The Electoral Connection in Court: How Sentencing Responds to Voter Preferences

Joshua Boston Bernardo S. Silveira

June 17, 2022

Abstract

Do elected judges tailor criminal sentences to the electorate's ideology? Utilizing sentencing data from North Carolina's Superior Courts—which transitioned from statewide to local elections in 1996—we study whether judges are obliging to voters' preferences. We find some evidence of responsiveness; judges from liberal districts were more lenient, while those from moderately conservative districts assigned harsher sentences. Judges from increasingly conservative districts did not change their sentencing patterns, which leads to lower re-election rates. These findings suggest that judges adapt their behavior to retain office, or else they are held accountable by the public.

Manuscript word count: 8,882

Abstract word count: 95

*Boston (jboston@bgsu.edu): Bowling Green State University, Department of Political Science; Silveira (silveira@econ.ucla.edu): University of California, Los Angeles, Department of Economics. The authors sincerely thank JB Duck-Mayr, Dino Hadzic, James L. Gibson, Sanford C. Gordon, Alessandro Lizzeri, Michael J. Nelson, Miguel M. Pereira, Andrew R. Stone, Nicholas W. Waterbury, and conference participants at the 2019 ALEA Annual Meeting for their thoughtful comments that have significantly improved this manuscript.

It is common across the American states for judges to be elected. Only 11 states do not hold any trial court judicial elections[1], leaving thousands of judges across the remaining states subject to various public accountability mechanisms. These range from infrequent retention elections to partisan primaries and general elections. To that end, a great deal of research has addressed the extent to which local and state judges are responsive to their constituents' preferences (e.g., Hall 1987; Huber and Gordon 2004; Hanssen 2004; Brace and Boyea 2008; Choi, Gulati and Posner 2010).

Much of the existing literature on judicial accountability focuses on the nature of selection: appointive vs. elective, partisan vs. non-partisan, competitive vs. retention, etc. (e.g., Taylor N.d.; Badas and Stauffer 2019; Canes-Wrone, Clark and Park 2012). Our study focuses, instead, on how trial court judges respond to changes in their constituencies. When an incumbent judge's geographic constituency changes, does sentencing behavior conform to the new electorate's preferences? We also examine whether—in subsequent elections—the public punishes judges who are not responsive. If judges do not conform their sentencing behavior to their new constituents' preferences, then the judges (1) might have incentive to retire, or (2) could be punished at the ballot box.

Our study exploits a change in the electoral rules that govern trial court judicial elections in North Carolina. Up until 1996, N.C. Superior Court judges had been elected in state-wide races. Reforms under the label Bill 41 localized Superior Court elections into 46 district-level competitions. We examine the degree to which judges exhibit responsiveness to local voters using 135,481 case-level terminations resulting in incarceration before and after the Bill 41 electoral reforms.[2]

Our results—which are conditional on the relative ideological polarization of each

[1] According to the Brennan Center for Justice, as of 2022, these states are Connecticut, Delaware, Hawaii, Maine, Massachusetts, New Hampshire, New Jersey, Rhode Island, South Carolina, Vermont, and Virginia. See: https://www.brennancenter.org/judicial-selection-map.

[2] We also employ a structural approach for judges' sentencing behavior to account for the selection processes determining which cases are settled and which result in a conviction at trial (Silveira 2017).

district—suggest that Bill 41 induced accountability to the narrower, district-level constituency among *some* of North Carolina's judges. Cumulative distribution plots show that judges across liberal and conservative districts assign noticeably different sentences after Bill 41. Our regression analyses—controlling for judge-level characteristics through fixed effects—reveal similar results. First, districts that closely approximate North Carolina's mean ideology do not yield a change in judicial behavior. Second, judges tend to sentence more leniently when assigned to districts that are increasingly liberal. Third, judges assigned to conservative-leaning districts—those that are ideologically adjacent to the statewide mean ideology—sentence more punitively. Fourth and surprisingly, districts that are considerably more conservative relative to the entire state do not show responsive sentencing behavior.

Despite the asymmetry of our results, further analysis shows that the electorate will still have its say; we find that unresponsiveness leads to electoral punishment or preemptive retirement. Those relatively unresponsive conservative judges—those who respond least to their narrower constituencies introduced by Bill 41—are more likely to end their judicial service sooner.

Electoral incentives for judicial behavior

Elections are meant to induce responsiveness to public preferences. That much has been clear in political science at least since the days of Miller and Stokes (1963), Fenno (1978), and Mayhew (1974). The electoral connection is concisely captured in the oft-repeated theorem that U.S. Members of Congress are "single-minded seekers of reelection," and that their "re-election quest establishes an accountability relationship with an electorate" (Mayhew 1974, 5-6). Indeed, scholars have provided extensive evidence regarding electoral accountability; candidates for public office adjust their platforms to suit the electorate's interests and, once elected, avoid making decisions that may elicit future rejection at the ballot box (e.g., Besley and Coate 2003; Barro 1973; Ferejohn 1986; Besley 2006; Bartels 1991; Glazer and Robbins 1985).

This line of inquiry is well-entrenched in the state judicial politics literature, where scholars have persistently examined how various institutions provide different incentives for judicial actors to conform to public opinion. From Hall's (1987; 1992) pioneering work, which establishes theoretical arguments and empirical evidence regarding the electoral incentives judges face in terms of public preferences, scholars have expanded their examinations of how judicial behavior systematically varies with institutional rules. These include variations across states in terms of (1) selection mechanisms—in terms of, merit selection (e.g., Hall 2001; Goelzhauser 2018), initial election (e.g., Bonneau 2006; Gill and Eugenis 2019) and retention elections (e.g., Hall and Brace 1999; Holmes and Emrey 2006)—(2) partisan signals to voters (e.g., Bonneau and Cann 2015; Canes-Wrone, Clark and Kelly 2014; Badas and Stauffer 2019), and (3) salience of legal policy areas (e.g., Canes-Wrone, Clark and Semet 2018; Carson et al. 2011).

Our study focuses on competitive elections for state trial court judgeships. Many prior studies of judicial responsiveness focus on state courts of last resort, given their policy impact on salient issues like capital punishment and abortion (Canes-Wrone, Clark and Park 2012; Canes-Wrone, Clark and Kelly 2014). Still, trial court judges face similar electorally induced incentives, especially in terms of criminal sentencing—a salient issue on which the public holds "well-defined policy preferences" (Cann and Wilhelm 2011); increases in public punitiveness over the last half century have corresponded to higher incarcerations rates across the United States (Enns 2014). But evidence remains mixed regarding what institutions tend to incentivize higher levels of judicial responsiveness among trial court judges.

For example, Gordon and Huber (2007) use Kansas's district-level variations across partisan-contested and non-partisan retention elections to provide evidence of greater punitiveness among judges elected in partisan systems. Lim (2013) finds evidence that sentencing is considerably more heterogeneous among Kansas judges who rely on partisan signals in their re-election efforts, as compared to those judges facing retention election, who must signal voters with their sentencing behavior. Looking at cross-state variations in sentencing, Taylor (N.d.) finds that trial judges facing retention and non-partisan elections tend to be marginally more punitive than those who face partisan election. Electoral incentives are also temporally conditional, where length of term offers the strategic judge an opportunity to conform to public preferences differently over time. Huber and Gordon (2004) find that trial judges sentence criminals more punitively as elections approach. Judicial incumbents—across the partisan divide—avoid instances of perceived under-punishment as public and media attention increases with proximate elections. Our research contributes to this line of inquiry by examining how trial court judges change their sentencing behaviors in response to their state's new electoral rules.

This research also examines the effect of constituency size on judicial responsiveness. Scholars have noted the empirical norm that incumbent vote shares tend to be decreasing as the size of the electorate increases (e.g., Hogan 2004). And those regularities hold in judicial elections, as several notable state court studies consider the effect of district-level vs. state-wide constituencies. Hall and Bonneau (2006) find that state high court partisan elections centered in districts rather than statewide elections result in fewer challengers than (a) non-partisan district elections or (b) partisan state-wide elections. Streb and Frederick (2009)—studying variations across states in whether intermediate appellate judgeships are selected on statewide or district-level ballots—find that district-level elections are less competitive, decreasing the likelihood of a challenger. At the same time, the literature suggests that district-level competitions lower the cost for candidate entry. In particular, scholars identify this regularity in the context of congressional redistricting, where district maps offer incentives for challenger entry (e.g., Carson et al. 2011; Williamson 2019). We believe our research is among the first to examine how judges respond to being assigned to localized constituencies after previously being elected in statewide competitions.

North Carolina' Institutional Changes in Bill 41

This study examines judicial responsiveness among North Carolina's Superior Court judges—the state's main trial court with jurisdiction over felony cases, civil cases involving more than $10,000, and misdemeanor cases appealed from North Carolina's District Courts.[3] On August, 2 1996, North Carolina ratified Bill 41, which changed the electoral institutions used to select Superior Court judges. Prior to Bill 41, judges (a) competed in statewide elections, and (b) rotated across the state's 46 districts.[4] After Bill 41, Superior Court elections localized into districts.

Legal commentary and analysis around this time suggests that North Carolina policymakers wanted to enhance the electoral connection for judges. Leading up to the introduction of Bill 41 in North Carolina's General Assembly, scholarly debate discussed how different selection mechanisms—including merit selection with retention elections—might achieve varying degrees of responsiveness to voters (e.g., Rosch and Rubin 1987; Helms 1987). One policy research memo discussed that, over the recent decades, "the nature of [North Carolina] judicial elections made it difficult to identify any real issues and any real reasons for ousting an incumbent judge" (Grimes 1997, 2287). While part of this stems from the Democrats dominance of North Carolina politics for the better part of the 20th century, Grimes (1997, 2314) notes that the emergence of genuine two-party competition in the 1980s—along with relaxed campaign and election rules—helped to resolve "past problems with voter apathy in

[3] Superior Court criminal trials require a twelve-member jury, whose verdict convicts the defendant or not. Judges determine a punishment as constrained by a set of structured sentencing guidelines, which consider the severity of the crime and the defendant's criminal history. The sentence may include incarceration, probation, community service, or alternative penalties. Plea bargains that meet the above criteria likewise terminate with a Superior Court judge.

[4] Though judges faced a statewide electorate, they served in district-based seats. The NC constitution mandates judges to rotate across districts—rotations that occurred roughly every six months.

North Carolina's judicial elections."

Bill 41 also instituted a second electoral reform, which would take effect in later elections: *primary* elections in 1998 and later would be non-partisan. When more than two individuals seek the same judicial office, a non-partisan primary clusters candidates of all partisan affiliations to run in a single "jungle" or "top-two" primary. In effect, this allows for two candidates of the same partisanship to run against one another for the same office in the general election, while the voters do not have a partisan signal across contests.

While the 1996 election remained a partisan competition, North Carolina's eventual change to non-partisan primaries in 1998 poses a threat to our inferences. Our goal is to study the effect of imposing district-level elections in 1996. While non-partisan and partisan elections differ in the information given to voters, prior scholarship provides empirical evidence that voters are sophisticated enough to perceive partisanship without labels at the ballot box (e.g., Bonneau and Cann 2015). Those findings might help to ameliorate any concerns we have about sequential, contaminating institutional changes. Even more, our identification strategy is driven by incumbents who were elected in 1996 or earlier under partisan labels, suggesting that those judges have a preexisting partisan reputation with voters.

Expectations for judicial responsiveness

Our research examines trial judges' substantive behavior vis-à-vis criminal sentencing, and how those choices might correspond to local preferences. As states have changed judicial selection mechanisms to induce varying degrees of public accountability, scholars continue to focus on the degree to which trial court judges are attentive to their constituents' preferences. North Carolina's judicial reforms in the mid-1990s are particularly ripe for such an endeavor, as the geographic narrowing of judicial electorates following Bill 41 alters the incentives trial judges face. In particular, Bill 41 changed North Carolina's trial court judicial elections from statewide to district-level constituencies. Given a variety of considerations, we develop expectations for how these electoral changes might impact the choices judges make.

We identify three realistic possibilities for what these electoral changes might mean for criminal sentencing. First, a judge—in anticipation of changes in the electorate—might always assign more punitive sentences in order to satisfy voters and deter negative media and interest group attention. This would be consistent with existing evidence of judicial accountability, especially in low-information retention elections (e.g., Huber and Gordon 2004; Canes-Wrone, Clark and Semet 2018). Second, judges' policy preferences may be strong enough that sentencing behavior would not systematically vary in response to changes in the electorate. Third, judicial sentencing could be a dynamic, strategic process, whereby judges attempt to retain office by tailoring their sentences to the preferences of their voters.

We believe the third possibility is most likely at play in North Carolina's Superior Courts following the electoral reforms. That is, the incentives created by localized elections tend to push judges toward responsiveness because local voters now can focus their attention on local judgeships. Electoral change incentivizes office-seeking and office-retaining judges to defer to their constituents' preferences; judges assigned to liberal districts will sentence criminals more leniently while judges assigned to conservative districts will sentence more punitively. As judges' district-level electorates are increasingly ideologically polarized, we expect to observe corresponding increases and decreases in judges' punitiveness. Judges assigned to moderate districts—that is, those districts approximating the state's mean ideology—should maintain their prior sentencing practices, as they had previously been elected by voters statewide. Finally, we believe those judicial actors who do not respond to their new constituents will lose re-election or opt not to continue their service on the Superior Court.

Existing studies of district size in judicial elections suggest countervailing pressures on incumbent judicial office holders. On one hand, smaller districts tend to yield larger vote margins for incumbents compared to larger or state-wide districts. On the other hand, smaller districts lower the cost for candidate entry and increase constituent attention on incumbent officeholders. We believe this last point is particularly important following an institutional change, as took place with Bill 41 in North Carolina. Furthermore, scholars also examine how interest groups and media impact constituents perceptions of incumbent office-holders. Localized elections might decrease the efficacy of outside groups in attracting the public's attention to incumbent judges as compared to statewide judicial races. Still, Bill 41 simplifies the informational environment for voters, who—after 1996—only vote for judicial candidates in their home districts.

While we expect the choices judges make to tend toward responsiveness, we anticipate several reasons why some judges may be unwilling to acquiesce to local voters' preferences. First, some judges may desire electoral or appointive promotions within or outside the state judiciary—ambition that requires attention to a broader constituency (e.g., Nelson 2014; Budziak 2013; Jensen and Martinek 2009). Second, judges have sincere policy preferences, and significant deviations from those preferences decreases a judge's overall utility (e.g., Brace, Hall and Langer 1998). Third, some judges may have realized that—after Bill 41—their electoral prospects were dim, and therefore decided not to cater to local voters. Fourth, trial court judicial behavior is intricately related to prosecutorial behavior and discretion—a consideration that we directly address below.

Measuring sentencing & district preferences

We expect that North Carolina's judges will strategically defer to their district voters' preferences in terms of criminal sentencing. Within that expectation, we measure (1) a judge's sentencing at the case level as our outcome variable, and (2) the post-Bill 41 district-level preferences as our primary explanatory variable. We begin by discussing our outcome variable.

Outcome Variable: Sentencing decisions

To measure a judge's sentencing, we utilize incarceration sentences assigned to an individual defendant in a given case. We obtained sentencing data from the North Carolina Administrative Office of the Courts. These data comprise every case filed at the North Carolina Superior Courts from January 1995 to October 2010. From these cases, we keep in our sample those in which (a) the defendant was found guilty, and (b) incarceration was a component of the sentence. Each case includes detailed information on case disposition, charged offenses, and characteristics of the defendants. We utilize 135,481 case-level terminations before and after Bill 41's passage in August 1996. We provide an overview of these data in Table 1 below.

As we summarize in Table 1 below, cases in our outcome variable can be terminated through (1) judge-assigned sentences or (2) judge-approved plea bargains. We include both for purposes of our outcome variable. First, judge-assigned sentences occur after jury or bench trial convictions using North Carolina's structured sentencing system. The system requires the judge to choose a sentence from a predetermined range, which depends on (a) the severity of the offense and (b) the offender's previous criminal record—variables for which we also control in our models. The sentence may generally consist of incarceration time or alternative punishments such as probation.[5] Although the system imposes constraints on the choices of judges, it still leaves considerable discretion for the assignment of sentences. For example, an offender with no prior criminal history who is convicted of assault with a deadly weapon with intention to kill may be may be sent to prison or be assigned an alternative punishment (e.g., probation). If the judge decides to assign incarceration time, the minimum sentence length ranges from 15 to 31 months.[6]

[5] The judge only sets a minimum incarceration length. The maximum length is determined according to a formula, and is roughly 120 percent of the minimum. See Silveira (2017) for details.

[6] We could not consider in the analysis cases in which it is not specified whether or not the offender was

Second, our outcome variable also includes sentences determined by a plea bargains, an empirical approach that is consistent with existing models of judge sentencing (e.g., Gordon and Huber 2007; Huber and Gordon 2004). As it happens in all American states, the vast majority of the cases are terminated through negotiated guilty pleas, which we observe in the topmost section of Table 1. Consistent with a well-established literature in law and economics, we contend that bargaining operates under the shadow of the judge—that is, the defendant and the prosecutor negotiate taking into consideration the harshness of the sentence to be assigned in the event of a conviction at trial (Elder 1989; LaCasse and Payne 1999; Kuziemko 2006; Boylan 2012; Bonneau and McCannon 2019). Hence the judge's expected sentencing behavior also affects cases decided by a plea bargain, although the effect is indirect. In our regressions, we include an indicator variable for bargained cases, along with dummy variables for the judge in each terminated case.[7]

In addition to detailing our outcome variable in terms of method of disposition, Table 1 also presents summary statistics of time of disposition (i.e., pre- and post-Bill 41), charged offense, defendant characteristics, and average sentence length across method of disposition. Again, these statistics only take into consideration cases that resulted in incarceration. The population of cases includes many more male than female defendants. Whites and African Americans defendants appear in similar proportions and represent most of the cases, and the average defendant age is 30.84 years. There are less observations for cases disposed in 1996 than for other years. Still, there are 4,950 cases disposed before the approval of Bill

incarcerated, as well as cases where incarceration time was assigned but the minimum sentence is not reported. Sentences in multiple count cases may be either consecutive or concurrent. We don't have this information in the data. In the analysis, we assume that all sentences are concurrent.

[7] To identify each judge, we use the annual editions of the North Carolina Manual to organize a complete list of Superior Court judges active in each year of our analysis. In the sentencing data, judges are only identified by their initials. In most cases, three initials are used. We match such initials to the full names of the judges as reported annually on the North Carolina Manual. In the period comprised by the sentencing data, only two pairs of judges had the same three initials. Cases decided by these judges were excluded from the data. We also excluded all the cases in which the judge was either identified by less than three initials or not identified at all.

Table 1: Descriptive statistics—Incarceration convictions

Distribution of cases by outcome		
Method of disposition	Observations	Frequency
Plea bargain	127,393	94.03%
Trial	8,088	5.97%
Total	135,481	100%
Date of disposition	Observations	Frequency
Prior to Bill 41	4,950	3.65%
After Bill 41	130,531	96.35%
Total	135,481	100%
Distribution of cases by severity of the charge		
Main charged offense	Observations	Frequency
Felony	121,222	89.43%
Other	14,259	10.57%
Total	135,481	100%
Defendant's characteristics		
	Observations	Frequency
African-American	80,733	59.59%
Hispanic	3,588	2.65%
Female	10,197	7.53%
	Mean	Standard deviation
Age (years)	30.85	9.94
Sentences' length by conviction method (months)		
	Mean	Standard deviation
Trial convictions	91.25	134.27
Settlement convictions	23.48	42.76

Notes: This table, which is based on data form the North Carolina Administrative Office for the Courts, refers to criminal cases decided at the North Carolina Superior Courts from January 1995 to October 2010. We exclude from the sample all homicide cases, as well as cases with missing information on any of the following: the sentence assigned, the method of disposition, the main charged offense or the defendant's age, gender or race/ethnicity.

41, which is roughly 3.65 percent of all cases.

Explanatory Variable: Electorate's conservativeness

As we summarize in Table 2, our primary explanatory variable is an approximation for district-level preferences on sentencing. Our variable measures the conservativeness of each district to which a given judge was assigned. Our ideal explanatory variable would directly measure district-level punitiveness among voters—how harsh a judge's constituents would like sentences to be. While scholars have developed national and state level punitiveness measures (e.g., Neill, Yusuf and Morris 2015; Enns 2014), we are unaware of localized measures regarding voter preferences on criminal justice issues. As such, we believe our conservativeness measure sufficiently captures what judges perceive from their constituents.

To measure each district on the liberal-conservative continuum, we utilize the two-party vote in the 2000 U.S. presidential election. Presidential voting outcomes are readily available at the precinct level for the 2000 election, which allows for an almost exact matching with the judicial districts.[8] Using the statewide two-party vote, we measure district conservativeness as the district-level Republican vote share—those votes cast for then-Gov. George W. Bush—in the 2000 election.

There are 66 judicial districts represented by at least one judge. Vote shares for Bush in the 2000 elections lie between 19.28% and 70.17%. The unweighted mean of vote shares for Bush across districts is 54.33%. As a basis for comparison, the statewide vote share for Bush was 56.47%. Thus voting records for the 2000 Presidential election indicate a high variance in the preferences of the electorate across districts. We welcome the high variance in district-level preferences, as it helps to identify whether judges with different electorates react differently to their constituents following the passage of Bill 41.

[8] Voting records at the district level are available online in the website of the North Carolina Board of Elections.

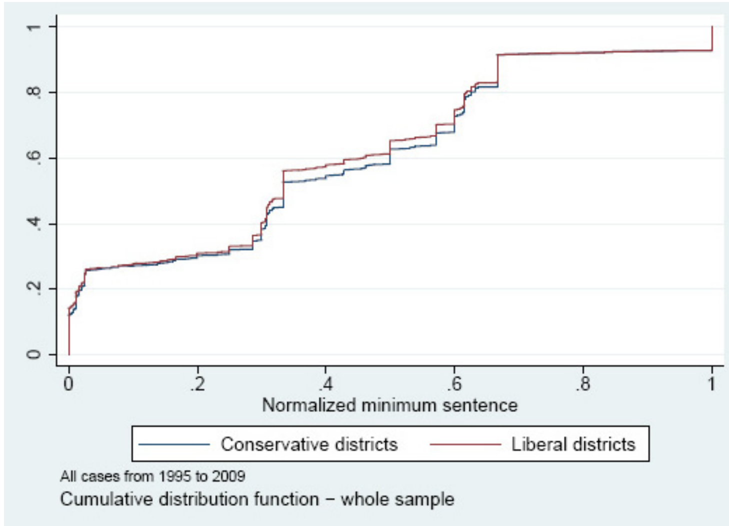
Table 2: Descriptive statistics—District conservativeness

District conservativeness		
Mean	54.33%	
Median	56.40%	
Min	19.28%	
Max	70.17%	
Judges by district conservativeness—full sample		
	Observations	Frequency
Conservative†	74	39.36%
Liberal††	114	60.64%
Total	188	100%
Judges by district conservativeness—judges active in 1996		
	Observations	Frequency
Conservative†	39	41.94%
Liberal††	54	58.06%
Total	93	100%

Notes: This table contains information on our measure of judicial district conservativeness, the Republican vote share in the 2000 Presidential elections. We obtained vote share data from the North Carolina Board of Elections. †. Republican vote share above 56.47% (statewide vote share). † † . Republican vote share below 56.47%.

We define a judge's district as conservative if its Republican vote share in the 2000 Presidential election was greater than 56.47%, the statewide vote share. Otherwise we classify a district as liberal. Table 2 presents the distribution of judges in the sample according to their districts' conservativeness. There are 188 judges in the whole sample. When all such judges are accounted for, those from liberal districts constitute a majority (60.64%). The distribution is very similar if we only consider the 93 judges in activity in 1996, when Bill 41 was approved.

Figure 1: Assigned sentences' length—Judges from liberal and conservative districts



Electorate's preferences and sentencing decisions

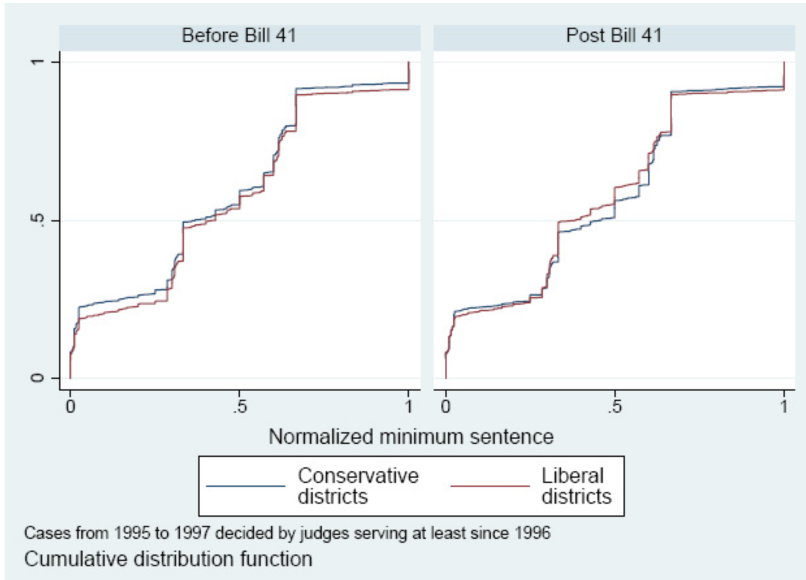
As a first endeavor toward examining our responsiveness hypothesis, we consider several baseline comparisons before and after Bill 41. Figures 1, 2, and 3 illustrate how judges diverge in sentencing behavior after Bill 41. Once judges are assigned to particular districts, we observe sentences that are more tailored to the exhibited policy preferences of those district-level voters.

Cross-sectional comparison

Figure 1 depicts the cumulative distribution functions of the incarceration sentence length assignments by judges from liberal and conservative districts in all years in the sample. The x-axis plots the normalized sentence.[9] The figure reveals that judges from liberal districts tend to assign sentences in the most lenient range of the scale (less than 1/3) more often than their counterparts from conservative districts. The latter group of judges assign rel-

[9] To compute the CDFs, we normalize each sentence, so that it equals zero if it is the lowest possible sentence under the structured sentencing guidelines and one if it is the maximum possible sentence. The normalization also holds for the few sentences assigned outside of the guideline bounds, so that such sentences are censored.

Figure 2: Reaction to Bill 41—Judges from liberal and conservative districts

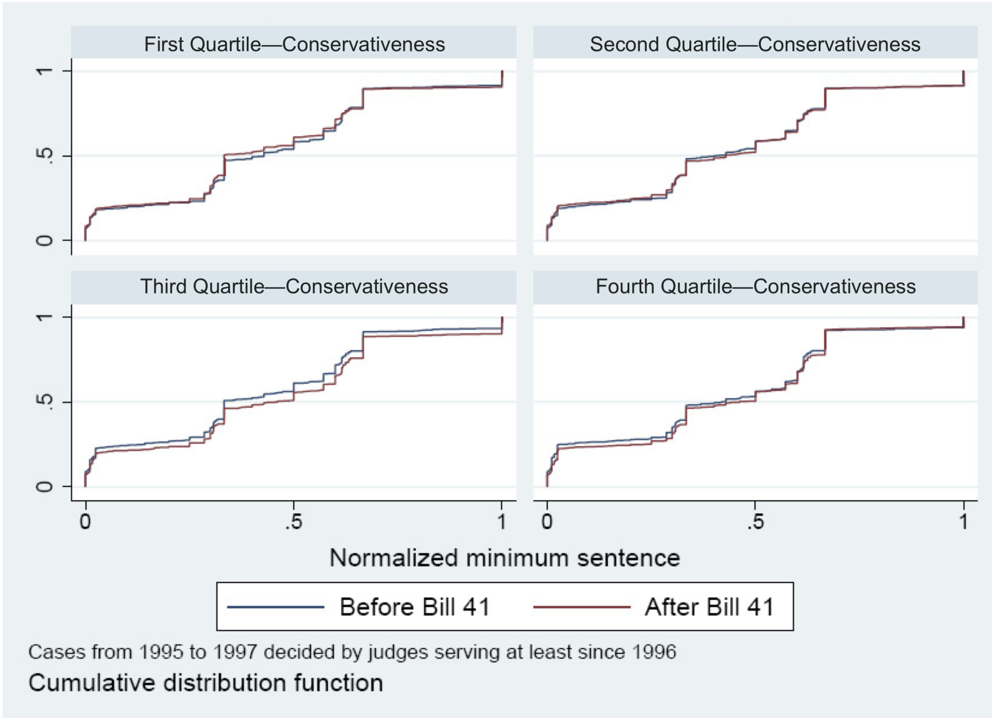


atively more sentences in an intermediate range (from 1/3 to 2/3). Both groups of judges assign sentences in the harshest range (from 2/3 on) with similar frequency. A two-sample Kolmogorov-Smirnov test rejects the null hypothesis of equality of the two CDFs at a confidence level of 1%.

Reaction to Bill 41

Figure 2 shows similar CDFs. This figure only consider judges that were serving at the time Bill 41 was approved. The left-hand plot depicts cases disposed by judges from liberal and conservative districts before Bill 41's passage. The right-hand plot depict cases disposed by the same two groups of judges after Bill 41's passage. The distinction between the two groups of judges is not very clear in the period prior to Bill 41 (left) but becomes more evident after (right). Judges from conservative districts start assigning sentences that are harsher than the ones chosen by judges from liberal districts. It is worth noticing that the CDFs in the right-hand side of Figure 2 are similar to the ones in Figure 1. Figure 2 thus suggests that judges responded to the change in the preferences of their electorate caused by

Figure 3: Reaction to Bill 41—Four groups of judges



Bill 41.

A more detailed examination of the data—however—reveals that the relationship between voters' preferences and judges' behavior is not as simple as the previous paragraphs suggest. Figure 3 separates judicial districts into quartiles, from least to most conservative (i.e., lowest to highest level of support for Bush in the 2000 election). The figure depicts CDFs of assigned incarceration time before and after Bill 41's approval within each quartile revealing an interesting pattern. Judges in the first quartile become more lenient after Bill 41's passage, while judges in the third quartile assign harsher sentences. Judges in the second quartile—who are elected by districts roughly as conservative as the whole state—seem to be less affected by the passage of the bill. All these effects are consistent with our expectations and with our observations in Figure 2. But judges in the fourth quartile, who have the most conservative districts, appear not to change their sentencing behavior at all after Bill 41.

A variety of confounding factors can make our simple analysis of histograms misleading. Taking from these figures some general impression of how judges reacted to Bill 41, our next section investigates the effects of Bill 41 in a more rigorous manner.

Empirical Analysis

To expand on the descriptive evidence we found in the figures above, this section utilizes several difference-in-difference models to examine what changes in judges' case-level sentencing behaviors occur following on Bill 41, which assigned judges to specific districts.

Measuring District Conservatism

As we discussed in our explanatory variable section above, our expectations require us to measure more than the dichotomy between liberal and conservative districts. To evaluate the impact of Bill 41 on sentencing behavior, we begin by defining a measure of the preferences of the judge's post-Bill 41 district electorate by each case i. Again, we utilize an approximation of punitiveness preferences using the district-level 2000 U.S. presidential election Republican vote share. The statewide Republican vote share in those elections was 56.47%. We define the following two variables:

$$\begin{aligned} &\text{Liberal}_i = \max\left\{0.5647 - \text{District Conservativeness}_i, 0\right\} \\ \text{and} \quad &\text{Conservative}_i = \max\left\{\text{District Conservativeness}_i - 0.5647, 0\right\}, \end{aligned} \tag{1}$$

where *District Conservativeness*$_i$ is the Republican vote share in the district of the judge in charge of case i. The variables *Liberal*$_i$ and *Conservative*$_i$ separately capture the distance from the statewide center in the liberal-conservative spectrum for liberal and conservative districts.

Model Specifications

To evaluate the impact of Bill 41 on sentencing behavior, we then estimate the following specification:

$$
\begin{aligned}
\text{Sentence's Length}_i = \; & \alpha + \gamma_C \text{ Conservative}_i * \text{Bill } 41_i + \gamma_L \text{ Liberal}_i * \text{Bill } 41_i \\
& + \delta \text{ Bill } 41_i + \lambda_{\text{judge}_i} + \beta X_i + \epsilon_i. \qquad (2)
\end{aligned}
$$

The dependent variable *sentence* is the log of the assigned sentence length. *Bill* 41 indicates if the disposition of case i took place after the bill's ratification, which we interact with our variables for district-level preferences: $liberal_i$ and $conservative_i$. The variables λ_{judge_i} are judge-specific dummies, which we include in place of a direct measure for judicial preferences. The vector X_i is a series of controls, which include a full set of dummies for the charged offense severity, as defined by North Carolina for the purpose of delimiting judicial discretion under structured sentencing. We regard these severity measures as reasonable approximations for case salience, where higher severity increases the likelihood of heightened media and public attention.[10] Other controls in X_i include several defendant characteristics: ethnicity, gender, previous criminal history, age and age squared. We also include dummies indicating the year of disposition and the county of prosecution of the case. Finally we include a dummy indicating whether the case was resolved by plea bargain. In our regression analysis we only consider cases in which incarceration time was assigned.[11]

The coefficients γ_C and γ_L are the main parameters of interest. A positive value for γ_C and a negative one for γ_L indicate that the sentencing behavior of Superior Court judges tends to correspond to the desires of their voters after the passage of the bill. Therefore, positive (negative) estimates for γ_C (γ_L) are consistent with the hypothesis that judges are

[10] See appendix section B for details on the distribution of offense severity across cases.

[11] We also investigated the relationship between the judges' electorate preferences and the decision of assigning any incarceration time, conditional on a conviction. We found no evidence of such effects.

responsive to the electorate's preferences.

We also estimate the following specification, which allows for non-linearities in the effects of the interactions between the passage of Bill 41 and the variables Conservative and Liberal:

$$
\begin{aligned}
\text{Sentence's Length}_i = \; & \alpha + \gamma_{C,1} \text{ Conservative}_i * \text{Bill } 41_i + \gamma_{C,2} \left[\text{Conservative}_i * \text{Bill } 41_i\right]^2 \\
& + \gamma_{L,1} \text{ Liberal}_i * \text{Bill } 41_i + \gamma_{L,2} \left[\text{Liberal}_i * \text{Bill } 41_i\right]^2 \\
& + \delta \text{ Bill } 41_i + \lambda_{\text{judge}_i} + \beta X_i + \epsilon_i. \qquad (3)
\end{aligned}
$$

The main parameters of interest in specification 3 are $\gamma_{C,1}$, $\gamma_{C,2}$, $\gamma_{L,1}$ and $\gamma_{L,2}$. The specification is flexible enough to allow judges from moderate and extreme districts to react differently to the passage of Bill 41.

Estimating Bill 41's Effect on Sentencing

In Table 3, we present results from OLS specifications 2 and 3, as defined above.[12] Column (1)—from specification 2 above—allows for a straightforward linear relationship between our interaction terms (i.e., county-level ideology $\times$ Bill 41) and criminal sentencing. Column (2) includes linear and quadratic terms for both of our interactions, which allows for a curvilinear relationship across the polynomial in specification 3 above. Column (3)—similar to column (2)—includes only one squared term for *conservative* $*$ *Bill 41*. Finally, columns (4) and (5) use subsetted versions of our data by district ideology; they provide straightforward—if incomplete—tests regarding whether conservative or liberal districts, respectively, lead to more lenient or punitive sentencing.

[12] In all regressions in this section, the reported standard deviations are robust to clustering at the judge level. The estimates reported in Table 3 are based on our entire sample, which comprises the years 1995-2010. Restricting our sample to a period closer to the passage of Bill 41 (for example, the years 1995-2002) does not qualitatively change our findings.

Table 3: Reaction to Bill 41

	(1)	(2)	(3)	(4)	(5)
$conservative * Bill\ 41$	-0.418	2.225**	1.829**	3.682*	-
	(0.292)	(1.122)	(0.876)	(1.986)	
$[conservative * Bill\ 41]^2$	-	-19.699***	-17.442***	-25.886**	-
		(6.781)	(5.572)	(11.637)	
$liberal * Bill\ 41$	-0.406***	0.257	-0.274*	-	-0.263*
	(0.131)	(0.539)	(0.150)		(0.120)
$[liberal * Bill\ 41]^2$	-	-1.620	-	-	-
		(1.391)			
$Bill\ 41$	0.071**	0.028	0.044	-0.072	0.079**
	(0.030)	(0.040)	(0.032)	(0.077)	(0.032)
$settled$	-0.519***	-0.518***	-0.518***	-0.514***	-0.522***
	(0.016)	(0.016)	(0.016)	(0.022)	(0.022)
age	0.013***	0.013***	0.013***	0.015***	0.011***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
age^2	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$female$	-0.172***	-0.172***	-0.172***	-0.160***	-0.184***
	(0.010)	(0.010)	(0.010)	(0.014)	(0.015)
$black$	0.008	0.008	0.008	0.020**	-0.006
	(0.007)	(0.007)	(0.007)	(0.010)	(0.009)
$hispanic$	-0.112***	-0.113***	-0.113***	-0.106***	-0.118***
	(0.016)	(0.016)	(0.016)	(0.025)	(0.020)
$attorney:\ private$	-0.047***	-0.047***	-0.047***	-0.054***	-0.041***
	(0.008)	(0.008)	(0.008)	(0.011)	(0.013)
$attorney:\ public\ defender$	-0.076***	-0.076***	-0.076***	-0.073***	-0.079***
	(0.008)	(0.008)	(0.008)	(0.013)	(0.009)
$convict\ history\ 2$	0.024**	0.024**	0.024**	0.034**	0.017
	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)
$convict\ history\ 3$	0.339***	0.339***	0.339***	0.339***	0.341***
	(0.012)	(0.012)	(0.012)	(0.017)	(0.016)
$convict\ history\ 4$	0.594***	0.594***	0.594***	0.585***	0.607***
	(0.015)	(0.015)	(0.015)	(0.021)	(0.020)
$convict\ history\ 5$	0.809***	0.808***	0.808***	0.826***	0.794***
	(0.017)	(0.017)	(0.017)	(0.022)	(0.024)
$convict\ history\ 6$	1.017***	1.017***	1.017***	1.017***	1.021***
	(0.019)	(0.019)	(0.019)	(0.026)	(0.027)
Observations	135,481	135,481	135,481	65,920	69,561
R^2	0.658	0.658	0.658	0.654	0.663

Notes: OLS estimates. The unit of observation is a case. Standard errors, provided in parentheses, are adjusted for two-way clustering at the judge-period levels, where period refers to pre- and post-Bill 41; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. In columns (1)-(3), we include all non-homicide criminal cases prosecuted from Jan 1995-Oct 2010, except for those cases in which critical information was missing, as explained in Table 1. Columns (4) and (5), includes only cases decided by judges from conservative and liberal districts, respectively. The variables *convict history* 2-6 indicate the defendant's previous number of criminal record points, as employed by the NC structured sentencing rules (2 = 1-4 points; 3 = 5-8 points; 4 = 8-14 points; 5 = 14-18 points; and 6 = 19 or more points). Further controls: judge dummies, county dummies, prosecution year and offense severity.

First we examine the results as they pertain to liberal districts, where we expect the judges assigned after Bill 41 will assign increasingly lenient sentences as the district constituents become more liberal. The linear terms in column (1) provide a straightforward test; given the concomitant effects of (i) the interaction variable *liberal* $*$ *Bill 41* paired with (ii) the term *Bill 41*, we observe a negative effect on sentence length. Therefore, as a North Carolina county is increasingly liberal (i.e. then-Gov. Bush's county-level vote-share is < 0.5647 and decreasing), judges are decreasingly punitive in their criminal sentencing following Bill 41's passage. These results for judges in liberal counties hold across Table 3's columns (1), (3), and (5) (subsetted data). Column (4) does not include data for liberal counties. Column (2) includes a quadratic term *liberal* $*$ *Bill* 41^2, which does not suggest that judicial responsiveness operates differently in districts that are more or less liberal.

Second, we examine whether judges assigned to more conservative districts sentence more punitively following Bill 41. Column (1)'s linear term, *conservative* $*$ *Bill 41*, is not statistically significant.[13] Columns (2), (3), and (4), which include a quadratic term, reveal the non-linear nature of judicial responsiveness in conservative districts; our results are consistent across those columns, the last of which uses subsetted data.

First, column (2)'s linear term *conservative* $*$ *Bill 41* is significant; when combined with the constitutive term *Bill 41*, it reveals that small increases in district-level conservativeness (i.e. counties that are above, but still reasonably proximate to then-Gov. Bush's statewide, county-level mean vote-share, 0.5647) correspond to more punitive sentences. At least 12 of North Carolina's 100 counties were within four percentage points of then-Gov. Bush's statewide vote-share. Second, column (2)'s squared term—*conservative* $*$ *Bill* 41^2—reveals that judges in increasingly conservative counties begin to sentence less punitively than their conservative colleagues in more moderate districts. Across our model specifications, we

[13] It is possible, given the findings with regard to judges assigned to more liberal districts, that, by maintaining prior sentencing practices, judges in conservative districts were nonetheless satisfying their respondents. Further below, we examine this possibility as it pertains to retention of judges in downstream elections.

find mixed evidence regarding our responsiveness expectations. While judges in moderately conservative districts sentence more punitively, we observe the opposite relationship among judges in increasingly conservative districts. We examine further below whether these judges end up (1) retiring from the bench early or (2) being punished at the ballot box due to their lack of responsiveness to district voters.

Regarding our other coefficient estimates, we note that they appear to support the reasonableness of our findings. Settled or bargained cases tend to result in shorter sentences across all columns, as do cases resolved by both public defenders or private attorneys. The coefficients for the defendant's age and age squared are positive and negative, respectively. Both are significant, suggesting that shorter sentences are assigned to very young and very old defendants. The results also indicate that female defendants tend to receive shorter sentences than males. Moreover, Hispanics tend to be assigned shorter sentences than non-Hispanic Whites, whereas the coefficients associated with African-American defendants are not significant in most specifications. Finally, criminal defendants with longer criminal histories are assigned more punitive sentences. Overall, we regard these findings as comporting with our overall understandings of the justice system. Additionally, we note that all of our specifications include judge, county, and year fixed effects. This allows for us to make claims regarding our explanatory variables of theoretical importance while controlling for any judge-level, county-level, or temporal idiosyncrasies.[14]

[14] The analysis in this section abstracts away from two sources of endogeneity problems. First, it only considers cases that resulted in an incarceration conviction. Second, it only addresses the difference between cases that were settled and those that were resolved at trial by incorporating a plea bargain indicator as a control variable in the regression specifications. However, most models of pre-trial negotiations suggest that both the likelihood of a successful plea bargain and that of an incarceration conviction depend on the severity of the trial sentence expected to be assigned by the judge in the event of a conviction at trial. For a detailed review of pre-trial negotiation models, see Daughety and Reinganum (2012). These problems could, in principle, be solved by resorting to instruments for an incarceration conviction and the plea bargain indicator. In practice, however, it is very challenging to obtain these instruments. In Appendix A, we propose a structural approach for dealing with the selection problems, based on the techniques developed by Silveira (2017). The findings from the structural analysis fully corroborate those from the present section.

Punishments for Non-Responsiveness

The results above suggest judges reacted differently to the passage of Bill 41. Judges assigned to more liberal districts exhibit responsiveness in our expect direction—assigning more lenient sentences. In districts that are marginally more conservative than the statewide mean, we observe judges assigning more punitive sentences. The sentencing behavior of judges whose districts are extremely conservative, however, do not fall in line with our expectation of responsiveness. While it is possible that these judges are simply more motivated by their individual preferences than by constituency pressures, it would seem odd that only judges assigned to more conservative districts exhibited sincere behavior.

Nevertheless, we expect that less responsive judges will be more likely to not seek reelection or to lose their re-election. We present evidence supporting this hypothesis in Table 4 below. Specifically we examine the performance of the 93 judges in office during the passage of Bill 41 in subsequent elections between 1996 and 2002.[15] If it is the case that judges from very conservative districts decided not to pander to voters, then the turnover for these judicial seats should be higher than that of their peers from moderate and more liberal districts.

In our analysis below, we set the unit of analysis to be a judge. Let the dummy $early\ exit_j$ indicate whether judge j served until before 2002.[16] The dummy indicates the success of judge j in the first election to take place after the passage of the bill. To identify whether judges from the more extreme districts performed worse than their counterparts from moderate districts in the wake of Bill 41, we consider the following specification:

$$\text{Early Exit}_j = \alpha + \theta_l \text{ Liberal}_j + \theta_c \text{ Conservative}_j + \epsilon_j, \tag{4}$$

[15] Since Superior Court judges serve eight-year terms, every judge serving during the passage of Bill 41 in August 1996 would be up for re-election at some point between the 1996 and the 2002 judicial elections.

[16] We obtained this information in the annual editions of the North Carolina Manual, which contain judicial directories listing every judge in activity for every year between 1996 and 2009.

Table 4: Electoral performance after Bill 41

	Probit
$liberal_j$	0.46
	(1.45)
$conservative_j$	7.08**
	(3.33)
$constant$	-0.40*
	(0.23)
Observations	93

Notes: This table reports OLS estimates. The unit of observation is a judge. Standard errors are provided in parentheses; $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The sample includes all judges active at the moment of the passage of bill 41. The dependent variable, $early\ exit_j$, indicates whether a judge stopped serving before 2002.

where Liberal_j and Conservative_j are defined as in (1). The parameters of interest are θ_l and θ_c. A positive θ_l indicates that, among judges from liberal districts, the turnover following the approval of Bill 41 is higher for judges whose districts are more extreme. Similarly, a positive θ_c indicates that, among judges from conservative districts, judges from extreme districts have higher turnover than those from relatively moderate districts.

Table 4 presents the results of Probit estimation of specification 4 above. The estimated coefficient for judges in liberal districts is not statistically significant. In contrast, the estimate for judges in conservative districts is 7.08 and significant at 5%. This suggests that judges assigned to conservative districts after Bill 41 were more likely to leave office or lose their reelection. These results are consistent with the evidence that judges from very conservative districts were the ones to react the least to the passage of Bill 41. Taken together the results from this section and the previous ones suggest that judges from the most conservative districts did not pander to voters and, as a consequence, were subsequently punished at the ballot box.

Above we suggested several possible reasons why a judge might not exhibit responsiveness. First, some judges may be attentive to different audiences due for career advancement.

Second, judges may prioritize their sincere policy preferences over satisfying voters. Third, a judge assigned to a particular district may realize that re-election is unlikely. Fourth and finally, judges operate in concert with prosecutors—a constraint that may be more severe on some judges in some locations. While we do not have sufficient data to adjudicate among these possibilities, be believe the empirical evidence paired with several plausible explanations provides a compelling story—and one that is likely worth further investigation.

Conclusions

This paper examines how the sentencing behavior of elected trial judges is affected by changes in electoral incentives. With this intent, we explore a unique change in the electoral rules for North Carolina's main trial court. Bill 41, passed in 1996, changed the selection method of Superior Court judges from statewide to district-level elections. We argue that the change in size and scope of the judges' constituencies pressures judges to change their sentencing behaviors. Judges assigned to more liberal districts would sentence more leniently, while those assigned to more conservative districts would sentence more punitively. For judges who do not tailor their criminal sentencing to local preferences, they risk losing their office.

We provide evidence that some judges adapted their sentencing decisions to suit their constituents' preferences. Specifically, judges from liberal districts became relatively more lenient, while those from moderately conservative districts started assigning harsher sentences. However, judges from the most conservative districts did not respond as we has expected to their new local constituents.

Our findings comport—in part—with our theory of responsive judicial behavior. But importantly, we note that further research must be done to examine why certain judges do not alter their official behavior in response to institutional changes. A possible line of inquiry includes whether some judges are ambitious office-seekers, who are less responsive to local preferences and instead cater to a broader audience or constituency—whether it be the statewide electorate, the governor, or certain federal officials. Additionally, while our models include judge-level fixed effects, we importantly note the astute reviewer comments—that it would be interesting to have a direct measure of judicial preferences.

We then explore one possible implication of these main results—namely, whether some judges' lack of responsiveness impacts their chances of re-election. We provide support for this hypothesis by comparing the electoral performance of judges from moderate and extreme districts in the wake of Bill 41. We show that judges from the most conservative districts, which are precisely the ones whose sentencing patterns were not affected by the bill, face lower chances of re-election than their counterparts from liberal and moderately conservative districts.

In this research, we strive to contribute to the literature on electoral connections between judges and voters. While it is well-established that variations in electoral institutions leads to disparate policy outcomes, it is not altogether clear that this fits with the expectations we have for objective or impartial courts. At the very least, this research presents a step forward in understanding the fine-grained nature of how judges' exhibit responsiveness to voters. Furthermore, the enduring judicial reform movement results in regular changes in judicial selection and retention. As such, we expect to observe many more systematic variations in how judges resolve disputes in court.

References

Badas, Alex and Katelyn E Stauffer. 2019. "Voting for women in nonpartisan and partisan elections." *Electoral Studies* 57:245–255.

Barro, Robert. 1973. "The control of politicians: An economic model." *Public Choice* 14:19–42.

Bartels, Larry M. 1991. "Constituency opinion and congressional policy making: The Reagan defense buildup." *American Political Science Review* 85(2):457–474.

Besley, Tim. 2006. *Principled Agents? The Political Economy of Good Government*. Oxford, UK: Oxford University Press.

Besley, Tim and Stephen Coate. 2003. "Elected Versus Appointed Regulators: Theory and Evidence." *Journal of the European Economic Association* 1:1176–1206.

Bonneau, Chris W. 2006. "Vacancies on the Bench: Open-Seat Elections for State Supreme Courts." *Justice System Journal* 27(2):143–159.

Bonneau, Chris W and Damon M Cann. 2015. "Party identification and vote choice in partisan and nonpartisan elections." *Political Behavior* 37(1):43–66.

Bonneau, Daniel and Bryan C McCannon. 2019. "Bargaining in the shadow of the trial? Deaths of law enforcement officials and the plea bargaining process." *Deaths of Law Enforcement Officials and the Plea Bargaining Process (September 21, 2019)* .

Boylan, Richard T. 2012. "The effect of punishment severity on plea bargaining." *The Journal of Law and Economics* 55(3):565–591.

Brace, Paul and Brent D Boyea. 2008. "State public opinion, the death penalty, and the practice of electing judges." *American Journal of Political Science* 52(2):360–372.

Brace, Paul, Melinda Gann Hall and Laura Langer. 1998. "Judicial choices and the politics of abortion: Institutions, context, and the autonomy of courts." *Alb. L. Rev.* 62:1265.

Budziak, Jeffrey. 2013. "Blind justice or blind ambition? The influence of promotion on decision making in the US courts of appeals." *Justice System Journal* 34(3):295–320.

Canes-Wrone, Brandice, Tom S Clark and Amy Semet. 2018. "Judicial Elections, Public Opinion, and Decisions on Lower-Salience Issues." *Journal of Empirical Legal Studies* 15(4):672–707.

Canes-Wrone, Brandice, Tom S Clark and Jason P Kelly. 2014. "Judicial selection and death penalty decisions." *American Political Science Review* 108(1):23–39.

Canes-Wrone, Brandice, Tom S Clark and Jee-Kwang Park. 2012. "Judicial independence and retention elections." *The Journal of Law, Economics, & Organization* 28(2):211–234.

Cann, Damon M and Teena Wilhelm. 2011. "Case visibility and the electoral connection in state supreme courts." *American Politics Research* 39(3):557–581.

Carson, Jamie L, Michael H Crespin, Carrie P Eaves and Emily Wanless. 2011. "Constituency congruency and candidate competition in US house elections." *Legislative Studies Quarterly* 36(3):461–482.

Choi, Stephen J, G Mitu Gulati and Eric A Posner. 2010. "Professionals or politicians: The uncertain empirical case for an elected rather than appointed judiciary." *The Journal of Law, Economics, and Organization* 26(2):290–336.

Daughety, Andrew F. and Jennifer F. Reinganum. 2012. Settlement. In *Encyclopedia of Law and Economics*, ed. Chris W. Sanchirico. second ed. Vol. 8 - Procedural Law and Economics Cheltenham, UK: Edward Elgar Publishing Co.

Elder, Harold W. 1989. "Trials and settlements in the criminal courts: An empirical analysis of dispositions and sentencing." *The Journal of Legal Studies* 18(1):191–208.

Enns, Peter K. 2014. "The public's increasing punitiveness and its influence on mass incarceration in the United States." *American Journal of Political Science* 58(4):857–872.

Fenno, Richard F. 1978. "Home style: Representatives in their districts." *Boston: Little and Brown* .

Ferejohn, John. 1986. "Incumbent performance and electoral control." *Public Choice* 50:5–25.

Gill, Rebecca D and Kate Eugenis. 2019. "Do voters prefer women judges? Deconstructing the competitive advantage in state supreme court elections." *State Politics & Policy Quarterly* 19(4):399–427.

Glazer, Amihai and Marc Robbins. 1985. "Congressional responsiveness to constituency change." *American Journal of Political Science* pp. 259–273.

Goelzhauser, Greg. 2018. "Does Merit Selection Work? Evidence from Commission and Gubernatorial Choices." *Journal of Law and Courts* 6(1):155–187.

Gordon, Sanford C. and Gregory A. Huber. 2007. "The Effect of Electoral Competitiveness on Incumbent Behavior." *Quarterly Journal of Political Science* 2:107–138.

Grimes, Samuel Latham. 1997. "Without Favor, Denial, or Delay: Will North Carolina Finally Adopt the Merit Selection of Judges." *NCL rev.* 76:2266.

Hall, Melinda Gann. 1987. "Constituent influence in state supreme courts: Conceptual notes and a case study." *The Journal of Politics* 49(4):1117–1124.

Hall, Melinda Gann. 1992. "Electoral politics and strategic voting in state supreme courts." *The Journal of Politics* 54(2):427–446.

Hall, Melinda Gann. 2001. "State supreme courts in American democracy: Probing the myths of judicial reform." *American political Science review* 95(2):315–330.

Hall, Melinda Gann and Chris W Bonneau. 2006. "Does quality matter? Challengers in state supreme court elections." *American Journal of Political Science* 50(1):20–33.

Hall, Melinda Gann and Paul Brace. 1999. State supreme courts and their environments: Avenues to general theories of judicial choice. In *Supreme Court decision-making: New institutionalist approaches*, ed. Howard Gillman and Cornell W. Clayton. University of Chicago Press pp. 281–300.

Hanssen, F Andrew. 2004. "Learning about judicial independence: Institutional change in the State courts." *The Journal of Legal Studies* 33(2):431–473.

Helms, H. Park. 1987. "Merit Selection: The Case For Judicial Election Reform." *North Carolina Insight* 9(4):22–27.

Hogan, Robert E. 2004. "Challenger emergence, incumbent success, and electoral accountability in state legislative elections." *The Journal of Politics* 66(4):1283–1303.

Holmes, Lisa M and Jolly A Emrey. 2006. "Court diversification: Staffing the state courts of last resort through interim appointments." *Justice System Journal* 27(1):1–13.

Huber, Gregory A. and Sanford C. Gordon. 2004. "Accountability and Coercion: Is Justice Blind when it Runs for Office?" *American Journal of Political Science* 48:247–263.

Jensen, Jennifer M and Wendy L Martinek. 2009. "The effects of race and gender on the judicial ambitions of state trial court judges." *Political Research Quarterly* 62(2):379–392.

Karunamuni, Rohana J. and Shunpu Zhang. 2008. "Some Improvements on a Boundary Corrected Kernel Density Estimator." *Statistics and Probability Letters* 78:499–507.

Kuziemko, Ilyana. 2006. "Does the threat of the death penalty affect plea bargaining in murder cases? Evidence from New Yorks 1995 reinstatement of capital punishment." *American Law and Economics Review* 8(1):116–142.

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

Lim, Claire S. H. 2013. "Preferences and Incentives of Appointed and Elected Public Officials." *American Economic Review* 103.

Mayhew, David R. 1974. *Congress: The electoral connection*. Yale University Press.

Miller, Warren E and Donald E Stokes. 1963. "Constituency influence in Congress." *The American Political Science Review* 57(1):45–56.

Neill, Katharine A, Juita-Elena Yusuf and John C Morris. 2015. "Explaining dimensions of state-level punitiveness in the United States: The roles of social, economic, and cultural factors." *Criminal Justice Policy Review* 26(8):751–772.

Nelson, Michael J. 2014. "Responsive Justice? Retention Elections, Prosecutors, and Public Opinion." *Journal of Law and Courts* 2(1):117–152.

Rosch, Joel and Eva R. Rubin. 1987. "Merit Selection: The Case Against Judicial Election Reform." *North Carolina Insight* 9(4):28–34.

Silveira, Bernardo S. 2017. "Bargaining with asymmetric information: An empirical study of plea negotiations." *Econometrica* 85(2):419–452.

Silverman, Bernard W. 1986. *Density Estimation for Statistics and Data Analysis*. London, UK: Chapman and Hall.

Streb, Matthew J and Brian Frederick. 2009. "Conditions for competition in low-information judicial elections: The case of intermediate appellate court elections." *Political Research Quarterly* 62(3):523–537.

Taylor, Travis N. N.d. "Judicial Selection and Criminal Punishment: Trial Court Elections, Sentencing, and Incarceration in the States." *Journal of Law and Courts*. Forthcoming.

Williamson, Ryan D. 2019. "Examining the Effects of Partisan Redistricting on Candidate Entry Decisions." *Election Law Journal: Rules, Politics, and Policy* 18(3):214–226.

A Dealing with selection: a structural approach

As explained in the main text, our reduced-form analysis abstracts away from two selection issues: First, it only considers cases that resulted in an incarceration conviction. Second, it only addresses the difference between cases that were settled and those that were resolved at trial by incorporating a plea bargain indicator as a control variable in the regression specifications. However, whether a case settles may depend on the severity of the expected trial sentence, creating a potential endogeneity problem.

In this Appendix, we explore a structural approach for dealing with the selection problems described above. Our strategy is based on the techniques proposed by Silveira (2017). It assumes that, for each case, the data generating process involves a defendant and a prosecutor who bargain over the sentence to be assigned. If bargaining is successful, the case is settled. Otherwise, it proceeds to the trial stage, in which the defendant is found guilty with a given probability. Both the potential trial sentence and the probability of conviction at trial are allowed to vary across cases. The bargaining protocol is take-it-or-leave-it: in every case, one of the bargaining parties offers to settle for a sentence. If the other party rejects the offer, the bargaining stage ends and a trial takes place.

As opposed to Silveira (2017), we do not fully specify a structural econometric model. As we argue in the remainder of the section, it is possible to investigate how the passage of Bill 41 affected the sentencing behavior of Superior Court judges by only partially recovering the primitives of the model. As a consequence, we do not need to take a stance on the identity of the proposal-maker at the bargaining stage or on the nature of the asymmetric information in the model. Accordingly, several of the assumptions that we make below are in terms of high-level objects. The structural econometric model presented in Silveira (2017), as well as the assumptions on its primitives discussed there, are fully consistent with the analysis developed here.

Assume that each case i is associated with characteristics Z_i, which are observable by the econometrician. These characteristics may, in principle, include the type of the main charge against the defendant, the defendant's demographics, variables related to the date and place of prosecution, etc. In our application below, Z_i consists of measures of the conservativeness of the judge responsible for the case and an indicator of whether the case was resolved after the passage of Bill 41. To every case correspond a trial sentence T_i. Such a sentence is assigned by the judge in the event that the case reaches trial and results in a conviction. We assume that the trial sentence is known by the prosecutor and the defendant at the plea bargaining stage but, from the econometrician's perspective, it is a random variable with a mixture distribution. Specifically, it may assume value zero with positive probability, which we allow to depend on Z_i. The interpretation for T_i is that, in the event of a conviction at trial, a non-incarceration sentence (such as a sentence to probation or community service) is assigned by the judge. Conditional on being strictly positive, T_i is described by the density function $g(\cdot|Z_i)$ with full support over $[\underline{t}, \bar{t}]$.

In every case i a settlement offer is made. Such an offer is represented by S_i, which, for the econometrician, is a random variable. Assume that there exists a strictly increasing continuous function $\tilde{s}(\cdot, \cdot)$ such that $S_i = \tilde{s}(T_i, Z_i)$. In other words, given realizations z_i and t_i of Z_i and T_i, respectively, the realization s_i of S_i satisfies $s_i = \tilde{s}(t_i, z_i)$. Assume that $\tilde{s}(0, Z_i) = 0$ for every Z_i. Under these assumptions, S_i is equal to zero with positive probability (the same probability that T_i is equal to zero) and, conditional on being strictly positive, S_i is described by a density function $b(\cdot|Z_i)$ with full support over $[\tilde{s}(\underline{t}, z), \tilde{s}(\bar{t}, z)]$.

For each case i the trial sentence T_i is only observed in the event of a conviction at trial. Similarly, the settlement offer S_i is only observed if it is accepted—i.e., the plea bargain is successful. Let Ψ_i denote a random variable indicating the way case i is resolved. Assume $\Psi_i = 1$ when the case is settled and $\Psi_i = 2$ if it results in a conviction at trial. We can then write the density of trial sentences, conditional on a conviction at trial, as

$$g(t|\Psi = 2, Z = z) = \frac{P[\Psi = 2|T = t, Z = z]g(t|Z = z)}{P[\Psi = 2|Z = z]}. \tag{5}$$

Also, we can write the density of settlement offers, conditional on a plea bargain, as

$$b(s|\Psi = 1, Z = z) = \frac{P[\Psi = 1|S = s, Z = z]b(s|Z = z)}{P[\Psi = 1|Z = z]}. \tag{6}$$

Both $g(\cdot|\Psi = 2, Z = z)$ and $b(\cdot|\Psi = 1, Z = z)$ are observed by the econometrician. Moreover the conditional probabilities $P[\Psi = 2|Z = z]$ and $P[\Psi = 1|Z = z]$ are also observed.

Let z' and z'' be two values of Z_i such that

$$\begin{aligned} & P[\Psi = 2|T = t, Z = z'] = P[\Psi = 2|T = t, Z = z''] \\ \text{and} \quad & P[\Psi = 1|S = s, Z = z'] = P[\Psi = 1|S = s, Z = z''] \end{aligned} \tag{7}$$

for all $t \in [\underline{t}, \bar{t}]$ and $s \in [\tilde{s}(\underline{t}, z), \tilde{s}(\bar{t}, z)]$. Then, the following equations hold:

$$\begin{aligned} & \frac{g(t|\Psi = 2, Z = z')}{g(t|\Psi = 2, Z = z'')} \frac{P[\Psi = 2|Z = z']}{P[\Psi = 2|Z = z'']} = \frac{g(t|Z = z')}{g(t|Z = z'')} \\ \text{and} \quad & \frac{b(s|\Psi = 1, Z = z')}{b(s|\Psi = 1, Z = z'')} \frac{P[\Psi = 1|Z = z']}{P[\Psi = 1|Z = z'']} = \frac{b(s|Z = z')}{b(s|Z = z'')}. \end{aligned} \tag{8}$$

The first equation in 8 shows that, using the observed unconditional probabilities of conviction at trial and the densities of trial sentences, conditional on a conviction at trial, we can recover some information on the unconditional distribution of trial sentences.[17] More precisely, we are able to identify the ratio of unconditional densities of trial sentences for z' and z''. Similarly, the second equation of 8 shows that, using the unconditional probabilities

[17] To be sure, when we refer to unconditional probabilities and densities, we mean to condition these objects on Z.

of successful plea bargain and the conditional densities of settlement offers, we can recover the ratio of densities of settlement offers for z' and z''.

The equations in 8 suggest a method for analyzing how the passage of Bill 41 affected the sentencing behavior of Superior Court judges in a way that accounts for the selection processes determining which cases are settled and which ones result in a conviction at trial. The vector z' can be set to indicate cases that were under the responsibility of a particular judge or group of judges and that were decided before the passage of the bill. Vector z'' can then be set to indicate cases under the same judges that were decided after the bill was approved. The conditional probabilities and densities on the left-hand side of the equations are observed and can be estimated. By examining the ratios of estimates of the conditional densities of trial sentences for z' and z'' (weighted by estimates of the ratios of unconditional probabilities of conviction at trial) we are able to estimate how the passage of Bill 41 affected the unconditional distribution of trial sentences. In the same way, we can estimate how the bill affected the unconditional distribution of settlement offers by recurring to an empirical version of the second equation in 8.[18] Since the trial sentence and the settlement offer of a case are related by the strictly monotonic function $\tilde{s}(\cdot, \cdot)$, the way Bill 41 affected the unconditional distribution of settlement offers is very informative of how it affected the sentencing behavior of judges.

The estimation of the conditional densities of settlement offers and trial sentences can, in principle, be done by kernel. But the supports of the distributions of trial sentences and settlement offers are bounded. Standard kernel density estimation techniques result in inconsistent estimates near the boundaries of the support. To overcome such a problem we employ a boundary correction strategy proposed by Karunamuni and Zhang (2008).[19]

[18] Of course simply employing the weighted ratios of conditional densities only identifies the ratios of unconditional densities of sentences and settlement offers. To independently obtain the unconditional distributions of trial sentences and settlement offers, a full structural estimation procedure, as the one implemented by Silveira (2017), is needed.

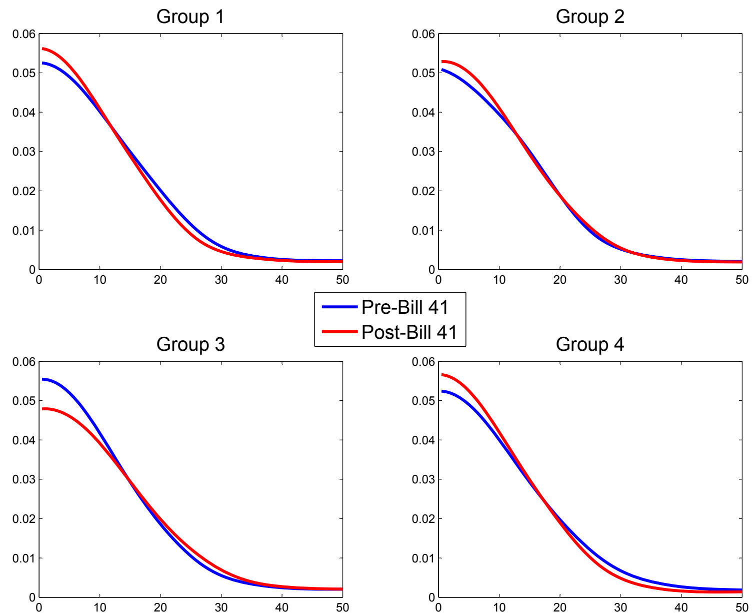
[19] Away from the boundaries, bandwidth selection follows Silverman's "rule-of-thumb" (Silverman 1986).

Since the vast majority of the cases in the sample are resolved by plea bargain, we focus our analysis on estimating the ratio of densities of accepted settlement offers before and after the passage of Bill 41. Due to the sample size, these densities are much more precisely estimated than those of trial sentence and, as argued above, they are still informative on the sentencing behavior of the Superior Court judges.

Figure 4 depicts the ratios of estimated densities of settlement offers before and after the passage of Bill 41 for four groups of judges. As described earlier, the judges are divided according to the conservativeness of their districts. The groups are numbered from the least to the most conservative, as implied by the Republican vote share in the 2000 Presidential elections. The plots on the top left and top right, which respectively refer to groups one and two, indicate that cases under judges from liberal districts started being settled for more lenient sentences after the passage of the bill. For settlement offers shorter than about ten months, the densities referring to the period after the passage of the bill are above those referring to the period prior to the bill. For settlement offers between ten and about 40 months, the order of the density functions switches. In contrast, the plot on the left bottom of figure 4 indicates that cases under judges from group three, whose districts are moderately conservative, settled for harsher sentences after Bill 41 passed. For settlement offers shorter than about 15 months, the density referring to the period after the bill is below that referring to the post-bill period. The order of the density functions is the inverse for settlement offers between 15 and about 40 months. The plot on the bottom left of the figure depicts cases under judges from group four, whose districts are the most conservative ones. This plot is very similar to the plots on the top of the figure, which refer to judges from the liberal districts. It indicates that cases under judges from extremely conservative districts started being settled for more lenient sentences after the passage of Bill 41.

Near the boundaries, a modified bandwidth is employed. The kernel function is the tri-cube. See Karunamuni and Zhang (2008) for details.

Figure 4: Reaction to Bill 41—Four groups of judges



A simple way of assessing the statistical significance of these results is to recur to two-sample Kolmogorov-Smirnov tests comparing the distribution of settlement offers made before and after the passage of Bill 41 for every group of judges. Table 5 presents the results of such tests. The null hypothesis is that the sample of pre-Bill 41 cases is drawn from the same distribution as the that of post-Bill 41 cases. The null is strongly rejected for groups one, two and four. For group three, the null is rejected at a confidence level of 5.57%.

The plots in figure 4 confirm the findings from the reduced form analysis. They suggest that, after the passage of Bill 41, judges from liberal and moderately conservative districts adapted their sentencing behavior to suit the preferences of their new electorate. The shock in sentencing patterns resulted in a change in the distribution of settlement offers, as captured by figure 4. Judges from the most conservative districts, however, did not alter their sentencing behavior in a way that was consistent with their voters' will. Instead,

Table 5: Kolmogorov-Smirnov tests: results per district group

	Group 1	Group 2	Group 3	Group 4
p-value	0.01%	0.00%	3.97%	0.02%

Notes: This table reports the results of Kolmogorov-Smirnof tests of equality of the settlement sentence distributions before and after the passage of Bill 41. The null hypothesis is that the two distributions are the same. The tests are conducted separately for cases decided by judges from district groups one (most liberal) to four (most conservative).

figure 4 suggests that these judges become more lenient after the bill passed. These results, therefore, corroborate the findings from the regression analysis presented in the main text.

B Accounting for Case Severity

Offense Severity in North Carolina Superior Court: 1994-2010

(1) Class	(2) Count	(3) Share	(4) Cumulative
Misdemeanor			
1	70,927	9.19	9.19
2	32,205	4.17	13.36
3	9,882	1.28	14.64
Felony			
A	2,619	0.34	14.98
A1	16,230	2.10	17.08
B1	8,715	1.13	18.21
B2	1,394	0.18	18.39
C	22,995	2.98	21.37
D	33,358	4.32	25.69
E	21,885	2.83	28.52
F	31,807	4.12	32.65
G	39,158	5.07	37.72
H	288,959	37.43	75.15
I	191,857	24.85	100.00