The Judge, the Politician, and the Press:
Newspaper Coverage and Criminal Sentencing Across Electoral Systems

Supplementary Material (Online Appendix)

Claire S.H. Lim* Cornell University
James M. Snyder, Jr.† Harvard University
David Strömberg‡ Stockholm University

Abstract

This supplementary material contains details of "The Judge, the Politician, and the Press: Newspaper Coverage and Criminal Sentencing Across Electoral Systems". It is composed of seven sections. Section A presents a model of media influence on judges. Section B presents examples of lenient criminal sentences that caused controversies and media coverage. Section C presents details of the National Judicial Reporting Program (NJRP) data. Section D lists ballot propositions used in the measurement of voters' penal preferences, *Harsh Vote Share*. Section E documents details of key control variables used in our main analysis and presents an instrumental variable regression version of our main result. Section F provides sensitivity analyses of our main results. Section G documents whether Democratic and Republican newspapers are differentially distributed across judicial selection systems or have different effects on sentencing.

*Department of Economics, Cornell University. e-mail: clairelim@cornell.edu
†Department of Government, Harvard University. e-mail: jsnyder@gov.harvard.edu
‡IIES, Stockholm University. e-mail: david.stromberg@iies.su.se

A A Model of Media Influence on Judges

In this section, we develop a simple model of media influence on courts under different judicial selection systems, starting with non-partisan elections. Theoretically, there are two main reasons why judicial selection mechanisms matter. They affect the *preference types of judges selected* ("selection effect") and the *incentives* of these judges once they are selected ("incentive effect"). The media may influence the functioning of judicial selection mechanisms through both channels. If there is more media coverage of the courts, then voters may acquire better information about judicial candidates, which makes the preference of judges selected better aligned with that of voters. Media coverage may also affect incentives by increasing the electoral penalty for judges who impose sentences that deviate from those desired by the voters.

Our model focuses on the selection effect, which substantially simplifies the analysis. The way that media coverage affects sentencing behavior through reelection incentives is similar to the effect through selection, because judges run in a series of elections, and the same variable that determines whether judges are elected also determines whether they are reelected.[1]

In our model, the media is not a strategic actor, but simply provides truthful information about the sentencing behavior of judges. Voters have two key heterogeneities: they differ in sentencing preferences (whether they prefer lenient or harsh sentences) and the intensity of their preferences (whether they are "ordinary" voters, who assign a small weight to sentencing in multi-issue elections, or "special interest" voters, who intensely care about sentencing). We assume that only a fraction of ordinary voters are informed about the sentencing behavior of the incumbent judge, while all the special interest voters are informed. Media coverage about judges increases the share of ordinary voters who are informed. Media coverage has a relatively larger influence in the non-partisan election system than in the partisan election system, because voters in the partisan system take into account information on the party affiliation of the incumbent and the challenger.[2] Media influence in the appointment system is also smaller than in the non-partisan system, because ordinary voters, influenced by media coverage, may care more intensely about other issues in gubernatorial elections.

[1]A model of the incentive effect would be one in which voters are imperfectly informed about judges' sentencing preferences, as in the model of the selection effect we lay out below, but incumbent judges potentially deviate from their own preferences to pretend (signal) that their preferences are better aligned with voters', in order to increase reelection probability.

[2]In the model, we assume that two parties have different distributions of sentencing preferences, which implies that the party affiliation of judicial candidates is informative about their future sentencing. This assumption is not absolutely necessary to derive our hypothesis that media influence is larger in the non-partisan than in the partisan system, thus we do not make this assumption in the conceptual framework in the main text. As long as voters have some reasons to vote based on party, it would explain our main empirical results. However, from the point of view of mathematical modeling, this assumption helps us to avoid making an ad hoc assumption that voters simply care about party affiliation for unclear reasons.

A.1 Preference and Timing

Sentencing preferences in the population are distributed as follows. Individual i, voter or judge, has utility

$$-v_i |h - \alpha_i|$$

from sentencing harshness h, where v_i is the intensity of his preference, and α_i is his preferred sentencing. The preferred sentencing, α, is either -1 (lenient) or 1 (harsh). A share π of the population has the harsh penal preference and the average preferred sentence in the population is $\overline{h}$. There are two parties, R and D. A share γ ($\gamma > \frac{1}{2}$) of people who belong to party R and a share $1-\gamma$ of people who belong to party D have the harsh penal preference.

In judicial elections, voters' preference also includes an idiosyncratic preference shock about other features of judges' performance, ε_i, uniformly distributed with mean zero and density φ, and a competence shock, η, that is uniformly distributed with mean β and density ψ. In sum, the utility that individual i gets from judges' behavior, in judicial elections, is as follows:

$$u(h, \varepsilon_i, \eta; v_i, \alpha_i) = v_i(-|h - \alpha_i| - \varepsilon_i - \eta).$$

For some voters ("special interest" voters), judges' behavior is a salient issue, so they care more and are better informed about it. Such voters could include, for example, people working in the state judicial or law enforcement systems such as lawyers, prosecutors, and police, as well as jurors, criminals, and victims. Special interest voters are a small share σ_s of the electorate. They have preference parameter $v_i = 1$ and are perfectly informed about the incumbent judge's sentencing harshness, h. The average sentencing preference among the special interest voters is denoted by $\overline{h}_s$. The rest of the voters are "ordinary" voters with preference parameter $v_i = v_L < 1$. They are not informed about sentencing in the absence of media coverage. The average sentencing preference among the ordinary voters is denoted by $\overline{h}_n$. Media coverage increases the share, ρ_n, of the ordinary voters that are informed about h. The share of the electorate that is informed about h is denoted by ρ $(= \sigma_s + \rho_n(1 - \sigma_s))$.

Our model has two periods. In the first period, an incumbent judge selects sentencing harshness, h. Then, the second-period judge is selected, by election or appointment. The winning judge selects sentencing in the second period. To focus entirely on the selection effect, we assume that a judge always selects his most preferred sentencing, $h = \alpha_i$. We analyze media influence through selection by focusing on sentencing in the second period.[3]

[3]This assumption allows us to abstract from judges' incentives to strategically transmit – i.e. signal – their preference through sentencing decisions in order to influence the reelection probability that would complicate the analysis without adding useful insights. With this assumption, our model can be regarded, in essence, as a one-shot model of selection where voters know the true preference of the incumbent and compare it with the expected preference of the challenger.

We analyze three judicial selection systems: non-partisan elections (with superscript NP), partisan elections (P), and gubernatorial appointments (A). In non-partisan and partisan elections, the incumbent is randomly drawn from the population in the first period. The primary difference between the two systems is that the party label of the incumbent judge is revealed to the voters only in partisan elections. The other difference is that in non-partisan elections, the incumbent runs against a challenger randomly drawn from the whole population, whereas in partisan elections the incumbent runs against a challenger randomly drawn from the other party. Under the appointment system, a governor is randomly drawn from the population in the first period. He appoints a judge of the same preference. Then, the governor runs against a challenger randomly drawn from the other party. Voters vote in the gubernatorial election, observing party labels. The elected governor may then appoint a new judge or re-appoint the incumbent.

A.2 Selection and Sentencing

A.2.1 Non-Partisan Elections

An informed voter i votes to reelect the incumbent if

$$v_i \left(\alpha_i \left(h - \overline{h}\right) - \varepsilon_i - \eta\right) > 0. \tag{1}$$

This happens with probabilities[4]

$$P_H = \frac{1}{2} + \varphi \left(h - \overline{h} - \eta\right) \text{ and } P_L = \frac{1}{2} + \varphi \left(-h + \overline{h} - \eta\right)$$

for a voter with harsh ($\alpha_i = 1$) and lenient ($\alpha_i = -1$) penal preferences, respectively.

In the non-partisan election system, uninformed voters have no information about sentencing preferences of the incumbent and the challenger. They vote for the incumbent judge with probability

$$P_u = \frac{1}{2} - \varphi\eta.$$

The incumbent is reelected if he or she receives more than half of the votes:

$$\rho \left(\sigma_{Hi} P_H + \sigma_{Li} P_L\right) + (1 - \rho) P_u \geq \frac{1}{2},$$

where σ_{Hi} and σ_{Li} are the shares of the voters who are informed and harsh, and informed and

[4] We assume that $2 + \beta + \frac{1}{2\psi} < \frac{1}{2\varphi}$. This ensures that P_L and P_H always lie between zero and one.

lenient, respectively, or, equivalently, if

$$\overline{h}_i \left(h - \overline{h}\right) \geq \eta,$$

where $\overline{h}_i$ is the average preference of the informed voters, $\overline{h}_i = \sigma_s \overline{h}_s + \rho_n \overline{h}_n$. The probability of reelection in non-partisan elections is [5]

$$P^{NP} = \frac{1}{2} + \psi\beta + \psi\overline{h}_i \left(h - \overline{h}\right).$$

A.2.2 Partisan elections

In the partisan election system, voters know the party label of the incumbent and the challenger. The expected sentencing harshness is $h_R^e = \gamma - (1 - \gamma) = 2\gamma - 1$ and $h_D^e = -h_R^e$ for Republican and Democratic judges, respectively. An uninformed voter now votes for an incumbent Republican judge with probability

$$P_i(\alpha_i) = \frac{1}{2} - \alpha_i \varphi \left(h_R^e - h_D^e\right) - \varphi\eta.$$

Republican and Democratic judges are reelected with probabilities

$$P_R^P = \frac{1}{2} + \psi\beta + \psi\overline{h}\left(h_R^e - h_D^e\right) + \rho\psi\overline{h}_i \left(h - h_R^e\right), \tag{2}$$

$$P_D^P = \frac{1}{2} + \psi\beta + \psi\overline{h}\left(h_D^e - h_R^e\right) + \rho\psi\overline{h}_i \left(h - h_D^e\right). \tag{3}$$

Note that voters now use the party labels. This accounts for the third terms in the expressions above. The final term is the effect of informed voters updating their beliefs from the expected sentencing of a judge with a given party label to the actual sentencing.

A.2.3 Appointments

In the gubernatorial election, there are two issues: criminal sentencing harshness (h) and an ideological issue which determines party labels R and D. We treat the ideological issue as exogenous. People vote for the incumbent governor if

$$v_i \alpha \left(h^e - h_2^e\right) - \varepsilon_{2i} - \eta_2 \geq 0, \tag{4}$$

where h^e and h_2^e are the expected sentencing preferences of the governor and the challenger in the gubernatorial election, respectively. Stochastic terms ε_{2i} and η_2 are idiosyncratic and

[5] We assume that $\psi(2 + \beta) < \frac{1}{2}$. This ensures that P^{NP} always lies between zero and one.

systemic preference shocks that favor party R's ideology, respectively. The idiosyncratic ideological preference shock, ε_{2i}, is uniformly distributed with mean zero and density φ_g. The systemic ideological preference shock, η_2, is uniformly distributed with mean β_g and density ψ_g. We assume that the systemic ideological preference shock in partisan elections matters more for voters' utility than the shock to judge competence in non-partisan elections. The distribution of η_2 is consequently wider, i.e., $\psi_g < \psi$. The parameter β_g captures the average ideological inclination of the population. Note that the expected share, p, of voters that are in favor of the Republican candidate is[6]

$$p = \frac{1}{2} + \varphi_g \beta_g. \tag{5}$$

In contrast to direct elections, the intensity of penal preference, v_i, now matters for the vote choice, which is evident from comparing the above inequality (4) with inequality (1). This is because the gubernatorial election bundles the judicial issue with the ideological issue, and voters trade off utility on one issue against the other. For this reason, it will be convenient to define the intensity-weighted aggregate penal preferences among the electorate and informed voters, respectively, as $\overline{h}_v = \sigma_s \overline{h}_s + v_L (1 - \sigma_s) \overline{h}_n$ and $\overline{h}_{vi} = \sigma_s \overline{h}_s + v_L \rho_n \overline{h}_n$. Because some voters care much more about sentencing, these may be very different from their unweighted counterparts, $\overline{h}$ and $\overline{h}_i$.

The probabilities of reelection for Republican and Democratic governors are, respectively,[7]

$$\begin{aligned} P_R^A &= \frac{1}{2} + \psi_g \varphi_g \beta_g + \psi_g \overline{h}_v \left(h_R^e - h_D^e\right) + \overline{h}_{vi} \psi_g \left(h - h_R^e\right), & (6) \\ P_D^A &= \frac{1}{2} - \psi_g \varphi_g \beta_g + \psi_g \overline{h}_v \left(h_D^e - h_R^e\right) + \overline{h}_{vi} \psi_g \left(h - h_D^e\right). & (7) \end{aligned}$$

The expressions have the same form as in partisan elections, equations (2) and (3). However, here the intensity-weighted aggregate preferences, $\overline{h}_v$ and $\overline{h}_{vi}$, take the place of the unweighted preferences, $\overline{h}$ and $\overline{h}_i$. Special interest voters matter more for sentencing because their votes in the gubernatorial election are more likely to be affected by sentencing.

A.2.4 Sentencing

Given the selection rules for the second period judge, it is now straightforward to compute the expected sentencing harshness for the second period.

Proposition 1. *Under the systems of non-partisan elections (NP), partisan elections (P),*

[6] We assume that $\varphi_g |\beta_g| < \frac{1}{2}$. This ensures that p always lies between zero and one.

[7] We assume that $\psi_g (3 + \varphi_g |\beta_g|) < \frac{1}{2}$. This ensures that P^A always lies between zero and one.

and appointment (A), the average second period sentencing harshness is

$$\begin{array}{lrcrcrcr} & & & & & \textit{party label} & & \textit{information} \\ \textit{Non-partisan elections:} & \overline{h}_2^{NP} = & & \overline{h} & & & + & 4\psi\pi\left(1-\pi\right)\overline{h}_i, \\ \textit{Partisan elections:} & \overline{h}_2^{P} = & & 2\beta\psi\overline{h} & + & 4\psi\left(2\gamma-1\right)^2\overline{h} & + & 4\psi\gamma\left(1-\gamma\right)\overline{h}_i, \\ \textit{Appointed judges:} & \overline{h}_2^{A} = & & \psi_g\overline{h} & + & 4\psi_g\left(2\gamma-1\right)^2\overline{h}_v & + & 4\psi_g\gamma\left(1-\gamma\right)\overline{h}_{vi}. \end{array}$$

Proof. This follows since in the non-partisan, partisan, and appointment systems, respectively:

$$\begin{aligned} \overline{s}_2^{NP} &= \sigma\left(P_1^{NP}+\left(1-P_1^{NP}\right)\overline{s}\right)+\left(1-\sigma\right)\left(-P_{-1}^{NP}+\left(1-P_{-1}^{NP}\right)\overline{s}\right), \\ \overline{s}_2^{P} &= p\gamma\left(P_{R1}^{P}+\left(1-P_{R1}^{P}\right)s_D^e\right)+p\left(1-\gamma\right)\left(-P_{R-1}^{P}+\left(1-P_{R-1}^{P}\right)s_D^e\right) \\ &\quad+\left(1-p\right)\left(1-\gamma\right)\left(P_{D1}^{P}+\left(1-P_{D1}^{P}\right)s_R^e\right)+\left(1-p\right)\gamma\left(-P_{D-1}^{P}+\left(1-P_{D-1}^{P}\right)s_R^e\right), \\ \overline{s}_2^{A} &= p\gamma\left(P_{R1}^{A}+\left(1-P_{R1}^{A}\right)s_D^e\right)+p\left(1-\gamma\right)\left(-P_{R-1}^{A}+\left(1-P_{R-1}^{A}\right)s_D^e\right) \\ &\quad+\left(1-p\right)\left(1-\gamma\right)\left(P_{D1}^{A}+\left(1-P_{D1}^{A}\right)s_R^e\right)+\left(1-p\right)\gamma\left(-P_{D-1}^{A}+\left(1-P_{D-1}^{A}\right)s_R^e\right). \end{aligned}$$

Plug in the definitions of the reelection probabilities, expected harshness of Republicans and Democrats, equation (5), and $p\gamma+(1-p)(1-\gamma)=\sigma$. □

The above expressions for sentencing harshness consist of three terms. The first term captures the random draw of the first period judge from the population with average preference $\overline{h}$ and the portion of selection effects unrelated to information. In the partisan election system, it includes the partisan ideological effect β. In the appointment system, it includes the effect of voting for governor on ideological grounds.

The second term arises because voters use party labels to infer the sentencing preferences of judges. Since everyone observes party labels, this term makes sentencing more aligned with the average preference of the electorate, $\overline{h}$, for partisan elected judges, and with the intensity-weighted average preference, $\overline{h}_v$, for appointed judges. For example, if $\overline{h}>0$, then the more informative party labels are, the harsher sentencing by partisan elected judges is. The term $(2\gamma-1)^2$ measures the information gain from learning the party label of the judge. When γ equals one half, the party labels are not informative and this term disappears. The larger γ is, the more informative party labels are and the more aligned sentencing is with voter preferences.

Our main interest is in the last term, which captures the effect of information from the media on sentencing. This term makes sentencing more aligned with the average preference of the informed voters, $\overline{h}_i$, for partisan elected judges, and with the intensity-weighted average preference of informed voters, $\overline{h}_{vi}$, for appointed judges. This effect is adjusted by a term of the form $x(1-x)$ that measures information gain from learning the judge's type perfectly.

Here, x is voters' prior that the judge is of the harsh preference. This is π if the judge is drawn from the population, γ if voters know that the judge is Republican, and $1-\gamma$ if voters know that the judge is Democrat. This gain is largest when the voters' prior is most uninformative, $x=\frac{1}{2}$, and falls monotonically as x rises towards 1 or falls towards 0. This information gain is smaller when voters already know the party label of the judge: $(1-\gamma)\,\gamma<(1-\pi)\,\pi$.[8] As party labels become more informative, their effect becomes stronger, and the effect of media becomes weaker for partisan elected and appointed judges.

Media coverage increases the share, ρ_n, of ordinary voters who are informed about the incumbent judge's sentencing. For simplicity, suppose that this relationship is linear, so that $\rho_n=\omega n$, where n is the number of newspaper articles covering the judge, and ω is a positive parameter. To see the effect of information from media on sentencing, differentiate the expressions in the proposition to get

$$\begin{aligned} \frac{d\overline{h}_2^{NP}}{dn} &= 4\omega\psi\,(1-\pi)\,\pi\overline{h}_n, \\ \frac{d\overline{h}_2^{P}}{dn} &= 4\omega\psi\,(1-\gamma)\,\gamma\overline{h}_n, \\ \frac{d\overline{h}_2^{A}}{dn} &= 4\omega\psi_g\,(1-\gamma)\,\gamma v_L\overline{h}_n. \end{aligned} \tag{8}$$

Media coverage makes sentencing more aligned with the preference of ordinary voters. We can now describe the effectiveness of media coverage by selection system.

Corollary 2. *The responsiveness of sentencing harshness to media coverage of sentencing is greatest for non-partisan elected judges, followed by partisan elected judges. The sentencing of appointed judges is least sensitive to media coverage.*

$$\frac{d\overline{h}_2^{NP}}{dn}>\frac{d\overline{h}_2^{P}}{dn}>\frac{d\overline{h}_2^{A}}{dn}. \tag{9}$$

[8] Because

$$\sigma=p\gamma+(1-p)\,(1-\gamma)\,,$$

we can write σ as

$$\sigma-\frac{1}{2}=x\,(2p-1)\,,$$

where x is the distance of γ to $\frac{1}{2}$, $x=\gamma-\frac{1}{2}$. Because $|2p-1|<1$, σ is closer to $\frac{1}{2}$ than γ. It follows that $(1-\sigma)\,\sigma>(1-\gamma)\,\gamma$ since the function $(1-x)\,x$ is monotonically decreasing with the distance to $\frac{1}{2}$.

Proof. By differentiation of the expressions in Proposition 1, we obtain

$$\begin{aligned}
\frac{d\bar{s}_2^{NP}}{d\rho} &= 4\psi\left(1-\sigma\right)\sigma\bar{s}_n, \\
\frac{d\bar{s}_2^{P}}{d\rho} &= 4\psi\left(1-\gamma\right)\gamma\bar{s}_n, \\
\frac{d\bar{s}_2^{A}}{d\rho} &= 4\psi_g\left(1-\gamma\right)\gamma v_L\bar{s}_n.
\end{aligned}$$

To prove the first inequality, we need to show that $(1-\sigma)\sigma > (1-\gamma)\gamma$. Because

$$\sigma = p\gamma + (1-p)(1-\gamma),$$

we can write σ as

$$\sigma - \frac{1}{2} = x(2p-1),$$

where x is the distance of γ to $\frac{1}{2}$, $x = \gamma - \frac{1}{2}$. Because $|2p-1| < 1$, σ is closer to $\frac{1}{2}$ than γ. It follows that $(1-\sigma)\sigma > (1-\gamma)\gamma$ since the function $(1-x)x$ is monotonically decreasing with the distance to $\frac{1}{2}$. The second inequality follows since $\psi_g < \psi$ and $v_L < 1$. $\square$

Media coverage matters less for sentencing by partisan elected than by non-partisan elected judges because voters in partisan elections have an additional factor, i.e., candidates' party affiliation, that affects their voting. Media coverage matters even less for sentencing by appointed judges. This is because the intensity of preference matters in multi-issue elections, and the media informs ordinary voters, who have relatively less intense preferences on sentencing. Formally, the second implication follows from the assumptions $\psi_g < \psi$ and $v_L < 1$. Thus, if these assumptions are violated, for example, if factors unrelated to sentencing also matter a great deal in partisan elections, then the second implication may not hold. This is a more general case, discussed in the conceptual framework in the main text.

Our model also has an implication for the interaction between media influence and severity of crimes. Cases involving severe, violent crimes are more likely to be covered by the media than others. We would expect that the change in the share of informed ordinary voters, ρ_n, by newspaper coverage of courts would be largest for severe, violent crimes. Similarly, if media are more likely to report cases with black defendants, then sentencing decisions on such cases would more closely follow the preference of ordinary voters.

The media plays two roles: providing information about the incumbent judge and increasing the importance of ordinary voters relative to special interest voters. More media coverage makes it less likely that judges with preferences different from those of the median ordinary voter get reelected. If the ordinary and special interest median voters both prefer harsh or lenient judges, the two groups of voters reinforce each other's influence. If the median voters have different preferences, then more media coverage makes sentencing more aligned with the preference of ordinary voters at the expense of special interest voters.[9]

From a welfare perspective, it is not clear whether this is good or bad. The welfare maximizing sentencing is maximum harshness if the preference-weighted average sentencing preference is positive ($\overline{h}_v > 0$), and minimum harshness if this is negative. This follows since utility is linear in sentencing. In the case that, for example, $\overline{h}_v < 0$ and $\overline{h}_n > 0$, more media coverage lowers welfare. The risk that media coverage reduces welfare is particularly strong in single-issue elections where media may enforce a "tyranny of the majority" of ordinary voters against the interest of a minority with much stronger preferences.

B Examples of Lenient Sentencing Causing Controversies

In this section, we provide examples of lenient sentences causing controversies and media coverage, some of which also caused judges' difficulties in reelection.

- Judge G. Todd Baugh in Montana (nonpartisan election):
  In 2013, Judge Baugh sentenced a man convicted of raping a 14-year-old girl to only 30 days in jail. The girl subsequently committed suicide. An uproar ensued, and Judge Baugh said that he had made a mistake in his sentencing. He also announced that he was retiring at the end of his term (he had been unopposed in his previous re-election attempts). The Montana Supreme Court overturned the sentence in 2014, and also censured Judge Baugh, suspending him without pay for his last month in office (December 2014). For an example of coverage, see http://www.cbsnews.com/news/judge-who-sentenced-rapist-to-30-days-faces-suspension/.

- Judges David S. Bruce and Rodney C. Warren in Maryland (nonpartisan election):
  In 2004, the two incumbent judges were defeated in reelection by challengers on the ground that their sentencing was too lenient. Challenger Paul G. Goetzke campaigned in the election that voters wanted tougher sentences and "criminals rehabilitated in jail, not in their neighborhoods." For an example of coverage, see http://www.washingtonpost.com /wp-dyn/articles/A20047-2004Nov2.html.

[9]Information from the media does not have the strongest effect in areas where everyone supports harsh (or lenient preferences) in non-partisan elections. Suppose that the non-salient voters are a negligible part of the electorate. Then $\overline{s}_n \approx \overline{s} = 2\sigma - 1$. The derivative of $\overline{s}_2^{NP}$ with respect to ρ then has its largest value at $\sigma = \frac{1}{2} + \frac{1}{6}\sqrt{3}$, which is approximately .8.

- Judge Joyce Karlin in California (nonpartisan election):
  In 1991, shortly after being appointed as judge, Judge Karlin gave a light sentence in a racially charged case in Los Angeles. A Korean-American grocer shot and killed a 15 year-old black teenager. The grocer was convicted of voluntary manslaughter, and the judge sentenced her to 5 years of probation combined with 400 days community service and $500 fine (the maximum possible penalty was 16 years in prison). The African American community in L.A. was outraged (there were even protests in front of the judge's home). In the next election, in 1992, four candidates challenged Judge Karlin in the primary. The trial and the sentence were heavily covered in the press, and the election also received more coverage than usual. The *Los Angeles Times* endorsed one of Judge Karlin's opponents. Judge Karlin was the only L.A. Superior court judge to be challenged that year. She won, but only received 51% of the vote. For an example of coverage, see http://www.nytimes.com/1992/06/01/us/los-angeles-votes-on-lenient-verdict.html.

- Judge Paula M. Martin in Kansas (appointment and retention election):
  In 2004, Judge Martin sentenced two 18-year-old men convicted of raping an intoxicated 13-year-old girl in June 2003 to only 60 days in jail, probation and community service. Voters, outraged over the leniency of her sentencing, organized a political action committee to unseat her in the 2004 retention election. In the 2004 election, she received 63% yes-votes, which is lower than her average share of yes-votes, 78%, in three other elections. For an example of newspaper coverage, see http://www2.ljworld.com/news/courts/judge_paula_martin/.

- Judge Edward Cashman in Vermont (appointment and re-appointment):
  In 2006, Judge Cashman sentenced a man to only 60 days in jail for raping a girl over a four-year period, beginning when she was six years old. This caused public outrage and demands for his resignation. He retired at age 67 in 2007 after completing his term. For details, see http://www.foxnews.com/story/0,2933,181498,00.html#ixzz2ZEsN5iH4.

C Details of the NJRP Data

In this section, we document details of the NJRP Data. In Tables A.1-A.2, we list the states and counties included in our sentencing data. In Table A.3, we show the number of judicial districts and sentences in the NJRP data by selection system and state. In Figure A.1, we show the number of NJRP sentences by judicial district. In Table A.4, we document the number of sentences and the mean sentence length by crime category.

[Tables A.1-A.3 here.]

[Figure A.1 here.]

[Table A.4 here.]

D Ballot Propositions Used for Measurement of Penal Preferences

In this section, we document the ballot propositions used to construct the measure of voters' penal preferences, *Harsh Vote Share*, in Tables A.5-A.6.

[Tables A.5-A.6 here.]

E Details of Key Variables and IV Regression

In this section, we document details of key variables used in our main analysis. In addition, we present instrumental variable regression of *Harshness* on the amount of media coverage, penal preferences, and selection systems, comparable to Table 8 in the main text.

First, in Table A.7, we document summary statistics of key variables in our analysis – the two media variables, and the demographic and political characteristics of judicial districts – by judicial selection system. The unit of observation is judicial district by year. The set of district-by-year observations included in the table is identical to that in Table 8. We conducted t-tests for each variable for the difference across selection systems. *, **, and *** on the mean of each variable denotes the statistical significance of the difference in the variable between the given system and the other systems. We find no statistically significant difference in the *Log Number of Articles* across systems. The mean level of *Congruence* is slightly lower in the appointment system than in the other systems, but the difference is small (.50 compared to .52 and .53) and statistically significant only at the 10% level. There are several demographic variables that have statistically significant differences across systems, but most of those differences are negligible.

Next, in Table A.8, we document how our key control variables are related to *Harsh Vote Share (HVS)*, *Congruence*, and *Harshness*. For all samples and for each judicial selection system, we present three specifications. In the first specification, we regress *HVS* on five key characteristics of judicial districts: *Democratic Vote Share (DVS)*, log population, log area size, log per capita income, and log number of crimes. The overall pattern of the relationship is similar across systems. *DVS* is strongly correlated with *HVS* in all three systems. In the second specification, we regress *Congruence* on *HVS* and the five key variables mentioned above. In the third specification, we regress *Harshness* on *Congruence*, *HVS*, and the five key control variables above. This regression is identical to Columns (2), (4), and (6) in Panel B of Table 8 except that we include only a small set of control variables to highlight relationships between key characteristics of judicial districts and *Harshness*. In the second and the third regressions, some variables such as log population and log area size are strongly correlated with *Congruence* and *Harshness*, which explains the change in the coefficient estimate of *Congruence* between Columns (4) and (5) in Panel A of Table 8. However, the smallness of the differences in the mean and variance of *Congruence* and demographic variables across systems in Table A.7 mitigates the concern that the variation in the effect of *Congruence* across systems in Table 8 captures variation in the distribution of unobserved heterogeneity across systems. Also note that the coefficient estimates of *Congruence* are close to that of the main result in Table 8, despite the fact that the set of control variables is much smaller. This robustness provides additional evidence of the credibility of our causal inference based on *Congruence*.

Finally, in Table A.9, we present instrumental variable regressions of *Harshness* on the amount of coverage, penal preferences, and selection systems, comparable to Table 8 where we use *Congruence* as a proxy rather than an instrumental variable. The main pattern of the result is similar to that in Table 8. The influence of media coverage is quantitatively and statistically significant only for non-partisan elected judges. Using the estimate of 0.062 in Columns (2) of Panel B, a one standard deviation (0.71) increase in the *Log Number of Articles* is estimated to increase the sentence length by about 8.6 months (5.2 percent) in the nonpartisan election system.[10]

[10] This calculation is similar to that in Footnote 43 in the main text: 194*0.062*0.71= 8.6 months.

F Sensitivity Analysis

In this section, we document robustness checks (sensitivity analyses) of the main results. In Table A.10, we document the robustness of the results presented in Columns (2) and (5) in Panel A of Table 8. In Table A.11, we reproduce in Column (1) the result in Panel B of Table 8, using interactions between *Congruence* and judicial selection systems rather than sample splits. *Congruence* and all the control variables are interacted with judicial selection systems. In Columns (2)-(6), we document the robustness of the result in Column (1). In both Tables A.10 and A.11, we present the following set of sensitivity analyses:

(a) OH: Ohio has a unique system with partisan primaries and nonpartisan general elections. In the baseline specification, we coded Ohio as a nonpartisan system. Columns labeled "OH" show the result from coding Ohio as a partisan system.

(b) MD: In Maryland, judges are initially selected by gubernatorial appointment. Then, they must run in the next major election for a 15-year term cross-filed in the Democratic and Republican primaries without party labels. If there are different winners in each primary, they will face off in the general election. In the baseline specification, we coded Maryland as a nonpartisan system. Columns labeled "MD" show results from coding Maryland as a appointment system.

(c) OHMD: In columns labeled "OHMD", we use alternative coding for both Ohio and Maryland. That is, Ohio is coded as a partisan system, and Maryland is coded as a appointment system.

(d) post90: In these columns, we drop sentencing data before 1991. We conduct this sensitivity analysis because we interpolated *Congruence* for the years 1983-1990.

(e) Texas: In these columns, we drop Texas, where the jury can decide the final sentence upon defendants' request.

[Tables A.10-A.11 here.]

In Table A.11, the coefficient of *Congruence* shows the estimated effect in the non-partisan system. The estimated effects of *Congruence* in the other selection systems are found by adding the main effect to the relevant interaction term. For example, the estimated effect in the partisan system can be obtained by adding the estimated coefficient of the relevant interaction (-.096) to the baseline estimate (0.098).

The rows labeled 'Congruence Partisan Elected' and 'Congruence Appointed' show the p-values from F-tests of the hypotheses that the effect of *Congruence* is zero in the partisan elected and appointed sub-samples, respectively. Likewise, 'Harsh Non-Partisan Elected' and 'Harsh Appointed' show the p-values from F-tests of the hypotheses that the effect of *Harsh Vote Share* is zero in the non-partisan elected and appointed sub-samples, respectively.

The estimated effects of *Congruence* on *Harshness* are not statistically significantly different from zero for appointed or partisan elected judges.

In Tables A.12-A.13, we present the following set of sensitivity analyses:

(a) We exclude all judicial districts where the share urban is 100%. This is around 7% of the observations in our sample used in Table 8.

(b) We control for the average population (level and log) of the headquarter (HQ) markets of a district's newspapers (weighted by circulation in the district).[11]

(c) We control for the average (level and log) circulation of a district's newspapers (weighted by circulation in the district).

(d) We control for demographics (population, race, education, etc.) of the metropolitan statistical area (MSA) a district is in, rather than demographics of the district itself. We aggregate the demographic controls from the county to the MSA level. We use the same set of controls as in the judicial district demographics (log population, log area size, log per capita income, log employment, the share of people in the district who are religious adherents, female, younger than 20, older than 65, black, white, Hispanic, urban, the share with high school education, the share with more than high school education, turnout in the most recent presidential election, and total newspaper penetration).

In each of Tables A.12 and A.13, we use the following sets of control variables:

- Controls=1: baseline set of controls
- Controls=2: baseline plus the four variables from points (b) and (c)
- Controls=3: baseline plus MSA demographics
- Controls=4: baseline plus the four variables from (b) and (c) and MSA demographics.

[11]We do not have the HQ county in our database, but used the county where the newspaper has the highest sales. In most cases, this is the HQ county.

Table A.12 shows that the coefficients of *Log Number of Articles* and *Congruence* are not affected much by any of these variations in control variables, although they marginally lose significance in the specifications with MSA demographics in the full sample (Columns (3) and (4)). Table A.13 shows the sensitivity analyses of the results in Table 8, using interactions between *Congruence* and judicial selection systems rather than sample splits. All controls are also interacted with the selection systems. The estimated effect of *Congruence* for the non-partisan elected is robust. The coefficient estimates of interacted *Congruence* are not affected much, either, although one coefficient loses significance ($Congruence * Appointed$) in Column (4).

[Tables A.12-A.13 here.]

G Partisan Newspapers

In this section, we investigate whether Democratic and Republican newspapers are differentially distributed across selection systems or have different effects on sentencing. To this end, we created an index of newspaper ideology using their endorsements of presidential candidates over four elections from 1992 to 2004. We have at least one endorsement for 631 newspapers, representing 65% of the newspapers and 73% of the sales of the newspapers used to compute our *Congruence* measure. We define a newspaper as non-endorsing if it did not make any endorsements, as Democratic if more than half of its two-party endorsements went to a Democratic candidate, and as Republican if more than half of its two-party endorsements went to a Republican candidate. In our sample, 55% of the endorsing papers are Republican by this definition.

A possible concern in the main results which we address here is that media effects differ across judicial selection systems because newspaper ideologies do. To investigate this, we regress the total number of endorsements for presidential candidates for the period from 1986 to 2002 and the share of two-party endorsements that went to Democrats candidates, on dummy variables for the judicial selection system in the judicial district where the newspaper has its highest sales (which is typically where the newspaper headquarters are located). The results are presented in Table A.14. As the dependent variable, Columns (1) and (2) use the total number of endorsements over the four elections, and Columns (3) and (4) use the share of endorsements for Democrats. Columns (2) and (4) control for judicial district demographic characteristics (log population, log area size, log per capita income, log employment, the share of people in the district who are religious adherents, female, younger than 20, older than 65, black, white, Hispanic, urban, the share with high school education, the share with more than high school education, and turnout in the most recent presidential election). We find no evidence that newspaper endorsements correlate with the judicial selection system.

[Table A.14 here.]

We next analyze whether Republican and Democratic papers have different effects on sentencing. They could have differential effects, for example, because Republican newspapers advocate longer sentences and it affects penal preferences, or because Republican newspapers cover the courts more than Democratic newspapers do. The latter turns out not to be true in our sample. Newspapers that endorse any two-party candidate carry twice as many articles about judges (20 compared to 10). However, Republican and Democratic papers cover judges to an equal extent. It seems that newspapers that engage in one type of political coverage (presidential endorsements) are also more likely to engage in others (covering judges).

[Table A.15 here.]

Since *Congruence* is constructed as a weighted sum of reader shares from different newspapers, we can split it into three parts, corresponding to the contributions to this sum by Democratic, Republican, and non-endorsing newspapers. The first column of Table A.15 shows a regression of number of newspaper articles per judge per year on *Congruence* (with the baseline definition), that of Democratic newspapers, and of Republican newspapers. The coefficient on endorsing papers is twice as large as that on non-endorsing papers.[12] The next column investigates whether Democratic and Republican *Congruence* affects penal preferences differently, using the specifications analogous to Table 4 in the main text. There is a split in the expected direction (Democratic *Congruence* and Republican *Congruence* are associated with lenient and harsh preferences, respectively), but these differences are small and statistically insignificant. The final three columns investigate sentencing, using the specification of Columns (4), (5), and (6) of Panel A in Table 8, respectively. Again, it seems that *Congruence* by endorsing papers is what matters, but the party being endorsed does not matter. This evidence is consistent with the *amount* of newspaper coverage being important for the accountability of judges.

[12] The coefficient is 31.277 (=15.089+16.188) for Democratic papers, 31.069 (=15.089+15.980) for Republican papers, and 15.089 for non-endorsing papers.

Table A.1
States and Counties Included in the NJRP Data

State	Counties
AL	Baldwin, Bullock, Clay, Dale, Fayette, Greene, Hale, Jefferson, Lauderdale, Macon, Madison, Mobile, Morgan, Pike, Tallapoosa
AR	Benton, Boone, Carroll, Dallas, Hot Spring, Mississippi, Newton, Phillips, Pulaski, Sharp, Union, Washington, White, Woodruff
AZ	Apache, Coconino, Gila, Maricopa, Pima, Yavapai, Yuma
CA	Alameda, Contra Costa, Fresno, Kern, Kings, Los Angeles, Modoc, Monterey, Orange, Sacramento, San Bernardino, San Diego, San Francisco, San Joaquin, San Luis Obispo, Santa Clara, Shasta, Sonoma, Tulare, Ventura
CO	Adams, Arapahoe, Boulder, Cheyenne, Denver, El Paso, Gunnison, Huerfano, Jackson, Jefferson, Kit Carson, La Plata, Otero, Weld
DC	District of Columbia
DE	New Castle, Sussex
FL	Alachua, Brevard, Broward, Charlotte, Citrus, Collier, Duval, Escambia, Franklin, Gilchrist, Gulf, Hillsborough, Indian River, Jackson, Jefferson, Lake, Lee, Leon, Madison, Manatee, Marion, Miami-Dade, Monroe, Okaloosa, Orange, Palm Beach, Pasco, Pinellas, Polk, Sarasota, St. Johns, St. Lucie, Volusia,
GA	Baldwin, Banks, Bryan, Bulloch, Calhoun, Chatham, Chattooga, Clarke, Clayton, Clinch, Cobb, Columbia, Coweta, De Kalb, Dooly, Dougherty, Evans, Fulton, Grady, Gwinnett, Hall, Liberty, Lowndes, McDuffie, Muscogee, Polk, Richmond, Rockdale, Walker
HI	Honolulu, Maui
IA	Benton, Butler, Des Moines, Johnson, Lucas, Mitchell, Polk, Scott, Woodbury, Wright
ID	Bannock, Bonner, Jefferson
IL	Champaign, Christian, Cook, Du Page, Ford, Hancock, Hardin, Jasper, Kane, Lake, Macon, McHenry, Rock Island, Sangamon, St. Clair, Whiteside, Williamson, Winnebago
IN	Adams, Allen, Decatur, Delaware, Franklin, Johnson, Lake, Marion, St Joseph, Vanderburgh, Wabash, Warrick
KS	Butler, Clay, Cowley, Douglas, Ford, Gray, Jefferson, Johnson, Kearny, Kiowa, Marion, Marshall, Montgomery, Norton, Pratt, Rawlins, Sedgwick, Shawnee, Sumner, Wilson, Wyandotte
KY	Adair, Barren, Bath, Boone, Bracken, Carter, Daviess, Estill, Fayette, Hancock, Hardin, Jefferson, Kenton, Lawrence, Logan, Marshall, Mason, Nicholas, Ohio, Owen, Owsley, Pendleton, Powell, Pulaski, Warren
LA	Ascension, Beauregard, Bossier, Caddo, Calcasieu, Concordia, East Baton Rouge, Iberia, Jefferson, Lafayette, Lafourche, Lincoln, Orleans, Ouachita, Rapides, Red River, Sabine, St. Mary, Tangipahoa, Terrebonne, Winn
MD	Anne Arundel, Baltimore, Baltimore City, Carroll, Charles, Dorchester, Frederick, Howard, Montgomery, Prince Georges, Washington
ME	Cumberland, Kennebec, Oxford, Somerset
MI	Antrim, Berrien, Calhoun, Chippewa, Clare, Eaton, Genesee, Grand Traverse, Isabella, Jackson, Kent, Lenawee, Livingston, Macomb, Marquette, Missaukee, Monroe, Montcalm, Oakland, Otsego, Ottawa, Saginaw, Sanilac, St. Clair, Washtenaw, Wayne
MN	Aitkin, Anoka, Blue Earth, Chippewa, Chisago, Dakota, Dodge, Douglas, Freeborn, Goodhue, Hennepin, Hubbard, Koochiching, Martin, McLeod, Morrison, Olmsted, Ramsey, Sherburne, St Louis, Stearns, Washington, Wright

Table A.2
States and Counties Included in the NJRP Data (cont'd)

State	Counties
MO	Andrew, Boone, Caldwell, Clay, Cole, Dunklin, Franklin, Howell, Jackson, Jasper, Jefferson, Johnson, Knox, Livingston, Madison, Oregon, Platte, Ray, Saline, Scotland, Scott, St. Charles, St. Louis, St. Louis City, Stone, Wright
MS	Benton, Copiah, Forrest, Hancock, Harrison, Holmes, Jackson, Lauderdale, Lowndes, Panola, Rankin, Sunflower, Walthall, Webster
MT	Carbon, Roosevelt, Yellowstone
NC	Alamance, Anson, Bladen, Buncombe, Caldwell, Catawba, Cherokee, Cleveland, Columbus, Cumberland, Davidson, Duplin, Durham, Forsyth, Franklin, Gaston, Guilford, Henderson, Iredell, Jackson, Mecklenburg, New Hanover, Orange, Pasquotank, Richmond, Rowan, Surry, Wake, Wilkes, Yadkin
NE	Cass, Dawson, Douglas, Lancaster
NH	Belknap, Carroll, Strafford
NJ	Bergen, Burlington, Camden, Cape May, Cumberland, Essex, Gloucester, Hudson, Hunterdon, Mercer, Middlesex, Monmouth, Ocean, Passaic, Salem, Somerset, Sussex, Union, Warren,
NM	Bernalillo, Cibola, Dona Ana, Lea, Los Alamos, Otero, Rio Arriba, San Juan, Santa Fe
NY	Albany, Allegany, Bronx, Broome, Chautauqua, Clinton, Cortland, Erie, Essex, Kings, Livingston, Madison, Monroe, Montgomery, Nassau, New York, Niagara, Oneida, Onondaga, Ontario, Orange, Oswego, Queens, Rensselaer, Richmond, Rockland, Schenectady, Steuben, Suffolk, Ulster, Wayne, Westchester, Wyoming
OH	Cuyahoga, Franklin, Hardin, Highland, Jefferson, Knox, Lake, Licking, Lucas, Mahoning, Montgomery, Putnam, Richland, Sandusky, Summit, Wood, Wyandot
OK	Atoka, Beckham, Canadian, Comanche, Garfield, Haskell, Jackson, Kay, Lincoln, Mayes, Mcclain, Oklahoma, Pittsburg, Rogers, Tulsa, Washington
OR	Benton, Clackamas, Columbia, Deschutes, Douglas, Lane, Linn, Malheur, Marion, Multnomah, Polk, Umatilla, Wallowa, Washington
PA	Allegheny, Armstrong, Beaver, Berks, Bucks, Butler, Carbon, Centre, Chester, Cumberland, Dauphin, Delaware, Erie, Fayette, Huntingdon, Lackawanna, Lancaster, Lawrence, Lebanon, Lehigh, Luzerne, Lycoming, Mercer, Monroe, Montgomery, Northampton, Northumberland, Philadelphia, Somerset, Sullivan, Venango, Wayne, Westmoreland, Wyoming, York
RI	Kent, Newport, Providence, Washington
SC	Anderson, Beaufort, Berkeley, Charleston, Colleton, Dillon, Edgefield, Florence, Greenville, Lancaster, Lexington, Marlboro, Richland, Saluda, Spartanburg, Williamsburg, York
SD	Aurora, Beadle, Pennington
TN	Blount, Davidson, Giles, Hamblen, Hamilton, Humphreys, Knox, Montgomery, Robertson, Rutherford, Shelby, Sullivan, Trousdale, Wayne
VA	Alexandria City, Alleghany, Appomattox, Arlington, Botetourt, Chesapeake City, Chesterfield, Danville City, Essex, Fairfax, Fairfax City, Fauquier, Hampton City, Henrico, Henry, Highland, King and Queen, Lancaster, Loudoun, Louisa, Middlesex, Nelson, New Kent, Newport News City, Norfolk City, Prince William, Radford, Richmond, Richmond City, Roanoke City, Rockingham, Smyth, Spotsylvania, Stafford, Virginia Beach City, Washington, Winchester City
VT	Caledonia, Chittenden, Essex
WA	Benton, Island, King, Kitsap, Lewis, Okanogan, Pierce, Skagit, Skamania, Snohomish, Thurston, Whatcom
WI	Brown, Crawford, Dane, Jackson, Jefferson, Langlade, Lincoln, Marathon, Milwaukee, Pierce, Rock, Wood
WV	Fayette, Mason, McDowell, Mineral, Monroe, Putnam, Taylor, Webster
WY	Sweetwater

Table A.3

Number of Judicial Districts and Sentences in the NJRP Data by Selection System and State

Partisan Election			Nonpartisan Election			Appointment		
State	Number of Districts	Number of Sentences	State	Number of Districts	Number of Sentences	State	Number of Districts	Number of Sentences
AL	13	631	AR	6	949	AZ	2	5,805
AR	9	699	AZ	5	150	CO	13	2,626
IL	14	16,772	CA	20	46,473	DC	1	1,532
IN	11	1,537	FL	20	23,964	DE	1	611
KS	9	1,315	GA	23	3,187	HI	2	1,199
LA	20	6,331	ID	3	53	IA	6	454
MO	18	325	IN	1	69	KS	9	351
MS	7	209	KY	20	1,859	ME	1	118
NC	30	1,895	MD	7	3,104	MO	5	2,685
NM	7	131	MI	26	10,238	NE	4	384
NY	12	30,972	MN	10	2,031	NH	1	137
PA	34	13,717	MS	12	580	NJ	15	6,999
TN	12	3,682	MT	1	35	RI	1	766
TX	37	15,589	NC	20	2,523	SC	12	1,612
WV	6	17	NV	2	790	UT	6	1,292
			OH	17	1,804	VA	25	2,586
			OK	15	1,336	VT	1	71
			OR	14	4,389	WY	1	6
			SD	3	150			
			WA	12	4,021			
			WI	12	1,709			

Table A.4
NJRP Offense Categories and Sentence Lengths

Category	Number of Sentences	Mean Sentence (months)
Homicide	26,759	486
Sexual Assault	59,578	122
Robbery	92,175	102
Aggravated Assault	171,280	41
Other Violent	31,257	26
Burglary	178,280	38
Larceny	225,241	17
Fraud	136,156	14
Drug Possession	266,458	13
Drug Trafficking	398,866	24
Weapon Offenses	87,183	25
Other Offenses	243,277	17
Total	1,916,510	36

Table A.5

Ballot Propositions Used to Measure Penal Preferences

State	Year	Prop No.	Percent Yes	Description
AL	1996	Amendment 3	70	To Remove the Prohibition on Guilty Pleas within 15 Days of Arrest in Non-Capital Felony Cases
AZ	1998	Proposition 301	48	Probation Eligibility For Drug Possession Or Use
AZ	2002	Proposition 103	80	Bailable Offenses; Prohibitions
AZ	2002	Proposition 302	69	Probation For Drug Crimes
AZ	2006	Proposition 100	77	Bailable Offenses
AZ	2006	Proposition 301	58	Probation for Methamphetamine Offenses
CA	2000	Proposition 18	72	Murder; Special Circumstances; Leg Initiative Amendment
CA	2000	Proposition 21	62	Juvenile Crime
CO	1992	Referendum A	80	Rights of Crime Victims
CO	1994	Referendum C	77	Post-Conviction Bail
FL	1998	Amendment 2	72	Preservation of Death Penalty; US Supreme Court Interpretation of Cruel And Unusual Punishment
HI	2002	Question 3	57	Initiation of Felony Prosecutions By Written Information
HI	2004	Amendment 1	65	Sexual Assault Crimes
HI	2004	Amendment 2	71	Public Access To Registration Information of Sex Offenders
HI	2004	Amendment 3	53	Rights of Alleged Crime Victims
HI	2004	Amendment 4	56	Initiation of Criminal Charges
HI	2006	Amendment 4	69	Sexual Assault Crimes Against Minors
IA	1998	Amendment 2	63	To Eliminate Limitation of Fines For Offenses That May Be Summarily Tried Without Indictment
ID	1994	H.J.R 16	80	To Provide for Rights of Crime Victims
IN	1996	Public Question 1	89	Victims' Rights
IN	2000	Public Question 1	65	Criminal Appeals Process
LA	1998	Amendment 4	69	To Provide for Rights of the Victim of a Crime
LA	1998	Amendment 6	68	To Make It Easier For Judges To Deny Bail
LA	1998	Amendment 14	62	To Require a Unanimous Verdict in Criminal Trials That Use Six-Member Jury
LA	1999	Amendment 1	59	To Provide That Governor May Not Commute Sentences or Pardon Persons Convicted Without A Favorable Recommendation By Board Of Pardons
LA	1999	Amendment 8	53	To Limit Automatic Pardon Provision To Persons Convicted of a Non-Violent Crime
MI	1994	Proposition B	74	A Proposal to Limit Criminal Appeals
MS	1998	Amendment 2		Victims' Rights
MT	1998	C-33	71	Criminal Laws Must Be Based on Principles Of Public Safety and Restitution For Victims As Well As Prevention And Reformation
NC	1996	Amendment 2	86	Probation, Restitution, Community Service, Work Programs and Other Restraints on Liberty May Be Imposed Upon Conviction of Criminal Offense
NC	1996	Amendment 3	78	Victims' Rights

Table A.6

Ballot Propositions Used to Measure Penal Preferences (con'd)

State	Year	Prop No.	Percent Yes	Description
NE	2006	Amendment 4	56	To Permit Supervision of Individuals Sentenced To Probation, Released on Parole, or Enrolled In Court Programs as Provided By Leg
NJ	2000	Public Question 2	79	To Permit Leg To Auth By Law Disclosure Of Information Concerning Sex Offenders
NV	1996	Question 2	74	To Provide Specifically For Rights of Victims of Crime?
OH	1997	Issue 1	73	Denial of Bail In Felony Offenses
OH	2002	Issue 1	32	Treatment in lieu of Incarceration for Drug Offenders
OK	1994	Question 664	91	To Allow the Legislature to set Minimum Prison Terms for All Convicted Felons
OR	1996	Measure 26	66	To Change Principles That Govern Laws for Punishment of Crime
OR	1996	Measure 40	58	To Give Crime Victims Rights, Expands Admissible Evidence, Limits Pretrial Release
OR	1999	Measure 68	58	To Allow Protecting Business, Certain Government Programs from Prison Work Programs
OR	1999	Measure 69	58	To Grant Victims Constitutional Rights In Criminal Prosecutions, Juvenile Court Delinquency Proceedings
OR	1999	Measure 71	58	To Limit Pretrial Release of Accused Person To Protect Victims
OR	1999	Measure 72	45	To Allow Murder Conviction by 11 to 1 Jury Verdict
OR	1999	Measure 73	46	To Limit Immunity from Criminal Prosecution of Person Ordered To Testify about his or her Conduct
OR	1999	Measure 74	53	To Require Terms of Imprisonment Announced in Court Be Fully Served, With Exceptions
OR	1999	Measure 75	57	To Prohibit Persons Convicted of Certain Crimes from Serving on Grand Juries, Criminal Trial Juries
OR	2000	Measure 3	67	To Require Conviction Before Forfeiture; Restricts Proceeds Usage; Requires Reporting, Penalty
OR	2000	Measure 94	26	To Repeal Mandatory Minimum Sentences for Certain Felonies, Requires Re-sentencing
OR	2008	Measure 57	61	To Increase Sentences for Drug Trafficking, Theft against Elderly and To Specify Repeat Property and Identity Theft Crimes
OR	2008	Measure 61	48	To Create Mandatory Minimum Prison Sentences for Certain Theft, Identity Theft, Forgery, Drug and Burglary Crimes
PA	1998	Joint Resolution 1	72	To Add Categories of Criminal Cases in Which Bail Is Disallowed
PA	1998	Joint Resolution 2	69	To Grant Commonwealth Right to Trial By Jury in Criminal Cases
PA	2003	Amendment 1	68	To Amend Right of Persons Accused of a Crime To Meet Witness against Them Face To Face
PA	2003	Amendment 2	80	To Authorize Legislature To Enact Laws Regarding Way That Children May Testify in Criminal Proceedings
SC	1996	Amendment 1 (A)	89	Victims' Rights
SC	1996	Amendment 1 (B)	87	To Allow Denial of Bail To Persons Charged With Violent Crimes
SC	1998	Amendment 1	48	To Allow Leg To Specify Which Crime Victims Are Protected By Victims Bill Of Rights
SD	2002	Amendment A	21	Criminal Defendants' Rights
TN	1998	Amendment 2	89	To Entitle Victims of Crime To Certain Basic Rights To Preserve and To Protect Their Rights To Justice, Due Process In All Cases including Criminal Cases
UT	1994	Proposition 1	69	Rights of Crime Victims
WA	1993	Initiative 593	76	Sentencing of Criminals
WI	2006	Question 2	55	To Reinstate Death Penalty

Table A.7

Summary Statistics of Key Variables by Judicial Selection System

Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A: All Sample					
Log Number of Articles	2363	0.71	0.77	-3.46	2.51
Congruence	2693	0.52	0.25	0.00	0.97
Harsh Vote Share	1982	0.00	0.04	-0.27	0.13
Democratic Vote Share	2694	0.48	0.12	0.15	0.90
Log Population	2699	5.41	0.53	3.24	7.00
Log Area	2697	3.08	0.52	1.18	4.49
Log Per Capita Income	2693	10.11	0.26	8.26	11.36
Log Number of Crimes	2587	3.57	0.29	1.43	4.43
Panel B: Non-partisan Election System					
Log Number of Articles	984	0.73	0.79	-3.46	2.22
Congruence	1130	0.52	0.25	0.00	0.97
Harsh Vote Share	798	0.00^{***}	0.05	-0.27	0.08
Democratic Vote Share	1126	0.48	0.11	0.19	0.83
Log Population	1134	5.43^{*}	0.53	3.97	7.00
Log Area	1132	3.16^{***}	0.47	1.67	4.41
Log Per Capita Income	1134	10.12	0.23	9.41	10.93
Log Number of Crimes	1082	3.60^{***}	0.25	1.62	4.16
Panel C: Appointment System					
Log Number of Articles	498	0.69	0.75	-3.07	1.63
Congruence	559	0.50^{*}	0.24	0.01	0.95
Harsh Vote Share	462	0.01^{***}	0.04	-0.08	0.13
Democratic Vote Share	564	0.49^{**}	0.11	0.15	0.90
Log Population	561	5.50^{***}	0.41	4.35	6.54
Log Area	561	3.04^{**}	0.70	1.18	4.49
Log Per Capita Income	555	10.17^{***}	0.33	8.26	10.88
Log Number of Crimes	552	3.64^{***}	0.22	2.61	4.23
Panel D: Partisan Election System					
Log Number of Articles	881	0.70	0.76	-1.85	2.51
Congruence	1004	0.53	0.26	0.00	0.97
Harsh Vote Share	722	0.00^{*}	0.04	-0.13	0.10
Democratic Vote Share	1004	0.48	0.12	0.18	0.89
Log Population	1004	5.34^{***}	0.58	3.24	6.73
Log Area	1004	3.01^{***}	0.44	1.36	4.17
Log Per Capita Income	1004	10.08^{***}	0.25	9.35	11.36
Log Number of Crimes	953	3.49^{***}	0.33	1.43	4.43

Note 1: This table is based on the data used in Table 8 in the main text. The unit of observation is judicial district by year.
Note 2: *, **, and *** denote statistical significance of the difference between the given system and the other systems at 10%, 5%, and 1% levels under t-test.

Table A.8

Relationships between *Harsh Vote Share*, *Congruence*, *Harshness*, and Key Control Variables

	All Sample			Non-partisan		
	HVS	Congruence	Harshness	HVS	Congruence	Harshness
	(1)	(2)	(3)	(5)	(5)	(6)
Congruence			0.057**			0.116***
			(0.026)			(0.040)
Harsh Vote Share (HVS)		-0.052	0.101		-0.627	0.178**
		(0.284)	(0.062)		(0.386)	(0.088)
Democratic Vote Share	-0.192***	-0.386***	0.012	-0.257***	-0.490**	0.033
	(0.022)	(0.138)	(0.040)	(0.038)	(0.192)	(0.058)
Log Population	0.001	0.170***	-0.041***	0.005	0.121***	-0.040***
	(0.006)	(0.029)	(0.009)	(0.012)	(0.045)	(0.011)
Log Area Size	0.001	0.138***	-0.026***	0.019	0.213***	-0.034**
	(0.010)	(0.034)	(0.010)	(0.019)	(0.059)	(0.016)
Log Per Capita Income	0.009	-0.182***	0.039***	0.013	0.094	0.058***
	(0.009)	(0.052)	(0.014)	(0.021)	(0.103)	(0.022)
Log Number of Crimes	0.019***	0.256***	0.026	-0.001	0.189***	0.036
	(0.007)	(0.049)	(0.020)	(0.015)	(0.066)	(0.028)
Observations	1,935	1,924	163,584	769	769	95,515
R^2	0.341	0.433	0.078	0.380	0.401	0.063
Unit of Observation	District-Year	District-Year	Case	District--Year	District--Year	Case

	Appointment			Partisan		
	HVS	Congruence	Harshness	HVS	Congruence	Harshness
	(7)	(8)	(9)	(10)	(11)	(12)
Congruence			-0.022			0.010
			(0.030)			(0.028)
Harsh Vote Share		0.531	0.060		0.900	0.072
		(0.610)	(0.247)		(0.588)	(0.157)
Democratic Vote Share	-0.107***	-0.490	0.060	-0.164***	-0.399**	-0.036
	(0.026)	(0.423)	(0.094)	(0.029)	(0.183)	(0.056)
Log Population	0.031***	0.098	-0.098***	-0.015**	0.233***	-0.028*
	(0.012)	(0.088)	(0.033)	(0.006)	(0.044)	(0.014)
Log Area Size	-0.010	0.082	0.006	0.026*	0.146*	-0.034**
	(0.007)	(0.054)	(0.020)	(0.015)	(0.087)	(0.016)
Log Per Capita Income	-0.001	-0.141**	-0.006	0.021	-0.509***	0.026
	(0.007)	(0.063)	(0.024)	(0.017)	(0.079)	(0.038)
Log Number of Crimes	0.011	0.235	0.010	0.043***	0.312***	0.029
	(0.013)	(0.156)	(0.038)	(0.009)	(0.087)	(0.028)
Observations	444	433	24,153	722	722	43,916
R^2	0.536	0.479	0.047	0.391	0.502	0.095
Unit of Observation	District-Year	District-Year	Case	District--Year	District--Year	Case

Note 1: OLS regression results. Standard errors, clustered by judicial district, in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%.

Note 2: Columns (1), (4), (7), and (10) include state fixed effects. The rest of the columns include state-by-year fixed effects.

TABLE A.9

Instrumental Variable Regression of *Harshness* on Coverage, Penal Preferences, and Selection Systems

Panel A: Overall Relationships between Media Coverage and Sentencing						
	All Sample			Year 2004		
	(1)	(2)	(3)	(4)	(5)	(6)
Log Number of Articles	0.043**	0.068*	0.050**	0.065**	0.069*	0.047***
	(0.018)	(0.040)	(0.023)	(0.028)	(0.036)	(0.018)
Harsh Vote Share		0.087	0.090		-0.088	-0.021
		(0.121)	(0.124)		(0.323)	(0.261)
Democratic Vote Share		0.040	0.056		0.053	-0.049
		(0.067)	(0.061)		(0.139)	(0.091)
Appointed		-0.113	-0.206		-0.502***	-0.097
		(0.079)	(383.800)		(0.165)	(0.071)
Partisan Elected		-0.053	-0.153		-0.257***	0.189
		(0.081)	(435.554)		(0.081)	(0.144)
Observations	212,837	147,497	92,229	27,922	19,340	13,140
R^2	0.132	0.126	0.114	0.081	0.092	0.085
Controls	No	Yes	Yes	No	Yes	Yes
Trimmed Sample	No	No	Yes	No	No	Yes

Panel B: Media Influence by Judicial Selection System						
	Non-partisan		Appointment		Partisan	
	(1)	(2)	(3)	(4)	(5)	(6)
Log Number of Articles	0.030*	0.062**	-0.034	-0.398	0.065*	0.003
	(0.016)	(0.025)	(0.035)	(6.576)	(0.035)	(0.055)
Harsh Vote Share		-0.019		-3.996		0.043
		(0.171)		(54.925)		(0.164)
Democratic Vote Share		-0.084		0.307		-0.062
		(0.082)		(5.767)		(0.078)
Observations	96,004	83,104	26,347	22,854	90,486	41,539
R^2	0.123	0.130	0.061		0.158	0.151
Controls	No	Yes	No	Yes	No	Yes
Trimmed Sample	No	No	No	No	No	No

Note 1: IV regression results. Standard errors, clustered by judicial district, are in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%.

Note 2: Columns (1)-(3) in Panel A and all six columns in Panel B use the entire data period (1986-2006).

Note 3: R^2 in Column (4) of Panel B is not shown because it is negative.

Table A.10
Sensitivity Analysis 1

Dependent Variable : *Harshness*						
Panel A: *Log Number of Articles* as Media Variable						
	Baseline	OH	MD	OHMD	post90	TX
	(1)	(2)	(3)	(4)	(5)	(6)
Log Number of Articles	0.017***	0.017***	0.017***	0.017***	0.017***	0.018***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Harsh Vote Share	0.113	0.113	0.113	0.113	0.086	0.093
	(0.090)	(0.090)	(0.090)	(0.090)	(0.094)	(0.098)
Democratic Vote Share	-0.013	-0.013	-0.013	-0.013	-0.026	-0.028
	(0.044)	(0.044)	(0.044)	(0.044)	(0.045)	(0.054)
Appointed	-0.086***	-0.086***	-0.086***	-0.086***	-0.090***	-0.088***
	(0.032)	(0.032)	(0.032)	(0.032)	(0.034)	(0.032)
Partisan Elected	-0.011	-0.011	-0.011	-0.011	-0.018	-0.014
	(0.025)	(0.025)	(0.025)	(0.025)	(0.027)	(0.024)
Observations	147,497	147,497	147,497	147,497	135,228	132,027
R^2	0.131	0.131	0.131	0.131	0.131	0.129
Panel B: *Congruence* as Media Variable						
	Baseline	OH	MD	OHMD	post90	TX
	(1)	(2)	(3)	(4)	(5)	(6)
Congruence	0.051**	0.051**	0.051**	0.051**	0.048**	0.059**
	(0.023)	(0.023)	(0.023)	(0.023)	(0.024)	(0.024)
Harsh Vote Share	0.142	0.142	0.142	0.142	0.090	0.104
	(0.091)	(0.091)	(0.091)	(0.091)	(0.093)	(0.099)
Democratic Vote Share	-0.021	-0.021	-0.021	-0.021	-0.044	-0.016
	(0.042)	(0.042)	(0.042)	(0.042)	(0.040)	(0.049)
Appointed	-0.102***	-0.102***	-0.102***	-0.102***	-0.110***	-0.100***
	(0.023)	(0.023)	(0.023)	(0.023)	(0.025)	(0.024)
Partisan Elected	-0.022	-0.022	-0.022	-0.022	-0.029	-0.020
	(0.027)	(0.027)	(0.027)	(0.027)	(0.028)	(0.027)
Observations	163,551	163,551	163,551	163,551	149,191	147,962
R^2	0.128	0.128	0.128	0.128	0.128	0.126

Note 1: Panels A and B present robustness checks of the results in Columns (2) and (5) in Panel A of Table 8 in the main text, respectively. Specifications of the robustness checks (OH, MD, OHMD, post90, and TX) are described on page 14 of this supplementary material.
Note 2: OLS regression results. Standard errors, clustered by judicial district, in parentheses: * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent. All specifications include state-by-year fixed effects.
Note 3: The unit of observation is individual felony case.

Table A.11
Sensitivity Analysis 2

	Dependent Variable: *Harshness*					
	Baseline	OH	MD	OHMD	post90	TX
	(1)	(2)	(3)	(5)	(6)	(6)
Congruence	0.098***	0.100***	0.098***	0.100***	0.085**	0.098***
	(0.036)	(0.036)	(0.036)	(0.036)	(0.037)	(0.036)
Congruence*Appointed	-0.120***	-0.123***	-0.120***	-0.123***	-0.098**	-0.120***
	(0.046)	(0.046)	(0.046)	(0.046)	(0.048)	(0.046)
Congruence*Partisan Elected	-0.096**	-0.109**	-0.096**	-0.109**	-0.094**	-0.106**
	(0.045)	(0.046)	(0.045)	(0.046)	(0.046)	(0.048)
Harsh Vote Share	0.092	0.191	0.092	0.191	0.036	0.009
	(0.134)	(0.137)	(0.134)	(0.137)	(0.132)	(0.218)
Harsh Vote Share*Non-Partisan Elected	0.185	0.078	0.185	0.078	0.163	0.268
	(0.202)	(0.206)	(0.202)	(0.206)	(0.200)	(0.265)
Harsh Vote Share*Appointed	-0.803***	-0.900***	-0.803***	-0.900***	-0.776***	-0.722**
	(0.281)	(0.283)	(0.281)	(0.283)	(0.286)	(0.330)
Democratic Vote Share	-0.053	-0.028	-0.053	-0.028	-0.080	-0.032
	(0.065)	(0.065)	(0.065)	(0.065)	(0.064)	(0.077)
Democratic Vote Share*Non-Partisan Elected	-0.020	-0.058	-0.020	-0.058	-0.015	-0.040
	(0.100)	(0.102)	(0.100)	(0.102)	(0.099)	(0.108)
Democratic Vote Share*Appointed	0.038	0.013	0.038	0.013	0.082	0.017
	(0.104)	(0.103)	(0.104)	(0.103)	(0.103)	(0.111)
Appointed	2.150**	2.074*	2.150**	2.074*	2.384**	2.148**
	(1.087)	(1.099)	(1.087)	(1.099)	(1.109)	(1.088)
Partisan Elected	1.603*	1.527	1.603*	1.527	1.476	0.187
	(0.933)	(0.944)	(0.933)	(0.944)	(0.989)	(0.996)
Constant	-1.537**	-1.464**	-1.537**	-1.464**	-1.584**	-1.539**
	(0.720)	(0.736)	(0.720)	(0.736)	(0.762)	(0.720)
Observations	163,551	163,551	163,551	163,551	149,191	147,962
R^2	0.132	0.132	0.132	0.132	0.132	0.131
Congruence Partisan Elected	.952	.749	.952	.749	.736	.812
Congruence Appointed	.441	.44	.441	.44	.655	.446
Harsh Non-Partisan Elected	.067	.081	.067	.081	.185	.066
Harsh Appointed	.004	.004	.004	.004	.004	.004

Note 1: This table presents robustness checks of the results in Panel B of Table 8 in the main text, using interactions instead of sample splits. Specifications of the robustness checks (OH, MD, OHMD, post90, and TX) are described on page 14 of this supplementary material. All the columns include control variables identical to those in Table 8.
Note 2: OLS regression results. Standard errors, clustered by judicial district, in parentheses: * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent. All specifications include state-by-year fixed effects.
Note 3: The unit of observation is individual felony case.

Table A.12
Sensitivity Analysis 3

Dependent Variable: *Harshness*					
Panel A: *Log Number of Articles* as Media Variable					
	(1)	(2)	(3)	(4)	(5)
Log Number of Articles	0.017***	0.015***	0.016***	0.014***	0.017***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Harsh Vote Share	0.113	0.111	0.180**	0.171*	0.123
	(0.090)	(0.090)	(0.089)	(0.090)	(0.093)
Democratic Vote Share	-0.013	-0.002	0.053	0.059	0.001
	(0.044)	(0.041)	(0.054)	(0.053)	(0.045)
Appointed	-0.086***	-0.092***	-0.076**	-0.077**	-0.077**
	(0.032)	(0.033)	(0.032)	(0.034)	(0.030)
Partisan Elected	-0.011	-0.014	-0.009	-0.005	0.002
	(0.025)	(0.024)	(0.025)	(0.026)	(0.025)
Observations	147,497	147,497	147,455	147,455	138,448
R^2	0.131	0.131	0.131	0.131	0.126
Controls	1	2	3	4	1
Sample	Full	Full	Full	Full	Non-Urban
Panel B: *Congruence* as Media Variable					
	(1)	(2)	(3)	(4)	(5)
Congruence	0.051**	0.052**	0.040	0.047*	0.056**
	(0.023)	(0.026)	(0.025)	(0.027)	(0.025)
Harsh Vote Share	0.142	0.143	0.180**	0.182**	0.176*
	(0.091)	(0.090)	(0.089)	(0.089)	(0.097)
Democratic Vote Share	-0.021	-0.011	0.012	0.027	0.001
	(0.042)	(0.040)	(0.046)	(0.045)	(0.042)
Appointed	-0.102***	-0.109***	-0.097***	-0.100***	-0.102***
	(0.023)	(0.022)	(0.027)	(0.025)	(0.023)
Partisan Elected	-0.022	-0.024	-0.023	-0.019	-0.014
	(0.027)	(0.026)	(0.028)	(0.029)	(0.026)
Observations	163,551	163,551	163,509	163,509	152,868
R^2	0.128	0.128	0.128	0.128	0.124
Controls	1	2	3	4	1
Sample	Full	Full	Full	Full	Non-Urban

Note 1: Panels A and B present robustness checks of the results in Columns (2) and (5) in Panel A of Table 8 in the main text, respectively.
Note 2: OLS regression results. Standard errors, clustered by judicial district, in parentheses: * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent. All specifications include state-by-year fixed effects.
Note 3: The unit of observation is individual felony case.
Note 4: Controls=1 means baseline set of controls ; Controls=2 means baseline plus the four variables from (b) and (c) on page 15; Controls=3 means baseline plus MSA demographics ; Controls=4 means baseline plus the four variables from (b) and (c) on page 15 and MSA demographics. Column (5) studies the sample excluding the 100% urban districts.

Table A.13
Sensitivity Analysis 4

Dependent Variable: *Harshness*					
	(1)	(2)	(3)	(4)	(5)
Congruence	0.098***	0.104**	0.103***	0.112***	0.107***
	(0.036)	(0.042)	(0.036)	(0.041)	(0.039)
Congruence*Appointed	-0.120***	-0.124**	-0.132**	-0.105	-0.095**
	(0.046)	(0.062)	(0.057)	(0.082)	(0.048)
Congruence*Partisan Elected	-0.096**	-0.097*	-0.117***	-0.101**	-0.111**
	(0.045)	(0.051)	(0.044)	(0.051)	(0.047)
Harsh Vote Share	0.092	0.067	0.158	0.128	0.140
	(0.134)	(0.132)	(0.125)	(0.125)	(0.133)
Harsh Vote Share*Non-Partisan Elected	0.185	0.218	0.181	0.217	0.105
	(0.202)	(0.198)	(0.202)	(0.205)	(0.202)
Harsh Vote Share*Appointed	-0.803***	-0.580**	-0.518	-0.463	-0.628**
	(0.281)	(0.259)	(0.366)	(0.371)	(0.267)
Democratic Vote Share	-0.053	-0.049	-0.061	-0.068	-0.055
	(0.065)	(0.066)	(0.061)	(0.062)	(0.067)
Democratic Vote Share*Non-Partisan Elected	-0.020	0.010	0.009	0.044	-0.000
	(0.100)	(0.098)	(0.117)	(0.115)	(0.099)
Democratic Vote Share*Appointed	0.038	0.050	0.095	0.106	0.119
	(0.104)	(0.109)	(0.134)	(0.139)	(0.102)
Appointed	2.150**	2.680**	-1.201	-1.640	1.852*
	(1.087)	(1.115)	(2.779)	(2.886)	(1.114)
Partisan Elected	1.603*	1.426	1.540	1.172	2.151**
	(0.933)	(0.973)	(1.139)	(1.158)	(1.017)
Observations	163,551	163,551	163,509	163,509	152,868
R^2	0.132	0.132	0.133	0.133	0.128
Controls	1	2	3	4	1
Sample	Full	Full	Full	Full	Non-Urban
Congruence Partisan Elected	.952	.812	.571	.728	.892
Congruence Appointed	.441	.659	.518	.922	.639
Harsh Non-Partisan Elected	.067	.053	.034	.035	.11
Harsh Appointed	.004	.020	.297	.339	.035

Note 1: This table presents robustness checks of the results in Panel B of Table 8 in the main text, using interactions instead of sample splits.
Note 2: OLS regression results. Standard errors, clustered by judicial district, in parentheses: * significant at 10 percent; ** significant at 5 percent; *** significant at 1 percent. All specifications include state-by-year fixed effects.
Note 3: The unit of observation is individual felony case.
Note 4: Controls=1 means baseline set of controls ; Controls=2 means baseline plus the four variables from (b) and (c) on page 15; Controls=3 means baseline plus MSA demographics ; Controls=4 means baseline plus the four variables from (b) and (c) on page 15 and MSA demographics. Column (5) studies the sample excluding the 100% urban districts.

Table A.14

Partisan Endorsements and Selection Systems

	Dependent Variable			
	Number of Endorsements	Number of Endorsements	Share of Democratic Endorsements	Share of Democratic Endorsements
	(1)	(2)	(3)	(4)
Partisan Elected	0.049	0.171	0.012	0.075
	(0.102)	(0.109)	(0.043)	(0.051)
Appointed	0.075	0.094	0.076	0.086
	(0.115)	(0.116)	(0.049)	(0.053)
Observations	631	629	615	613
R^2	0.001	0.207	0.004	0.064
Controls	No	Yes	No	Yes

Note 1: OLS regression results. Standard errors, clustered by newspaper, are in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%.
Note 2: The unit of observation is newspaper.
Note 3: Control variables are log population, log area size, log per capita income, log employment, the share of people in the district who are religious adherents, female, younger than 20, older than 65, black, white, Hispanic, urban, the share with high school education, the share with more than high school education, and turnout in the most recent presidential election.

TABLE A.15

Regression of Coverage, Penal Preferences, and *Harshness* on *Congruence*, by Partisan Endorsement

	Dependent Variable				
	Number of Articles	Harsh Vote Share	Harsh-ness	Harsh-ness	Harsh-ness
	(1)	(2)	(3)	(4)	(5)
Congruence	15.089***	-0.005	0.001	0.023	0.039
	(2.844)	(0.008)	(0.019)	(0.024)	(0.026)
Congruence, Democratic papers	16.188***	-0.013*	0.040**	0.031*	0.030*
	(3.922)	(0.007)	(0.017)	(0.017)	(0.017)
Congruence, Republican papers	15.980***	0.002	0.037**	0.044***	0.031**
	(3.747)	(0.006)	(0.018)	(0.014)	(0.015)
Observations	1,413	1,177	232,470	163,551	100,983
R^2	0.198	0.880	0.128	0.128	0.115
Controls	Yes	Yes	No	Yes	Yes
Congruence Dem.	.000	.076	.056	.021	.010
Congruence Rep.	.000	.748	.052	.007	.011
Trimmed Sample	No	No	No	No	Yes

Note 1: OLS regression results. The unit of observation is judicial district in Column (1) and (2), and individual case in Columns (3)-(5). Standard errors, clustered by the unit of observation, are in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%.
Note 2: In Column (1), the set of control variables is identical to that in Table 6 in the main text. In Column (2), it is identical to that in Table 4 in the main text. In Columns (3)-(5), it is identical to that in Table 8 in the main text.

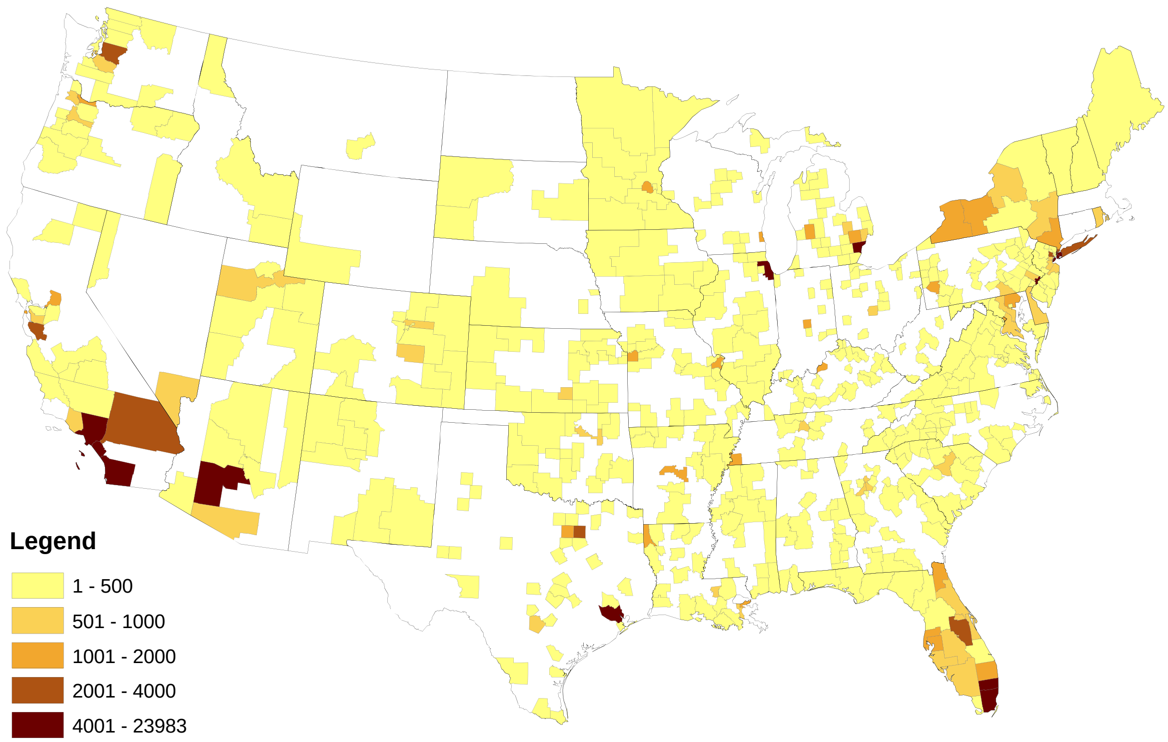

Figure A.1
Number of Sentences by Judicial District