

Accountability and Coercion: Is Justice Blind when It Runs for Office?

Gregory A. Huber Yale University
Sanford C. Gordon New York University

Through their power to sentence, trial judges exercise enormous authority in the criminal justice system. In 39 American states, these judges stand periodically for reelection. Do elections degrade their impartiality? We develop a dynamic theory of sentencing and electoral control. Judges discount the future value of retaining office relative to implementing preferred sentences. Voters are largely uninformed about judicial behavior, so even the outcome of a single publicized case can be decisive in their evaluations. Further, voters are more likely to perceive instances of underpunishment than overpunishment. Our theory predicts that elected judges will consequently become more punitive as standing for reelection approaches. Using sentencing data from 22,095 Pennsylvania criminal cases in the 1990s, we find strong evidence for this effect. Additional tests confirm the validity of our theory over alternatives. For the cases we examine, we attribute at least 1,818 to 2,705 years of incarceration to the electoral dynamic.

In the United States, only courts can authorize the incarceration of individual citizens. Because appellate courts review only a tiny fraction of criminal cases, trial court judges occupy an enormously significant role in administering criminal justice. There are nearly 5,000 state trial court judges. In 1998, they sentenced almost one million convicted felons to a total of more than two million years in state jails and prisons (Pastore and Maguire 2002, Tables 1.81 and 5.40–5.44). Prosecutors charge and juries convict, but trial judges' sentencing authority ultimately governs how this coercive element of state power is brought to bear on individual defendants.¹

In 39 states, trial judges stand for reelection.² Do elections, which ostensibly assure accountability to citizens, degrade judicial impartiality? The near consensus among legal scholars is that this tradition—particularly in the form of partisan, competitive contests—is politically unassailable but insidious in its potential for compromising judicial independence (e.g., ABA 1997, 2000; Croley 1995; Grodin 1988). At the same time, empirical studies suggest that voters are almost entirely uninformed about judge behavior (Dubois 1984; Hall 1999; Mathias 1990; Sheldon and Lovrich 1983) and that trial judges retain office at high rates (Aspin 1999; Baum 1983). But if this is

Gregory A. Huber is Assistant Professor of Political Science and Fellow at the Institution for Social and Policy Studies, Yale University, PO Box 208209, New Haven, CT 06520-8209 (gregory.huber@yale.edu). Sanford C. Gordon is Assistant Professor of Politics, New York University, 726 Broadway, 7th floor, New York, NY 10003-9580 (sanford.gordon@nyu.edu).

This is one of several joint papers by the authors on criminal justice institutions and politics; the ordering of names reflects a principle of rotation. We wish to thank the Institution for Social and Policy Studies at Yale and the Criminal Justice Research Center at The Ohio State University (where Gordon served as assistant professor of political science in the beginning stages of this project) for their generous financial support. We also gratefully acknowledge Brandon Bartels, Kevin Eirich, and Shaun Holness for superlative research assistance and Paul Brace, Larry Baum, Laura Langer, Todd Lochner, Deborah Schildkraut, and three anonymous reviewers for valuable comments. Earlier drafts of this article were presented in seminars at Columbia, George Washington, Northwestern, NYU, Princeton, and Yale where we received additional constructive feedback. Finally, we would like to express our gratitude to the staff of the Pennsylvania Commission on Sentencing for providing us with (and answering questions about) the sentencing data employed in this analysis.

¹Trial judges retain broad discretion over sentencing in nearly all noncapital cases. Defendants can usually appeal these sentences only on very narrow grounds of law. Also, even though most cases are resolved via plea bargain, the presiding judge must approve the proposed sentence.

²There is substantial institutional variation in this regard. Incumbent trial judges run in competitive partisan elections in eight states, in competitive nonpartisan elections in 21 states, and in noncompetitive retention elections in 10 states. (In a retention election, the incumbent judge's name appears on the ballot with no opponents listed; she must receive more "yes" than "no" votes to keep her position.) In seven other states, the legislature, governor, or a judicial nominating commission periodically evaluates incumbent judges for retention.

American Journal of Political Science, Vol. 48, No. 2, April 2004, Pp. 247–263

the case, might not judges feel *unconstrained* by electoral pressures?

This article develops and tests a theory specifying the conditions under which trial judges will alter their sentencing behavior to improve their electoral prospects. Monitoring judicial behavior is difficult for voters, so even a single instance of apparent judicial malfeasance can decisively influence an election. Certain characteristics of the informational environment surrounding trial judge elections, however, make perceived underpunishment easier to observe than perceived overpunishment—even though voters may care about both (Gordon and Huber 2002). If judges discount the future value of retaining office, they will take greater pains to minimize the electoral consequences of underpunishment later in their terms. Thus, the theory predicts *unidirectional convergence*: Trial judges will become more punitive as their terms proceed.

Existing studies of electoral incentives predict *bidirectional convergence*: officials more liberal than their constituencies will become more conservative as election approaches, while those more conservative will become more liberal (e.g., Ahuja 1994; Elling 1982; Thomas 1985; Wright and Berkman 1986). We predict that judges will become more punitive irrespective of their position relative to their constituents' preferences. Our theory is not unique in anticipating unidirectional convergence, which could also emerge as a consequence of a bias in the selection of judges that produces incumbents uniformly less punitive than their constituents. Accordingly, we devise a set of critical tests to distinguish among these different accounts.

Our empirical analysis focuses on Pennsylvania, where incumbent trial judges must face voters in noncompetitive retention elections every ten years. Examining over 22,000 Pennsylvania trial court sentences for aggravated assault, rape, and robbery convictions in the 1990s, we confirm that sentences for these crimes are significantly longer the closer the sentencing judge is to standing for reelection. Our results also suggest the superiority of our theoretical account over others. Finally, we impute a baseline estimate of the aggregate increase in prison time that occurs as a consequence of electoral incentives for the cases we examine. We can attribute at least 1,818 to 2,705 years of additional prison time to this electoral dynamic.

Judicial Behavior and Electoral Accountability

The method of choosing trial court judges is a matter of substantial controversy.³ There are obvious democratic

³See, e.g., Citizens for Independent Courts (2000), Dagger (1993), Price (1996), and Reidinger (1987).

concerns associated with removing voters from the choice of officials who exercise enormous authority. At the same time, electing trial judges may compromise other aspects of judicial integrity. Foremost, elections may tie judges too closely to the whims of public opinion (Croley 1995). Elected judges presiding over controversial cases may base their decisions on the potential political effects of those decisions instead of legal precepts or an unbiased reading of the facts of a case.

Representation and the Electoral Connection: The Bidirectional Convergence Hypothesis

To our knowledge, no systematic empirical research directly examines the effect of standing for election on the behavior of criminal court trial judges.⁴ We build on theoretical and empirical work that conceives of electoral control as an agency problem. In this framework, elections serve two fundamental roles. First, as selection devices, they give voters the opportunity to pick agents whose preferences closely mirror their own (Fearon 1999). Limited information, however, implies that voters will typically not fully solve the adverse selection problem. To the extent that the preferences of officials diverge from those of their constituents, elections can serve a second role as incentive mechanisms, inducing officials to approximate constituents' preferences (Barro 1973; Downs 1957; Ferejohn 1986; Mayhew 1974; Miller and Stokes 1963).

In addition to implementing their policy preferences, officials wish to retain office, either to enjoy its perquisites or to influence policy in the future. Consequently, the effectiveness of electoral incentives will increase with the official's perception of the value of retention. This perception, however, is not static. For several reasons, over the course of their terms officials will continuously reevaluate the balance between their own preferences and electoral concerns. First, they may discount the future value of retaining office. At the beginning of their terms, when the need to secure reelection and retain office is a far-off prospect, they will place greater weight on implementing their own preferred policies given their ideologies and the information available to them. Toward the end of a term, however, retaining office becomes a paramount concern.

⁴Gibson (1980, 365–67) demonstrates that elected trial judges who have experienced defeat at the polls are more responsive to the sociopolitical characteristics of their districts than undefeated ones. Kuklinski and Stanga (1979) examine the effect on aggregate judicial sentencing in different counties of voter support for a California referendum endorsing less stringent penalties for certain drug crimes. A substantial literature addresses the behavior of elected state supreme court justices (e.g., Brace and Hall 1997 and Hall 1987, 1992, 1995).

Second, because elections are only periodic, voter evaluation of candidate performance, if it occurs at all, is likely to be temporally proximate to each election (Popkin 1991). As such, officials may ignore constituent preferences when voters are inattentive. Voters may also simply have short memories. (In a competitive electoral environment, these effects are mitigated by a challenger's ability to audit and advertise an incumbent's historical performance.)

Under ordinary circumstances, the incentive effects of elections and their variation over time imply *bidirectional convergence*: as election approaches, officials will moderate their behavior to more closely approximate the wishes of some pivotal constituent (e.g., the median voter). In its starkest form, this hypothesis suggests that officials more liberal than that constituent will drift rightward over the electoral cycle, while more conservative officials will drift leftward. Further, convergence should be *constituency specific*: For example, a senator who is conservative relative to the rest of the chamber may still be liberal relative to her state; bidirectional convergence would predict a rightward shift as election approached (Ahuja 1994; Wright and Berkman 1986).⁵

The Electoral Incentives of Trial Judges: Unidirectional Convergence

The informational and institutional environment in which trial judge elections occur differs in fundamental respects from those of other elected officials. Voters are uninformed about the most basic aspects of these officials' behavior and responsibilities (Hall 1999; Mathias 1990; Sheldon and Lovrich 1983). Exacerbating this paucity of information in certain institutional settings is a lack of contextual cues like party labels (DuBois 1984), because in many states judicial elections are nonpartisan. Further, where judges preserve office via retention election, there are no challengers to provide voters with information about incumbent performance (Aspin 1998; Volcansek 1981). Also, until recently, many states imposed restrictions on position taking by judicial candidates.

This paucity of information does not necessarily imply that elected trial judges will feel entirely unconstrained by the pressures of public opinion. Rather, it alters the manner in which they are constrained. Under conditions of near absolute voter ignorance, information about the adverse consequences of a *single case*, when publicized, can be decisive in swaying voter opinion against a presiding judge. A Chicago trial judge, for example, lost

an election bid in 1986 as a consequence of acquitting a defendant who had allegedly attacked a police officer (Mount 1988). Our consideration of the availability of information to voters in these settings draws on McCubbins and Schwartz's (1984) distinction between types of oversight. Sustained, active "police patrol" oversight by voters is costly. In the current context, passive "fire alarm" oversight occurs when asymmetrically well-informed actors such as victims' rights groups, police officers, and entrepreneurial legislators publicize individual instances of perceived judicial malfeasance.

In the context of criminal justice, fire alarms when sounded will almost always correspond to perceived instances of *underpunishment*, not overpunishment. Nearly all convicts claim their punishment is too severe, and newsworthy cases of wrongful punishment by definition come to light years after their imposition. Underpunishment, however, is more easily observed. News accounts of recidivism make voters aware of convicts who committed additional crimes after (seemingly) brief periods of incarceration. Further, victims' families and the groups mentioned above have clearly defined incentives to publicize specific instances of underpunishment. This asymmetry would motivate even moderate or liberal voters to assume the worst about defendants and judges.⁶

Certain stylized facts are sometimes misapplied to suggest the ineffectiveness or nonexistence of fire alarm oversight. First, we rarely see fire alarms pulled. A perfect fire alarm, however, would never be pulled, because the relevant agent would be deterred from malfeasance by the mere threat of the alarm. In that case, the fact of voter ignorance in trial judge elections would be a consequence of trial judge compliance with public opinion, rather than a cause of judicial autonomy. Likewise, the fact that incumbent trial judges retain office at very high rates is not by itself evidence of judicial independence from public opinion. A high retention rate may signal total autonomy or total subservience—it is impossible to tell absent additional information.

Strong evidence that elected judges do not ignore electoral concerns despite the low probability of loss comes from judges themselves. Based on a survey of 645 trial and appellate judges in 10 states with retention elections, Aspin and Hall conclude, "[T]he behavior of retained judges is shaped by the existence of retention elections even though the probability of losing is low" (1994, 315). At the same time, while judges may be attentive to public opinion, it is unlikely that they are totally beholden to it. As with other officials, judges have their own preferences

⁵Elling (1982) and Thomas (1985) discuss convergence to constituency, but their statistical analyses do not measure constituency preferences.

⁶For a different treatment of asymmetric information revelation, see Canes-Wrone, Herron, and Schotts (2001).

over criminal justice issues. Given these preferences and access to additional information about specific cases that the public lacks, judges are likely to desire to judge; that is, to make distinctions among individual defendants based on their perceived culpability, remorse, or likelihood of recidivism.

If incumbent judges, like all officials, continuously reevaluate the balance between the value of office and implementing their own preferences, the balance will shift increasingly toward the former as election approaches. Further, the fire-alarm nature of trial judge oversight suggests that because voters typically have access only to information about sentencing perceived as overly lenient, the judge's response will be *unidirectional convergence*. In other words, our theory predicts that judges will become more *punitive*, not more representative, over the course of their terms.

Comparing our approach with that taken by Hall (1992) is illuminating. She finds that liberal state supreme court justices in states with short terms are less likely to dissent from decisions upholding the imposition of the death penalty. Her underlying causal story is similar to ours in its informational basis: dissent reveals ideological extremity because voters can compare dissenters with their peers. Such comparisons are impossible in the trial judge environment, however, because each trial judge presides over different cases. Further, a sentence by itself is generically uninformative because of a dearth of public information about individual trials. Only when police unions, victims' rights groups, legislators, and reporters publicize cases are voters likely to conclude that a decision was in some sense "incorrect."

An alternative theory yielding a prediction of unidirectional convergence is preference-based: judges in all districts may be uniformly more liberal than their constituents, either by training, or, in the eyes of conservative critics, because of a proclivity for "coddling criminals."^{7,8} This "uniform judicial liberalism" argument seems at best incomplete, however, given the dissatisfaction of policy *liberals* with criminal sentencing and the fact that voters

⁷Hall (1995) finds evidence that state supreme court justices are more likely to uphold death sentences in the last two years of their terms. While she does not assume all justices are more liberal than their constituents, her prediction rests on a preference-based causal mechanism: the overwhelming support for the death penalty among citizens of states included in her analysis. In a related study, Brace and Hall (1997) find Democratic justices more likely to uphold death penalty convictions in states with short terms lengths.

⁸In the 1990s roughly 85% of American respondents to the General Social Survey answered "not harshly enough" when asked whether they thought local courts dealt too harshly or not harshly enough with criminals. Conservatives are only slightly more likely than liberals to take this position (88% vs. 80%). See also Warr (1995).

themselves (or their elected representatives) put judges in office in the first place. An experiment by Roberts and Edwards (1989) suggests that voters' seemingly punitive tendencies are primarily a consequence of their informational environment. When randomly selected respondents were shown a newspaper account of an assigned sentence, they almost uniformly preferred a more punitive sentence. In contrast, among respondents asked to read a more detailed account of the courtroom proceedings in the case, a much smaller proportion believed the assigned sentence was too lenient. Similar results are reported by Roberts and Doob (1990). In other words, media coverage of criminal proceedings may explain the perception of judicial leniency, a perception that would differ if voters more closely monitored court proceedings.

In the empirical analysis that follows, we develop a set of tests that distinguish not only between the bidirectional and unidirectional convergence hypotheses, but also between a preference-based causal story for unidirectional convergence and our information-centered one.

Data and Method

Pennsylvania's general jurisdiction trial courts are the Courts of Common Pleas. We linked information about sentencing in these courts, state elections, and judges' backgrounds to create the dataset for our analysis.⁹ When a judgeship vacates, replacements are selected via a partisan competitive election. (Mid-term vacancies may be filled via gubernatorial appointment; the seat is considered open in the subsequent election.) In the primary election, judges compete for one or (87% of the time) both of the major party nominations. In the general election, the top vote getter(s) will fill the one or more open seats in a particular judicial district. Judicial districts correspond to counties, although in sparsely populated areas some judicial districts encompass two counties. Once elected, judges stand for reelection every ten years on the basis of a noncompetitive retention vote. Importantly, not all judges in a district are on the same electoral calendar.

Pennsylvania is in an ideal setting for our analysis. Conventional wisdom concerning different methods of electing judges suggests that judges will be most divorced from the electoral connection when they serve long terms and run in nonpartisan retention elections. If this is the case, Pennsylvania trial judges operate in an institutional

⁹Sentencing data were obtained from the Pennsylvania Commission on Sentencing, elections data from the Pennsylvania Department of State, and judges' backgrounds from the *Pennsylvania Manual*. Data on judges' backgrounds was verified using data reported in Steffensmeier and Hebert (1999).

context that will render them least sensitive to periodic voter review. At the same time, however, these very institutional conditions are the ones our theory predicts will produce unidirectional convergence toward punitiveness.

As in all states, the manner in which criminal cases wind their way through the judicial system in Pennsylvania is enormously complex. Consequently, a full accounting of the intricacies of the criminal justice system and the interrelated and strategic behavior of each actor in this system is impractical in the current context. (We point out some implications of the strategic behavior of prosecutors and defendants in the discussion section.) Here, we discuss the range of options available to judges once a defendant pleads guilty or a jury (or judge in a bench trial) finds a defendant guilty of a misdemeanor or felony.¹⁰

Common Pleas judges generally exercise enormous discretion in imposing sentences. Several constraints exist, however. All crimes carry with them statutory maximum sentences, and some have associated mandatory minima as well. Additionally, the Pennsylvania Commission on Sentencing (PCS) offers voluntary sentencing guidelines for most felonies and misdemeanors. Judges are obliged to take account of PCS instructions, but not to abide by them.¹¹ The Pennsylvania guidelines work as they do in many other states: PCS classifies crimes by offense gravity and defendants by prior record. Given those two variables, a judge can determine the recommended sentencing range by referring to a sentencing matrix. The guidelines expand the recommended penalty range upward or downward by 12 months in the presence of aggravating or mitigating factors. Additional matrices exist for separate sentence enhancements such as possession of a deadly weapon during the commission of a crime.

For a given conviction, sentencing judges in Pennsylvania hand down both a minimum and maximum sentence. In cases involving incarceration, the defendant is obliged to spend at least the minimum term in prison before becoming eligible for parole. Subsequently, a state parole board may or may not grant the defendant parole up to the release time specified by the judge as the maximum sentence. Overall, the manner in which Pennsylvania incarcerates and releases defendants falls between the

extreme case of fully indeterminate sentencing (in states where a parole board is granted large discretion to reduce judicially imposed sentences) and fully determinate sentencing (in states where parole has been abolished).

Analyzing Sentencing Behavior

The details of the Pennsylvania criminal justice system suggest a need to avoid several pitfalls in our analysis. First, we must account for a judge's discretion in a given case. We restrict attention to a class of felonies for which judges both always have some discretion in sentencing and typically assign prison time. There are a number of such felonies, including rape, sexual assault, arson, robbery, theft, and possession with intent to distribute Schedule I and II narcotics. To keep the analysis as simple as possible, we focus our attention on all convictions in which the highest count was some form of aggravated (felony) assault, robbery, or rape. These encompass nearly all cases with high offense gravity scores under the Pennsylvania guidelines. We have 22,095 observations for discretionary sentences imposed from 1990 to 1999 according to guidelines issued in 1988, 1994, and 1997.¹²

As noted above, judges assign two sentences for each case. The dependent variable in our analysis is the smaller of these two quantities, measured in months of incarceration. (Summary statistics for model variables appear in Table 1.) This represents the determinate portion of the judge's discretion over sentencing, as defendants must spend at least the smaller sentence behind bars before becoming eligible for parole.¹³

By statute, the smaller sentence imposed by the judge cannot exceed one-half the larger sentence, which itself cannot be greater than the statutory maximum. Additionally, for certain crimes, the law mandates a minimum prison sentence. Together, these rules place upper and lower boundaries on the range of a judge's sentencing options, creating a censoring problem. When a judge's smaller assigned sentence is the statutory minimum, the dependent variable is left-censored. She may have preferred an even lower sentence, but was prohibited by law from imposing it. When a judge's smaller sentence is one-half the statutory maximum, the dependent variable is right-censored; the law prevents her from being more

¹⁰As in all state courts, most cases are settled via plea bargain. Prosecutors and defendants in such instances negotiate a settlement, *subject to the approval of the presiding judge*, whereupon the defendant pleads guilty to reduced charges or in exchange for a recommendation to the judge by the prosecutor of a reduced sentence.

¹¹42 Pa.C.S. § 9781 sets conditions for the appeal of an assigned sentence. If a sentence is within the guideline range, the state or defendant can appeal only for clerical mistakes or if the "application of the guidelines would be clearly unreasonable." If outside the guideline range, the standard for appeal is reasonableness.

¹²The 1988 guidelines were revised in 1991. We have accounted for these changes as well as alterations to the criminal code during the period under study. In a given case, the authoritative guideline is the one in place when the crime was committed. For example, if a crime was committed in 1993 and the trial occurred in 1995, the judge would use the 1991 guidelines, not the 1994 ones.

¹³Our results are nearly identical if we examine the larger sentence imposed.

TABLE 1 Summary Statistics for Model Variables

Variable	Mean	Standard Deviation	Minimum	Maximum
Assigned "Smaller" Sentence (months)	24.83	28.73	0	240
Guideline Minimum Sentence	20.07	19.22	0	120
Guideline Maximum Sentence	32.55	24.27	6	120
1988 Guidelines in Force	0.54	0.50	0	1
1994 Guidelines in Force	0.26	0.44	0	1
Defendant Male (1 = Yes)	0.91	0.28	0	1
Defendant Nonwhite (1 = Yes)	0.66	0.48	0	1
Defendant Age (years)	28.57	9.06	14.73	88.42
Nonnegotiated Guilty Plea (1 = Yes)	0.20	0.40	0	1
Negotiated Guilty Plea (1 = Yes)	0.45	0.50	0	1
Deadly Weapon Enhancement (1 = Yes)	0.11	0.31	0	1
Deadly Weapon Use (1 = Yes)	0.02	0.14	0	1
Counts in Conviction	2.17	2.29	1	90
Rape (1 = Yes)	0.06	0.24	0	1
Robbery (1 = Yes)	0.52	0.50	0	1
Electoral Proximity	0.44	0.28	0	1
Judge Conservatism (Estimated)	13.77	1.35	10.05	15.03
Judge Age (years)	53.07	9.37	35	83
Judge Male (1 = Yes)	0.78	0.41	0	1
Judge Prosecution Experience (1 = Yes)	0.36	0.48	0	1
Republican Percentage of Vote for Attorney General	0.43	0.15	0.20	0.74

N = 22,095, except for judge biographical data and judge conservatism, for which N = 21,776 due to missing data.

punitive. OLS regression produces biased coefficient estimates in the presence of censoring. In order to compensate for censoring problems while retaining the OLS assumption of normally distributed errors, we employ a two-limit tobit model with observation-specific left and right censoring points (Maddala 1983, 160–62; Tobin 1958).¹⁴ Employing this model also allows us to address a second problem created by the 16% of cases in which no prison time was imposed. In these cases, defendants were placed on probation, forced to pay a fine, or given some other form of limited restrictive punishment. We treat these cases as left-censored, assuming they represent punishment less than the minimum jail time.¹⁵

The next issue we confront is that factors other than electoral proximity and statutory limits may explain as-

signed sentences. A failure to control for these will only bias our inferences if the omitted variables are correlated with the included ones. This can occur if different judges—or the same judges at different points in their electoral cycles—preside over different types of cases. In other words, the threat of omitted variables is intimately related to the possibility of nonrandom case assignment, an issue to which we return below. Accordingly, employing case- and judge-level controls is a conservative strategy. If case assignment is nonrandom, these controls minimize the threat of omitted variables bias. If it is random, they improve predictive efficiency.

Crimes, defendants, and cases vary independently in ways that will affect judges' use of their discretion. The Pennsylvania Sentencing Commission's recommended minimum and maximum sentences provide the ideal measures to control for the severity of the offense committed and the defendant's prior criminal record. Guideline sentences reflect consensus within the state about appropriate punishments and the latitude judges should enjoy given the defendant's criminal history and the nature of the crime. They incorporate an enormous amount of information, including victim age, the crime's location, and

¹⁴The tobit model assumes the existence of a latent variable representing, in this case, the sentence the judge would prefer to assign. It is fully observed only when the judge's sentence falls between the censoring boundaries, in this case the statutory minimum sentence and one-half the statutory maximum.

¹⁵This need not imply that nonincarcerative punishments are non-punishment, only that they are, from the perspective of judges and their constituents, less punishment than prison time.

the level of violence.¹⁶ As such, they vary considerably even when one restricts attention to a single crime. Additionally, we employ dummy variables for the applicable sentence guideline regime (1988, 1994, or 1997).¹⁷ As supplementary controls for the nature of particular crimes, we employ indicator variables that distinguish the type of crime (rape and robbery—the baseline category is aggravated assault) and whether it involved the possession or use of a deadly weapon. Finally, we account for the possibility that judges distinguish among defendants based on demographic characteristics. These include age and age-squared (because judges may treat young and old defendants more leniently than others), race, and sex.

Second, we control for variation in the disposition of cases. In 51.5% of the cases in our sample, the defendant was convicted on more than one count. (In only 14% of cases was the defendant convicted on more than three counts.) Because judges can decide whether to impose sentences consecutively or concurrently and which counts to issue sentences on, we examine only the sentences associated with the most severe count on which the defendant was convicted and control for the number of counts. For most cases, this is the only count accompanied by a sentence (we omit from the sample the handful of cases in which the defendant was given prison time only for less severe counts).¹⁸ We also include indicator variables for negotiated and nonnegotiated guilty pleas (the baseline category is conviction at trial).

Finally, judges have their own sentencing ideology. Note that in addressing the unidirectional convergence hypothesis, we are not interested in the primitive preferences of the judges per se except to the extent that they are needed as controls in our statistical models. Measuring judicial ideology is difficult, so we take three approaches. The first is to note that judges' time-invariant ideological proclivities cannot be systematically correlated with where they happen to be in their own electoral cycles. This obviates the need to control for those proclivities. Second, as a robustness check, we control for characteristics of

the judges as proxies for their punitive tendencies. The measures we employ are the judge's age and age-squared, whether the judge was male or female, and whether the judge had prosecution experience (see Goldman 1975; Tate 1981).¹⁹ Third, as a more comprehensive robustness check, we employ judge-specific fixed effects (i.e., one dummy variable per judge—425 variables total) to control for all time-invariant characteristics of the sentencing judge. This approach is the most conservative because it requires no a priori assumptions about how judges' preferences are derived.

Our primary hypothesis concerns the effect of electoral proximity. We code proximity as the number of days elapsed in the judge's term at the date of sentencing divided by 3,653. The measure is thus scaled from zero to one, with zero representing 10 years until the next election, and one an imminent retention vote (Election Day).²⁰ We expect a positive coefficient on this measure: as proximity increases, so should assigned sentences. One other variable, whose relevance we explain below, appears in the summary statistics. It is a measure of district political conservatism on criminal justice issues. Lacking a perfect measure, we employ the district Republican share of the two-party vote in the previous statewide attorney general race.

Results

The presentation of empirical findings proceeds in two stages. First, we provide statistical results that confirm our primary hypothesis. Second, because these results are consistent with several rival explanations, we devise and implement a series of critical tests to distinguish the underlying causal mechanism.

Assessing the Unidirectional Convergence Hypothesis: Initial Results

The unidirectional convergence hypothesis predicts an increase in punitiveness associated with an increase in electoral proximity. Table 2 displays the first round of tobit estimates. The coefficient estimates in column (1) come from a regression that includes the electoral proximity

¹⁶Alternatively, one may employ as controls the defendant's prior record score and the offense gravity score. Because these are ordinal scales that change between guideline regimes and map nonlinearly to the guideline sentences, we strongly prefer using the guidelines themselves. Substituting the former for the latter, however, does not alter our substantive findings.

¹⁷The 1994 and 1997 guidelines did not result in unambiguous increases (or decreases) in punitiveness for the cases we study. Rather, they altered the relative classification of case severity across crimes (e.g., aggravated assault versus robbery) and within crimes (e.g., aggravated assault with significant bodily injury versus simple aggravated assault). Similarly, they revised the scaling of previous offense history.

¹⁸The results do not differ substantially if we confine our analysis to cases with only a single count.

¹⁹We were unable to use judge's race or religion because it is not reliably reported. Additionally, it is impossible to identify the partisan affiliation of most Pennsylvania judges. 87% of judges run in both primaries, and a search of newspaper accounts and judicial biographies failed to reveal partisanship in all but a tiny handful of cases. We do not consider this a major concern, because the party labels of local officials have different meanings in different parts of the state.

²⁰For midterm appointments, we exclude all cases a judge hears before her initial competitive election.

TABLE 2 The Effect of Electoral Proximity on Sentencing: Two-Limit Tobit Models

	(1) Year Effects	(2) Year Effects	(3) Year and Judge Effects
Guideline Minimum	0.66 (11.18)	0.67 (11.35)	0.67 (16.20)
Guideline Maximum	0.23 (5.11)	0.22 (4.97)	0.21 (6.36)
1988 Guideline	3.29 (2.19)	3.43 (2.24)	2.77 (2.35)
1994 Guideline	2.92 (3.27)	3.11 (3.45)	2.51 (3.32)
Defendant Male	8.75 (16.81)	8.67 (16.59)	8.18 (14.80)
Defendant Nonwhite	0.43 (1.34)	0.71 (2.18)	2.71 (7.72)
Defendant Age	0.14 (1.69)	0.15 (1.73)	0.14 (1.83)
Defendant Age Squared	-0.0036 (2.99)	-0.0037 (3.00)	-0.0034 (3.13)
Nonnegotiated guilty plea	-6.40 (14.25)	-6.32 (14.04)	-5.14 (11.30)
Negotiated guilty plea	-6.77 (19.05)	-6.86 (19.16)	-7.50 (19.61)
Deadly Weapon Enhancement	18.93 (25.15)	18.82 (24.98)	17.89 (32.35)
Deadly Weapon Use	-1.07 (0.63)	-0.88 (0.52)	-1.17 (0.95)
Counts in Conviction	1.68 (9.76)	1.70 (9.63)	1.73 (25.39)
Rape	14.03 (15.28)	14.01 (15.21)	12.81 (18.69)
Robbery	6.33 (19.59)	6.40 (19.61)	6.90 (21.19)
Electoral Proximity	4.29 (7.72)	3.70 (6.36)	2.94 (4.23)
Judge Age		0.27 (1.55)	
Judge Age Squared		-0.0025 (1.62)	
Judge Male		2.91 (7.45)	
Judge Prosecution Experience		0.42 (1.33)	
Intercept	-19.47 (8.61)	-23.49 (4.65)	-50.90 (3.80)
Standard error	21.40 (82.42)	21.38 (81.91)	20.55 (179.17)
Log-likelihood	-79893.83	-78774.01	-79034.04

Notes: Dependent variable is the minimum (smaller) sentence assigned by the judge. Absolute values of parameter t-ratios are in parentheses. Eight year dummies omitted from all columns. 425 judge dummies omitted from column (3). N = 22,095 (4,527 left-censored cases, 690 right-censored) for columns (1) and (3). N = 21,776 (4,440 left-censored, 686 right-censored) for column (2). Robust standard errors are employed to estimate t-ratios in columns (1) and (2). All models significant at the .001 level or better.

variable, the controls discussed above, and eight year-specific fixed effects whose coefficient estimates we do not report. The year variables allow us to account for global changes in sentencing practices over time (e.g., those caused by uniform responses to changing state conditions).

The estimates reported in the second and third columns confirm our initial model's insensitivity to changes in model specification. Column (2) estimates come from a regression that includes variables from the first specification plus judge background variables. Column (3) estimates come from a model that includes variables from the first specification plus 425 judge-specific fixed effects whose coefficients we do not report. As stated above, this last specification permits us to control for all static characteristics of the sentencing judge (e.g., ideology).²¹

Before proceeding to our test of the primary hypothesis, we examine the parameter estimates for the control variables. The results are without exception consistent with our expectations and prior research on criminal sentencing (e.g., Adelstein 1978; Albonetti 1997; Bushway and Piehl 2001; Curran 1983; Landes 1971; Mieth and Moore 1985; Zingraff and Thompson 1984) and suggest the validity of our specification and coding. Increases in the guideline sentences lead to increases in actual sentences. A nonmonotonic relationship exists between age and length of incarceration. (Per the estimates from the first specification, 20 year-old defendants can expect to see the most prison time.) Men receive longer sentences than women, and the specification in column (1) suggests that all else equal, judges hand down sentences 13 days longer for nonwhite defendants than for their white counterparts.^{22, 23} One of the strongest predictors of additional

punishment is possession of a deadly weapon. Once one controls for possession of a weapon, however, using it does not significantly add to the sentence. (This is not surprising because using a weapon typically increases the classified severity of the crime, in turn raising the guideline minimum sentence—which is already controlled for.)

We also point out an interesting finding concerning the disposition of cases. The negative coefficient on the negotiated plea variable provides information about the way that plea-bargaining typically occurs (Taha 2001). A positive coefficient would suggest that the dominant form of negotiation between prosecutors and defendants concerns charge reduction: a defendant might plead guilty to a lesser charge and receive a penalty that, while stiff for the reduced charge, is nonetheless lighter than the penalty for the higher count. The negative coefficient that we observe across specifications suggests that negotiation typically concerns the length of the sentence that the prosecutor recommends to the judge given a particular charge, more often than on the charge itself.

Next we consider the effect of electoral proximity. It is important to consider how one should interpret a null finding for this variable. A failure to detect unidirectional convergence is consistent with three accounts of judicial behavior. First, it might demonstrate judicial autonomy. Judges may behave in a manner totally independent of the preferences of their constituents and sentence as they see fit. Second, a null finding could indicate judicial subservience: judges may prioritize the desires of their constituents throughout the electoral cycle. Finally, it could indicate bidirectional convergence, wherein some judges become more conservative as election nears while others become more liberal, leading to zero net effect.

We need not confront this indeterminacy, however. In all specifications, the parameter estimate for electoral proximity is positive and highly statistically significant (one can reject the null hypothesis at above the 0.001 level in all three specifications). All else equal, the sentence imposed by a judge whose election is imminent is likely to be about three to four-and-one-fourth months longer (depending on specification) than if the judge were recently elected or retained. A standard deviation shift in electoral proximity raises an assigned sentence by about 25 to 36 days. The magnitude of the result is substantial. The median sentence in the sample is about 12 months. Consequently, a standard deviation increase in electoral proximity produces a 7 to 10% increase from the median sentence, while a change from zero to one produces an increase of 24 to 37%.

An even more useful measure of the impact of electoral proximity can be derived by imputing an estimate of the aggregate increase in prison time stemming from

²¹We believe this fixed-effect approach is asymptotically consistent, since the number of convictions will approach infinity faster than the number of judges (which is fixed) and years (see Heckman and MaCurdy 1980). In the current setting, consistency appears to be only a theoretical issue. Fixed-effects OLS (whether restricted to uncensored observations or for all cases) produces nearly identical results.

²²Throughout, we report the effect of changes in independent variables on the latent rather than on the observed dependent variable. Because it reflects the sentence a judge *would* hand down in the absence of statutory constraints, this quantity is the more easily interpretable of the two. To calculate for a particular observation the estimated effect of a change in an independent variable on the observed dependent variable, multiply its coefficient by the probability the observation is in the censoring region (Greene 2000, 909). Evaluated at the data sample means and based on the current set of estimates, that probability is approximately 0.68.

²³The effect is larger and significant in columns (2) and (3). We consider this to be disturbing evidence of sentencing disparity, an important topic but one beyond the scope of this article.

judges' desire to secure reelection. Assume that on the first days of their terms, judges feel completely unconstrained by the electoral consequences of their sentencing decisions. For each sentencing decision in our dataset, we can calculate an estimate of the sentence the judge would have imposed had she just been elected or retained, and compare that with an unconstrained prediction. Adjusting for statutory constraints on sentencing, the coefficient estimates in column (1) suggest that the proximity effect augmented sentences for the cases in our dataset by 2,705 years (+/-681), or 5.9% of total prison time.²⁴ For two reasons, this is a conservative estimate. First, even a judge serving her first day in office may feel somewhat constrained by the future electoral consequences of her actions. Given lifetime tenure, she might sentence even less punitively. Second, these figures correspond only to the cases in our Pennsylvania dataset, which comprise only a fraction of total convictions in the state.

Distinguishing the Underlying Causal Mechanism

Findings of unidirectional convergence constitute preliminary evidence in favor of our informational theory. However, an alternative causal mechanism may have generated these estimates. As discussed above, unidirectional convergence is consistent with an alternative account ("uniform judicial liberalism") in which all judges are more lenient than their constituents. Also, bidirectional convergence may be at work, but with a sufficiently large proportion of judges to the left of their constituents to lend the *appearance* of unidirectional convergence ("lopsided bidirectional convergence").

To distinguish our informational story from these accounts, we conduct a series of critical tests. If lopsided bidirectional convergence generated our finding, then we would anticipate that at least *some* judges would become more lenient over the course of their terms. Our theory, by contrast, predicts that *no* judges will become more lenient. The most punitive judges will simply exhibit minimal or no change during their terms, because their sentences are already sufficiently punitive to minimize the risk of a fire alarm being pulled. The approach we adopt to distinguish these accounts modifies the one suggested by Segal, Songer, and Cameron (1994). In a study of appellate court jurisprudence, they first conduct a logit analysis of judges' votes on time-invariant judge characteristics, employing

²⁴Parameter estimates from column (2) suggest a total sentence augmentation of 2,313 years (+/-698); estimates in column (3) suggest 1,818 years (+/-845). Simulated 95% confidence intervals are in parentheses.

the linear prediction from the model as a judge's ideology score. They then employ these scores in subsequent analyses of appeals court decisions.

Our approach differs in several respects. First, the theoretically relevant quantity in our estimator is the *interaction* of ideology and electoral proximity. If our theory is valid, then judges of all ideological stripes should have nonnegative proximity effects. Second, they employ a vector of judge characteristics appropriate for the study of federal appellate judges. In contrast, we employ the trial judge characteristics discussed in the previous section. Further, while they employ the ideology of the appointing president as a component of judicial ideology, our need to distinguish the incentive and selection effects of elections makes the inclusion of a district ideology measure inappropriate for this test. (Such a measure is appropriate for a separate test, discussed below.) Third, a two-stage approach will tend to produce biased standard errors because the linear prediction used in the second stage is a stochastic regressor. We therefore estimate the model using both the two-stage and a full information maximum likelihood (FIML) approach.²⁵ Parameter estimates from each model are displayed in Table 3.

The second specification employs county fixed effects as additional controls for unexplained heterogeneity. Note the similar results across specifications: the conditional effect of electoral proximity given an ideology score of zero is positive in each, although statistical significance is reduced slightly in the FIML estimation. Also, the coefficient on the interaction with the ideology score is consistently negative, implying that the proximity effect is reduced for more punitive judges. If lopsided bidirectional convergence is at work, however, then the effect of electoral proximity on more punitive judges should not only be smaller—it should be negative. For the estimates for columns (1) and (2), the *most* punitive judge has an ideology score of 15.03. Substituting, we find that the net proximity effect for that judge is *positive* 2.72 or 2.76 months

²⁵We thank an anonymous reviewer for the initial idea. A judge's age, sex, and prosecutorial experience were used to construct our proxy measure of judicial preferences (e.g., Goldman 1975; Tate 1981). See also *supra*, note 19. In the two-stage approach, stage one consisted of running a tobit model predicting judicial sentences including the judge data and additional control variables (The results reported here are robust to alternative statistical specifications). We then used the observed coefficients on the judge variables to construct the ideology measure: $Judge\ Conservatism = -0.4058058 * judge\ age - 0.0035203 * judge\ age\ squared + 2.944905 * judge\ male + 0.3898156 * prosecution\ experience$. The FIML estimator works as follows: Let y_i^* be the latent punishment variable, z_i a vector of judge-specific characteristics, and x_i a vector of other variables (including electoral proximity). The latent regression model, estimated as a tobit, is $y_i^* = x_i'\beta + z_i'\gamma + \xi\ proximity_i(z_i;\gamma) + \varepsilon_i$, with ξ the relevant interaction coefficient. Code is available upon request.

TABLE 3 Judicial Preferences and Electoral Proximity: Two-Limit Tobit Models

	(1) Year Effects 2-Stage	(2) Year and County Effects, 2-Stage	(3) Year Effects, FIML
Guideline Minimum	0.67 (11.38)	0.67 (11.43)	0.66 (15.54)
Guideline Maximum	0.22 (4.94)	0.23 (5.23)	0.22 (6.52)
1988 Guideline	3.43 (2.24)	3.91 (2.55)	3.48 (2.88)
1994 Guideline	3.10 (3.43)	3.21 (3.56)	3.13 (4.06)
Defendant Male	8.66 (16.58)	8.18 (15.91)	8.65 (15.21)
Defendant Nonwhite	0.71 (2.19)	2.85 (8.32)	0.71 (2.17)
Defendant Age	0.15 (1.73)	0.11 (1.32)	0.15 (1.86)
Defendant Age Squared	-0.0037 (3.00)	-0.0032 (2.60)	-0.0037 (3.30)
Nonnegotiated guilty plea	-6.34 (14.11)	-5.47 (11.62)	-6.33 (14.72)
Negotiated guilty plea	-6.86 (19.20)	-7.85 (20.57)	-6.88 (19.59)
Deadly Weapon Enhancement	18.79 (24.93)	18.14 (24.04)	18.80 (33.56)
Deadly Weapon Use	-0.88 (0.52)	-0.83 (0.49)	-0.95 (0.76)
Counts in Conviction	1.69 (9.64)	1.75 (9.56)	1.69 (24.71)
Rape	14.01 (15.21)	12.91 (14.18)	14.03 (20.04)
Robbery	6.42 (19.67)	6.78 (20.76)	6.39 (19.31)
Electoral Proximity	13.44 (2.39)	14.68 (2.62)	16.65 (1.64)
Judge Age			0.58 (1.87)
Judge Age Squared			-0.0053 (1.90)
Judge Male			3.79 (5.97)
Judge Prosecution Experience			0.87 (1.71)
Judge Conservatism	1.20 (6.52)	0.71 (3.76)	
Judge Conservatism * Electoral Proximity	-0.71 (1.77)	-0.79 (1.97)	-0.69 (2.43)
Intercept	-30.54 (9.92)	1.94 (0.19)	-32.54 (3.76)
Standard error	21.38 (81.91)	21.01 (80.54)	21.32 (178.16)
Log-likelihood	-78773.69	-78377.44	-78850.52

Notes: Dependent variable is the minimum (smaller) sentence assigned by the judge. Absolute values of parameter t-ratios are in parentheses. Eight year dummies omitted from all columns. 66 county dummies omitted from column (2). N = 21,776 (4,440 left-censored cases, 686 right-censored). Robust standard errors are employed to estimate t-ratios in columns (1) and (2). All models significant at the .001 level or better.

(depending on specification). In the FIML estimation, the most punitive judge has an ideology score of 20.34. The estimated net proximity effect for that judge is 2.5 months. If even the most conservative judge in the sample becomes more punitive over the course of the term, then lopsided bidirectional convergence cannot explain our findings.

It is possible, however, that judges are all more liberal than their constituencies. If uniform judicial liberalism generated our findings, then it should be the case that the effect of electoral proximity will be largest in the most *conservative* districts: Liberal judges would have to traverse a greater ideological distance when sentencing in order to appeal to their constituents. If our informational story is correct, however, the effect of electoral proximity should be largest in the most *liberal* districts. On average, judges will reflect the preferences of their constituents with some error. All will face similar incentives to become more punitive, but the judges coming from the most conservative districts will be punitive to begin with, and will consequently have less distance to traverse.

The uniform judicial liberalism account, in other words, predicts a positive coefficient on the interaction between electoral proximity and district conservatism, while our informational account predicts a negative coefficient. Initial estimates, employing our measure of district conservatism (Republican vote share in the previous statewide attorney general election), are displayed in columns (1) and (2) of Table 4. As anticipated by our informational theory, the coefficient on the interaction is negative and statistically significant in both specifications.²⁶

To make our results more interpretable, we display the net effect of electoral proximity in months (and its 95% confidence interval) as a function of district conservatism in the top two panels of Figure 1. Viewed across judicial districts, the average Republican vote share in an attorney general race was 58.1%. In a district one standard deviation more liberal than the average, a judge facing an imminent retention vote will sentence 2.21 to 2.66 months longer than one who has just been elected or retained. At the same time, in a county one standard deviation more conservative than average, the proximity effect is -3.37 to -5.1 months.

This last finding is troubling at first glance. Our theory predicts that no judges should become more lenient over their electoral cycles, yet the model implies that judges in a number of counties will do just that. This apparent evidence against our theory, however, is an artifact of the model's specification. As is evident from the

top panels of Figure 1, the specification of the interaction effect is linear, which forces the proximity effect to be negative for some values of district conservatism. A more flexible approach is to employ a nonlinear interaction specification such as the quadratic, as we do in columns (3) and (4). The lower panels of Figure 1 display the estimates of the proximity effect given the more flexible specification. These plots confirm our theory's predictions: in liberal districts, the effect of electoral proximity is large and statistically significant, while in moderate to conservative districts, it is statistically indistinguishable from zero. At no point is the effect negative and significant.

To summarize, all judges, even the most punitive, increase their sentences as reelection nears, demonstrating that our finding is not attributable to bidirectional convergence with a preponderance of lenient judges. Similarly, the proximity effect is largest in the least punitive counties, thereby ruling out the possibility that uniform judicial liberalism explains the observed relationship. Overall, these tests provide strong support for our informational model.

Discussion

Four plausible objections may be raised against our findings. First, perhaps judges simply "learn" to become more punitive as they grow older and more experienced. We offer two responses. First, the model estimates reported in column (2) of Table 2 control for judge age. Second, we reestimated the models reported in Table 2 accounting for the number of years a judge has been on the bench (and that quantity squared). The proximity effect remains positive and statistically significant, although collinearity reduces the magnitude of the coefficient to between 1.51 and 2.35, depending on specification.

We also considered whether the proximity effect changes from term to term. On the one hand, there are reasons to suspect that judges will respond to impending electoral review less after they have been through the process at least once. Judges' terms in Pennsylvania last 10 years, and the median judge in our dataset first achieves her position at age 45. Half of all judges, then, decide whether to run for retention a second time at age 65 or older. Since retirement is a likely option, the value of retaining the office may have declined by that point. Additionally, judges learn over time what they must do to be retained. A comfortable retention margin the first time around may prompt a decrease in concern about appearing soft on crime. On the other hand, these same effects may actually *magnify* the electoral proximity effect. Judges who have won retention before may recognize that only what they do late in their terms is noticed and consequently become even less

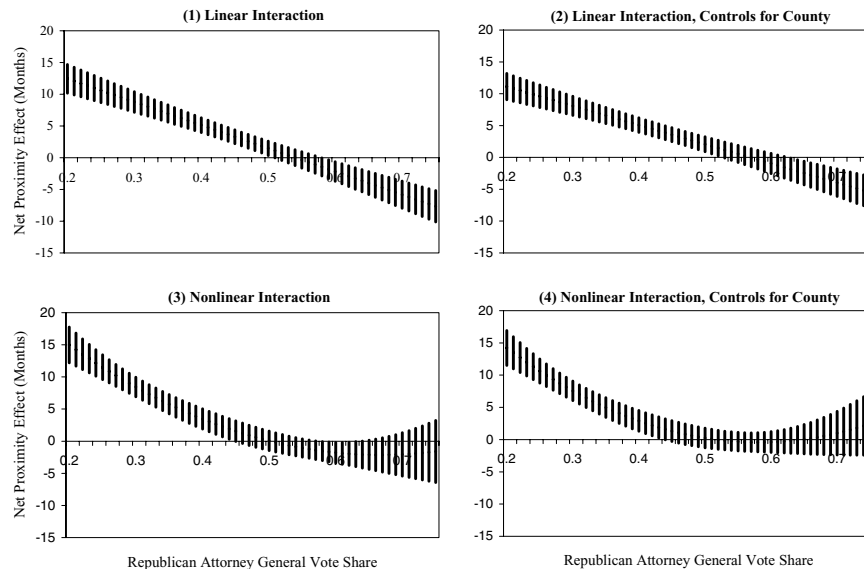
²⁶As a robustness check, we also employed a more traditional measure of conservatism: the Republican margin in the two-party vote for President. The results were nearly identical.

TABLE 4 Constituent Preferences and Electoral Proximity: Two-Limit Tobit Models

	(1) Year Effects	(2) Year and County Effects	(3) Year Effects	(4) Year and County Effects
Guideline Minimum	0.65 (11.12)	0.65 (11.17)	0.65 (11.11)	0.65 (11.12)
Guideline Maximum	0.23 (5.13)	0.24 (5.41)	0.23 (5.18)	0.24 (5.48)
1988 Guideline	4.06 (2.70)	3.63 (2.41)	3.75 (2.49)	3.41 (2.27)
1994 Guideline	2.98 (3.33)	2.84 (3.21)	2.82 (3.16)	2.73 (3.09)
Defendant Male	8.75 (16.93)	8.30 (16.23)	8.67 (16.79)	8.29 (16.20)
Defendant Nonwhite	1.70 (5.22)	2.85 (8.41)	1.68 (5.16)	2.87 (8.45)
Defendant Age	0.17 (2.01)	0.13 (1.53)	0.17 (2.04)	0.13 (1.49)
Defendant Age Squared	-0.0040 (3.27)	-0.0034 (2.80)	-0.0040 (3.30)	-0.0034 (2.76)
Nonnegotiated guilty plea	-6.24 (13.93)	-5.34 (11.40)	-6.25 (13.99)	-5.26 (11.23)
Negotiated guilty plea	-6.99 (19.73)	-7.60 (20.10)	-7.03 (19.88)	-7.55 (19.96)
Deadly Weapon Enhancement	18.49 (24.49)	17.94 (23.83)	18.27 (24.22)	17.81 (23.68)
Deadly Weapon Use	-0.97 (0.57)	-0.94 (0.56)	-0.82 (0.49)	-0.87 (0.53)
Counts in Conviction	1.71 (9.85)	1.75 (9.73)	1.69 (9.68)	1.75 (9.67)
Rape	14.01 (15.31)	12.99 (14.34)	13.94 (15.22)	13.05 (14.39)
Robbery	6.75 (20.76)	6.89 (21.26)	6.77 (20.85)	6.92 (21.34)
Electoral Proximity	19.75 (10.73)	17.16 (9.26)	32.75 (6.26)	33.59 (6.40)
Republican Percentage of Vote for Attorney General	30.45 (15.97)	0.67 (0.15)	12.23 (0.92)	9.45 (0.64)
Republican Percentage of Vote for Attorney General Squared			21.02 (1.42)	-6.84 (0.41)
Republican Percentage of Vote for Attorney General * Electoral Proximity	-36.55 (9.58)	-30.19 (7.87)	-104.53 (4.20)	-116.90 (4.66)
Republican Percentage of Vote for Attorney General Squared * Electoral Proximity			78.27 (2.84)	100.45 (3.60)
Intercept	-34.71 (13.94)	19.25 (2.10)	-30.61 (8.36)	-23.34 (1.52)
Standard error	21.31 (82.17)	20.97 (81.14)	21.28 (82.22)	20.96 (81.25)
Log-likelihood	-79762.75	-79424.8	-79733.28	-79412.66

Notes: Dependent variable is the minimum (smaller) sentence assigned by the judge. Absolute values of parameter t-ratios are in parentheses. Eight year dummies omitted from all columns. 66 county dummies omitted from columns (2) and (4). N = 22,095 (4,527 left-censored cases, 690 right-censored). Robust standard errors are employed to estimate t-ratios in all columns. All models significant at the .001 level or better.

FIGURE 1 The Net Effect of Electoral Proximity as a Function of Constituent Preferences



punitive early on (if they so desire). Thus, the net proximity effect—the difference between a judge’s sentencing just after retaining office and just before facing the voters again—might actually increase because some veteran judges’ early term sentencing is even less punitive.

We reproduced the specifications in Table 2, this time including a measure of how many terms a judge has served and an interaction of this quantity with the proximity variable. We continue to find a statistically significant electoral proximity effect, but this effect declines somewhat in a judge’s subsequent terms. The coefficient on electoral proximity becomes larger, ranging from 3.91 to 6.22, while the coefficient on the interaction between term and proximity receives a negative coefficient that ranges between -0.46 and -1.28 . (We also find that the number of terms a judge serves leads to larger sentences in the column (1) and (2) specifications [1.47 and 1.79 months per term, respectively, both statistically significant] and a statistically insignificant and negative effect [-0.66] in the column (3) specification.) These results confirm that electoral proximity is not strictly a first-term effect, although the magnitude of the effect does appear to decline somewhat in later terms.

A second concern may be raised about our measure of district punitiveness, which we employ to reject the uniform judicial leniency hypothesis. This measure, the Republican vote share in the last statewide attorney general’s race, is highly correlated with one other salient feature of judicial districts: the number of judges voters are asked to review (the correlation is -0.88 for the cases

in our dataset).²⁷ In districts with fewer judges, voters might have an easier time monitoring judges throughout the course of their entire terms via active “police-patrol” oversight.²⁸ If this is the case, our finding concerning district punitiveness might actually be caused by variation in the ease of voter monitoring. To eliminate this alternative explanation for the findings reported in Table 4, we reestimated those models including the number of judges in a district and that number squared plus these quantities interacted with electoral proximity. With the inclusion of these additional variables, we still find that the proximity effect is largest in the most liberal districts. We also find that the proximity effect is larger in districts with more judges.

Third, any nonrandom assignment of cases to judges may bias our results. Whether or not different judges hear systematically different cases, however, will not create the observed proximity effect unless individual judges hear different types of cases at different points during their terms and we did not control for this variation. Before considering case assignment methods directly, it is important to remember that all of our statistical models account for variation in case seriousness, judicial discretion, case

²⁷All other static features of counties that might create unobserved heterogeneity are controlled for in our judge fixed-effects specification. Additionally, the estimates reported in columns (1) and (2) of Table 2 are robust to the inclusion of county fixed effects.

²⁸Alternatively, in districts with many judges, a judge might believe that a “bad” decision would be drowned out by other news or unlikely to be attributed to her at election time.

disposition, and defendant characteristics directly. To be sure that no additional assignment-related factors were influencing our results, we investigated how cases are actually assigned. In small counties with only one or two judges, this is not a concern. However, selection may play a role in large counties such as Philadelphia or Allegheny.

We contacted the Administrative Office of the Philadelphia Courts to inquire about case assignment methodology. In Philadelphia, cases are divided into three pools: homicides, other major cases (which encompass those that we consider), and minor (“list room”) cases. During their first two to three years on the bench, judges are generally restricted to list room cases. After that period, judges are randomly selected to trials, with slightly reduced selection probabilities afforded to historically slow judges with large open caseloads.²⁹ As an additional supplementary test, we therefore reran the models reported in Table 2 for all judges in Philadelphia after discarding cases heard by judges in their first three years of their first term.³⁰ The proximity coefficient remains large and statistically significant.

Finally, one might believe that the electoral proximity effect is incorrectly attributing to judges the strategic behavior of other officials in the criminal justice system. Prosecutors, for instance, might become more punitive as their reelection nears. The nature of judicial elections in Pennsylvania, however, allows us to reject this explanation. In a given district, a prosecutor serves a fixed four-year term. Judges, however, serve ten-year terms and within a district judges are on different electoral calendars. In other words, if prosecutors did ratchet up their efforts to secure more punitive sentences in a given year, this would not correspond to higher values of the proximity measure because judges serve longer terms and judges are in different places in their electoral cycles. Also, our case disposition controls and year fixed effects would capture any uniform changes in how cases are handled across time. Finally, we can directly control for the effect of prosecutor behavior in our analysis within Philadelphia County reported above. There is only a single elected prosecutor in Philadelphia, and the year-effects will control for variation in her effort to secure different sentences over time. With these controls in place, the proximity effect for judges persists.

Additionally, suppose prosecutors and defendants anticipated a judge’s electoral concerns and altered

their courtroom strategies accordingly. Defense attorneys might seek to delay sentencing until after an election, whereas prosecutors might seek to accelerate it. If defense tactics in this regard dominate the sequence, this would bias against our findings, as judges would have fewer cases with which to demonstrate punitiveness toward the ends of their terms. If prosecutor tactics dominate, this would contribute to our finding, but not reject our basic story. Such anticipative action would still constitute evidence that judges take electoral considerations into account. Electoral effects might also filter into the plea bargaining between defendants and prosecutors, with prosecutors able to extract higher sentences from defense counsel when cases are before a judge whose reelection is near. This would again confirm our finding—bargaining about sentencing takes place in the shadow of a judge’s increasing punitiveness as reelection nears. Insofar as the other factors that influence bargaining are unlikely to vary systematically over the course of a judge’s term, they cannot explain our results.

Conclusion

Our analysis provides insight into two important areas of concern for political scientists: the state’s use of coercion and the nature of representation. Judges assign sentences to convicted criminals, determining in part how governments use their authority to deny liberty. We provide evidence that judges become significantly more punitive the closer they are to standing for reelection. In Pennsylvania, for the time period and crimes we analyze, we can attribute more than two thousand years of additional incarceration to this dynamic. This may imply judges sentence too harshly near elections, or too leniently early in their terms. In either case, it implies a downside to electoral control of judges. The power to incarcerate is applied on a case-by-case basis, and we can attribute substantial inconsistency in the exercise of this power to the electoral connection.

Critics of judicial election might be tempted to seize on this result to justify removing judges from direct citizen review. It is not clear, however, that the same phenomenon would not occur with any other form of periodic review—for example, by governors or special commissions. Moreover, there may be larger costs associated with removing judges from such review. The electoral connection may have pernicious effects on consistency, but for some, this may be an acceptable side effect of ensuring that judges’ decisions are at least partially representative of citizen preferences.

Our research also provides insight into the relationship between citizens and elected officials. Elected judges

²⁹We have confirmed that there is no correlation between case seriousness (measured using the guideline minimum sentence) and electoral proximity after the first three years of a judge’s first term.

³⁰The findings for Philadelphia County are also robust to the inclusion of the judge experience measures discussed in this section.

in Pennsylvania are bound by the (weak) threat of losing office, and alter their behavior accordingly. We can thus say with near certainty that at least in this case, elections are not simply a method of selecting “good types” (Fearon 1999). Further, our analysis suggests an important point about information flows in electoral environments where, under ordinary circumstances, voters know almost nothing about officials’ behavior. Because voters are more likely to learn about perceived instances of underpunishment than overpunishment, reelection-minded trial judges might take steps to sentence more harshly than they would if they were not bound by periodic review. Our statistical tests demonstrate that it is this mechanism, and not a bias in selection that makes judges unusually liberal relative to their constituents, that generates the observed sentencing variation.

This research sets the stage for more extensive inquiry into the comparative politics of judicial selection (cf. Brace and Hall 1995). Even in the low information setting created by nonpartisan retention elections, and despite the ten-year terms that afford judges significant distance from electoral review, Pennsylvania trial judges appear to respond to the potential electoral consequences of sentencing leniently by becoming more punitive as reelection approaches. At the very least, we can conclude that the retention method does not remove politics from the sentencing process.

No method of selection is perfect, but at present we know little about the trade-offs associated with mechanisms other than the one studied here. We wish to build on this project to compare across term length, informational environment (competitive versus noncompetitive, partisan versus nonpartisan), immediate political principals (voters, governors, legislatures), and court (general jurisdiction versus appellate). Parsing out the effects of this variation requires additional theoretical and empirical work.

This, of course, is not an easy task. Gathering data about the behavior of trial court judges is time consuming and expensive. Understanding the restrictions (formal and informal) placed on judicial discretion is similarly complicated. Comparative analysis of electoral systems is made possible, however, by the enormous institutional variation afforded by the American states, which serve in this regard as institutional “laboratories,” to borrow Louis Brandeis’ well-known metaphor. This variation permits us to test theoretically derived claims about responsiveness, representation, and fairness in different settings while holding fixed the broad contours of the legal system. Further, differences in the methods employed to select trial judges will contribute to our understanding of how the techniques used to choose and subsequently evaluate other elected officials influence their behavior.

References

- Adelstein, Richard P. 1978. “The Plea Bargain in Theory: A Behavioral Model of the Negotiated Guilty Plea.” *Southern Economic Journal* 44(3):488–503.
- Ahuja, Sunil. 1994. “Electoral Status and Representation in the Senate: Does Temporal Proximity to Election Matter?” *American Politics Quarterly* 22(1):104–18.
- Albonetti, Celesta A. 1997. “Sentencing under the Federal Sentencing Guidelines: Effects of Defendant Characteristics, Guilty Pleas, and Departures on Sentence Outcomes for Drug Offenses, 1991–1992.” *Law & Society Review* 31(4):789–822.
- American Bar Association, Governmental Affairs Office. 1997. *An Independent Judiciary: Report of the Commission on Separation of Powers and Judicial Independence*. Washington: American Bar Association.
- American Bar Association, Standing Committee on Judicial Independence. 2000. “Standards on State Judicial Selection” [Online]. Available: <http://www.abanet.org/judind/publ/reformat.pdf>.
- Aspin, Larry T. 1998. “Campaigns in judicial retention elections: do they make a difference?” *Justice System Journal* 20(1):1–15.
- Aspin, Larry T. 1999. “Trends in Judicial Retention Elections, 1964–1998.” *Judicature* 83(September/October):79–81.
- Aspin, Larry T., and William K. Hall. 1994. “Retention Election and Judicial Behavior.” *Judicature* 77(May/June):306–15.
- Barro, Robert J. 1973. “The Control of Politicians: An Economic Model.” *Public Choice* 14(Spring):19–42.
- Baum, Lawrence. 1983. “The Electoral Fate of Incumbent Judges in the Ohio Court of Common Pleas.” *Judicature* 66(April):420–30.
- Brace, Paul R., and Melinda Gann Hall. 1995. “Studying Courts Comparatively: The View from the American States.” *Political Research Quarterly* 48(1):5–29.
- Brace, Paul R., and Melinda Gann Hall. 1997. “The Interplay of Preferences, Case Facts, Context, and Rules in the Politics of Judicial Choice.” *Journal of Politics* 59(4):1206–31.
- Bushway, Shawn D., and Anne Morrison Piehl. 2001. “Judging Judicial Discretion: Legal Factors and Racial Discrimination in Sentencing.” *Law & Society Review* 35(4):733–64.
- Canes-Wrone, Brandice, Michael C. Herron, and Kenneth W. Shotts. 2001. “Leadership and Pandering: A Theory of Executive Policymaking.” *American Journal of Political Science* 45(3):532–50.
- Citizens for Independent Courts. 2000. *Uncertain Justice: Politics and America’s Courts*. Washington: Century Foundation.
- Croley, Steven P. 1995. “The Majoritarian Difficulty: Elective Judiciaries and the Rule of Law.” *University of Chicago Law Review* 62(Spring):689–794.
- Curran, Debra A. 1983. “Judicial Discretion and Defendant’s Sex.” *Criminology* 21(1):41–58.
- Dagger, Richard. 1993. “Playing Fair with Punishment.” *Ethics* 103(3):473–88.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.
- Dubois, Philip L. 1984. “Voting Cues in Nonpartisan Trial Court Elections: A Multivariate Assessment.” *Law & Society Review* 18(3):395–436.

- Elling, Richard C. 1982. "Ideological Change in the United State Senate: Time and Electoral Responsiveness." *Legislative Studies Quarterly* 7(1):75–92.
- Fearon, James D. 1999. "Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance." In *Democracy, Accountability, and Representation*, ed. Bernard Manin, Adam Przeworski, and Susan Stokes. New York: Cambridge University Press, pp. 55–97.
- Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50(1–3):5–25.
- Gibson, James L. 1980. "Environmental Constraints on the Behavior of Judges: A Representational Model of Judicial Decision Making." *Law & Society Review* 14(2):343–70.
- Goldman, Sheldon. 1975. "Voting Behavior on the United States Courts of Appeals Revisited." *American Political Science Review* 69(2):491–506.
- Gordon, Sanford C., and Gregory A. Huber. 2002. "Information, Evaluation, and the Electoral Incentives of Criminal Prosecutors." *American Journal of Political Science* 46(2):334–51.
- Greene, William H. 2000. *Econometric Analysis*, 4th ed. Upper Saddle River, NJ: Prentice Hall.
- Grodin, Joseph H. 1988. "Developing a Consensus of Constraint: A Judge's Perspective on Judicial Retention Elections." *Southern California Law Review* 61(September): 1969–83.
- Hall, Melinda Gann. 1987. "Constituent Influence in State Supreme Courts: Conceptual Notes and a Case Study." *Journal of Politics* 49(4):1117–24.
- Hall, Melinda Gann. 1992. "Electoral Politics and Strategic Voting in State Supreme Courts." *Journal of Politics* 54(2):427–46.
- Hall, Melinda Gann. 1995. "Justices as Representatives: Elections and Judicial Politics in America." *American Politics Quarterly* 23(4):485–503.
- Hall, Melinda Gann. 1999. "Ballot Roll-Off in Judicial Elections: Contextual and Institutional Influences on Voter Participation in the American States." Presented at the annual meeting of the American Political Science Association.
- Heckman, James J., and Thomas E. MaCurdy. 1980. "A Life Cycle Model of Female Labour Supply." *Review of Economic Studies* 47(1):47–74.
- Kuklinski, James H., and John E. Stanga 1979. "Political Participation and Government Responsiveness: The Behavior of California Superior Courts." *American Political Science Review* 73(4):1090–99.
- Landes, William M. 1971. "An Economic Analysis of the Courts." *Journal of Law and Economics* 14(1):61–107.
- Maddala, G.S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. New York: Cambridge University Press.
- Mathias, Sara. 1990. *Electing Justice: A Handbook of Judicial Election Reforms*. Chicago: American Judicature Society.
- Mayhew, David R. 1974. *Congress, the Electoral Connection*. New Haven: Yale University Press.
- McCubbins, Mathew D., and Thomas Schwartz. 1984. "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms." *American Journal of Political Science* 28(1):165–79.
- Miethe, Terence D., and Charles A. Moore. 1985. "Socioeconomic Disparities under Determinate Sentencing Systems: A Comparison of Preguideline and Postguideline Practices in Minnesota." *Criminology* 23(2):337–64.
- Miller, Warren E., and Donald E. Stokes. 1963. "Constituency Influence in Congress." *American Political Science Review* 57(1):45–56.
- Mount, Charles. October 15, 1988. "Judge Taken Off Bench by Voters Makes Bid to Return." *Chicago Tribune*, sec. 1, p. 5.
- Pastore, Ann L., and Kathleen Maguire, ed. 2002. *Sourcebook of Criminal Justice Statistics* [Online]. Available: <http://www.albany.edu/sourcebook/>.
- Popkin, Samuel L. 1991. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.
- Price, Polly J. 1996. "Selection of State Court Judges." In *State Judiciaries and Impartiality: Judging the Judges*, ed. Roger Clegg. Washington: National Legal Center for the Public Interest, pp. 9–38.
- Reidinger, Paul. 1987. "The Politics of Judging." *ABA Journal* 73(April): 52–58.
- Roberts, Julian V., and Anthony N. Doob. 1990. "News Media Influences on Public Views of Sentencing." *Law & Human Behavior* 14(5):451–68.
- Roberts, Julian V., and Don Edwards. 1989. "Contextual Effects in Judgments of Crimes, Criminals, and the Purposes of Sentencing." *Journal of Applied Social Psychology* 19(11):902–17.
- Segal, Jeffrey, Donald Songer, and Charles Cameron. 1994. "The Hierarchy of Justice: Testing A Principal-Agent Model of Supreme Court-Circuit Court Interactions." *American Journal of Political Science* 38(3):673–96.
- Sheldon, Charles H., and Nicholas P. Lovrich. 1983. "Knowledge and Judicial Voting: The Oregon and Washington Experience." *Judicature* 67(November):234–45.
- Steffensmeier, Darrell, and Chris Hebert. 1999. "Women and Men Policymakers: Does the Judge's Gender Affect the Sentencing of Criminal Defendants?" *Social Forces* 77(3):1163–96.
- Taha, Ahmed E. 2001. "The Equilibrium Effect of Legal Rule Changes: Are the Federal Sentencing Guidelines Being Circumvented?" *International Review of Law and Economics* 21(3):251–69.
- Tate, C. Neal. 1981. "Personal Attribute Models of the Voting Behavior of U.S. Supreme Court Justices: Liberalism in Civil Liberties and Economics Decisions, 1946–1978." *American Political Science Review* 75(2):355–67.
- Thomas, Martin. 1985. "Election Proximity and Senatorial Roll Call Voting." *American Journal of Political Science* 29(1):96–111.
- Tobin, James. 1958. "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26(1):24–36.
- Volcansek, Mary L. 1981. "An Exploration of the Judicial Election Process." *Western Political Quarterly* 34(4):572–77.
- Warr, Mark. 1995. "Poll Trends: Public Opinion on Crime and Punishment." *Public Opinion Quarterly* 59(2):296–310.
- Wright, Gerald C., and Michael B. Berkman. 1986. "Candidates and Policy in United States Senate Elections." *American Political Science Review* 80(2):567–88.
- Zingraff, Matthew, and Randall Thompson. 1984. "Differential Sentencing of Women and Men in the U.S.A." *International Journal of the Sociology of Law* 12(4):401–13.