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

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

December 18, 2012

Abstract

We study the influence of media coverage on the behavior of public officials, focusing on elected and appointed U.S. state trial court judges. We develop a model where media coverage increases the responsiveness of sentencing decisions to ordinary citizens' preferences, at the expense of judges and special interests. This effect is largest for non-partisan elected judges, followed by partisan elected and appointed judges. We test this using data on 2 million sentences in the National Judicial Reporting Program from 1986 to 2006 and newly collected data on the coverage of 9,828 trial court judges in 1,400 newspapers. Since newspaper coverage may be endogenous, we use the match between the newspaper markets and the judicial districts to identify effects. We find that press coverage significantly increases sentence length. This is driven by non-partisan elected judges in cases with violent crimes. For partisan elected and appointed judges, there are no significant effects. Additionally, we find that that newspaper coverage does not affect the public's penal preferences.

Keywords: Court, Media, Sentencing, Crime, Appointment, Election, Voter Information

JEL Classification: H1, H7, K4, L8

---

*This paper was previously circulated under the title "Measuring Media Influence on U.S. State Courts." We thank seminar participants at UC-Berkeley, Northwestern-Kellogg School, Stanford GSB, U.Illinois, U. Chicago Harris School, and conference participants at the ASSA annual meeting 2010, CIPREE Political Economy Workshop, and Erasmus Political Economy Workshop for their suggestions and comments. We also thank Evelina Persson and Christopher Stanton for excellent research assistance. The research conducted for this article was supported by a research grant from the American Bar Association (ABA) Section of Litigation; however, the views expressed here are not intended to represent ABA positions or policies.

†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

1 Introduction

This paper investigates how the media affects the different accountability structures put in place to select public officials. This issue can be studied in detail in the case of the U.S. state trial court judges, where a variety of selection systems exists. Some judges are appointed by the governor, some are elected in normal partisan elections after being nominated by political parties, and, most commonly, some are elected in non-partisan elections where they compete without party-identification on the ballot.[1]

State trial court judges exercise enormous power in the U.S. judicial system. State courts deal with more than 90 percent of civil and felony cases in the U.S. In 2006, they convicted over one million felons to a total of over two million years in prison (Bureau of Justice Statistics, 2009). In the state judicial system, while juries have the power to convict, judges have the authority to impose sentences, and only a small fraction of felony cases are reviewed by appellate courts. Consequently, the decisions of trial court judges are of paramount importance, and therefore, so are the selection and incentive structures these judges face.

The media may matter because the citizens who are to monitor judges – and in most cases also elect them – have little reason to gather information unless they are personally involved with the courts. The vast majority of voters say that they have insufficient information about judicial candidates (Sheldon and Lovrich, 1999). The media sometimes provide this information, but often does not. As we show empirically, the amount of press coverage about judges varies tremendously, from none to hundreds of articles per newspaper, judge and year in our sample.

Our paper begins with a simple model of political accountability. The model suggests that variations in media coverage alter the functioning of different selection methods, and ultimately affect sentencing decisions. Briefly, non-partisan elected judges behave more like appointed judges when there is little media coverage and more like partisan-elected judges when there is ample media coverage. In our model, voters try to select judges with preferences similar to their own. For this they

[1] The influence of judicial selection systems on rulings in lawsuits has also been a major public policy concern. For example, a U.S. Supreme Court case, *Caperton v. A.T. Massey Coal Co.*, illustrates how campaigning for elections might create biases in judicial decisions that affect businesses. The case deals with a situation where a judge presided over a case in which one of the litigants was a company that provided campaign funds to the judge in his first election. For details, see http://www.supremecourt.gov/opinions/08pdf/08-22.pdf. In response to the public concern, the former U.S. Supreme Court Justice Sandra Day O'Connor has campaigned to remove direct elections of judges. See http://www.nytimes.com/2009/12/24/us/24judges.html.

need information, and hence better informed voters are more influential. There are two types of voters: special interest voters who care intensely about sentencing (e.g. bar associations), and ordinary voters. Special interest voters are always well informed.

The media reduces the informational advantage of the special interest voters by informing the ordinary voters. Thus, the media increases the policy influence of the ordinary voters in all selection systems. However, the impact of the media varies depending on the system used to select and retain judges. It is largest in the case of judges chosen in non-partisan elections, smaller for judges chosen in partisan elections, and smallest for appointed judges. The media has more influence in non-partisan elections than partisan elections, because in the latter case there exists an additional factor, party affiliation of the judicial candidates, that affects voters' decisions. Therefore, the influence of information on judges' behavior from the media is smaller. It may seem self-evident that appointed officials are the least affected by media. However, those who appoint the officials are themselves elected, and media scrutiny of bad practices by the latter would reflect poorly on the former. In our model, media coverage of appointed officials matters less because the media mainly inform the ordinary voters, and these voters matter less than special interest voters in the multi-issue election of the appointing politician. As argued by Besley and Coate (2003), in multiple-issue elections, voters trade off utility in one area against another, whereas in single-issue elections they can always vote according to their preferences on each issue. In the absence of media coverage, non-partisan elected judges cater more to the special interests (as do appointed judges), and increasing media coverage make them cater more to the interest of ordinary voters (as do partisan elected judges).

The model provides several testable hypotheses. There is substantial survey evidence that the ordinary voters believe that criminal sentences are too lenient. Assuming this to be the case, the following hypotheses hold. First, our main hypothesis is that media coverage increases sentence length, and the effect is largest for non-partisan elected judges, second-largest for partisan elected judges, and smallest for appointed judges. Second, our model suggests that the media's effect on sentencing is increasing in the overall amount of media coverage. This implies that sentences for the most serious violent crimes – which receive the most coverage – should be particularly affected by the media.

We investigate these hypotheses using data on 2 million sentences between 1986 and 2006 collected within the National Judicial Reporting Program. We combine this with newly collected data on the coverage of 9,828 trial court judges in 1,400 newspapers during 2004 and 2005. We find an average of 9 newspaper articles covering each judge each year. We also find that the variation is very large; one standard deviation is 21 articles. Since newspaper coverage of judges may be endogenous, we use the "match" between newspaper markets and judicial districts to identify effects. This match has a large effect on coverage (as we show), and is more likely to be exogenous than newspaper coverage. Under the weaker assumption that the correlation between this match and the error term does not depend on the judicial selection system, we can also consistently estimate a lower bound for the effect for non-partisan elected judges.

We find that press coverage significantly increases sentence length. The effects of coverage are sizable. A one standard deviation increase in the match between the media market and the judicial district – which translates into 7 more articles per judge in the judicial district – is estimated to increase the average sentence length for murders, rapes and robberies by about 6 months. Thus, a uniform increase in this match by one standard deviation is estimated to increase total sentence length in 2006 by over 40,000 prison years. We find that the media effects are monotonically increasing in the ratio of newspaper articles to convictions: highest for the most violent crimes, followed by all violent crimes, property crimes and drug related crimes. Also, the results are driven by non-partisan elected judges. The estimated effects are significantly lower for appointed and partisan elected judges; in fact, they are not significantly different from zero for either of these two sub-samples. This suggests that information about the party affiliation of judicial candidates has such a strong affect on voters' behavior that the effect of additional information by the media is insignificant. Finally, we find no evidence that newspaper coverage of the courts affects the public's penal preferences.

Our results are closely related to Maskin and Tirole (2004), who argue that elected public officials may become too responsive and select policies that they know are bad in order to pander to public opinion. Our results indicate that the media may enforce this tendency among non-partisan elected public officials, while leaving other officials relatively less affected. Also, it is not clear that increasing accountability to the ordinary voters via media coverage is generally welfare enhancing. In single-issue elections, there is a risk that the media may enforce a "tyranny of the majority" against the interests of a minority with more intense preferences on the issue.

Our paper contributes to the literature on the functioning of judicial selection mechanisms. Hall (2001) and Bonneau and Hall (2009) document statistics of various types of judicial elections, such as the defeat rate of incumbents, the average vote share of winners, and the amount of campaign spending. Several studies also document the empirical relationship between selection mechanisms and court decisions, e.g., Hanssen (1999, 2000), Huber and Gordon (2004, 2007), Lim (2012) and Tabarrok and Helland (1999).[2] Our paper enriches the understanding of the functioning of these institutions by uncovering their interaction with the media environment, which (to our knowledge) has not been done.

Our paper also contributes to the literature on how the media influences elections and the behavior of policy makers. Dyck, Moss, and Zingales (2008) find that U.S. House Representatives were more pro-consumer in their voting when the circulation of muckraking magazines in the early 1900s was larger. In other words, they find that media increased the political influence of ordinary voters at the expense of the industry special interests. Snyder and Strömberg (2010) find that newspaper coverage makes House Representatives more accountable, and Besley and Burgess (2002) and Strömberg (2004) find that the introduction of radio influenced government expenditures and voter turnout. Finally, Gentzkow (2006) find that the introduction of television affected voter turnout and DellaVigna and Kaplan (2007) find that Fox News influenced the 2000 presidential election. We add to this literature by documenting the impact of media exposure on judicial decisions.

The remainder of this paper is organized as follows. In the next section, we discuss the institutional background of the U.S. state court system. In Section 3, we present our model. In Section 4, we describe our data. In Section 5, we document the main results. The final section concludes.

2 Institutional background

Table I shows the judicial selection mechanisms used by state trial courts. Currently, there are three major judicial selection mechanisms. The most common is the non-partisan election system, where multiple candidates compete without party identification on the ballot, and the top two vote-getters compete against each other in general elections (i.e., there are runoff elections). In the partisan election system, judges are selected by usual competitive elections. That is, judicial candidates seek nomination from political parties in primaries, and candidates nominated by parties compete in

[2] Huber and Gordon (2004, 2007) document the effect of electoral proximity on sentencing harshness. In contrast, we do not find evidence of the influence of electoral proximity. For a richer discussion on electoral cycles in court decisions, see Berdejó and Chen (2012) in which they argue that electoral cycles may not necessarily be evidence of reelection incentives but rather a consequence of ideological priming.

general elections. Finally, some judges are initially appointed by the governor or legislature, and when their terms expire they either must either be re-appointed by the governor, or they must run in non-competitive, "retention" elections and be approved by a majority of voters in a yes-or-no vote. A few states use systems that do not fall into one of the above three categories. For example, in Illinois, New Mexico, and Pennsylvania, judges must run in partisan elections for their initial term, and then run in retention elections for subsequent terms. There are also three states in New England region, New Hampshire, Rhode Island, and Massachusetts, in which judges are selected by gubernatorial appointment and life-tenured.[3]

[Table I here.]

This variation in judicial selection procedures has emerged over the nation's history. For the first 50 years after U.S. independence, all states appointed their judges; subsequently, partisan elections became increasingly popular, followed by non-partisan elections. One key driver of judicial reform has been changes in beliefs about the desired degree of judicial independence, and how each system delivers this (Hanssen, 2004a, 2004b). Thus, although many states have changed their selection procedures at some point, the time that a state entered the Union is a strong predictor of the type of selection system used today.

3 Model

We develop a simple model of media influences on courts under the 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 ("reelection incentive effect").[4] The media may

[3]We abstract from the difference between appointed judges with life-tenure and those who run for retention elections. Although the two procedural rules may seem quite different, in practice judges almost never fail in retention elections. Hall (2001), Lim (2012) and Lim and Snyder (2012) document that incumbent judges win retention elections more than 99% of the time.

[4]Lim (2012) shows that both the selection and reelection incentives affect judges' decisions substantially, using a structural analysis of sentencing behavior by Kansas trial court judges. She finds significant differences in the intrinsic preferences of appointed and elected judges. She also finds that reelection incentives result in substantial variation

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.[5] The media in our model is not a strategic actor, but simply provides truthful information about the sentencing behavior of judges. In the model some voters are informed and some are uninformed, and an increase in media coverage of judges increases the share of voters who are informed.

3.1 Preferences and Timing

The 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 preference type and the average preferred sentence in the population is $\overline{h}$. There are two parties, R and D. A share $\gamma > \frac{1}{2}$ of people who belong to party R have the harsh penal preference, and a share $1-\gamma$ of people who belong to party D have the harsh preference.

In judicial elections, voters' preference also include 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

in their sentencing decisions, even for judges with the same preferences. Huber and Gordon (2007) also analyze the sentencing behavior by Kansas judges. They focus primarily on average sentencing harshness across selection systems, while Lim (2012) primarily focuses on the variability of sentencing harshness. They argue that partisan-elected judges are harsher than appointed judges and that the reelection incentive is the dominant factor that explains this difference, by documenting electoral cycles in the sentencing behavior by partisan-elected judges.

[5] Judges run in a series of elections, and the same variable that determines whether judges are elected also determines whether they are reelected; selection and incentive effects are thus rendered identical.

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, judges' behavior is a salient issue, so they care more and are better informed about it. These special interest voters could, for example, include people working in the state judicial or law enforcement systems such as lawyers, prosecutors and police, but also jurors, criminals and victims. Special interest voters are a small share σ_s of the total 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 ordinary voters have 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 whole population 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. In order to abstract from reelection incentives and focus entirely on the selection effect, we assume that judges only care about sentencing and always select $h = \alpha_i$. That is, judges always reveal their penal preference truthfully.[6]

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 gubernatorial appointments, the governor is randomly drawn from the population in the first period. He appoints a judge of the same preference. Then, governor runs against a challenger randomly drawn from the other party.

[6]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 which would complicate the analysis without adding useful insights.

Voters vote in the gubernatorial election, observing party labels. The elected governor may then appoint a new judge or re-appoint the incumbent.

3.2 Selection and Sentencing

3.2.1 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 [7]

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

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

Under non-partisan elections, uninformed voters have no information regarding the difference in sentencing preferences of incumbent and challenger judges. 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 lenient, respectively) or, equivalently, if

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

where $\overline{h}_i$ is the aggregate preference of the informed voters, $\overline{h}_i = \sigma_s \overline{h}_s + \rho_n \overline{h}_n$. The probability of

[7] 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.

reelection in non-partisan elections is[8]

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

3.2.2 Partisan elections

Under partisan elections, voters know the party label of the incumbent judge 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 if

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

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 second terms in the expressions above. The final term is the effect of informed voters updating from the expected sentencing of a judge with a given party label to the actual sentencing.

3.2.3 Appointments

In the gubernatorial election, there are two issues, the judicial sentencing harshness (h) and the ideological issue which determines the 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 is the expected sentencing preference of the governor and h_2^e is the expected sentencing preference of the challenger in the gubernatorial race. Stochastic terms ε_{2i} and η_2 are idiosyncratic and systemic preference shocks favoring party R's ideology, respectively. The idiosyncratic ideological

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

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 ideological preference shock matters more for voters' utility than the shock to judge competence. The distribution of η_2 is consequently wider, and $\psi_g < \psi$. 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[9]

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

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 in one issue against the other. For this reason, it will be convenient to define the intensity-weighted aggregate penal preferences among the population and the 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 the unweighted counterparts, $\overline{h}$ and $\overline{h}_i$.

The probabilities of reelection for Republican and Democratic governors are, respectively,[10]

$$P_R^A = \frac{1}{2} + \psi_g \varphi_g \beta_g + \psi_g \overline{h}_v \left(h_R^e - h_D^e\right) + \overline{h}_{vi} \psi_g \left(h - h_R^e\right) \tag{6}$$

$$P_D^A = \frac{1}{2} - \psi_g \varphi_g \beta_g + \psi_g \overline{h}_v \left(h_D^e - h_R^e\right) + \overline{h}_{vi} \psi_g \left(h - h_D^e\right). \tag{7}$$

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

3.2.4 Sentencing

Given the selection rules for the second period judges, it is now straightforward to compute the expected sentence harshness in the second period.

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

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

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

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

For proof, see Appendix.

The above expressions for sentencing harshness consist of three terms. The first term captures the random draw of the incumbent from a population with average preference $\overline{h}$ and the part of selection effects unrelated to information. In the case of the partisan elections, this term includes the partisan ideological effect β. In the case of appointed judges, 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 the judges. Since everyone observes the party labels, this term has the effect of making sentencing more aligned with the average preference in the population, $\overline{h}$, in the case of partisan elected and the intensity-weighted average preference, $\overline{h}_v$, for appointed judges. For example, if $\overline{h} > 0$, then the more informative are party labels, the harsher is sentencing for partisan elected judges. 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 is γ the more informative are party labels and the more aligned is sentencing with voter preferences.

Our main interest is in the last term, which identifies the effect of information from the media on sentencing. This term has the effect of making sentencing more aligned with the average preference of the informed voters, $\overline{h}_i$, in the case of partisan elected and the intensity-weighted average preference of informed voters, $\overline{h}_{vi}$ in the case of appointed judges.

This effect is modified by a term of the form $x\left(1-x\right)$ that measures information gain from learning the judge's type perfectly. Here x is the voters prior that the judge is of the harsh type. This is π if the judge is drawn from the population and γ if the voter know that the judge is Republican ($1 - \gamma$ if the judge is Democrat). This gain is largest when the voter's 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, $(1-\gamma)\gamma < (1-\pi)\pi$.[11] As party labels become more informative, the effect of party label 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 judges sentencing. For simplicity, suppose that this relationship is linear, so that $\rho = \omega n$, where n is the number of newspaper articles n covering the judge and ω is a positive parameter. To see the effect of information from media about 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 preferences of the ordinary voters. We can now describe the effectiveness of media coverage by selection system.

Corollary 2. *The responsiveness of sentence 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.*

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

Information matters less for sentencing of partisan elected than non-partisan elected judges. The reason is that voters in partisan elections have an additional factor, i.e., candidates' party affiliation, that affects their voting, so the informational gain from knowing the sentencing of the judge is smaller. Information matters even less for sentencing of appointed judges. The reason is that the media informs the ordinary voters, who matter less for the sentencing of appointed than partisan elected judges. This is because intensity of preference matters in multi-issue elections. Formally, it

[11] 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 $\frac{1}{2}$.

follows since $\psi_g < \psi$ and $v_L < 1$.

Cases involving violent crimes are more likely to be covered in the media than others. We would expect that the change in the share of informed ordinary voters, ρ_n, due to the existence of a local newspaper covering the court would be largest for sentencing decisions on these types of cases. This implies that sentencing would be more affected by the presence of local coverage when the crime is more violent. Similarly, if media are more likely to report about cases with black defendants, then sentencing decisions on these cases would more closely follow the sentencing preferences of the ordinary voters.

The media plays two roles, increasing the accountability of judges and increasing the importance of ordinary voters relative to the special interest. More media coverage makes it less likely that judges with preferences different from that 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 ordinary voters preferences at the expense of the special interest voters.[12]

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.

[12]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.

4 Data

4.1 Unit of Analysis: Judicial Districts in the State Trial Court System

We study sentencing decisions in state trial courts.[13] In most states, the state trial court is divided into multiple judicial districts. There are approximately 1,700 judicial districts for state trial courts nationwide, with an average population of just under 170,000. Each district typically has multiple judges. On average, there are 6.6 judges per district.

Judicial districts typically consist of a collection of counties.[14] To construct data on the composition of each judicial district over time, we have collected information on the geographic boundaries of these judicial districts for the entire data period, using *The American Bench*.[15] In total, we have data on 1,413 courts.[16]

4.2 Sentencing

We use sentencing data from the National Judicial Reporting Program (NJRP). This program collects felony sentencing data from a nationally representative stratified sample of state courts.[17] The data includes jail time sentenced, as well as offense category and penal codes applied, and demographic characteristics of offenders such as age, race and gender. Data has been collected every 2 years since 1986 by the Census Bureau.[18]

[13]The state court systems typically have three layers: state supreme court, state appellate court, and state trial court. State trial courts are courts of general jurisdiction: they handle general civil and felony crime cases. State trial courts are often called district courts, circuit courts, or superior courts.

[14]Small states in New England (e.g., Maine, New Hampshire) tend to have just one judicial district covering the whole state. States in Pacific region (e.g., California) and Mid-Atlantic region (e.g., New Jersey, Pennsylvania) tend to have one judicial district covering one or two counties. The Southern and Midwestern states have judicial districts covering multiple (three or four) counties.

[15]We first allocated each county to a court using *The American Bench* 2004-2005 edition. To find out if and when each state's judicial district lines were redrawn, we contacted various state officials, typically the director of the administrative office of the judicial branch. We then used the data in the annual series of *The American Bench* to track each such change. We did not collect data on Alaska, Connecticut, Massachusetts, and Virginia, where the county is not the primary geographical unit of the judicial districts.

[16]The number of courts used in our data is smaller than the total number of courts in the nation for the following reasons. First, we exclude Alaska, Virginia, and Massachusetts, in which judicial districts are not completely county-based. Second, for Texas, we use 254 counties rather than 432 judicial districts as the main geographic unit, because multiple judicial districts can overlap for the same county.

[17]The states and counties included in the NJRP data are listed in Tables A.I-A.II in the appendix.

[18]Each survey year, approximately 300 counties are sampled, except in 1986 when 100 counties were sampled. The counties are selected through stratified sampling. Within each county, cases are randomly sampled within crime types. Since many counties are repeatedly included in the sample, the combined data set has an unbalanced panel structure.

4.3 Newspaper Coverage

We merge the sentencing data with data on the amount of press coverage of state court judges that we collected using content analysis. Our sample of judges consists of 9,828 state trial court judges in the U.S. in 2004 and 2005.[19] Our sample of newspapers consists of all 1,400 newspapers for which the articles published in 2004 and 2005 are searchable through NewsLibrary.com. For each judge in our sample and each newspaper with positive sales in the state where the judge presides, we count the number of articles that appeared in 2004 and 2005 that mention the name of judges in our sample. We use the search string { "judge N1" OR "judge N2" }, where N1 is the judge's full name including middle initial, and N2 is the judge's first and last name only. This yields the frequency of coverage for approximately 1 million judge-newspaper combinations. Since our key variables vary at the judicial district level, we aggregate the frequency of coverage to the judicial district-newspaper level. Summary statistics of the basic data are shown in Table II. On average, a newspaper in our sample writes 9 articles about each judge per year. Coverage varies considerably – the standard deviation in coverage is 21 articles.

[Table II here.]

A few other comments about coverage are noteworthy. First, to estimate the degree to which coverage of judges focuses on especially violent crime, we ran searches that included the search string {AND (murder* OR rape*)}. In our sample, about 20% of the stories contain the added string. Thus, while murder and rape are over-represented in newspapers relative to the share of criminal acts they represent, they do not dominate the coverage. Second, to estimate the degree to which coverage of judges focuses on their sentencing behavior, we ran searches that included the search string {AND sentenc*}. About 33% of the stories contain this added string. Third, inspection of a sample of 200 articles reveals that stories that are not about sentencing cover a wide range of topics, including: election campaigns; candidates' backgrounds, qualifications, and endorsements; election results; judicial procedures and reforms; prison overcrowding and building new prisons or jails; crime rates; laws on the statute of limitations; appellate court rulings; other judicial decisions such as restraining orders; and articles describing ongoing court proceedings in particular high-profile cases.

[19] We obtained the list of judges from *The American Bench*.

Fourth, we also investigated coverage of courts by local television, using stories in the *Local TV News Media Project*.[20] There appears to be very little coverage of state court judges on local television news. Searching for news stories using the word "judge" yielded just 12 hits, none of which were about sentencing.[21] Searching for the word "sentence" or "sentenced" or "sentencing" yielded 35 stories about criminal sentencing decisions or appeals, but none of these mentioned the name of the judge who passed the sentence.

4.4 Congruence

A key concern in identifying the causal effect of newspaper coverage on sentence harshness is that both may be driven by the seriousness of the crime. For this reason, we use a measure based on newspaper sales called *Congruence* to capture the intensity of newspaper coverage of the courts. We will argue that this measure is exogenous to sentencing.

Congruence measures the match between the newspaper market and the judicial district. When this match is better, in the sense that each newspaper has most of its sales in one district, then newspapers cover the judicial district more (see Snyder and Strömberg (2010) for the application of congruence to the analysis of media influence on U.S. congressmen). Formally,

$$Congruence_D = \sum_{m=1}^{M} MarketShare_{md}\, ReaderShare_{md}, \tag{10}$$

where the $ReaderShare_{md}$ is the share of newspaper m's sales who are in district d, and $MarketShare_{md}$ is newspaper m's share of total newspaper sales in district d. The logic behind this measure is that the larger the share of a newspaper's readers that live in a judicial district, the more likely is the newspaper to cover sentencing in that district. The influence of different newspapers is proportional to their market share in the district.

[20] The *Local TV News Media Project*, at the University of Delaware, contains a database with over 10,600 individually digitized stories from over 600 broadcasts of 61 stations in 20 local television markets around the country that aired during the spring of 1998. See http://www.localtvnews.org/index.jsp for more information.

[21] One of these stories was about election judges rather than trial or appellate judges, and one was about a judge's funeral, so only ten stories concerned judges' actions or decisions, or judicial elections. Of these, three concerned a judge who was sentenced to jail for fraud, two were about whether a candidate met the residency requirements to run for a judicial office (the candidate was not a sitting judge), one was about a federal judge's decision to strike down Chicago's ban on tobacco and alcohol billboards, one was about a state supreme court's decision that a judge had not violated a state ethics law but had simply exercised his free speech, one was about a judge's decision not to quit a trial against tobacco companies, one was about the dismissal of a complaint against a judge for using a racial slur, and one was a retraction by the station of an error in an earlier broadcast.

We use variation in $Congruence_d$ to identify effects of newspaper coverage of judges. Note that since *Congruence* is defined using market shares, it is not dependent on the total newspaper penetration in the judicial district. This is important since total newspaper readership in an area is related to characteristics such as education and income levels.

To measure *Congruence*, we use county-level newspaper sales data. Each year, the Audit Bureau of Circulation (ABC) collects data on each newspaper's circulation in each county for almost all U.S. newspapers. We have this data for 1982 and for the period 1991-2004. For the years 1983-1990 when we do not have circulation data, we interpolate *Congruence*.[22] We complemented this with county-circulation data for non-ABC newspapers for 1991 and 2004, and interpolated values between those years. The non-ABC data is mainly for small newspapers.[23] In our data, the average number of newspaper copies sold in a year is 56 million. The average number of copies sold per household is 0.58, falling from about 0.70 in 1982 to 0.50 in 2004.

4.5 Local Penal Attitudes and Controls

We use two measures of the voters' penal preferences. One is the share of voters who vote for the Democratic Party in the presidential election.[24] This measure reflects the general liberalness of voters and is negatively related to the harshness of penal preferences. The other is the share of voters who vote for harsher crime punishment on various ballot propositions. Specifically, we use all available statewide ballot propositions that deal mainly with the punishment of criminals, the rights of the accused, and victim's rights.[25] In virtually all cases, a majority of voters voted for an increase in harshness towards criminals or the accused, or in favor of victim's rights. On average, more than 65% of voters took the harsher position. This is consistent with the widespread view that most Americans believe the criminal justice system is too lenient. We collected county-level voting data on these ballot propositions from states' election web sites and/or election officials. We code all propositions so that higher vote-shares represent greater support for increased harshness towards criminals or the accused. For states with more than one proposition, we average the vote shares

[22] In Section D of the appendix, we check the robustness of our empirical results to dropping the period before 1991.

[23] The non-ABC data was provided by SRDS. On average there are about 10,900 observations each year in the ABC data, and about 500 observations in the non-ABC data. There are about 3,000 counties in the U.S., so the average number of observations per county in each year is slightly less than 4.

[24] For the years without the presidential election, we use linear interpolation.

[25] These propositions are listed in Tables A.III-A.IV in the appendix.

across the available propositions. We then de-mean the vote shares so that in each state the mean score is zero. We call the resulting variable *Harsh Vote Share*.

We also use data on a number of demographic characteristics at the court level. These have been aggregated from the county level, using data from the U.S. Census Bureau.

5 Analysis

In this section, we first analyze how our congruence measure is related to the observed media coverage, as well as how important people say that newspapers are for their information about courts. We then turn to the main analysis of the effects of media coverage on sentencing.

5.1 Specification

We now discuss two specifications: (i) simply regressing $Harshness$ on $Congruence$, and (ii) estimating the difference between the coefficients in two samples, where we expect a large effect in one sub-sample and a small effect in the other. The difference gives a lower bound on the estimate in the high impact sub-sample. This procedure is related to the procedure in Altonji, Elder and Taber (2005). Those authors desire an estimate of the effect of Catholic schooling, but their initial estimates are potentially biased because identifying as a Catholic is likely to correlated with the error term. They estimate this bias in a sub-sample with low probability of going to Catholic schools (public school 8th graders). In a similar fashion, we estimate the bias in the sample of appointed judges, where we expect that the effect of media is small. In addition, we test whether the bias produced by the observable part of the error term is similar in the two samples, and discuss this test below.

Our main specification is of the form

$$H_{it} = \alpha C_{it} + \beta' d_{st} + \gamma' x_{it} + \varepsilon_{it}, \tag{11}$$

where H is our sentencing harshness measure, C is $Congruence$, d_{st} are a set of state-by-year dummy variables and x_{it} contains our control variables. Define $\widetilde{C}$ as the residual from regression C_{it} on the state-year dummy variables and the controls in x_{it}. Then

$$\widehat{\alpha} \underset{p}{\longrightarrow} \alpha + \frac{cov(\widetilde{C}_{it}, \varepsilon_{it})}{var(\widetilde{C}_{it})}.$$

Our key identifying assumption is that *Congruence* is uncorrelated with the error, after adding controls, so that $cov(\widetilde{C}_{it}, \varepsilon_{it}) = 0$.

The main threat to identification is that *Congruence* might be related to sentencing harshness for reasons other than newspaper coverage. One way to mitigate this problem is to use temporal variation from redistricting, or within district-and-year variation, as in Snyder and Strömberg (2010). We cannot follow this strategy, however, because judicial district lines are rarely redrawn, and newspaper entries or exits are too few to generate sufficient variation to precisely identify *Congruence* effects. Thus, we are forced to rely primarily on variation across judicial districts, and we must worry more about the correlation between *Congruence* and other variables. For example, *Congruence* tends to be lower in densely populated areas. Since density might also be correlated with sentencing harshness, we control for population and area size (in logs). In addition, we sometimes use a trimmed sample in which the probability that *Congruence* is above the median is larger than 10% and smaller than 90% based on population and area. We can control for other variables in a similar fashion, but omitted-variable bias remains a concern.

We can, however, consistently estimate the *difference* in the effect of *Congruence* for, say, non-partisan elected judges and appointed judges under the weaker assumption that *Congruence* is correlated with the error term in the same way in these subsamples.[26] Assume that the correlation $cov(\widetilde{C}, \varepsilon)/var(\widetilde{C})$ is independent of judicial selection system. Under these conditions, the coefficient estimate on C from equation (11) in the subsample of appointed judges will converge to $\hat{\alpha}^A \underset{p}{\longrightarrow} \alpha^A + cov(\widetilde{C}, \varepsilon)/var(\widetilde{C})$ while the corresponding estimate in the subsample of non-partisan elected judges will converge to $\hat{\alpha}^{NP} \underset{p}{\longrightarrow} \alpha^{NP} + cov(\widetilde{C}, \varepsilon)/var(\widetilde{C})$. Consequently, we can consistently estimate

$$\hat{\alpha}^A - \hat{\alpha}^{NP} \underset{p}{\longrightarrow} \alpha^A - \alpha^{NP}$$

This difference can be estimated by interacting *Congruence* with a dummy variable for the judicial selection system being appointment, as follows:

$$H_{it} = \alpha_1 C_{it} + \alpha_2 d^A C_{it} + \beta' d_{st} + \gamma' x_{it} + \varepsilon_{it} \tag{12}$$

[26]The discussion below is similar to Altonji, Elder and Taber (2005).

In this case

$$\widehat{\alpha}_1 \underset{p}{\longrightarrow} \alpha^{NP} + \frac{cov(\widetilde{C}, \varepsilon)}{var(\widetilde{C})}$$

and

$$-\widehat{\alpha}_2 \underset{p}{\longrightarrow} \alpha^{NP} - \alpha^{A}.$$

Under the restriction that the effect of *Congruence* on sentencing is in the same direction for non-partisan elected and appointed judges, $-\widehat{\alpha}_2$ is a lower bound on the effect of *Congruence* on sentencing for non-partisan elected judges.

We cannot test the identifying assumption that $cov(\widetilde{C}, \varepsilon)/var(\widetilde{C})$ is independent of judicial selection system, since ε is not observable. However, it is possible to test the equivalent condition for the observable part of the error term, captured by our controls in the term $\gamma' x_{it}$. Define $\widetilde{\widetilde{C}}$ as the residual from regression C_{it} on the state-year dummy variables only. Then we can test whether $cov(\widetilde{\widetilde{C}}, \gamma' x_{it})/var(\widetilde{\widetilde{C}})$ depends on the electoral system, for example, by running the regression

$$\widehat{\gamma}' x_{it} = \alpha_1 C_{it} + \alpha_2 d^A C_{it} + \beta_2' d_{st} + u_{it}.$$

In this regression, α_1 measures $cov(\widetilde{\widetilde{C}}, \gamma' x_{it})/var(\widetilde{\widetilde{C}})$ in the sample of non-partisan elected judges and α_2 measures the difference in correlation in the non-partisan elected and appointed samples.

5.2 Newspaper Coverage, Congruence and Voter Information

We begin by investigating the relationship between *Congruence* and newspaper coverage of the courts. There is a strong positive relationship between *Congruence* and the number of articles written about judges in a district. This is shown in Figure 1, which displays the binned averages of these two variables. Each dot contains 0.5% of the observations, sorted by *Congruence* so that the leftmost dot contains the observations with the lowest 0.5% of the observations.

[Figure 1 here]

We next investigate this relationship more closely. Column 1 of Table III shows the results from a set of regressions of the number of articles per judge in a court district on *Congruence*. An increase in *Congruence* from zero to one is associated with an additional 24 newspaper articles per judge in the judicial district. This relationship is highly statistically significant. The next column adds a set of control variables: state-fixed effects, crime rates for 9 crime categories, population, per capita income, average education levels (share with 1-11 years, share with 12 years, and share with more than 12 years), share black, share urban, area in square miles, employment, turnout in presidential election, the Democratic vote share in the presidential election, and the share religious adherents. This does not significantly affect the estimated coefficient on *Congruence*. Regarding the controls, the number of newspapers stories is positively related to the share of crimes reported to police that are violent, log of population, log of income, and the share religious adherents. After adding these controls, the number of newspaper stories per judge is not related to whether the judges are elected or appointed, or to the number of judges on the court.

[Table III here.]

Our estimates suggest that an increase in *Congruence* from zero to one is associated with an additional 24 newspaper articles per judge in the judicial district, after adding controls. What does this number imply for the expected number of articles that an average person actually reads? A back-of-the-envelope calculation is illuminating. Household newspaper penetration rates and estimated readership are both about 60% in our period of study.[27] Also, readers typically read between 1/3 and 1/4 of all articles in a newspaper (Graber (1988) and Garcia and Stark (1991)). Thus, an increase in *Congruence* from zero to one would be associated with an average person reading about 4 more newspaper articles about their judge each year (24*.6*.25). A one standard deviation increase in *Congruence* implies effects about 1/3 as large, or about 1.3 more articles read.

We next investigate how voter information about courts is related to our measures of newspaper coverage. The National Center for State Courts conducted a survey in 2000 where they asked a random sample of U.S. respondents the following question: "How important to you are the following sources of information to your overall impression of how the courts in your community work?" They

[27] In our sample, the average number of newspapers sold per household is 0.58. The average total U.S. daily newspaper readership reported by the Newspaper Association of America is 60% of people aged above 18 for the period 1982 to 2004. Readership is measured by the share of survey respondents who reported reading a newspaper yesterday. See the Newspaper Association of America, "Daily Newspaper Readership Trend - Total Adults (1964-1997)," and "Daily Newspaper Readership Trend - Total Adults (1998-2007)."

were given ten alternatives, including newspapers and television news, but also their own experience in court, friends, relatives, their job, etc. We regressed a dummy variable for whether the respondent said that newspapers were very or somewhat important on the number of newspaper articles about the judges and on *Congruence*. The results are shown in Table IV. People who live in areas where *Congruence* is high are about 35% more likely to cite newspapers as an important source of information about the courts, and this is highly statistically significant and robust to the inclusion of controls.

[Table IV here.]

People who live in areas where we identified many newspaper articles covering judges are also more likely to cite newspapers as a source of information. However, this effect is only marginally statistically significant. One reason for the weaker result could be that the survey was conducted in 2000, while we collected data on the number of newspaper articles covering the judges in 2004 and 2005.

In sum, we have found that *Congruence* is strongly and positively correlated with the number of newspaper articles written about the judges on the court and with people stating that newspapers are an important source of information about the courts. We now turn to the effects on sentencing.

5.3 Effects of Media Coverage on Sentencing

Our model suggests that media coverage would make sentencing more aligned with voter preferences, and that such effect would be largest for non-partisan elected judges and for crimes that are more likely to attract media attention. We now investigate how media coverage influences sentencing. In Section 5.3.1, we focus on the three most serious offense types – homicides, sexual assaults and robberies – because these are most likely to attract media attention. We look at the average effects across all judicial selection systems. Section 5.3.2 analyzes the effect separately for appointed, non-partisan elected and partisan elected judges. Section 5.3.3 looks at the effect for less severe crimes.

In all cases, our main dependent variable is a measure of the harshness of sentencing, relative to other sentences in the same state and year and with the same penal code citation. Once a felon is convicted under a certain penal code citation, it is typically under the discretion of the judge to set the sentence. Our measure is supposed to capture the discretionary part of sentencing by judges. To construct this measure, we first generate a variable, penal code, that takes the same value for all crimes in each state in each year that have the same penal code citation for the 1st, 2nd, and 3rd most serious offenses. We then identify the minimum and maximum sentence given for that penal code. The variable $Harshness$ is defined as

$$Harshness = \frac{sentence - minimum}{maximum - minimum}.$$

This variable is bounded between zero and one, where one means that the judge imposed the highest sentence for this penal code citation in this state and year, and 0 means that the lowest sentence was imposed. The sentencing data does not give the identity of the sitting judge in each case, so our analysis of sentencing harshness focuses on the relationship between $Congruence$ and sentencing at the judicial district level.

5.3.1 All Selection Systems – Most Severe Crimes

We now restrict attention to homicides, sexual assaults and robberies. Table V shows that the average sentence length for these crimes is considerably higher than other offense categories. The average sentence length for murder is 486 months, compared to 122 months for sexual assault and 102 months for robbery. The average sentence length is less than half of this in each remaining crime category, ranging from an average 41 months for aggravated assault to an average 13 months for drug possession.

[Table V here.]

Table VI shows the summary statistics of the main variables. Our measure of sentencing harshness has mean of .26 and a standard deviation of .34. So, there is a substantial variation in $Harshness$, but more sentences are closer to the minimum than the maximum harshness.

[Table VI here.]

Table VII presents estimates from regressions of *Harshness* on our media variables. All specifications include state-by-year fixed effects. The first four columns use the Log Number of Articles about judges in 2004 and 2005 as the main independent variable.[28] The first column shows that the number of newspaper articles are positively and significantly associated with harsher sentencing. Of course, this correlation could exist for a variety of reasons. One is that more severe crimes attract media attention. However, the correlation does not seem to be driven by specific cases during the time period of our newspaper data. In fact, the correlation is virtually unchanged after when we remove 30,000 observations from the year 2004 (coefficient estimate = .0255, s.e. = .0044).

[Table VII here.]

Another possible reason for the correlation between sentencing and newspaper coverage and sentencing harshness is that newspapers are likely to locate in more populous and densely populated areas, and these areas might also have different crime rates and sentencing patterns. The specification in the second column addresses this by including log of population and log of area as controls. Including these two variables alone causes the estimated coefficient on Log Number of Articles to fall substantially.

Areas with ample newspaper coverage may of course differ in other respects. Crime rates may be higher, and the types of crimes may be different as well as the characteristics of the defendants. The areas could also differ in terms of important demographics, such as age, race, ethnicity, and income. The third specification includes an extensive set of controls: (i) fixed effects for the type of crime; (ii) defendant characteristics – gender, race, age (in years), and age squared; (iii) crime-related district characteristics – the log total number of convictions in the district, the share of those convictions that involved violent crimes, and the share of that involved drug related crimes, the log of the total number of crimes in the district reported to the police, and the share of those crimes that were violent; and (iv) other district characteristics – log income in the district, log employment, the share of people in the district who are religious adherents, female, younger than 20, older than 65, black,

[28] Since we have the number of articles only for 2004 and 2005 but use sentencing data from 1986 to 2006, the results from the first four columns yield only a rough, overall relationship between media coverage and sentencing, rather than a precise relationship. Moreover, we are mainly interested in the relationship between sentencing and *Congruence*, which is less subject to endogeneity. Thus, results from the first four columns are useful primarily for the purpose of comparison with results from using *Congruence* as the key independent variable.

white, Hispanic, and urban, the share with high school education and the share with more than high school education, turnout in the most recent presidential election, and total newspaper penetration. The inclusion of these controls reduces the estimated coefficient on Log Number of Articles further.

The last specification uses a trimmed sample. This excludes all observations which, based on population and area size, have a less than 10% or above 90% probability of having an above median number of log articles. The coefficient estimate on Log Number of Articles is consistently positive and significant across these specifications. However, it falls as we include more controls. Although most of this drop is caused by population and area, we still worry that there is some bias even in the final estimate.

The next four columns repeat the same specifications, but use *Congruence* as the key independent variable. We use *Congruence* as the independent variable rather than an instrumental variable for the number of articles, because we have *Congruence* measure for most of the years in the data period while we have the number of articles only for 2004 and 2005. After controlling for log population size and log area, the estimated coefficient on *Congruence* is positive, significant and stable. When we add the same extensive set of controls as above to the regression, the estimated coefficient unchanged. The final specification trims the sample by excluding all observations which, based on population and area size, have a less than 10% or above 90% probability of having above median *Congruence*. The estimated coefficient is of similar size and significance in this trimmed sample.

In terms of magnitudes, the estimated coefficients are sizeable. A one standard deviation increase in *Congruence* is estimated to increase the sentence length by about 1/2 of a year (6.2 months).[29] To get a sense of the aggregate implications of changes in the media market, consider a uniform increase in *Congruence* of one standard deviation affecting all the 83,610 estimated convictions in these crime categories in 2006. The estimates imply that this change would have increased the aggregate sentence length for these crimes by more than 40,000 years.[30] Given the results presented in Table III, we would expect the coefficient on *Congruence* to be about 2.2 times larger than the estimated coefficient on the log number articles. The actual estimates of the effect of *Congruence* are larger, but not significantly so.[31]

[29] To relate sentence length to *Harshness*, we regress the former on the latter yielding a coefficient of 408 (controlling for state-by-year fixed effects and offense class). This is multiplied by the estimated effect of *Congruence* on *Harshness* (0.051), and one standard deviation in *Congruence* (0.30), yielding 408*.051*.3 = 6.2.

[30] The effect is 83610*6.2/12 = 43,493 years. The number convictions is from the Bureau of Justice Statistics (2009).

[31] We do not have a strong prior on the direction of bias with Log Number of Articles. On one hand, areas with more severe crimes may get more articles about judges and longer sentences producing a positive bias. On the other hand,

We interpret the consistently positive estimates on sentence length as evidence that those informed by the media prefer harsher sentences, and that media helps them enforce this preference. That is, the media acts as a "fire alarm" (McCubbins and Schwartz (1984)). The potential threat of negative coverage keeps judges in check, and voters are largely uniformed about judicial elections so even one publicized case can decide their votes.[32]

A standard finding in surveys is that the ordinary voters want tougher sentencing. For example, the National Annenberg Election Survey interviewed 79,458 US residents living in 2,898 counties in connection with the 2000 Presidential election. The survey asked: "The number of criminals who are not punished enough – is this an extremely serious problem, a serious problem, not too serious or not a problem at all?" An overwhelming majority responded that this was an extremely serious (34%) or serious (47%) problem, while only 17% answered that this was not too serious or not a problem at all.

Several empirical papers in the literature argue that elected judges are generally harsher than appointed. For example, Huber and Gordon (2007) compare elected and appointed judges in Kansas and argue that the former impose longer sentences. We find similar but statistically weaker differences. In Table VII, appointed judges are associated with less harsh sentencing, although this is only statistically significant in the untrimmed sample. There is no discernible difference between partisan and non-partisan elected judges. We can identify the coefficients on partisan elected and appointed judges because some states have both appointed and elected judges. The reason that the estimates change when we look at the trimmed sample is because appointed judges are often in the more populous districts and these are removed by trimming.

We estimated a specification that includes an interaction term between *Congruence* and *Harsh Vote Share*, but this term is not statistically significant. Our model suggests that the effect of *Congruence* on *Harshness* may not necessarily increase in *Harsh Vote Share*. The reason is that if everyone in a locality agrees that harsh penalties are good, and if judges are from this locality, then there is no agency problem.

we only measure newspaper coverage in two years covering 13 percent of the observations in our sample. So there is most likely considerable measurement error and consequent attenuation bias.

[32] In other words, our results do *not* require that voters read newspaper articles about judges carefully or remember their names. Rather, the presence of media coverage about judges and judges' large disutility from getting negative press coverage are enough to induce them to avoid unpopular decisions.

Regarding the controls, male and black convicts receive significantly harsher sentences. There is a strong age profile where convicts of age 44 get maximally long sentences. *Harshness* is not correlated with the number or the types of crimes dealt with in the courts or reported to police, given that we compare sentencing only within the group of cases that have common penal code citations. *Harshness* is higher in districts that are rich and geographically small.

5.3.2 Effects By Judicial Selection System

We now investigate the hypotheses that the influence of the media depends on the selection system. We hypothesized that newspapers would have the largest effects on sentencing for non-partisan elected judges, followed by partisan elected and appointed judges (see Corollary 2). We also hypothesized that partisan elected judges would be most sensitive to local penal preferences.

To study these differences, we interact our media coverage variables with indicator variables for the type of selection systems. As in the previous subsection, we focus on the three most severe crime categories (homicides, sexual assaults, and robberies). The result from this regression is shown in Table VIII. Columns I and II use Log Number of Articles, and Columns III and IV use *Congruence* as the key independent variable. All specifications include controls, and the Columns II and IV use the trimmed samples.

[Table VIII here.]

The main effects of the log number of articles and *Congruence* show the effects for non-partisan elected judges (the omitted category in the interactions). These effects are all statistically significant. For *Congruence*, the effect is about twice as large as the average effect across selection systems that we measured in Table VII. In the non-trimmed sample, it is 0.098 compared to 0.051 in Table VII. This implies that a one standard deviation increase in *Congruence* increases sentence length by an estimated 12 (= 408*.098*.3) months for non-partisan elected judges.

The estimated effects of the media variables in the other selection systems are found by adding the main effect to the relevant interaction term. For example, using the specification in Column I, the estimated effect of Log Number of Articles for partisan elected judges is 0.021-0.011 = 0.010. The row labeled Congruence Partisan Elected shows the p-value from an F-test of the hypothesis that the effect of *Congruence* is zero in the partisan elected sub-sample, and similarly for the appointed sub-sample. The estimated effects of our media variables on $Harshness$ are not statistically significantly different from zero for appointed or partisan elected judges. Perhaps this zero effect is more surprising for the partisan elected judges. This indicates that the use of partisan labels in voting is sufficiently strong that the additional information in newspaper articles plays little role to help voters select and discipline judges.

The effect of media coverage on sentencing therefore seems to be entirely driven by the non-partisan elected judges. The zero result for appointed and partisan elected judges obviously cannot be generalized to higher salience elections with significantly more voluminous media coverage. However, they do indicate an interesting limit to media effects in low salience elections.

The estimates above correspond to $\widehat{\alpha}_1$ in equation (12). As discussed, we can consistently estimate the effect of these variables, conditional on the judicial selection system, given that they are uncorrelated with the error term. The coefficient on the interaction between media variables (Log Number of Articles and *Congruence*) and selection systems correspond to $\widehat{\alpha}_2$ in equation (12). We can consistently estimate the differential effect of our media variables on $Harshness$ in the different selection systems if the correlation between them and the error term is independent of the judicial selection system.

Consistent with our model, the effect of the number of articles and *Congruence* is higher for the non-partisan judges than in other judicial systems, although this difference is only significant for *Congruence*. For example, the estimated effect of *Congruence* is .117 lower for appointed than non-partisan elected judges. We further hypothesize that partisan elected judges would be more influenced by media coverage than appointed judges. The point estimates are consistent with this hypothesis, but the difference between them (-0.095 and -0.117) is not statistically significant.

Finally, our estimates provide a lower bound on the effect of *Congruence* on sentencing in non-partisan elected systems under the additional assumption that this effect is of the same sign but of lower magnitude for partisan elected or appointed judges than for non-partisan elected judges. This lower bound of the effect of *Congruence* on $Harshness$ is .095. Note that this is almost identical to the effect estimated under the assumption that *Congruence* is uncorrelated with the error term, .098.

Table IX investigates whether the observable part of the error is correlated with the number of newspaper articles covering the court and *Congruence*, and whether it is differentially correlated with these variables depending on the judicial selection system. The dependent variable is the predicted *Harshness*, based on all our included regressors, excluding the media variables and the fixed effects for state-and-year and offense class. Columns I and II show that log number of articles correlates as much with the predicted sentencing as it does with the actual sentencing. It also correlates differentially with predicted *Harshness* across selection systems. This is consistent with our result that the specifications with the actual number of articles in Table VII are sensitive to the inclusion of controls. However, this is not true for *Congruence*. The first row shows that *Congruence* is not significantly correlated with the predicted harshness in the non-partisan elected sample. It also does not correlate differentially with the predicted *Harshness* across selection systems. However, in the trimmed sample, the *Congruence* correlates differently with predicted *Harshness*, indicating that estimation is more sensitive to the inclusion of controls in this subsample.

[Table IX here.]

5.3.3 Effects by Severity

Some types of criminal cases attract more media attention than others. For example, in one content analysis of news reports, murder accounted for 25% of crime stories, although it constitutes less than 1 percent of all reported crimes (Graber (1980)). Similarly, in a search in Newslibrary.com, we identified more than 2 million newspaper articles mentioning "judge" and "sentenc*". Although the most serious violent crimes – murder, rape, robbery – constitute only 7% of all felony convictions in 2006,[33] 42% of the newspaper stories mention these types of crimes. Consequently, the newspaper coverage to convictions ratio for these crimes is $42/7 = 6$. This ratio is 2.89 for all violent crimes, 0.86 for property crimes (burglary, theft, fraud) and 0.76 for drug related crimes.[34]

Our model suggests that the effect of newspapers on sentencing is increasing in the amount of media coverage (see equation (8)). This implies that newspaper coverage should have the greatest

[33] Source: Bureau of Justice Statistics (2009).

[34] Violent cries are mentioned in 52% of the articles and constitute 18% of convictions. Property crimes (burglary, theft or fraud) are mentioned in 24 percent of the articles and constitute 28 percent of the convictions. Drugs are mentioned in 25% of the articles and drug crimes constitute 33 percent of the convictions.

effect on the behavior of non-partisan elected judges choosing sentences for the most serious violent crimes. If the ordinary voters prefer longer sentences, then *Congruence* and the total number of articles should correlate more with harsh sentencing for these cases.

To investigate this, we run separate regressions by type of crime. The results are shown in Table X. The specifications include controls and the full (non-trimmed) sample. They are similar to Columns III and VII in Table VII.

[Table X here.]

The estimated coefficients on *Congruence* are higher for more newsworthy crime categories, with higher ratios of newspaper coverage to convictions. The estimated effect is statistically significant for violent crimes and property crimes, although smaller than for the most serious violent crimes in the first column. The correlation between the number of newspaper articles covering the court and sentence harshness is driven by violent crimes. The size of this correlation drops sharply and becomes insignificant for property, drug or weapons related crimes.

There is also some evidence that black defendants are overrepresented in media coverage. For example, Dixon and Linz (2000) find that while 25 percent of all felony perpetrators according to crime reports were black, 44% of the perpetrators on television news were black. By comparison, 23% of the felony perpetrators were white, while only 18% of felony perpetrators on television news were white. Again, if effects on sentencing are increasing in media coverage, we would expect to see a stronger correlation between *Congruence* and sentencing in cases involving black defendants. However, we found no evidence of this; see Table XI.

[Table XI here.]

5.4 Effects on Penal Preferences

It is also possible that newspaper coverage of the courts affects voters' penal preferences. The media's focus on violent crimes, for example, may induce a belief that these crimes are more prevalent than they actually are, and that harsher sentences are appropriate. This of course requires that the public does not fully understand the media's news selection criteria.

We find no evidence that newspaper coverage of the courts affect penal attitudes. We regressed the share voting for harsher sentencing measures in ballot propositions – *Harsh Vote Share* – on our media variables. The results are shown in Table XII. We find that *Harsh Vote Share* is strongly and positively related to the number of crimes known to police, population size, the share urban, falling unemployment, and the Democratic vote share in the presidential election. However, neither of our measures of media coverage intensity are correlated with the *Harsh Vote Share*.

[Table XII here.]

6 Conclusion

This study provides theoretical foundations and empirical evidence of how the media influences the functioning of selection systems for public officials. It has several important implications. First, whether the election system strengthens the relationship between policy outcomes and voter preferences critically depends on the media environment. A common argument is that election systems are better than appointment systems because they are more responsive to voters' preferences. In practice, however, voters are often ignorant of candidates in elections for most of state and local public offices. Our results show that the degree to which an election system results in a stronger relationship between policy outcomes and voter preferences depends on the media environment. Second, the effects of the media differ across partisan and non-partisan electoral systems; in particular, the effects appear to be much stronger under non-partisan elections. The partisan election system is used for numerous public offices because of the information provided by political parties, as in Snyder and Ting (2002), despite various issues that the party system causes, such as entry barriers to politics. Our results show that information from the media can sometimes substitute for information generated via party labels and partisan competition, but this depends on the functioning of the media environment.

While this study provides useful insights for understanding how media environments affect functioning of accountability structures, there are a number of remaining issues that require further research. First, it is not obvious that an increase in responsiveness to ordinary voters' preferences is welfare enhancing. This might be especially true for offices with narrow policy jurisdictions, such as judicial offices. Elections for these types of offices are unlikely to incorporate variation in the intensity of preferences in an optimal manner. Thus, there is a risk that media coverage may help enforce a "tyranny of the majority" of ordinary voters against the interest of a minority with much stronger preferences. More research is needed to assess the conditions under which it is desirable for public policy to reflect the views of a majority of voters, and how this varies across policy areas. Second, this study abstracts from other features of political environments, such as the competitiveness of state and local politics, which can affect the degree of media influence on the behavior of public officials. On one hand, the competitiveness of electoral races may be a substitute for media influence, since campaigning by the candidates may increase the amount of voter information. On the other hand, it may also be complementary, because the candidates's campaigns may actively disseminate the information provided by media. Theoretical and empirical research on such issues will improve our understanding of the interaction between the media and the of electoral institutions.

Claire S.H. Lim, Cornell University

James M. Snyder, Jr., Harvard University

David Strömberg, Stockholm University

Appendix

A Proofs

Proof. Proposition 1. This follows since in the non-partisan case

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

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

Proof. Corollary 2. By direct differentiation of the expressions in the proposition,

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

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 $v_L < 1$. □

B States and Counties in the NJRP Data

In Tables A.I-A.II, we list the states and counties included in our sentencing data.

C Ballot Propositions Used for Measurement of Penal Preferences

In Tables A.III-A.IV, we list the ballot propositions used to measure penal preferences.

D Sensitivity Analysis

In Tables A.V-A.VIII, we document the robustness of the results presented in Tables VII and VIII. We conducted the following set of sensitivity analyses:

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

- 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 the nonpartisan system. Columns labeled "MD" show results from coding Maryland as the appointment system.

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

- post90: In these columns, we drop sentencing data before 1991. We conduct this sensitivity analysis because we interpolated congruence variable for the years 1983-1990. See page 18 for details.

References

Altonji, Joseph G., Todd Elder, and Christopher Taber, "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 2005, *113* (1), 151–184.

Berdejó, Carlos and Daniel L. Chen, "Priming Ideology? Electoral Cycles Without Electoral Incentives Among U.S. Judges," 2012. unpublished manuscript.

Besley, Timothy and Robin Burgess, "The Political Economy of Government Responsiveness: Theory and Evidence from India," *Quarterly Journal of Economics*, 2002, *117* (4).

_ **and Stephen Coate**, "Elected Versus Appointed Regulators: Theory and Evidence," *Journal of the European Economic Association*, Sep. 2003, *1* (5), 1176–1206.

Blume, John and Theodore Eisenberg, "Judicial Politics, Death Penalty Appeals, and Case Selection: An Empirical Study," *Southern California Law Review*, 1999, *72*, 464–503.

Bonneau, Chris W. and Melinda G. Hall, *In Defense of Judicial Elections*, Routeledge, 2009.

Bureau of Justice Statistics, "Felony Sentences in State Courts, 2006 - Statistical Tables," 2009. http://bjs.ojp.usdoj.gov/content/pub/pdf/fssc06st.pdf.

DellaVigna, Stefano and Ethan Kaplan, "The Fox News Effect: Media Bias and Voting," *Quarterly Journal of Economics*, 2007, *122*, 1187–1234.

Dyck, Alexander, David Moss, and Luigi Zingales, "Media versus Special Interests," 2008. NBER Working Paper No. 14360.

Garcia, Mario R. and Pegie Stark, *Eyes on the News*, St. Petersburg, FL: Poynter Institute, 1991.

Gentzkow, Matthew, "Television and Voter Turnout," *Quarterly Journal of Economics*, 2006, *121* (3).

Graber, Doris A., *Processing the News: How People Tame the Information Tide*, New York: Longman., 1988.

Hall, Melinda Gann, "State Supreme Courts in American Democracy: Probing the Myths of Judicial Reform," *American Political Science Review*, June 2001, *95* (2), 315–330.

Hanssen, F. Andrew, "The Effect of Judicial Institutions on Uncertainty and the Rate of Litigation: The Election versus Appointment of State Judges," *Journal of Legal Studies*, 1999, *28* (1), 205–232.

_ , "Independent Courts and Administrative Agencies: An Empirical Analysis of the States," *Journal of Law, Economics, and Organization*, 2000, *16*, 534–571.

_ , "Is There a Politically Optimal Level of Judicial Independence?," *American Economic Review*, 2004, *94*, 712–729.

_ , "Learning about Judicial Independence: Institutional Change in State Courts," *Journal of Legal Studies*, 2004, *33*, 431–474.

Huber, Gregory A. and Sanford C. Gordon, "Accountability and Coercion: Is Justice Blind When It Runs for Office?," *American Journal of Political Science*, Apr. 2004, *48* (2), 247–263.

_ **and** _ , "The Effect of Electoral Competitiveness on Incumbent Behavior," *Quarterly Journal of Political Science*, May 2007, *2*, 107–138.

Lim, Claire S.H., "Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Court Judges," 2012. forthcoming in the *American Economic Review*.

Maskin, Eric and Jean Tirole, "The Politician and the Judge: Accountability in Government," *American Economic Review*, September 2004, *94* (4), 1034–1054.

McCubbins, Mathew D. and Thomas Schwartz, "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms," *American Journal of Political Science*, 1984, *28* (1), 165–179.

Sheldon, Charles H. and Nicholas P Lovrich Jr., "Voter knowledge, Behavior and Attitudes in Primary and General Judicial Election," *Judicature*, 1999, *82* (5), 216–223.

Snyder, James M. and Michael M. Ting, "An Informational Rationale for Political Parties," *American Journal of Political Science*, January 2002, *46* (1), 90–110.

Strömberg, David, "Radio's Impact on Public Spending," *Quarterly Journal of Economics*, 2004, *119* (1).

_ , "News Floods, News Droughts, and U.S. Disaster Relief," *Quarterly Journal of Economics*, 2007, *122* (2).

_ **and Jr. James M. Snyder**, "Press Coverage and Political Accountability," *Journal of Political Economy*, 2010, *118* (2).

Tabarrok, Alexander and Eric Helland, "Court Politics: The Political Economy of Tort Awards," *Journal of Law and Economics*, Apr. 1999, *42* (1), 157–88.

Table I
Selection and Retention Rules for the State Trial Courts

No. of States	Initial Selection	Reelection	Set of States
9	Partisan Election	Partisan Election	AL, IN, KS, LA, MO NY, TN, TX, WV
22	Non-partisan Election	Non-partisan Election	AR, AZ, CA, FL, GA ID, IN, KY, MD, MI MN, MS, MT, NV NC, ND, OH, OK OR, SD, WA, WI
3	Partisan Election	Retention Election	IL, NM, PA
10	Appointment	Retention Election	AZ, AK, CO, IA, IN, KS, MO, NE, UT, WY
11	Appointment		CT, DE, HI, ME MA, NH, NJ, RI, SC, VA, VT

Note 1: The selection systems can be divided into five groups. There are four states (Arizona, Indiana, Kansas, and Missouri) that have a within-state variation of two different systems (partisan or non-partisan election and appointment-retention election) at the district level. These states are included in both categories. For more details, see the website on judicial selection systems by the American Judicature Society (http://www.judicialselection.us/). In New Mexico judges are first appointed by the governor, then they must run in a partisan election, and subsequent elections are retention elections. In Maryland judges are either initially appointed by the governor or run in a non-partisan election.

Note 2: We classify a state as having non-partisan elections if party labels do not appear on the general election ballot. In Arizona (in some counties), Maryland, and Ohio, nominations are partisan but the general election ballot is non-partisan.

Note 3: Arkansas changed its selection system from partisan to nonpartisan, effective in 2002. North Carolina switched from partisan to nonpartisan, effective in 1998. These rule changes are reflected in our analysis of sentencing data.

Note 4: Illinois, New Mexico, and Pennsylvania, where partisan-elected judges face retention elections, are classified as the partisan system in our analysis. All the key results in subsequent tables are robust to the exclusion of these three states.

TABLE II
Newspaper Summary Statistics

Variable	Obs	Mean	Std. Dev	Min	Max
Total Number of Articles per Court[a]	1413	122.81	297.44	0	4566
Number of Judges per Court	1413	6.38	17.15	0	389
Number of Articles per Judge	1413	16.82	28.40	0	421
Articles per Judge and Year[b] (circulation weighted average)	1413	9.30	21.35	0	421
Congruence	1413	0.22	0.30	0	1

Note: Outside of Texas, the minimum number of judges is 1. In Texas, we use county as the unit of observation and allocate judges proportional to populations.

[a] The unit of observation in this table is court, not court-year. The total number of articles per court is the sum of all articles covering any judge on a court for the two-year period from 2004 to 2005.

[b] If we divide the total number of articles per court, 122.8, by the average number of judges per court which is 6.38, we get 19.2 articles per judge for two-year period or 9.6 articles per judge and year. If we instead use the number of articles per judge and year *weighted by circulation*, we get 9.3. It is more correlated with exposure of an individual since people typically only read one newspaper.

Table III

Relationship between the Amount of Coverage and Congruence

Dependent Variable: Articles per Judge (market share weighted)			
Variables	I Articles	II Articles	III Log Articles
Congruence	22.872***	23.760***	2.208***
	(1.804)	(2.266)	(0.199)
Appointed	-4.366**	-5.691	-0.177
	(1.758)	(4.963)	(0.451)
Number of Judges		-0.023	-0.014**
		(0.083)	(0.007)
Log number crimes known to police		2.960	0.029
		(2.420)	(0.227)
Share violent crimes known to police		29.306**	1.089
		(11.662)	(1.262)
Log population		5.988*	0.674**
		(3.119)	(0.291)
Log income		9.855**	1.567***
		(4.745)	(0.437)
Education (share with 12 years)		-23.197	-0.304
		(17.443)	(1.602)
Education (share with 12+ years)		0.004	2.173**
		(11.141)	(1.018)
Share Black		7.511	0.793
		(7.499)	(0.702)
Share Urban		-1.722	0.506
		(3.604)	(0.336)
Log land area in square miles		-0.920	0.062
		(0.987)	(0.092)
Log total employment (millions)		-4.644	-0.846***
		(3.247)	(0.299)
Voter Turnout		-14.873	1.063
		(11.544)	(0.983)
Democratic Vote Share		-9.206	-0.227
		(6.640)	(0.603)
Share Religious Adherents		10.056**	0.559
		(4.564)	(0.428)
Fixed Effects	No	State	State
Observations	1,413	1,413	1,169
R^2	0.102	0.175	0.278

Note 1: Results from OLS regressions. Standard errors in parentheses: *** significant at 1%, ** significant at 5% and * significant at 10%.

Note 2: The unit of observation is court by year. The dependent variable is the newspaper circulation (market share) weighted average number of newspaper stories per judge in the court. The number of observations is reduced in Column III because of observations with zero articles and the logarithm.

Table IV

Relationship between Voter Information, Congruence, and the Amount of Coverage

Dependent variable: Newspapers Important for Information about Courts					
Variables	I	II	III	IV	V
Congruence	0.345***	0.394***			0.436***
	(0.095)	(0.103)			(0.130)
Log Number of Articles			0.021*	0.024	0.026*
			(0.012)	(0.015)	(0.015)
Controls	No	Yes	No	Yes	Yes
Observations	533	531	475	473	473
R^2	0.123	0.155	0.115	0.146	0.172

Note 1: Results from OLS regressions. All specifications include state-fixed effects. Standard errors, clustered by court, in parenthesis: * significant at 10%; ** significant at 5%; *** significant at 1%.

Note 2: The number of observations is reduced in Columns III-V because of observations with zero articles and the logarithm.

Table V
NJRP Offense Categories and Sentence Lengths

Category	Number of sentences	Mean sentence (months)
murder	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 VI
Summary Statistics

Variable	Obs	Mean	Std. Dev	Min	Max
Harshness	232470	0.26	0.34	0.00	1.00
Log Number of Articles	212837	0.50	0.71	-3.46	2.51
Log Number of Articles	215750	8.90	16.62	0.00	322.42
Congruence	232470	0.57	0.21	0.00	0.97
Harsh Vote Share	166604	-0.01	0.05	-0.27	0.13
Democratic Vote Share	230520	0.57	0.14	0.15	0.90

Note 1: These statistics are based on severe crimes (homicides, sexual assaults, and robberies).
Note 2: *Harsh Vote Share* has a significantly smaller number of observations because ballot propositions are not available in all states.

TABLE VII

Regression of *Harshness* on Media Variables, Penal Preferences, and Selection Systems

Variables	I	II	III	IV	V	VI	VII	VIII
Log Number of Articles	0.026***	0.020***	0.017***	0.015***				
	(0.004)	(0.003)	(0.004)	(0.004)				
Congruence					0.027	0.046**	0.051**	0.061**
					(0.019)	(0.022)	(0.023)	(0.025)
Harsh Vote Share			0.114	0.103			0.140	0.167*
			(0.090)	(0.103)			(0.091)	(0.098)
Democratic Vote Share			-0.012	-0.016			-0.019	0.015
			(0.044)	(0.051)			(0.042)	(0.041)
Appointed			-0.069**	-0.070*			-0.084***	0.017
			(0.029)	(0.041)			(0.022)	(0.030)
Partisan Elected			0.006	0.010			-0.025	0.070*
			(0.025)	(0.024)			(0.026)	(0.039)
Observations	212,837	211,595	147,497	142,363	232,470	231,228	163,551	100,983
R^2	0.133	0.134	0.131	0.131	0.128	0.129	0.128	0.115
Controls	No	Pop., Area	Yes	Yes	No	Pop., Area	Yes	Yes
Trimmed Sample	No	No	No	Yes	No	No	No	Yes

Note 1: OLS regression results. Standard errors, clustered by court, in parenthesis. All specifications include state-by-year fixed effects.
Note 2: The number of observations is smaller in Columns III, IV, VII, and VIII, because *Harsh Vote Share*, based on ballot propositions, is not available in all states.
Note 3: Tables A.V and A.VI in Appendix D show results from various sensitivity analyses of Columns III, IV, VII, and VIII.

Table VIII

Regression of *Harshness* on Media Variables, Selection Systems, Penal Preferences, and Their Interactions

Variable	I	II	III	IV
Log Number of Articles	0.021***	0.017***		
	(0.007)	(0.006)		
Log Number of Articles*Partisan Elected	-0.011	-0.011		
	(0.010)	(0.010)		
Log Number of Articles*Appointed	-0.008	-0.003		
	(0.010)	(0.010)		
Congruence			0.098***	0.100***
			(0.036)	(0.038)
Congruence*Partisan Elected			-0.095**	-0.062
			(0.044)	(0.051)
Congruence*Appointed			-0.117**	-0.129**
			(0.045)	(0.062)
Harsh Vote Share	0.078	0.023	0.083	0.099
	(0.146)	(0.190)	(0.135)	(0.145)
Harsh Vote Share*Non-Partisan Elected	-0.008	-0.044	0.197	0.073
	(0.203)	(0.252)	(0.202)	(0.212)
Harsh Vote Share*Appointed	-0.654**	-0.482	-0.830***	-0.646*
	(0.270)	(0.300)	(0.278)	(0.375)
Democratic Vote Share	-0.045	-0.077	-0.051	-0.010
	(0.069)	(0.075)	(0.067)	(0.079)
Democratic Vote Share*Non-Partisan Elected	-0.050	-0.102	-0.021	0.048
	(0.116)	(0.140)	(0.101)	(0.104)
Democratic Vote Share*Appointed	-0.001	0.078	0.024	-0.042
	(0.104)	(0.103)	(0.102)	(0.134)
Partisan Elected	1.575	2.709**	1.622*	2.040*
	(1.093)	(1.317)	(0.936)	(1.141)
Appointed	1.286	3.592***	2.030*	0.358
	(1.160)	(1.320)	(1.065)	(1.459)
Observations	147,497	142,363	163,551	100,983
R^2	0.135	0.137	0.132	0.118
Controls	Yes	Yes	Yes	Yes
Trimmed Sample	No	Yes	No	Yes
Congruence Partisan Elected	.159	.427	.913	.262
Congruence Appointed	.109	.092	.489	.547
Harsh Non-Partisan Elected	.616	.902	.064	.264
Harsh Appointed	.011	.048	.002	.106

Note 1: OLS regression results. Standard errors, clustered by court, in parenthesis. All specifications include state-by-year fixed effects.

Note 2: Tables A.VII and A.VIII in Appendix D show results from various sensitivity analyses of these results.

Table IX

Regression of Predicted *Harshness* on Media Variables, Selection Systems, and Their Interactions

Variable	Media Variable Used			
	Log Number of Articles		*Congruence*	
	I	II	III	IV
Media Variable	0.016***	0.018***	-0.049	-0.002
	(0.003)	(0.003)	(0.030)	(0.023)
Media Variable*Appointed	-0.025***	-0.027***	0.004	-0.042
	(0.003)	(0.004)	(0.031)	(0.028)
Media Variable*Partisan Elected	-0.009**	-0.010**	0.026	-0.059**
	(0.004)	(0.005)	(0.032)	(0.028)
Observations	147,497	142,363	163,551	100,983
R^2	0.433	0.508	0.406	0.422
Trimmed Sample	No	Yes	No	Yes

Note: OLS regression results. Standard errors, clustered by court, in parenthesis. All specifications include state-by-year fixed effects.

Table X

Heterogenous Effects by Offense Category for Non-partisan Elected Judges

	Homicide, Sexual Assault, Robbery	Violent	Property	Drug	Weapons
	I	II	III	IV	V
Log Number of Articles	0.021***	0.013**	0.006	0.010*	-0.001
	(0.007)	(0.005)	(0.004)	(0.005)	(0.005)
Observations	83,136	205,021	311,595	393,447	192,274
R^2	0.132	0.096	0.055	0.082	0.042
	VI	VII	VIII	IX	X
Congruence	0.099***	0.077***	0.069***	0.052*	0.039*
	(0.036)	(0.029)	(0.026)	(0.029)	(0.023)
Observations	95,547	232,284	354,864	452,247	215,863
R^2	0.125	0.093	0.057	0.086	0.050

Note 1: OLS regression results. Standard errors, clustered by court and year, in parenthesis. All specifications include state-by-year fixed effects and control variables.
Note 2: Violent crimes include murder, sexual assault, robbery, aggravated assault and other violent crimes. Property crimes include burglary, larceny and fraud. Drug crimes include drug possession and drug trafficking. Weapons include weapon offenses and other offenses.
Note 3: The number of observations is smaller in Columns I-V because of observations with zero articles and the logarithm.

Table XI

Heterogenous Effects by Offense Category and Race (Non-partisan Elected Sample)

Crime Type	homicide, sexual assault, robbery	violent	property	drug	weapon & other
Variable	I	II	III	IV	V
Congruence	0.103***	0.079***	0.071**	0.045	0.040*
	(0.037)	(0.030)	(0.028)	(0.029)	(0.024)
Congruence*	-0.016	-0.009	-0.008	0.024	-0.006
Black Defendant	(0.023)	(0.019)	(0.022)	(0.029)	(0.020)
Observations	95,547	232,284	354,864	452,247	215,863
R^2	0.126	0.093	0.057	0.086	0.050

Note: OLS regression results. Standard errors, clustered by court, in parenthesis. All specifications include state-by-year fixed effects.

Table XII

Regression of Penal Preferences (*Harsh Vote Share*) on Media Variables

Variables	I	II
Congruence	-0.006	
	(0.007)	
Log Number of Articles		0.000
		(0.001)
Log Newspaper Circulation	0.001	-0.003
	(0.004)	(0.003)
Log Number Crimes Known to Police	0.032***	0.031***
	(0.006)	(0.007)
Log Population	0.023***	0.028***
	(0.008)	(0.008)
Log Income	0.017	0.026**
	(0.012)	(0.013)
Education (share with 12 years)	-0.013	-0.007
	(0.036)	(0.041)
Education (share with 12+ years)	-0.018	-0.062**
	(0.024)	(0.027)
Share Urban	0.017**	0.027***
	(0.007)	(0.008)
Log Total Employment (millions)	-0.025***	-0.027***
	(0.007)	(0.007)
Voter Turnout	-0.020*	-0.019
	(0.012)	(0.012)
Democratic Vote Share	-0.159***	-0.183***
	(0.014)	(0.015)
Observations	1,177	955
R^2	0.879	0.877
Fixed Effects	State	State
Controls	Yes	Yes

Note 1: OLS regression results. Standard errors, clustered by court, in parenthesis.
Note 2: Specifications include fixed effects for state and appointed and non-partisan elected judges. It also includes land area, the share violent crimes known to police, the share black, and the share religious adherents, none of which are significantly correlated with *Harsh Vote Share*.
Note 3: The number of observations is smaller in Column II because of observations with zero articles and the logarithm.

TABLE A.I
States and Counties Included in the NJRP Data

State	Counties
AK	Anchorage, Fairbanks North Star, Matanuska Susitna, Nome, Prince Wales Ketchikan
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
CT	Fairfield, Hartford, Litchfield, New Haven, New London, Windham
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
MA	Barnstable, Berkshire, Essex, Hampden, Middlesex, Norfolk, Worcester
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.II
States and Counties Included in the NJRP Data (con'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
ND	Mountrail, Renville, Stutsman
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.III
Ballot Propositions Used to Measure Penal Preferences

State	Year	Prop No.	Percent Yes	Description
AL	1996	Amendment 3	70	Removing the Prohibition on Guilty Pleas within 15 Days of Arrest in Non-Capital Felony Cases
AZ	1998	Proposition 301	48	Relating To 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	Eliminate Limitation of Fines For Offenses That May Be Summarily Tried Without Indictment
ID	1994	H.J.R 16	80	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	Provides for Rights of the Victim of a Crime
LA	1998	Amendment 6	68	Make It Easier For Judges To Deny Bail
LA	1998	Amendment 14	62	Require a Unanimous Verdict in Criminal Trials That Use Six-Member Jury
LA	1999	Amendment 1	59	Provide That Governor May Not Commute Sentences or Pardon Persons Convicted Without A Favorable Recommendation By Board Of Pardons
LA	1999	Amendment 8	53	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.IV
Ballot Propositions Used to Measure Penal Preferences (con'd)

State	Year	Prop No.	Percent Yes	Description
NE	2006	Amendment 4	56	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	Allow the Legislature to set Minimum Prison Terms for All Convicted Felons
OR	1996	Measure 26	66	Changes Principles That Govern Laws for Punishment of Crime
OR	1996	Measure 40	58	Gives Crime Victims Rights, Expands Admissible Evidence, Limits Pretrial Release
OR	1999	Measure 68	58	Allows Protecting Business, Certain Government Programs from Prison Work Programs
OR	1999	Measure 69	58	Grants Victims Constitutional Rights In Criminal Prosecutions, Juvenile Court Delinquency Proceedings
OR	1999	Measure 71	58	Limits Pretrial Release of Accused Person To Protect Victims
OR	1999	Measure 72	45	Allows Murder Conviction by 11 to 1 Jury Verdict
OR	1999	Measure 73	46	Limits Immunity from Criminal Prosecution of Person Ordered To Testify about his or her Conduct
OR	1999	Measure 74	53	Requires Terms of Imprisonment Announced in Court Be Fully Served, With Exceptions
OR	1999	Measure 75	57	Person Convicted of Certain Crimes Cannot Serve on Grand Juries, Criminal Trial Juries
OR	2000	Measure 3	67	Requires Conviction Before Forfeiture; Restricts Proceeds Usage; Requires Reporting, Penalty
OR	2000	Measure 94	26	Repeals Mandatory Minimum Sentences for Certain Felonies, Requires Re-sentencing
OR	2008	Measure 57	61	Increase Sentences for Drug Trafficking, Theft against Elderly and Specified Repeat Property and Identity Theft Crimes
OR	2008	Measure 61	48	Creates Mandatory Minimum Prison Sentences for Certain Theft, Identity Theft, Forgery, Drug and Burglary Crimes
PA	1998	Joint Resolution 1	72	Adding Categories of Criminal Cases in Which Bail Is Disallowed
PA	1998	Joint Resolution 2	69	Granting Commonwealth Right to Trial By Jury in Criminal Cases
PA	2003	Amendment 1	68	Amending Right of Persons Accused of a Crime To Meet Witness against Them Face To Face
PA	2003	Amendment 2	80	Auth Leg 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	Allows Denial of Bail To Persons Charged With Violent Crimes
SC	1998	Amendment 1	48	Allow Leg To Specify Which Crime Victims Are Protected By Victims Bill Of Rights
SD	2002	Amendment A	21	Relating To A Criminal Defendant's Rights
TN	1998	Amendment 2	89	Entitles Victims of Crime To Certain Basic Rights To Preserve and 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	Reinstate Death Penalty

TABLE A.V

Sensitivity Analysis 1

	Dependent Variable : Harsh							
	Variation in Specification and Data							
Variable	OH	MD	OHMD	post90	OH	MD	OHMD	post90
	I	II	III	IV	V	VI	VII	VIII
Log Number of Articles	0.017***	0.017***	0.017***	0.017***	0.015***	0.015***	0.015***	0.015***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Harsh Vote Share	0.114	0.114	0.114	0.130	0.103	0.103	0.103	0.114
	(0.090)	(0.090)	(0.090)	(0.099)	(0.103)	(0.103)	(0.103)	(0.114)
Democratic Vote Share	-0.012	-0.012	-0.012	-0.007	-0.016	-0.016	-0.016	-0.002
	(0.044)	(0.044)	(0.044)	(0.048)	(0.051)	(0.051)	(0.051)	(0.056)
Appointed	-0.069**	-0.069**	-0.069**	-0.072**	-0.070*	-0.070*	-0.070*	-0.069
	(0.029)	(0.029)	(0.029)	(0.032)	(0.041)	(0.041)	(0.041)	(0.045)
Partisan Elected	0.006	0.006	0.006	0.006	0.010	0.010	0.010	0.021
	(0.025)	(0.025)	(0.025)	(0.029)	(0.024)	(0.024)	(0.024)	(0.031)
Observations	147,497	147,497	147,497	127,600	142,363	142,363	142,363	123,182
R^2	0.131	0.131	0.131	0.103	0.131	0.131	0.131	0.104
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimmed Sample	No	No	No	No	Yes	Yes	Yes	Yes

Table A.VI

Sensitivity Analysis 2

	Dependent Variable : Harsh							
	Variation in Specification and Data							
Variable	OH	MD	OHMD	post90	OH	MD	OHMD	post90
	I	II	III	IV	V	VI	VII	VIII
Congruence	0.051**	0.051**	0.051**	0.047*	0.061**	0.061**	0.061**	0.066**
	(0.023)	(0.023)	(0.023)	(0.025)	(0.025)	(0.025)	(0.025)	(0.026)
Harsh Vote Share	0.140	0.140	0.140	0.120	0.167*	0.167*	0.167*	0.133
	(0.091)	(0.091)	(0.091)	(0.097)	(0.098)	(0.098)	(0.098)	(0.098)
Democratic Vote Share	-0.019	-0.019	-0.019	-0.035	0.015	0.015	0.015	-0.013
	(0.042)	(0.042)	(0.042)	(0.042)	(0.041)	(0.041)	(0.041)	(0.043)
Appointed	-0.084***	-0.084***	-0.084***	-0.098***	0.017	0.017	0.017	0.034
	(0.022)	(0.022)	(0.022)	(0.027)	(0.030)	(0.030)	(0.030)	(0.033)
Partisan Elected	-0.025	-0.025	-0.025	-0.022	0.070*	0.070*	0.070*	0.098**
	(0.026)	(0.026)	(0.026)	(0.029)	(0.039)	(0.039)	(0.039)	(0.038)
Observations	163,551	163,551	163,551	140,437	100,983	100,983	100,983	86,957
R^2	0.128	0.128	0.128	0.101	0.115	0.115	0.115	0.084
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimmed Sample	No	No	No	No	Yes	Yes	Yes	Yes

Table A.VII

Sensitivity Analysis 3

Dependent Variable: Harsh								
	Variation in Specification and Data							
Variable	OH	MD	OHMD	post90	OH	MD	OHMD	post90
	I	II	III	IV	V	VI	VII	VIII
Log Number of Articles	0.023***	0.021***	0.023***	0.021***	0.019***	0.017***	0.019***	0.015**
	(0.006)	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)
Log Number of Articles*Appointed	-0.010	-0.008	-0.010	-0.002	-0.005	-0.003	-0.005	0.006
	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)	(0.010)	(0.010)	(0.011)
Log Number of Articles*Partisan Elected	-0.013	-0.011	-0.013	-0.014	-0.011	-0.011	-0.011	-0.012
	(0.009)	(0.010)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
Democratic Vote Share	-0.016	-0.045	-0.016	-0.049	-0.048	-0.077	-0.048	-0.075
	(0.067)	(0.069)	(0.067)	(0.085)	(0.072)	(0.075)	(0.072)	(0.088)
Harsh Vote Share	0.162	0.078	0.162	0.041	0.130	0.023	0.130	-0.015
	(0.149)	(0.146)	(0.149)	(0.153)	(0.198)	(0.190)	(0.198)	(0.201)
Democratic Vote Share*Non-Partisan Elected	-0.091	-0.050	-0.091	-0.069	-0.168	-0.102	-0.168	-0.129
	(0.118)	(0.116)	(0.118)	(0.129)	(0.143)	(0.140)	(0.143)	(0.143)
Harsh Vote Share*Non-Partisan Elected	-0.083	-0.008	-0.083	-0.014	-0.141	-0.044	-0.141	-0.118
	(0.210)	(0.203)	(0.210)	(0.222)	(0.262)	(0.252)	(0.262)	(0.288)
Appointed	1.176	1.286	1.176	2.083*	3.569***	3.592***	3.569***	4.635***
	(1.176)	(1.160)	(1.176)	(1.203)	(1.341)	(1.320)	(1.341)	(1.327)
Partisan Elected	1.386	1.575	1.386	1.687	2.831**	2.709**	2.831**	2.812*
	(1.109)	(1.093)	(1.109)	(1.199)	(1.333)	(1.317)	(1.333)	(1.508)
Democratic Vote Share*Appointed	-0.029	-0.001	-0.029	-0.014	0.048	0.078	0.048	0.054
	(0.102)	(0.104)	(0.102)	(0.113)	(0.101)	(0.103)	(0.101)	(0.115)
Harsh Vote Share*Appointed	-0.739***	-0.654**	-0.739***	-0.639**	-0.591*	-0.482	-0.591*	-0.422
	(0.271)	(0.270)	(0.271)	(0.267)	(0.305)	(0.300)	(0.305)	(0.313)
Observations	147,497	147,497	147,497	127,600	142,363	142,363	142,363	123,182
R^2	0.135	0.135	0.135	0.108	0.137	0.137	0.137	0.110
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimmed Sample	No	No	No	No	Yes	Yes	Yes	Yes
Congruence Partisan Elected	.131	.159	.131	.367	.305	.427	.305	.748
Congruence Appointed	.113	.109	.113	.022	.091	.092	.091	.018
Harsh Non-Partisan Elected	.591	.616	.591	.864	.949	.902	.949	.519
Harsh Appointed	.011	.011	.011	.006	.047	.048	.047	.068

Table A.VIII
Sensitivy Analysis 4

Dependent Variable: Harsh								
	Variation in Specification and Data							
Variable	OH	MD	OHMD	post90	OH	MD	OHMD	post90
	I	II	III	IV	V	VI	VII	VIII
Congruence	0.100***	0.098***	0.100***	0.081**	0.102***	0.100***	0.102***	0.085**
	(0.036)	(0.036)	(0.036)	(0.040)	(0.038)	(0.038)	(0.038)	(0.037)
Congruence*Appointed	-0.120***	-0.117**	-0.120***	-0.085*	-0.131**	-0.129**	-0.131**	-0.112*
	(0.046)	(0.045)	(0.046)	(0.049)	(0.062)	(0.062)	(0.062)	(0.061)
Congruence*Partisan Elected	-0.108**	-0.095**	-0.108**	-0.091*	-0.067	-0.062	-0.067	-0.040
	(0.046)	(0.044)	(0.046)	(0.049)	(0.051)	(0.051)	(0.051)	(0.049)
Democratic Vote Share	-0.024	-0.051	-0.024	-0.049	0.020	-0.010	0.020	0.001
	(0.066)	(0.067)	(0.066)	(0.084)	(0.079)	(0.079)	(0.079)	(0.099)
Harsh Vote Share	0.182	0.083	0.182	0.075	0.208	0.099	0.208	0.010
	(0.137)	(0.135)	(0.137)	(0.141)	(0.153)	(0.145)	(0.153)	(0.152)
Democratic Vote Share*Non-Partisan Elected	-0.060	-0.021	-0.060	-0.057	0.020	0.048	0.020	0.005
	(0.103)	(0.101)	(0.103)	(0.115)	(0.106)	(0.104)	(0.106)	(0.125)
Harsh Vote Share*Non-Partisan Elected	0.089	0.197	0.089	0.141	-0.023	0.073	-0.023	0.140
	(0.207)	(0.202)	(0.207)	(0.211)	(0.220)	(0.212)	(0.220)	(0.219)
Appointed	1.947*	2.030*	1.947*	2.430**	0.188	0.358	0.188	-0.212
	(1.077)	(1.065)	(1.077)	(1.124)	(1.474)	(1.459)	(1.474)	(1.932)
Partisan Elected	1.551	1.622*	1.551	1.224	1.721	2.040*	1.721	1.614
	(0.947)	(0.936)	(0.947)	(1.031)	(1.176)	(1.141)	(1.176)	(1.216)
Democratic Vote Share*Appointed	-0.002	0.024	-0.002	0.044	-0.072	-0.042	-0.072	-0.104
	(0.102)	(0.102)	(0.102)	(0.110)	(0.134)	(0.134)	(0.134)	(0.127)
Harsh Vote Share*Appointed	-0.930***	-0.830***	-0.930***	-0.849***	-0.759**	-0.646*	-0.759**	-0.608*
	(0.280)	(0.278)	(0.280)	(0.283)	(0.378)	(0.375)	(0.378)	(0.364)
Observations	163,551	163,551	163,551	140,437	100,983	100,983	100,983	86,957
R^2	0.132	0.132	0.132	0.106	0.118	0.118	0.118	0.087
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Trimmed Sample	No	No	No	No	Yes	Yes	Yes	Yes
Congruence Partisan Elected	.786	.913	.786	.706	.295	.262	.295	.142
Congruence Appointed	.490	.489	.490	.875	.554	.547	.554	0.574
Harsh Non-Partisan Elected	.078	.064	.078	.164	.244	.264	.244	.338
Harsh Appointed	.002	.002	.002	.001	.104	.106	.104	.073

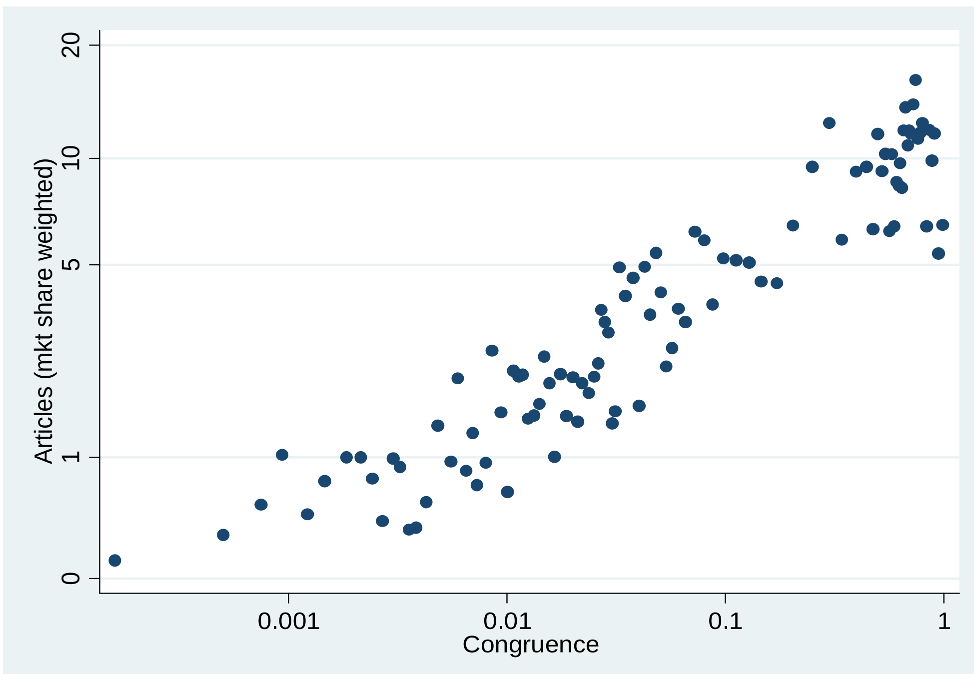

FIGURE I
Newspaper Articles and Congruence