Judicial Decision Making: A Dynamic Reputation Approach*

Alma Cohen† Alon Klement‡ Zvika Neeman§

November 27, 2013

Abstract

We present a theoretical model that suggests that strategic judges who are concerned about their reputation would tend to "decide against their prior." Namely, judges who imposed a large number of severe sentences in the past and so are believed to be objective would impose less severe sentences on average, and judges who imposed a large number of light sentences in the past and so are suspected of being lenient would impose more severe sentences on average. Empirical analysis of sentencing data from the Pennsylvania Commission on Sentencing (PCS) between the years 2001 and 2010 is consistent with our theoretical predictions.

1 Introduction

In their book, *The Behavior of Federal Judges*, Epstein, Landes and Posner (henceforth, ELP) contrast two alternative theories of judicial behavior: One, the legalistic theory, assumes that judges decide cases in accordance with orthodox norms of judicial decision-making. The other, the realistic theory, conjectures that judges, like other agents, decide cases according to their preferences and incentives, subject to various constraints - physical, ethical, institutional and others. 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 judicial behavior.

Papers that belong to the 'realistic theory' literature often proceed by hypothesizing an external, non-legalistic, factor that may affect judges' decisions, and by examining whether this hypothesis can be supported by empirical findings. They typically try to correlate judges' decisions with some non-legalistic factors. Finding of such a correlation lends support to the realistic theory of adjudication, and refutes the purely legalistic approach.

In this paper we suggest a dynamic approach to the analysis of judicial decision making. Instead of aggregating the decisions of each judge and examining various statistics over the aggregate set of decisions, we propose instead to examine a judge's individual decisions as a function of her past decisions. This approach, we suggest, promises valuable theoretical and empirical insights.

*We thank Zvika Messing and Roni Shani for excellent research assistance. Other acknowledgements to be added.

†Berglas School of Economics, Tel Aviv University, Israel; NBER and Harvard Law School

‡Radzyner School of Law, Interdisciplinary Center Herzliya, Israel

§Berglas School of Economics, Tel Aviv University, Israel

Our approach is based on a *dynamic* reputation model of judicial behavior, as contrasted with the *static*, labor market model, that is proposed by ELP. Reputation models assume that an agent (in this case a judge) holds private information about some of his 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 over the agent's characteristics is updated based on the agent's observed actions.

The agent takes this into consideration when choosing his actions – as he tries to distinguish himself from other types of agents, which are undesirable for the principal. The agent has long term incentives based on his expected repeated interaction with the principal. He may sacrifice his short term payoff, only to establish a reputation that would increase his long term expected payoff. In this setting, agents' actions and principals' beliefs are determined in an equilibrium of the dynamic game between them. Since decisions are repeatedly made over time, the agent's present decision depends, among other factors, on her 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 offered to arbitrate a specific case, try to establish a reputation for being neutral and unbiased. To establish such a reputation, arbitrators may deliver decisions they know to be incorrect. We examined this reputation effect under alternative arbitrator selection regimes. In particular, we identified a possible conflict between the positive *selection* effect that is induced by allowing litigants to veto arbitrators based on information about their past Win/Loss ratios, and the negative *incentive* effect that is induced by such selection. Whereas selection allows screening of biased arbitrators, incentive effects may cause unbiased arbitrators to deliver incorrect decisions, only to avoid a bad reputation for being biased. [2]

We suggest that similar reputational considerations may implicate judicial decision making, especially in those circumstances where 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 either impose a light or a severe sentence. The judge knows what the correct sentence should be in each specific case, given its factual and legal circumstances.[3] Judges are either *lenient*, or *objective*. Lenient judges tend to give lighter sentences, because of their idiosyncratic beliefs and preferences. The general public does not know what is the 'correct' sentence, and it only observes the actual sentence handed by the judge. The public prefers *objective* judges over *lenient* ones. This preference is either due to the public's preference for harsher sentences, or its preference for correct ones.

As we show, in equilibrium, both *lenient* and *objective* judges' sentencing decisions are affected by their concern for reputation. Both types of judges tend to deliver more severe sentences than they would have delivered otherwise: Lenient judges deliver more correct severe sentences, compared to the (incorrect) lenient sentences they would have delivered absent reputation concerns; Objec-

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

[2] As we explained, a more comprehensive model of arbitrator reputation would also account for favorable reputation effects, which may induce biased arbitrators to deliver correct decisions in an effort to favorably impress future litigants and convince them they are unbiased.

[3] A sentence may be either correct from a 'legalistic perspective', or correct from the public's perspective, had it known all the relevant facts and considerations in the case.

tive judges,on the other hand, are driven by reputation concerns to deliver more incorrect severe sentences, to distinguish themselves as objective.

Nevertheless, a judge's reputation is affected by her history of sentencing decisions. As a judge's reputation for being lenient becomes stronger, she would give more severe sentences. On the other hand, if a judge's reputation for being objective is sufficiently strong, she would impose less severe sentences, as she would be less concerned about 'convincing' the public that she is not leneint. Thus, we expect to find that sentencing decisions would be negatively serially correlated. Importantly, although the public is assumed to have a preference against lenient judges, this does not imply that judges would render as high sentences as they can. Judges are expected to 'balance' two types of preferences - their preference over the sentence they should give in a specific case, and their interest in getting reelected. As a judge's reputation for being objective becomes stronger, her sentencing behavior is expected to be less affected by her reputational concerns and more by her conviction about the appropriate sentence in the specific case before her.

We test these theoretical conjectures empirically, by examining sentencing decisions 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 her retention elections. Furthermore, we find that this correlation is significant only for judges in the first ten years of their judicial term. After being reelected, judges do not feature any such correlation. In our reputation model terminology, judges who have already established reputation for being objective, are less concerned that any single decision would change the public's belief over them. Those judges therefore impose sentences according to their true judicial convictions, without trying to convince ' the public' that they are not lenient. These findings are therefore consistent with the predictions of the dynamic reputation model.

From a normative perspective, our analysis features two opposing potential effects of judicial *accountability*[4] - welfare enhancing or 'good' judicial reputational effect and welfare reducing or 'bad' judicial reputational effect.[5] In both cases, building and sustaining a reputation requires the judge to manipulate the beliefs of the public about her type. In the case of 'good reputation' the actions taken by the judge are welfare enhancing, as would be the case when a lenient judge makes a correct severe decision in order to foster her reputation for being objective. In the case of 'bad reputation' the actions taken by the judge are welfare reducing, for example when an objective judge renders an *incorrect* severe decision in order to foster her reputation for being objective.[6]

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

[4]For a discussion of the effect of judicial election on judicial accountability see Saphire and Moke, 2008.

[5]For the formal model of *Bad Reputation* see Ely and Valimaki, 2003.

[6]As we show, an objective judge is subject to both good and bad reputational forces. In contrast, a lenient judge is only subject to good reputational forces.

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 governor, an appointment committee, or the public at large. Both initial appointment of judges, and their retention or reappointment may serve to screen out judges whose attitudes do not sufficiently conform with the electing body's approach. *Incentive* effects imply that judges would decide cases in ways that increase their chances to be reappointed. If judges expect to be screened based on 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. Shephard, 2011, provides a concise review of this literature. Three hypotheses emerge from the literature: First, that judges' decisions conform with the preferences of their appointing agents (see, e.g., Hall 1992 and 1995; Brace and Hall 1997; and Shephard 2009, 2009a, 2009b). Second, that the degree of such conformity depends on various institutional features of the method of reappointment (whehter reappointment decisions are made by some gubernatorial or legislative body or through public elections, e.g. 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 non-partisan basis, e.g. Helland and Tabarrock, 2002; Gordon and Huber 2007; Caldarone et. al. 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 e.g. Brooks and Raphael, 2002; Huber and Gordon, 2004; Gordon and Huber, 2007; Shephard 2009, 2009a, 2009b; 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 sentening 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, 2007a, and Lim, 2013. Huber and Gordon, 2007a presents a multi-stage dynamic model in which legislators set sentencing boundaries first, then judges impose their preferred sentences, and finally, the elctorate 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 in order to increase their probability of getting reelected. These features are applied also in our model. However, unlike Huber and Gordon, our model analyzes the effect of different prior public beliefs on judges' sentencing behavior. This allows us to postulate about the dependence between a judge's sentence and her prior sentencing history, and to examine this postulate 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, which has within state variation of judicial selection methods, that are either by appointment or by 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 significant role when judges are elected. Her model, too, does not analyze the dependence of present sentencing decisions on previous sentencing history.

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 judge is believed by the public to be either *lenient* or *objective*. An objective judge obtains a payoff of one if she renders the correct sentence (light or severe) and a payoff of zero otherwise. A lenient judge finds it hard to impose a severe sentence. We assume she obtains a payoff of one if she renders a light sentence and a payoff of zero otherwise.[7] We assume that the correct sentence is severe with an ex-ante probability $q \in [0,1]$, independently across different cases. Judges can observe the correct sentence; the public cannot.

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 by $V > 1$ and assume that it is commonly believed that the likelihood of reappointment, denoted $f(\pi_O)$, is increasing and concave in the probability that the judge is objective, denoted π_O.[8] Furthermore, we assume that the value of reappointment V is sufficiently high to induce judges to sometimes render an incorrect (or correct, in the case of a lenient judge) decision when it increases their chance of reappointment. Thus, the long term interests of a judge in reappointment may be sufficiently strong to overcome her short term interests in rendering a correct or a light sentence, depending on her type.

We denote the prior probability with which the public believes the judge to be objective and lenient by π_O, $\pi_L = 1 - \pi_O$, respectively. This prior probability is determined by the judge's individual characteristics (such as race, gender, family situation, her professional record prior to being elected to the judicial office, etc.) 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 would 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.

[7] We assume only two types of judges to keep the analysis simple. Adding another, *harsh*, type of judge (like Huber and Gordon , 2007a) would not change our qualitative results, but would render the analysis less tractable.

[8] We discuss these assumptions in Section 3.3 below.

For simplicity, we abstract away from this consideration in the formal analysis below.

3.2 Analysis

A judge has four pure strategies available: impose a severe sentence independent of the correct sentence, impose a light sentence independent of the correct sentence, impose a correct sentence, and impose an (always) incorrect 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 correct sentence when a light and severe sentence is appropriate, respectively. A strategy of an objective judge is denoted by a pair $\left(p_l^O, p_s^O\right)$, and a strategy of a lenient judge is denoted by a pair $\left(p_l^L, p_s^L\right)$.

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

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

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 type. Since a lenient judge derives a larger *immediate* benefit from a lighter sentence, she will impose a (weakly) lighter sentence than an objective judge. ■

Corollary. For judges to always decide correctly, regardless of their type, cannot be sustained as equilibrium behavior.

Proof. If judges always decide correctly, then the public learns nothing about a judge's type from the judge's 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 his deviation. ■

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

Claim 2. For every prior belief, $\pi_O \in (0,1)$, if V is sufficiently high then a lenient judge would sometimes render a severe sentence and an objective judge would sometimes render an incorrect severe sentence.

Proof. Fix $\pi_O \in (0,1)$. By claim 1, imposition of a light sentence implies a (weakly) lower posterior probability that the judge is objective. Therefore, if an objective judge cares enough about reappointment, she would prefer to give an incorrect severe sentence than a correct light sentence.

Similarly, a lenient judge that cares enough about reappointment would prefer to render a severe sentence than a light sentence. ■

We describe a mixed strategy equilibrium that gives rise to both good and bad reputational forces as described above.[9] For large values of the prior belief π_O, greater than a threshold value $\overline{\pi}_O$, objective judges decide correctly 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 π_O, that fall below the threshold $\overline{\pi}_O$, 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, objective judges always correctly impose a severe sentence when it is due (that is, $p_s^O = 1$) but when a light sentence is due, they impose a light sentence with probability $p_l^O < 1$; lenient judges always correctly 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 indeed a part of an equilibrium,objective and lenient judges have to be indifferent between rendering a severe and light sentence when a light and a severe sentence is due, respectively. This means that the probabilities p_l^O and p_s^L have to satisfy the following equation:

$$1 + f\left(\pi_O'\right) V = f\left(\pi_O''\right) V$$

where $\pi_O' = \frac{\pi_O(1-q)p_l^O}{\pi_O(1-q)p_l^O + \pi_L(1-q+q(1-p_s^L))} < \pi_O$ denotes the posterior probability that the judge is objective after a light sentence, and $\pi_O'' = \frac{\pi_O\left(q+(1-q)\left(1-p_l^O\right)\right)}{\pi_O\left(q+(1-q)\left(1-p_l^O\right)\right)+\pi_L q p_s^L} > \pi_O > \pi_O'$ denotes the posterior probability that the judge is objective after a severe sentence.[10]

It is possible to construct other mixed strategy equilibria in which competent 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 $1 + \epsilon$ for some small $\epsilon > 0$ from rendering a light decision when it is correct and a payoff of 1 when a light decision is incorrect, would eliminate these other equilibria.[11]

[9]Since each type of judge has four pure strategies available, there are sixteen different potential pure strategy equilibrium combinations. Only two out of these sixteen 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. 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 less interesting because in practice judges' decisions do vary with the particulars of the cases they are deciding. Furthermore, these two equilibria are inconsistent with the assumption that the probability of a judge's reappointment is increasing in π_N because in these equilibria the judge's sentence is independent of π_N.Moreover, the equilibrium where judges always impose a light sentence fails a version of the intuitive criterion (Cho and Kreps, 1987).

[10]Generally, the indifference conditions for an objective and a lenient judge would be different. Their identity here is an outcome of our assumptions about judges' payoffs.

[11]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 a light sentence, and that the posterior belief that a judge is objective is larger than or equal

The next proposition summarizes the empirical implications of the mixed strategy equilibrium described above.

Proposition. Ceteris Paribus, judges exhibit a tendency to decide "against their prior." Specifically,

1. An objective judge who has imposed a large number of severe sentences in the past and so has a strong reputation for being objective (i.e., $\pi_O > \overline{\pi}_O$) decides correctly, which on average implies she imposes shorter sentences than she does after a large number of light sentences in which case $\pi_O < \overline{\pi}_O$. An objective judge who has imposed a large number of light sentences in the past and so has a strong reputation for being lenient (i.e.,$\pi_O < \overline{\pi}_O$) mixes and imposes a severe sentence when it is due and sometimes also when it is not due. Thus, on average she imposes longer sentences than she does after a large number of severe sentences, where she decides correctly.

2. A lenient judge who has imposed a large number of severe sentences in the past and so has a strong reputation for being objective (i.e., $\pi_O > \overline{\pi}_O$) always imposes a light sentence, which implies she imposes shorter sentences than she does after a large number of light sentences in which case $\pi_O < \overline{\pi}_O$. A lenient judge who has imposed a large number of light sentences in the past and so has a strong reputation for being lenient (i.e., $\pi_O < \overline{\pi}_O$) mixes and sometimes imposes a severe sentence when it is due. Thus, on average she imposes longer sentences than she does after a large number of severe sentences where she always imposes a light sentence.

3. The tendency to decide "against the prior" should decrease, the lower is the value of reappointment.

Proof. The proof follows from the construction of the mixed strategy equilibrium. An objective judge with prior $\pi_O > \overline{\pi}_O$ imposes a correct sentence, which is shorter, on average, than the average sentence she imposes in equilibrium for $\pi_O < \overline{\pi}_O$. A lenient judge with prior $\pi_O > \overline{\pi}_O$ imposes a light sentence, which is shorter, on average, than the average sentence she imposes in equilibrium for $\pi_O < \overline{\pi}_O$. And, by Claim 1, objective judges render more severe sentences than 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 objective, or such that $\pi_O > \overline{\pi}_O$, and a judge who has imposed a large number of light sentences in the past has a reputation for being objective, or such that $\pi_O < \overline{\pi}_O$.

The threshold $\overline{\pi}_O$ that determines the point at which judges decide according to their short term interest only is determined by the equation $1 + f(\pi'_O) V = f(\pi''_O) V$. The concavity of the function

to the prior belief following a severe sentence. It therefore follows that an objective 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 an objective 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 main text it cannot be that both lenient and objective judges are indifferent between imposing a light and a severe sentence when a light sentence is due.

f implies that the threshold is increasing in V. That is, the lower is V, the lower is the weight that the judges puts on her likelihood of reelection, the bigger the weight she puts on maximization of her short term interests, and the lower is the threshold prior 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 and for whom the information conveyed by one additional decision is small, and judges who have short histories of decisions and for whom the information conveyed by one additional decision is large. Intuitively, since the effect of any single sentence on the reputation of judges is decreasing the more previous decisions a judge has made, a judge's tendency to decide "against the prior" 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 zero if her decision has no effect on the probability of reelection. This implies that the tendency to decide against the prior, which hinges on the difference in judges' behavior below and above the threshold loses much of its power because the chance that it would be manifested becomes smaller.

3.3 Are Reputation Effects Welfare Enhancing?

The mixed equilibrium implies that reputation considerations affect both the probability that a judge renders a severe sentence and the probability that the sentence she imposes is the 'correct' one. When judges are concerned about reelection then they impose larger sentences on average. But, are these sentences more 'correct' on average? On the one hand, lenient (and thereofre 'biased') judges, deliver more correct sentences due to their reputational considerations, as they sometime impose harh sentences when those are due. On the other hand, objective judges are sometimes induced to impose an incorrect harsh sentence, to facilitate their reputation as objective. There is no way to tell whether the overall probability of a correct sentence is larger when judges are concerned about reelection, as compared to when they don't have such considerations.

Furthermore, the analysis above assumed that the likelihood of a judge's reappointment is increasing and concave in the public's belief that the judge is objective. This assumption implies that the public cares about the probability that the judge is objective, and that it cares about it more the higher the probability that the judge is lenient. This may reflect the public's preference either for *severe* sentences, or for *correct* ones. Notice, though, that these probabilities are not necessarily increasing in the probability that the judge is objective.

Below, we explain the reason that the probability that a judge renders a correct decision is not necessarily increasing in the probability that the judge is perceived to be objective. The explanation for the reason that the probability that a judge renders a severe decision is also not necessarily increasing in the probability that the judge is perceived to be objective is similar, and so omitted.

The probability of a correct decision in this equilibrium is

$$\pi_O \left(q + (1-q)\, p_l^N \right) + (1-\pi_O) \left(1 - q + q p_s^L \right).$$

For $\pi_O > \overline{\pi}_O$ the probability of a correct decision is equal to $\pi_O + (1 - \pi_O)(1 - q) = 1 - q(1 - \pi_O)$, which is increasing in π_O.

However, for $\pi_O \leq \overline{\pi}_O$ the probability of a correct decision is equal to

$$\pi_O \left(q + (1 - q) p_l^N (\pi_O) \right) + (1 - \pi_O) \left(1 - q + q p_s^L (\pi_O) \right).$$

This expression is not necessarily increasing in π_N. First, it is not clear which type of judge decides correctly with a higher ex-ante probability, so it is not clear that a higher π_O will necessarily lead to a higher probability of a correct decision.[12] Second, the probability that an objective judge renders a correct decision is increasing in π_O, but the probability that a lenient judge renders a correct decision is decreasing in π_O. In order for the overall probability of a correct decision to increase in π_O, the probability of a correct decision by an objective judge has to rise sufficiently fast to compensate for the decrease in the probability of a correct decision by a lenient judge. Unfortunately, there is no easy way to verify that this is indeed the case generally. It is easy enough to produce examples where this is the case, but a general proof is difficult and probably requires additional restrictions on the model.

4 Empirical Examination of Pennsylvania Sentencing Decisions

4.1 Data

We use sentencing data from the Pennsylvania Commission on Sentencing (PCS). Our dataset covers the ten-year period between 2001 and 2010, and it includes all felony and misdemeanor offenses in which the offender was convicted and then sentenced by the Pennsylvania Court of Common Pleas.[13] Criminal cases are assigned among the 60 judicial districts following Pennsylvania venue rules, according to the place where the crime occurred.[14] Ordinarily, every sentence must be imposed within 90 days of conviction or the entry of a plea of guilty or nolo contendere,[15] and reported by the judge to the Commission on Sentencing.

The data includes rich information about the characteristics of each case. In particular, it includes information about the offender (age, gender, race, past criminal history), the offense (severity of the offense, weapons enhancement, type of offense, 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 felony or a misdemeanor offense.[16] The guidelines assign two scores for each offense: (1) an offense grav-

[12] Even if $p_l^N (\pi_O)$ and $p_s^L (\pi_O)$ are both independent of π_O, the probability of a correct decision would decrease in π_O if $q + (1 - q) p_l^N (\pi_O) \ll 1 - q + q p_s^L (\pi_O)$ as would be the case if q is near zero.

[13] 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, § 5, 42 Pa.C.S. § 931.

[14] Pennsylvania Rules of Criminal Procedure, Rule 130(A).

[15] 234 Pa. Code Ch. 7, Rule 704.

[16] Our data contains three different sets of sentencing guidelines: the 5th, the 6th and the 6th revised guideline editions. The sentencing guidelines that apply to each offense are the ones that were in force on the date the offense was committed. Effective dates of the guidelines can be found on http://pcs.la.psu.edu/guidelines/sentencing/sentencing-guidelines-

ity score (OGS), which ranges between 1 and 14 (where 14 represents the highest severity); and (2) a prior record score (PRS), which is based on the offender's prior criminal record and divided into eight reported categories.[17] Both scores are reported for each offense. Based on 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 sentence.[18] The sentencing recommendation guideline sets a range for the *minimum* sentence between a lower and an upper limit, both stated in months of incarceration. 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.[19] In every case where the judge imposes a sentence outside the sentencing guidelines, she must provide a contemporaneous written statement of the reason or reasons for the deviation from the guidelines. 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 max sentencing guideline level (based on offense gravity score and prior record score) among all counts of a specific case, is the highest, namely level five. We focus on these cases since they are the most visible to the public, and are more likely to earn some report in the media. Our data contains about 38,000 such cases (which accounts for for about 5% of the total cases). For these offenses the average sentencing is about 67 months as compared to 18.7, 5.5, 1.3 and 0.2 months for offenses with level four, three, two and one respectively.

Information on the judges was obtained from the Pennsylvania Manual.[20] The Manual is published biennially, for years in which municipal elections take place. For each judge, the Manual provides information on when the judge was initially elected, and when he was re-elected. Once elected, judges stand for retention election every 10 years.[21] In such elections, the judge's name is submitted to the electors without party designation, at the municipal election immediately preceding the expiration of the term of the judge, to determine only the question whether he shall be retained in office. If a majority favors retention, the judge serves an additional ten-year term.[22]

and-implementation-manuals/sentencing-guideline-effective-dates.

[17]The categories are Repeat Violent Offender [REVOC], Repeat Felony 1 and Felony 2 Offender [RFEL], and point-based categories of 0, 1, 2, 3, 4 and 5. We code REVOC and RFEL as 6 point category.

[18]See 42 Pa.C.S.A. § 9756. The minimum sentence may not be reduced through parole prior to its expiration.

[19]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 statutory limit then that limit is the longest guideline sentence recommendation. The court has no authority to impose a sentence lower than that required by a mandatory minimum provision established in 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 shall consider the guideline sentence recommendation.

[20]The Pennsylvania Manual is published by the Pennsylvania Bureau of Publications. Electronic version of vol. 116-121 can be found on: http://www.portal.state.pa.us/portal/server.pt/community/pa_manual/1294.

[21]PA constitution, Article V, s. 15(a).

[22]PA constitution, Article V, s. 15(b).

Judges retire on the last day of the calendar year in which they attain the age of 70.[23]

Using both the Pennsylvania Manual and the PCS, we were able to identify judge election and retention information for about 81% of the judges appearing in the PCS with offenese of level 5 (410 judges out of 506), which also accounts for 91% of all offenses (we managed to merge 34,527 out of 37,886 level five offenses).

4.2 Variables

Our dependent variable is *Judge Sentence*, which is defined to be 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.[24] Therefore, our dependent variable is the sum of all consecutive sentences with the max of all the concurrent sentences imposed on the same offender in the case. Since the data contains some cases in which offenders were sentenced for hundreds of years, and we want to avoid biases resulting from outliers, we winsorize our dependent variable at the ninety-ninth percentile.

To test whether judges are affected by their past sentencing history, we define the variable *Diff_Judge_Dist*. This variable is equal to the difference between (i) the average sentences that the judge has imposed in the preceding year for all cases whose sentencing guideline level equals five and (ii) the average sentences imposed in the previous year in the judge's district for all offenses whose sentencing guideline level equals five.

To test whether judges behave differently when they are close to re-election, we define a variable labeled *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 re-election, divided by 3650. Thus, the variable values range from 0 and 1, where 1 stands for most proximate elections to sentence.[25] We also define an interaction term between the *Prox_to_Elect* variable and the *Diff_Judge_Dist* variable to test whether judges' current sentencing is affected by their previous year's sentences in a different way when they are close to their re-election.

Table 1 provides summary statistics of our dependent variable, our two main variables of interest and all of the covariates used in our analysis. Our data contains about 31,000 offense-year observations (for which we have full information) out of which 47% are observations for offenses that were dealt by judges whose experience by the time of sentencing is less than 10 years. The offenses that are included in our database were treated by the judges for which reelection was relevant, meaning that their age in the next reelections was younger than retirement age (which is 70). Our dependent variable, *Judge Sentence*, ranges between 0 and 366 with a mean of 67 months. An offender's age is on average 31 years old, 93% of the crimes were done by male offender, 40% by white offenders. The average offense gravity score (which is the max OGS among all counts per case) is

[23] PA constitution, Article V, s. 16(b).

[24] See, for example, Commonweath v. Gonzalez DeJusus, 2010 PA Super 62, describing the range of discretion held by the trial court in determining whether sentences will be imposed concurrently or consecutively.

[25] As defined by Huber and Gordon, 2004.

10, and the offender's prior record score is on average 2. In 16% of the cases an enhanced weapon was involved, and on average there were 3 counts per case. With respect to the disposition of the case prior to sentencing, 49% of the cases result in a non-negotiated guilty plea, 18% in negotiated guilty plea and the rest 33% in conviction on trial. On average, in each district there are about 64 cases with sentencing guideline level equal five, and about 6 different judges. Also on average each judge gets to handle about 14 such cases each year. It should also be noted that the mean vote for judges who stand for reelection is 75%, with variance equal to 0.073. In fact, only two judges in our sample have failed to win a majority vote on reelection. As we show below, judges are nevertheless affected by retention elections in their sentencing decisions.

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 run the following three models:

$$(1)\ JudgeSentence = \alpha_1 + \beta ProxtoElect_{j,t} + \gamma X_{i,t} + \delta_j + \lambda_t + \varepsilon_{i,j}$$

$$(2)\ JudgeSentence = \alpha_1 + \beta DiffJudgeDist_{j,t} + \gamma X_{i,t} + \delta_j + \lambda_t + \varepsilon_{i,j}$$

$$\begin{aligned}(3)\ JudgeSentence \quad = \quad & \alpha_1 + \beta_1 ProxtoElect_{j,t} + \beta_2 DiffJudgeDist_{j,t} + \beta_3 ProxtoElect_{j,t} \times DiffJudgeDist_{j,t} \\ & + \gamma X_{i,t} + \delta_j + \lambda_t + \varepsilon_{i,j}\end{aligned}$$

where i is an indicator for the case/offender, j for the judge, and t for the year of sentence. The matrix $X_{i,t}$ includes offender's characteristics such as: age, gender, an indicator if the offender is white, and prior record score; offense's characteristics such as: gravity of the offense, an indicator whether weapon was involved, the number of counts in each case, indicator variables for negotiated and non-negotiated guilty plea (where the baseline is conviction on trial); and two measures for the district (1) the number of judges in a specific district, and (2) the number of cases with sentencing guideline level equal to five handled last year in that district. We also generate indicators for the sentencing guideline editions.

To control for judge's ideology we employ 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 controls also for the unobserved heterogeneity in districts. We also control for year and for offense statutory grade dummies. Our standard errors are clustered by judges.

Table 2, Columns (1)-(3) provide the results of our first model. The regression is run first for all judges and then separately for judges with less than 10 years of experience and for judges with more than 10 years of experience. We find the coefficient on *Proximity to Election* to be indistinguishable from zero for all judges as well as for inexperienced and experienced judges. Our findings suggest

that judges' current sentencing are not affected by proximity to election regardless of the judges' experience.

Columns (4)-(6) provide the results of our second model. As in the previous model, the regression is first run for all judges and then it is run separately for inexperienced and experienced judges. We find the coefficient on *Diff_Judge_Dist* to be indistinguishable from zero for all judges as well as for inexperienced and experienced judges. Hence, we find no effect of a judge's last year's sentencing on her current sentence.

Columns (7)-(9) provide the results of our third model. Here we introduce an interaction term between Proximity to Election and *Diff_Judge_Dist*. Column 7 shows that the coefficients on *Proximity to Election*, *Diff_Judge_Dist*, and on their interaction term are all indistinguishable different from zero when we run the regression on all judges, regardles of their experience. However, once we run this model separately for inexperienced judges we get that although *Proximity to Election* and *Diff_Judge_Dist* is positive but insignificant the interaction term is negative and statistically significant (with a coefficient of -0.328 significant at the 5% 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 run the regression separately for experienced judges we do not get such an association,

This effect is not small. Table 5, Column 1 provides the magnitude of the effect in months on sentencing, for judges who are inexperienced, a year before election (with *Proximity to Election* equal 0.9) and with *Diff_Judge_Dist* values in the 10th, 25th, 75th and 90th quantile (with *Diff_Judge_Dist* equals -40, -16, 12 and 29 respectively). Our prediction suggests that a year before standing for reelection, a judge with *Diff_Judge_Dist* in the 10th (25th) quantile would raise the length of incarceration by 9 (5) months. Similarly judges with *Diff_Judge_Dist* in the 90th (75th) quantile would lower the length of incarceration by 2 months (1) , respectively.

Most of the other covariates have signs that are consistent with findings by prior literature. Male offenders receive longer sentences than female offenders, and white offenders receive lower sentences than non-whites. As expected, when the offense has a higher OGS or a higher PRS the length of incarceration is longer. When enhanced weapon is used sentences are higher. Negotiated plea and non-negotiated plea have a negative effect on sentencing. Guideline editions have a negative effect on judge sentences (suggesting that the 6th guideline editions are less stringent than the 5th guideline edition) and the number of counts on each case has a positive and statistically significant effect. The number of judges in a district has negative and significant effect on sentencing only for experienced judges. We find offender age, and the number of similar cases handled last year in the same district to have no statistically significant effect.

4.4 Alternative Specifications of the Model

Table 3 provides the results of regressing *Judge Sentence* on *Proximity to Election Squared*. This expresses our expectation that the effect of proximity to election should not only 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 elections approximate . The results show similar pattern to the one we observed before. Table 5, Column 2 provides the magnitude of

the effect for a judge who is one year before election with *Diff_Judge_Dist* value in the 10th, 25th, 75th and 90th quantile (in a similar maner to what we did for Table 2 Column 8). We find that one year beofre reelection, a judge with *Diff_Judge_Dist* in the 10th and 25th quanitle would rasie the length of incarceration by 8 and 3 months respectively and would lower it by 3 and 6 months if his *Diff_Judge_Dist* is in the 75th and 90th quantile.

On average, judges in our data have about 14 different cases with sentencing guideline whose level equals five. However, 6% of the judges decided less than 1 'level five' case in a specific year and 14% decided less than 3 'level five' cases in a given year. Therefore, Table 4 repeats Table 3, once for all cases where judges handled more than one case in the previous year (Columns 1-3), and once for all cases where judges handled more than three cases in the previous year (Columns 4-6). The last three columns (Columns 7-9) present the results when we weight our regression by the number of cases that judges handled in the previous year. We get similar patterns to the ones obtained in Table 2 and Table 3. Table 5, Column 3,4 and 5 provide the magnitue of the effect for inexperienced judges. We find the magnitude of the effects is a bit larger, with the largest effect obtained in the weighted regressions.

5 Conclusion

This paper demonstrated the theoretical and empirical implications of a dynamic reputation approach for the study of judicial behavior. This approach formalizes a judge's reputation as a belief held by the public or any other appointing agent, about that judge's attitude, political and moral convictions. The public's belief is updated based on the judge's decisions, and consequently affects the judge's prospects of reappointment. Since the judge cares about being reappointed, her decision in each case may be affected by her effort to manipulate the public's belief.

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

Importantly, the model's prediction depends on information asymmetry between the judge and the public, not only with respect to the judge's characteristics, but also with respect to the exact merits of the case. Put differently, to manipulate the public's belief through her judicial decision, only the judge must know what the correct decision should be. Thus, the model would have more predictive power, the less information is held by the public (or any appointing body) about the merits of each case.

The reputation model casts some doubt over the value of past decisions as a predictor of future decisions.[26] Judges' past decisions do not necessarily predict how they will decide future cases. On the contrary - if judges act strategically, subject to reputation considerations, their decisions may feature negative correlation with their past decisions.

[26] See review in Brace, Langer and Hall (2000) p. 390-391

The reputation model also implies that appointment prospects may not necessarily be monotonic in the public's belief about the judge's expected bias. A judge who is believed to be more biased on any issue, may in fact decide cases in a less biased manner, compared to another judge who is believed to be less biased. If the public can predict such strategic behavior, it may take it into consideration when deciding whether to appoint the judge. Whether this is indeed how the public decides, or should one assume, instead, the public's bounded rationality in this respect, remains open for future research.

From an empirical perspective, the dynamic approach offers a way to test the legalistic conjecture. If a judge's decisions depend only on the factual and legal merits of each case, then they must be stochastically independent of each other. Any statistical dependence between a judge's decision and her previous history of decisions undermines the legalistic theory, as it can be explained only based on extra-legal motivations. Furthermore, testing 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.

Appendix

Claim 3. There are only two pure strategy equilibria. One where both types of judges always impose a light sentence and another where both types of judges impose a severe sentence.

Proof. As explained above, both of the equilibria described above 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 objective 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 would benefit from deviating and rendering the same decision that the objective type would render. Finally, for both types of judges to always impose an incorrect decision cannot be an equilibrium for the same reason that for both judges to always render the correct decision cannot be an equilibrium. Namely, in such a case the judges' decisions don't affect the pubic's belief about their type, and so both types of judges would benefit from deviating and rendering the decision they prefer. ■

References

[1] Boyea Brent D., 2010. Does Seniority Matter? The Conditional Influence of State Methods of Judicial Retention. *Social Science Quarterly* 91: 209-227

[2] 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: 1206-1231

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

[4] 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: 609-640.

[5] Caldarone Richard P., Brandice Canes-Wrone, Tom S. Clark, 2009. Partisan Labels and Democratic Accountability: An Analysis of State Supreme Court Abortion Decisions. *The Journal of Politics* 71: 560-573.

[6] Canes-Wrone Brandice, Clark Tom S., and Jee-Kwang Park. 2012. Judicial Independence and Retention Elections. *Journal of Law, Economics and Organization* 28: 211-234.

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

[8] 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.

[9] Ely, J., and Valimaki, J. 2003. Bad Reputation. *Quarterly Journal of Economics* 118: 785-814.

[10] Epstein, Lee, William M. Landes, and Richard A. Posner. 2013. *The Behavior of Federal Judges* (Oxford).

[11] Hall, Melinda G.. 1992. Electoral Politics and Strategic Voting in State Supreme Courts. *Journal of Politics* 54: 427-446.

[12] Hall, Melinda G.. 1995. Justices as Representatives: Elections and Judicial Politics in the American States. *American Politics Quarterly* 23: 485-503.

[13] Gordon Sanford C. and Gregory A. Huber. 2007. The Effect of Electoral Competitiveness on Incumbent Behavior. *Quarterly Journal of Political Science*. 2: 107-138.

[14] Helland, Eric, and Alexander Tabarrok. 2002. The Effect of Electoral Institutions on Tort Awards. *American Law and Economics Review*. 4:341-470.

[15] 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: 247-263.

[16] Huber, Gregory A., and Sanford C. Gordon. 2007a. Directing Retribution: On the Political Control of Lower Court Judges. *Journal of Law, Economics and Organization*. 23: 386-420.

[17] Klement Alon and Zvika Neeman. 2013. Does Information about Arbitrators' Win/Loss Ratio Improve Their Accuracy?. *Journal of Legal Studies* 42: 369-397

[18] Lim, Claire S.H.. 2013. Preferences and Incentives of Appointed Public Officials: Evidence from State Trial Court Judges. *American Economic Review* 103: 1360-1397.

[19] Miceli, Thomas J., and Metin M. Cosgel. 1994. Reputation and Judicial Decision-Making. *Journal of Economic Behavior and Organization* 23:31-51.

[20] Mailath, George J., and Larry Samuelson. 2006. *Repeated Games and Reputations: Long-Run Relationships* (Oxford).

[21] Shepherd, Joanna M.. 2009. The Influence of Retention Politics on Judges' Voting. *Journal of Legal Studies* 38:169-206.

[22] Shepherd, Joanna M.. 2009a. Money, Politics and Impartial Justice. *Duke Law Journal* 58: 623-685.

[23] Shepherd, Joanna M.. 2009b. Are Appointed Judges Strategic Too?. *Duke Law Journal* 58: 1589-1626.

[24] Shepherd, Joanna M. 2011. Measuring Maximizing Judges: Empirical Legal Studies, Public Choice Theory, and Judicial Behavior. *University of Illinois Law Review* 2011: 1753-1766.

Table 1: Summary Statistics

	(1) All	(2) 0-10	(3) 10+
Judge Sentence	67.228	64.090	69.994
	(70.594)	(68.649)	(72.155)
Proximity to Election	0.495	0.488	0.501
	(0.275)	(0.245)	(0.298)
Diff Judge-District	0.372	-1.834	2.135
	(27.964)	(30.149)	(2553)
Offender Age	30.873	30.572	31.139
	(10.839)	(10.726)	(10.932)
Offender Male	0.927	0.928	0.926
	(0.260)	(0.258)	(0.262)
Offender White	0.404	0.376	0.429
	(0.491)	(0.484)	(0.495)
Max 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	0.157	0.142	0.171
	(0.365)	(0.349)	(0.377)
Counts per Case	3.140	2.957	3.302
	(4.946)	(3.693)	(5.826)
No.cases by Judge Last Year	14.188	13.609	14.641
	(19.234)	(23.147)	(15.499)
Non Negotiated	0.178	0.196	0.162
	(0.382)	(0.397)	(0.368)
Negotiated Guilty	0.491	0.473	0.508
	(0.500)	(0.499)	(0.499)
No.cases in District	63.591	58.681	67.376
	(165.748)	(142.652)	(181.732)
No. of Judges in Dis	5.708	5.870	5.582
	(7.340)	(7.531)	(7.201)
Guideline Ed. 6th	0.441	0.464	0.421
	(0.497)	(0.499)	(0.494)
Guideline Ed. 6th Re	0.115	0.142	0.091
	(0.319)	(0.349)	(0.287)
Number of Obs.	31,040	14,543	16,497

Table 2: Association between judge sentence, proximity to election, judge-district diff and its interaction

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All	0-10	10+	All	0-10	10+	All	0-10	10+
Proximity to election	0.902	-2.036	1.784				1.253	2.798	1.745
	(1.794)	(4.569)	(2.061)				(1.905)	(7.031)	(2.023)
Diff Judge-District				0.016	-0.018	-0.006	0.061	0.138	-0.098
				(0.024)	(0.042)	(0.024)	(0.047)	(0.092)	(0.064)
Prox X Diff							-0.104	-0.328**	0.125
							(0.076)	(0.143)	(0.099)
Offender Age	0.326*	0.272	0.416*	0.301	0.109	0.456**	0.211	0.121	0.298
	(0.184)	(0.277)	(0.221)	(0.189)	(0.320)	(0.224)	(0.203)	(0.324)	(0.240)
Offender Age Square	-0.006**	-0.005	-0.006**	-0.005**	-0.003	-0.007**	-0.004	-0.003	-0.005
	(0.002)	(0.004)	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)
Offender Age Missing	10.182	-0.915	11.785	0.300	-21.629	10.156	-2.213	-22.325	5.780
	(12.190)	(24.372)	(12.612)	(14.938)	(32.837)	(16.074)	(14.615)	(32.150)	(15.049)
Offender Male	14.560***	14.409***	14.526***	14.090***	14.503***	14.022***	14.502***	14.404***	14.610***
	(1.221)	(1.634)	(1.666)	(1.204)	(1.734)	(1.600)	(1.289)	(1.747)	(1.788)
Offender White	-5.700***	-7.344***	-4.722***	-5.693***	-7.319***	-4.975***	-5.827***	-7.048***	-5.177***
	(0.732)	(1.205)	(0.921)	(0.758)	(1.364)	(0.969)	(0.776)	(1.360)	(0.977)
Max OGS	21.952***	21.668***	21.884***	22.712***	22.556***	22.614***	22.764***	22.628***	22.590***
	(0.477)	(0.622)	(0.657)	(0.463)	(0.673)	(0.603)	(0.498)	(0.662)	(0.678)
Offender PRS	9.715***	9.826***	9.545***	9.790***	9.968***	9.606***	9.874***	9.963***	9.724***
	(0.228)	(0.293)	(0.323)	(0.232)	(0.319)	(0.309)	(0.251)	(0.327)	(0.350)
Enhanced Weapon	33.466***	33.941***	32.161***	34.252***	35.038***	33.054***	34.811***	34.991***	33.730***
	(1.544)	(2.173)	(1.675)	(1.522)	(2.320)	(1.667)	(1.658)	(2.299)	(1.866)
Counts per Case	2.882***	4.414***	2.350***	2.839***	4.418***	2.393***	2.730***	4.387***	2.214***
	(0.668)	(0.679)	(0.696)	(0.670)	(0.764)	(0.695)	(0.667)	(0.762)	(0.669)
Non Negotiated	-10.033***	-9.386***	-10.385***	-10.835***	-11.216***	-10.614***	-10.452***	-10.724***	-9.938***
	(1.330)	(2.082)	(1.791)	(1.316)	(2.392)	(1.454)	(1.453)	(2.381)	(1.717)
Negotiated Guilty	-14.396***	-12.891***	-15.563***	-14.388***	-12.913***	-15.275***	-14.280***	-12.832***	-15.210***
	(1.233)	(1.742)	(1.598)	(1.179)	(1.831)	(1.443)	(1.244)	(1.857)	(1.544)
No.cases in District Last Year	0.002	-0.002	0.000	0.002	0.001	-0.002	0.002	-0.002	0.005
	(0.005)	(0.008)	(0.006)	(0.004)	(0.007)	(0.005)	(0.004)	(0.007)	(0.006)
No. of Judges in District	-0.341	0.029	-0.335	-0.289	0.147	-0.383*	-0.219	0.261	-0.312
	(0.211)	(0.326)	(0.238)	(0.191)	(0.336)	(0.221)	(0.200)	(0.334)	(0.232)
Guideline Ed. 6th	-7.957***	-8.863***	-6.866***	-7.965***	-7.920***	-7.683***	-7.605***	-8.087***	-6.870***
	(1.624)	(2.192)	(2.232)	(1.577)	(2.370)	(2.023)	(1.668)	(2.384)	(2.233)
Guideline Ed. 6th Revised	-13.507***	-11.809***	-15.186***	-13.654***	-11.156***	-15.378***	-13.361***	-11.406***	-15.225***
	(2.335)	(3.084)	(3.159)	(2.287)	(3.214)	(3.046)	(2.355)	(3.188)	(3.178)
Adjusted R-squared	0.4938	0.5159	0.4869	0.5009	0.5194	0.4987	0.5014	0.5214	0.4987
No. of Obs	31,040	14,543	16,497	30,190	12,098	18,092	26,849	11,923	14,926

Note: Standard errors in parentheses and are clustered by Judge. Stars denote the level of statistical significance * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. We control for Judge, Year and Offense Statutory Grade dummies.

Table 3: Different Specification for Proximity to Election

	(1)	(2)	(3)
	All	0-10	10+
Proximity to Election Squared	1.981	-0.277	1.838
	(1.878)	(8.207)	(1.927)
Diff Judge-District	0.042	0.072	-0.066
	(0.038)	(0.066)	(0.049)
Prox Square X Diff	-0.104	-0.337**	0.089
	(0.077)	(0.143)	(0.096)
Adjusted R-squared	0.5014	0.5214	0.4987
No. of Obs	26,849	11,923	14,926

Note: Standard errors in parentheses and are clustered by Judge. Stars denote the level of statistical significance * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. We control for Judge, Year and Offense Statutory Grade dummies. We also control for all other offender and offense covariates (not shown in the table)

Table 4: Sub samples defined by number of cases handled by judge last year using proximity squared

	More than one case			More than three cases			weighted by no of cases last year		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	All	0-10	10+	All	0-10	10+	All	0-10	10+
Proximity to Election Squared	1.826	0.386	1.297	1.850	3.155	1.291	0.950	2.037	0.115
	(1.932)	(8.626)	(2.012)	(1.928)	(9.225)	(2.010)	(2.122)	(17.296)	(2.260)
Diff Judge-District	0.038	0.087	-0.083	0.081*	0.170*	-0.056	0.142***	0.272**	-0.017
	(0.042)	(0.072)	(0.052)	(0.047)	(0.092)	(0.057)	(0.053)	(0.110)	(0.074)
Prox Squared X Diff	-0.122	-0.375**	0.079	-0.197**	-0.547***	0.036	-0.342***	-0.813***	-0.139
	(0.083)	(0.150)	(0.102)	(0.096)	(0.186)	(0.116)	(0.127)	(0.282)	(0.146)
Adjusted R-squared	0.5003	0.5205	0.4974	0.4994	0.5222	0.4945	0.4987	0.5124	0.4935
No. of Obs	26,287	11,643	14,644	24,976	10,913	14,063	26,849	11,923	14,926

Note: Standard errors in parentheses and are clustered by Judge. Stars denote the level of statistical significance * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. We control for Judge, Year and Offense Statutory Grade dummies. We also control for all other offender and offense covariates (not shown in the table)

Table 5: How Large is the Effect? (in months for inexperienced judges)

	(1) Table 2 Column 8	(2) Table 3 Column 2	(3) Table 4 Column 2	(4) Table 4 Column 5	(5) Table 4 Column 8
10 Percentile (Diff Judge-District=-40)	9	8	9	14	17
25 Percentile (Diff Judge-District=-16)	5	3	4	7	8
75 Percentile (Diff judge-District=12)	1	-3	-2	-1	-3
90 Percentile (Diff judge-District=12)	-2	-6	-6	-5	-10