an individual can make. It is also an extremely important source of unwarranted disparity in the criminal justice system. This paper undertakes the most detailed empirical study to date of the plea bargain decision. We use a dataset of approximately 375,000 observations from 12 years of court cases across North Carolina. In order to overcome omitted variables bias, we make use of an instrumental variables strategy. As expected, In North Carolina a defendant seeking to minimize their sentence length or chance of incarceration is generally better off accepting a plea. Behind this result, however, we find substantial heterogeneity. Most importantly, we find significant differences between the treatment of black and white defendants, even when addressing unobservable factors, Also in spite of North Carolina's efforts to create judicial homogeneity, we find substantial regional variation. These findings have important implications both for optimal decision-making by defendants and for structure and equity of the criminal justice system.

[*]University of Pennsylvania Law and Wharton. Email: dabrams@law.upenn.edu

[†]University of Pennsylvania Economics

The authors would also like to thank Bernardo Silveira and others for their assistance with this paper.

1 Introduction

For those interested in decision-making under uncertainty or high-stakes negotiation, the criminal plea bargain should be of paramount importance. Here is a situation where an individual literally bargains for years of his life to be spent largely as he chooses or in a prison. Theorists have recognized the importance of this situation and a number of important papers have modeled it (Landes, Reinganum, Daugherty and Reinganum, Grossman and Katz)[1]. But given the immensely high stakes of the defendant decision to accept or reject a plea offer, it has been severely understudied by empirical economists. Nor is the plea bargain a rare scenario rather, it is one that individuals in the U.S. face millions of times each year.

In criminal proceedings, prosecutors often make defendants an offer of a specific recommended sentence or sentencing range in exchange for the defendants guilty plea. The defendant, usually through his attorney, may accept the offer, reject it, or make a counter-offer. Well over 90% of sentences in the U.S. are determined through plea bargaining[2].

This paper makes the broadest and most in-depth empirical economic study of plea bargaining to date. We examine data on approximately 375,000 cases prosecuted in North Carolina state courts from 1998 to 2010. The depth of data allows us to uncover more nuanced phenomena about plea bargaining than previously possible.

If defendants are risk-neutral sentence-minimizers, we should expect the expected sentence to be approximately the same regardless of whether a defendant accepts or rejects a plea bargain. In Section III we elaborate on this simple equilibrium model of sentencing. There is also an out-of-equilibrium belief in a trial penalty that is widely held by practitioners in the criminal justice system. This view holds that defendants who reject a plea offer receive higher expected sentences as there are penalized for using extra judicial resources. We are able to test both of these theories in several ways.

We run OLS regressions using both statewide data and that from Charlotte, the largest metro area in North Carolina. We are able to control for defendant, case, attorney, judge, and jurisdiction characteristics. We find that on average defendants receive longer sentences when they reject pleas, providing some support for the "trial penalty" belief. But the findings are more complex; the expected sentencing gap (between accepting and rejecting a plea) varies substantially by type

[1] William M Landes *An Economic Analysis of the Courts* (1971), Jennifer F. Reinganum, *Plea Bargaining and Prosecutorial Discretion* (1988), Andrew F. Daughety and Jennifer F. Reinganum *Economic Theories of Settlement Bargaining* (2005), and Gene M. Grossman and Michael L. Katz *Plea Bargaining and Social Welfare* (1983)

[2] See Bureau of Justice Statistics *Plea and Charge Bargaining: Research Summary* (2011).

of crime, and by the race of the defendant.

Additionally, OLS regressions using statewide data shows that accepting a plea bargain significantly reduces the probability of being incarcerated by over 6%.[3] Again, this significant reduction in incarceration after pleading is not uniform. As shown in 4b, the impact of pleading on incarceration varies significantly by crimes types.

Much has been written about racial disparities in the criminal justice system. Thus it may come as little surprise that the plea decision also varies substantially by race. Black defendants on average obtain 10-month shorter sentences from pleading compared to 4-month shorter sentences for white defendants. For incarceration rate, however, the racial difference reverses; black defendants receive a 5.8% reduction in incarceration rates from pleading, while white defendants see a reduction of nearly 7.5%. Importantly, this manifests even though black defendants are incarcerated nearly 7% more than white defendants.

Attorneys and judges have enormous impact on case outcomes; court appointed private attorneys are associated with significantly longer average sentences than either public defenders or individually retained private attorneys[4]

Additionally, we explore the efficacy of North Carolina's significant efforts to create a homogenous court system state wide. One fundamental principle of North Carolina's structured sentencing system is is that "[s]entencing policies should be consistent: Offenders convicted of similar offenses, who have similar prior records, should generally receive similar sentences.[5]". We find that these efforts have not homogenized outcomes. There is tremendous variation across court divisions in terms of both discounts from pleading how the courts treat defendants of different races. For example, defendants in the third division expect a sentence discount from pleading that is over 5 months longer than similar defendants in the first or eighth divisions.

Of course, even when including controls for all observable characteristics there is the possibility that unobservable variables will bias the results. This we make use of an IV strategy to address this. We use the fact that agreement between defense and prosecution attorneys is more likely when there are less diffuse priors about the likely outcome at trial. Since attorneys gain information about a judges sentencing tendencies with experience, the distribution of priors should

[3] When we discuss incarceration, we reduce the sentence length a defendant receives into a binary variable which equals 100 if the sentence is greater than zero, and equals 0 if the sentence is 0.

[4] We estimate a fixed effect of 0.25 years associated with being represented by a court appointed attorney rather than a private attorney. This estimate controls for all observable characteristics.

[5] A Citizen's Guide to Structured Sentencing (2014) pg. 1 `http://www.nccourts.org/Courts/CRS/Councils/spac/Documents/citizenguide2014.pdf`

become tighter as the judge gains more public experience. This tightening of priors should then lead to a greater likelihood of a plea. We use the number of cases and trials argued before a judge as our measure of judicial experience, and find a positive and linear relationship with plea rate.[6]

The statewide IV estimates are larger than the corresponding OLS estimates, a result consistent with unobservably better cases being more likely to be plea than go to trial. Most interestingly, we see stark differences in our estimates of the effect of pleading on sentence length depending on the defendant's race. While we see no significant difference in sentence length for white defendants who plea rather than go to trial, the difference for black defendants is actually substantially larger than the OLS estimates. The difference in OLS estimates suggests that the court may treat defendants differently based on their race. The increased difference in our IV estimates additionally suggests that black defendants may be more inclined to go to trial than white defendants.[7]

The findings presented in this paper substantially extend the closest previous work on the topic, presented in Abrams (2011,2013). These papers found that rejecting plea bargains in Chicago lowered average sentences, but the data examined was not detailed enough to allow for separate estimates by type of crime, race, location, or attorney type.[8]. In Sections 4 and 6, we discuss several possible explanations for these findings, most notably North Carolina's use of a sophisticated structured sentencing scheme. This structured sentencing scheme reduces judicial leeway and may constrain the ability to negotiate substantial sentence length reductions.

This paper takes important steps forward in our understanding of defendant decision-making and clarifies the important role played by plea bargaining in the criminal justice system. The fact that both expected sentence and probability of incarceration are unequal depending on the decision to plea is evidence that the risk-neutral defendant is too simple. Rather, it is likely that defendant risk-aversion, agency problems, racial bias, and behavioral phenomena are behind some of our findings. We explore these further in Section VI.

The remainder of this article proceeds as follows: Section 2 situates our article within the relevant literature and Section 3 presents our model of the plea decision. Section 4 introduces the data and discusses features specific to the North Carolina court system and Section 5 presents the empirical findings. We end with a discussion of our results in Section 6 and conclude in Section 7.

[6] We use the percentile of the number of cases rather than the raw number of cases seen by a specific judge. As shown in Figure 6, the relationship between percentile of cases seen and plea rates is positive and linear.

[7] If black defendants are more inclined to go to trial, we should see on average worse cases for black defendants than white defendants at trial, and on average better cases for black defendants than white defendants who plea.

[8] As will be discussed further in Section 2, Abrams (2011) found evidence that plea bargains produce sentences between 4 and 21 months longer than those produced from trial.

2 Background

To situate this article within the relevant literature, we must understand how authors understand the choice of defendant to go to trial or take a plea bargain. As Jones (1978) describes this choice,

> Plea bargaining results from an agreement between the prosecutor, defense attorney, and occasionally the defendant. The prosecutor offers the defendant a *quid pro quo* (charge reduction or sentence recommendation) for pleading guilty. ... Moreover, the participants wield such discretion that in most cases they serve as the final arbiters in the sentencing decision.[9]

This negotiation described by Jones is fundamentally similar to the theory of The Shadow of the Law proposed by Mnookin and Kornhauser in their 1979 article discussing divorce settlements. The authors assert that despite the vast majority of divorces resolving with a negotiated settlement, the possibility of trial inherently frames any negotiation. It is natural to extend this theory to criminal cases — despite the vast majority of criminal cases resolving with plea bargains, the outcomes are inherently framed by the possibility of jury trials[10].

One important difference in Jones's story and the Shadow of the Law theory is the question of timing. Jones describes the trade as largely a discussion about *ex-post* outcomes. By pleading, the defendant obtains a reduced sentence as compared to what the defendant would have received were they convicted at trial (i.e. conditional on a finding of guilt). Similarly, the prosecutor willingly gives up sentence length in order to avoid to cost (both in time and effort) of going to trial. The Shadow of the Law theory, however, considers the defense and prosecution having an *ex-ante* negotiation. The defendant obtains a sentence from a plea bargain less than the *expected* sentence they receive from declining the plea bargain[11], and the prosecution both avoids the cost of trial and secures the benefit from a conviction with certainty.

The 'correct' way to understand outcome of plea bargaining is objectively contextually depen-

[9] J. B. Jones *Prosecutors and the Disposition of Criminal Cases: An Analysis of Plea Bargaining Rates* (1978)

[10] The shadow of the law theory has been employed by many others in a variety of contexts. For examples, see Cooter, Marks, and Mnookin *Bargaining in the Shadow of the Law: A Testable Model of Strategic Behavior* (1982) and Priest and Klein *The Selection of Disputes for Litigation* (1984).

[11] When we say expected sentence, this effectively is the sentence a defendant expects conditional on being found guilty at trial weighted by the probability that they would be found guilty at trial. Mathematically:

$$E\left[Sentence\right] = Sentence \cdot Pr(Found\ Guilty) + 0 \cdot Pr(Found\ Not\ Guilty)$$

dent. The *ex-post* discussion would be more appropriate for a lawmaker trying to determine the appropriate statutory punishment for a crime. Similarly, someone interested in understanding the punishments faced by individuals convicted of a certain crime should analyze the ex-post outcomes. When a defense attorney is advising their client on whether or not to accept a plea bargain, however, they arguably should focus on the ex-ante choice. Even if facing a potential 3 years in prison given conviction, no defendant would accept a plea bargain of 1 year in jail if they believe there is only a 1% chance of conviction at trial. A defendant who believes they face a 75% chance of conviction at trial, however, would be much more inclined to accept the offer.

More recently, the application of the Shadow of the Law theory to criminal cases has faced some criticism. Bibas discusses how institutional features of the court system as well as well-documented Behavioral Economics phenomena complicate the Shadow of the Law [12]. Stuntz points out how prosecutor discretion in case selection and differing prosecutor goals may be problematic for the Shadow of the Law theory[13].

While imperfect, the Shadow of the Law theory still provides a natural entry point to understand the plea bargaining decision. Additionally, because this article focuses on the defendant's choice to accept a plea bargain of a certain length, we will exclusively use the *ex-ante* perspective. This choice of focus distinguishes our work from the majority of the existing empirical literature on plea bargaining, and puts us theoretically squarely in the Shadow of the Law mindset.

The earliest empirical discussion of the plea bargaining process was published by Rhodes (1979), who examine ex-post sentences from plea bargains compare to ex-post sentence from trial[14]. Using data on arrests in Washington, D.C. in 1974, Rhodes shows no distributional differences for charges of larceny, burglary, and assault. Rhodes does find that probation is a more common sentence and jailtime of 3+ years is less common after pleading guilty to the charge of robbery (as opposed to sentences condition on conviction at trial).

This ex-post analysis of plea bargaining continued with the work of Brereton and Casper[15]. Using data from three California jurisdictions in the late 1970s, the authors test for the equality of rates of imprisonment between those who plead guilty and those who go to trial. While the authors control for some relevant observable measures[16], they examine only the fraction of defendants who were imprisoned. This is problematic on two fronts: 1) by looking at imprisoned rather than

[12] Stephanos Bibas *Plea Bargaining Outside the Shadow of Trial* (2004).

[13] W. Stuntz *Plea Bargaining and Criminal Laws Disappearing Shadow* (2004)

[14] W.N. Rhodes *Plea Bargaining: Its Effect on Sentencing and Convictions in the District of Columbia* (1979)

[15] Brereton and Casper *Does it Pay to Plead Guilty?* (1982)

[16] The authors include controls for: jurisdiction, two types of crime (robbery and burglary), prior record, number of charges, and seriousness of charge.

incarcerated, the authors ignore the sentences less than a year, and 2) the authors only examine the rate of imprisonment and not sentence lengths. The first problem causes sentences of less than a year to be discounted unnecessarily. The second, more severe problem, prohibits the authors from assessing the more plausible margin to observe a difference in outcomes; differences in outcomes reasonably manifest in differences in sentence lengths and not necessarily a difference in the rate of jailtime[17]

The discussion of ex-post outcomes from trials and plea bargains continued to be a popular topic, with authors analyzing the dynamics of a variety of different court systems. Schulhofer and Nagel (1989, 1992, and 1997) discuss how the federal sentencing guidelines build in a 35 percent sentence reduction in sentence length for pleading[18]. This happens as a consequence of the U.S. Sentencing Guidelines Manual permitting a reduction in sentence for "accepting responsibility" for one's action[19]. Despite the possibly 35 percent ex-post discount, Ulmer et al. (2010) estimate only between a 3 and 15 percent penalty[20] reduction in sentences from plea bargains.

Prior to Abrams (2011), the only author to empirically discuss how defendants make the ex-ante choice to plea or go to trial was Smith (1986), who examined whether a defendant receives a prison sentence of at least a year[21]. Controlling for several observable case and defendant characteristics, he finds that there was no statistically significant difference between the unconditional expected sentence from trial and that obtained after a plea bargain. As the author does not address potentially unobserved selection issues, we still may worry of bias in these estimates.

It is in this conceptual position in Abrams (2011,2013) that the literature must empirically understand the ex-ante choice to plea. Using data from Cook County, Illinois state courts[22], he finds direct evidence that defendants are not minimizing expected sentences through plea bargains. In fact, his OLS estimates provide evidence that a risk-neutral defendant could expect a shorter sentence by declining a plea bargain than by accepting. Unlike previous literature, Abrams (2011) used IV regression to control for unobservable selection issues. These estimates provide no evidence against his first surprising result.

[17] For example, as North Carolina uses a structured sentencing scheme with minimum active sentences proscribed given guilt of many crimes, we would expect no difference in ex-post incarceration rates.

[18] See Schulhofer and Nagel *Negotiated Pleas Under the Federal Sentencing Guidelines: The First Fifteen Months* (1989), Nagal and Schulhofer *Tale of Three Cities: An Empirical Study of Charging and Bargaining Practices Under the Federal Sentencing Guidelines* (1992), and Schulhofer and Nagel *Plea Negotiations Under the Federal Sentencing Guidelines: Guideline Circumvention and its Dynamics in the Post-Mistretta Period* (1997).

[19] See U.S. Sentencing Guidelines Manual 3E1.1 (2010).

[20] Ulmer, Eisenstin, and Johnson *Trial Penalties in Federal Sentencing: Extra-Guidelines Factors and District Variation* (2010).

[21] Douglas Smith *"The Plea Bargaining Controversy"* (1986).

[22] See Abrams (2011,2013)

Abrams (2011) presented a counter-intuitive result, that defendants would on average be better off going to trial, and unsurprisingly was received with some skepticism[23]. The most compelling critique comes from Alschuler (2013), who criticized (among other features) the construction of the dataset from Cook County, Illinois[24]. In particular, Alschuler commented "He [Abrams] was confident that had he actually compared post-trial sentences to post-guilty-plea sentences, the posttrial sentences would have been less severe. Abrams might be correct, and it would be worth the effort to find out. The following section of this article considers circumstances that might explain findings like Abrams if these findings were replicated in a better conceived and better executed study " (691).

With Alschuler's request squarely in mind, the current paper proceeds as a conceptual descendant of Abrams (2011,2013). The current article uses a significantly larger dataset involving all felony cases in the state of North Carolina between 1995 and 2009. After significant cleaning of the data to ensure it truly characterizes the observed incidents, we retain over 300,000 observations[25].

3 Theory

We begin with a simple mathematical model describing a criminal defendant's choice to go to trial rather than accept a plea bargain. This conceptual framework was introduced by Priest and Klein (1994) and has been subsequently used in Abrams (2011) in analyzing the choice to take a plea in Chicago.

Assume that for every criminal defendant, all information relevant to the case is known, and that this information can be summarized by a single variable Y. We may think of Y as the weight of evidence that a given defendant is guilty of their charged crimes. Let Y^* represent the burden of proof necessary for a finding of guilt. We denote the realized value of Y for a particular case as

[23] It is worth mentioning A.C. Kim *Underestimating the Trial Penalty: An Empirical Analysis of the Federal Trial Penalty and Critique of the Abrams Study* (2014). This paper seems to fundamentally misunderstand the conceptual difference between the ex-ante and ex-post outcomes. Additionally statements such as "Abrams's methodology implicitly assumes that defendants who pled guilty would have had the same odds of being acquitted as those defendants who actually went to trial" suggests a lack of understanding about the entire discussion of observable and unobservable endogeous group selection. This is explicitly not the assumption that is made in Abrams (2011,2013). This assumption would obviate the need for instrumental variables thereby rendering pages 206-207 and 214-218 unnecessary.

[24] See Albert Alschuler *Lafler and Frye: Two Small Band-Aids for a Festering Wound* (2013).

[25] It is important to acknowledge, and will be described in Section 4 that these remaining observations are not capturing the entire universe of felony cases in North Carolina. For example we exclude homicide cases, cases only involving the violation of probation, and crucially for the identification strategy, cases where we cannot identify the judge.

Y'. Therefore, the outcome of a trial will be:

$$\text{Defendant is found guilty if } Y' > Y^*$$
$$\text{Defendant is found not guilty if } Y' \leq Y^*$$

As in any situation, however, there is uncertainty about the true value of Y'. We will model this uncertainty as if the defendants and prosecutors receive a noisy signal about the true value of Y':

$$\hat{Y}'_p = Y' + \varepsilon_p$$
$$\hat{Y}'_d = Y' + \varepsilon_d$$

Where ε_p and ε_d are the prosecution and defendant's uncertainty about Y' respectively. We will assume that ε_p and ε_d are drawn from the same distribution, specifically:

$$\varepsilon_p, \varepsilon_d \sim N(0, \sigma^2)$$

As $E\left[\varepsilon_p\right] = E\left[\varepsilon_d\right] = 0$, this means that $E\left[\hat{Y}'_p\right] = E\left[\hat{Y}'_d\right] = Y'$. Thus, in expectation, both the prosecution and defendant's beliefs about Y' are correct (i.e. they have rational expectations). Without any loss of generality, we can normalize $Y^* = 0$.

As both the prosecution and the defendant know $\varepsilon_p, \varepsilon_d \sim N(0, \sigma^2)$, given their respective signals $\hat{Y}'_p$ and $\hat{Y}'_p$, they form beliefs about the probability of a finding of guilt at trial:

Prosecution	*Defendant*
$P_p = Pr(Y' > 0 \mid \hat{Y}'_p)$	$P_d = Pr(Y' > 0 \mid \hat{Y}'_d)$
$= Pr(\hat{Y}'_p - \varepsilon_p > 0)$	$= Pr(\hat{Y}'_d - \varepsilon_d > 0)$
$= Pr(\hat{Y}'_p > \varepsilon_p)$	$= Pr(\hat{Y}'_d > \varepsilon_d)$
$= F(\hat{Y}'_p)$	$= F(\hat{Y}'_d)$[26]

Now, for a given charge, let us temporarily assume that the length of a jail sentence is fixed upon conviction is fixed at a length of J. Given this assumption, the expected jail time that the defendant and prosecution can expect given a trial is:

$$E\left[J \text{ at trial}\right] = E\left[P_p \cdot J\right] = E\left[P_d \cdot J\right] = F(Y') \cdot J$$

[26] Because we have assumed that $\varepsilon_p, \varepsilon_d \sim N(0, \sigma^2)$, we know that $F(\hat{Y}'_p) = \Phi(\frac{\hat{Y}'_p}{\sigma})$ and $F(\hat{Y}'_d) = \Phi(\frac{\hat{Y}'_d}{\sigma})$, where $\Phi(\cdot)$ corresponds to the standard normal distribution.

3.1 Choice to Plea vs. Trial

Let us define:

Prosecution	*Defendant*
$C_p \equiv$ cost of going to trial	$C_d \equiv$ cost of going to trial
$S_p \equiv$ cost of settling	$S_d \equiv$ cost of settling

If we assume that both the prosecutor and the defendant are risk neutral and linearly value jail time, then the prosecutor would be willing to accept any sentence length A, so long as:

$$A - S_p \geq P_p \cdot J - C_p$$

Thus, we can define the minimum sentence, $\underline{A}$, the prosecutor would accept as:

$$\underline{A} \equiv P_p \cdot J - C_p + S_p$$

Similarly, we can define the maximum sentence the defendant would be willing to accept, $\overline{B}$ as:

$$\overline{B} \equiv P_d \cdot J + C_p - S_p$$

This allows us to conclude that the two parties will go to trial if:

$$\underline{A} > \overline{B} \text{ or } P_p - P_d > \frac{C - S}{J}$$

where $C = C_d + C_p$ and $S = S_d + S_p$.

From this logic, we conclude that a trial will occur if the minimum sentence a prosecutor will accept is greater than the maximum sentence that a defendant will accept. This model has features that match intuition. Trials are more likely in cases where prosecutors have a high expectation of conviction, while defendants have a lower expectation. Increasing the cost of going to trial makes a trial less likely to occur, while increasing the cost of settling makes a trial more likely to occur. Trials are also more likely as the length of sentence given conviction increases (assuming settlement costs are less than trial costs). This occurs because the relative cost savings from settling rather than going to trial decease in sentence length[27]

If, however, $\underline{A} \leq \overline{B}$, then any negotiated sentence length $J' \in \left[\underline{A}, \overline{B}\right]$ from a plea bargain would be preferable to going to trial for both the prosecution and defendant. If we assume that $C_p = S_p$ and $C_d = S_d$, then the defense and prosecution would be willing to accept any negotiated sentence

[27] This is directly visible from the term equation $P_p - P_d > \frac{C-S}{J}$. Assuming $C - S > 0$, then $\frac{C-S}{J}$ falls as J increases. This increases the likelihood that $P_p - P_d$, or that the two parties go to trial.

length $J' \in \left[P_p \cdot J, P_d \cdot J\right]$. By our assumption of rational expectations, we have that $E\left[P_p\right] = E\left[P_d\right]$, and therefore:

$$E\left[J \text{ from plea}\right] = E\left[P_p \cdot J\right] = E\left[P_d \cdot J\right] = J \cdot F(Y')$$

Inspecting above, we see that this produces the result that $E\left[J \text{ from plea}\right] = E\left[J \text{ at trial}\right]$. So, if there is no cost difference between going to trial and settling, we would expect to see no difference between the expected negotiated sentence and the expected sentence at trial.

Allowing now the costs to differ, we can discuss the "shadow of the law" result. Let us define:

$$D \equiv \frac{C - S}{J}$$

Assuming $D > 0$, our decision to go to trial can be reframed as[28]:

$$\text{Trial if } F(\hat{Y}'_p) - F(\hat{Y}'_d) > D \tag{1}$$

Additionally, because of our assumption of rational beliefs,

$$E\left[F(\hat{Y}'_p) - F(\hat{Y}'_d)\right] = 0 < D \tag{2}$$

Equation (2) tells us that in expectation, cases will be settled with a plea bargain and not a trial, a feature that matches both intuition and reality. Equation (1) tells us that *some* cases will go to trial, and that these cases will be those with the greatest divergence between prosecutor confidence and defendant pessimism.

This paper seeks to determine whether the expected sentence length differs between a plea bargain and a trial. We as researchers cannot possibly observe $\hat{Y}'_p$ and $\hat{Y}'_d$. Moreover, we cannot even observe a complete panel of variables necessary to fully determine $\hat{Y}'_p$ and $\hat{Y}'_d$, as there invariable are unobservable characteristics. As such, we will almost certainly have estimation bias if we ignore the fact that the choice to take a plea bargain is a function of unobservable case quality[29]

To properly handle this estimation bias, we will need to construct an instrument for the probability of taking a plea bargain that is unrelated to the probability of guilt at trial. Given our above

[28] This assumption is not strictly necessary. If $C < S$, then $D < 0$ for all possible sentence lengths. Since the prosecutor and defendant will end at trial if $P_p - P_d > \frac{C-S}{J} = D$, and because $E\left[P_p - P_d\right] = 0$, having $D < 0$ would imply that in expectation each case will go to trial. Moreover, it will imply that, in actuality, a majority of cases will go to trial. This result would be wildly inconsistent with reality where fewer than 10% of cases go to trial. Thus, we can safely restrict our attention to the case where $D > 0$.

[29] One unobservable characteristics is the underlying true guilt of the defendant. It is possible that in situations where the defendant is actually guilty, the defendant and prosecutor gets a higher values for $\hat{Y}'_p$ and $\hat{Y}'_d$. This would raise both P_p and P_d causing a change in defendant and prosecutor behavior that is completely unmeasurable to us as researchers.

model, any factor that randomly disturbs settlement costs, trial costs, sentence lengths, the distribution of Y, or the distribution of the ε's would be suffice – all variables that enter into Equation (1).

Any variable that increases (decreases) $Var\left[F(\hat{Y}'_p) - F(\hat{Y}'_d)\right]$ increases (decreases) the probability that Equation (1) is satisfied (i.e. that the parties go to trial). Thus, any such variable that satisfies our exclusion restriction[30] could serve as a good instrument for whether the case goes to trial.

Up to this point, we have introduced uncertainty only insofar as the defense and prosecutor receive a noisy signal about the true case quality Y'. We could easily change the source of uncertainty to be Y^*, the bar for conviction. This uncertainty, ϵ_p, $\epsilon_d \sim N(0, \sigma^2)$ would manifest because different judges will have idiosyncratically different definitions of "reasonable doubt" in a case, and the formulae above will be functionally unchanged[31]. Given this new source of error, any information about judge proclivities will decrease the variance of ϵ_p and ϵ_d. This will directly decease σ^2, which reduces $Var\left[F(\hat{Y}'_p) - F(\hat{Y}'_d)\right]$.

The primary way where prosecutors and defendants learn about judge proclivities is through inspection of prior decisions. Judges new to the bench will have few prior decisions, thereby providing little information about σ^2 for that judge. Conversely, an experienced judge will have a large amount of prior casework, allowing both parties to develop accurate assessments of σ^2. We will now show that changing σ^2 changes, $Var\left[F(\hat{Y}'_p) - F(\hat{Y}'_d)\right]$, and can serve as an instrument for the likelihood of pleading:

First, let us denote $Var\left[F(\hat{Y}'_p) - F(\hat{Y}'_d)\right]$ as σ^2_τ. Then, by definition:

$$\sigma^2_\tau = E\left[\left(F(\hat{Y}'_p) - F(\hat{Y}'_d)\right)^2\right] - \underbrace{E^2\left[F(\hat{Y}'_p) - F(\hat{Y}'_d)\right]}_{=0 \text{ from Equation 2}}$$

$$\sigma^2_\tau = E\left[\left(F(\hat{Y}'_p) - F(\hat{Y}'_d)\right)^2\right]$$

$$\sigma^2_\tau = E\left[F(\hat{Y}'_p)^2 - 2F(\hat{Y}'_p)F(\hat{Y}'_d) + F(\hat{Y}'_d)\right]$$

As ϵ_p and ϵ_d are drawn from the same distribution, we have:

$$\sigma^2_\tau = 2E\left[F(Y')^2\right] - 2E\left[F(\hat{Y}'_p)F(\hat{Y}'_d)\right]$$

[30] We have a situation where the choice to go to trial is correlated with unobserved case characteristics. Prosecutors are more likely to require a trial for cases with smaller ε_p (ie. easier cases to prosecute), as such $E\left[\varepsilon_p | trial\right] \neq 0$. Thus, our exclusion restriction is that we need a variable that is related to the choice to go to trial that is uncorrelated with ε_p.

[31] We have chosen to present the above model with the error on the case quality and not judge proclivity for ease of exposition only. The intuition is slightly more natural, but the logic is identical.

Lastly, since ϵ_p and ϵ_d are independent, we have:

$$\sigma_\tau^2 = 2E\left[F(Y')^2\right] - 2E^2\left[F(Y')\right]$$
$$\sigma_\tau^2 = 2Var\left[F(Y')\right]$$

As $F(Y') = \Phi(\frac{Y'}{\sigma})$, we have that σ_τ^2 is monotonically increasing in σ. This result is shown using numerical methods.

3.2 Comparative Statics

The above model allows us to also make predictions about how case outcomes should vary with changes to parameters. By identifying circumstances with predictable differences in parameters levels, we can empirically evaluate the consistency of our model. We will first focus on how the plea rate changes with our various parameters, and then talk about the rate of guilt at trial.

Equation (1) tells us that for a given Y_p' and Y_d', a trial will happen if $F(\hat{Y}_p') - F(\hat{Y}_d') > \frac{C-S}{J}$. Accordingly, the probability that a trial happens is:

$$\begin{aligned} Pr(trial) &= Pr\left[F(\hat{Y}_p') - F(\hat{Y}_d') > \frac{C-S}{J}\right] \\ &= Pr\left[\Phi(\frac{\hat{Y}_p'}{\sigma_p}) - \Phi(\frac{\hat{Y}_d'}{\sigma_d}) > \frac{C-S}{J}\right] \end{aligned}$$

Without needing to specify the exact form of this probability, we can directly see how this probability will change with the parameters. So, let us consider how the parameters impact the probability a case goes to trial:

1. $\frac{\delta Pr(trial)}{\delta C} < 0$: Increasing C (either C_d or C_p) will make the inequality harder to satisfy.
2. $\frac{\delta Pr(trial)}{\delta S} > 0$: Increasing S (either S_d or S_p) will make the inequality easier to satisfy.
3. $\frac{\delta Pr(trial)}{\delta J} > 0$: Increasing J will make the inequality easier to satisfy.
4. $\frac{\delta Pr(trial)}{\delta \sigma_d} > 0$: This effect is harder to see as raising σ_d increases $\Phi(\frac{\hat{Y}_d'}{\sigma_d})$ if $Y_d' < 0$ and decreases $\Phi(\frac{\hat{Y}_d'}{\sigma_d})$ if $Y_d' > 0$. Overall, since $\Phi(\frac{\hat{Y}_p'}{\sigma_p}) - \Phi(\frac{\hat{Y}_d'}{\sigma_d}) > \frac{C-S}{J}$ is hardest to satisfy when $\hat{Y}_d'$ is large, these two effects net to increase $Pr(trial)$.
5. $\frac{\delta Pr(trial)}{\delta \sigma_p} > 0$: Similarly, since $\Phi(\frac{\hat{Y}_p'}{\sigma_p}) - \Phi(\frac{\hat{Y}_d'}{\sigma_d}) > \frac{C-S}{J}$ is hardest to satisfy when $\hat{Y}_p'$ is small, increasing $\Phi(\frac{\hat{Y}_p'}{\sigma_p})$ increases $Pr(trial)$.
6. $\frac{\delta Pr(trial)}{\delta Y'} < 0$: If we assume the pool of defendants have Y' coming from a symmetric distribution centered at 0, then increasing the mean above 0 will reduce the probability Equation

(1) will be satisfied.

Given these predictions, we can now ask under what circumstances can we reasonably make predictions about changes to our parameters. We have identified five margins along which comfortable making predictions about the impact on the parameters: 1) defendant race, 2) defendant sex, 3) defense attorney ability, 4) defense attorney type, and 5) average sentence for the crime.

1. Race: The presence of any manifestation of taste based discrimination against black defendants would manifest as if Y' for a black defendant were increased. Thus, the distribution of Y' for black defendants would have their mean increased, which makes satisfying Equation (1) harder. This lowers the probability of trial for black defendants relative to white.
2. Female: We might expect a female defendant to be harder to convict at trial than an equal male defendant. This manifests with lower average Y' for female defendants, which would lower the rate of trials relative to male defendants.
3. Defense attorney ability: We might expect a better defense attorney to be able to secure a lower sentence at trial J. This should make trials more appealing, thereby raising the rate of trials relative to lower skilled attorneys.
4. Defense attorney type: The pay-structure and other caseload of public defenders, privately appointed, and privately retained attorneys all differ. The substantial caseloads of public defenders should make going to trial a more costly. The higher C_d for public defenders should mean a lower rate of trials.
5. Average Sentence for Crime: A low-level offense should have a lower J than a more serious crime. As such, we expect more trials for more serious offenses.

Now let us turn our attention to the probability of guilt at trial. To analyze how the rate of guilt changes with our parameters, we are going to make a simplifying assumption — $F(\hat{Y}'_p) = g$. That is, the prosecution doesn't receive a noisy signal about guilt but simply has a fixed belief about the likelihood of guilt at trial. As such, we modify Equation (1) and have:

$$\begin{aligned} Pr(trial) &= Pr\left[g - F(\hat{Y}'_d) > \frac{C-S}{J}\right] \\ &= Pr\left[g - \Phi(\frac{\hat{Y}'_d}{\sigma_d}) > \frac{C-S}{J}\right] \end{aligned}$$

This means that the probability of guilt at trial can be expressed as:

$$Pr(guilty \mid trial) = \mathbb{E}\left[Pr\left[Y' > 0 \mid Y'_d, trial\right]\right]$$

Again, without needing to specify the exact probability we can characterize the changes to the probability of guilt at trial:

1. $\frac{\delta Pr(guilty \mid trial)}{\delta C} < 0$: Increasing C (either C_d or C_p) will make the inequality harder to satisfy, so only lower Y'_d go to trial. As probability of guilt is lower if Y'_d is lower, the average rate of guilt at trial should decline..
2. $\frac{\delta Pr(guilty \mid trial)}{\delta S} > 0$: Increasing S (either S_d or S_p) will make the inequality easier to satisfy, which should raise the average Y'_d at trial.
3. $\frac{\delta Pr(guilty \mid trial)}{\delta J} > 0$: Increasing J will make the inequality easier to satisfy, which should raise the average Y'_d at trial.
4. $\frac{\delta Pr(guilty \mid trial)}{\delta Y'} > 0$: If we assume the pool of defendants have Y' coming from a symmetric distribution centered at 0, then increasing the mean above 0 will raise the probability of guilt at trial. Raising Y's mean should increase the share of Y' at trial that are greater than 0.

So, given these predictions, we can again ask what we would predict across our five margins:

1. Race: Raising Y' for black defendants should increase the probability of guilt at trial.
2. Female: Lowering Y' for female defendants should decrease the probability of guilt at trial.
3. Defense attorney ability: Lower J should make worse cases go to trial, thereby raising the rate of guilt.
4. Defense attorney type: The higher C_d for public defenders should mean a lower guilt rates at trial.
5. Average Sentence for Crime: A low-level offense should have a lower J than a more serious crime. As such, we expect more guilt at trials for more serious offenses.

Lastly, let us examine how the average sentence from a plea bargain should change with our parameters. To do so, we continue with our above simplified Equation (1). Without specifying exactly how a plea bargain is reached when it is, let us denote $q(C_d, S_d, C_p, S_p, J, Y'_d)$ the length. Given this, the average sentence from plea is:

$$\mathbb{E}_{Y'}\left[\ \mathbb{E}\left[q(\cdot, Y'_d)\right]\ \mid Y' \text{ and } plea\right]$$

That is, the average sentence at plea can be thought of as the average sentence at plea for each Y' averaged over all Y'. If we assume that Y' is symmetrically distributed around 0, then the following can be shown[32]

[32] Given our modification to Equation (1), then for each Y', plea bargains will happen based on the following cutoff rule: Plea if: $Y'_d \geq \sigma_d \Phi^{-1}\left[\frac{S-C}{J} + g\right] = \underline{Y}_d$. Given this, if we assume $Y' \sim N(0, \sigma_y)$, then the average sentence from

1. $\frac{\delta q(\cdot)}{\delta C} > 0$: Increasing C makes trials less appealing, so defense attorneys will accept higher sentences for any Y'_d.
2. $\frac{\delta q(\cdot)}{\delta S} < 0$: Increasing S makes pleading less appealing, so defense attorneys will demand lower sentences for any Y'_d.
3. $\frac{\delta q(\cdot)}{\delta J} > 0$: Increasing J makes trials more appealing, so defense attorneys will accept higher sentences for any Y'_d.
4. $\frac{\delta q(\cdot)}{\delta Y'} > 0$: If we shift the distribution of Y' up, we will increase the likelihood of large Y'. This increases the likelihood of observing larger Y'_d. Since $q(\cdot)$ is increasing in Y'_d, this raises our expected $q(\cdot)$.

Given these predictions, we can again hypothesize that we'll see:

1. Race: Raising Y' for black defendants should raise the average sentence from plea for black relative to white defendants..
2. Female: Lowering Y' for female defendants should lower the average sentence from plea relative to male defendants.
3. Defense attorney ability: Lower J should lower the average sentence from plea bargains
4. Defense attorney type: The higher C_d for public defenders should mean a higher average sentences from plea for public defender clients
5. Average Sentence for Crime: A low-level offense should have a lower J than a more serious crime. Accordingly, we expect average sentence from plea to increase with charge severity.

4 Data

In this paper, we analyze data from the North Carolina courts. With nearly 10 million residents, North Carolina is a tremendously diverse state with substantial geographic and socio-economic variation. The state is broken into 100 counties, including eight metropolitan areas,[33] with county

pleas can be expressed as:

$$\int_{-\infty}^{\infty} \left[\int_{\underline{Y}_d}^{\infty} \frac{q(\cdot, Y'_d) \cdot \phi(\frac{Y'_d - Y'}{\sigma_d})}{1 - \Phi(\underline{Y}_d)} dY'_d \right] \phi(\frac{Y'}{\sigma_y}) dY'$$

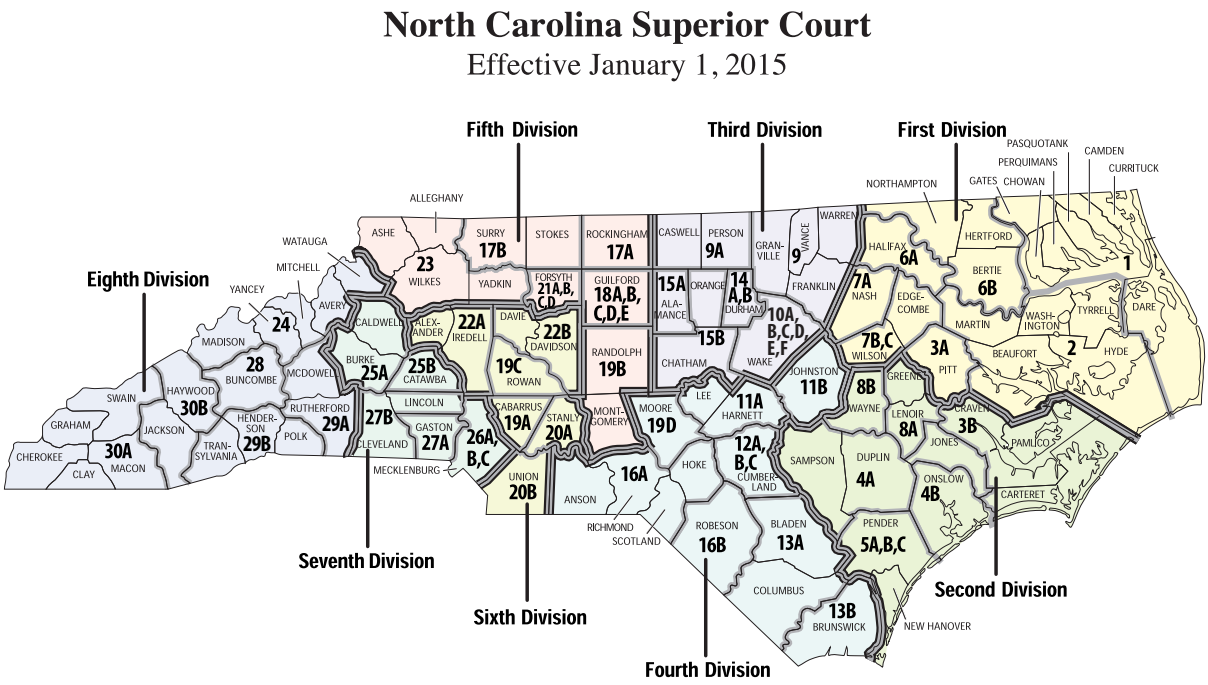
The above partial derivatives can be verified from this with assumptions about the sign of $\frac{\delta q(\cdot)}{\delta X}$ for various relevant X.

[33] We use the Office of Management and Budget's definition of a combined statistical area (CSA) as our definition of a metropolitan area. The eight CSAs in North Carolina are: 1) Charlotte-Gastonia-Salisbury, 2) Raleigh-Durham-Cary, 3) Greensboro-Winston-Salem-High Point, 4) Rocky Mount-Wilson, 5) Asheville-Brevard, 6) Fayetteville-Lumberton-Laurinburg, 7) Greenville-Washington, and 8) New Bern-Morehead City.

populations ranging from 4,115 people in Tyrrell County to 1,012,539 people in Mecklenburg County (Charlotte).[34]

North Carolinas court system has two trial courts: the District Court Division for misdemeanors and the Superior Court Division for felonies. The Superior Court is broken into 50 districts, which are aggregated into eight divisions. These divisions and districts are shown in Figure 1.[35] In the 2005-6 fiscal year, 109,815 felony cases were filed in Superior courts. There were a total of 109 Superior Court judges, 514 prosecutors, and 197 public defenders.

Figure 1: North Carolina Superior Courts



Note: Districts that have more than one letter associated with the district number (i.e., 10A, B, C, D) are divided into separate districts for electoral purposes. For administrative purposes, they are combined into a single district.

[34] Population estimates from the U.S. Census Bureau's 2014 population estimates. See: `http://quickfacts.census.gov/qfd/maps/north_carolina_map.html`.

[35] This map is current as of 2015 and is available at http://www.nccourts.org/Courts/Trial/District/Documents/SuperiorCourtmap.pd

Twice annually, the Administrative Office of the Courts publishes a publicly accessible Master Schedule detailing the intended judge service for the coming six months. The schedule lays out the intended location for each judge on a weekly basis, and whether the judge will be handling criminal, civil, or administrative issues. Unfortunately, while these Master Schedules are publicly available, the court does not publish a finalized schedule of where the judges actually held court each week. Informal conversations with court staff suggest that changes to the schedule are not uncommon, with changes occurring due to illness, vacations, and case-load needs.

Another feature of the North Carolina court system is the use of a structured sentencing scheme. Implemented through the Structured Sentencing Act of 1993, North Carolina's structured sentencing program separates felony charges into ten classes (Class A as the most severe and Class I the least) and convicted felons into six different criminal history levels (Level I the least severe and Level VI the most). Over the 20 years of North Carolina's structured sentencing scheme's use, several small tweaks have been implemented, including a revision to the criminal history levels, and the formalization of aggravating and mitigating factors.

Table 14 at the end of this document presents the current sentencing guidelines. [36]. Table 14 presents the current sentencing guidelines.

The main data set includes 374,066 cases[37] with a final date of disposition between 1998 and 2010. In order to ensure that our estimates are correct, we have extensively cleaned the data. The entire data cleaning procedure is described in Appendix 1.3. Of this initial pool of cases, 358,826 or 95.9 percent were resolved through plea bargain. The other 4.1% are either resolved at trial or dismissed.

In the ensuing sections we use the term "No plea" rather than "Trial" as the alternative to "Plea" since foregoing a plea offer may result in a trial or dismissal. Thus the analysis takes the perspective of an ex-ante comparison before the decision to plea is made. This is in contrast to a number of prior papers that have focused on comparing plea sentences to those from a guilty verdict at trial[38].

Table 1 presents summary statistics for the statewide data by method of case resolution. Cases that resolved with a plea have 0.36 more charges. Plea bargained cases have somewhat younger

[36] For an overview of North Carolina's sentencing, see "The North Carolina Sentencing and Policy Advisory Commission: A History of its Creation and its Development of Structured Sentencing" available at `http://www.nccourts.org/Courts/CRS/Councils/spac/Documents/commission_history_aug2011.pdf`.

[37] We use cases to mean all charges that were resolved for an individual on a single date. See Appendix 1.3 for further detail.

[38] Examples of such papers include Rhodes (1979), Brereton and Casper (1982), and Schulhofer and Nagel (1989,1992, and 1997).

defendants, a higher share of female defendants, and a lower share of black defendants. First offenders are far less likely to plea, accounting for 41% of the plea bargain cases and 68% of the rest.

Slightly more than a quarter of defendants are incarcerated overall, although the rate is 5 percentage points higher for those who agree to a plea bargain. Incarceration is a dummy variable indicated any non-zero active prison sentence beyond time served. Sentences for those who take a plea bargain are lower on average, although we will see shortly that there is substantial variation based on crime type and other characteristics. The sentencing disparity is even greater when restricting attention to non-zero sentences.

Cases resolved via plea have substantially higher prior points (a measure of criminal history). This fact, along with the higher number of charges and lower rate of first offense would seem to suggest that those who plea face more serious consequences. We examine plea rates by type of charge below to delve into this further. The distribution of attorney type also varies by plea decision, with public defenders comprising a higher share of cases resolved via plea and private attorneys the opposite. This could be due to resource constraints or other factors or selection. We explore this further later in the paper. Table 11 presents summary statistics for just the Charlotte CSA, mirroring what we see statewide.

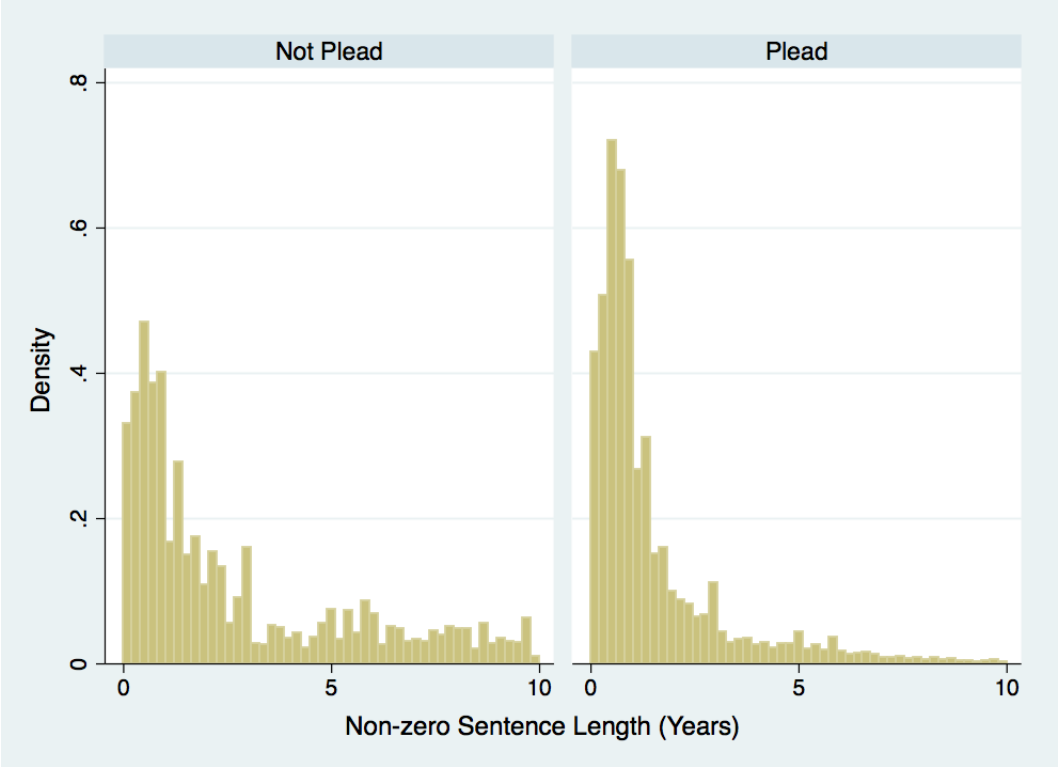
It is useful to compare the full sentencing distribution and not just look at averages. Figure **??** shows the distribution of nonzero sentences determined by our two methods of resolution and is truncated at 10 years for ease of display. Sentence length excludes time served. We immediately observe that both sentencing distributions are highly skewed, and the distribution of sentence lengths for cases settled through a plea is substantially more concentrated between 0 and 2 years, while the sentencing distribution for those cases settled without a plea is more dispersed.

Table 1 also illustrates that we have substantial variation in the method of case resolution by type of crime. We see that kidnapping cases are only plea 90% of the time, while forgery cases are plea 99%. These results in Table 1 are consistent with the idea that more serious crimes are less likely to be resolved with a plea bargain. Kidnapping is the crime with the highest unconditional sentence length of 2.20 years when plead, while forgery the lowest with 0.09 years.

Our hypothesis that $\frac{\delta Pr(trial)}{\delta J} > 0$ is further supported by Figure 3, which plots the average plea rate against the minimum sentence a defendant is facing. We can see a strong decline in the rate of pleading as the minimum sentence given conviction rises.

We see evidence that these properties may differ when we inspect incarceration rates for different types of crimes. As shown in Table 2, incarceration rates are generally higher after a plea bargain than after a trial. Some categories of crimes appear to, unconditionally on any selection issues,

Figure 2: Distribution of Non-Zero Sentences by Method of Resolution - Full State



have shorter sentences after a plea bargain is accepted while others have longer sentences.

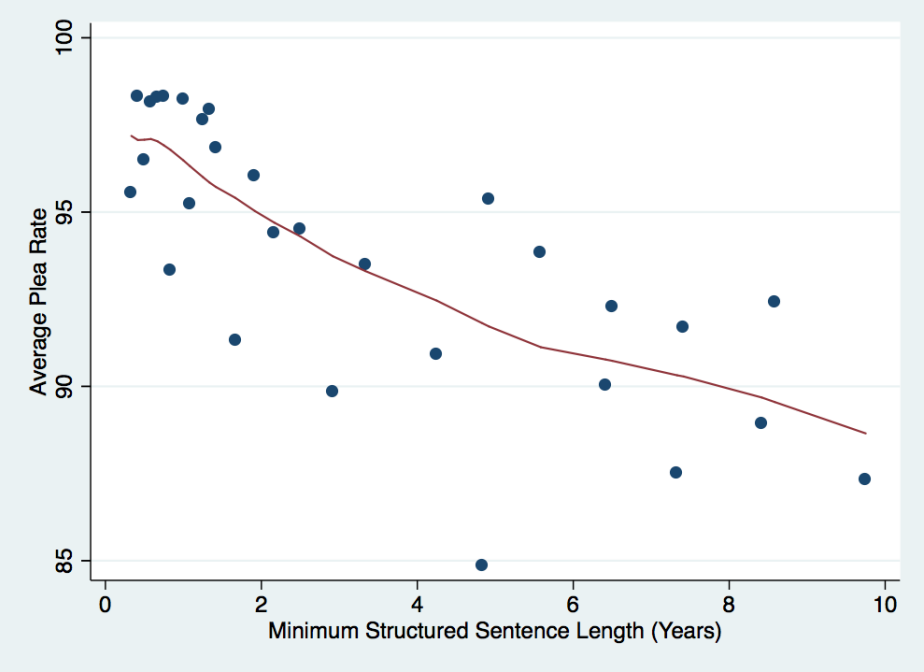
Another feature evident in Table 2 is that there is a negative relationship between the average sentence after declining a plea bargain and the rate at which defendants accept a plea. That is, the lowest rate of pleading are seen in the crimes with the highest average sentence. This is consistent with the sign of $\frac{\delta Pr(trial)}{\delta J} > 0$ discussed earlier. We also see weak evidence that $\frac{\delta Pr(guilt \mid trial)}{\delta J} > 0$

Table 2 also shows higher rates of incarceration after a plea bargain. This could be the directly product of the structured sentencing scheme used by North Carolina coupled with the fact that defendants who choose to plea have significantly more prior points (more felony convictions)[39]. This combination produces a situation where those accepting plea bargains are inherently facing longer jail sentences even conditional on committing identical crimes with identical observable and unobservable characteristics (except prior points).

Thus, Tables 1 to 2 provide strong evidence that we are going to need to control for both observable variation (e.g. a defendant's prior points) and unobservable variation (e.g. presence of eye witnesses) in cases to measure a causal relationship between method of case resolution and

[39] This fact can be seen in Table 15.

Figure 3: Plea Rate by Minimum Structured Sentence Length



expected sentence length. In order to handle observable variation, we use ordinary least squares (OLS) regression specifications, and to handle unobservable variation, we use instrumental variable (IV) regressions (discussed in Section 5).

Table 1: Summary Statistics - Full State

Variable	*Overall Mean*	*Overall SD*	*No Plea Mean*	*Plea Mean*	*t-Statistic*
Offender & Case Characteristics					
Charges	2.16	1.34	1.81	2.17	-32.7
Age	30.62	10.43	32.61	30.54	23.71
Race (black = 1)	0.54	0.5	0.56	0.54	4.11
Sex (female = 1)	0.17	0.37	0.14	0.17	-9.76
First Offense	0.43	0.49	0.68	0.41	64.17
Prior Points	3.99	4.97	2.44	4.06	-39.55
Structured Sentence Level (1=A, 10=I)	8.46	1.57	7.86	8.49	-48.27
All Charges Dismissed	0	0.03	0.02	0	95.27
Sentence (Years)	0.46	1.48	1.26	0.42	69.2
Incarceration	0.27	0.44	0.32	0.27	14.97
Non-zero Sentence	1.69	2.44	3.9	1.57	66.62
Guilty	0.98	0.16	0.53	0.99	-443.47
Lawyer Types					
Court Appointed Private Attorney	0.49	0.5	0.52	0.49	6.15
Public Defender	0.21	0.41	0.16	0.21	-16.46
Private Attorney	0.24	0.43	0.28	0.24	12.47
Waived	0.06	0.23	0.04	0.06	-7.43
Observations	374,066		15,240	358,826	

Notes: This table reports summary statistics on 374,066 observations of felony cases in across NC from 1998-2010. Homicides and some sexual assault are excluded due to idiosyncratic nature of these offenses. For other cuts, see description in the text.

Table 2: Case Outcomes by Offense and Method of Disposition - Full State

				No Plea			*Plea*	
Offense Characteristics	**Share of Total Cases**	**Plea Rate**	**Pct. Black**	*Sentence*	*Incarceration*	*Finding of Guilt*	*Sentence*	*Incarceration*
Kidnapping	1 %	90 %	55 %	4.58	54.91 %	1 %	1.94	52.52 %
Robbery	7 %	94 %	72 %	4.02	59.65 %	1 %	1.63	54.91 %
Sexual Assault	3 %	93 %	39 %	1.81	33.88 %	0 %	0.89	37.98 %
Assault	5 %	91 %	57 %	1.38	39.28 %	1 %	0.68	35.19 %
Arson	1 %	95 %	36 %	1.86	34.21 %	1 %	0.51	29.88 %
Weapons	3 %	95 %	69 %	0.75	39.67 %	1 %	0.4	33.1 %
Dangerous Drugs	34 %	96 %	64 %	0.87	26.76 %	0 %	0.32	22.93 %
Burglary	16 %	97 %	42 %	0.7	34.4 %	1 %	0.31	30.28 %
Other	6 %	93 %	46 %	0.4	17.39 %	0 %	0.22	16.92 %
Traffic Offenses	2 %	97 %	59 %	0.27	19.76 %	1 %	0.2	26.81 %
Flee Arrest	1 %	96 %	50 %	0.44	41.86 %	1 %	0.19	27.46 %
Fraud	8 %	98 %	42 %	0.25	16.48 %	1 %	0.17	17.49 %
Larceny	8 %	97 %	44 %	0.28	21.73 %	1 %	0.15	21.58 %
Embezzlement	2 %	97 %	39 %	0.28	11.49 %	0 %	0.14	7.61 %
Forgery	3 %	99 %	44 %	0.09	12.34 %	0 %	0.09	13.79 %
Total	374066			15240			358826	

Notes: This table reports summary statistics on 374,066 observations of felony cases in across NC from 1998-2010. Homicides and some sexual assault are excluded due to idiosyncratic nature of these offenses. For other cuts, see description in the text.

5 Results

In the previous Section, we discuss the myriad of observable and unobservable ways where cases settled via plea bargain differs substantially from those resolved by other means. As a first step for controlling for these differences, we begin with an OLS specification to control for observable characteristics. Specifically, we estimate:

$$Sentence_{i,j} = \alpha + \beta_1 \cdot plea_i + \sum \gamma X_{i,j} + \varepsilon_{ij} \tag{3}$$

where $X_{i,j}$ includes: 1) $race_i$, $gender_i$, age_i, $charges_i$, and $first\ offense_i$

2) $Crime\ Class_i$ (See Table 2)

3) $Year_i$

4) $Defense\ Lawyer\ Type_i$

5) $Prior\ Points\ Level_i$ (See Table 14)

6) $Crime\ Level_i$ (See Table 14)

7) $Judge_j$

Where $Sentence_{i,j}$ is the non-probation sentence in years, $plea_i$ is a indicator for whether the case was resolved via a plea bargain. The subscript i indexes the incident, and the subscript j the superior court judge. The term $\sum \gamma X_{i,j}$ is the set of case, defendant, and judge controls that differ depending on the specific regression specification.

The results of estimating Equation 3 are presented in Table 3. Specifications (1) includes controls for observed defendant and case characteristics including: race, sex, age, number of charges, and a flag for whether it is the defendant's first offense. Additionally, we include district, year, and judge, and crime type fixed effects. Specification (2) introduces controls for the type of lawyer representing the defendant. Specification (3) includes two variable for the structured sentencing cell the defendant faces: one variable representing the minimum specified sentence given conviction and the other a punishment style indicator[40].

Across all specifications we see a consistently negative and significant effect for pleading guilty on expected sentence length, with a coefficient around -0.7 years and a standard error near 0.01. This result suggests that on average, defendants who plead guilty can expect an 8-month shorter sentence than those who do not. The decrease in our estimated coefficient from specification (1) to (3) matches intuition; as shown in Tables 1 and 15, defendants who take plea bargains tend to have more serious cases (based on observable features such as the number of charges, the class of

[40] As described in Table 14, some crimes have active, inactive or community punishments. The variable for punishment style indicates whether we have: 1) only active punishment available, 2) both active and inactive punishment available, or 3) only inactive punishment available.

the charge, and the level of prior points). Therefore, by controlling partially for measures of case severity should remove some of the impact of this selection, and make the choice to plea seem more appealing.

Table 3: Effect of Pleading on Sentence Length and Incarceration (OLS)

	Sentence Length			Incarceration		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Plea	-0.75 ***	-0.74 ***	-0.67 ***	-7.20 ***	-6.98 ***	-6.24 ***
	(0.01)	(0.01)	(0.01)	(0.35)	(0.35)	(0.34)
Black	-0.01	-0.02 ***	-0.07 ***	1.91 ***	0.98 ***	-0.8 ***
	(0.01)	(0.01)	(0.00)	(0.15)	(0.15)	(0.14)
Female	-0.15 ***	-0.16 ***	-0.13 ***	-8.35 ***	-9.04 ***	-7.7 ***
	(0.01)	(0.01)	(0.01)	(0.19)	(0.19)	(0.19)
Age	0.00 ***	0.00 ***	0.00 ***	0.06 ***	0.08 ***	0.03 ***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Number of Charges	-0.04 ***	-0.04 ***	-0.03 ***	-3.70 ***	-3.41 ***	-3.52 ***
	(0.00)	(0.00)	(0.00)	(0.05)	(0.05)	(0.05)
First Offense	-0.28 ***	-0.27 ***	-0.16 ***	-22.84 ***	-21.92 ***	-18.81 ***
	(0.00)	(0.00)	(0.00)	(0.15)	(0.14)	(0.14)
Minimum Structured Sentence Length (Months)			0.03 ***			0.07
			(0.00)			(0.01)
Appointed		0.34 ***	0.25 ***		20.03 ***	17.41 ***
		(0.01)	(0.01)		(0.31)	(0.30)
Public Defender		0.12 ***	0.08 ***		12.28 ***	10.87 ***
		(0.01)	(0.01)		(0.34)	(0.33)
Waived Defense		0.17 ***	0.12 ***		8.17 ***	6.76 ***
		(0.01)	(0.01)		(0.32)	(0.31)
Observations	366752	366446	365,229	366,752	366,446	365,229
Adjusted R^2	0.12	0.13	0.23	0.16	0.18	0.24
Mean Dependent	0.46	0.46	0.46	27.17	27.18	27.24

Notes: (***) refers to significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.
Dependent variable in sentence length regressions is the minimum length of active sentence in excess of time served in years.
Dependent variable in incarceration regressions is a binary variable indicating whether an active sentence in excess of time served as received (yes = 100 and no = 0).

Control variables not displayed explicitly include: 1) district, 2) year fixed effects, 3) judge fixed effects, 4) crime type fixed effects, and 5) structured sentence punishment style fixed effects by sentence style. Privately retained lawyers is the reference category for the lawyer type fixed effects.

Table 4a: Effect of Pleading on Sentence Length By Type of Crime (OLS) - Full State

Crime Type	**Coefficient**	**SE**	**p-value**	**Observations**	**Adjusted R^2**
Kidnapping	-2.31	0.2	0	4167	0.24
Robbery	-1.95	0.07	0	26465	0.24
Sexual Assault	-0.59	0.1	0	11458	0.1
Assault	-0.75	0.04	0	18888	0.21
Arson	-1.53	0.19	0	2108	0.14
Weapons	-0.42	0.03	0	12846	0.21
Dangerous Drugs	-0.44	0.02	0	123943	0.25
Burglary	-0.36	0.02	0	57406	0.16
Other	-0.3	0.03	0	19595	0.15
Traffic Offenses	-0.09	0.04	0.03	6453	0.29
Flee Arrest	-0.19	0.04	0	5279	0.17
Faud	-0.14	0.02	0	30621	0.15
Larceny	-0.15	0.02	0	29655	0.15
Embezzlement	-0.13	0.06	0.03	5564	0.23
Forgery	-0.03	0.03	0.44	10781	0.12
Total	-0.67	0.01	0	365229	0.23

Notes: (***) refers to significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.
Dependent variable in all regressions is the minimum length of active sentence in excess of time served in years.
The form of all regressions matches regression (3) from Table 5, but only the coefficient on pleading is reported.
Offenses are ordered by expected sentence length (see Table 2).

Table (3) also allows us to examine several of our hypotheses presented in Section (4). First, our prediction that female defendants should accept lower average sentences seems robust to the sentencing scheme. Second, our hypothesis that defendants represented by public defenders may accept longer sentences is supported; defendants with public defenders accept statistically longer sentences than private attorneys, but court appointed attorneys accept the longest sentences, even controlling for different defendant backgrounds. In general, we see broadly similar results when we strict our attention to just the Charlotte CSA.

Table 3 also shows OLS estimates with incarceration and not sentence length as the dependent variable. In specifications (1) and (2), we see a 7% lower incarceration rate after pleading. In specifications (3), we see that controlling for North Carolina's structured sentencing scheme lowers the discount on incarceration slightly to 6%. In Charlotte (Table 13), we initially see a 8% lower incarceration rate after pleading, which reduces to 7% when controlling for criminal record.

This result should be unsurprising given North Carolina's structured sentencing scheme. We ex-

Table 4b: Effect of Pleading on Incarceration By Type of Crime (OLS) - Full State

Crime Type	**Coefficient**	**SE**	**p-value**	**Observations**	**Adjusted R^2**
Kidnapping	3.56	2.28	0.12	4167	0.28
Robbery	3.33	1.17	0	26465	0.24
Sexual Assault	5.49	1.69	0	11458	0.21
Assault	-7.95	1.12	0	18888	0.24
Arson	-9.91	4.26	0.02	2108	0.23
Weapons	-13.07	1.77	0	12846	0.2
Dangerous Drugs	-6.42	0.58	0	123943	0.22
Burglary	-9.2	1.06	0	57406	0.24
Other	-7.41	1.08	0	19595	0.2
Traffic Offenses	-3.81	3.15	0.23	6453	0.25
Flee Arrest	-14.17	2.93	0	5279	0.21
Fraud	-5.47	1.35	0	30621	0.19
Larceny	-6.92	1.29	0	29655	0.21
Embezzlement	-3.95	2.03	0.05	5564	0.14
Forgery	-3.49	2.49	0.16	10781	0.27
Total	-6.24	0.34	0	365229	0.24

Notes: (***) refers to significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.
Dependent variable in all regressions is a binary variable indicating whether an active sentence in excess of time served as received (yes = 100 and no = 0).
The form of all regressions matches regression (3) from Table 5, but only the coefficient on pleading is reported.
Offenses are ordered by expected sentence length (see Table 2).

pect defendants with longer criminal histories to be more likely to face jail time, so controlling for criminal history should account for much of the unconditional difference. In specifications (1) and (2), the flag for first offense serves as a proxy for the amount of prior points[41], and in (3), we fully control for the lead charge and structured sentencing cell. In fact, without either the first-offense flag or the full sentencing controls, no significant difference in incarceration rates manifests. Thus, this is suggestive that much of the difference in incarceration rates evident in Tables 1 and 11 is attributable to observable case characteristics and North Carolina's structured sentencing scheme.

Tables 4a and 4b summarize our OLS analysis done separately for different types of crime. These analyses, which use our full panel of controls (i.e. specification 3), suggest substantial heterogeneity. We see large and statistically significant trial penalties for robbery, arson, and kidnapping, with the smallest for fraud, embezzlement, and forgery. While this heterogeneity is predictable without any controls as different crimes have wildly disparate sentencing guidelines, it is impor-

[41] The first offense variable is functionally an indicator variable that equals 1 if the defendant has 0 or 1 prior points.

Table 5: Effect of Pleading on Sentence Length By Race (OLS)

	Black			White		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Plea	-0.96 ***	-0.96 ***	-0.85 ***	-0.37 ***	-0.37 ***	-0.35 ***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Female	-0.16 ***	-0.16 ***	-0.13 ***	-0.10 ***	-0.12 ***	-0.1 ***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age	0.00 ***	0.00 ***	0.00 **	0.00 ***	0.01 ***	0 ***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0)
Number of Charges	-0.06 ***	-0.06 ***	-0.05 ***	-0.02 ***	-0.02 ***	-0.01 ***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0)
First Offense	-0.38 ***	-0.37 ***	-0.24 ***	-0.24 ***	-0.22 ***	-0.15 ***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Minimum Structured Sentence Length (Months)			0.03 ***			0.03 ***
			(0.00)			(0)
Appointed		0.33 ***	0.24 ***		0.28 ***	0.21 ***
		(0.01)	(0.01)		(0.01)	(0.01)
Public Defender		0.10 ***	0.06 ***		0.13 ***	0.10 ***
		(0.02)	(0.02)		(0.02)	(0.02)
Waived Defense		0.20 ***	0.14 ***		0.1 ***	0.06 ***
		(0.02)	(0.01)		(0.01)	(0.01)
Observations	199060	198903	198,352	144,762	144,629	144,051
Adjusted R^2	0.16	0.16	0.26	0.09	0.10	0.20
Mean Dependent	0.51	0.51	0.51	0.35	0.35	0.35

Notes: (***) refers to significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.
Dependent variable in all regressions is the minimum length of active sentence in excess of time served in years.
Control variables not displayed explicitly include: 1) district, 2) year fixed effects, 3) judge fixed effects, 4) crime type fixed effects, and 5) structured sentence punishment style fixed effects by sentence style. Privately retained lawyers is the reference category for the lawyer type fixed effects.

tant to understand that this heterogeneity persist even with our controls.

Similarly Tables 5 and 6 present our OLS analysis separate for black and white defendants. Table 5 presents evidence that black defendants face between a 5 to 6-month longer trial penalty than similar white defendants. Even controlling for the structured sentencing scheme, black defendants who plea receive 10-month shorter sentences than those who go to trial, while white defendants only 4-months.

Table 6 shows that with full controls (specification 3), black defendants who plead guilty are in-

Table 6: Effect of Pleading on Incarceration By Race (OLS)

	Black			White		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Plea	-7.57 ***	-7.21 ***	-5.81 ***	-7.22 ***	-7.18 ***	-7.49 ***
	(0.47)	(0.47)	(0.46)	(0.53)	(0.53)	(0.52)
Female	-9.43 ***	-9.66 ***	-8.04 ***	-5.46 ***	-6.65 ***	-5.92 ***
	(0.29)	(0.29)	(0.28)	(0.26)	(0.26)	(0.25)
Age	0.03 ***	0.04 ***	-0.02 *	0.11 ***	0.13 ***	0.09 ***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Number of Charges	-5.39 ***	-4.95 ***	-5.04 ***	-1.99 ***	-1.84 ***	-1.93 ***
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.07)
First Offense	-25.97 ***	-25.23 ***	-21.50 ***	-21.49 ***	-20.26 ***	-17.87 ***
	(0.20)	(0.20)	(0.20)	(0.22)	(0.22)	(0.21)
Minimum Structured Sentence Length (Months)			0.09 ***			0
			(0.01)			(0.01)
Appointed		20.91 ***	18.09 ***		16.73 ***	14.87 ***
		(0.42)	(0.41)		(0.45)	(0.44)
Public Defender		11.75 ***	10.28 ***		12.11 ***	11.02 ***
		(0.47)	(0.45)		(0.53)	(0.52)
Waived Defense		9.15 ***	7.51 ***		5.2 ***	4.37 ***
		(0.45)	(0.43)		(0.46)	(0.45)
Observations	199060	198903	198,352	144,762	144,629	144,051
Adjusted R^2	0.18	0.20	0.26	0.14	0.16	0.21
Mean Dependent	29.92	29.93	29.98	22.49	22.51	22.55

Notes: (***) refers to significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.
Dependent variable in all regressions is a binary variable indicating whether an active sentence in excess of time served as received (yes = 100 and no = 0).
Control variables not displayed explicitly include: 1) district, 2) year fixed effects, 3) judge fixed effects, 4) crime type fixed effects, and 5) structured sentence punishment style fixed effects by sentence style. Privately retained lawyers is the reference category for the lawyer type fixed effects.

carcerated 5.8% less than those who go to trial. White defendants who plea, however, are incarcerated 7.5% less than those who go to trial. This over 1.5% lower incarceration discount for white defendants is especially surprising given the base rates of incarceration. White defendants are on average incarcerated 23% of the time while black defendants 30%. In fact, unconditional on any controls, white defendants who plea are incarcerated 1.6% less than those who do not plea, while black defendants who plea are incarcerated 6.0% less.

The results of these two tables suggests that substantial racial differences may exist in the North Carolina court system. Controlling for observable characteristics, we see a greater disparity in sentences between black defendants who plea and go to trial than white defendants. Simultaneously, we see white defendants receiving 1.5% lower incarceration rate after pleading than black defendants, while simultaneously being incarcerated overall at a substantially lower rate.

In addition to the above results, we see substantial regional variation across the eight divisions in the North Carolina court system. As shown in Figure 4(a), the eight divisions exhibit significant heterogeneity in the discounts received from pleading. We see the urban center of the state has larger discounts from pleading, as compared to the rural sides. Defendants in the third division face a 10.7 month shorter sentence from pleading instead of trial, while a similar defendant in the first and eighth divisions face only a 5.5 month shorter sentence if they plea.

We see similar variation in discounts for incarceration in Figure 4(b). Again, defendants in the central urban mass of North Carolina face incarceration rate discounts that are substantially larger than their rural peers. Defendants in the sixth division expect a 15% lower chance of incarceration after pleading rather than trial, while defendants in the eighth division see no significant incarceration rate reduction from pleading[42] .

At first glance, we might be surprised that the heterogeneity we see in Figure 4 is not mirrored in Figure 5, but there is no reason to assume it should. Figure 4 shows an urban-rural divide with urban areas providing larger discounts to defendants who plea. These discounts might reflect differences in cost of trial, with overburdened urban areas encouraging more pleading. These greater discounts in urban areas, however, do not speak to how black and white defendants are treated differently in the same division.

Figures 5(a) and 5(b) show exactly this, how black and white defendants are treated differently in the same division. In both Figures, we do not see the same uniform urban-rural divide. Figure 5(b) shows that black defendants in the urban third district receive larger incarceration discounts from pleading relative to white peers, while in the neighboring urban fifth division, they receive much smaller discounts relative to white peers.

Interestingly, 5(a) and 5(b) do not show common treatments of black and white defendants in the same divisions. For example, Figure 5(a) shows that black defendants in the fifth division receive a sentence discount over 6-months larger than their white peers who plea, while Figure 5(b) shows that white defendants in the fifth division receive an incarceration rate discount over 1.5

[42] From the sample just from division eight and using equation (3) in Table 3, we estimate an incarceration rate discount of -1.2% with a standard deviation of 1.27.

percentage points larger than their black peers.

These regional differences do not appear related to obvious regional demographic differences. The black population in North Carolina is primarily concentrated in Charlotte (sixth and seventh divisions), Greensboro (fifth division), Raleigh-Durham (third division), and Fayetteville (fourth division)[43], but we can immediately see that these divisions run the gamut of treatment of black and white defendants. Similarly, black individuals represent a larger share of the population in the eastern half of North Carolina, which again does not exhibit any systematic results in Figure 5[44]. Lastly, these regional differences do not seem related to the share of black defendants in the division's courts, with nearly identical representation in the first and third divisions, but opposite results. [45]

This regional variation is surprising given the lengths that North Carolina goes to homogenize their court system. The North Carolina Administrative Office of the Courts specifically states:

> The constitution requires superior court judges to rotate, or 'ride the circuit,' from one district to another in their division. Judges are assigned to a judicial district for a six-month period and then rotated to another district for the same time period. ... On the other hand, the rotation system helps avoid any favoritism that might result from always having a judge hold court where he or she lives, has close friends among the lawyers and might be more personally familiar with and interested in the particular cases tried. It also contributes to uniformity of procedure. The frequent changes of judges tend to discourage the development of local rules that are unique to that area.[46]

Additionally, a fundamental principal of North Carolina's structured sentencing system is that "[s]entencing policies should be consistent: Offenders convicted of similar offenses, who have similar prior records, should generally receive similar sentences."[47] Admittedly the judicial rotation is within a division, and not specifically designed to create homogeneity across divisions. That said, the magnitude of the differences are surprising given the tremendous effort the court system exerts to avoid regional differences, and seem to indicate a failure to obtain 'consistency' in sentencing.

[43] 2010 Census population estimates can be found: `https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk`.

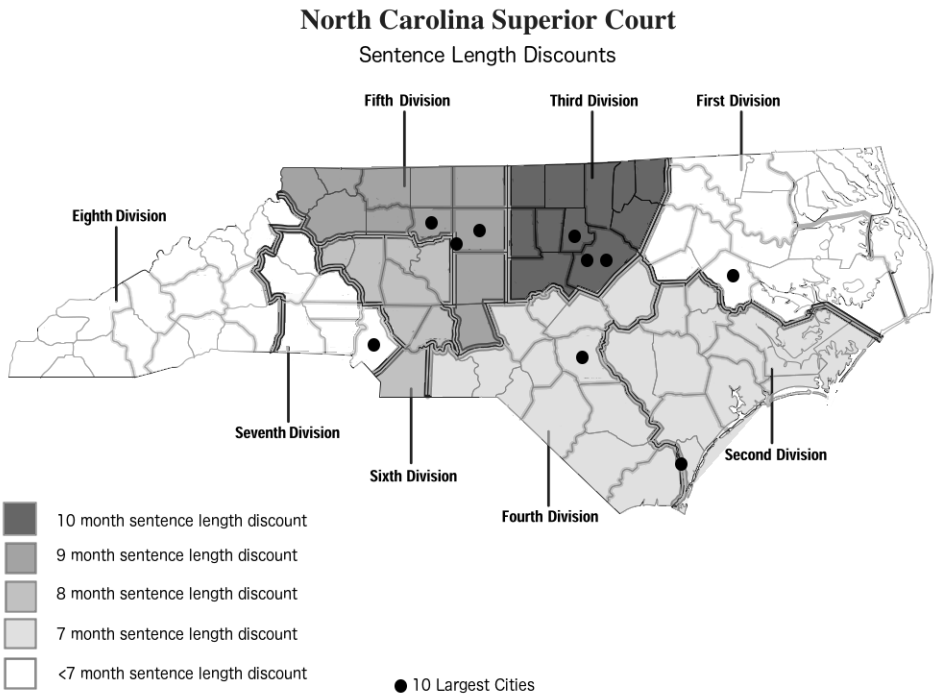
[44] For a good visual representation, see `http://www.indexmundi.com/facts/united-states/quick-facts/north-carolina/black-population-percentage#map`.

[45] Defendants are 65.4% black in division 1, 57.4% black in division 2, 64.5% black in division 3, 52.6% black in division 4, 54.1% black in division 5, 41.2% black in division 6, 53.1% black in division 7, and 23.6% black in division 8.

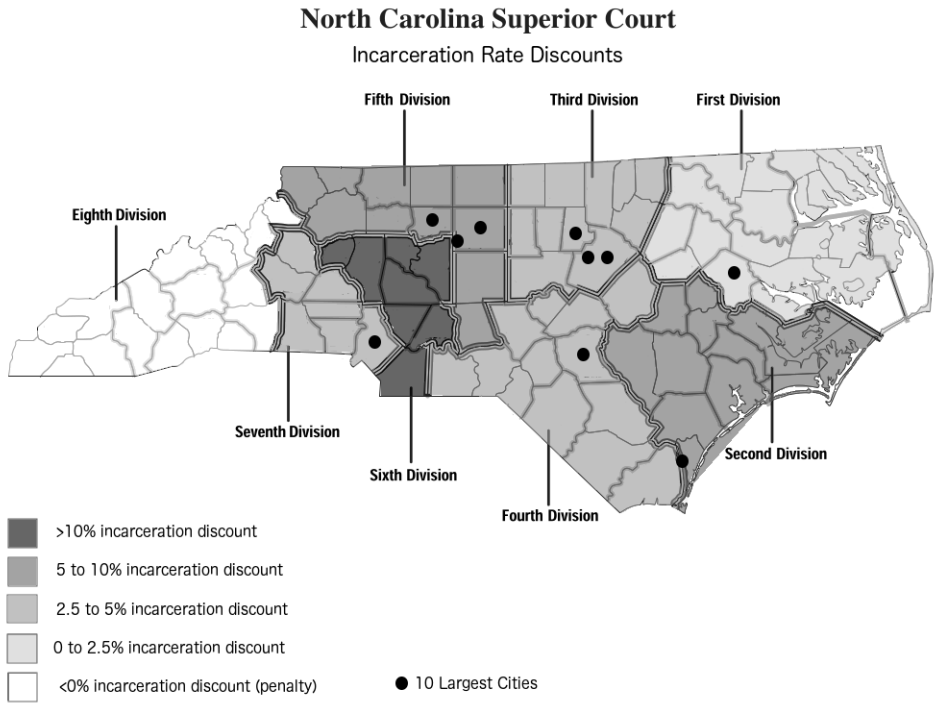
[46] The Judicial System in North Carolina (2007) pg. 5 `http://www.nccourts.org/news/documents/judicialsystem.pdf`

[47] A Citizen's Guide to Structured Sentencing (2014) pg. 1 `http://www.nccourts.org/Courts/CRS/Councils/spac/Documents/citizenguide2014.pdf`

Figure 4: Regional Variation in Sentencing and Incarceration Discounts

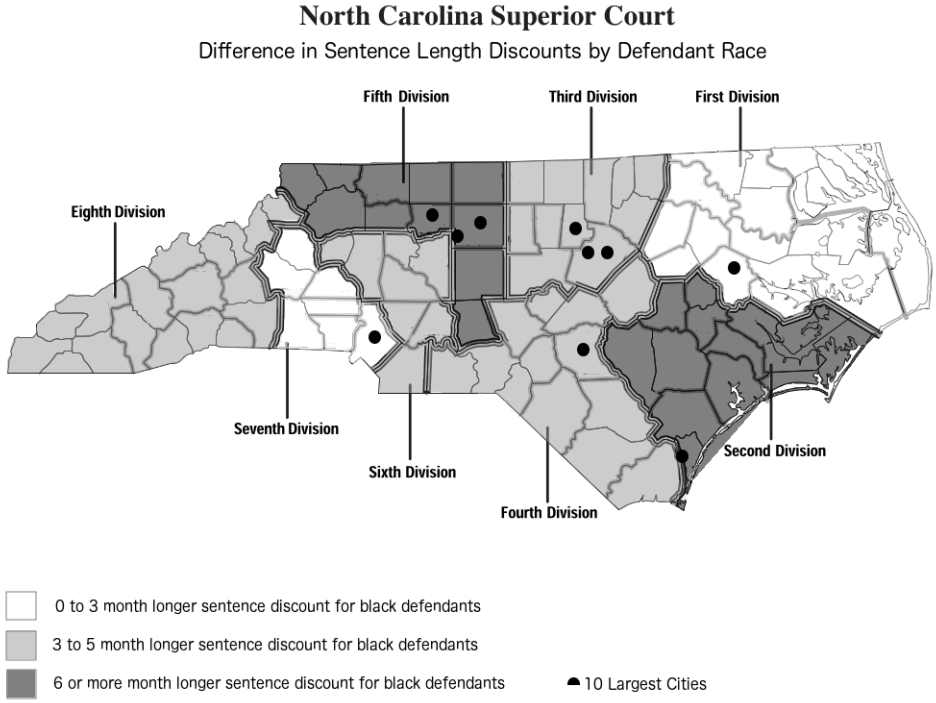


(a) Sentence Discounts

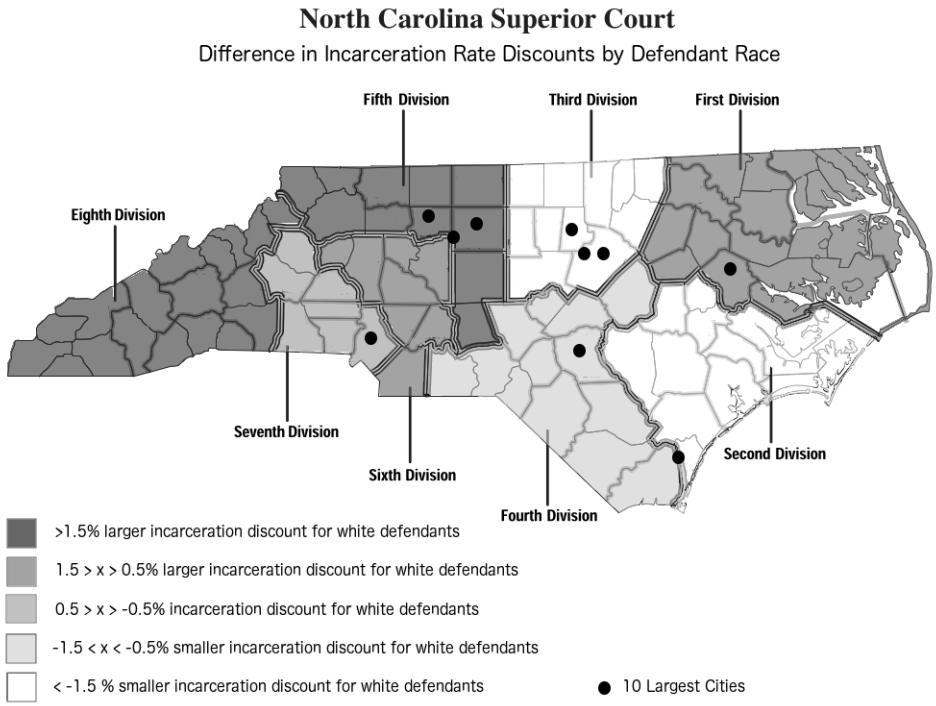


(b) Incarceration Discounts

Figure 5: Regional Variation in Sentencing and Incarceration Discounts by Race



(a) Sentence Discounts by Race



(b) Incarceration Discounts by Race

5.1 Addressing Unobservable Variation

Up to this point, we have explored only those observable differences between cases resulting with and without a plea. As originally described Section 3.1, we have reason to believe that there are important selection effects occurring based on unobservable heterogeneity. This selection on unobservables may bias the coefficient on plea and prohibit causal interpretation of our OLS results.

As described in Table 15, we have reason to think potentially worse cases go to trial, while defendants with longer criminal histories are more likely to plea[48] We could easily expect to see similar trends with unobservables. District attorneys may be less inclined to plea unobservably worse cases in order to not seem "light on crime." Were such unobserved selection occuring, our estimate of the trial penalty would be inflated (i.e. our coefficient on plea would be overly negative). On the other hand, we could also construct a story where defendants with unobservably worse cases may plead guilty knowing that they have an idiosyncratically higher chance of losing at trial. If this form of selection were occurring, we would expect a lower trial penalty (i.e our coefficient on plea should be more negative).

Regardless of the direction that bias manifests, we have sufficient fear of unobservable selection to motivate the use of instrumental variable regression. To do so we use the number of judge and defense attorney interactions as our instrument[49]. The IV specification is as follows:

$$\text{First Stage: } plea_i = \tilde{\alpha} + \tilde{\beta}_1 \cdot number_cases_{i,j} + \sum \tilde{\gamma} X_{i,j} + \nu_{ij} \tag{4}$$

$$\text{Second Stage: } Sentence_{i,j} = \alpha + \beta_1 \cdot \widehat{plea_i} + \sum \gamma X_{i,j} + \varepsilon_{ij} \tag{5}$$

Where $\widehat{plea_i}$ is the estimated value of $plea_i$ that comes from first first stage regression, $number_cases_{i,j}$ is the instrument(s) used in the specific regression[50], and $X_{i,j}$ is the same as in equation 3.

As we discuss at the end of Section 3.1, we require an instrument that is related to the propensity for a defendant to plead guilty but unrelated to the idiosyncratic probability of guilt at trial. As such, a variable such as the defendant's prior points would not work as an instrument as it is suggestive of a criminal propensity that may be related to the probability of guilt at trial.

We propose an instrument that relates to the knowledge of defense attorneys about judge pro-

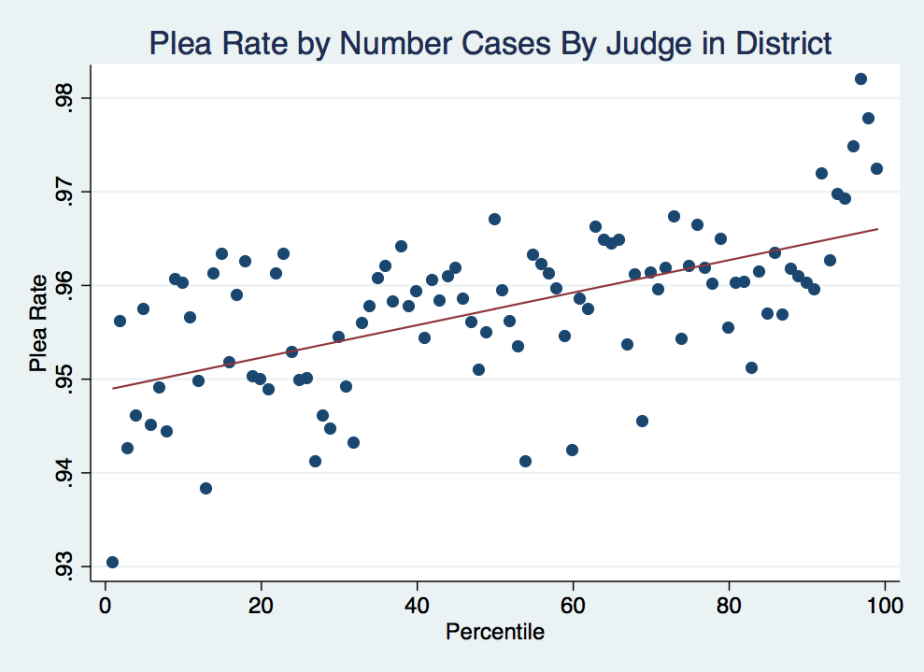
[48] For example, 36% of trials involve a lead charge of a class above H whereas only 32% of plea bargains are above class H. Based on the observables, it appears that more severe cases are not resolved with a plea bargain. Cases resolving with a plea involve a defendant with a prior points level of 2 or higher 67% of the time, as compared to 15% in cases that do not plea.

[49] This is not the only possible instrument. Abrams (2013) used judicial tenure as an instrument. We believe that the number of defense attorney-judge interactions more accurately captures the knowledge of the attorney.

[50] Our estimates are obtained using two stage least squares regression (2SLS), but are qualitatively similar if we use generalized method of moments (GMM) to estimate.

clivities. Defense attorneys primarily learn about judge behavior through direct experience. Accordingly, we look at the prior number of cases and trials the specific judge has seen in the district. To create this measurement, we restrict our attention to cases where: 1) the judge saw no cases before 1999 anywhere in North Carolina, or 2) the judge saw no cases in the specific district before 2001. We make this restriction to ensure that we are accurately counting the number of cases and trials the judge supervises in the district. If we did not make these cuts, we would only be measuring the number of interactions observed in our data and not that the judge has *ever* had.

Figure 6: Instrument: Percentile of Number of Cases for Judge in District



Additionally, because curvature is evident in the relationship between our instrument and the propensity to plea, we use the percentile of the number of cases/trials rather than the raw count. We additionally include the squared number of cases and trials. As such, equation (2) in Tables 7 to 10 has a first stage with squared versions of the instruments. This alternative specification takes the form:

$$\text{Second Stage: } Sentence_{i,j} = \alpha + \beta_1 \cdot \widehat{plea_i} + \sum \gamma X_{i,j} + \varepsilon_{ij} \tag{4}$$

$$\text{First Stage: } plea_i = \tilde{\alpha} + \tilde{\beta}_1 \cdot number_cases_{i,j} + \beta_2 \cdot number_cases^2_{i,j} + \sum \tilde{\gamma} X_{i,j} + \nu_{ij} \tag{5a}$$

Lastly, it is important for our identification strategy, cases are randomly assigned to judges. Using a similar methodology to that employed in Abrams et al. (2012), we simulate the distribution of observable case characteristics under random assignment. [51]. We then compare the empirical interjudge dispersion for observable case characteristics with the simulations. Appendix section

[51] See Abrams, Bertrand, and Mullainathan *Do Judges Vary in Their Treatment of Race?* (2012).

1.5 reports results for charge severity, defendant race, sex, age, and criminal history. The results are consistent with randomly assigned cases, which is consistent with multiple conversations with officials in the North Carolina court system.

Table 7: Effect of Pleading on Sentence Length - IV - Full and Charlotte

	Full Sample		**Charlotte**	
	(1)	(2)	(1)	(2)
First Stage				
Cases	0.001 ***	0.000 ***	0.001 ***	0.001 ***
	(0.000)	(0.000)	(0.000)	(0.000)
$Cases^2$		0.000 ***		0.000
		(0.000)		(0.000)
Trials	-0.001 ***	-0.001 ***	-0.002 ***	-0.003 ***
	(0.000)	(0.000)	(0.000)	(0.000)
$Trials^2$		0.000		0.000 ***
		(0.000)		(0.000)
F-Stat	34.091	33.854	16.267	16.065
Adjusted R^2	0.037	0.037	0.042	0.042
Second Stage				
Plea	-1.153 ***	-0.825 **	-0.559	-0.657
	(0.333	(0.324)	(0.424)	(0.414)
Adjusted R^2	0.214	0.217	0.274	0.275
Observations	162000	162000	39500	39500
Mean Dependent	0.444	0.444	0.42	0.42
OLS Coefficient	-0.669	-0.669	-0.678	-0.678

Notes: (***) refers to significance at the 1% level and (**) at the 5% level, and (*) at the 10% level.

Dependent variable in second stage regressions and in the OLS regressions is the minimum length of active sentence in excess of time served in years.

Control variables include: 1) demographic controls from Table 5, 2) district fixed effects, 3) year fixed effects, 4) judge fixed effects, 5) crime type fixed effects, and 6) structured sentence punishment style fixed effects by sentence style.

"Corresponding OLS" estimates report the results from regression (3) in Table 5 when run on the subset of the data in each column.

Table 8: Effect of Pleading on Sentence Length - IV - Class E-I Felonies

	Full Sample		**Class E-I Felonies**	
	(1)	(2)	(1)	(2)
First Stage				
Cases	0.001 ***	0.000 ***	0.001 ***	0.000 *
	(0.000)	(0.000)	(0.000)	(0.000)
$Cases^2$		0.000 ***		0.000 ***
		(0.000)		(0.000)
Trials	-0.001 ***	-0.001 ***	-0.001 ***	-0.001 ***
	(0.000)	(0.000)	(0.000)	(0.000)
$Trials^2$		0.000		0.000 *
		(0.000)		(0.000)
F-Stat	34.091	33.854	28.068	27.937
Adjusted R^2	0.037	0.037	0.034	0.034
Second Stage				
Plea	-1.153 ***	-0.825 **	-0.402	-0.155
	(0.333	(0.324)	(0.242)	(0.233)
Adjusted R^2	0.214	0.217	0.135	0.135
Observations	162000	162000	148000	148000
Mean Dependent	0.444	0.444	0.272	0.272
OLS Coefficient	-0.669	-0.669	-0.277	-0.277

Notes: (***) refers to significance at the 1% level and (**) at the 5% level, and (*) at the 10% level.

Dependent variable in second stage regressions and in the OLS regressions is the minimum length of active sentence in excess of time served in years.

Control variables include: 1) demographic controls from Table 5, 2) district fixed effects, 3) year fixed effects, 4) judge fixed effects, 5) crime type fixed effects, and 6) structured sentence punishment style fixed effects by sentence style.

"Corresponding OLS" estimates report the results from regression (3) in Table 5 when run on the subset of the data in each column.

Table 9: Effect of Pleading on Sentence Length - IV - Race

	White		Black	
	(1)	(2)	(1)	(2)
First Stage				
Cases	0.001 ***	0.000 **	0.001 ***	0.000 **
	(0.000)	(0.000)	(0.000)	(0.000)
$Cases^2$		0.000 *		0.000 ***
		(0.000)		(0.000)
Trials	-0.001 ***	-0.001 ***	-0.001 ***	-0.001 ***
	(0.000)	(0.000)	(0.000)	(0.000)
$Trials^2$		0.000		0.000
		(0.000)		(0.000)
F-Stat	15.923	15.775	20.176	20.044
Adjusted R^2	0.04	0.04	0.041	0.041
Second Stage				
Plea	0.032	0.132	-2.023 ***	-1.481 ***
	(0.48	(0.475)	(0.482)	(0.462)
Adjusted R^2	0.184	0.182	0.221	0.235
Observations	67442	67442	84340	84340
Mean Dependent	0.34	0.34	0.493	0.493
OLS Coefficient	-0.343	-0.343	-0.854	-0.854

Notes: (***) refers to significance at the 1% level and (**) at the 5% level, and (*) at the 10% level.

Dependent variable in second stage regressions and in the OLS regressions is the minimum length of active sentence in excess of time served in years.

Control variables include: 1) demographic controls from Table 5, 2) district fixed effects, 3) year fixed effects, 4) judge fixed effects, 5) crime type fixed effects, and 6) structured sentence punishment style fixed effects by sentence style.

"Corresponding OLS" estimates report the results from regression (3) in Table 5 when run on the subset of the data in each column.

Table 10: Effect of Pleading on Sentence Length - IV - First Time Vs. Repeat Offenders

	Repeat Offender		**First Time Offender**	
	(1)	(2)	(1)	(2)
First Stage				
Cases	0.001 ***	0.000 ***	0.001 ***	0.000 *
	(0.000)	(0.000)	(0.000)	(0.000)
$Cases^2$		0.000 **		0.000 ***
		(0.000)		(0.000)
Trials	-0.001 ***	-0.001 ***	-0.002 ***	-0.002 ***
	(0.000)	(0.000)	(0.000)	(0.000)
$Trials^2$		0.000		0.000
		(0.000)		(0.000)
F-Stat	12.647	12.546	17.723	17.613
Adjusted R^2	0.023	0.023	0.045	0.045
Second Stage				
Plea	-3.187 ***	-2.857 ***	0.479	0.7 *
	(0.59	(0.579)	(0.373)	(0.363)
Adjusted R^2	0.26	0.269	0.116	0.105
Observations	94790	94790	67282	67282
Mean Dependent	0.556	0.556	0.287	0.287
OLS Coefficient	-1.698	-1.698	-0.139	-0.139

Notes: (***) refers to significance at the 1% level and (**) at the 5% level, and (*) at the 10% level.

Dependent variable in second stage regressions and in the OLS regressions is the minimum length of active sentence in excess of time served in years.

Control variables include: 1) demographic controls from Table 5, 2) district fixed effects, 3) year fixed effects, 4) judge fixed effects, 5) crime type fixed effects, and 6) structured sentence punishment style fixed effects by sentence style.

"Corresponding OLS" estimates report the results from regression (3) in Table 5 when run on the subset of the data in each column.

As shown in Tables 7 to 10, our IV regressions suggest a significant difference between sentences received as a result of a plea bargain from those received after choosing not to plea. Our point estimates in Table 7 of 10-months for the full state and 8-months for Charlotte, are both slightly larger than their corresponding OLS estimates, suggesting unobservable selection. The larger point estimate point to unobservably better cases being plead (or worse cases going to trial), thereby expanding the differences.

The point estimates in Table 9 suggest that private information may play different roles for black and white defendants. The point estimate for the discount black defendants receive for pleading are over twice as large as the corresponding OLS estimates (17.7 months vs. 10.3 months), but the point estimates for the white defendants fall slightly and are indistinguishable from no discount. While not conclusive, this provides further evidence of racial disparities in the treatment on black and white defendants.

6 Discussion

The results presented above provide the first evidence of the existence of an *ex-ante* trial penalty. From simple summary statistics, we that the average sentence from a plea bargain is 9 months shorter than the average sentence of those who did not plea. After controlling for observable heterogeneity (OLS regressions in Table 3), this falls to 8 months. Our use of instrumental variables to control for additional unobserved heterogeneity provides evidence that the true difference in sentence lengths is even longer.

Importantly, this work sheds light into possible racial disparities in treatment by the North Carolina court system. One rationale for North Carolina's structured sentencing scheme was to provide homogeneity across similarly situated defendants. Our OLS and results suggest that this uniformity is lacking; black defendants see greater reductions in sentence lengths after pleading than white defendants, while white defendants see greater reductions in incarceration rates. Our IV results amplify the differences observed in our OLS estimates.

The results presented here are consistent with prior empirical literature measuring sentences received from pleading or trial. As shown in Table 1, we see a substantially longer sentence after a finding of guilt at trial than after a plea bargain. This corresponds to the *ex-post* trial penalty, and is robust to the inclusion of our panel of controls for observable case differences. That said, despite push-back against the analysis of the ex-ante trial penalty, we still believe it to be the proper measurement for understanding the implications of the shadow of the law theory.

The theory of The Shadow of the Law was first proposed by Mnookin and Kornhauser in their

1979 article discussing divorce settlements. The authors propose the notion that despite the vast majority of divorces resolving with a negotiated settlement, the law shapes these negotiations. As either party is always free to opt out of the negotiations and request judicial intervention, the divorce negotiations are inherently framed by this outside option. The shadow of the law has proven compelling, with many other authors employing the framework[52] It is straightforward to extend Mnookin and Kornhauser's original idea to both criminal and civil cases. This paper and most every paper concerning a trial penalty fundamentally rests on the shadow of the law argument to provide a logical connection between plea bargains and trials.

In recent years, the shadow of the law argument has come under pragmatic scrutiny. Bibas addresses many features of the criminal justice system that complicate the real-world function of the shadow of the law model[53]. Bibas also invokes well-documented Behavioral Economics results as problematic for the assumptions of the model. Stuntz's criticism focuses on the motivation and discretion of prosecutors[54]. Stuntz questions whether prosecutors' goal is to maximize sentences, and questions the functioning of the shadow of the law model when prosecutors have discretion in what cases to see.

At this point, let us consider how the shadow of the law argument translates into the trial penalty. Essentially the formal model presented in Section 3 mathematically characterizes the shadow of the law; defendant and prosecutors are able to negotiate any plea they choose[55], and their willingness to accept a given plea deal is made only in comparison to the opt-out value of a trial. Recalling the notation of Section 3, if $\underline{A} \leq \overline{B}$, then any negotiated sentence length $J' \in \left[\underline{A}, \overline{B}\right]$ from a plea bargain would be preferable to going to trial for both the prosecution and defendant. That is, prosecutors and defendants are willing to accept any plea bargain sentence length that is above the minimum required by the prosecutor, $\underline{A}$, and below the maximum permitted by the defendant, $\overline{B}$.

We must then ask: what will the outcome of the negotiated plea be? If the defendant has all power in negotiation, then the prosecutor will be pushed to accept his least desired outcome, $\underline{A}$. If, on the other hand, the prosecutor has all bargaining power, the defendant will be driven up to their least desired outcome, $\overline{B}$. Translating in terms of trial penalty, the trial penalty is largest when the defendant has all the bargaining power, and smallest when the prosecutor has all the

[52] For two early examples of its use see: Cooter, Marks, and Mnookin *Bargaining in the Shadow of the Law: A Testable Model of Strategic Behavior* (1982) and Priest and Klein *The Selection of Disputes for Litigation* (1984).

[53] Stephanos Bibas *Plea Bargaining Outside the Shadow of Trial* (2004).

[54] W. Stuntz *Plea Bargaining and Criminal Laws Disappearing Shadow* (2004)

[55] In the model only. Innumerable constraints exist on plea bargain negotiations in the real world.

power[56].

Given the enormous time cost to trial relative to a plea bargain for prosecutors and the cost to failing to obtain a conviction at trial, one may suspect that the defendant has *some* power in the negotiation. Given this, we may then ask: "what impact would structured sentencing have on the outcome of the negotiation?" Temporarily ignoring the option of allowing the defendant to plea to a lesser charge, we may think of a structured sentencing scheme as imposing bounds on the range of acceptable plea bargains. An unconstrained prosecutor may be willing to accept a sentence length of $\underline{A}$, but the structured sentencing scheme may not allow any sentence less than $\underline{J} > \underline{A}$ for the given crime. Were this the case, the imposition of the structured sentencing scheme would have the same effect as raising prosecutor power in negotiation[57].

Now, let us permit the prosecutor and defendant to agree to a plea to a lesser charge. Introducing this power increases the span of feasible bargains on the lower end of sentence lengths only. Consider a first-time defendant currently facing a charge of assault with intent to kill in North Carolina. Given the structured sentencing scheme, this defendant if convicted at trial faces 44 to 92 months in jail. If he were to plea instead to assault with intent to inflict serious injury, he faces only 15 to 31 months in jail.

Now, let us consider the case where $\underline{A} = 30$ and $\overline{B} = 60$, and let us say the defendant's bargaining power is sufficient to always force the prosecutor to agree to a sentence 20% of the way between the minimum and maximum available. Without the ability to plea to a lesser charge, this means that the defendant and prosecutor would settle on a sentence of 53.6 months[58]. With the ability to plea to a lesser charge, the minimum possible sentence length is reduced without any change to the maximum. This then means that the defendant and prosecutor will settle on a charge of only 30.4 months[59].

Without the structured sentencing guidelines, we might think that the negotiated plea would be more of a direct product of $\underline{A}$ and $\overline{B}$. For example, we might think that a similar defendant's bargaining power would be sufficient to always force the prosecutor to agree to a sentence 20% of the way between the minimum and maximum available. Without the formality of the structured sentencing scheme, this would yield a sentence from pleading of 36 months in jail[60]. Now, assuming

[56] To see this, we assume that $J * Pr(guilt\ at\ trial)$ doesn't depend on power. Then, we have $J * Pr(guilt\ at\ trial) - \overline{B} < J * Pr(guilt\ at\ trial) - \underline{A}$. That is the trial penalty is maximized with full defendant bargaining power.

[57] Theoretically, it could also be the case that the upper sentence limit is binding. That is, $\overline{J} < \overline{B}$. In this case, the structured sentencing scheme would have the effect of reducing prosecutor power.

[58] The math behind this is straight forward. The minimum sentence is 44 months and the maximum 92. The sentence 20% of the way between 44 and 92 months is 53.6 months.

[59] The underlying math: $((92 - 15) \cdot 0.2 + 15) = 30.4$.

[60] The underling math: $((60 - 30) \cdot 0.2 + 30) = 36$.

the expected sentence from trial is unchanged, at $E[J]$, we can see that the imposition of a structured sentencing scheme actually increases the trial-penalty: $E[J] - 30.4 > E[J] - 36$, regardless of $E[J]$.

This discussion is not meant to necessarily reconcile the positive ex-ante trial penalty observed in North Carolina with previous literature. Rather, this discussion is important to understand the immeasurable ways in which the structure of the criminal law may intentionally or unintentionally shape the shadow of the law.

6.1 Moving Forward - Decomposition of Differences

We find evidence not just of an ex-ante trial penalty, but of racial disparities in treatment by the North Carolina court system. We also find evidence that North Carolina's structured sentencing scheme does not provide uniformity in sentencing across the state. The obvious question after these results is how do the racial and regional differences manifest? There are four potential sources of racial and regional differences:

1. **Judges:** Conditional on conviction for a crime, do sentences imposed by judges vary by race or region?
2. **Juries:** Conditional on choosing a trial, do juries determine guilt rates differently by race or region?
3. **Prosecutors:** Do prosecutors charge defendants differently based on their race alone, and is there regional variation in charging decisions? Conditional on an initial charge, do prosecutors plea bargain differently based on defendants race or region?
4. **Defense Attorneys / Defendants:** Independent of all other actors, do defense attorneys and defendants plea bargain differently based on defendants race or region?

While these actors all operate in one environment, the each may be responsible independently for the differences we find. To see why, consider two defendants with no prior criminal records who are arrested on suspicion of breaking and entering. From the outset, prosecutors have the initial decision charge the individuals with a class H felony, a class 1 misdemeanor, or to not bring charges at all. If both defendants are charged with a class H felony, then do the prosecutor and defense attorney agree to plea to the felony or to a lesser charge. If they cannot agree on a plea bargain, do two juries find guilt at the same rate?

Regardless of whether by plea or trial, the final step of the process is how judges choose to sentence the defendants. North Carolina's structured sentencing scheme constrains the latitude available to judges; it has discretized the judges choices. Does the defendant get community punishment, in-

termediate punishment, or an active jail sentence? If a defendant gets an active sentence, is it in the mitigated (shortest), presumptive, or aggravated ranges? Is the sentence imposed concurrently or consecutively if there are multiple charges? Any small differences in judges' predilections to use probation over jail can produce systematically different sentences by race.

Of these four actors, separating the behavior of the prosecution from the defense is the hardest, as the two actors consistently make simultaneous choices. For example, the identification strategy used in this paper cannot be used, because any judicial idiosyncrasies are felt by both parties. Similarly, any defendant idiosyncrasy cannot be used for identification, as it too effect both parties. As such, we need to find a factor that influences either the prosecution or the defense but not both.

Two potential identification strategies include:

1. **Election Cycle:** District attorneys are elected for a 4-year term, while public defenders are unelected. Accordingly, we might think that contested elections represent shocks to the prosecution but not the defense.
2. **District Attorney Staffing:** The North Carolina legislature has frequently changed the budgets, staffing, and regional coverage of district attorney offices. When these staffing changes occur, we have a unilateral shock to the prosecution. Adding a new attorney lowers the caseload pressure on the other prosecutors, which might decrease the cost of trial.

Currently we are in the process of gathering data on these two factors. With data on these factors, we believe we can properly attribute the differences found in this paper to their appropriate agent.

7 Conclusion

The article seeks to further our understanding of whether defendants are better off accepting or rejecting a plea bargain. Unlike the vast majority of previous scholarship, we focus on comparing the unconditional average sentence from taking a plea bargain or declining to do so. While asking ex-post whether taking a plea bargain would have yielded a lower sentence is certainly of some interest, we assert that the shadow of the law model is predicated on the ex-ante decision. Using OLS regression to control for issues of selection based on observable case characteristics, we find evidence suggesting that on average the expected sentence is longer in North Carolina's for those who reject a plea, but that there is substantial heterogeneity. Two particularly interesting forms of heterogeneity is the type of crime committed and the race of the defendant.

There are real concerns that unobservable case characteristics may bias these estimates. For example, we might worry that the underlying true guilt of a defendant may influence the choice

to take a plea bargain and may be (hopefully) related to the ex-post sentence lengths, but is not directly observable to even the courts. As such, we employ instrumental variable regression using judge experience as the instrument for likelihood of pleading. The underlying logic justifying this instruments is straight-forward: uncertainty about case outcomes changes the propensity to go to trial. Knowledge of judicial temperament increases with judicial experience and reduces uncertainty about case outcomes. As such, we can use experience to instrument for the propensity plea.

We obtain statistically significant estimates of a positive trial penalty. Our OLS estimates suggest that defendants who plead guilty receive 8-month shorter sentences, with IV estimates of 10 months. Importantly, we see substantial racial heterogeneity. Both OLS and IV estimates suggest black defendants see greater reductions in sentence lengths after pleading than white defendants, but our OLS estimates suggest that white defendants see reductions in incarceration rates from pleading than black defendants do not. Lastly, counter to the stated goal of North Carolina's court system, we see tremendous geographic heterogeneity in sentence discounts and treatment of defendants of different races.

1 Appendix

1.1 Charlotte CSA Additional Tables

Table 11: Summary Statistics - Charlotte

Variable	*Overall Mean*	*Overall SD*	*No Plea Mean*	*Plea Mean*	*t-Statistic*
Charges	1.91	1.12	1.64	1.93	-13.49
Age	30.41	10.38	32.44	30.32	10.76
Race (black = 1)	0.55	0.5	0.59	0.55	4.44
Sex (female = 1)	0.16	0.37	0.12	0.17	-6.47
First Offense	0.44	0.5	0.64	0.43	22.15
Prior Points	3.97	5.11	2.8	4.02	-12.73
Structured Sentence Level (1=A, 10=I)	8.46	1.56	7.79	8.49	-23.8
All Charges Dismissed	0	0.02	0.02	0	33.65
Sentence (Years)	0.43	1.31	1.31	0.4	37.48
Incarceration	0.25	0.43	0.35	0.25	12.51
Non-zero Sentence	1.73	2.17	3.78	1.61	32.05
Guilty	0.98	0.14	0.61	1	-175.63
Lawyer Types					
Court Appointed Private Attorney	0.41	0.49	0.42	0.41	1.91
Public Defender	0.34	0.47	0.28	0.35	-7.15
Private Attorney	0.22	0.41	0.26	0.22	5.49
Waived	0.03	0.17	0.03	0.03	1.08
Observations	76,430		2,965	73,465	

Notes: This table reports summary statistics on 255,357 observations of felony cases in across NC from 1998-2010. Homicides and some sexual assault are excluded due to idiosyncratic nature of these offenses. For other cuts, see description in the text.

Table 12: Case Outcomes by Offense and Method of Disposition - Charlotte

				No Plea			*Plea*	
Offense Characteristics	**Share of Total Cases**	**Plea Rate**	**Pct. Black**	*Sentence*	*Incarceration*	*Finding of Guilt*	*Sentence*	*Incarceration*
Kidnapping	1 %	89 %	53 %	3.21	51.69 %	67 %	1.78	48.88 %
Robbery	8 %	94 %	73 %	3.8	61.56 %	71 %	1.49	52.43 %
Sexual Assault	3 %	94 %	37 %	1.39	34.53 %	50 %	0.89	37.95 %
Assault	5 %	90 %	54 %	1.6	38.62 %	58 %	0.65	33.32 %
Arson	1 %	94 %	35 %	0.78	27.27 %	50 %	0.46	23.4 %
Weapons	4 %	95 %	74 %	0.84	36.18 %	57 %	0.34	28.93 %
Dangerous Drugs	35 %	97 %	66 %	0.98	28.64 %	57 %	0.29	20.57 %
Burglary	15 %	97 %	41 %	0.6	33.69 %	73 %	0.28	26.43 %
Other	4 %	94 %	44 %	0.64	24.58 %	46 %	0.18	15.27 %
Traffic Offenses	2 %	98 %	66 %	0.07	12.5 %	63 %	0.15	21.75 %
Flee Arrest	1 %	95 %	42 %	0.52	48.78 %	80 %	0.18	23.1 %
Fraud	7 %	98 %	39 %	0.12	18.1 %	53 %	0.14	14.02 %
Larceny	8 %	97 %	42 %	0.4	23.96 %	64 %	0.14	18.79 %
Embezzlement	1 %	98 %	38 %	0.27	8 %	36 %	0.1	6 %
Forgery	3 %	98 %	45 %	0.23	14.71 %	53 %	0.08	11.47 %
Total	76430			2965			73465	

Notes: This table reports summary statistics on 255,357 observations of felony cases in across NC from 1998-2010. Homicides and some sexual assault are excluded due to idiosyncratic nature of these offenses. For other cuts, see description in the text.

Table 13: Effect of Pleading on Sentence Length and Incarceration (OLS) - Charlotte

	Sentence Length			Incarceration		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
Plea	-0.70 ***	-0.70 ***	-0.62 ***	-8.01 ***	-8.00 ***	-7.24 ***
	(0.02)	(0.02)	(0.02)	(0.75)	(0.74)	(0.71)
Black	-0.01 **	-0.02 ***	-0.07 ***	1.86 ***	1.28 ***	-0.67 ***
	(0.01)	(0.01)	(0.01)	(0.32)	(0.32)	(0.31)
Female	-0.14 ***	-0.14 ***	-0.12 ***	-7.78 ***	-8.26 ***	-7.11 ***
	(0.01)	(0.01)	(0.01)	(0.42)	(0.42)	(0.4)
Age	0.00 ***	0.00 ***	0.00 ***	0.13 ***	0.14 ***	0.06 ***
	(0.00)	(0.00)	(0.00)	(0.02)	(0.01)	(0.01)
Number of Charges	-0.12 ***	-0.11 ***	-0.10 ***	-7.50 ***	-6.97 ***	-6.73 ***
	(0.00)	(0.00)	(0.00)	(0.14)	(0.14)	(0.13)
First Offense	-0.29 ***	-0.28 ***	-0.16 ***	-21.33 ***	-20.65 ***	-17.24 ***
	(0.01)	(0.01)	(0.01)	(0.31)	(0.31)	(0.3)
Minimum Structured Sentence Length (Months)			0.03 ***			0.06 ***
			(0.00)			(0.02)
Appointed		0.20 ***	0.15 ***		11.98 ***	10.24 ***
		(0.01)	(0.01)		(0.4)	(0.38)
Public Defender		-0.05 ***	-0.05 ***		-1.04 ***	-1.20 ***
		(0.01)	(0.01)		(0.48)	(0.46)
Private		-0.14 ***	-0.10 ***		-5.26 ***	-4.25 ***
		(0.03)	(0.03)		(0.91)	(0.88)
Observations	74187	74158	74,001	74,187	74,158	74,001
Adjusted R^2	0.15	0.16	0.27	0.18	0.19	0.26
Mean Dependent	0.43	0.43	0.43	25.26	25.26	25.29

Notes: (***) refers to significance at the 1% level, (**) at the 5% level, and (*) at the 10% level.
Dependent variable in sentence length regressions is the minimum length of active sentence in excess of time served in years.
Dependent variable in incarceration regressions is a binary variable indicating whether an active sentence in excess of time served as received (yes = 100 and no = 0).

Control variables not displayed explicitly include: 1) district, 2) year fixed effects, 3) judge fixed effects, 4) crime type fixed effects, and 5) structured sentence punishment style fixed effects by sentence style. Waiving representation is the reference category for the lawyer type fixed effects.

1.2 Process of a Court Case

The general process for a court case in North Carolina is as follows:

- **Magistrate**
  Once charged, the defendant is arrested and brought before a magistrate judge. This magistrate will make a preliminary determination of whether the defendant can be released on bond or personal recognizance. Additionally, the magistrate will set the date of first appearance (one-to-three days after arrest) in District Court.

- **District Court**
  Within three days of seeing the magistrate (or at first session of district court), the defendant is arraigned in District Court. The District Court judge advises the defendant of the charges and rights, and may consider the bond amount or terms set by the magistrate.

  If requested, the District Court judge appoints an attorney to represent the defendant. This appointed attorney is not necessarily a public defender. Only 26 counties (16 districts) have a Public Defender office. When a public defender is unavailable, the judge appoints a local private attorney to act as the defendants council. While defendants must complete an Affidavit of Indigency to demonstrate that they cannot afford to retain private council, public defense is not free; if the defendant is found guilty or pleas guilty to their charged crimes, they must pay for their counsel.[61]

  Lastly, the district court judge will set a date for a hearing of probable cause, where the District Attorney is required to prove probable cause (i.e., showing that there is enough evidence to go forward with the case in Superior Court). The defendant has the right to waive this hearing, which sends the case directly to Superior Court. In general, most defendants choose to waive, sending the case to a grand jury in Superior Court. Choosing to waive the probable cause hearing speeds up discovery (allowing the defense to begin to understanding the quality of the states case) and may produce reductions in bail.

- **Superior Court**
  A defendant charged with a felony follows a multi-step process in Superior Court:
  - Voluntary Dismissal:

[61] North Carolinas court system explicitly states: "Court-appointed attorneys are not free attorneys. If you plead guilty or are found guilty, the judge will order you to reimburse the State of North Carolina for the value of services provided by a court-appointed attorney." `http://www.nccourts.org/County/_Common/Documents/ncids-StatePublicDefenderOffices-NCPDDirectory.pdf`.

The District Attorney may choose to dismiss any or all charges with a voluntary dismissal. A voluntary dismissal of any charges by the District Attorney does not prohibit the DAs office from refiling the charge. A second voluntary dismissal ends to states ability to bring further action.

- Grand Jury:
  The District Attorney must submit a bill of indictment charging the offense to the grand jury for its consideration. If the grand jury does not return a bill of indictment, the charges against the defendant are dismissed. If the grand jury returns a true bill of indictment, then a trial date is scheduled in Superior Court. The grand jury may also send the case back to the lower court to be handled as a misdemeanor.

  As with the hearing on probable cause, the defendant has the right to waive the grand jury hearing. To do so, the defendant submits a bill of information, agreeing to proceed to felony arraignment. Additionally, it is typical to waive the grand jury hearings in order to expedite the process, especially when a plea bargain has been negotiated.

- Bond Hearing:
  After the grand jury returns a true bill of indictment or the defendant waives the grand jury hearing, a hearing is held to determine the bond again for the defendant.

- Arraignment:
  The arraignment hearing in Superior Court is analogous to that in District Court. The Superior Court judge explains the charges against the defendant and his rights. Arraignment provides the defendant the opportunity to plead guilty to the charges. If the defendant pleads guilty or no contest, the judge will either immediately impose sentencing or schedule a sentencing hearing. If the defendant pleads not guilty, the judge will schedule a pre-trial conference.

- Pretrial Conference:
  The District Attorney and the defendants lawyer will meet to discuss the case. There may be hearings to discuss various pre-trial matters. At this point, the defense may submit pre-trial motions that require the State to dismiss the charge because there is no longer evidence that will be admissible at trial or because of procedural reasons. Assuming the case may continue, the judge will set a trial date.

– Trial:
If the defense and prosecution cannot reach a plea bargain, the case goes to trial. Until a voter referendum in 2014 opened the possibility of bench trials, all felony cases in North Carolina that were not plead out had to be settled through a jury trial. With the exception of the most serious felonies, trials last less than a week. Importantly for the analysis in this paper, while different judges may have been involved up-to this point in the process, judges are constant through the length of a trial.[62]

– Sentencing:
Unlike many other states, North Carolina has a complex structured sentencing scheme. Implemented through the Structured Sentencing Act of 1993, North Carolinas structured sentencing program separates felony charges into ten classes (Class A as the most severe and Class I the least) and convicted felons into six different criminal history levels (Level I the least severe and Level VI the most).[63] At each level, the structured sentencing program proscribes a minimum and maximum range of potential active jail time.

For example, the structured sentencing scheme proscribes that a defendant convicted of a class I felony (least severe) with no prior felony convictions be sentenced to community punishment only.[64] All else being equal, a defendant facing a class I felony with 19+ prior points has a presumptive sentence of 8 to 10 months of intermediate or active punishment.[65]

Lastly, if a defendant is found or pleads guilty to more than one crime, the judge has the discretion to impose the sentence consecutively rather than the default concurrent sentence. The decision to impose a consecutive sentence is often motivated by the defendant having committed multiple distinct offenses or having multiple victims. Additionally, judges may impose inactive but consecutive sentences as a deterrent to violating the terms of probation. [66]

[62] In a conversation with a North Carolina court staffer, we were told that only in the most extreme case of illness could a case have the judge switched during trial, but she could only think of a single instance where this happened.

[63] For a detailed discussion of the structured sentencing scheme, see: `http://www.nccourts.org/Courts/CRS/Councils/spac/Documents/statisticalrpt_fy12-13.pdf`.

[64] Community punishment include: probation, outpatient drug/alcohol treatment, community service, restitution, or fines.

[65] Intermediate punishment includes: special probation or a split sentence (some jail followed by probation), living at a residential treatment facility, daily reporting to a treatment facility, intensive probation, house arrest, or drug treatment court. An active sentence is incarceration.

[66] This description of the use of consecutive sentencing is described in detail at `http://www.nccourts.org/courts/crs/councils/spac/documents/disparityreportforwebr_060209.pdf`.

- Active Sentence:
  If a defendant receives an active sentence, they must serve the entire minimum sentence.[67] A defendant may earn time-credits that reduce the maximum sentence range, but never the minimum sentence. North Carolinas Secretary of Correction has the obligation to define how earned time may be earned or lost, and the Department of Correction or the custodian of a local confinement facility must supervise each individual inmates earned credit. If the defendant does not earn time, they will be released after satisfying the maximum range of their sentence.[68]

While we have frequently discussed plea bargaining, we have not placed plea negotiation at any specific stage of this process. This is because despite approximately 95% of court cases resolving with a guilty plea, there is substantial heterogeneity in the plea negotiation process.

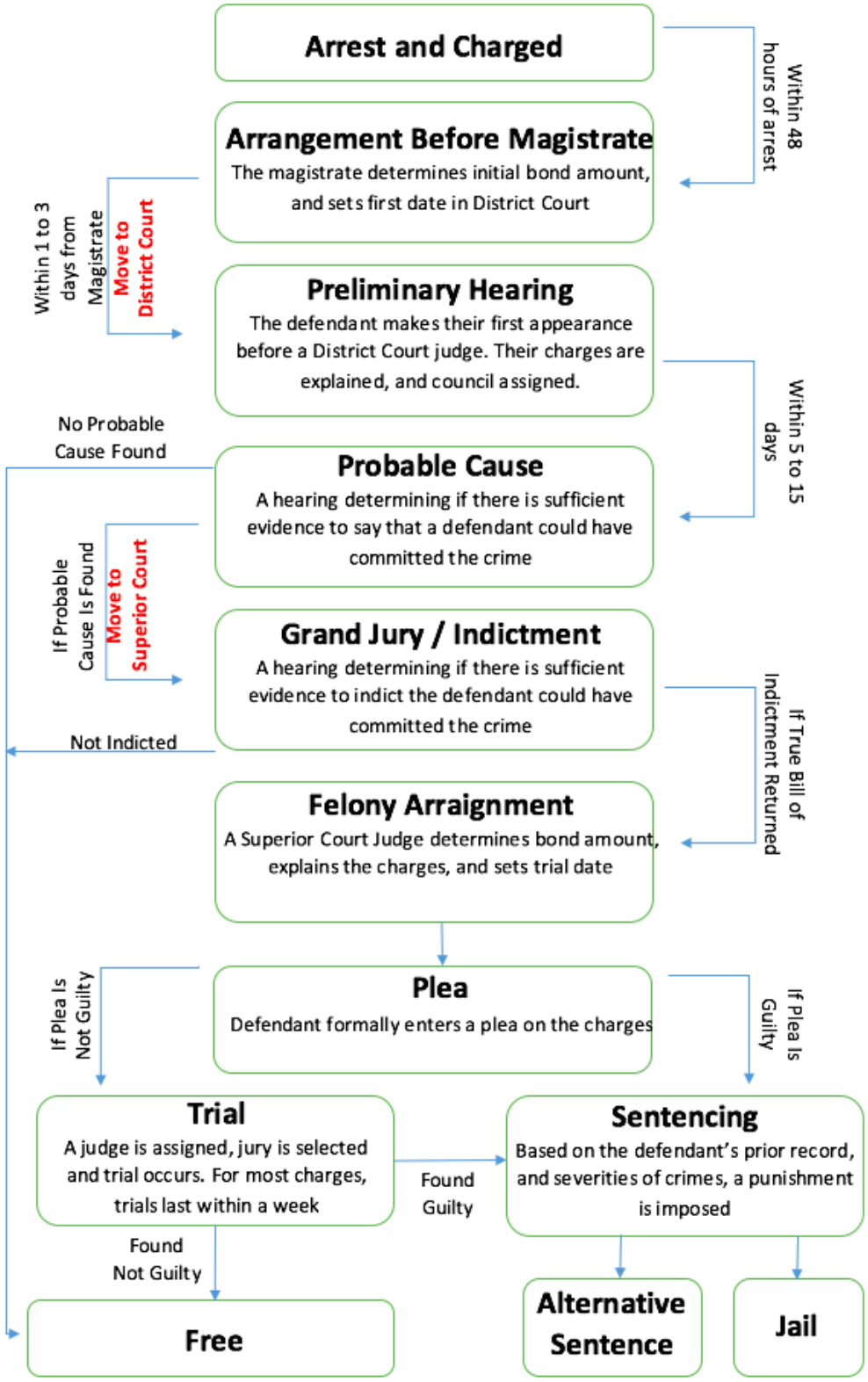
A Forsythe County public defender described how plea negotiations generally begin when the case is calendared in Superior Court and discovery is received. Similar sentiment was expressed by Assistant ADA Calvin King in Beaufort County, who described that negotiations begin after Superior Court indictment. A Guilford County public defender, however, said that if the defendant wants to plea, the negotiations can begin earlier in District Court.

Moreover, depending in part on the severity of the crime, a defendant may plead guilty to the lowest level offenses (H & I) before a District Court judge. This does not mean that the case process stops before reaching Superior Court, and that if the defendant chose not to plea, the case would be tried in District Court. Rather, it simply changes who the judge is that listens to the plea and imposes sentencing.

[67] This exact mechanics of determining a sentence length described in detail by the North Carolina Sentence and Policy Advisory Commission in the Structured Sentencing Training and Reference Manual `http://www.nccourts.org/Courts/CRS/Councils/spac/Documents/sstrainingmanual_09.pdf`.

[68] This policy was confirmed with the Combined Records Section of the NC Department of Public Safety at `http://www.ncdps.gov/combined-records`.

Figure 7: Diagram of the Process of a Case in North Carolina



1.3 Description of Data Cleaning

After fully cleaning the data, we analyze 359,652 cases with a final date of disposition between 1998 and 2010.[69] In order to ensure that our analysis and estimates are accurate, we have both extensively cleaned the data and restricted our attention away from the complete population of all court records. Below we describe the cleaning process and all restrictions made to the data.

1. **Restrict Attention to Felony Cases**
   While we have data on misdemeanors and traffic infractions, we are interested in a defendants choices when they potential jail time. As such, we first removed all non-felony charges that were not grouped with a felony charge. We do not remove all misdemeanor and traffic infractions at this stage because otherwise we might have situations where probation violation appears as the only charge. These remaining misdemeanors and traffic infractions constitute 9.27% of the fully cleaned data, and will be removed before analysis.

2. **Clean Defendant Information**
   The first step of the cleaning process uses the defendants personal information. We have information on the defendants: 1) names, 2) dates of birth, 3) race, 4) zip code, and 5) sex.

   For the defendants name, the custom of the court system is to write names as Last Name, First Name, Additional Names. While the records appear consistent in following the Last Name, First Name, the inclusion of middle names or initials is inconsistent. Accordingly, we define the defendants name to be the first and second name listed.

   Reducing defendants down to just a first and last name results in many duplicates. Only 81.68% of first and last name combinations are unique. For example, as expected, there are 76 John Smith in the records, with the most common name being James Williams. To separate our our different John Smiths, we used the defendants birthday and zip code. If the defendant listed in two different cases has the same full name[70] as well as the same birthday, we assume this is a unique person who interacted with the court system twice. If the defendant listed in two different cases has the same full name, has a birthday within a month of each other, and have the same zip code, we assume the difference in birthdates is a typo and that there is only one unique person.[71] With this assumption, we define a unique person to be a unique first name, last name, and date of birth.

[69] As explained later in the text, use the term cases to mean all charges that were disposed of for an individual on a single date.

[70] Full name being "Fackler Ryan Kanof" rather than "Fackler Ryan" which is just the first and last.

[71] This assumption effects 3,134 individuals out of the 450,302 unique defendant name, birthday combinations.

Having identified a person that we can follow through possibly numerous interactions with the court system, we cleaned the defendants races and sex. Sometimes an individuals race and gender is unknown or missing in one charge, but identified elsewhere. In this case, we assumed the identified race and sex is correct and extend to the other observations.

3. **Clean the Date of Disposition for a Case**
Ultimately, this project looks to put ourselves in the mindset of a defendant deciding to take a plea bargain or go to trial. To do so, we need to see the full gamut of charges a defendant faces when considering his or choice to plead guilty or proceed to trial. While charges are frequently grouped together with a common file-number,[72] a defendant may have charges with multiple file-numbers disposed of at the same time. As such, it is important to have an accurate identification of the final date of disposition for a case.

This identification faces two main complications. First, at any stage in the lifecycle of a case, it is possible to have any or all of charges dismissed, withdrawn, superseded, or otherwise modified. When such a modification occurs, the charge is given a date of disposition,[73] even if the remainder of the case progresses.[74] Second, if a defendant violates their probation, the probation violation may listed with the same file-number as the original case. For example, if a defendant received probation after pleading guilty to a robbery in 2006, and violated the terms of his probation in 2008, we may see charges with the same file-number with disposition dates years apart.

These multiple dates pose a problem; we want to consider all charges a defendant faces as a single incident. In our example above, the probation violation should not be grouped with the original incident, because the violation had simply not yet occurred at the time the defendant was deciding to plea. Accordingly, while for the majority of cases, there is a single date of disposition across all charges (94.21%), we determine the date of disposition for

[72] The court system store charges filed together with a common file-number. We cannot simply use this file-number as our definition of a case, because a defendant may have multiple file-numbers at once. For example, person 319298 in our data had 212 counts of uttering a forged instrument and forgery of an instrument separated across 102 distinct file-numbers that were all disposed of on the same day. If we used the courts file-number as a case, this defendant would have made 102 separate decisions to plea or go to trial, when in reality they made a single choice across all charges.

[73] While there are many potential dates to use, we believe the date of disposition the date that a charge finishes its lifecycle in the court system is the right date to use to group charges. In North Carolina, all outstanding charges, even if unrelated in their origin, are handled together where possible. While unrelated charges may have different starts into the judicial system, they will all have the same final date of disposition. In addition to grouping charges, we use the date of disposition to match identify the judge who imposes a defendants sentence.

[74] In the raw data, 445,711 (13.36%) charges are dismissed by any of several methods, and 946,743 (28.39%) charges were superseded by another charge.

charges grouped with a file-number as follows.

(a) For 93.02% of charges, there is a unique date of disposition for all non-probation charges that are adjudicated.[75]

(b) For 1.97% of charges, there is not a unique date of disposition for all non-probation charges that are adjudicated. In this case we take the earliest date of disposition.

(c) For 1.19% of charges that do not have a date of disposition on a non-probation charge that is adjudicated, there is a unique date of disposition listed for charges that eventually are dismissed or superseded.

(d) For 0.13% of charges that do not have a date of disposition on a non-probation charge that is adjudicated, there are multiple dates of disposition listed for charges that eventually are dismissed or superseded. Here we take the latest listed date of disposition.

(e) For 1.69% of charges, a probation violation is the only listed charge in Superior Court, but other charges are listed in District Court. In this case, we take the earliest date of disposition in District Court.

(f) For 2.00% of charges, probation is the only listed charge. In this case, we take the earliest date of disposition.

Using the above hierarchy, we defined a common date of disposition across all charges for each defendant. With this, we can now define a case as a unique person-date of disposition combination. This definition consolidates 3,335,121 charges listed across 750,033 file-numbers into 592,075 cases.

4. **Defining the Lead Charge**

Having now connected potentially multiple charges into a single case, we next collapse this collection of charges down to a single observation. We do this because we are interested in understanding how a defendant chooses to plea, and despite facing potentially many unique charges, the decision to plea is made over the full set.

One complication stemming from this decision to collapse multiple charges to a single observation is that we must decide which charges information should remain with the case. To do this, we first calculated the average sentence received when a defendant is convicted with each crime, and then defined the lead charge of an incident to be that charge with the

[75] By all non-probation charges that are adjudicated we mean any charge that finishes with a guilty plea, or a trial. We are excluding charges that: 1) are probation violations, 2) are eventually dismissed, or 3) are superseded by other charges.

highest associated sentence length.

To illustrate, consider an individual charged with two counts of felony possession with intent to sell or deliver cocaine and two counts of selling or delivering cocaine. The average minimum sentence conditional on conviction for selling or delivering cocaine is 1.78 years, and the average minimum sentence conditional on conviction for the possession charge is 1.02 years. Accordingly, we randomly assign one of the two charges of selling or delivering cocaine to be the lead charge, as if convicted, it is the longest expected sentence.

5. **Sentence Length and Incarceration**
When a defendant is found or pleads guilty to a felony, North Carolina imposes a sentencing range. If the judge determines the sentence should be active, the defendant is required to serve the full minimum of the range, and may serve less than the maximum with good behavior. For the purpose of this paper, we assume that the defendant serves the minimum sentence length imposed.[76]

In most (89.35% of raw data) of cases with an active sentence imposed, the sentences are to be served concurrently. This has the practical effect of making the charge with the highest minimum sentence range the defendants final sentence length. That is, if a defendant receives sentences of 2-4 months, 3-5 months, and 4-6 months to be served concurrently, their sentence is the same as had they just received 4-6 months only. In the remaining 10.65% of cases, the defendants sentence is served consecutively. In these cases, we determine a new sentence range by adding the minimum sentences imposed by the courts.

Lastly, our analysis does not involve the defendant anticipating a probation violation. This means that a defendant who receives a 4-6 month inactive sentence will consider this as having received no sentence at all. One could alternatively assume that a defendant places a probability of violating the probation, and discount the received sentence accordingly.

6. **Clean Judicial Information**
The identification strategy in this paper relies on having knowledge of what judge presided over a given case. Unfortunately, the court records do not naturally lend themselves to identifying the judge on a given case. When judicial information is present, the judge is listed as a 2 or 3 letter acronym (e.g. SCT or PJ), and no look-up table exists. These acronyms appear

[76] According to Jamie Markham, the average sentence length served is slightly above the minimum (102%) of the minimum for class B1 felonies to 113% of the minimum for class I felonies `http://nccriminallaw.sog.unc.edu/time-actually-served/`. We unfortunately do not have data on time served, so cannot independently produce this information.

idiosyncratic to the specific courthouse in which a case was heard.

To translate these acronyms into specific judges, we use the Master Schedules described above. The schedules provide us a week-by-week snapshot over which judicial district and division[77] judges may be presiding. Using this information, we construct one, two, and three letter acronyms for each judge in the schedule, and match this against our case data.

We have a hierarchy when merging the schedules against the case datas judicial acronyms and week of the date of disposition.[78] We say that an acronym corresponds to a specific judge if we can match:

(a) Unique two or three letter acronym matched at a judicial district for a given week

(b) Unique first letter of their last name matched at a judicial district for a given week

(c) Unique two or three letter acronym matched at a judicial division for a given week

(d) Unique first letter of their last name matched at a judicial district for a given week

This matching scheme allows us to identify a specific judge in 75.82% of cases where a unique acronym exists between 1998 and 2010.[79] In order to improve this match rate, we manually attempted to determine which judges specific acronyms may correspond to which name. With our manually cleaned acronym information, we matched acronyms against the master schedule again and identified a specific judge if we had a match with:

(a) Unique two letter acronym matched at a judicial district for a given week

(b) Unique two letter acronym matched at a judicial division for a given week

This allowed us to identify a specific judge in 84.32% of cases where we have a judicial acronym. The additional failure rate can stem from multiple sources including, 1) missing judge acronym information from the case-data, 2) multiple judge acronyms in the case data, 3) idiosyncratic acronyms (e.g. a clerk entering BR for Bob Robertson rather than RR for Robert Robertson), or 4) an acronym corresponding to a District Court judge.

In the cases where we cannot identify a specific judge, we constructed a composite of the

[77] As explained above, while judges are elected from a specific judicial district and division, they rotate districts within their elected division every six months.

[78] We merge the Master Schedule on a cases judicial acronym and the Sunday of the cases date date of disposition. The date of disposition of a case is when the case terminates in the court system (ignoring appeals), and is entered by the clerks at the same time as judge information is entered.

[79] A judicial acronym exists in 90.24% of cases.

expected judge according to the schedule. To explain, consider we have a case with a date of disposition in the week of July 7, 2008 in District 1, but we cannot identify a judge for the case. Looking at the master schedule, we see that Judges Tillet and Hinton were scheduled to be in residence in District 1 in that week. For the analysis in the paper, we assume one of these two judges is the judge who heard the case, and place half weight on the information and/or fixed effects from either judge.

The justification for the creation of this expected judge comes by putting ourselves in the mind of the involved parties. Before the actual plea or trial, there is uncertainty about which judge may eventually preside over a case were it to go to trial. But, the parties can form beliefs based on the master schedule, and these beliefs should influence the plea bargaining process a la the Shadow of the Law argument.

7. **Cut Data to Final Subsample**
With the cleaning of the data complete, we made several cuts to the data in order to obtain a final sample that is suitable for the analysis done in this paper. Starting with 584,971 cases:

   (a) We restricted attention to cases with a final date of disposition between 1998 and 2010. Despite having data on cases ranging from 1994-2010, we only have scheduling information from 1998 onward. This restriction causes us to drop 134,527 (23.00%) cases whose date of disposition was before 1998.

   (b) We excluded any case where a probation violation included in the list of charges, but we cannot identify the original case that the probation violation is connected to. We do this because the sentence corresponding to a violation of probation relates to both the original conviction that resulted in probation and to the act that caused the probation to be violated. This cut removes 38,152 cases (8.47% of remaining cases) where probation is at least one of the listed charges.

   (c) We excluded any case where all charges were superseded by another case, but we could not identify the superseding case. This should not occur with perfectly cleaned data, and in fact only occurs 20 times (0.01%) of the data.

   (d) We excluded all homicide and high-level (B2 and B1 Class offenses) sexual assault and offenses. These cases are idiosyncratic in many ways: often receiving significant media attention, extended-length trials, and having victims who suffer harm disproportionate to other crimes. The restriction on any level of homicide removes 8,066 (1.96%) of cases, and the restriction on severe sexual assault or offenses removes 8,282 (2.04%) more cases.

   (e) We remove any remaining cases where the lead charge is not at the felony level. These

few cases remain because the felony violation. This removes 3,439 (0.87%) of the remaining cases.

(f) We remove any cases where habitual felony is a charge in the case,[80] as the final sentence disconnected from the underlying lead offense because of the presence of the habitual felony charge. This removes 15,905 (4.05%) of remaining cases.

(g) Lastly, we remove any cases where neither a judge could be identified, nor an expected judge could be created. This could happen under several circumstances including if the observation comes from a district-week where the only judges are listed as "TBA". This cut removes 16,928 (4.50%) of remaining cases, leaving us with 359,652 cases.

In addition to the above cuts, we made one additional and significant cut to our sample for our IV strategy to be successful. We used the number of cases and the number of trials that a specific judge has presided over in a district as a proxy for the defense and prosecutions knowledge about judge idiosyncrasies.

To accurately measure how many cases a judge has seen, we necessarily must restrict our attention to newer judges. To see why, consider if we simply started counting the number of cases a judge presides over starting in 1998 with out data. All judges, regardless of whether they have heard 5,000 cases before 1998 or if they were hearing their very first case starting in 1998 would be considered to have only heard one case in the first time they appear in our data.

Accordingly, we only include judges who: 1) hear their first case any time after 1998, or 2) hear their first case in the specific district after 2000. If we do not have a match for a specific judge, we only use the expected judge if all judges first these criteria. While potentially over-restrictive, this choice is designed to avoid the problem of incorrectly counting the experience of existing judges. After this cut, our final dataset has 148,000 (losing 211,652 cases or 58.85% of cases).

[80] Habitual felony is a North Carolina specific charge that allows for sentencing modification for repeat offenders. See `http://sogpubs.unc.edu/electronicversions/pdfs/aojb0804.pdf`.

1.4 Structured Sentencing

Table 14: Share of Cases: Crime Class Vs. Prior Points Table

		Level 1	*Level 2*	*Level 3*	*Level 4*	*Level 5*	*Level 6*
	Factor	0-1	2-5	6-9	10-13	14-17	18+
A		Death or Life Without Parole					
B1	Aggravated	240-300	276-345	317-397	365-456	Life Without Parole	
	Presumptive	192-240	221-276	254-317	292-365	336-420	386-483
	Mitigated	144-192	166-221	190-254	219-292	252-336	290-386
B2	Aggravated	157-196	180-225	207-258	238-297	273-342	314-393
	Presumptive	125-157	144-180	165-207	190-238	219-273	251-314
	Mitigated	94-125	108-144	124-165	143-190	164-219	189-251
C	Aggravated	73-92	83-104	96-120	110-138	127-159	146-182
	Presumptive	58-73	67-83	77-96	88-110	101-127	117-146
	Mitigated	44-58	50-67	58-77	66-88	76-101	87-117
D	Aggravated	64-80	73-92	84-105	97-121	111-139	128-160
	Presumptive	51-64	59-73	67-84	78-97	89-111	103-128
	Mitigated	38-51	44-59	51-67	58-78	67-89	77-103
E	Aggravated	25-31	29-36	33-41	38-48	44-55	50-63
	Presumptive	20-25	23-29	26-33	30-38	35-44	40-50
	Mitigated	15-20	17-23	20-26	23-30	26-35	30-40
F	Aggravated	16-20	19-23	21-27	25-31	28-36	33-41
	Presumptive	13-16	15-19	17-21	20-25	23-28	26-33
	Mitigated	10-13	15-19	13-17	15-20	17-23	20-26
G	Aggravated	13-16	14-18	17-21	19-24	22-27	25-31
	Presumptive	10-13	12-14	13-17	15-19	17-22	20-25
	Mitigated	8-10	9-12	10-13	11-15	13-17	15-20
H	Aggravated	6-8	8-10	10-12	11-14	15-19	20-25
	Presumptive	6-8	6-8	8-10	9-11	12-15	16-20
	Mitigated	4-5	4-6	6-8	7-9	9-12	12-16
I	Aggravated	6-8	6-8	6-8	8-10	9-11	10-12
	Presumptive	4-6	4-6	5-6	6-8	7-9	8-10
	Mitigated	3-4	3-4	4-5	4-6	5-7	6-8

The only categories that do not automatically yield an active sentence are: E(1-2), F(1-3), G(1-4), H(1-5), and I(1-6). Community punishment is an option only for H(1), and I(1-2). An active sentence isn't available for only I(1-3). Community punishment is the only option for I(1).

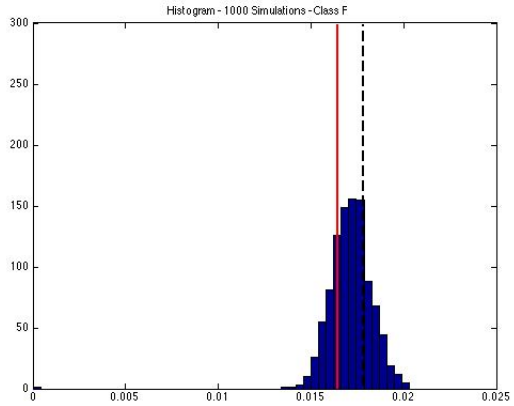
Table 15: Share of Cases: Crime Class Vs. Prior Points Table By Plea - Full State

	Level 1	*Level 2*	*Level 3*	*Level 4*	*Level 5*	*Level 6*	*Total*
Plead							
Crime C	1.07	0.95	0.48	0.23	0.06	0.05	2.84
Crime D	2.32	1.9	1.07	0.63	0.17	0.14	6.24
Crime E	1.42	1.33	0.6	0.26	0.07	0.05	3.73
Crime F	2.05	1.57	0.98	0.47	0.13	0.11	5.31
Crime G	3.92	4.79	3.08	1.83	0.41	0.3	14.34
Crime H	13.11	14.81	8.14	4.65	1.3	0.94	42.96
Crime I	7.48	9.21	4.4	2.19	0.59	0.47	24.33
Total	31.47	34.66	18.8	10.28	2.73	2.06	100
Not Plead							
Crime C	4.48	1.43	1.25	0.77	0.18	0.17	8.27
Crime D	5.47	2.18	1.69	1.25	0.36	0.27	11.22
Crime E	4.24	1.31	0.85	0.35	0.18	0.08	7
Crime F	4.48	1.04	0.77	0.51	0.12	0.14	7.07
Crime G	6.6	2.07	1.57	0.89	0.31	0.18	11.61
Crime H	19.82	5.2	3.41	1.87	0.52	0.47	31.28
Crime I	14.87	3.49	1.79	0.75	0.26	0.14	21.31
Total	62.1	16.75	11.36	6.4	1.94	1.45	100

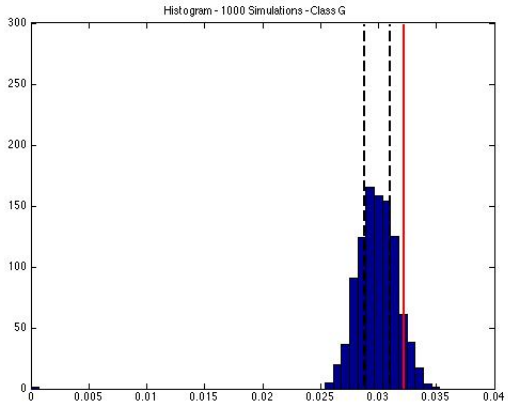
Notes: This table reports summary statistics on 255,357 observations of felony cases in across NC from 1998-2010. Homicides and some sexual assault are excluded due to idiosyncratic nature of these offenses. For other cuts, see description in the text.

1.5 Judicial Randomization Check

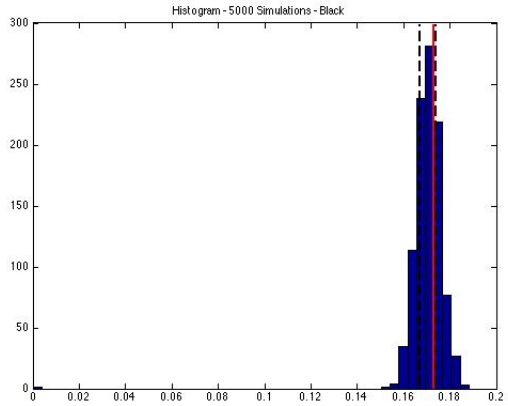
Figure 7: Randomization Check Through Simulations



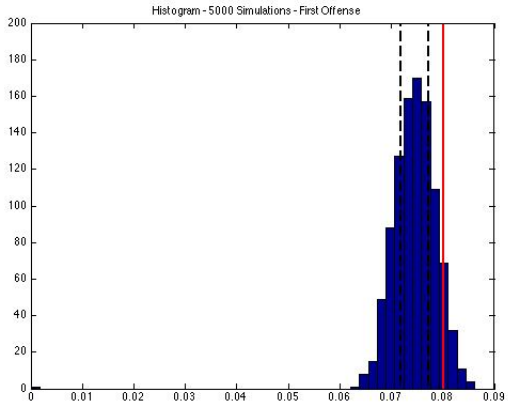
(a) Class F



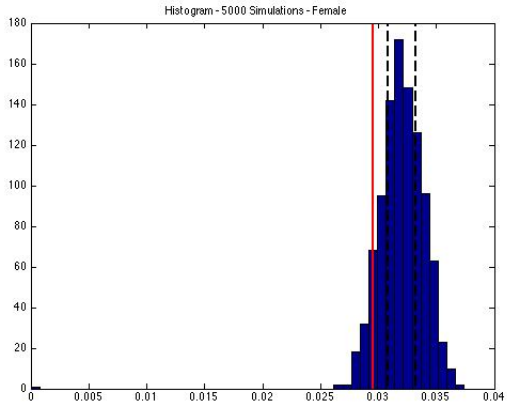
(b) Class G



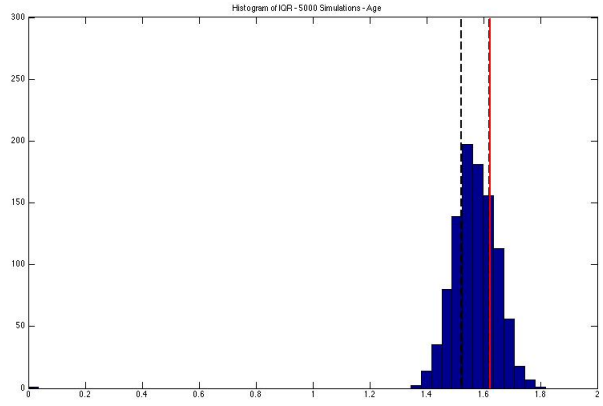
(c) Black



(d) First Offender



(e) Female



(f) Age

References

Abrams, D., Bertrand, M., and Mullainathan, S. (2012). Do Judges Vary in Their Treatment of Race? *Journal of Legal Studies*, 41(2):347–383.

Abrams, D. S. (2011). Is Pleading Really a Bargain? *Journal of Empirical Legal Studies*, 8(s1):200–221.

Abrams, D. S. (2013). Putting the Trial Penalty on Trial. *Duquesne University Law Review*, 51.

Alschuler, A. (2013). Lafler and Frye: Two Small Band-Aids for a Festering Wound. *Duquesne Law Review*, 51.

Bibas, S. (2004). Plea Bargaining Outside the Shadow of Trial. *Harvard Law Review*, pages 2463–2547.

Brereton, D. and Casper, J. (1982). Does It Pay to Plead Guilty? Differential Sentencing and the Functioning of Criminal Courts. *Law & Society Review*, 16(1):45–70.

Cooter, R., Marks, S., and Mnookin, R. (2012). Bargaining in the Shadow of the Law: A Testable Model of Strategic Behavior. *The Journal of Legal Studies*, 11(2):225–251.

Daughety, A. F. and Reinganum, J. F. (2005). Economic Theories of Settlement Bargaining. *Annual Review of Law and Social Sciences*, 1:35–59.

Grossman, G. M. and Katz, M. L. (1983). Plea Bargaining and Social Welfare. *The American Economic Review*, 73(4):749–757.

Jones, J. (1978). Prosecutors and the Disposition of Criminal Cases: An Analysis of Plea Bargaining Rates. *Journal of Criminal Law and Criminology*, 69(3).

Landes, W. M. (1971). An Economic Analysis of the Courts. *The Journal of Law & Economics*, 14(1):61–107.

Nagel, I. and Schulhofer, S. (1989). Negotiated Pleas Under the Federal Sentencing Guidelines: The First Fifteen Month. *American Criminal Law Review*, 231.

Priest, G. L. and Klein, B. (1984). The Selection of Disputes for Litigation. *The Journal of Legal Studies*, 13(1):1–55.

Reinganum, J. F. (1988). Plea Bargaining and Prosecutorial Discretion. *The American Economic Review*, pages 713–728.

Rhodes, W. (1979). Plea Bargaining: Its Effect on Sentencing and Convictions in the District of Columbia. *Journal of Criminal Law and Criminology*, 70(3).

Schulhofer, S. and Nagel, I. (1996). Plea Negotiations Under the Federal Sentencing Guidelines: Guideline Circumvention and its Dynamics in the Post–Mistretta Period. *Northwestern University Law Review*, 91.

Smith, D. A. (1986). The Plea Bargaining Controversy. *The Journal of Criminal Law and Criminology*, pages 949–968.

Stuntz, W. J. (2004). Plea Bargaining and Criminal Law's Disappearing Shadowl. *Harvard Law Review*, pages 2548–2569.