Judicial Decision Making: A Dynamic Reputation Approach

Alma Cohen, Alon Klement, and Zvika Neeman

ABSTRACT

We seek to contribute to an understanding of how judicial elections affect the incentives and decisions of judges. We develop a theoretical model suggesting that judges who are concerned about their reputation will tend to decide against their prior decisions as they approach elections. That is, judges who imposed a large number of severe sentences in the past and are thus perceived to be strict will tend to impose less severe sentences prior to elections. Conversely, judges who imposed a large number of light sentences in the past and are thus perceived to be lenient will tend to impose more severe sentences prior to elections. Using data from the Pennsylvania Commission on Sentencing, we test, and find evidence consistent with, the predictions of our model.

1. INTRODUCTION

In their book *The Behavior of Federal Judges*, Epstein, Landes, and Posner (2013) contrast two alternative theories of judicial behavior: the legalistic theory, which assumes that judges decide cases in accordance with orthodox norms of judicial decision making, and the realistic the-

ory, which conjectures that judges, like other economic agents, decide cases according to their preferences and incentives. Following the latter approach, the book suggests a labor market model for judicial behavior and presents a comprehensive empirical examination of this model. The book concludes that the model provides a convincing account of actual judicial behavior.

In this paper, we follow the realistic approach as well, and we seek to contribute to an understanding of how incentives and preferences influence judicial decisions. In particular, we put forward a dynamic approach to the analysis of decision making by judges who face elections. Instead of aggregating the decisions of each judge over time, we propose instead to consider how different decisions by the same judge vary over the election cycle and, in particular, how decisions are influenced by the judge's past decisions. This approach, we suggest, can generate theoretical and empirical insights.

Our approach is based on a dynamic reputation model of judicial behavior, which differs from the static labor market model that is proposed by Epstein, Landes, and Posner. Reputation models assume that an agent (in this case, a judge) holds private information about some of her own characteristics. A principal (in this case, the general public or any other audience that affects the judge's future welfare) does not know this information. The principal's belief about the agent's characteristics is updated on the basis of the agent's observed actions. In such a setting, the agent may make choices aimed at establishing a reputation that will increase her long-term payoffs. Since decisions are made repeatedly over time, the agent's present decision depends on past observed decisions.[1]

In a previous paper (Klement and Neeman 2013), two of us showed that arbitrators, who may be vetoed by the litigants upon being selected to arbitrate a specific case, may try to establish a reputation for being neutral and unbiased. To establish such a reputation, arbitrators may deliver decisions that they know to be incorrect. In particular, we identified a possible conflict between the positive selection effect that is induced by allowing litigants to veto arbitrators on the basis of information about their past win/loss ratios and the negative incentive effect that is induced by such selection. Whereas selection allows screening of biased arbitra-

---

1. For a comprehensive review of reputation models in economics, see Mailath and Samuelson (2006).

tors, incentive effects may cause unbiased arbitrators to deliver incorrect decisions to avoid a reputation of being biased.[2]

We suggest that similar reputational considerations may affect judicial decision making in circumstances in which judges stand for reelection or reappointment. We present a simple model in which a judge is asked to apply her discretion in sentencing a convicted felon. The judge may impose either a light or a severe sentence. The judge knows what the general public's preferred sentence would be, given the factual and legal circumstances of the case. Judges are either lenient or strict, representing their idiosyncratic beliefs and sentencing preferences. Thus, lenient judges prefer to give lighter sentences, and strict judges prefer to give more severe sentences. The general public does not know the exact details of each case and observes only the sentence handed down by the judge. The public has some preferences over the type of judges, which is implied by its preference regarding sentences. In the model, we assume that the preferences of the public and of strict judges are aligned and therefore that the public prefers strict judges. Our conclusions, however, continue to hold under the opposite assumption.

We show in this paper that, in equilibrium, both lenient and strict judges' sentencing decisions are affected by their concern for reputation. As may be expected, given the assumption that the public prefers strict judges, both types of judges tend to deliver more severe sentences than they would have delivered absent reputation concerns. However, a judge's sentencing decisions depend on her prior sentencing history in a way that is more nuanced. As a judge's reputation for being lenient becomes stronger, she will choose more severe sentences prior to elections. On the other hand, if a judge's reputation for being strict is sufficiently strong, she will impose less severe sentences prior to elections, as she will be less concerned about convincing the public that she is not lenient. Thus, our analysis suggests that sentencing decisions will be negatively serially correlated. Interestingly, even if the public is assumed to have a preference against lenient judges, this does not imply that judges will render sentences that are as severe as allowed. Judges are expected to balance two types of preferences—their preference over the sentence

2. For a general model of bad reputation, see Ely and Valimaki (2003). As explained in Klement and Neeman (2013), however, a more comprehensive model of arbitrator reputation would also account for favorable reputation effects, which may induce biased arbitrators to deliver appropriate decisions in an effort to favorably impress future litigants and convince them that they are unbiased.

they would give in a particular case and their interest in getting reelected. As a judge's reputation for being strict becomes stronger, her sentencing behavior is expected to be less affected by her reputational concerns and more by her preference regarding the appropriate sentence in the case before her.

We test empirically the theoretical predictions produced by our model. To this end, we use data on sentencing decisions made by elected Pennsylvania Common Pleas Court judges between 2001 and 2010. We indeed find significant negative correlation when we interact a judge's prior decisions with her proximity to reelection. Furthermore, we find that this correlation is significant only for judges in the first 10 years of their judicial term. We do not find a correlation for judges after they have been reelected. In our reputation model terminology, judges who have already established a reputation for being strict are less concerned that any single decision will change how they are perceived by the public. Those judges therefore impose sentences according to their judicial preferences without trying to convince the public that they are not lenient. Our empirical findings are therefore consistent with the predictions of the dynamic reputation model that we develop.

The paper proceeds as follows. Section 2 presents a short review of the literature on the effect of judicial elections on judicial decisions. Section 3 presents the theoretic reputation model. Section 4 provides an empirical examination of the model, and Section 5 concludes.

2. RELATED LITERATURE

Election and appointment of judges may impact their decisions in two ways—indirectly, through selection, and directly, by affecting their incentives. Selection effects imply that judges' attitudes resemble those of their appointing body, be it the state's governor, an appointment committee, or the public at large. Both the initial appointment of judges and their retention or reappointment may serve to screen out judges whose attitudes do not conform sufficiently to the electing body's approach. Incentive effects imply that judges will decide cases in ways that increase their chances of being reappointed. If judges expect to be screened on the basis of their judicial decisions, they may consider this when rendering those decisions. The two effects are not easy to distinguish empirically.

Many studies have demonstrated the effect of reelection and reappointment on judicial decision making. Shepherd (2011) provides a con-

cise review of this literature. Three hypotheses emerge: The first is that judges' decisions conform to the preferences of their appointing agents (see, for example, Hall 1992, 1995; Brace and Hall 1997; Shepherd 2009a, 2009b, 2009c). The second is that the degree of such conformity depends on various institutional features of the method of reappointment (whether reappointment decisions are made by some gubernatorial or legislative body or through public elections) (see, for example, Brace and Boyea 2008; Boyea 2010; Choi, Gulati and Posner 2010; Lim 2013) and, if made through elections, whether these are competitive and whether they are held on a partisan or nonpartisan basis (see, for example, Helland and Tabarrock 2002; Gordon and Huber 2007; Caldarone, Canes-Wrone, and Clark 2009). These two hypotheses do not distinguish between selection and incentive effects. The third hypothesis does. It suggests that the conformity between the preferences of the judge and her appointing agent becomes more pronounced as reappointment or reelection approaches and that it becomes less pronounced when judges are serving their last term in office (see, for example, Brooks and Raphael 2002; Huber and Gordon 2004; Gordon and Huber 2007; Shepherd 2009a, 2009b, 2009c; Lim 2013).

Our analysis focuses on the incentive effect of judicial elections. To do so, we examine how a judge's sentencing decision depends on her history of previous sentencing decisions. We suggest that such dependence may be explained by a reputation model. Two papers that are most proximate in their theoretical approach to ours are Huber and Gordon (2007) and Lim (2013). Huber and Gordon (2007) present a multistage dynamic model in which legislators set sentencing boundaries, judges then impose their preferred sentences, and finally the electorate decides whether to retain the judge for another judicial term. Their model assumes that judges have private information regarding their type and the appropriate sentence. In equilibrium, judges deviate from their preferred sentence to increase their probability of getting reelected. These features are applied also in our model. However, unlike Huber and Gordon (2007), our model analyzes the effect of different prior public beliefs on judges' sentencing behavior. This allows us to develop subtler conjectures about the dependence between a judge's sentence and her prior sentencing history and to examine these conjectures empirically.

Lim (2013) presents a dynamic structural empirical model in which she estimates judges' preferences over sentencing against their benefits from being reelected. She uses sentencing data from Kansas that contain

within-state variation of two judicial selection methods, appointment and election. The structural model allows her to distinguish selection effects, which are more pronounced when judges are appointed by government officials, from incentive effects, which play a significant role when judges are elected. Her model does not analyze the dependence of present sentencing decisions on previous sentencing history.

Finally, there is some literature on Supreme Court justices' preference changing (or ideological drift) over time (for example, Epstein et al. 1998; Martin and Quinn 2002, 2007). This literature, however, should be distinguished from the current paper, which analyzes dynamic decision making, namely, the dependence of a judge's decision on her prior decisions.

3. A MODEL OF JUDICIAL REPUTATION

3.1. The Model

Consider the problem of a judge who has to sentence a convicted offender. For simplicity, we assume that the sentence can be either light or severe. The general public has some preference over the severity of a sentence in each case—and thus the appropriate sentence in this case[3]—depending on the details of the case.[4] However, the public is not informed about the details of each case, and therefore it knows that the appropriate sentence is severe only with an ex ante probability $q \in [0, 1]$, independently across all cases. Thus, q reflects the public's preference for severe sentences over all cases.

Judges observe the details of each case and therefore know the public's preferred sentence. There are two types of judges—strict and lenient. A strict judge's preferences are aligned with the public's preferences, and she therefore obtains a payoff of one if she renders the appropriate sentence (light or severe) and a payoff of zero otherwise. A lenient judge finds it hard to impose a severe sentence. We assume that she obtains a

---

3. The terminology is borrowed from the Pennsylvania Sentencing Guidelines: "The court shall consider the sentencing guidelines in determining the appropriate sentence for offenders convicted of, or pleading guilty or nolo contendere to, felonies and misdemeanors" (Pa. Code sec. 303.1).

4. Notice that the public is not assumed to always want the sentence to be severe. See Cullen, Fisher, and Applegate (2000).

payoff of one if she renders a light sentence and a payoff of zero otherwise.[5]

We assume that in addition to the payoff that a judge derives from the sentence she renders, she also cares about her chances of reappointment. We denote the value of being reappointed as $V > 1$ and assume that it is commonly believed that the likelihood of reappointment, denoted $f(\pi_S)$, is increasing and concave in the probability with which the judge is believed to be strict, denoted π_S. Furthermore, we assume that the value of reappointment V is sufficiently high to induce judges to sometimes render a decision that does not conform with their sentencing preferences when such a decision increases their chance of reappointment. Thus, the judge's long-term interests in reappointment may be sufficiently strong to overcome her short-term interests in rendering her preferred sentence, depending on her type.

We denote the prior probability with which the public believes the judge to be strict and lenient by π_S and $\pi_L = 1 - \pi_S$, respectively. This prior probability is determined by the judge's individual characteristics (such as race, gender, family situation, professional record prior to being elected to the judicial office, and so on) and by her history of past decisions. In general, the information gleaned from one additional decision depends not only on the prior probability but also on the length of history on which this prior probability is based. The information inferred from one additional decision will not lead to a posterior probability that is significantly different from the prior probability if the number of the judge's previous decisions is large but may lead to a significantly different posterior belief if the number of the judge's previous decisions is small. In other words, the prior belief about a judge who has made a large number of previous decisions can be thought of as strong and more resistant to change, while the prior belief about a judge who has made a small number of previous decisions can be thought of as weak and more susceptible to change. For simplicity, we abstract away from this consideration in the formal analysis below.

---

5. The assumption that the public's preferences are aligned with those of the strict judge is made only for expositional convenience and expresses no normative judgment as to what the sentence should be. Our results also hold if we assume the opposite—namely, that the lenient judge's preferences are aligned with the public's preferences, whereas the strict judge prefers to always render a severe judgment, except that instead of distorting their decisions upward, judges in this case would tend to distort their decisions downward.

3.2. Analysis

A judge has four pure strategies available: impose a severe sentence independent of the publicly preferred—appropriate—sentence, impose a light sentence independent of the appropriate sentence, impose an appropriate sentence, and impose an (always) inappropriate sentence. A judge also has mixed strategies available. A mixed strategy can be described by two probabilities, p_l and p_s, which denote the probability of imposing the appropriate sentence when a light and severe sentence is appropriate, respectively. A strategy of a strict judge is denoted by a pair (p_l^S, p_s^S), and a strategy of a lenient judge is denoted by a pair (p_l^L, p_s^L).

The next observation follows from our assumption about judges' preferences.

Claim 1.   In equilibrium, a lenient judge always renders a (weakly) lighter sentence than a strict judge, for every prior belief $\pi_S \in [0, 1]$.

Proof.   A judicial decision generates an immediate benefit and a future benefit that depends on its effect on the judge's chances of reappointment. The public's posterior belief about the judge's type depends on the sentence imposed by the judge but not on the judge's type, which is not known by the public. This implies that the future benefit from any decision is independent of the judge's true type. Since a lenient judge derives a larger immediate benefit from a lighter sentence, she will impose a (weakly) lighter sentence than will a strict judge.

Corollary 1.   The publicly preferred judicial behavior, in which judges always give the publicly preferred—appropriate—sentence regardless of their type, cannot be sustained as equilibrium behavior.

Proof.   If judges always give the appropriate sentence, then the public learns nothing about a judge's type from her decision. This means that judges' decisions do not affect their chances of being reappointed. Thus, a lenient judge who deviates and imposes a light sentence when a severe sentence is due is not sanctioned for her deviation.

Next we show that reputational considerations imply that simply following the judge's preferred sentence, given the appropriate sentence in a particular case, cannot be sustained in equilibrium either.

Claim 2.   For every prior belief $\pi_S \in (0, 1)$, if V is sufficiently high, then a lenient judge will sometimes render a severe sentence and a strict judge will sometimes render a severe sentence where the publicly preferred—appropriate—sentence is light.

Proof. Fix $\pi_S \in (0, 1)$. By claim 1, imposition of a light sentence implies a (weakly) lower posterior probability that the judge is strict. Therefore, if a strict judge cares enough about reappointment, she will prefer to render a severe sentence rather than an appropriate light sentence. Similarly, a lenient judge who cares enough about reappointment would prefer to render a severe sentence rather than a light sentence.

We describe a mixed-strategy equilibrium that gives rise to both good and bad reputational forces as described above.[6] For large values of the prior belief π_S that are greater than a threshold value $\bar{\pi}_S$, strict judges impose the publicly preferred—appropriate—sentence, and lenient judges always impose a light sentence. In this range, the concavity of f implies that the inference that the judge is more likely to be lenient after a light sentence has a smaller effect on the likelihood of being reappointed. For small values of the prior belief that fall below the threshold $\bar{\pi}_S$, the inference that the judge is more likely to be lenient after a light sentence has a large negative effect on the likelihood of being reelected. In this range, strict judges always impose a severe sentence when it is due (that is, $p_s^S = 1$), but when a light sentence is due, they impose a light sentence with probability $p_l^S < 1$; lenient judges always impose a light sentence when it is due (that is, $p_l^L = 1$), but when a severe sentence is due, they impose a severe sentence only with probability $p_s^L < 1$.

For this behavior to be a part of an equilibrium, strict and lenient judges have to be indifferent between rendering a severe and a light sentence when a light and a severe sentence is due, respectively. This means that the probabilities p_l^S and p_s^L have to satisfy the following equation:

$$1 + f(\pi_S')V = f(\pi_S'')V,$$

where

$$\pi_S' = \frac{\pi_S(1 - q)p_l^S}{\pi_S(1 - q)p_l^S + \pi_L[1 - q + q(1 - p_s^L)]} < \pi_S$$

6. Since each type of judge has four pure strategies available, there are 16 potential pure-strategy equilibrium combinations. Only two of these 16 combinations can be sustained as a pure-strategy equilibrium—for both types of judges to always impose a light sentence or for both of them to always impose a severe sentence. See claim 3 in the Appendix. Both equilibria are sustained by the public's (off-the-equilibrium) belief that any deviation indicates that the judge is lenient. We find these two pure-strategy equilibria to be less interesting because in practice judges' decisions do vary with the particulars of the cases that they are deciding. Furthermore, these two equilibria are inconsistent with the assumption that the probability of a judge's reappointment is increasing in π_S because in these equilibria the judge's sentence is independent of π_S. Moreover, the equilibrium in which judges always impose a light sentence fails a version of the intuitive criterion (Cho and Kreps 1987).

denotes the posterior probability that the judge is strict after a light sentence and

$$\pi_S'' = \frac{\pi_S[q + (1-q)(1-p_l^S)]}{\pi_S[q + (1-q)(1-p_l^S)] + \pi_L q p_s^L} > \pi_S > \pi_S'$$

denotes the posterior probability that the judge is strict after a severe sentence.[7]

It is possible to construct other mixed-strategy equilibria in which lenient judges mix both when a severe and a light sentence are due. If, however, a lenient judge prefers to mix when a severe sentence is due rather than when a light sentence is due, all such equilibria are eliminated. Indeed, imposition of this as an additional assumption—so that instead of deriving a payoff of 1 from rendering a light decision, a lenient judge would obtain a payoff of $1 + \varepsilon$ for some small $\varepsilon > 0$ from rendering a light decision when it is appropriate and a payoff of 1 when the appropriate sentence is severe—would eliminate these other equilibria.[8]

Proposition 1 summarizes the empirical implications of the mixed-strategy equilibrium described above.

Proposition 1. Ceteris paribus, judges exhibit a tendency to decide against their prior decisions. In particular,

i) A strict judge who has imposed a large number of severe sentences in the past and so has a strong reputation for being strict ($\pi_S > \bar{\pi}_S$) imposes the publicly preferred—appropriate—sentence, which on average implies that she imposes shorter sentences than she does after a large number of light sentences, in which case $\pi_S < \bar{\pi}_S$. A strict judge who has imposed a large number of light sentences in the past and so has a strong reputation for being lenient ($\pi_S < \bar{\pi}_S$) mixes and imposes a severe sentence when she

7. In general, the indifference conditions for a strict and a lenient judge would be different. Their identity here is an outcome of our assumptions about judges' payoffs.

8. The argument is as follows: Recall that in a mixed-strategy equilibrium, a judge must be indifferent between rendering a light and a severe sentence. Claim 1 implies that the posterior belief that a judge is lenient is smaller than or equal to the prior belief following the imposition of a light sentence, and that the posterior belief that a judge is strict is larger than or equal to the prior belief following the imposition of a severe sentence. It therefore follows that a strict judge cannot be indifferent between imposing a light sentence and a severe sentence when a severe sentence is due. Thus, it must be that in equilibrium a strict judge always renders a severe sentence when it is due, and she mixes only when a light sentence is due. In equilibrium, a lenient judge cannot mix when a light sentence is due because once we change the payoffs slightly, as explained in the text, it cannot be that both lenient and strict judges are indifferent between imposing a light and a severe sentence when a light sentence is due.

thinks it is the publicly preferred—appropriate—sentence and sometimes also when she thinks it is not. Thus, on average, she imposes longer sentences than she does after a large number of severe sentences.

ii) A lenient judge who has imposed a large number of severe sentences in the past and so has a strong reputation for being strict ($\pi_S > \bar{\pi}_S$) always imposes a light sentence, which implies that she imposes shorter sentences than she does after a large number of light sentences, in which case $\pi_S < \bar{\pi}_S$. A lenient judge who has imposed a large number of light sentences in the past and so has a strong reputation for being lenient ($\pi_S < \bar{\pi}_S$) mixes and sometimes imposes a severe sentence when such a sentence is appropriate according to the public's preferences. Thus, on average, she imposes longer sentences than she does after a large number of severe sentences in which she always imposes a light sentence.

iii) The tendency to decide against the prior decisions should become weaker as the value of reappointment becomes smaller.

Proof. The proof follows from the construction of the mixed-strategy equilibrium. A strict judge with prior $\pi_S > \bar{\pi}_S$ imposes the publicly preferred—appropriate—sentence, which is shorter, on average, than the average sentence that she imposes in equilibrium for $\pi_S < \bar{\pi}_S$. A lenient judge with prior $\pi_S > \bar{\pi}_S$ imposes a light sentence, which is shorter, on average, than the average sentence that she imposes in equilibrium for $\pi_S < \bar{\pi}_S$. And by claim 1, strict judges render more severe sentences than do lenient judges, so Bayesian updating implies that a judge who has imposed a large number of severe sentences in the past has a reputation for being strict, such that $\pi_S > \bar{\pi}_S$, and a judge who has imposed a large number of light sentences in the past has a reputation for being strict, such that $\pi_S < \bar{\pi}_S$.

The threshold $\bar{\pi}_S$ that determines the point at which judges decide according to their short-term interest is determined only by $1 + f(\pi_S')V = f(\pi_S'')V$. The concavity of the function f implies that the threshold is increasing in V. That is, the lower V is, the lower the weight that the judges puts on her likelihood of reelection, the greater the weight she puts on maximization of her short-term interests, and the lower the threshold prior π_S beyond which she ignores the effect of her decisions on her probability of reelection.

As mentioned above, for simplicity we did not distinguish between judges who have long histories of decisions, for whom the information conveyed by one additional decision is small, and judges who have short

histories of decisions, for whom the information conveyed by one additional decision is large. Intuitively, since the effect of any single sentence on a judge's reputation is decreasing the more previous decisions she has rendered, her tendency to decide against her prior decisions should also decrease in the number of sentences she renders. This is due to the fact that if a judge's decision has only a small effect on her likelihood of reelection, the threshold beyond which she decides according to her short-term interests is very low and approaching 0 if her decision has no effect on her probability of reelection. This implies that the tendency to decide against her prior decisions, which hinges on the difference in judges' behavior below and above the threshold, loses much of its power because the chance that it will be manifested becomes smaller.

4. EMPIRICAL EXAMINATION OF PENNSYLVANIA SENTENCING DECISIONS

4.1. Data

We use sentencing data from the Pennsylvania Commission on Sentencing. Our data set covers the 10 years between 2001 and 2010 and includes all felony and misdemeanor offenses in which an offender was convicted and then sentenced by the Pennsylvania Court of Common Pleas.[9] Criminal cases are assigned to the 60 judicial districts following Pennsylvania venue rules, according to the place where the crime was committed (Pa. R. Crim. Proc. 130[A]). Ordinarily, every sentence must be imposed within 90 days of conviction or the entry of a plea of guilty or nolo contendere (234 Pa. Code ch. 7, rule 704) and reported by the judge to the Commission on Sentencing.

The data include rich information about the characteristics of each case. In particular, they include information about the offender (age, gender, race, and past criminal history), the offense (severity of the offense, weapons enhancement, type of offense, and number of counts), the sentence imposed by the judge, and the name of the judge who imposed it.

Pennsylvania provides sentencing guidelines that judges must consider when sentencing a defendant for a felony or a misdemeanor offense.[10]

9. The courts of common pleas are the courts of general jurisdiction in Pennsylvania. As such, they have jurisdiction over all felony and misdemeanor offenses (see Pa. Const., art. V, sec. 5; 42 Pa. Cons. Stat. sec. 931).

10. Our data contain three different sets of sentencing guidelines: the fifth, sixth, and sixth revised editions. The sentencing guidelines that apply to an offense are the ones that were in force on the date the offense was committed. For the effective dates of the guide-

The guidelines assign two scores for each offense: an offense gravity score (OGS), which ranges from 1 to 14 (where 14 represents the greatest severity), and a prior record score (PRS), which is based on the offender's prior criminal record and is divided into eight categories.[11] Both scores are reported for each offense. Using the combination of these two scores, the guidelines provide a sentence recommendation.

Each case may include several counts (offenses). For each count, the judge must determine the minimum and the maximum sentences (42 Pa. Const. Stat. Ann. sec. 9756).[12] The sentencing guidelines set a range for the minimum sentence between a lower and an upper limit, both stated in number of months. The judge has discretion whether to impose a sentence within the guidelines range or not, subject to mandatory minimum and maximum statutory sentences, when those apply.[13] In every case in which the judge imposes a sentence outside the sentencing guidelines, she must provide a contemporaneous written statement of the reason or reasons for the deviation. The judge's discretion, however, goes beyond sentencing on each count, as she must decide whether to impose the sentence concurrently or consecutively with sentences for other counts in the same case. We exclude cases in which one of the offenses is subject to a mandatory life or death sentence, since the judge has no discretion in sentencing these offenses.

The sentencing guidelines categorize all offenses into five levels. Our analysis focuses on cases in which the maximum sentencing guidelines level (based on the OGS and the PRS) of all counts in a case is the highest, namely, level 5. We focus on these cases since they are the most

---

lines, see Pennsylvania Commission on Sentencing, Sentencing Guideline Effective Dates (http://pcs.la.psu.edu/guidelines/sentencing/sentencing-guidelines-and-implementation -manuals/sentencing-guidelines-effective-dates/view).

11. The categories are repeat violent offender, repeat felony 1 and felony 2 offender, and point-based categories of 0–5. We code repeat violent offender and repeat felony 1 and felony 2 offender as 6-point categories.

12. The minimum sentence may not be reduced through parole prior to its expiration.

13. The guideline sentence cannot exceed the longest legal minimum sentence, which is one-half the maximum allowed by law. When the guideline sentence exceeds such a statutory limit, then that limit is the longest sentence recommendation. The court has no authority to impose a sentence less than that required by the mandatory-minimum provision established in the statute. When the guideline range is lower than that required by a mandatory-sentencing statute, the mandatory-minimum requirement supersedes the sentence recommendation. When the sentence recommendation is higher than that required by a mandatory-sentencing statute, the court must consider the guideline sentence recommendation.

visible to the public and are more likely to be reported in the media and thus to affect the judge's reputation. Our data contain about 38,000 such cases (which account for about 5 percent of all cases). For these offenses, the average sentence is about 67 months as compared with 18.7, 5.5, 1.3, and .2 months for offenses categorized as level 4, 3, 2, and 1, respectively.

Information about the judges was obtained from the *Pennsylvania Manual*.[14] The *Manual* is published biennially for years in which municipal elections take place. For each judge, the *Manual* provides information about when a judge was initially elected and when she was reelected. Once elected, judges stand for retention election every 10 years (Pa. Const., art. V, sec. 15[a]). In such elections, the name of the judge is submitted to the electors without party designation, to determine only the question of whether she shall be retained in office. This vote takes place at the municipal election immediately preceding the expiration of the term of the judge. If a majority favors retention, the judge serves an additional 10-year term (Pa. Const., art. V, sec. 15[b]). Judges retire on the last day of the calendar year in which they reach the age of 70 years (Pa. Const., art. V, sec. 16[b]).

Using both the *Pennsylvania Manual* and the Pennsylvania Commission on Sentencing, we were able to identify information about election and retention for about 81 percent of the judges appearing in the Pennsylvania Commission on Sentencing who sentenced defendants who had an offense of level 5 (410 of 506 judges), which also accounts for 91 percent of all offenses (we managed to merge 34,527 of 37,886 level 5 offenses).

4.2. Variables

Our dependent variable is Judge Sentence, which is defined as the minimum sentence assigned to the offender by the judge in a specific case. Each case may include more than one count. The judge has discretion whether to impose the sentence on each count concurrently or consecutively with other sentences imposed in the same case.[15] Therefore, our dependent variable is the sum of all consecutive sentences and the maximum of all the concurrent sentences imposed on the same offender in the case. Since the data contain some cases in which offenders were sentenced

---

14. For volumes 116–121, see Pennsylvania Department of General Services, Publications: *The Pennsylvania Manual* (http://www.portal.state.pa.us/portal/server.pt/community/pa_manual/1294).

15. See, for example, *Commonwealth v. Gonzalez-DeJesus* (2010 Pa. Super. 62), which describes the trial court's range of discretion in determining whether sentences are imposed concurrently or consecutively.

for hundreds of years, and we want to avoid biases resulting from outliers, we winsorize our dependent variable at the 99th percentile.

To test whether judges are affected by their past sentencing history, we use the variable Diff_Judge_Dist. This variable is equal to the difference between the average sentences that the judge imposed the preceding year for all cases whose sentencing guidelines level is level 5 and the average sentences imposed in the previous year in the judge's district for all offenses whose sentencing guidelines level is level 5.

To test whether judges behave differently when they are close to reelection, we use the variable Prox_to_Elect to measure proximity to election. This variable is equal to 1 minus the number of days from the date of sentencing to the judge's nearest reelection, divided by 3,650. Thus, the variable values range from 0 and 1, where 1 stands for the most proximate elections to sentence.[16] We also define an interaction term Prox_to_Elect × Diff_Judge_Dist to test whether judges' current sentencing is affected by their previous year's sentences differently when they are close to their reelection.

Table 1 provides summary statistics for our dependent variable, our two main variables of interest, and all of the covariates used in our analysis. Our data contain about 31,000 offense-year observations (for which we have full information), of which 47 percent are observations for offenses that were handled by judges who had less than 10 years of experience by the time of sentencing. The offenses that are included in our database were handled by the judges for whom reelection was relevant, which means that they were younger than 70 (the retirement age) during the next reelections. Our dependent variable, Judge Sentence, ranges from 0 to 366 months, with a mean of 67 months. An offender is on average 31 years old; 93 percent of the crimes were committed by male offenders, and 40 percent were committed by white offenders. The average OGS (which is the maximum OGS among all counts per case) is 10, and the offender's PRS is on average 2. In 16 percent of the cases, a weapon was involved, and on average there were 3 counts per case. With respect to the disposition of the case prior to sentencing, 49 percent of the cases resulted in a negotiated guilty plea, 20 percent resulted in a negotiated guilty plea, and the rest (31 percent) resulted in conviction at trial. On average, in each district there are about 64 cases with a sentencing guidelines level of 5 and about six different judges. Also, on average each judge handled about 14 such cases each year. It should also be noted that the mean vote for judges who stand for reelection is 75 percent, with vari-

---

16. As defined by Huber and Gordon (2004).

Table 1. Summary Statistics

	All Judges (1)	Inexperienced Judges (2)	Experienced Judges (3)
Judge Sentence	67.228	64.090	69.994
	(70.594)	(68.649)	(72.155)
Prox_to_Elect	.495	.488	.501
	(.275)	(.245)	(.298)
Diff_Judge_Dist	.372	−1.834	2.135
	(27.964)	(30.149)	(25.53)
Offender Age	30.873	30.572	31.139
	(10.839)	(10.726)	(10.932)
Offender Male	.927	.928	.926
	(.260)	(.258)	(.262)
Offender White	.404	.376	.429
	(.491)	(.484)	(.495)
Maximum OGS	10.156	10.081	10.223
	(1.637)	(1.595)	(1.671)
Offender PRS	1.996	1.976	2.013
	(2.214)	(2.199)	(2.227)
Enhanced Weapon	.157	.142	.171
	(.365)	(.349)	(.377)
Counts per Case	3.140	2.957	3.302
	(4.946)	(3.693)	(5.826)
Cases Handled by Judge in Past Year	14.188	13.609	14.641
	(19.234)	(23.147)	(15.499)
Nonnegotiated	.198	.215	.183
	(.382)	(.397)	(.368)
Negotiated Guilty	.491	.472	.508
	(.500)	(.499)	(.499)
Cases in District	63.591	58.681	67.376
	(165.748)	(142.652)	(181.732)
Judges in District	5.708	5.870	5.582
	(7.340)	(7.531)	(7.201)
Sentencing Guidelines, 6th Edition	.441	.464	.421
	(.497)	(.499)	(.494)
Sentencing Guidelines, 6th Revised Edition	.115	.142	.091
	(.319)	(.349)	(.287)
N	31,040	14,543	16,497

Note. Inexperienced judges have less than 10 years of experience, and experienced judges have 10 or more years of experience. OGS = offense gravity score; PRS = prior record score.

ance equal to .073. In fact, only two judges in our sample failed to win a majority vote on reelection. As we show below, judges' sentencing decisions are nevertheless affected by retention elections.

4.3. Testing the Hypothesis

Our hypothesis suggests that we should expect inexperienced judges (defined as judges with less than 10 years of experience as judges) to be affected by the difference between the average sentences they imposed last year and the sentences imposed by their colleagues in the same district. We also expect the effect to become more pronounced as judges approach election. To test this hypothesis, we use the following three models:

$$\text{Judge Sentence} = \alpha_1 + \beta \text{Prox\_to\_Elect}_{j,t} + \gamma X_{i,t} + \delta_j + \lambda_t + \varepsilon_{i,j}, \quad (1)$$

$$\text{Judge Sentence} = \alpha_1 + \beta \text{Diff\_Judge\_Dist}_{j,t} + \lambda X_{i,t} + \delta_j + \lambda_t + \varepsilon_{i,j}, \quad (2)$$

and

$$\begin{aligned}\text{Judge Sentence} = \alpha_1 &+ \beta_1 \text{Prox\_to\_Elect}_{j,t} + \beta_2 \text{Diff\_Judge\_Dist}_{j,t} \\ &+ \beta_3 \text{Prox\_to\_Elect}_{j,t} \times \text{Diff\_Judge\_Dist}_{j,t} \\ &+ \gamma X_{i,t} + \delta_j + \lambda_t + \varepsilon_{i,j}, \end{aligned} \quad (3)$$

where i is an indicator for the case or offender, j for the judge, and t for the year of sentence. The matrix $X_{i,t}$ includes an offender's characteristics such as age, gender, an indicator if the offender is white, and PRS; the offense's characteristics, including the gravity of the offense, an indicator for whether a weapon was involved, the number of counts in each case, indicator variables for negotiated and nonnegotiated guilty pleas (where the baseline is conviction on trial); and two measures for the district: the number of judges in a district and the number of cases with a sentencing guidelines level of 5 that were handled in the past year in that district. We also generate indicators for the sentencing guidelines editions.

To control for a judge's ideology, we employ a judge fixed effect, which accounts for all judges' characteristics that do not change over time. Since judges do not move from one district to another, controlling for judge fixed effect also controls for unobserved heterogeneity in districts. We also control for year and for offense statutory-grade dummies. Our standard errors are clustered by judges.

The first three columns of Table 2 provide the results for our first model. The regression is performed first for all judges and then separately for judges with less than 10 years of experience and judges with 10 or more years of experience. We find the coefficient on Prox_to_Elect to be indistinguishable from 0 for all three groups of judges. Our findings sug-

Table 2. Association between Judges' Sentences and Proximity to Election

	Model 1			Model 2			Model 3		
	All Judges (1)	Inexperienced Judges (2)	Experienced Judges (3)	All Judges (4)	Inexperienced Judges (5)	Experienced Judges (6)	All Judges (7)	Inexperienced Judges (8)	Experienced Judges (9)
Prox_to_Elect	.913 (1.798)	−1.974 (4.564)	1.763 (2.064)				1.229 (1.910)	3.103 (6.992)	1.638 (2.021)
Diff_Judge_Dist				.016 (.024)	−.018 (.042)	−.007 (.024)	.062 (.047)	.140 (.091)	−.098 (.065)
Prox_to_Elect × Diff_Judge_Dist							−.105 (.076)	−.332* (.143)	.126 (.101)
Offender Age	.332+ (.184)	.283 (.276)	.417+ (.222)	.316+ (.188)	.127 (.319)	.470* (.224)	.218 (.203)	.133 (.323)	.302 (.240)
Offender Age²	−.006** (.002)	−.005 (.004)	−.006* (.003)	−.005* (.002)	−.003 (.004)	−.007* (.003)	−.004 (.003)	−.004 (.004)	−.005 (.003)
Offender Age Missing	10.830 (12.429)	−.546 (24.441)	12.500 (13.039)	1.400 (15.400)	−20.979 (33.038)	11.352 (16.786)	−1.217 (15.047)	−21.826 (32.339)	6.891 (15.695)
Offender Male	14.431** (1.221)	14.371** (1.628)	14.299** (1.678)	13.955** (1.200)	14.427** (1.721)	13.854** (1.603)	14.360** (1.288)	14.340** (1.733)	14.403** (1.800)
Offender White	−5.621** (.731)	−7.239** (1.211)	−4.668** (.916)	−5.622** (.757)	−7.230** (1.372)	−4.914** (.963)	−5.740** (.775)	−6.931** (1.367)	−5.108** (.972)
Maximum OGS	21.909** (.477)	21.633** (.620)	21.829** (.657)	22.671** (.463)	22.510** (.669)	22.572** (.604)	22.718** (.498)	22.582** (.659)	22.539** (.678)
Offender PRS	9.673** (.227)	9.785** (.290)	9.501** (.322)	9.752** (.230)	9.935** (.315)	9.566** (.308)	9.831** (.249)	9.918** (.322)	9.681** (.349)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Enhanced Weapon	33.285**	33.763**	31.975**	34.064**	34.827**	32.874**	34.621**	34.781**	33.554**
	(1.537)	(2.162)	(1.656)	(1.516)	(2.302)	(1.660)	(1.651)	(2.281)	(1.853)
Counts per Case	2.875**	4.401**	2.345**	2.832**	4.402**	2.387**	2.724**	4.373**	2.210**
	(.667)	(.676)	(.695)	(.668)	(.761)	(.694)	(.665)	(.759)	(.669)
Nonnegotiated	−13.051**	−12.002**	−14.092**	−13.521**	−13.922**	−13.461**	−13.637**	−13.881**	−13.480**
	(1.473)	(2.288)	(1.960)	(1.416)	(2.584)	(1.507)	(1.573)	(2.637)	(1.789)
Negotiated Guilty	−16.215**	−14.515**	−17.715**	−16.121**	−14.699**	−17.070**	−16.201**	−14.778**	−17.273**
	(1.373)	(1.937)	(1.797)	(1.291)	(2.065)	(1.538)	(1.374)	(2.095)	(1.683)
Cases in District	.002	−.002	−.001	.001	.001	−.002	.002	−.002	.004
	(.005)	(.008)	(.006)	(.004)	(.007)	(.005)	(.004)	(.007)	(.006)
Judges in District	−.371+	.025	−.393+	−.320+	.126	−.424+	−.250	.258	−.369
	(.211)	(.328)	(.237)	(.193)	(.342)	(.221)	(.201)	(.336)	(.234)
Sentencing Guidelines, 6th Edition	−7.930**	−8.810**	−6.836**	−7.934**	−7.818**	−7.676**	−7.559**	−8.009**	−6.825**
	(1.626)	(2.201)	(2.234)	(1.578)	(2.377)	(2.024)	(1.670)	(2.390)	(2.235)
Sentencing Guidelines, 6th Revised Edition	−13.374**	−11.578**	−15.177**	−13.521**	−10.843**	−15.379**	−13.204**	−11.117**	−15.200**
	(2.327)	(3.074)	(3.146)	(2.280)	(3.195)	(3.043)	(2.347)	(3.172)	(3.167)
Adjusted R^2	.4953	.5173	.4886	.5023	.5210	.5001	.5029	.5231	.5003
N	31,040	14,543	16,497	30,190	12,098	18,092	26,849	11,923	14,926

Note. Inexperienced judges have less than 10 years of experience, and experienced judges have 10 or more years of experience. Standard errors, clustered by judge, are in parentheses. All regressions control for judge, year, and offense statutory grade. OGS = offense gravity score; PRS = prior record score.

$^+$ $p < .1$.

* $p < .5$.

** $p < .01$.

gest that judges' current sentencing is not affected by proximity to election, regardless of their level of experience.

Columns 4–6 provide the results for our second model. As in the first model, the regression is first performed for all judges and then for inexperienced and experienced judges. We find the coefficient on Diff_Judge_Dist to be indistinguishable from 0 for all three groups of judges. Hence, we find no effect of a judge's sentencing in the past year on the current sentence that she imposes.

Columns 7–9 provide the results of our third model. Here we introduce an interaction term between Prox_to_Elect and Diff_Judge_Dist. Column 7 shows that the coefficients on Prox_to_Elect, Diff_Judge_Dist, and their interaction term are all indistinguishable from 0 when we perform the regression on all judges, regardless of their experience. However, for inexperienced judges, we see that although Prox_to_Elect and Diff_Judge_Dist are positive but insignificant, the interaction term is negative and statistically significant (with a coefficient of −.332, significant at the 5 percent level). This suggests that judges tend to impose sentences against their prior record, but they do so only when they approach retention elections. When we perform the regression separately for experienced judges, we do not find such an association.

This effect is not small. Table 3, column 1, shows the magnitude of the effect in months on sentencing, for inexperienced judges, a year before election (with Prox_to_Elect equal to .9) and with Diff_Judge_Dist values in the 10th, 25th, 75th, and 90th quantiles. Our prediction suggests that a year before standing for reelection, a judge with Diff_Judge_Dist in the 10th (25th) quantile would increase the length of incarceration by 9 (5) months. Similarly, judges with Diff_Judge_Dist in the 90th (75th) quantile would reduce the length of incarceration by 2 months (1 month).

Most of the other covariates have signs that are consistent with findings in previous literature. Male offenders receive longer sentences than female offenders, and white offenders receive shorter sentences than nonwhites. As expected, when the offense has a higher OGS or a higher PRS, the length of incarceration is longer. When a weapon is used, sentences are longer. Negotiated pleas and nonnegotiated pleas have a negative effect on sentencing. Guidelines editions have a negative effect on sentences (which suggests that the sixth edition is less stringent than the fifth edition), and the number of counts on each case has a positive and statistically significant effect. The number of judges in a district has a negative and significant effect on sentencing only for experienced judges. We find

Table 3. Size of the Effect for Inexperienced Judges

Percentile	Table 2, Column 8 (1)	Table 4, Column 2 (2)	Table 5, Column 2 (3)	Table 5, Column 5 (4)	Table 5, Column 8 (5)
10th: Diff_Judge_Dist = –40	9	9	9	14	18
25th: Diff_Judge_Dist = –16	5	3	4	7	9
75th: Diff_Judge_Dist = 12	1	–3	–2	–1	–2
90th: Diff_Judge_Dist = 12	–2	–7	–6	–5	–9

Note. Values are in months.

that an offender's age and the number of similar cases handled the previous year in the same district to have no statistically significant effect.

4.4. Discussion and Alternative Specifications of the Model

Table 4 provides the results of regressing Judge Sentence on Prox_to_Elect^2. This tests our expectation not only that the effect of proximity to election should become larger when approaching election but also that this effect is convex. Put differently, we expect any change in the effect of proximity to election to become more pronounced as the time for elections nears. The results show a pattern similar to the one we observed previously. Column 2 of Table 3 shows the magnitude of the effect for a judge 1 year before an election with Diff_Judge_Dist values in the 10th, 25th, 75th, and 90th quantiles (as in Table 2, column 8). We find that 1 year before reelection, a judge with Diff_Judge_Dist values in the 10th and 25th percentiles will increase sentence lengths by 9 and 3 months, respectively, and will reduce them by 3 and 7 months if Diff_Judge_Dist is in the 75th and 90th quantiles.

On average, judges in our data have about 14 cases with a sentencing guidelines level equal to 5. However, 6 percent of the judges decided less than one level 5 case in a year, and 14 percent decided less than three level 5 cases in a year. Therefore, Table 5 repeats Table 4, once for all cases in which judges handled more than one level 5 case in the previous year and once for all cases where judges handled more than three level 5 cases in the previous year. Columns 7–9 present the results when we weight our regression by the number of level 5 cases that judges handled in the previous year. We obtain patterns similar to those shown in Tables 2 and 5.

Columns 3–5 of Table 3 present the magnitude of the effect for inex-

Table 4. Alternative Specification for Proximity to Election

	All Judges (1)	Inexperienced Judges (2)	Experienced Judges (3)
Prox_ to_Elect^2	1.936	−.100	1.730
	(1.882)	(8.215)	(1.930)
Diff_Judge_Dist	.043	.073	−.066
	(.038)	(.066)	(.049)
Prox_to_Elect^2 × Diff_Judge_Dist	−.106	−.342*	.091
	(.078)	(.142)	(.098)
Adjusted R^2	.5030	.5231	.5002
N	26,849	11,923	14,926

Note. Inexperienced judges have less than 10 years of experience, and experienced judges have 10 or more years of experience. Standard errors, clustered by judge, are in parentheses. All regressions control for judge, year, and offense statutory grade; regressions also control for all other offender and offense covariates (not shown).
   * $p < .5$.

perienced judges. We find that the magnitude of the effects is a bit larger, with the largest effect obtained in the weighted regressions.

Finally, to the extent that prosecutors and defendants expect approaching judicial elections to affect the expected sentence, such expectations might affect prosecutors' decisions in filing charges with minimum sentences or their plea bargain decisions (for recent literature discussing the effects of prosecutorial discretion on sentencing, see Bjerk 2005; Ulmer, Kurlychek, and Kramer 2007; Starr and Rehavy 2013; Yang 2014). This might raise a concern about a selection bias effect. However, we find no evidence of such an effect in our data. In particular, focusing on judges with less than 10 years of tenure—the judges who we show to be most influenced by the timing of judicial elections—we find that the percentage of plea deals is .503 for cases that are in the bottom decile in terms of proximity to elections and .48 for cases that are in the top decile in terms of proximity to elections. Furthermore, in a regression of the likelihood of a plea bargain on all the characteristics included in the regression in Table 2, the coefficients on proximity to election and its interaction with Diff_Judge_Dist are not statistically significant.

Moreover, even if the election cycle does somewhat change the mix of cases with and without plea bargains, the presence of such an effect does not undermine our identification of the effect of the election cycle on judges' sentencing in a given case. We checked and verified that judges are facing a mix of cases with and without plea bargains throughout the

Table 5. Alternative Specification for Proximity to Election by the Number of Cases Handled

	More Than One Case in Past Year			More Than Three Cases in Past Year			Weighted by Number of Cases in Past Year		
	All Judges (1)	Inexperienced Judges (2)	Experienced Judges (3)	All Judges (4)	Inexperienced Judges (5)	Experienced Judges (6)	All Judges (7)	Inexperienced Judges (8)	Experienced Judges (9)
Prox_to_Elect^2	1.763	.604	1.183	1.800	3.535	1.179	.995	3.298	.158
	(1.938)	(8.637)	(2.016)	(1.937)	(9.204)	(2.017)	(2.152)	(17.211)	(2.311)
Diff_Judge_District	.039	.088	-.084	.081$^+$	.171$^+$	-.056	.142**	.272*	-.018
	(.042)	(.072)	(.052)	(.047)	(.092)	(.058)	(.053)	(.109)	(.075)
Prox_to_Elect^2 ×									
Diff_Judge_Dist	-.124	-.379*	.080	-.199*	-.553**	.036	-.347**	-.812**	-.144
	(.084)	(.150)	(.105)	(.096)	(.185)	(.118)	(.126)	(.280)	(.149)
Adjusted R^2	.5019	.5223	.4990	.5011	.5241	.4961	.5008	.5150	.4953
N	26,287	11,643	14,644	24,976	10,913	14,063	26,849	11,923	14,926

Note. Inexperienced judges have less than 10 years of experience, and experienced judges have 10 or more years of experience. Standard errors, clustered by judge, are in parentheses. All regressions control for judge, year, and offense statutory grade; regressions also control for all other offender and offense covariates (not shown).

$^+$ $p < .1$.
* $p < .5$.
** $p < .01$.

judicial election cycle. We also checked and verified that the same is true for cases with and without mandatory minimum sentencing. Furthermore, our analysis controls for various characteristics of cases tried, including whether there was a plea bargain. Therefore, our analysis shows that, for any given mix of cases, the judicial election cycle affects sentencing decisions. To be sure, our finding of an effect of the judicial election cycle on sentencing decisions suggests that it may be worthwhile for future work to examine how the judicial election cycle affects the choices of agents other than judges, such as prosecutors and defendants. Such future work may be useful for understanding how judicial election cycles affect sentencing outcomes.

5. CONCLUSION

This paper developed a dynamic reputation model of judicial decisions over the election cycle. This approach models a judge's reputation as a belief held by the public (or an appointing principal) concerning the judge's attitudes and preferences. The public's beliefs are updated on the basis of the judge's decisions and consequently affect the judge's prospects of reelection or reappointment. Since the judge cares about being reelected or reappointed, her decision in each case may be affected by her effort to influence the public's perceptions.

The dynamic reputation approach predicts a negative serial correlation between a judge's past decisions and the present decision she makes on any issue that is of importance to the public. Our empirical examination of Pennsylvania courts' criminal sentences demonstrates a negative correlation for less experienced judges as they approach retention elections. These findings are consistent with the predictions of the reputation model.

The reputation model casts some doubt over the value of past decisions as a predictor of future decisions.[17] Judges' past decisions do not necessarily predict how they will decide future cases. On the contrary, current decisions may exhibit negative correlation with judges' past decisions.

Because our findings are consistent with the view that approaching judicial elections influence judges' decisions, they are broadly consistent with the realistic approach to judging. Furthermore, our empirical evidence of

---

17. See the review in Brace, Langer, and Hall (2000, p. 390–91).

the correlation between a judge's current and past decisions identifies the dynamic incentive effects of judicial elections, as distinguished from their static selection effects. Our work suggests that analysis of dynamic effects is necessary for a complete picture of how judicial elections influence the decisions of judges.

APPENDIX: PURE-STRATEGY EQUILIBRIA

Claim 3. There are only two pure-strategy equilibria, one in which both types of judges always impose a light sentence and another in which both types of judges always impose a severe sentence.

Proof. Both of these equilibria are sustained by the public's belief that any deviation indicates that the judge is lenient. Claim 1 implies that all the combinations in which the lenient type sometimes imposes a more severe sentence than the strict type can be ruled out. All other separating equilibria can also be ruled out because they imply that a lenient type is identified as such with a positive probability. This implies that if V is large enough, then the lenient type will benefit from deviating and rendering the same decision that the strict type would render. Finally, for both types of judges to always impose an inappropriate decision cannot be an equilibrium for the same reason that for both judges to always render the appropriate decision cannot be an equilibrium. Namely, in such a case, the judges' decisions do not affect the public's belief about their type, and so both types of judges would benefit from deviating and rendering the decision they prefer.

REFERENCES

Bjerk, David. 2005. Making the Crime Fit the Penalty: The Role of Prosecutorial Discretion under Mandatory Minimum Sentencing. *Journal of Law and Economics* 48:591–625.

Boyea, Brent D. 2010. Does Seniority Matter? The Conditional Influence of State Methods of Judicial Retention. *Social Science Quarterly* 91:209–27.

Brace, Paul, and Brent D. Boyea. 2008. State Public Opinion, the Death Penalty, and the Practice of Electing Judges. *American Journal of Political Science* 52:360–72.

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* 59:1206–31.

Brace, Paul R., Laura Langer, and Melinda Gann Hall. 2000. Measuring the Preferences of State Supreme Court Judges. *Journal of Politics* 62:387–413.

Brooks, Richard R. W., and Steven Raphael. 2002. Life Term or Death Sentences:

The Uneasy Relationship between Judicial Elections and Capital Punishment. *Journal of Criminal Law and Criminology* 92:609–40.

Caldarone, Richard P., Brandice Canes-Wrone, and Tom S. Clark. 2009. Partisan Labels and Democratic Accountability: An Analysis of State Supreme Court Abortion Decisions. *Journal of Politics* 71:560–73.

Cho, In-Koo, and David M. Kreps. 1987. Signaling Games and Stable Equilibria. *Quarterly Journal of Economics* 102:179–221.

Choi, Stephen J., G. Mitu Gulati, and Eric A. Posner. 2010. Professionals or Politicians: The Uncertain Empirical Case for an Elected Rather than Appointed Judiciary. *Journal of Law, Economics, and Organization* 26:290–336.

Cullen, Francis T., Bonnie S. Fisher, and Brandon K. Applegate. 2000. Public Opinion about Punishment and Correction. *Crime and Justice* 27:1–79.

Ely, Jeffrey C., and Juuso Välimäki. 2003. Bad Reputation. *Quarterly Journal of Economics* 118:785–814.

Epstein, Lee, Valerie Hoekstra, Jeffrey A. Segal, and Harold J. Spaeth. 1998. Do Political Preferences Change? A Longitudinal Study of U.S. Supreme Court Justices. *Journal of Politics* 60:801–18.

Epstein, Lee, William M. Landes, and Richard A. Posner. 2013. *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice.* Cambridge, MA: Harvard University Press.

Gordon, Sanford C., and Gregory A. Huber. 2007. The Effect of Electoral Competitiveness on Incumbent Behavior. *Quarterly Journal of Political Science* 2:107–38.

Hall, Melinda G. 1992. Electoral Politics and Strategic Voting in State Supreme Courts. *Journal of Politics* 54:427–46.

———. 1995. Justices as Representatives: Elections and Judicial Politics in the American States. *American Politics Quarterly* 23:485–503.

Helland, Eric, and Alexander Tabarrok. 2002. The Effect of Electoral Institutions on Tort Awards. *American Law and Economics Review* 4:341–70.

Huber, Gregory A., and Sanford C. Gordon. 2004. Accountability and Coercion: Is Justice Blind When It Runs for Office? *American Journal of Political Science* 8:247–63.

———. 2007. Directing Retribution: On the Political Control of Lower Court Judges. *Journal of Law, Economics, and Organization* 23:386–20.

Klement, Alon, and Zvika Neeman. 2013. Does Information about Arbitrators' Win/Loss Ratio Improve Their Accuracy? *Journal of Legal Studies* 42:369–97.

Lim, Claire S. H. 2013. Preferences and Incentives of Appointed Public Officials: Evidence from State Trial Court Judges. *American Economic Review* 103:1360–97.

Mailath, George J., and Larry Samuelson. 2006. *Repeated Games and Reputations: Long-Run Relationships.* New York: Oxford University Press.

Martin, Andrew D., and Kevin M. Quinn. 2002. Dynamic Ideal Point Estimation

via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999. *Political Analysis* 10:134–53.

———. 2007. Assessing Preference Change on the U.S. Supreme Court. *Journal of Law, Economics, and Organization* 23:365–85.

Shepherd, Joanna M. 2009a. Are Appointed Judges Strategic Too? *Duke Law Journal* 58:1589–1626.

———. 2009b. The Influence of Retention Politics on Judges' Voting. *Journal of Legal Studies* 38:169–206.

———. 2009c. Money, Politics, and Impartial Justice. *Duke Law Journal* 58: 623–85.

———. 2011. Measuring Maximizing Judges: Empirical Legal Studies, Public Choice Theory, and Judicial Behavior. *University of Illinois Law Review*, pp. 1753–66.

Starr, Sonja B., and M. Marit Rehavi. 2013. Mandatory Sentencing and Racial Disparity: Assessing the Role of Prosecutors and the Effect of *Booker*. *Yale Law Journal* 123:2–80.

Ulmer, Jeffery T., Megan C. Kurlychek, and John H. Kramer. 2007. Prosecutorial Discretion and the Imposition of Mandatory Minimum Sentences. *Journal of Research in Crime and Delinquency* 44:427–58.

Yang, Crystal S. 2014. Have Inter-judge Sentencing Disparities Increased in an Advisory Guidelines Regime? Evidence from *Booker*. *New York University Law Review* 89:1268–1342.