

The Effects of Trial Judge Gender and Public Opinion on Criminal Sentencing Decisions

*Christina L. Boyd**
*Michael J. Nelson***

We explore the effects of a trial judge's gender in criminal sentencing decisions by addressing two unsettled questions. First, do female and male trial judges sentence criminal offenders differently from one another? While numerous qualitative and quantitative scholars have examined this question, the results lack consistency. Second, are female trial judges' sentencing practices differentially affected by public opinion compared to male judges' behavior? Little research exists on this second question, but existing theory on how females and males make decisions and operate as judges is informative.

To provide new empirical insight into these questions, we rely on two sources of data: judge sentences stemming from Colorado trial court marijuana-related drug cases filed from 2004 to 2009 and local public opinion on marijuana from a 2006 Colorado general election initiative on whether to legalize marijuana possession. These data permit us to analyze judges' baseline sentencing practices (pre-2006 initiative) and the effect that public opinion has on the sentences (pre- vs. post-2006 initiative). The statistical modeling indicates that while male and female judges in Colorado generally do not sentence defendants differently from one another, there is one exception. Namely, female judges are more lenient than male judges when sentencing female defendants. Our empirical results also indicate that while Colorado trial judges were responsive to local public opinion following the 2006

* Associate Professor, Department of Political Science, University of Georgia. cLboyd@uga.edu.

** Assistant Professor, Department of Political Science, Pennsylvania State University. mjn15@psu.edu. We thank Tracey George, Michael Kang, Joanna Shepherd, Albert Yoon, David Lyle, the American Constitution Society for Law and Policy, and the *Vanderbilt Law Review* Board for organizing and sponsoring this Symposium and inviting us to participate. We appreciate the excellent feedback from the Symposium participants. We are grateful to the Colorado judicial branch for providing the data and answering questions about them.

marijuana initiative, that responsiveness was not more potent for female judges than it was for male judges. Together, these empirical results provide important new insights into the behavior of male and female trial court judges.

INTRODUCTION.....	1820
I. GENDERED JUDGING AND CRIMINAL SENTENCING	1822
A. <i>A Female Judge Difference in Sentencing?</i>	1822
B. <i>Data and Research Design</i>	1825
C. <i>Findings</i>	1828
II. GENDERED JUDICIAL RESPONSIVENESS TO PUBLIC	
OPINION?	1832
A. <i>Public Opinion and Judicial Behavior</i>	1832
B. <i>A Gendered Connection Between Public Opinion and</i> <i>Judicial Behavior?</i>	1836
C. <i>Data and Research Design</i>	1838
D. <i>Findings</i>	1839
CONCLUSION	1842

INTRODUCTION

“State trial judges have a great deal of authority and discretion over criminal prosecutions.”¹ What a judge does with that authority and discretion in criminal cases—including plea bargains, bench trials, evidentiary motions, the content of jury instructions, and sentencing—is likely to depend greatly on the judge’s background, preferences, and biases.² Moreover, most state trial judges are selected and/or retained through elections, meaning that there is a direct connection to and constraint from the public as well.³ In other words, the identity of the trial judge and the local culture in which she operates are almost certainly important in determining criminal case outcomes and rulings.

We turn our focus to how one particular judicial characteristic—a judge’s sex—might affect judicial behavior in criminal defendant sentencing decisions. In isolation, the fact that female and male judges

1. TRACEY E. GEORGE & ALBERT H. YOON, AM. CONSTITUTION SOC’Y FOR LAW & POLICY, THE GAVEL GAP: WHO SITS IN JUDGMENT ON STATE COURTS? 12, <http://gavelgap.org/pdf/gavel-gap-report.pdf> (last visited Oct. 2, 2017) [<https://perma.cc/67RE-AJ58>].

2. See, e.g., LAWRENCE BAUM, THE PUZZLE OF JUDICIAL BEHAVIOR (1997); LEE EPSTEIN & JACK KNIGHT, THE CHOICES JUSTICES MAKE 9–10 (1998); JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED 312–403 (2002).

3. E.g., Michael J. Nelson, *Uncontested and Unaccountable? Contestation Rates in Trial Court Elections*, 94 JUDICATURE 208, 210 (2011).

may sentence criminal defendants differently from one another has critical implications. As George and Yoon report, gender representativeness on state courts is low.⁴ As a result, a defendant's likelihood of drawing a female judge is greatly affected. This, in turn, can systematically lead to disproportionately lenient or harsh sentencing of defendants. New empirical analysis can help provide insight into the severity of this concern.

We also consider if and how a trial judge's sex intersects with public opinion pressures. Elected judges generally fear being viewed as "soft on crime,"⁵ particularly when their electorate has directly spoken on the issue through a ballot initiative.⁶ While it may be that all elected judges—male and female—equally fear the public's wrath, female judges may be particularly likely to seek voter approval and consensus in their criminal sentencing behavior. Once again, we look to empirical analysis to provide the answer.

We examine differences in responsiveness to public opinion among male and female judges through a study of trial court sentencing in marijuana cases in Colorado. In 2006, voters in Colorado rejected a proposed constitutional amendment that would have legalized the possession of small amounts of marijuana, leaving the legal status quo but providing local judges with constituency-specific information about the public's views toward marijuana legalization. With this unique vote, we have access to a constituency-level measure of public opinion on an issue that regularly comes before trial court judges and was publicly available to those judges as part of the normal process of reporting vote totals. This vote, therefore, sidesteps issues with measuring local-level public opinion that have plagued earlier studies, enabling us to assess responsiveness to public opinion with a measure of public opinion that is both valid and of particular relevance to elected judges, since it represents the views of those citizens who actually do turn out on election day.

In what follows, we tackle these questions in order. In Part I, we examine the behavior of female and male judges in sentencing criminal defendants. This includes an exploration of the theory and prior empirical analyses in this area, an overview of the empirical research

4. GEORGE & YOON, *supra* note 1, at 8.

5. Joanna Shepherd & Michael S. Kang, *Skewed Justice: Citizens United, Television Advertising and State Supreme Court Justices' Decisions in Criminal Cases*, SKEWED JUSTICE, <http://skewedjustice.org> (last visited Oct. 2, 2017) [<https://perma.cc/VR5D-78XT>].

6. James H. Kuklinski & John E. Stanga, *Political Participation and Government Responsiveness: The Behavior of California Superior Courts*, 73 AM. POL. SCI. REV. 1090, 1090–91 (1979); Michael J. Nelson, *Responsive Justice? Retention Elections, Prosecutors, and Public Opinion*, 2 J.L. & CTS. 117, 123–24 (2014).

design, and a presentation of our empirical results. In Part II, we continue with our dual examination of whether public opinion has a differential effect on male and female judges in their sentencing decisions. This begins with a general background on the effects of public opinion on judicial behavior followed by a more specific theoretical exploration of the judge gender–public opinion question. We then turn to a discussion of our data and research design and conclude with our original empirical analysis and findings. Finally, in our Conclusion, we discuss the implications of our findings for criminal defendants, state trial court judging, and future research in this area.

I. GENDERED JUDGING AND CRIMINAL SENTENCING

We begin our examination of whether male and female judges behave differently when sentencing criminal defendants by first detailing the theoretical arguments and prior empirical efforts on the subject. We then turn to an original empirical analysis.

A. *A Female Judge Difference in Sentencing?*

Scholars frequently study the question of whether female judges behave differently from male judges. The reasons are clear. A judiciary with a composition that descriptively represents society’s makeup “has a positive legitimizing effect on the functioning of a democracy.”⁷ Additionally, representative courts help to ensure that our laws are interpreted “by judges who can understand the circumstances of the communities which they serve.”⁸ Moreover, if male and female judges behave differently from one another while serving on the bench, important effects on judicial outputs may result. As Boyd argues:

Judges, and particularly trial judges, are well positioned to affect their assigned cases and the way that they progress. This has tremendous implications in trial courts for outcomes, settlements, costs, appeals, the distribution of resources after a case, and even the decision of litigants to file their cases and seek adjudicated remedies at all. In short, if diverse trial judges behave differently from their colleagues, we should expect substantive differences in the outputs of the judiciary.⁹

Numerous theoretical accounts have emerged to explain if, when, and why female judges will behave differently from their male

7. Thomas G. Walker & Deborah J. Barrow, *The Diversification of the Federal Bench: Policy and Process Ramifications*, 47 J. POL. 596, 597 (1985).

8. GEORGE & YOON, *supra* note 1, at 3.

9. Christina L. Boyd, *Representation on the Courts? The Effects of Trial Judges’ Sex and Race*, 69 POL. RES. Q. 788, 789 (2016).

colleagues. Notable accounts include different voice,¹⁰ informational,¹¹ representational,¹² and organizational theories.¹³ Each theory has been thoroughly discussed and debated in the literature (generally and as applied to judging).¹⁴

Empirical scholarship looking for gender-based results in judging behavior in appellate courts has been relatively uniform in its findings that judicial gender matters, but generally only in cases involving women's issues, like employment-based sexual discrimination and sexual harassment.¹⁵ On the trial court side, however, much less empirical consensus exists. Recent trial court work has highlighted the need to focus on individual issue areas and a variety of judge decisions beyond those published in the *Federal Supplement* or analogous state outlets when testing for the presence of a judge-gender effect.¹⁶

What about trial judge behavior in criminal cases, particularly the sentences handed down by trial court judges? Does a judge's gender systematically affect outcomes in these judicial decisions?¹⁷ The

10. Gilligan's different voice theory argues that males and females have distinct worldviews. Under this theory, males are committed to masculinity, logic, and justice, and females tend to value obligations, relationships, and personal communication. CAROL GILLIGAN, IN A DIFFERENT VOICE: PSYCHOLOGICAL THEORY AND WOMEN'S DEVELOPMENT (1993). As applied to judging, the different voice theory expects that these female-specific values affect female judges' decisions across "all aspects of society, whether or not they affect men and women differently." Suzanna Sherry, *The Gender of Judges*, 4 LAW & INEQ. 159, 160 (1986).

11. Under this theory, female judges bring unique knowledge and experience to the bench based on their shared professional backgrounds. See, e.g., Gerard Gryski et al., *Models of State High Court Decision Making in Sex Discrimination Cases*, 48 J. POL. 143, 145 (1986).

12. Under the representational theory, female judges make decisions to advance their interests and liberate other women. Beverly B. Cook, *Will Women Judges Make a Difference in Women's Legal Rights? A Prediction from Attitudes and Simulated Behaviour*, in WOMEN, POWER, AND POLITICAL SYSTEMS 216, 217 (Margherita Rendel ed., 1981).

13. Organizational theory asserts that all judges undergo the same professional training and experience before joining the bench and rely on the same laws and norms while making decisions. See Herbert M. Kritzer & Thomas M. Uhlman, *Sisterhood in the Courtroom: Sex of Judge and Defendant in Criminal Case Disposition*, 14 SOC. SCI. J. 77, 86 (1977).

14. See Christina L. Boyd et al., *Untangling the Causal Effects of Sex on Judging*, 54 AM. J. POL. SCI. 389, 390–91 (2010), for a review of the literature and empirical studies.

15. *Id.* at 389; see, e.g., SUSAN B. HAIRE & LAURA P. MOYER, DIVERSITY MATTERS: JUDICIAL POLICY MAKING IN THE U.S. COURTS OF APPEALS 34–35 (2015); Sean Farhang & Gregory Wawro, *Institutional Dynamics on the U.S. Court of Appeals: Minority Representation Under Panel Decision Making*, 20 J.L. ECON. & ORG. 299, 300 (2004); Jennifer L. Peresie, Note, *Female Judges Matter: Gender and Collegial Decisionmaking in the Federal Appellate Courts*, 114 YALE L.J. 1759, 1761 (2005); Nancy E. Crowe, *The Effects of Judges' Sex and Race on Judicial Decision Making on the U.S. Courts of Appeals, 1981–1996* (June 1999) (unpublished Ph.D. dissertation, University of Chicago) (on file with author).

16. Boyd, *supra* note 9, at 791.

17. Prior research highlights the high degree of prosecutorial discretion, including in making sentencing recommendations to judges. See, e.g., Nancy King, Commentary, *Three Directions for Future Research into Sentencing Discretion*, 30 JUST. Q. 223 (2013). Future empirical projects should seek to examine whether prosecutors behave differently and make different sentencing recommendations before male and female judges.

empirical research here continues the trend of other trial court–judicial diversity studies: it is a hodgepodge of findings. While some scholars find that there is little to no difference between male and female judge behavior in trying or sentencing defendants,¹⁸ a nontrivial number of studies in recent years have found that female judges are more punitive than male judges when deciding these cases¹⁹ or that female judges are more lenient than male judges when deciding these cases.²⁰

One explanation for the current state of the empirical criminal trial court–judge gender literature may lie with the disparities present in the underlying theoretical accounts and their expectations for female judges deciding criminal cases. For some issue areas, like sexual discrimination or harassment cases, three of the four notable theoretical accounts consistently expect female judges to behave differently from male judges.²¹ These accounts may vary in the underlying explanation of why a difference is expected, but the empirical implications are the same. By contrast, when considering judicial decisionmaking in criminal cases, the theoretical accounts are less optimistic in predicting a female judge effect. There is no reason to expect that female judges uniformly possess unique and valuable information about criminal cases (the informational account). As in other circumstances, the organizational account of judging continues to expect no differences among female and male judge behavior. Indeed, it is only the different voice account of judging that might anticipate different outcomes in criminal cases for all defendants based on the presiding judge’s sex. Recall that this theory’s “feminine perspective” to

18. *E.g.*, David S. Abrams et al., *Do Judges Vary in Their Treatment of Race?*, 41 J. LEGAL STUD. 347, 372–73 (2012); Richard Fox & Robert Van Sickle, *Gender Dynamics and Judicial Behavior in Criminal Trial Courts: An Exploratory Study*, 21 JUST. SYS. J. 261, 271 (2000); Claire S.H. Lim et al., *Do Judges’ Characteristics Matter? Ethnicity, Gender, and Partisanship in Texas State Trial Courts*, 18 AM. L. & ECON. REV. 302, 305 (2016); Gregory C. Sisk et al., *Charting the Influences on the Judicial Mind: An Empirical Study of Judicial Reasoning*, 73 N.Y.U. L. REV. 1377, 1453–54 (1998).

19. Lim et al., *supra* note 18; Cassia Spohn, *Decision Making in Sexual Assault Cases: Do Black and Female Judges Make a Difference?*, 2 WOMEN & CRIM. JUST. 83, 94 (1990); Darrell Steffensmeier & Chris Hebert, *Women and Men Policymakers: Does the Judge’s Gender Affect the Sentencing of Criminal Defendants?*, 77 SOC. FORCES 1163, 1181 (1999); Lydia Tiede et al., *Judicial Attributes and Sentencing-Deviation Cases: Do Sex, Race, and Politics Matter?*, 31 JUST. SYS. J. 249, 263–64 (2010); *see also* Fox & Van Sickle, *supra* note 18, at 270.

20. Brian D. Johnson, *Judges on Trial: A Reexamination of Judicial Race and Gender Effects Across Modes of Conviction*, 25 CRIM. JUST. POLY REV. 159, 174 (2014); Kenneth L. Manning, *¿Cómo Decide?: Decision-Making by Latino Judges in the Federal Courts*, Paper Presented at 2004 Midwest Political Science Association Annual Meeting (Apr. 15, 2004), http://citation.allacademic.com/meta/p_mla_apa_research_citation/0/8/3/3/9/pages83393/p83393-1.php [<https://perma.cc/2A9H-DSRQ>]. In related work, Bryna Bogoch, *Judging in a ‘Different Voice’: Gender and the Sentencing of Violent Offenses in Israel*, 27 INT’L J. SOC. L. 51, 62–65 (1999), finds that female Israeli trial judges are more lenient in their sentencing behavior.

21. In particular, different voice, representational, and informational theories predict this.

judging “encompasses all aspects of society.”²² In the context of criminal law, different voice theory would likely predict that “women’s emphasis on caring and their recognition of the importance of relationships would be expressed in a therapeutic rather than a punitive model.”²³ A more conditional prediction yields from the representational account of judging. Here, female judges are expected to behave in ways designed to protect the interests of women more broadly. In the context of criminal sentencing, this is likely to result in more lenient sentences, but only for female defendants.

Another potential explanation for the mixed bag of criminal trial court–judge gender findings may rest with previous data limitations. While a number of studies have tackled this topic, the same small set of data sources continue to be utilized. That has generally meant only studying Pennsylvania trial courts, Texas trial courts, local city trial courts,²⁴ or federal trial judge decisions published in the *Federal Supplement*. Turning to new data sources may provide additional insight.

B. Data and Research Design

In this study, we seek to provide new empirical insight into the question of whether female and male trial judges behave differently from one another. To do this, we look to novel data from the Colorado state trial courts.²⁵ Our data are unique in two major ways. First, unlike many prior studies of sentencing, which examine all criminal cases or violent crimes, we confine our analysis to marijuana cases, a set of cases about which the public has broadly differing views. Whereas there is broad public agreement that violent crime should be punished, the public differs widely about the appropriateness of criminalizing marijuana use and possession. Second, we are able to tie this sample of marijuana cases to a unique vote, explained below, that provided judges with information about their constituents’ opinions on marijuana.

Colorado has a two-tiered trial court system: county courts and district courts. Each county has its own county court (64 in total). The counties are also grouped into 22 judicial districts, with each judicial district housing a district court. There is some jurisdictional overlap between the two types of trial courts, but the district courts generally

22. Sherry, *supra* note 10, at 160.

23. Bogoch, *supra* note 20, at 68.

24. John Gruhl et al., *Women as Policymakers: The Case of Trial Judges*, 25 AM. SCI. 308, 311–12 (1981) (studying “Metro City” felonies); Spohn, *supra* note 19, at 88 (studying Detroit Recorder’s Court cases).

25. Nelson, *supra* note 6, at 118.

hear the more severe cases.²⁶ County and district trial judges are initially selected for a two-year term, via merit criteria, through a commission process.²⁷ Thereafter, the judges face uncontested retention elections. Subsequently, the judges face similar retention elections at the end of each four-year (for county court judges) or six-year (for district court judges) term.²⁸

We focus on the 2,589 cases decided in 2004 and 2005 in Colorado county and district courts alleging a violation of Colorado's prohibition on the usage of marijuana.²⁹ The data for this study come from records kept by the Colorado judicial branch.³⁰

We test for a gender effect in sentencing in three ways. First, we analyze the full set of sentences handed down by Colorado trial court judges, using as our dependent variable a traditional sentence severity scale developed by the United States Bureau of Justice.³¹ The scale assigns point values to different types of sentences. For example, fines and unsupervised probation receive zero points. A sentence of supervised probation is worth one point if it is less than one year in length, two points if it is between one and three years in length, and four points if it is over three years in length. Similarly, a prison sentence of less than one month is worth one point, while a prison sentence of seven to twelve months is worth five points.³² The measure is facially valid, declining in value with the severity of the charge: the average value of this variable for felonies, misdemeanors, and petty offenses, respectively, is 12.06, 7.01, and 3.71.

Second, for a more intuitive metric of sentence severity, we subset the data into only the most severe sentences. Examining only felonies, we model the number of days of incarceration in the sentence received by the defendant. Because this variable is highly skewed, we

26. U.S. DEPT OF JUSTICE, BUREAU OF JUSTICE STATISTICS, NCJ 212351, STATE COURT ORGANIZATION 2004, at 273 (2006).

27. *Judicial Selection in the States: Appellate and General Jurisdiction Courts*, AM. JUDICATURE SOC'Y (2013), http://www.judicialselection.us/uploads/documents/Judicial_Selection_Charts_1196376173077.pdf [<https://perma.cc/K2BB-WDQP>].

28. *Id.*

29. COLO. REV. STAT. § 18-18-406 (2016), *amended by* 2017 Colo. Sess. Laws ch. 401 (titled "Offenses relating to marijuana and marijuana concentrate").

30. The data analyzed do not include any cases from the city of Denver. The data received from the Colorado judicial branch do not include cases heard by the Denver County Courts. These courts are under the purview of the city of Denver and not the Colorado judicial branch.

31. This is the same scale used by Nelson, *supra* note 6, at 142. For other uses of this scale, see Beverly B. Cook, *Public Opinion and Federal Judicial Policy*, 21 AM. J. POL. SCI. 567, 572 (1977); James L. Gibson, *Judges' Role Orientations, Attitudes, and Decisions: An Interactive Model*, 72 AM. POL. SCI. REV. 911, 913 (1978); and Herbert M. Kritzer, *Federal Judges and Their Political Environments: The Influence of Public Opinion*, 23 AM. J. POL. SCI. 194, 205 (1979).

32. For the full scale, see Nelson, *supra* note 6, at 142.

follow standard practice and use as the dependent variable $\ln(1 + \text{Days of Incarceration})$.³³

Finally, we examine the least severe sentences in the data—the petty offenses—separately. As Nelson notes, such an examination has the benefit of examining the class of offenders who are least likely to have a prior criminal record, thereby offering some insurance against the inability of the other statistical analyses to control for variation in the defendant's prior criminal record.³⁴ Here, we take a similar strategy to that used for the felonies, modeling $\ln(1 + \text{Fine Levied})$ as the dependent variable.

Our primary independent variable for this study is the sentencing judge's sex. Coded using the judges' first names, this variable is measured 1 for female judges and 0 for male judges. The data contain sentences handed down by 106 male judges and 40 female judges.

We also control for a variety of political and legal factors that may affect a judge's behavior when sentencing a defendant. For example, trial judges may be swayed in their decisionmaking by the public's overwhelming liberal (or conservative) tendencies³⁵ or their own ideological proclivities.³⁶ To measure the public's ideological position, we include the percentage of the vote won by the Democratic candidate in the closest U.S. presidential election (Constituent Ideology). To capture the judge's ideological position, we include another variable (Republican) to indicate whether the judge was appointed by a Republican (= 1) or a Democratic governor (= 0).³⁷ The predictions from both political variables are straightforward: judges with more conservative constituencies and judges appointed by Republican governors should sentence defendants more harshly in marijuana cases.

Turning to legal factors that may affect judicial behavior in these sentencing cases, we include indicator variables for the severity of the

33. Sanford C. Gordon & Gregory A. Huber, *The Effect of Electoral Competitiveness on Incumbent Behavior*, 2 Q.J. POL. SCI. 107, 130 (2007).

34. Nelson, *supra* note 6, at 146.

35. We explore in detail below why public opinion may affect judicial behavior in these cases. Note, however, that prior to 2006, Colorado state judges had little direct and specific information on how their local publics felt about marijuana-related issues.

36. See, e.g., SEGAL & SPAETH, *supra* note 2, at 86–111 (detailing how judicial ideology affects judicial behavior).

37. It would have been ideal to have judicial-ideology measures similar to Lee Epstein et al., *The Judicial Common Space*, 23 J.L. ECON. & ORG. 303, 306–10 (2007), or the PAJID scores created by Paul Brace et al., *Measuring the Preferences of State Supreme Court Judges*, 62 J. POL. 387, 398 (2000). However, no such measure exists for local-level state actors.

offense: felony, misdemeanor, or petty.³⁸ We also control for whether the case was heard by a district court or a county court. We also account for important defendant characteristics, including age, gender, and minority status.³⁹ We expect that younger, nonwhite, and male defendants will, on average, receive higher sentences.⁴⁰ To determine whether female judges sentence female defendants differently than male defendants, we include a multiplicative interaction term between the gender of the judge and the gender of the defendant. Sixteen percent of the sentences in the data were given to female defendants. We model the data using linear regressions with random intercepts for county and judge.⁴¹

C. Findings

Table 1 displays the results of all three linear regressions. The first column of the table displays the model estimates for all sentences using the sentence severity scale as the dependent variable. The second column of the table models the amount of incarceration time received by defendants convicted of felonies, and the final column of the table models the amount of the fine received by defendants convicted of petty offenses.

In all three models, the coefficient for female judge fails to attain statistical significance. As such, the conclusion is clear: for the average case, there is no evidence that male and female judges in Colorado systematically sentence differently from one another. The average defendant in our data is a twenty-four-year-old white male convicted of a petty offense. For him, our results indicate that the predicted sentence severity is 2.75 points when he is sentenced by a male judge and 2.58 points when he is sentenced by a female judge.⁴²

38. See, e.g., King, *supra* note 17, at 227, for a discussion on the need to differentiate misdemeanors from other offenses. King notes that “[o]nly a small percentage of all misdemeanor convictions result in an initial sentence to incarceration.” *Id.*

39. This study does not control for the defendant’s prior criminal history. Unfortunately, the Colorado judicial system did not keep these records with their other rich data.

40. For more on the importance of defendant characteristics in the criminal justice system and the case outcomes that emerge, see, e.g., Debra A. Curran, *Judicial Discretion and Defendant’s Sex*, 21 CRIMINOLOGY 41, 54–56 (1983); Darrell Steffensmeier et al., *The Interaction of Race, Gender, and Age in Criminal Sentencing: The Punishment Cost of Being Young, Black, and Male*, 36 CRIMINOLOGY 763, 763–66, 788 (1998); and James D. Unnever et al., *Race Differences in Criminal Sentencing*, 21 SOC. Q. 197, 204–05 (1980).

41. The data suggest substantively identical conclusions for a variety of other modeling strategies, including negative binomial regressions.

42. These predictions are based on the estimates from Model 1, Table 1.

TABLE 1: MODELING SENTENCE SEVERITY IN MARIJUANA
CONVICTIONS, COLORADO 2004–2005

	All Sentences	Felonies	Petty Offenses
Female Judge	-0.18 (0.27)	0.26 (0.47)	0.15 (0.11)
Republican Appointee	-0.54* (0.23)	-0.47 (0.37)	-0.20 (0.11)
Defendant's Age	0.00 (0.01)	0.01 (0.01)	0.00 (0.00)
Constituent Ideology	0.00 (0.02)	-0.03 (0.03)	0.00 (0.00)
District Court	0.23 (0.28)	2.33 (2.89)	-0.44* (0.11)
Minority Defendant	0.48* (0.16)	0.48 (0.37)	-0.02 (0.06)
Female Defendant	0.03 (0.15)	-0.75* (0.43)	-0.01 (0.06)
Misdemeanor	-4.80* (0.19)	— —	— —
Petty Offense	-8.00* (0.26)	— —	— —
Constant	11.74* (0.75)	1.49 (3.11)	5.26* (0.26)
σ_{Judge}	0.95	0.30	0.13
σ_{District}	0.72	0.52	0.11
N	2589	379	1436
BIC	12928.80	1934.17	3485.42

Standard errors are in parentheses, and * indicates statistical significance at $p < 0.05$. Felonies are the baseline category in the first model (column 1).

TABLE 2: MODELING SENTENCE SEVERITY IN MARIJUANA
CONVICTIONS, COLORADO 2004–2005

	All Sentences	Felonies	Petty Offenses
Female Judge	-0.01 (0.27)	0.63 (0.50)	0.16 (0.11)
Female Defendant	0.26 (0.17)	-0.17 (0.50)	0.01 (0.06)
Female Judge x Female Defendant	-1.03* (0.36)	-2.26* (0.98)	-0.11 (0.14)
Defendant's Age	0.00 (0.01)	0.01 (0.01)	0.00 (0.00)
Minority Defendant	0.48* (0.16)	0.50 (0.36)	-0.02 (0.06)
Constituent Ideology	0.00 (0.01)	-0.03 (0.03)	0.00 (0.01)
District Court	0.21 (0.28)	2.61 (2.88)	-0.44* (0.11)
Republican Appointee	-0.52* (0.23)	-0.45 (0.37)	-0.20 (0.11)
Misdemeanor	-4.80* (0.19)	— —	— —
Petty Offense	-8.02* (0.26)	— —	— —
Constant	11.73* (0.76)	1.05 (3.11)	5.26* (0.26)
σ_{Judge}	0.97	0.32	0.13
σ_{District}	0.74	0.53	0.11
N	2589	379	1436
BIC	12928.52	1933.00	3494.15

Standard errors are in parentheses, and * indicates statistical significance at $p < 0.05$. Felonies are the baseline category in the first model (column 1).

The control variables behave inconsistently across models. There is some evidence that female defendants receive lighter

sentences—about one year less—when convicted of felonies and that nonwhite defendants receive harsher sentences (about half of a point on the sentence severity scale) overall. Surprisingly, Republican-appointed judges tend to sentence *less* punitively than judges appointed by Democratic governors, though this effect is only statistically significant in one model.

A follow-up question to ask of these data concerns the representational effects of female judges; perhaps female defendants are advantaged when they appear before a female judge because female judges, on average, give lighter sentences to female defendants. To this end, Table 2 replicates the models from Table 1, including a multiplicative interaction term between the gender of the judge and the gender of the defendant.

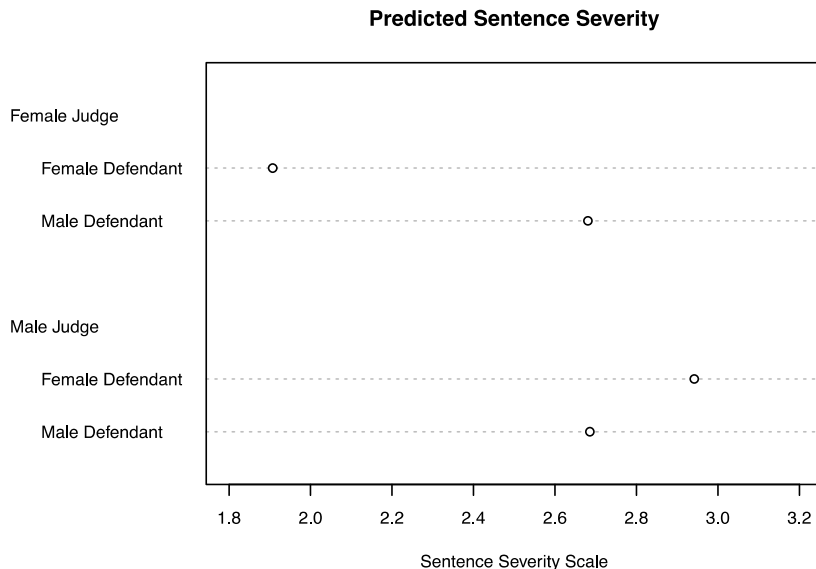
Both overall and in felony cases, the multiplicative interaction term is statistically significant, providing some evidence of a differential effect. However, the substantive effect is not readily apparent from Table 2. To provide more insight into this substantive effect, Figure 1 plots the marginal effects of judge and defendant gender on sentence severity.

Looking first at all charges, Figure 1's marginal effect of being sentenced by a female judge is 0.00 for male defendants. In other words, male and female judges, all else equal, sentence male defendants identically. Of course, this effect is not statistically significant. For female defendants, the marginal effect of being sentenced by a female judge is -1.03, or a full point less on the sentence severity scale. This effect is statistically significant ($p = 0.01$).⁴³

From these analyses, we draw two major conclusions. First, there is no evidence that male and female judges differ in their average punitiveness. This is true for all charges as well as for the most and least severe offenses. Second, strong evidence for a gendered effect in sentencing emerges from a representational aspect: here, female judges are significantly more lenient on female defendants than male judges.

43. For felonies specifically, the pattern is similar: the marginal effect of being sentenced by a female judge is very close to zero for male defendants but is larger and negative, accounting for a difference of about one year of incarceration, for female defendants. The marginal effect of this difference is statistically significant at the $p = 0.09$ level.

FIGURE 1: MARGINAL EFFECTS OF JUDGE AND DEFENDANT GENDER ON SENTENCE SEVERITY



This figure displays predicted values of sentence severity for combinations of defendant and judge sex. The model estimates come from Model 1 in Table 2 and hold all other covariates at their median (for interval-level variables) or modal (for categorical variables) values. The modal charge type is a petty offense. The figure demonstrates that female judges sentencing female defendants tend to give the lightest sentences.

II. GENDERED JUDICIAL RESPONSIVENESS TO PUBLIC OPINION?

Using the above baseline for the judicial behavior of male and female trial judges in criminal sentencing cases, we now assess whether male and female judges respond differently to signals about their constituents' opinions. In what follows, we discuss the general connection between public opinion and judicial behavior and whether this effect is likely to be stronger for female judges than male judges. We then turn to an empirical test of this potential effect on Colorado state trial judges.

A. Public Opinion and Judicial Behavior

Federal and state judges are affected by public opinion, regardless of the institution used to select or retain them. At the federal level, where judges will never directly face the electorate, scholars have argued that judges attempt to follow public opinion to bolster their

legitimacy.⁴⁴ These studies generally assess the extent to which the liberal or conservative nature of court outcomes coincide with changes in the general ideological leanings of the public.⁴⁵

At the state level, there is an even stronger judicial connection to the public. The vast majority of state judges face voters to retain their positions on the bench.⁴⁶ This creates a strong electoral connection and attendant incentive to remain alert to and follow public opinion. Scholars of the state judiciary have found that elected judges are generally responsive to the people in their state.⁴⁷ This is true even in states where judges are retained using uncontested retention elections.⁴⁸

While judges' individual decisions in cases are generally hidden from widespread public scrutiny, there is still reason to believe that the public's preferences will constrain judicial behavior. Within congressional elections and legislative politics, where incumbents have at least a 90% chance of reelection, scholars argue that the mere possibility of electoral defeat colors nearly every action that legislators take.⁴⁹ The same is true for judicial elections. Interest groups organize

44. See Bryan Calvin et al., *On the Relationship Between Public Opinion and Decisionmaking in the U.S. Courts of Appeals*, 64 POL. RES. Q. 736, 743 (2011); Lee Epstein & Andrew D. Martin, *Does Public Opinion Influence the Supreme Court? Possibly Yes (But We're Not Sure Why)*, 13 U. PA. J. CONST. L. 263, 280 (2010); Roy B. Flemming & B. Dan Wood, *The Public and the Supreme Court: Individual Justice Responsiveness to American Policy Moods*, 41 AM. J. POL. SCI. 468, 492–94 (1997); Kevin T. McGuire & James A. Stimson, *The Least Dangerous Branch Revisited: New Evidence on Supreme Court Responsiveness to Public Preferences*, 66 J. POL. 1018, 1023 (2004).

45. See Calvin et al., *supra* note 44, at 739; Epstein & Martin, *supra* note 44, at 271–72; Flemming & Wood, *supra* note 44, at 472–74; McGuire & Stimson, *supra* note 44, at 1027–28.

46. Nelson, *supra* note 3, at 209.

47. See, e.g., Paul Brace & Brent D. Boyea, *State Public Opinion, the Death Penalty, and the Practice of Electing Judges*, 52 AM. J. POL. SCI. 360, 370 (2008). Notably, however, the type of judicial retention institution used conditions the magnitude of the effect. See, for example, Richard P. Caldarone et al., *Partisan Labels and Democratic Accountability: An Analysis of State Supreme Court Abortion Decisions*, 71 J. POL. 560, 571 (2009), which finds that judges retained through nonpartisan elections are more likely than partisan-elected judges to issue abortion decisions that comport with public opinion.

48. See Brandice Canes-Wrone et al., *Judicial Independence and Retention Elections*, 28 J.L. ECON. & ORG. 211, 229 (2012) (demonstrating that, in abortion cases, retention elections induce a responsiveness to public opinion similar to that experienced by judges in contestable elections); Carol Ann Traut & Craig F. Emmert, *Expanding the Integrated Model of Judicial Decision Making: The California Justices and Capital Punishment*, 60 J. POL. 1166, 1176–79 (1998) (finding that public opinion plays a role in California Supreme Court justices' death penalty decisions even though those judges face uncontested retention elections). For state trial courts, trial court judges retained through uncontested retention elections are generalizable to most trial judges in partisan and nonpartisan contested election states. See, for example, Nelson, *supra* note 3, at 212, which indicates that state trial judges (on courts of general jurisdiction) in these states rarely face a challenger in their primary or general elections.

49. See, e.g., DAVID R. MAYHEW, CONGRESS: THE ELECTORAL CONNECTION (2d ed. 2004) (asserting that politicians are primarily motivated by the goal of reelection); R. Douglas Arnold, *Can Inattentive Citizens Control Their Elected Representatives?*, in CONGRESS RECONSIDERED 401,

campaigns against judicial retention or reelection.⁵⁰ Judges are aware when their colleagues are not retained and, in response, are more cognizant of the political environment in their future behavior.⁵¹

When it comes to crime-related issues, the public–judicial behavior connection is particularly strong. Baum has argued that “creating the impression that a judge is soft on crime can have great electoral impact.”⁵² Shepherd and Kang similarly assert that “[a]t the margin, whether consciously or unconsciously, [elected judges] prefer to avoid a judicial vote in a criminal case that can be the basis for attack advertisements funded by independent expenditures.”⁵³ Hall has similarly asserted that “[i]n judicial elections, arguably the most important policy focus for voters is crime.”⁵⁴ Why might this be the case? By their very nature, criminal cases are “brought in the name of the government on behalf of the community.”⁵⁵ In other words, these cases place the mass public and the shared values of society on trial.

The empirical evidence supports this constraint. Hall finds that state supreme court justices overturn fewer death penalty sentences when the murder rate is high.⁵⁶ When it comes to trial court judicial sentencing practices, previous work finds a strong public opinion constraint. Huber and Gordon find that, as the date of their retention election approaches, Pennsylvania judges hand down more punitive sentences.⁵⁷ Similarly, Berdejo and Yuchtman demonstrate that Washington state trial court judges who stand for contestable nonpartisan elections issue sentences that are around 10% longer at the

411–15 (Lawrence C. Dodd & Bruce I. Oppenheimer eds., 5th ed. 1993) (arguing that, due to a fear of electoral retribution, legislators consider the potential preferences of even inattentive voters).

50. See Nelson, *supra* note 6, at 122 (noting that interest groups may inform voters about a judge’s unpopular decisions even in uncontested elections); Roy A. Schotland, *Iowa’s 2010 Judicial Election: Appropriate Accountability or Rampant Passion?*, 46 CT. REV. 118, 119–20 (2011) (outlining the increase in interest group spending aimed at informing voters about a controversial judicial decision made by judges who face retention elections).

51. James L. Gibson, *Environmental Constraints on the Behavior of Judges: A Representational Model of Judicial Decision Making*, 14 LAW & SOC’Y REV. 343, 365–67 (1980) (asserting that judges who had previously been defeated in an election were more receptive to environmental influences).

52. Lawrence Baum, *Judicial Elections and Judicial Independence: The Voter’s Perspective*, 64 OHIO ST. L.J. 13, 35 (2003).

53. Shepherd & Kang, *supra* note 5.

54. Melinda Gann Hall, *State Supreme Courts in American Democracy: Probing the Myths of Judicial Reform*, 95 AM. POL. SCI. REV. 315, 322 (2001).

55. STEPHEN A. SALTZBURG & DANIEL J. CAPRA, *AMERICAN CRIMINAL PROCEDURE* 2 (5th ed. 1996).

56. Melinda Gann Hall, *Justices as Representatives: Elections and Judicial Politics in the American States*, 23 AM. POL. Q. 485, 497 (1995).

57. Gregory A. Huber & Sanford C. Gordon, *Accountability and Coercion: Is Justice Blind When It Runs for Office?*, 48 AM. J. POL. SCI. 247, 255 (2004).

end of a judge's term than at the beginning.⁵⁸ Kuklinski and Stanga find that the sentencing behavior of California trial court judges changed drastically after a failed 1972 marijuana legalization initiative in the state.⁵⁹ Much more recently, Nelson finds that Colorado state trial judges adjusted their sentencing behavior in response to local public opinion about marijuana revealed by a 2006 legalization initiative in the state.⁶⁰

While theory leads us to expect that elected judges should reflect public opinion in their decisions, they can only do so to the extent that they know what that public opinion actually is. In the absence of local (e.g., county) polling data, judges must look to other sources. Nelson describes judges' typical alternative sources for this information and their inherent limitations:

Through conversations with friends, colleagues, the local bar association, and the alleged criminals with whom they work, judges and prosecutors gather some information about how their geographic constituents feel about issues. Still, while they may try to ascertain their constituents' preferences by talking to their friends and neighbors, they risk the chance that those individuals are a homogeneous segment of their constituency with a biased view of overall constituency opinion.⁶¹

On rare occasions—such as those leveraged in the Kuklinski and Stanga study⁶² using 1972 California data, or the Nelson study⁶³ using the 2006 Colorado data—judges have high-quality, widely available local precinct public opinion information to rely on. In these situations, judges' quests to follow public opinion are undoubtedly eased and improved.

58. Carlos Berdejo & Noam Yuchtman, *Crime, Punishment, and Politics: An Analysis of Political Cycles in Criminal Sentencing*, 95 REV. ECON. & STAT. 741, 742 (2013).

59. Kuklinski & Stanga, *supra* note 6, at 1093. Kuklinski and Stanga compare the aggregate sentences produced by each California county's superior court before and after the 1972 initiative. They find that courts whose constituents favored the legalization of marijuana sentenced defendants more leniently than the other trial courts. *Id.*

60. Nelson, *supra* note 6, at 118. Like Kuklinski and Stanga, *supra* note 6, at 1091, Nelson's study utilizes a failed marijuana legalization initiative to provide a strong, issue-specific signal of public preferences. Nelson, *supra* note 6, at 126. Unlike Kuklinski and Stanga, *supra* note 6, at 1092, Nelson's data provide judge-level voting behavior rather than more aggregated and noisy court-level statistics. Nelson, *supra* note 6, at 136.

61. Nelson, *supra* note 6, at 123–24.

62. Kuklinski & Stanga, *supra* note 6, at 1091.

63. Nelson, *supra* note 6, at 126.

B. A Gendered Connection Between Public Opinion and Judicial Behavior?

To our knowledge, little prior empirical work assesses whether similarly situated female and male elites respond differently to public opinion. Given the importance of public opinion to judges, as described above, we believe this is an oversight in need of remedy. We see two related theoretical avenues that may anticipate that female judges will be more likely than their male counterparts to court the public's favor.

First, female judges may be more likely to actively seek approval and acceptance from voters than male judges. Previous work finds that females are more likely to adopt democratic management styles, while males favor autocratic ones.⁶⁴ In the context of studying federal district court judges, Boyd argues that

a female style of management is characterized by a decision-making environment that is more likely to encourage participation among subordinates, democratic communication, collaboration, consensus building, and the drawing of multiple voices into deliberative processes. . . . Conversely, a traditional male style of management is one that disfavors this type of participative environment and instead sees the male decision maker seizing opportunities to exercise his authority over outcomes.⁶⁵

While these gender-specific styles are most likely to manifest themselves when judges are in leadership and management positions, like serving as a court's chief judge or managing a trial court case, they may also emerge more broadly. In other words, female judges may be more likely to seek democratic consensus in their behavior (via public opinion), whereas male judges may prefer to "exercise their authority" over a decision by following their own preferences to the exclusion or limitation of public opinion's influence.

Second, a bevy of research studies find that females are frequently devalued in job application and evaluation processes compared to equally qualified men.⁶⁶ This work indicates, for example, that "evaluators may subtly shift the criteria they use to make hiring decisions to benefit gender or race typical applicants."⁶⁷ These studies

64. Alice H. Eagly & Blair T. Johnson, *Gender and Leadership Style: A Meta-analysis*, 108 PSYCHOL. BULL. 233, 236 (1990).

65. Christina L. Boyd, *She'll Settle It?*, 1 J.L. & CTS. 193, 196–97 (2013) (citations omitted) (citing Eagly & Johnson, *supra* note 64; Judy B. Rosener, *Ways Women Lead*, 68 HARV. BUS. REV. 119 (1990); Cindy Simon Rosenthal, *Determinants of Collaborative Leadership: Civic Engagement, Gender or Organizational Norms?*, 51 POL. RES. Q. 847 (1998)).

66. *E.g.*, Alice H. Eagly & Steven J. Karau, *Role Congruity Theory of Prejudice Toward Female Leaders*, 109 PSYCHOL. REV. 573 (2002); Madeline E. Heilman & Michelle C. Haynes, *No Credit Where Credit Is Due: Attributional Rationalization of Women's Success in Male-Female Teams*, 90 J. APPLIED PSYCHOL. 905 (2005).

67. Julie E. Phelan et al., *Competent Yet Out in the Cold: Shifting Criteria for Hiring Reflect Backlash Toward Agentic Women*, 32 PSYCHOL. WOMEN Q. 406, 407 (2008).

also find that “when both the in-group [here, males] and out-group [here, females] are assessed on the same criteria, discrimination that favors the in-group is likely,”⁶⁸ and that “status stereotypes often result in diminished expectations of competence” for women and racial minorities.⁶⁹ Among other settings, this gender-specific bias has been known to materialize in elections. Sanbonmatsu finds that voters rely on “stereotypes about men and women in society” to develop voting preferences.⁷⁰

In the face of this potential evaluation bias among voters, elected female judges may double their efforts (or “jump through more hoops”⁷¹) to prove their judicial capabilities and avoid failure at the ballot box. This exceptional female effort can result in many different types of judicial outputs. One such output may be that female judges are more cognizant of public opinion. In the context of uncontested retention trial court elections, this may lead female trial judges to be more attentive to the public’s preferences than their male colleagues. And, in the face of strong evidence of those preferences, they may be more responsive to those preferences in their decisionmaking behavior.

Despite these pro-difference explanations, there are plenty of reasons to expect that, all other things equal, male and female judges will be equally responsive to public opinion pressures. One such powerful reason lies with the constraining effect of the public for all elected judges—males and females. All elected judges, even those who face uncontested retention elections, are likely to be mindful of the public’s preferences.⁷² This also fits with the larger literature that finds incumbent elected officials are still strongly “reelection minded” even when the odds of losing reelection are slim.⁷³ While this literature has not separately tested this effect for male and female politicians, historical trends in those elected offices indicate that most of those office holders were male.⁷⁴

68. Janis V. Sanchez-Hucles & Donald D. Davis, *Women and Women of Color in Leadership: Complexity, Identity, and Intersectionality*, 65 AM. PSYCHOLOGIST 171, 177 (2010).

69. Robert K. Christensen et al., *Race and Gender Bias in Three Administrative Contexts: Impact on Work Assignments in State Supreme Courts*, 22 J. PUB. ADMIN. RES. & THEORY 625, 627 (2012).

70. Kira Sanbonmatsu, *Gender Stereotypes and Vote Choice*, 46 AM. J. POL. SCI. 20, 22 (2002).

71. Monica Biernat & Diane Kobrynowicz, *Gender- and Race-Based Standards of Competence: Lower Minimum Standards but Higher Ability Standards for Devalued Groups*, 72 J. PERSONALITY & SOC. PSYCHOL. 544, 554 (1997).

72. See, e.g., Traut & Emmert, *supra* note 48.

73. MAYHEW, *supra* note 49, at 37; Arnold, *supra* note 49, at 408–09.

74. See, e.g., ROGER H. DAVIDSON ET AL., CONGRESS AND ITS MEMBERS 68 (16th ed. 2017).

Another reason not to expect a public opinion difference rests with the above-described organizational theory.⁷⁵ Organizational theory expects that all judges, male and female, underwent the same legal training and background experiences before receiving their judgeships. Those things, rather than sex, are likely to affect behavior and decisionmaking responsiveness for elected judges.

C. Data and Research Design

To test the effect of public opinion on trial judge behavior with a specific focus on whether female and male trial judges respond differently to that public opinion, we turn once again to Nelson's Colorado data on trial judge behavior in issuing marijuana-related criminal sentences in the 2000s.⁷⁶ Here, we focus specifically on whether Colorado's 2006 marijuana legalization initiative results affected male and female judges in distinct ways.

The Colorado general election ballot in 2006 included the following initiative: the ballot asked, "Shall there be an amendment to section 18-18-406 (1) of the Colorado revised statutes making legal the possession of one ounce or less of marihuana for any person twenty-one years of age or older?"⁷⁷ The proposal failed (with 41% support).⁷⁸ However, as depicted in Figure 2, county-level support for the measure varied considerably.

We follow the same three-pronged approach to estimate differences in the responsiveness of male and female judges. First, we examine all sentences handed down by Colorado trial court judges, using the sentence severity scale as the outcome variable for that model. Second, we subset the data to felonies alone and model the amount of incarceration time as the dependent variable. Finally, we look to petty offenses, modeling the amount of the fine received by the defendant as the dependent variable.

Because our concept of interest is differential responsiveness to the initiative results by gender, the key independent variable is a three-way interaction between the gender of the judge, whether the sentence was handed down before or after the initiative vote (Post-Initiative),

75. Kritzer & Uhlman, *supra* note 13, at 87.

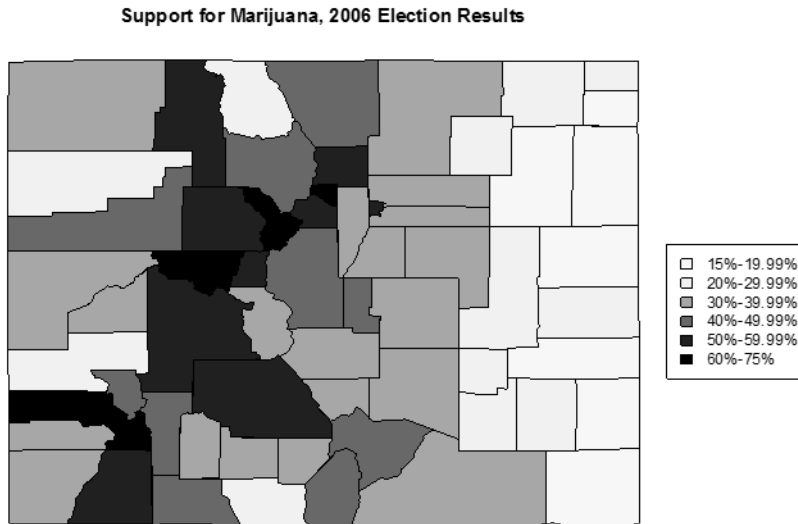
76. Nelson, *supra* note 6, at 127–33.

77. *Id.* at 126; *Official Publication of the Abstract of Votes Cast for the 2005 Coordinated, 2006 Primary, 2006 General, ST. COLO. 156*, <https://www.sos.state.co.us/pubs/elections/Results/Abstract/pdf/2000-2099/2006AbsractBook.pdf> (last visited Oct. 3, 2017) [<https://perma.cc/5R4T-5YNK>] [hereinafter *Colorado Vote Summary*].

78. Nelson, *supra* note 6, at 126. Participation and turnout were high. Nearly 98% of ballots included a vote on the marijuana measure, and two-thirds of active Colorado voters turned out in the election. *Id.*; *Colorado Vote Summary, supra* note 77, at 144, 163.

and the results of the initiative (Initiative Support). Because the initiative sought to *legalize* small amounts of marijuana, higher values of this latter variable indicate more pro-marijuana views.

FIGURE 2: COLORADO COUNTY-LEVEL SUPPORT FOR MARIJUANA BASED ON THE 2006 BALLOT MEASURE



Like in our statistical modeling above, we continue to control for other political and legal factors that may affect judicial sentencing behavior. This includes the following previously discussed measures: Constituent Ideology, Republican, severity of the offense, whether the case was heard by a district or county court, and defendant characteristics.

D. Findings

Table 3 displays the results of the analysis. Again, the three columns in the table correspond to the three separate linear regressions: the sentence severity scale for all charges, incarceration time for felonies, and fines for petty offenses.

TABLE 3: MODELING SENTENCE SEVERITY IN MARIJUANA CONVICTIONS,
COLORADO 2004–2005

	All Sentences	Felonies	Petty Offenses
Female Judge	0.20 (0.90)	4.35 (4.15)	0.76 (0.56)
Post-Initiative	1.42* (0.36)	4.19* (2.08)	1.30* (0.18)
Initiative Support	-0.05* (0.02)	-0.10 (0.06)	0.00 (0.01)
Initiative Support x Post-Initiative	-0.02 (0.01)	-0.09 (0.05)	-0.01* (0.01)
Initiative Support x Female Judge	0.00 (0.02)	-0.09 (0.10)	-0.01 (0.01)
Post-Initiative x Female Judge	-0.01 (0.62)	-3.40 (4.02)	-0.13 (0.29)
Initiative Support x Post-Initiative x Female Judge	0.00 (0.01)	0.06 (0.10)	0.00 (0.01)
Republican Appointee	-0.66* (0.18)	-0.47 (0.51)	-0.29* (0.13)
Constituent Ideology	0.04* (0.01)	0.04 (0.05)	0.02* (0.01)
District Court	0.04 (0.20)	8.45 (6.28)	-0.54* (0.11)
Minority Defendant	0.27* (0.09)	1.03* (0.40)	-0.04 (0.05)
Female Defendant	-0.05 (0.09)	-1.54* (0.53)	0.09 (0.05)
Defendant's Age	0.00 (0.00)	0.04* (0.02)	0.00 (0.00)
Misdemeanor	-5.07* (0.11)	— —	— —
Petty Offense	-8.67* (0.16)	— —	— —
Constant	12.29*	4.62	2.28*

	All Sentences	Felonies	Petty Offenses
	(0.71)	(6.64)	(0.47)
σ_{Judge}	0.85	1.38	0.31
σ_{District}	0.56	1.05	0.67
N	9,099	1,497	5,108
BIC	47547.62	9837.94	17379.42

Standard errors are in parentheses, and * indicates statistical significance at $p < 0.05$. Felonies are the baseline category in the first model (column 1).

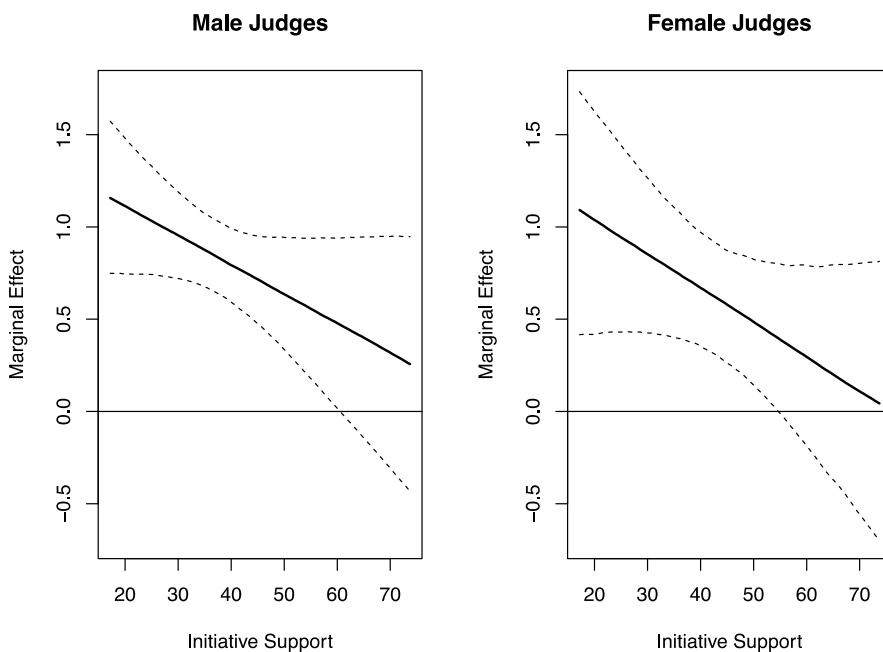
With so many multiplicative interaction terms in the model, it is easiest to understand the results of the analysis with a figure. Figure 3 displays the marginal effect of deciding a case after the initiative for both male and female judges across the range of initiative support. Just as Nelson previously showed,⁷⁹ there is strong evidence that judges did, on the whole, respond to the initiative by sentencing more punitively overall (as evidenced by the positive and statistically significant coefficients on Post-Initiative and the fact that the marginal effect is positive and statistically significant for most values in the figure).⁸⁰ The fact that the marginal effect is downward sloping indicates that the amount of responsiveness was less for judges whose constituents supported the legalization of marijuana (the highest values of initiative support) than for those judges whose constituents did not support the measure (the lowest values of initiative support).

The second resounding conclusion from Figure 3 is the lack of a difference between male and female judges in their response for the initiative. In the table, the coefficient on the three-way multiplicative interaction term does not reach statistical significance. This suggests that the extent to which judges reflected issue-specific public opinion (here, the support for the initiative) before and after the initiative does not change based upon whether the judge is male or female. This is easily confirmed in Figure 3: the two panels of the figure look nearly identical.

79. Nelson, *supra* note 6, at 125, 128.

80. There is one slight difference from Nelson's finding as shown in Figure 3: there is evidence that, using a linear model rather than a count model and controlling for the gender of the judge, the judges with the most liberal constituencies did not change their behavior after the initiative. *Id.* at 143–44.

FIGURE 3: MARGINAL EFFECTS OF POST-INITIATIVE ON SENTENCE SEVERITY, BY JUDGE GENDER



The model estimates come from Model 1 in Table 3 and hold all other covariates at their median (for interval-level variables) or modal (for categorical variables) values. The figure demonstrates the lack of differential responsiveness to the public opinion signal by the gender of the judge.

The control variables suggest similar conclusions to those drawn earlier. Again, there is some evidence (albeit limited to felonies) that female defendants receive lighter sentences (by about one year) while nonwhite defendants appear to get harsher sentences overall and when convicted of a felony. This corresponds to a 1.8-year increase in the average incarceration time for a nonwhite defendant convicted of a felony. Likewise, there is some evidence that judges in Colorado who were appointed by Republican governors, all else equal, are less punitive in their sentences.

CONCLUSION

As we noted at the outset of this project, because female judges represent but a small fraction of U.S. trial judges, systematic judicial behavior differences between male and female judges can have

important implications for criminal defendants. As our empirical findings reveal, this difference is born out in the data for one group of defendants—namely females. In other words, our results indicate that while male and female judges sentence male defendants equally harshly, the sentences of female defendants differ based on the gender of the judge. Compared to male defendants, female defendants are sentenced significantly more leniently by female judges, while female defendants tend to receive harsher sentences than a similarly situated male defendant when a male judge is behind the bench. The intricacy of these results may explain the hodgepodge state of the prior empirical literature on the effects of judge gender on criminal sentencing: only by examining the interactive effect of judge gender and defendant gender does a difference emerge. Of course, this needs to be carefully tested for in other trial court settings before we can attest to its generalizability.

By contrast, our results show no evidence that female and male trial judges in Colorado respond differently to public opinion. While some theories expected that, because of social stereotypes that question female competence, female judges would be more likely to be cognizant of and responsive to public opinion, it may well be that the perception of vulnerability extends to all elected judges, regardless of their gender. We hope to see the effects of public opinion on female elites further tested in future empirical projects on other judicial and political settings.

More generally, we believe that the continued empirical evaluation of female and male judicial behavior is fruitful. There are many other judicial decisionmaking opportunities that may yield different outputs between male and female judges. These include, for example, opinion writing, collegial interaction with fellow judges, oral argument behavior, and trial court case management, including the potential for active encouragement of settlement and plea bargaining. Only by examining these different arenas of judicial behavior across courts and time can we truly gain an understanding of the substantive importance of diversifying the judiciary.