UNIVERSITY of York

Discussion Papers in Economics

No. 22/04

Lady Justice:
The impact of female judges on trials' verdicts in US

Alessandra Foresta

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Lady Justice: The impact of female judges on trials' verdicts in US

Foresta Alessandra*†[1]

[1]University of York

September 12, 2022

Abstract

This work evaluates the role of judges' gender on jury trials verdicts in the US state of North Carolina. My identification strategy is based on judges' rotation across districts and fixed effects. The results indicate that, in trials presided by female judges, juries are more likely to express guilty verdicts. I implement a series of robustness checks (different models' specifications, defendants' characteristics, district sizes, judges' types, judges' experience and workloads) and a series of heterogeneity checks (judges' characteristics, types of crimes and jurors' gender). Finally, I discuss the possible mechanisms behind these findings and I explore the impact of the jury selection process, the role of judges' toughness and the attitudes of women towards courts and sentencing.

JEL Codes: K10, K40, J16
Keywords: Gender, Judge, Trials behaviours

*alessandra.foresta@york.ac.uk

†I thank the North Carolina Sunshine Project, Ron Wright, Francis Flanagan, Shea R. Denning for the help with the dataset and the information over the North Carolina Judicial system, Andrew Pickering, Cheti Nicoletti, Emma Tomminey, Neil Cummins, the participants of the AME cluster seminar in York, the participants of the SES2022 and the participants of the EPCS2022 for the many useful suggestions.

1 Introduction

In many countries, the concepts of *Justice* is represented by *Lady Justice*, a blindfold woman holding a scale and a sword. The sword symbolizes authority and the scale measures the support and opposition to the cases. The blindfold represents impartiality, the principle according to which justice should be applied without regard to power, wealth or status. Impartiality is a key element in many justice systems and, in many cases, judges are described as its "embodiment". However, theory and practice are not always the same and judges are not always as neutral as they are supposed to be. For example, works like Spitzer and Talley (2013) and Johnson (2014) indicate that certain judges' characteristics, such as political affiliation or gender, might influence the sentences duration in bench trials.[1] Moreover, in jury trials, Lenehan and O'Neill (1981) and Halverson et al. (1997) find that judges non-verbal behaviours can sway juries in one direction or in the other.

In this work, I explore the role of judges' gender on jury verdicts. The previous literature is mainly focused on bench trials and suggests that judges' gender can influence the sentence (e.g. Spohn (1991), Steffensmeier and Hebert (1999) and Johnson (2014)) or their behaviours during the trials (e.g. Fox and Van Sickel (2000)). This article expands the previous literature and investigates the effect of judges' gender on jury trial outcomes. My empirical analysis uses data on felony trials in North Carolina between 2010 and 2012. To ensure the causality of my findings, I exploit the judges' rotation across different districts. The findings suggest that female judges increase the probability of guilty verdicts in juries' trials. Additionally, I implement a series of robustness checks based on different models' specifications, defendants' characteristics, district sizes, judges' types, judges' experience and workloads. The results are consistent with the main outcomes. I also investigate the sensitivity of my findings to judges' characteristics, types of crimes and jurors' gender through a series of heterogeneity checks. I find that a relevant role is played by female white judges, democratic judges and the proportion of women in jury pools.

Finally, I investigate the possible mechanisms behind my findings. While the jury selection process does not seem to have a (statistically significant) impact, there is a (statistically significant) effect of differences in toughness between female and male judges and different attitudes of women and men towards crimes and sentencing.

This article relies on the previous literature in many ways. First, it is connected with the vast

[1] Bench trials are trials by judge, where the judges play the role of finder of facts in addition to sentencing. The type of trials examines in this paper are jury trials, where the jury is the finder of facts and the judge presides the trial and emits the sentence after the jury reach a verdict. For more details, see Section 2.

gender literature, especially with the works exploring gender differences in the decision-making process. For example, Gilligan (1982) suggests that men and women make decision through different criteria. According to the author, women tend to make decisions through an "ethic of care", based on ideals such as "correctness", responsibility and attention to the context, while men tend to make a decision through an "ethics of rights", based on abstracts rules and principles, such as absolute individual liberty and hierarchy. These gender differences can lead to variations in judges' decision making and behaviours, as proposed by Fox and Van Sickel (2000). The authors indicate that female judges are more likely to side with the prosecution, while male judges are more likely to side with the defence. Moreover, female judges are more likely to adopt an inclusive and procedural sentencing style. On the contrary, male judges are more likely to use the consensual and authoritarian sentencing style.

Second, this article is linked to the literature investigating the relationship between judges' characteristics and trials outcomes. The majority of previous works is focused on bench trials and sentencing, while jury trials receive less attention. According to the previous literature, the effect of judges' gender is not as clear as the effect of other judges' characteristics, such as political affiliation (e.g. Schanzenbach and Tiller (2008), Fischman and Schanzenbach (2011) and Spitzer and Talley (2013)). On one hand, Johnson (2014) finds that female judges are more lenient. On the other hand, Steffensmeier and Hebert (1999) and Spohn (1991) indicate that women impose harsher sentences and Songer et al. (1994) suggest that there are some differences in sentencing behaviours between male and female judges only in employment discrimination cases.

Third, this study also relies on the previous works on judges' influences on juries. As suggested in Section 1, judges are supposed to be "neutral" actor during the trials and do not influence the jury with their opinions and expectations. However, the literature suggests that juries know exactly what judges think about the case (e.g Hart (1995) and Hart (1992)), with an important played by judges non-verbal behaviours (e.g. Ekman and Friesen (1969)). The authors investigate the role of judges' non-verbal behaviours and their impact on the juries and suggest that even the most restraint and disciplined judge might transmit biased information through non-verbal behaviours. The role of judges' non-verbal behaviours is extremely relevant and there are multiple cases in which courts rule that defendants' rights have been violated by judges' non-verbal behaviours (e.g. Blanck et al. (1985)). However, to the best of knowledge, there is nothing about the effect of judges' characteristics on jury trials outcomes. This article tries to (partially) fill the gap and investigates the role of judges' gender in the relationship between judges and juries.

The remainder of the paper is organised as follows: in Section 2 I focus on the legal system in

North Carolina, in Section 3 on the dataset and in Section 4 on the empirical strategy. In Section 5, I present the main results, in Section 5.2 and Section 5.3 the robustness and heterogeneity checks. Finally, in Section 6 I speculate over the possible explanations behind the results and, in Section 7, I list my conclusions.

2 North Carolina Legal system and Judge

Criminal trials in the US legal system involve different "actors": the defendant, the defence attorney, the prosecutor, a (seated) jury and a judge. The defendant is the person charged with the offence/offences while the defence attorney is the lawyer whose job is to counsel and represent the defendant. Prosecutors represent the state and they present the case against the defendant. They also have significant influence in deciding if and when the trials take place, as suggested by Bandyopadhyay and McCannon (2014). Overall, prosecutors in the US are called *Assistant District Attorneys*,[2] and the chief of the prosecution office, called *District Attorney*, is an elected officials with a four year term.[3]

The most classical element of Anglo-American legal system is the jury, whose main tasks are to assess the facts and to reach a verdict. The jury that presides the trial is called *seated jury*, while the *jury pool* is the group of potential jurors.[4] The seated jury in North Carolina is composed by 12 jurors plus alternatives.[5] The process that "transforms" jury pools into seated juries is called jury selection process. Through this process, judges, defence attorneys and prosecutors remove biased individuals from the jury pool and select the seated jury (more details in Section 6.1).

The person with the greatest power in the courtroom is the judge. In jury trials, judges are tasked with presiding over trials and with issuing a sentence after the jury reach a verdict.[6] They ensure that laws and procedures are always respected during the trials. For example, they review whether there are any illegality issues in the submitted evidences and they are responsible to provide jury instructions.[7] Finally, judges are supposed to behave impartially during trials, regardless of their personal opinions and expectations, backgrounds and characteristics. The trials in my sample are presided by a specific group of state judges, North Carolina's Superior Court

[2] This is not true in all states but it holds for North Carolina.

[3] North Carolina is divided in 48 prosecutorial districts in the analysed period.

[4] According to North Carolina legislation, a qualified potential juror must be; U.S. citizen, a resident of the county where the summons was issued, at least 18 years old, able to understand English and physically/mentally competent. Potential jurors must not: have served as a juror during the previous two years; have served a full term as a grand juror in the last six years and have been convicted of a felony (unless citizenship rights have been restored). Source: North Carolina Jury Service.

[5] Alternatives are those jurors who replace a seated juror in case she/he can no longer continue to serve in this role. The number of alternatives per trial is decided by the judge.

[6] For a more detail exploration of the different types of trials, see note 1.

[7] Jury instructions are the guidelines provided by the judge to a jury at the end of the evidence presentation to explain to the jury what the applicable laws are.

Judges. These judges are elected officials and serve a eight year-term. A small minority of judges are appointed by the Governor and they serve five year-terms.[8] All judges must be attorneys but they are prohibited to practice the law privately during their tenure on the bench. There are no terms limits but judges need to be under 72 years. In North Carolina, Superior Court Judges rotate across districts every six months (more details in Section 4).[9]

3 Data

The dataset consists of felony trials in North Carolina between 2010 and 2012.[10] These data are collected by the research team of North Carolina Jury Sunshine Project.[11] The dataset includes judges' information such as name, ethnicity, gender, political affiliation, elections[12] and experience.[13] Moreover, it also includes trials information such as the list of charges, verdict, sentence, defendant's characteristics (age, ethnicity[14] and gender) and characteristics of potentials and seated jurors (political affiliation, gender and ethnicity).[15] The sample in this article is limited to non-capital felony trials,[16] with only one defendant and one charge, with a known verdict of guilty/non-guilty[17] and the main variables are presented in Table 1.

[TABLE 1 APPROXIMATELY HERE]

Table 1 presents the summary statistics of my sample. Overall, 11 % of the defendants is female, 63 % is black and 34 % is white. The majority of defendants, 65 %, is over 30. Following Anwar et al. (2012, 2014) and Flanagan (2018), I generate a set of dummies for the different charges. Overall, the most common offences are property offences (30%), drug offences (20%) and other violent offences (14%). On average, juries pools[18] are (more or less) gender balanced and show a higher percentage of people identifies as white (61 %) and only 17 % that identifies as black. On

[8] The account for around 14% of the sample. In my main analysis I insert a control for elected judges (see Section 5) and in Section 5.2 I implement a robustness check by removing non-elected judges. There seems to be no effects.

[9] Source: North Carolina Judiciary branch site.

[10] There are 8 trials in the data that are set in 2009 or 2013 due to delays in the court system and lags between the jury selections and the conclusion of the trials. Following Flanagan (2018), I include them in the analysis.

[11] Sunshine Project Website.

[12] In some cases, elections data are missing. I used Ballotpedia to fill the gaps.

[13] There are some unknown judges. By looking at trials date and court calendars, I found part of the missing judges. Those judges that I was not able to identify were drop from the sample (22 obs.).

[14] Given the limited number of observations in some categories, I re-frame ethnicity as white, black, unknown and other, which includes the categories Asian, Native American or indigenous, Hispanic and other.

[15] For a more detailed description of the data collection, see Flanagan (2018).

[16] I exclude those murder charges that are classified as A1 felonies in North Carolina (33 obs.).

[17] Given the structure of judicial courts in North Carolina, I also remove trials with less than 12 jurors in the pool and in the seated jury, trials with more than 20 seated jurors and with more than 60 jurors in the jury pool (37 obs.). I also remove trials with unknown judges' gender (17 obs.), mistrials (57 obs.) and trials with defendant of unknown gender and/or ethnicity (18 obs.).

[18] I use as controls the jury pools characteristics, because, as mentioned by Anwar et al. (2012), there could be some indirect effect from the potential jurors, due to pretrial interactions, that can alter the seated jurors' attitudes.

average, Republican and Democratic jurors represent the 26 % and 32 % of jury pools. 72 % of the defendants are convicted.

Trials presided by a female judges count for 8 % of my sample, by a non-white judge 15 %. Democrats judges preside 62 % of trials while Republicans judges only 27 %. *Experience (per 100)* is the number of years passed from the "*Juris Doctor*" degree, which is a graduate-entry professional degree in law. I use this measure as a proxy of experience in law.[19] On average, judges have 30 years of experience and around 3.55 trials. Given the possible relevance of these characteristics on my main outcomes, I investigate if these variables are, on average, statistically different between judges' gender. In Table 2, I report the p-values for the t-test for different judges characteristics between female and male judges.

[TABLE 2 APPROXIMATELY HERE]

The (preliminary) findings provided by Table 2 indicate that there are some statistically significant differences between female and male judges. Overall, women tend to have less experience, be less white, have less cases and to be up for election next year. Given these systematic and significant differences between female and male judges and their possible impact on the main outcomes, I include these variables in my main regressions (see Eq. 2).

4 Research Design

In this Section, I present my empirical strategy, which relies on judges' rotation and fixed effects. Judges' rotation[20] is mandate by North Carolina Constitution (Article 4, Section 11).[21] In more details, North Carolina is divided in 100 counties, 50 districts[22] and 8 divisions, as shown in Fig 1.

[FIGURE 1 APPROXIMATELY HERE]

Superior Court Judges are assigned to one of the 50 districts and every six months they rotate across different districts within the same division (N.C. Gen. Stat. sec. 7). For example, let's assume that judge A is assigned to Robson county, which coincides to district 16B in the fourth division. After six months, judge A leaves district 16B and goes to district 16A, which is still part of the forth division and is composed by two counties, Scotland and Hoke. The main purpose

[19] To be fully authorize to practice law, in North Carolina individuals must pass also a bar examination. However, this information is not present in my database.

[20] The rotation has been suspended due to budgetary constraints on at least three occasions (1990, 2002 and 2009), but these years are not included in my sample. Source: Article.

[21] The Article specifically states: "[t]he principle of rotating Superior Court judges among various districts of a division is a salutary one and shall be observed.".

[22] A reform that took place in 2016 reduced the number of districts from 50 to 48. All the trials presented in my sample took place before this reform.

behind this process is to avoid conflict of interest and judges' corruption. Judges' calendars are generally decided in advance and they are available on-line for consultation.[23] This mechanism ensures that judges cannot choose the trials they are presiding.

Moreover, my identification also includes a series of fixed effects, in line Flanagan (2018). Fist, my research design includes year and district fixed effects, to account for elements such as judges' availability and crimes and economic trends. Second, I include prosecutor fixed effects. The previous literature mentions the (strong) influence of prosecutors on trials' schedule. According to Bandyopadhyay and McCannon (2014), prosecutors have significant power in allocating trials and, as a consequence, they could influence the choice of judge. For example, prosecutors could use their knowledge of trials cases and judges' leanings to "manipulate" the calendar and allocate certain trials to a specific judge. Judges personal preferences could be connected with judges' gender and prosecutors' strategic allocation could generate an omitted variable bias. To ensure the causal interpretation of my results, I include prosecutors fixed effects in my regressions.[24]

Finally, following the previous literature (e.g. Anwar et al. (2012, 2014) and Flanagan (2018)), I test my identification strategy by regressing judges' characteristics on trials' observable characteristics, as shown in Eq. 1.

$$FemaleJudge_i = \alpha + \beta_1 Trial_i + \beta_2 Def_i + \beta_3 Jury_i + DistFE_i + ProsFE_i + YearFE_i + \epsilon_i \quad (1)$$

where $FemaleJudge_i$ is a dummy for judges' gender, $Trial_i$ contains the trials characteristics, Def_i the defendant characteristics and $Jury_i$ the jury characteristics. $DistFE_i$, $ProsFE_i$ and $YearFE_i$ represent the district, prosecutor and year fixed effects. If my research design was truly randomly working, then there should be low correlation between the characteristics in Eq. 1. In Table 3, I present the results with and without fixed effects in Columns (1) and (2), respectively. If judges' rotation is sufficient to ensure the randomization of judges' gender, the coefficients in Column (1) would be non-significant. However, if the rotation is not enough and fixed effects are required, some coefficients in Column (1) would be statistically significant, while the statistically significance should disappear in Column (2).

[TABLE 3 APPROXIMATELY HERE]

The findings in Table 3 are quite promising. In Column (1) there is only one (slightly) significant coefficient and the significance disappears when fixed effects are included (Column (2)). Although the results of Table 3 cannot rule out that the judges' selection is related to characteristics that I

[23] Source: North Carolina Courts Website.

[24] I also replicate my results without prosecutors fixed effects with similar findings. Available upon request.

cannot observe, they suggest that this should not be a major concern.

5 Results

5.1 Main results

In this Section, I evaluate the causal impact of female judges on juries' trials outcomes using a linear probability model. As mentioned in Sections 3 and 4, I include in my regressions districts fixed effects, prosecutors fixed effects and year effects as well as a series of controls variables, as presented in Eq. 2.

$$Guilty_i = \alpha + \beta_1 FemaleJudge_i + \beta_2 Controls_i + DistrictsFE_i + ProsFE_i + YearFE_i + \epsilon_i \quad (2)$$

where $Guilty_i$ is the dependent variable, a dummy for guilty/non-guilty, $FemaleJudge_i$ is a dummy for judges' gender. $Controls_i$ is the set of controls based on jury pool,[25] trials, judges and defendant characteristics. $DistrictsFE_i$, $ProsFE_i$ and $YearFE_i$ represent the fixed effects for districts, prosecutors and year. Finally, trials are not evenly distributed across judges with some judges having very few observations. To avoid possible over-rejections (e.g. Flanagan (2018), Cameron, Gelbach, et al. (2008), Carter et al. (2017) and Cameron and Miller (2015)), I use heteroskedasticity-robust standard errors rather that clustered standard errors at judge levels.[26] The findings for Eq. 2 are presented in Table 4.[27]

[TABLE 4 APPROXIMATELY HERE]

In Table 4, the coefficients in Columns (1)-(4) for female judges are always positive and statistically significant. The presence of female judges systematically increases the probability of guilty outcomes, even when fixed effects and controls are included. Moreover, the other judges characteristics do not seem to have a statistically significant effect on the dependent variables, with the only exception of judges' experience in Column (4).[28] In terms of magnitude, having a female judge increases of 13.9 percentage points in the probability of a guilty verdict in Column (1) and of 26.9 in my benchmark model of Column (4). The magnitude of my findings is in line with the previous literature. For example, Anwar et al. (2012) find that there is a 16-percentage point conviction gap between all white juries and juries where there is at least one black member in the jury pool.

[25] I also implement the analysis using the characteristics in the seated juries rather than the jury pools. The findings are similar to these presented in Table 4. Available upon request.

[26] I also implement the analysis using cluster standard errors at judge level. The results are similar to those in Table 4. Available upon request.

[27] In the Appendix A1, I present the findings including also the coefficients for all controls variables.

[28] I also implement the analysis using a set of dummies about experience of the judge. The findings are non-significant. Available upon request.

To assess whether my estimated gender effect is statistically significant as a result of pure chance, I implement a permutation test, following the previous literature (e.g. Nagler et al. (2020), Bertrand et al. (2004) and Fujiwara and Wantchekon (2013)). This test randomly reassigns judges' gender (my treatment) in the sample and re-estimates β using this placebo assignment multiple times (1,000 in my case).[29] The randomization inference test for my benchmark model (Column (4) of Table 4) indicates that my estimated coefficient is statistically significant at 1% and larger in magnitude than almost all simulated effects, as shown in Figure 2.[30]

[FIGURE 2 APPROXIMATELY HERE]

5.2 Robustness checks

In this Section, I implement a series of checks to evaluate the reliability of the results presented in Table 4. Specifically, in Table 5 I control the robustness of my findings to changes in my model or to the removal of (possible) outliers' observations from the sample.

[TABLE 5 APPROXIMATELY HERE]

Column (1) of Table 5 is my benchmark specification (Column (4) of Table 4). Column (2) of Table 5 reports the margin estimated by a Logit model. Given that my dependent variable is a dummy, a linear probability model is not always the perfect choice. To control that my findings are not driven by choice of model, I replicate the benchmark regression using a Logit model. This specification has a significant drop in observations, but the finding remains positive and significant, in line with the results in Table 4.[31]

In Columns (3) to (10) of Table 5 I evaluate the robustness of my findings to the exclusion of some outliers' observations. In the first groups of checks, I remove female defendants (Column (3)) and appointed judges (Column (4)), two categories that have a very low number of observations. In order to avoid possible small sample issues, I remove these groups rather than interact them with the right side of Eq. 2. The coefficient for female judges in Column (3) is positive and significant and suggests that my findings in Table 4 are not driven by the small group of female defendants.[32] Similarly, the coefficient for female judges in Column (4) is still positive and significant even after removing non-elected judges.[33]

[29] To implement this analysis, I use the randomization inference test (*ritest*) proposed by Heß (2017).

[30] I replicate this test also for the other specifications presented in Table 4 with similar results. Graphs available upon request.

[31] Although the findings for the Logit model are larger, once I replicate the linear probability model in the Logit sub-sample, I obtain a similar magnitude. Findings available upon request.

[32] As shown in Table 1, female defendants count only for the 11% of my sample and the interaction between female defendants and female judges include only 5 obs.

[33] As shown Table 1, non-elected judges are only 14% and the interaction between appointed judges and female judge includes only 8 obs.

In the second group of robustness checks, I control that my findings are not driven by outliers' observations in districts dimensions, judges' workloads and judges' experience. In Columns (5) and (6) of Table 5, I remove the districts that have very few or very high number of trials. Specifically, I remove the low 10 percentiles (Column (5)) and high 10 percentiles (Column (6)) of the distribution of total trials per districts.[34] The removal of "small" an "big" districts do not seem to affect the coefficients for female judges, which remain positive and significant. Columns (5) and (6) indicate that my results are not driven by small number of trials that takes place in a very "small" or "big" districts.

In Columns (7) and (8) of Table 5 I evaluate the robustness of my results with respect to judges' workloads. As mentioned in Section 5.1, some judges are disproportionally over-represented or under-represented in my sample. In Columns (7) and (8) of Table 5, I remove judges with the low 10 percentiles and top 10 percentile of the workloads' distribution. Also, in this case, the coefficients for female judges are similar to my benchmark model and they remain positive and significant. These findings indicate that the results in Table 4 do not seem to be driven by judges with very high/low workloads.

Finally, the previous literature (e.g Johnson (2014) and Steffensmeier and Hebert (1999)) indicates that tenure on the bench can have an impact on sentences duration. The differences in attitudes between junior and senior judges could have an impact also on my findings which could br driven by very junior or very senior judges. In Columns (9) and (10) of Table 5 I remove the low 10 percentiles and high 10 percentiles of the distribution of *Experience (per 100)*. The coefficients for female judges remain positive and significant and the results in Table 4 does not seem to be driven by very junior/very senior judges.

5.3 Heterogeneity checks

5.3.1 Judges' Characteristics

In this section I implement a series of heterogeneity checks based on the different judges' characteristics. In Table 6, I interact female judges with my dummy for ethnicity (Column (2)), judges' political affiliation (Column (3)) and electoral cycle (Column (4)). In Column (1) of Table 6 I report my benchmark model.

[TABLE 6 APPROXIMATELY HERE]

The findings presented in Columns (2)-(4) of Table 6 are interesting. In Column (2) of Table 6

[34] I also implement the analysis by removing the top and low 5 percentiles of the distribution of total trials per districts. The findings are similar to those in Table 5. Available upon request.

the interaction between female judges and non-white judges is negative but non-significant, while the coefficients for female white judges is positive and significant. These findings indicate that, at least for female judges, ethnicity matters, with white judges more sway to influence juries towards guilty verdicts. In Column (3) of Table 6, surprisingly, the coefficient indicates that the interaction between Republican and female judges is non-significant. Similarly, interaction between independent and female judges is non-significant while the result for democratic female judges is positive and significant. These results are contrary to the main literature where Republican judges are considered harsher with respect to Democratic judges. Finally, in Column (4) of Table 6 I examine the possible presence of an electoral cycle effect. The interaction between the dummy for next year elections and female judges is non-significant, suggesting that career concern do not seem to play a statistically significant role.[35]

5.3.2 Type of Crimes

In this Section, I implement a series of heterogeneity checks based on different types of crimes. First, I generate a set of dummies for the different types of crimes: *Violent crimes*,[36] *Property Crimes* and *Drug*. Second, I interact female judges with the crime variables. The results are represented in Table 7, with Column (1) represented my benchmark model.[37]

[TABLE 7 APPROXIMATELY HERE]

The interactions between female judges and violent crimes (Column (2)), drug crimes (Column (3)), and property crimes (Column (4)) are non-significant. Table 7 indicates that, at least for female judges, the type of crimes do not seem to have statistical significant impact on the probability of a guilty verdict.

5.3.3 Jurors' gender

As mentioned in Section 2, juries have the responsibility of assessing the facts and reaching a verdict. Previous literature, such as Anwar et al. (2012), indicates that juries characteristics matters in term of jury outcomes. In this Section, I explore how the interaction between female judges and juries gender composition can affect trials outcomes using two different (yet complementary) strategies. First, I interact the dummy about judges' gender with a dummy for female jurors being the majority in jury pools (Column (2) of Table 8). Second, I interact the dummy about judges'

[35] I also interact my variable for the different years of the election cycle. The results are similar to those in Table 6. Available upon request.

[36] I define as *Violent Crimes* as a dummy equal to one if there is at least one of the following offences: murder, robbery, drug, sex and other violent crimes.

[37] I also replicate, with similar findings, this exercise using the crime dummies as defined by Table 1. Available upon request.

gender with a set of dummies about the proportion of women in the jury pools (Column (3) of Table 8). The results are presented in Table 8.[38]

[TABLE 8 APPROXIMATELY HERE]

Column (1) of Table 8 represents the benchmark model (Column 4 in Table 4). In Column (2) of Table 8, the interaction *Female Judge X Female Majority (jury pool)* is positive and (slightly) significant. This result indicates that, when there is a female judge, jury gender composition plays a role although the coefficient is not extremely robust.

In Column (3) of Table 8 I take a deeper look to the gender dynamics between the jury and the judge by interacting the variable about judges' gender with a set of dummies about female jurors' proportion in the jury pools. These dummies are based on the percentiles of the distribution of *Jury pool: women (Prop.)*. The coefficients in Column (3) of Table 8 suggest that trials with the top 20 percentile female jurors are positive and significant, while the rest of coefficients are negative and mostly non-significant. My results in Table 4 seem to be strongly influenced by trials with the highest percentage of women in the jury pools.

Table 8 indicates that the impact of judges' gender on trials outcomes is affected by the gender composition of jury pools. The importance of gender dynamics is already presented in previous works, evaluating the teacher-student relationship. For example, Carrell et al. (2010) indicate that, although professors' gender has little impact on male students, it has a powerful effect on female students' performance in math and science classes. Another work is Dee (2007), which investigates the role of same-gender teachers on students' achievement. Within-student comparisons indicate that assignment to a same-gender teacher significantly improves the achievement of both girls and boys as well as teacher perceptions of student performance and student engagement with the teacher's subject. Overall, these works indicate that in the student-teacher relationship gender matters. This literature can provide useful information to interpret my results because the student-teacher relationship has many points in common with the juror-judge relationship, such as the imbalance in power and knowledge. Moreover, Table 8 indicates that gender plays a role also in juror-judge dynamics.

6 Why does judges' gender play a role?

The previous sections document the importance of judges' gender on jury trials outcomes. In this section, I speculate over the possible mechanisms behind these findings. First, I investigate the role

[38] I replicate them also for jurors' ethnicity and political affiliation. The findings are (mostly) non-significant. Results available upon request.

of judges' gender on the seated jury composition in Section 6.1. Second, in Section 6.2, I evaluate the possible differences in attitudes and behaviours of female judges and jurors.

6.1 Jury selection process and seated jury composition

Previous literature, such as Anwar et al. (2012, 2014) and Flanagan (2018), suggests that jury composition influences trials outcomes. For example, higher proportions of white, white male and old jurors in the jury pools are more likely to generate guilty verdicts. As mentioned in Section 2, judges play a role in the jury selection. A possible mechanism behind my findings in Table 4 is that, during the jury selection process, female judges systematically select jurors that are more likely to be associated with more guilty verdicts.

In more details, the jury selection process allows judges, defence attorneys, and prosecutors to remove those jurors that they deem to be biased. Potential jurors can be excluded either for cause or thanks to peremptory challenges. A removal for cause is when the judge decides to struck a juror for apparent bias or hardship.[39] These removals are unlimited but are left to judges' discretion. Prosecutors and defence attorneys can exclude jurors through the "peremptory challenges". These removals do not require any explanations, but they are limited in numbers and cannot be based on ethnicity or gender (*Batson v. Kentucky*, 476 U.S. 79 [1986]; *J.E.B. v. Alabama*, 511 U.S. 127 [1994]).[40] In North Carolina, prosecutors and defence attorneys have six peremptory challenges plus one for every alternative juror each.

If female and male judges behave differently in their choices of jurors during the jury selection process, this could influence the seated juries' composition, and, as a consequence, the verdicts. To evaluate the relationship between exclusions for cause, characteristics of jurors and judges, I implement a series of OLS regressions.[41] As dependent variables, I use different dummies for the characteristics of the potential jurors while, as independent variables, I use the removal from court and the interaction between the removal form court and female judges. I also include specific set of controls and trial fixed effects both interacted with female judges. The findings are presented in Table 9.[42]

[TABLE 9 APPROXIMATELY HERE]

Overall, the interaction in Table 9 between removal from court and female judges is almost

[39] Some examples are medical or financial hardship.

[40] If one side is suspected of gender/race discrimination, the opposing side may object using the so-called Batson's challenge. In practice, successful Batson challenge are extremely rare.

[41] Additionally, there is a literature about the role of number of jurors in jury outcomes. I replicate my analysis using the choice in the number of alternative jurors. However, the results were not statistically significant. Available upon request.

[42] Full regressions is presented in Section A2 in the Appendix.

always non-statistically significant. Table 9 indicates that judges' gender do not seem to influence the jury selection process and the characteristics of the seated jury.

6.2 Judges Behaviours

As mentioned in Section 2, while the judge is the most important person in the courtroom, it is the jury role to reach a verdict. The judge is supposed to be a neutral actor. However, my findings indicate that judges' gender can sway the jury towards a guilty verdict. Moreover, the previous literature suggests that judges have opinions and expectations over the trials' outcomes and that these ideas "leak" towards the juries, with verbal and non-verbal behaviours (e.g. Lenehan and O'Neill (1981), Ekman and Friesen (1969) and Burnett and Badzinski (2005)). If female judges, on average, have stricter attitude or a more likely to expect a guilty verdict and these attitudes and expectations "leak" to the juries, this could be one of the mechanisms behind my findings in Table 4.

In this Section, I investigate the existence of different attitudes between men and women in general and in my sample. First, I use the *General Social Survey* (GSS)[43] dataset and explore the attitude of women, in general, towards the legal system (Section 6.2.1). Second, I create a measured judges' toughness in my sample and evaluate the possible differences between female and male judges (Section 6.2.2).

6.2.1 GSS analysis

As proposed Anwar et al. (2014), I explore the attitude of US women towards the judiciary system by extracting some information about the general attitudes towards courts and legal system from the GSS:[44]

- Courts and Courts behaviours: extrapolated from a series of questions about courts' harshness/leniency[45] and stiffness of sentences.[46]
- Law and the legal system: extrapolated from a series of questions about the importance of the law[47] and the importance of obeying to the law.[48]

[43] Source: GSS.
[44] Descriptive statistics are presented in Table A3 in the Appendix.
[45] The variables are extracted from those answering "Too harsh", "too lenient" to the question "Courts dealing with criminals".
[46] The variables are extracted from those answering "strongly agree" or "agree" to the question "Lawbreakers should get stiffer sentences".
[47] The variables are extracted from those answering "very important" or "fairly important" to the question "How important to respect America's laws etc".
[48] The variables are extracted from those individuals answering "very important" to the question " Obligation–jury duty".

I implement a linear probability model to explore the relationship between these variables and gender. My results are presented in Table 10.[49]

[TABLE 10 APPROXIMATELY HERE]

The findings in Table 10 indicate that women have harsher attitudes and are more likely to consider obeying to the law important. Specifically, women are more likely to find courts not harsh enough (Columns (1) and (2)). Moreover, the coefficients in Columns (3) and (4) are positive and statistically significant, indicating that women are more likely to consider the law and its obedience important. Overall, these findings suggest that women consider the law important and are stricter in term of sentences. These differences in attitudes between men and women can affect both judges and jurors. First, this is a first indication that female judges could be stricter than their male counterparts (more in Section 6.2.2). Second, also female jurors might be stricter than male jurors and more like to agree with the strictness "leaked" by the judges. Some findings supporting this mechanisms are present in Table 10.

6.2.2 Judges Toughness

In this section, I focus on the female judges in my sample. Overall, my data over judges' attitudes during trials is scarce. However, I notice that female judges, on average, give (slightly) longer sentences with respect of male judges (three additional months). Using data about sentences, I generate an (imperfect) measure for "toughness". I consider a judge tough if his/her sentences duration are in the 10% toughest sentences and I generate a dummy based on this definition. In Table 11, I implement a series of descriptive statistics about the distribution of "toughness" across gender.

[TABLE 11 APPROXIMATELY HERE]

Table 11 indicates that female judges are tougher with respect to male judges (Table 11(a)). Moreover, in Table 11(b) when a tough woman presides the trial, it is more likely that the jury expresses a guilty verdict.

The findings in Table 11 combined with those in Table 10 are coherently indicating that female judges could have stricter attitudes than their male counterparts, which could "leak" to the jury and influence the verdict. According to teh previous literature, the possible causes of these different attitudes are many. First, as mentioned before, the previous literature indicates that women and men have different decision-making process, also in the judiciary system (e.g. Gilligan (1982)

[49] Summary statistics and full regressions are presented in Section A3 in the Appendix.

and Fox and Van Sickel (2000)). Second, Croson and Gneezy (2009) and Byrnes et al. (1999) indicate that women are more risk adverse. This could make female judges less likely to release potential (and potentially violent) criminals in society. Third, female judges are elected officials and, similar to female politicians, have to deal with positive/negative stereotypes. For example, Sanbonmatsu (2003) suggests that female political candidates are usually considered better to handle "compassionate" issues (such as elderly care), while male politicians' candidates are better with crime and military. It can be safely assumed that female judges face similar stereotypes during the electoral campaigning. To overcome the negative environment, only women with a more than average severe attitude might have a possibility of winning.

7 Conclusion

This article investigates the impact of female judges in jury trials' outcomes. The analysis is implemented in North Carolina between 2010-2012 and the research design relies on fixed effects and judges' rotation. My findings indicate that female judges are more likely to preside jury trials that end with a guilty verdict. The results do not depend on the choice of model and they are robust to the exclusion of female defendants, appointed judges, small and big districts, very low and high experience and low and high judges' workloads. I also implement a series of heterogeneity checks based on judge characteristics, types of crimes and proportion of women in the jury pools. The findings indicate that female democratic judges and white judges play an important role in my main results. Moreover, the interaction between female judges and high number of female jurors in the jury pools is positive and significant.

Finally, I explore the possible mechanisms behind my findings. The data does not seem to support the presence of an effect of judges' gender on the jury selection process. On the contrary, they indicate that female judges are tougher than their male counterpart. The previous literature suggest that this stricter attitude "leaks" to the jury.

References

Anwar, S., Bayer, P., & Hjalmarsson, R. (2012). The impact of jury race in criminal trials. *The Quarterly Journal of Economics*, *127*(2), 1017–1055.

Anwar, S., Bayer, P., & Hjalmarsson, R. (2014). The role of age in jury selection and trial outcomes. *The Journal of Law and Economics*, *57*(4), 1001–1030.

Bandyopadhyay, S., & McCannon, B. C. (2014). The effect of the election of prosecutors on criminal trials. *Public Choice*, *161*(1-2), 141–156.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, *119*(1), 249–275.

Blanck, P. D., Rosenthal, R., & Cordell, D. H. (1985). The appearance of justice: Judges'verbal and nonverbal behaviour in criminal jury trials. *Stanford Law Review*.

Burnett, A., & Badzinski, D. M. (2005). Judge nonverbal communication on trial: Do mock trial jurors notice? *Journal of Communication*, *55*(2), 209–224.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, *125*(3), 367.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The review of economics and statistics*, *90*(3), 414–427.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of human resources*, *50*(2), 317–372.

Carrell, S. E., Page, M. E., & West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *The Quarterly journal of economics*, *125*(3), 1101–1144.

Carter, A. V., Schnepel, K. T., & Steigerwald, D. G. (2017). Asymptotic behavior of at-test robust to cluster heterogeneity. *Review of Economics and Statistics*, *99*(4), 698–709.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, *47*(2), 448–74.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human resources*, *42*(3), 528–554.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, *32*(1), 88–106.

Fischman, J. B., & Schanzenbach, M. M. (2011). Do standards of review matter? the case of federal criminal sentencing. *The Journal of Legal Studies*, *40*(2), 405–437.

Flanagan, F. X. (2018). Race, gender, and juries: Evidence from north carolina. *The Journal of Law and Economics*, *61*(2), 189–214.

Fox, R., & Van Sickel, R. (2000). Gender dynamics and judicial behavior in criminal trial courts: An exploratory study. *Justice System Journal*, *21*(3), 261–280.

Fujiwara, T., & Wantchekon, L. (2013). Can informed public deliberation overcome clientelism? experimental evidence from benin. *American Economic Journal: Applied Economics*, *5*(4), 241–55.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Harvard University Press.

Halverson, A. M., Hallahan, M., Hart, A. J., & Rosenthal, R. (1997). Reducing the biasing effects of judges' nonverbal behavior with simplified jury instruction. *Journal of Applied Psychology*, *82*(4), 590.

Hart, A. J. (1995). Naturally occurring expectation effects. *Journal of Personality and Social Psychology*, *68*(1), 109.

Hart, A. J. (1992). On the sobriety of judges: Nonverbal influence in the courtroom.

Heß, S. (2017). Randomization inference with stata: A guide and software. *The Stata Journal*, *17*(3), 630–651.

Johnson, B. D. (2014). Judges on trial: A reexamination of judicial race and gender effects across modes of conviction. *Criminal Justice Policy Review*, *25*(2), 159–184.

Lenehan, G. E., & O'Neill, P. (1981). Reactance and conflict as determinants of judgment in a mock jury experiment. *Journal of Applied Social Psychology*, *11*(3), 231–239.

Nagler, M., Piopiunik, M., & West, M. R. (2020). Weak markets, strong teachers: Recession at career start and teacher effectiveness. *Journal of Labor Economics*, *38*(2), 453–500.

Sanbonmatsu, K. (2003). Political knowledge and gender stereotypes. *American Politics Research*, *31*(6), 575–594.

Schanzenbach, M. M., & Tiller, E. H. (2008). Reviewing the sentencing guidelines: Judicial politics, empirical evidence, and reform. *The University of Chicago Law Review*, *75*(2), 715–760.

Songer, D. R., Davis, S., & Haire, S. (1994). A reappraisal of diversification in the federal courts: Gender effects in the courts of appeals. *The Journal of Politics*, *56*(2), 425–439.

Spitzer, M., & Talley, E. (2013). Left, right, and center: Strategic information acquisition and diversity in judicial panels. *The Journal of Law, Economics, & Organization*, *29*(3), 638–680.

Spohn, C. (1991). Decision making in sexual assault cases: Do black and female judges make a difference? *Women & Criminal Justice*, *2*(1), 83–105.

Steffensmeier, D., & Hebert, C. (1999). Women and men policymakers: Does the judge's gender affect the sentencing of criminal defendants? *Social forces*, *77*(3), 1163–1196.

Tables

Table 1: Summary Statistics

	count	mean	sd	min	max
Defendant characteristics:					
Def. woman	598	0.1120	0.32	0.00	1.00
Def. white	598	0.3361	0.47	0.00	1.00
Def. black	598	0.6288	0.48	0.00	1.00
Def. over 30	598	0.6472	0.48	0.00	1.00
Trials' characteristics:					
Murder charge	598	0.0234	0.15	0.00	1.00
Robbery charge	598	0.0803	0.27	0.00	1.00
Drug charge	598	0.1973	0.40	0.00	1.00
Sex charge	598	0.1221	0.33	0.00	1.00
Other violent crime charge	598	0.1421	0.35	0.00	1.00
Property crime charge	598	0.2993	0.46	0.00	1.00
Other crime charge	598	0.0769	0.27	0.00	1.00
Dependent variable:					
Guilty	598	0.7174	0.45	0.00	1.00
Jury seated characteristics:					
Jury Pool: women (Prop.)	598	0.5108	0.11	0.14	0.79
Jury Pool: men (Prop.)	598	0.4571	0.11	0.18	0.81
Jury Pool: black (Prop.)	598	0.1686	0.14	0.00	0.67
Jury Pool: white (Prop.)	598	0.6123	0.18	0.10	1.00
Jury Pool: Republicans (Prop.)	598	0.2589	0.12	0.00	0.63
Jury Pool: Democrats (Prop.)	598	0.3217	0.14	0.00	0.83
Judge characteristics:					
Experience (per 100)	598	0.2985	0.08	0.13	0.44
Non-White Judge	598	0.1538	0.36	0.00	1.00
Nr. trials per Judge	598	3.5452	2.47	1.00	10.00
Democratic Judge	598	0.6187	0.49	0.00	1.00
Republican Judge	598	0.2659	0.44	0.00	1.00
Election (Next year)	598	0.0836	0.28	0.00	1.00
Non-Elected judges	598	0.1405	0.35	0.00	1.00
Female Judge	598	0.0769	0.27	0.00	1.00

Notes: *Def. woman*, *Def. white*, *Def. black* and *Def. over 30* are dummies for defendants' gender, ethnicity and age. *Murder charge*, *Robbery charge*, *Drug charge*, *Sex charge*, *Other violent crime charge*, *Property crime charge* and *Other crime charge* are a set of dummies for the type of crimes. *Guilty* is a dummy for the trials outcomes. *Jury Pool: women (Prop.)*, *Jury Pool: men (Prop.)*, *Jury Pool: Black (Prop.)*, *Jury Pool: White (Prop.)*, *Jury Pool: Republicans (Prop.)* and *Jury Pool: Democrats (Prop.)* are a series of variables representing the proportion of women, men, black and white, Republicans and Democrats in jury pools. *Experience (per 100)* is the number of years since the J.D. of the judge and it is divided by 100. *Nr. trials per Judge* indicates the number of trials per different judges. *Non-White Judge*, *Democratic Judge*, *Republican Judge*, *Election (Next year)*, and *Female Judge* are a series of dummies for ethnicity, political affiliation, elections and gender. *Non-Elected Judge* is a dummy for non-elected judges.

Table 2: Differences between Judges by gender

	Female (Mean)	Male (Mean)	Diff.	Std. Error	Obs.
Experience (per 100)	0.2374	0.3036	0.0662***	0.0120	598
Non-White Judge	0.5870	0.1178	-0.4692***	0.0520	598
Nr. trials per Judge	2.3043	3.6486	1.3442***	0.3760	598
Democratic Judge	0.6957	0.6123	-0.0833	0.0746	598
Republican Judge	0.2174	0.2699	0.0525	0.0679	598
Election (Next year)	0.1957	0.0743	-0.1214***	0.0423	598
Non-Elected judges	0.1739	0.1377	-0.0362	0.0534	598

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. *Experience (per 100)* is in year divided by 100. *Nr. trials per Judge* indicates the number of trials per different judges. *Non-White Judge*, *Democratic Judge*, *Republican Judge*, *Election (Next year)* and *Female Judge* are a series of dummies for ethnicity, political affiliation, elections and gender. *Non-Elected Judge* is a dummy for non-elected judges.

Table 3: Relationship between Judges demographics and Trials Characteristics

dep. var.:	Female Judge					
	(1)			(2)		
	b	se	t	b	se	t
Murder charge	-0.0881*	(0.05)	[-1.78]	-0.0877	(0.08)	[-1.07]
Robbery charge	-0.0527	(0.06)	[-0.95]	-0.00438	(0.05)	[-0.09]
Drug charge	-0.00205	(0.05)	[-0.04]	0.0188	(0.05)	[0.36]
Sex charge	0.00531	(0.06)	[0.09]	0.0767	(0.07)	[1.17]
Other violent crime charge	-0.0184	(0.06)	[-0.33]	0.00425	(0.05)	[0.09]
Property crime charge	-0.0114	(0.05)	[-0.22]	0.0410	(0.04)	[0.93]
Other crime charge	0.0438	(0.07)	[0.64]	0.0759	(0.08)	[0.92]
Jury Pool: women (Prop.)	0.215	(0.24)	[0.90]	-0.0429	(0.26)	[-0.16]
Jury Pool: men (Prop.)	0.300	(0.25)	[1.19]	0.0763	(0.27)	[0.28]
Jury Pool: black (Prop.)	-0.0227	(0.17)	[-0.13]	0.0852	(0.23)	[0.37]
Jury Pool: white (Prop.)	-0.215	(0.13)	[-1.64]	-0.131	(0.16)	[-0.80]
Jury Pool: Republicans (Prop.)	-0.0361	(0.12)	[-0.31]	-0.0288	(0.14)	[-0.20]
Jury Pool: Democrats (Prop.)	0.0448	(0.10)	[0.44]	0.0997	(0.13)	[0.78]
Def. woman	0.00175	(0.03)	[0.05]	0.0304	(0.03)	[0.88]
Def. white	0.0249	(0.05)	[0.49]	0.0767	(0.07)	[1.03]
Def. black	0.0341	(0.05)	[0.70]	0.0480	(0.07)	[0.66]
Def. over 30	0.00696	(0.02)	[0.31]	0.0160	(0.02)	[0.78]
Constant	-0.0655	(0.23)	[-0.29]	0.00589	(0.26)	[0.02]
Observations			598			598
F-stat			1.971			0.764
Year, Prosecutor, District FE			No			Yes

Notes: OLS regressions. Robust standard errors in parenthesis and t statistics in square brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. *Female Judge* is a dummy for female judges. *Def. woman*, *Def. white*, *Def. black* and *Def. over 30* are dummies for defendants' gender, ethnicity and age. *Murder charge*, *Robbery charge*, *Drug charge*, *Sex charge*, *Other violent crime charge*, *Property crime charge* and *Other crime charge* are a set of dummies for the type of crimes. *Jury Pool: women (Prop.)*, *Jury Pool: men (Prop.)*, *Jury Pool: Black (Prop.)*, *Jury Pool: White (Prop.)*, *Jury Pool: Republicans (Prop.)* and *Jury Pool: Democrats (Prop.)* are a series of variables representing the proportion of women, men, black and white, Republicans and Democrats in the jury pools. *Excluded category*: weapon.

Table 4: Main results

dep. var.:	Guilty			
	(1)	(2)	(3)	(4)
Female Judge	0.139**	0.178**	0.251***	0.269***
	(0.06)	(0.08)	(0.09)	(0.10)
	[2.47]	[2.13]	[2.65]	[2.80]
Experience (per 100)			-3.275	-3.740*
			(2.26)	(2.26)
			[-1.45]	[-1.66]
Experience sq. (per 100)			0.000603	0.000674*
			(0.00)	(0.00)
			[1.56]	[1.73]
Non-White Judge			-0.0799	-0.117
			(0.08)	(0.08)
			[-1.01]	[-1.49]
Nr. trials per Judge			0.0122	0.0143
			(0.01)	(0.01)
			[1.03]	[1.24]
Democratic Judge			0.0766	0.0584
			(0.08)	(0.09)
			[0.96]	[0.68]
Republican Judge			0.0402	0.0170
			(0.09)	(0.09)
			[0.47]	[0.19]
Election (Next year)			0.0216	0.0424
			(0.09)	(0.09)
			[0.23]	[0.47]
Non-Elected judges			-0.0472	-0.0619
			(0.07)	(0.07)
			[-0.68]	[-0.91]
Year, Prosecutor, District FE	No	Yes	Yes	Yes
Controls	No	No	No	Yes
Observations	602	602	598	598
N. District	33	33	33	33
Mean Guilty	0.719	0.719	0.717	0.717
Mean Fem Judge	0.0764	0.0764	0.0769	0.0769

Notes: Robust standard errors in parenthesis and t statistics in square brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. *Guilty* is a dummy for a guilty verdict. *Female Judge* is a dummy for female judges. *Experience (per 100)* is the number of years since the J.D. of the judge and it is divided by 100. *Nr. trials per Judge* indicates the number of trials per different judges. *Non-White Judge*, *Democratic Judge*, *Republican Judge*, *Election (Next year)*, and *Female Judge* are a series of dummies for ethnicity, political affiliation, elections and gender. *Non-Elected Judge* is a dummy for non-elected judges. *Controls* includes: Murder charge, Robbery charge, Drug charge, Sex charge, Other violent crime charge, Property crime charge, Other crime charge, Jury Pool: women (Prop.), Jury Pool: men (Prop.), Jury Pool: Black (Prop.), Jury Pool: White (Prop.), Jury Pool: Republicans (Prop.), Jury Pool: Democrats (Prop.), Def. woman, Def. white, Def. black and Def. over 30.

Table 5: Robustness checks

dep. var.:	Guilty									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female Judge	0.269***	0.471***	0.253**	0.320***	0.272**	0.278**	0.294**	0.261***	0.258**	0.274***
	(0.10)	(0.15)	(0.11)	(0.11)	(0.12)	(0.12)	(0.14)	(0.10)	(0.10)	(0.10)
	[2.80]	[3.06]	[2.22]	[2.83]	[2.35]	[2.39]	[2.16]	[2.64]	[2.47]	[2.80]
Observations	598	387	517	495	508	509	428	528	523	517
N. District	33	25	32	33	23	33	27	33	31	33
Removal: Female Defendants			Yes							
Removal: Non-Elected Judge				Yes						
Removal: "Small" Districts					Yes					
Removal: "Big" Districts						Yes				
Removal: Small Workload							Yes			
Removal: Big Workload								Yes		
Removal: Junior Judge									Yes	
Removal: Senior Judge										Yes
Year, Prosecutor, District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parenthesis and t statistics in square brackets. * p < 0.1, ** p < 0.05, *** p < 0.01. *Guilty* is a dummy for a guilty verdict. *Female Judge* is a dummy for female judges. In Column (1) I present my benchmark model as in Column (4) in Table (4). In Column (2) I present the margin from a Logit regression. In Columns (3)-(10) I remove from my sample different elements to check the robustness of my findings. *Removal: Female Defendants* indicates the removal of female defendants, *Removal: Non- Elected Judge* the removal of non-elected judges. *Removal: "Small" Districts* and *Removal: "Big" Districts* indicate the removal of the lower 10 percentiles and higher 10 percentiles of the distribution of total trials per districts. *Removal: Small Workload* and *Removal: Big Workload* indicate the removal of the lower 10 percentiles and higher 10 percentiles of the distribution of workload per judge. *Removal: Junior Judges* and *Removal: Senior Judges* indicate the removal of the lower 10 percentiles and higher 10 percentiles of the distribution of tenure on the bench per judge. *Controls* includes: Experience (per 100), Nr. trials per Judge, Non-White Judge, Democratic Judge, Republican Judge, Election (Next year), Non-Elected Judge, Murder charge, Robbery charge, Drug charge, Sex charge, Other violent crime charge, Property crime charge, Other crime charge, Jury Pool: women (Prop.), Jury Pool: men (Prop.), Jury Pool: Black (Prop.), Jury Pool: White (Prop.), Jury Pool: Republicans (Prop.), Jury Pool: Democrats (Prop.), Def. woman, Def. white, Def. black and Def. over 30. *Excluded category*: weapon.

Table 6: Heterogeneity checks over judges characteristics

dep. var.:	Guilty			
	(1)	(2)	(3)	(4)
Female Judge	0.269***	0.348***	0.243**	0.201*
	(0.10)	(0.13)	(0.12)	(0.10)
	[2.80]	[2.71]	[2.07]	[1.96]
Non-White Judge		-0.0915		
		(0.09)		
		[-1.04]		
Female Judge x Non-White Judge		-0.170		
		(0.20)		
		[-0.85]		
Republican Judge			-0.0198	
			(0.06)	
			[-0.34]	
Independent Judge			-0.0448	
			(0.09)	
			[-0.50]	
Female Judge x Republican Judge			-0.0507	
			(0.21)	
			[-0.24]	
Female Judge x Independent Judge			-0.00197	
			(0.34)	
			[-0.01]	
Election (Next year)				-0.0206
				(0.10)
				[-0.20]
Election (Next year) x Female Judge				0.149
				(0.16)
				[0.90]
Observations	598	598	598	598
N. District	33	33	33	33
Year, Prosecutor, District FE	Yes	Yes	Yes	Yes
Judges' Controls	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parenthesis and t statistics in square brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. *Guilty* is a dummy for a guilty verdict. *Female Judge* is a dummy for female judges. *Controls* includes: Experience (per 100), Nr. trials per Judge, Non-White Judge, Democratic Judge, Republican Judge, Election (Next year), Non-Elected Judge, Murder charge, Robbery charge, Drug charge, Sex charge, Other violent crime charge, Property crime charge, Other crime charge, Jury Pool: women (Prop.), Jury Pool: men (Prop.), Jury Pool: Black (Prop.), Jury Pool: White (Prop.), Jury Pool: Republicans (Prop.), Jury Pool: Democrats (Prop.), Def. woman, Def. white, Def. black and Def. over 30. *Excluded category*: weapon.

Table 7: Heterogeneity checks over crime characteristics

dep. var.:	Guilty			
	(1)	(2)	(3)	(4)
Female Judge	0.269***	0.237*	0.366***	0.392***
	(0.10)	(0.13)	(0.11)	(0.14)
	[2.80]	[1.81]	[3.29]	[2.89]
Violent Crimes		1.593		
		(1.90)		
		[0.84]		
Violent Crimes x Female Judge		0.278		
		(0.25)		
		[1.12]		
Drug Crimes			-1.530	
			(1.73)	
			[-0.89]	
Drug Crimes x Female Judge			-0.147	
			(0.39)	
			[-0.37]	
Property Crimes				1.999
				(1.91)
				[1.05]
Property Crimes x Female Judge				0.00812
				(0.24)
				[0.03]
Observations	598	598	598	598
N. District	33	33	33	33
Year, Prosecutor, District FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parenthesis and t statistics in square brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. *Guilty* is a dummy for a guilty verdict. *Female Judge* is a dummy for female judges. *Controls* includes: Experience (per 100), Nr. trials per Judge, Non-White Judge, Democratic Judge, Republican Judge, Election (Next year), Non-Elected Judge, Jury Pool: women (Prop.), Jury Pool: men (Prop.), Jury Pool: Black (Prop.), Jury Pool: White (Prop.), Jury Pool: Republicans (Prop.), Jury Pool: Democrats (Prop.), Def. woman, Def. white, Def. black and Def. over 30. *Excluded category*: weapon.

Table 8: Heterogeneity checks over jurors' gender

dep. var.:	Guilty		
	(1)	(2)	(3)
Female Judge	0.269***	0.156	0.527***
	(0.10)	(0.12)	(0.15)
	[2.80]	[1.34]	[3.59]
Female Majority (jury pool)		0.00334	
		(0.05)	
		[0.07]	
Female Judge X Female Majority (jury pool)		0.222*	
		(0.13)	
		[1.69]	
Female jurors below 20 cent. (prop. pool)			-0.0261
			(0.07)
			[-0.36]
Female jurors btw 20-40 cent. (prop. pool)			-0.0616
			(0.07)
			[-0.89]
Female jurors btw 40-60 cent. (prop. pool)			-0.0844
			(0.08)
			[-1.10]
Female jurors btw 60-80 cent. (prop. pool)			0.0317
			(0.07)
			[0.45]
Female Judge X Female jurors below 20 cent. (prop. pool)			-0.386*
			(0.21)
			[-1.86]
Female Judge X Female jurors btw 20-40 cent. (prop. pool)			-0.358*
			(0.20)
			[-1.80]
Female Judge X Female jurors below btw 40-60 (prop. pool)			-0.165
			(0.20)
			[-0.81]
Female Judge X Female jurors btw 60-80 cent. (prop. pool)			-0.382*
			(0.22)
			[-1.76]
Observations	598	598	598
N. District	33	33	33
Year, Prosecutor, District FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Robust standard errors in parenthesis and t statistics in square brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. *Guilty* is a dummy for a guilty verdict. *Female Judge* is a dummy for female judges. *Controls* includes: Experience (per 100), Nr. trials per Judge, Non-White Judge, Democratic Judge, Republican Judge, Election (Next year), Non-Elected Judge, Murder charge, Robbery charge, Drug charge, Sex charge, Other violent crime charge, Property crime charge, Other crime charge, Jury Pool: Black (Prop.), Jury Pool: White (Prop.), Jury Pool: Republicans (Prop.), Jury Pool: Democrats (Prop.), Def. woman, Def. white, Def. black and Def. over 30. *Excluded category*: weapon. *Majority* indicates over 50% of the jury is composed by female. In Columns (3)-(7) *Year FE, Prosecutor FE, District FE* and *Controls* are interacted with the variables.

Table 9: Jury Selection Process and Female Judges

dep. var.:	Republican Juror		Democrats Juror		Independent Juror		Female Juror		White Juror		Black Juror	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Removal Court	-0.0757***	-0.0461***	0.0270*	0.0287**	0.00425	0.0182	0.0159	0.0104	-0.0433**	-0.0168	0.0433**	0.0168
	(0.014)	(0.013)	(0.016)	(0.014)	(0.013)	(0.013)	(0.018)	(0.018)	(0.017)	(0.015)	(0.017)	(0.015)
Female Judge	-0.143	-0.130	-0.127	-0.0778	0.280**	0.273**	-0.272*	-0.292*	0.0497	-0.0733	-0.0497	0.0733
	(0.137)	(0.129)	(0.106)	(0.096)	(0.130)	(0.121)	(0.145)	(0.149)	(0.091)	(0.079)	(0.091)	(0.079)
Female Judge X Removal Court	0.0443	0.0545	0.0737	0.0685	-0.0417	-0.0343	-0.104	-0.112*	-0.0584	-0.0322	0.0584	0.0322
	(0.048)	(0.042)	(0.058)	(0.046)	(0.042)	(0.040)	(0.065)	(0.065)	(0.069)	(0.055)	(0.069)	(0.055)
Trail FE X Female Judges	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls1 X Female Judges	No	Yes	No	Yes	No	Yes	No	No	No	No	No	No
Controls2 X Female Judges	No	No	No	No	No	No	No	Yes	No	No	No	No
Controls3 X Female Judges	No	No	No	No	No	No	No	No	No	Yes	No	Yes
Observations	13268	13268	13268	13268	13268	13268	12831	12831	10269	10269	10269	10269
N. District												
N.Trials	602	602	602	602	602	602	602	602	602	602	602	602

Notes: robust standard errors in parenthesis. * p < 0.1, ** p < 0.05, *** p < 0.01. *Excluded category*: whether the prospective jurors was seated on the jury. *Controls1* includes: removal from state, removal from defence, removal of unknown origin, female jurors, black jurors and white jurors. *Controls2* includes: removal from state, removal from defence, removal of unknown origin, Republican jurors, Democrats jurors, jurors with unknown political affiliation, black jurors and white jurors. *Controls3* includes: removal from state, removal from defence, removal of unknown origin, Republican jurors, Democrats jurors, jurors with unknown political affiliation and female jurors. In Columns (9)-(12) I keep only black or white jurors.

Table 10: GSS regressions results

dep. var.:	Courts:		Law: Importance of	
	Too harsh	Too lenient	Always obey	To respect
	(1)	(2)	(3)	(4)
Female	-0.020***	0.039***	0.090***	0.019**
	(0.00)	(0.00)	(0.01)	(0.01)
	[-8.69]	[11.23]	[7.09]	[2.45]
Observations	54056	54056	5802	3742
Controls	Yes	Yes	Yes	Yes

Notes: robust standard errors in parenthesis and t-statistics in brackets. * p < 0.1, ** p < 0.05, *** p < 0.01. Dependent variables are a dummies. *Controls*: years fixed effects, a dummy for unaffiliated and democrats, a dummy for white, a variable about the higher year of education completed, a variable for the number of children and a dummy for married. *Omitted category*: Republicans.

Table 11: Toughness, Judges gender and guilty verdicts

(a)

	Non Tough	Tough
Female Judges	58.7%	41.30%
Men Judges	64.03%	35.97%

(b)

	Non Guilty	Guilty
Non Tough	Female 18.52%	Female 81.48%
	Male 29.78%	Male 70.22%
Tough	Female 10.53%	Female 89.47%
	Male 28.00%	Male 72.00%

Figures

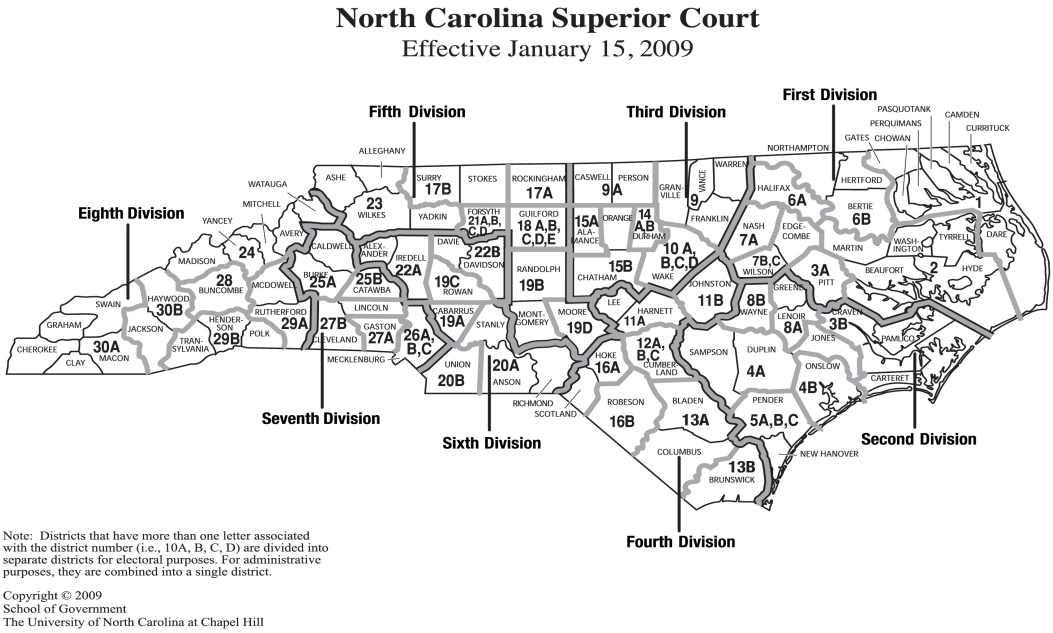

Figure 1: North Carolina Superior Court Map

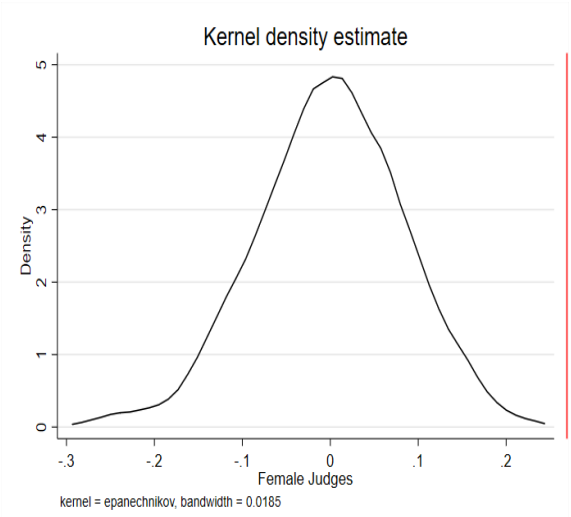

Figure 2: Permutation test. Shown it the Kernel density plot of a randomization inference test for simulated judges' gender assigned using 1000 replications. The red vertical line shows the benchmark model estimate (Column (4) in Table 4).

Appendices

A1 Main results: Additional results

Table A1: Main results

dep. var.:	Guilty			
	(1)	(2)	(3)	(4)
Female Judge	0.139**	0.178**	0.251***	0.269***
	(0.06)	(0.08)	(0.09)	(0.10)
	[2.47]	[2.13]	[2.65]	[2.80]
Experience (per 100)			-3.275	-3.740*
			(2.26)	(2.26)
			[-1.45]	[-1.66]
Experience sq. (per 100)			0.000603	0.000674*
			(0.00)	(0.00)
			[1.56]	[1.73]
Non-White Judge			-0.0799	-0.117
			(0.08)	(0.08)
			[-1.01]	[-1.49]
Nr. trials per Judge			0.0122	0.0143
			(0.01)	(0.01)
			[1.03]	[1.24]
Democratic Judge			0.0766	0.0584
			(0.08)	(0.09)
			[0.96]	[0.68]
Republican Judge			0.0402	0.0170
			(0.09)	(0.09)
			[0.47]	[0.19]
Election (Next year)			0.0216	0.0424
			(0.09)	(0.09)
			[0.23]	[0.47]
Non-Elected judges			-0.0472	-0.0619
			(0.07)	(0.07)
			[-0.68]	[-0.91]
Jury Pool: women (Prop.)				0.0553
				(0.60)
				[0.09]
Jury Pool: men (Prop.)				-0.250
				(0.60)
				[-0.42]
Jury Pool: black (Prop.)				-0.572
				(0.38)
				[-1.50]
Jury Pool: white (Prop.)				0.124
				(0.29)
				[0.43]
Jury Pool: Republicans (Prop.)				0.364
				(0.29)
				[1.26]
Jury Pool: Democrats (Prop.)				0.371
				(0.26)
				[1.44]
Def. woman				-0.0149
				(0.07)
				[-0.21]
defrace_nowhite				0.0145
				(0.05)
				[0.27]
Def. over 30				-0.00175
				(0.05)
				[-0.03]
Murder charge				0.181
				(0.24)
				[0.77]
Robbery charge				0.181
				(0.13)
				[1.44]
Drug charge				0.195
				(0.12)
				[1.57]
Sex charge				0.213*
				(0.12)
				[1.74]
Other violent crime charge				0.103
				(0.12)
				[0.90]
Property crime charge				0.232**
				(0.11)
				[2.17]
Other crime charge				0.396***
				(0.14)
				[2.85]
Constant	0.709***	0.706***	1.015***	0.790
	(0.02)	(0.02)	(0.35)	(0.64)
	[36.71]	[36.36]	[2.94]	[1.24]
Year, Prosecutor, District FE	No	Yes	Yes	Yes
Observations	602	602	598	598
N. District	33	33	33	33
Mean Guilty	0.719	0.719	0.717	0.717
Mean Fem Judge	0.0764	0.0764	0.0769	0.0769

Notes: Robust standard errors in parenthesis and t statistics in square brackets. * p < 0.1, ** p < 0.05, *** p < 0.01. *Guilty* is a dummy for a guilty verdict. *Experience of Judge* is in years. *Excluded category*: weapon

A2 Jury selection process and seated jury composition: Additional results

Table A2: Jury Selection Process and Female Judges

dep. var.:	Republican Juror		Democrats Juror		Independent Juror		Female Juror		White Juror		Black Juror	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Removal Court	-0.0757***	-0.0461***	0.0270*	0.0287**	0.00425	0.0182	0.0159	0.0104	-0.0433**	-0.0168	0.0433**	0.0168
	(0.014)	(0.013)	(0.016)	(0.014)	(0.013)	(0.013)	(0.018)	(0.018)	(0.017)	(0.015)	(0.017)	(0.015)
Removal State	-0.0940***	-0.0441***	0.105***	0.0472***	-0.0236**	-0.00479	-0.0605***	-0.0757***	-0.161***	-0.106***	0.161***	0.106***
	(0.013)	(0.012)	(0.015)	(0.013)	(0.012)	(0.012)	(0.017)	(0.017)	(0.016)	(0.014)	(0.016)	(0.014)
Removal Defense	0.0380***	-0.00340	-0.0528***	-0.00506	0.0204*	0.00430	-0.00150	0.00802	0.117***	0.0955***	-0.117***	-0.0955***
	(0.013)	(0.012)	(0.012)	(0.011)	(0.011)	(0.011)	(0.014)	(0.014)	(0.011)	(0.010)	(0.011)	(0.010)
Removal Unknown	-0.00719	-0.00850	0.0226	0.0130	0.00424	0.00270	-0.0273	-0.0295	-0.00781	0.000254	0.00781	-0.000254
	(0.017)	(0.016)	(0.018)	(0.016)	(0.014)	(0.014)	(0.020)	(0.020)	(0.017)	(0.016)	(0.017)	(0.016)
Female Judge	-0.143	-0.130	-0.127	-0.0778	0.280**	0.273**	-0.272*	-0.292*	0.0497	-0.0733	-0.0497	0.0733
	(0.137)	(0.129)	(0.106)	(0.096)	(0.130)	(0.121)	(0.145)	(0.149)	(0.091)	(0.079)	(0.091)	(0.079)
Female Judge X Removal Court	0.0443	0.0545	0.0737	0.0685	-0.0417	-0.0343	-0.104	-0.112*	-0.0584	-0.0322	0.0584	0.0322
	(0.048)	(0.042)	(0.058)	(0.046)	(0.042)	(0.040)	(0.065)	(0.065)	(0.069)	(0.055)	(0.069)	(0.055)
Female Judge X Removal State	0.0459	0.0573	-0.0295	-0.0437	0.0334	0.0464	0.130**	0.137**	-0.0231	-0.0180	0.0231	0.0180
	(0.040)	(0.035)	(0.054)	(0.037)	(0.040)	(0.039)	(0.055)	(0.055)	(0.057)	(0.044)	(0.057)	(0.044)
Female Judge X Removal Defense	0.0751*	0.0627	-0.0687	0.0151	-0.0319	-0.0425	-0.0473	-0.0499	0.116***	0.0707**	-0.116***	-0.0707**
	(0.042)	(0.040)	(0.044)	(0.039)	(0.034)	(0.034)	(0.049)	(0.049)	(0.038)	(0.035)	(0.038)	(0.035)
Female Judge X Removal Unknown	0.00284	0.0176	0.104	0.0866*	-0.0445	-0.0361	0.0160	0.00384	-0.0669	-0.00166	0.0669	0.00166
	(0.056)	(0.051)	(0.064)	(0.050)	(0.050)	(0.049)	(0.074)	(0.074)	(0.074)	(0.060)	(0.074)	(0.060)
Female juror		-0.0132*		0.0604***		-0.0114*				-0.0170**		0.0170**
		(0.007)		(0.007)		(0.007)				(0.007)		(0.007)
White juror		0.373***		0.172***		0.186***		-0.0616***				
		(0.007)		(0.008)		(0.007)		(0.017)				
Black juror		0.00775		0.733***		0.0532***		-0.0180				
		(0.007)		(0.011)		(0.009)		(0.021)				
Female Judge X Female juror		-0.00370		0.00325		0.00297				0.0304		-0.0304
		(0.024)		(0.025)		(0.022)				(0.026)		(0.026)
Female Judge X White juror		0.0141		-0.0200		0.0258		0.0355				
		(0.025)		(0.030)		(0.025)		(0.061)				
Female Judge X Black juror		0.00932		0.0314		-0.0303		-0.0239				
		(0.016)		(0.034)		(0.026)		(0.071)				
Dem juror								0.120***		-0.358***		0.358***
								(0.017)		(0.013)		(0.013)
Rep juror								0.0416**		0.0406***		-0.0406***
								(0.018)		(0.011)		(0.011)
Ind juror								0.0373**		-0.0270**		0.0270**
								(0.018)		(0.013)		(0.013)
Female Judge X Dem juror								0.0105		-0.0557		0.0557
								(0.062)		(0.051)		(0.051)
Female Judge X Rep juror								-0.00286		0.0437		-0.0437
								(0.066)		(0.044)		(0.044)
Female Judge X Ind juror								0.00575		0.0516		-0.0516
								(0.067)		(0.049)		(0.049)
Constant	0.409***	0.133	0.195**	-0.00127	0.159**	0.0220	0.591***	0.591***	0.953***	1.033***	0.0472	-0.0326
	(0.100)	(0.089)	(0.084)	(0.075)	(0.076)	(0.073)	(0.102)	(0.103)	(0.054)	(0.051)	(0.054)	(0.051)
Trail FE X Female Judges	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13268	13268	13268	13268	13268	13268	12831	12831	10269	10269	10269	10269
N. District												
N.Trials	602	602	602	602	602	602	602	602	602	602	602	602

Notes: robust standard errors in parenthesis. * p < 0.1, ** p < 0.05, *** p < 0.01. *Excluded category*: whether the prospective jurors was seated on the jury. In Columns (9)-(12) I keep only black or white jurors.

A3 Judges Behaviours: Additional results

Table A3: Summary Statistics GSS

	count	mean	sd	min	max
Courts are too harsh	54572	0.08	0.27	0.00	1.00
Courts are too lenient	54572	0.78	0.41	0.00	1.00
Very/Fairly important to respect America's laws etc	3772	0.94	0.24	0.00	1.00
Unaffiliated	64814	0.38	0.48	0.00	1.00
Republicans	64396	0.25	0.43	0.00	1.00
Democrats	64396	0.37	0.48	0.00	1.00
Female	64814	0.56	0.50	0.00	1.00
White	64814	0.80	0.40	0.00	1.00
School (years)	64637	12.87	3.18	0.00	20.00
N. children	64613	1.94	1.78	0.00	8.00
Married	64814	0.08	0.27	0.00	1.00
Year	1972-2018				

Table A4: GSS regressions results: DV

dep. var.:	Courts:				Law: Importance of			
	Too harsh		Too lenient		Always obey		To respect	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	-0.018***	-0.020***	0.039***	0.039***	0.100***	0.090***	0.021***	0.019**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)
	[-7.80]	[-8.69]	[11.11]	[11.23]	[7.68]	[7.09]	[2.74]	[2.45]
Unaffiliated		0.041***		-0.069***		-0.094***		-0.055***
		(0.00)		(0.00)		(0.02)		(0.01)
		[16.09]		[-16.20]		[-5.88]		[-6.14]
Democrats		0.037***		-0.067***		-0.071***		-0.025***
		(0.00)		(0.00)		(0.02)		(0.01)
		[13.92]		[-15.34]		[-4.23]		[-2.80]
White		-0.085***		0.074***		-0.120***		0.005
		(0.00)		(0.01)		(0.02)		(0.01)
		[-22.03]		[14.82]		[-7.24]		[0.44]
School (years)		0.001*		-0.009***		-0.034***		0.002*
		(0.00)		(0.00)		(0.00)		(0.00)
		[1.86]		[-14.25]		[-16.33]		[1.84]
N. children		-0.003***		0.003***		0.003		0.006**
		(0.00)		(0.00)		(0.00)		(0.00)
		[-4.39]		[3.07]		[0.68]		[2.40]
Married		-0.046***		0.036***		-0.013		0.027*
		(0.01)		(0.01)		(0.03)		(0.01)
		[-5.52]		[3.44]		[-0.51]		[1.87]
Constant	0.063***	0.115***	0.845***	0.899***	0.391***	0.969***	0.920***	0.896***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.04)	(0.01)	(0.02)
	[8.12]	[10.15]	[72.09]	[51.93]	[19.12]	[25.85]	[101.64]	[37.06]
Observations	54572	54056	54572	54056	5856	5802	3772	3742
Years	No	Yes	No	Yes	No	Yes	No	Yes

Notes: robust standard errors in parenthesis and t-statistics in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables are a dummies. *Omitted category*: republicans.