UNIVERSITY OF BIRMINGHAM

Department of Economics

# The Effect of the Election of Prosecutors on Criminal Trials

Department of Economics Discussion Paper 11-08

Siddhartha Bandyopadhyay
Bryan C McCannon

# The Effect of the Election of Prosecutors on Criminal Trials*

Siddhartha Bandyopadhyay
University of Birmingham

Bryan C. McCannon
Wake Forest University

March 10, 2011

**Abstract**

We examine if elections of public prosecutors (as is common in the U.S.) influence the way they handle cases. In particular, does it affect which cases are taken to trial? A theoretical model is constructed where voters use outcomes of the criminal justice system as a signal of prosecutor's quality. This leads to a distortion of the mix of cases they take to trial. Our results imply that when re-election pressures are high (i) prosecutors take too many cases to trial. This increases the number of convictions from trial and reduces the amount of plea bargaining so that (ii) the proportion of convictions stemming from trial increases. Consequently, (iii) the average sanction obtained in both jury trials and plea bargains decreases. A detailed dataset from North Carolina is used to identify empirical evidence of such distortions. Our empirical findings verify that elections do affect the decision of which cases to take to trial and confirms our predictions.

**JEL codes**: K41, D82

**Keywords**: elections, prosecutor, trials

## 1 Introduction

Local prosecutors in the United States exercise enormous discretion in how they handle criminal cases.[1] There are 2344 local prosecutor offices in the U.S., which collectively handle 2.3 million felony cases each year (Perry, 2006). This represents approximately 95% of all criminal prosecutions (Simmons, 2004). A

*We thank Anindya Banerjee, Marco Barassi, Matthew Cole, Allin Cottrell, Valentina Dimitrova-Grajzl, Rob Elliott, Rosa Ferrer, Amanda Griffith, Andy Hanssen, Toby Kendall, Clare Leaver, Tito Pietra, Peter Postl, and seminar participants at the University of Birmingham, the University of Sassari, St. Bonaventure University, and the Workshop on Law, Economics, and Institutions for their comments and suggestions.

[1] See Mellon, Jacoby, and Brewer (1981) for a detailed analysis of how prosecutors exercise this discretion across the U.S.

common feature of almost all of these local offices is that the "chief" prosecutor is elected by popular vote.[2] Given their importance in the functioning of the criminal justice system, one would like to understand the impact that popular elections have on how they choose to exercise their discretion. Unfortunately, this issue has not been formally analyzed. We attempt to fill this void by asking: do elections of prosecutors influence the way they handle cases? In particular, does it affect which cases they take to trial and which cases are plea bargained? We use a model of asymmetric information to show how prosecutors could signal their quality to voters by distorting the mix of cases they take to trial. A detailed dataset from North Carolina is used to assess if there is empirical evidence of such distortions.

The model we use is a probabilistic voting version of our recent theoretical paper (Bandyopadhyay and McCannon, 2010) where we find that such distortions can indeed arise from signaling motives. It is shown that voters can use information on the aggregate sentences obtained over a term in office to judge the quality of prosecutors. Evaluating prosecutors on the basis of their being able to obtain stiff sentences leads to an insufficient amount of plea bargaining and an excessive number of trials. Desiring to be retained, prosecutors increase the aggregate sentence lengths obtained by taking cases to trial where the difference between the expected sanction from trial and that obtained via plea bargaining is too small to justify the costs. The equilibria are characterized and the following testable predictions arise. If re-election pressures are high: (i) the number of cases brought to trial goes up, (ii) the number of convictions from trial increases and the number of guilty pleas decreases so that the proportion of convictions that are obtained from jury trials increases, and (iii) the average sentence obtained in both trials and plea bargaining goes down as marginal cases with weaker evidence are brought to trial.

These hypotheses are tested using a district-level, panel dataset from North Carolina. The data cover all forty-three prosecutorial districts between the years 1997 and 2009. Data on election contests over this period are collected to measure election pressures.

[2] The three states that do not elect the public prosecutors are Alaska, Connecticut, and New Jersey (Perry, 2006).

The results receive empirical validity. The number of convictions obtained via jury trials (relative to the total number of convictions) increases in the year before an incumbent runs for re-election. In the year of re-election if the incumbent has a competitor, either in the primary or the general election, a further increase in the relative number of convictions obtained from jury trials is observed. If a district does not have a contested election in any year within the sample, fewer convictions arise from trials. Even in the most conservative estimate there is a 24.3% increase in the total number of trial convictions in a district if the incumbent is running against a challenger.

Additionally, as predicted by the theory, the average maximum sanction obtained in a year decreases. Also, the percentage of convictions resulting in community punishments, which is an indirect measure of the prosecutor's leniency, decreases.

Early analysis of prosecutors focused on the allocation of the office's budget and the role of plea bargaining (Landes, 1971; Forst and Brosi, 1977). While not about local prosecutors, Boylan (2005) analyzes data of chief federal prosecutors and their subsequent careers. He presents results that suggest prosecutors maximize sentence length, which supports our choice of evaluation metric used in the theory presented here. Glaesar, Kessler, and Piehl (2000) find evidence that career concerns affect the decisions of U.S. Attorneys and, specifically, has led to an increase in the federalization of drug crimes.

The focus in this paper is on state-level prosecutors who experience a different retention mechanism, namely election by the voting public. The work closest to ours is Rasmusen, Raghav, and Ramseyer (2009) who investigate the decision to prosecute a case (either through plea bargaining or trial) or dismiss it. An environment where a prosecutor is interested in allocating the office's budget and effort to obtaining convictions, improving conviction rates, and obtaining "personal goals" is considered. A cross-sectional dataset of 1625 local prosecutors' offices is used to assess the impact of an increase in the office's budget on outcomes. They show that increased budgets increase both conviction rates and the number of convictions. They do not consider the effect of the signaling problem arising from retention motivation and do not consider the decision to engage in plea bargaining versus proceeding to trial, as is done here.

Our work complements their findings, though, since they show that elected prosecutors have higher conviction rates, while we show that the increased success of elected prosecutors coincides with an expanded use of the courtroom. The retention of local prosecutors has also been qualitatively analyzed by Wright (2009), who presents stylized facts on media coverage of these elections.

This is not the first work to present a theoretical model of signaling in the pretrial decisions. Reinganum (1988) models the prosecutor's choice between dismissal, plea bargaining, and prosecution at trial by considering the asymmetric information regarding the strength of the evidence. The plea offer made by the prosecutor acts as a signal. Reinganum (2000) considers the opportunity for a defendant to use his offer during plea bargaining to signal the severity of the harm he caused, which may be revealed during the trial. These models are used to evaluate the effect of policy (restrictions to prosecutorial discretion and sentencing guidelines respectively) on the plea bargaining process. Here we are interested on how retention policies affect such choices.

The work is also related to the growing literature on how career concerns affect the behavior of public officials. Leaver (2009) also considers an environment of asymmetric information on the quality of public servants, but instead considers bureaucrats who differ in their ability to avoid making regulatory mistakes. An otherwise publicly-spirited regulator is interested in developing a reputation that improves after-office opportunities. She analyzes the impact of the length of term in office on how diligently the bureaucrat regulates. Shepherd (2009) considers the voting of state supreme court judges and shows that the behavior of judges is influenced by the preferences of the retention agents. Hanssen (2000) provides evidence that independent judges, less influenced by political motives, provide administrative agencies the incentive to spend more effort attempting to protect their actions from judicial review. Hanssen (1999) argues that knowledge of how court decisions will affect powerful groups, who provide support for elected judges, narrows the range of likely rulings. This diminishes the uncertainty and, consequently, decreases the amount of litigation. While not about prosecutorial behavior, these papers demonstrate that the retention motives affect the outcome of the justice system.

The theoretical model is presented in Section 2, while Section 3 derives the

equilibria and testable predictions when asymmetric information constrains the retention agent. Section 4 describes the data from North Carolina and Section 5 presents the results. Section 6 concludes.

# 2 Model

Consider a two-period model. In the first period there is a single prosecutor of unknown quality who is to decide how to handle cases brought before her. The prosecutor takes one of two quality types $q \in \{H, L\}$, referred to as high and low respectively. She is high quality with probability $\gamma \in (0, 1)$. Let the parameter $\theta$ be a summary measure of the quantity and quality of evidence she has against the defendant in any given case. It may also include information the prosecutor has on the skills of the defendant's attorney. Assume $\theta \in [0, \theta_M]$ where $\theta_M < \infty$. Observing $\theta$ for a particular case the prosecutor has two options. She may either take the case to trial or she may engage in plea bargaining.[3] Assume that a large number of cases come up in the first period. One may think, then, of the first time period as a term in office.

Denote $s$ as the sanction received if successful in the courtroom. For example, the sanction may represent the length of time the criminal is incarcerated. The sanction is known and is exogenously set by, for example, sentencing guidelines. Alternatively, with judicial discretion, uncertain parole outcomes, and appeals this is best thought of as the expected sanction and, for simplicity, is assumed to be fixed.[4] Additionally, the probability the prosecutor is successful at trial depends on the quality of the prosecutor and the evidence. A high-quality prosecutor wins at trial with probability $p_H(\theta)$, while if she is low quality she wins with probability $p_L(\theta)$. Assume $1 > p_H(\theta) > p_L(\theta) > 0 \ \forall \theta$, $p_q(0) = 0$, and $\frac{dp_q}{d\theta} \geq 0 \ \forall q$. Finally, if she takes the case to trial, a cost $c > 0$ is experienced, which for simplicity is assumed not to depend on the quality of the evidence or the prosecutor's ability.

With regards to plea bargaining for a guilty plea, denote $g(\theta)$ as the agreement that arises. As the prosecutor type is private information the outcome

[3] We assume that the choices of whether to file charges and which charges to file have already been made.

[4] The analysis is unaffected if one assumes $\frac{dx}{d\theta} \geq 0$.

under plea bargaining cannot be conditioned on type. Since there are costs associated with the trial then, absent any asymmetric information (as in Reinganum, 1988) or optimism bias[5], plea bargaining is successful. One might expect better evidence against the defendant to result in a tougher plea agreement. Hence, assume $\frac{dp_q}{d\theta} s > \frac{dg}{d\theta} \geq 0$ and $g(\theta) > 0 \; \forall \theta$. Additionally, assume that if the prosecutor is indifferent between the two options she chooses to plea bargain.

Denote $w(g(\theta))$ as the welfare generated from a case that results in $g(\theta)$ and $w(p_q(\theta) s) - c$ as the welfare generated from a case that goes to trial. Assume $w$ is a strictly increasing function. Denote $W(q)$ as the welfare generated over the entire term, or rather, the first-period welfare if the prosecutor is of quality $q$. To link the two concepts let $F : [0, \theta_M] \to [0, 1]$ denote the distribution function in which the quality/quantity of evidence from each case is (independently) drawn. Assume a large number of cases arise in the term[6] and the number of cases disposed of does not depend on the manner in which they are handled[7] so that the expected welfare from a case equals the average welfare generated from each case over the course of the term. Given this, assume first-period welfare equals the expected welfare from a randomly selected case. Thus, if a prosecutor chooses to take every case to trial where $\theta \geq \overline{\theta}$ and plea out those with $\theta < \overline{\theta}$ , then first-period welfare is

$$W(q) = \int_{\theta=0}^{\overline{\theta}} w(g(\theta)) \, dF(\theta) + \int_{\theta=\overline{\theta}}^{\theta_M} [w(p_q(\theta) s) - c] \, dF(\theta). \tag{1}$$

The assumption of welfare being monotonic in sanctions is based on the presumption that the sanction has been optimally chosen by society. This does not necessarily mean that society benefits from having every defendant punished as harshly as possible, but rather given the option to impose the high sanction deemed appropriate by the judge (or chosen by the sentencing board or legislature) or to accept a plea offer, welfare is greater if the prosecutor achieves the

[5] See Burke (2007) for how such psychological factors affect plea bargaining outcomes. See Garoupa and Stephen (2006) for a more cautionary view of the efficiency of plea bargaining.

[6] In fact, the average number of guilty felony convictions obtained by a district attorney in North Carolina over the period 1997-2009 is 799 per year, or rather, 2876 felony convictions in a term. This does not include the misdemeanor cases either.

[7] The correlation coefficient between the total number of convictions and the fraction of total convictions that arise from jury trials is only -0.133. The null hypothesis of no correlation fails to be rejected at the 1% confidence level. Thus, this assumption is justified in our data. While interesting and important issues, case backlogs, resources, and judicial incentives are not considered here.

higher of the two. Additionally, one may think of welfare as the expected welfare allowing for the possibility of wrongful convictions. Finally, denote $V(q')$ as the second-period welfare if the prosecutor is of quality $q'$ with $\frac{dV}{dq'} > 0$. Consequently, total welfare is $W(q) + V(q')$.

## 2.1 First Best

In this environment the first-best outcome can be described. Taking a case to trial is best if $w(p_q(\theta)s) - c > w(g(\theta))$. Trial costs are assumed to not be too great, $c < w(p_L(\theta_M)s) - w(g(\theta_M))$. Otherwise, if the costs are excessively high, then even with very good evidence the expected welfare of going to trial is insufficient. Since it is assumed that $\frac{dp}{d\theta}s > \frac{dg}{d\theta}$ and $p_q(0)s = 0 < c$, it follows that there exists a threshold value of $\theta$, denoted $\widetilde{\theta}_q$, where welfare is equal between the two options.[8] The threshold value depends on the prosecutor's quality since the probability of conviction depends on her abilities. Consequently, $\widetilde{\theta}_L > \widetilde{\theta}_H > 0$. If $\theta > \widetilde{\theta}_q$, then $W(q)$ is improved if the case proceeds to trial. The evidence is so substantial that the expected sanction is high enough to make the trial preferable. The plea bargain, which divides the surplus between the two parties, is insufficient.[9] If $\theta < \widetilde{\theta}_q$, then the best outcome is for the case to be decided by plea bargaining. The expected sanction is small and with the cost of the trial the resulting plea generates a better outcome. Figure 1 depicts the determination of these thresholds.

Additionally, the first-best outcome requires that a prosecutor of high quality is retained since $V(H) > V(L)$. Also, if the prosecutor is low quality, then a new one should be selected since $EV = \gamma V(H) + (1-\gamma)V(L) > V(L)$.

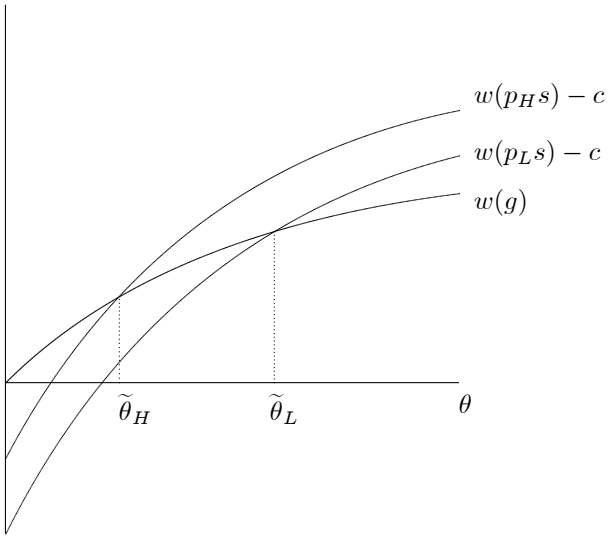
## 2.2 The Model with Asymmetric Information

Consider the principal-agent problem that arises where the retention official (e.g. median voter) wants to maximize total welfare, but does not know the

[8] With the assumption that $p_q(\theta)s$ and $g(\theta)$ are continuous functions and $c$ is bounded from above, the intermediate value theorem guarantees that these thresholds (and all thresholds derived throughout the analysis) exist.

[9] Clearly, an agreement of $g(\theta) = p_q(\theta)s$ is always strictly preferred. While, for example, we do not explicitly model the bargaining process we presume the process divides the surplus. Consequently, in cases with better evidence the expected sanction obtained in court is greater even taking into consideration the costs of trial than what can be achieved at the negotiation table.

Figure 1: First-Best Prosecution




quality of prosecutor or the evidence in each particular case. With perfect information only high-quality prosecutors, who are more likely to be successful in the courtroom, are retained. However, if there is a bonus for being retained, the prosecutor's payoff no longer coincides with the median voters. Thus, all types of prosecutors may try to signal they are of a high type. We are interested in understanding how such signals distort the outcomes in the criminal justice system.

Suppose the benefit received by the prosecutor from a particular case is either $u\left(p_q\left(\theta\right)s\right) - c$ or $u\left(g\left(\theta\right)\right)$. Her total utility, $U_q$, then, aggregates the benefit derived in each case. As in the derivation of first-period welfare, assume the expected utility of a randomly selected case equals the average utility generated over the term. Also, let $b > 0$ denote the bonus received by the prosecutor if and only if she is retained by the retention agents. One may think of the bonus as future wages earned, but could also represent the future gains from an altruistic prosecutor who if retained gains utility from prosecuting future cases.[10] Thus,

[10] This setup is similar to the 'incumbent challenger' models of political agency, which consider retention motivations of political leaders; discussed in detail by Besley (2006). In Besley's terminology $b$ can be considered the dissonant component of the prosecutor's payoff function,

if a prosecutor takes a case to trial if and only if $\theta \geq \overline{\theta}_q$, then

$$U_q = \int_{\theta=0}^{\overline{\theta}_q} u\left(g\left(\theta\right)\right) dF\left(\theta\right) + \int_{\theta=\overline{\theta}_q}^{\theta_M} \left[u\left(p_q\left(\theta\right) s\right) - c\right] dF\left(\theta\right) + b \tag{2}$$

if retained and $U_q - b$ if not retained. Hence, the prosecutor cares about obtaining the strongest punishment possible (net of any costs) as well as behaving to be retained. To simplify the analysis assume, absent compensation, the preferences of the prosecutor are directly related to welfare, $u\left(P\right) = \alpha w\left(P\right)$ for $\alpha > 0$. Thus, the environment may be best thought of as addressing the decisionmaking of the chief prosecutor who is directly accountable to the voting public so that, other than the desire to be retained, she has preferences that are aligned with societal well-being.

With asymmetric information if $\theta \in \left(\widetilde{\theta}_H, \widetilde{\theta}_L\right]$, then the first-best outcome is for a prosecutor of low quality to plea bargain and one of high quality to the case to court. The aggregate sanctions achieved by the two types will differ and retention agents will be able to identify the skill of the prosecutors. Thus, the prosecutor has the incentive to distort the mix of plea bargain versus trial to be retained. The Perfect Bayesian equilibria of this game will be identified.

# 3 Prosecutor Evaluation

The question becomes how does this incentive to misrepresent affect the behavior of the prosecutor. The retention agents, who do not have access to information on her true type, may use a number of measures to evaluate the prosecutor. One may legitimately ask if voters can use information from trials to effectively screen prosecutors. Thus, we use a particular metric, aggregate sentences obtained. Evidence has been presented that this variable is one of the primary measurements used to assess incumbent prosecutors in media coverage of elections (Wright, 2008). Boylan (2005), as stated earlier, provides evidence suggesting that prosecutors maximize it.

---

i.e. the component that causes a divergence between her preferences and that of the voters. It is possible to make $b$ type dependent as well i.e. a better prosecutor could gain higher utility in period 2 as well as she can serve society better. As it does not change the primary predictions of the model, such complications are not presented.

Hence, suppose retention agents assess the quality of the prosecutor conditioning retention on the length of the sentences she obtains. Sanctions may be achieved either through plea bargaining or successes in the courtroom. Over the course of her term many cases with differing amounts of evidence come before her. As a result, the aggregate sanction imposed over the course of her term can potentially be used to assess her quality.

As described, for each trial an independent draw of $\theta$ is taken from the distribution function $F$. The prosecutor selects a threshold value of $\theta$, denoted $\overline{\theta}_q$, for her term. Any trial with a value of $\theta$ greater than $\overline{\theta}_q$ is taken to trial, while one with a value less results in a plea bargain. This arises because the difference between the expected sanction from trial and the plea bargain grows as the quality and quantity of evidence improves. The threshold she chooses may depend on her quality. The expected sanction generated is

$$S_q = \int_{\theta=0}^{\overline{\theta}_q} g(\theta)\, dF(\theta) + \int_{\theta=\overline{\theta}_q}^{\theta_M} p_q(\theta)\, sdF(\theta). \tag{3}$$

Since it is assumed that a large number of cases are decided in a term, the expected sanction of a randomly selected case represents the actual aggregate sentence when the number of cases is normalized to unity.

Thus, the retention agent observes the aggregate sanction achieved by the prosecutor and makes a decision whether to retain or replace. If the aggregate sentence length exceeds a threshold, then the prosecutor is retained. If the threshold is not achieved, then assume the prosecutor is replaced with probability $z$. This allows for uncertainty in the process and can be a summary measure of the likelihood of a suitable replacement being found. For example, $z$ may represent the probability that the opposing party organizes a campaign to replace the incumbent. Thus, the parameter $z$ serves as a measure of competitiveness of the electoral environment.

## 3.1 Separating Equilibria

Consider, first, separating equilibria. In such outcomes the retention agents offer to keep the prosecutor if $S$ matches or exceeds $S^e$ supported by the belief that a prosecutor whose aggregate sanction is greater than or equal to $X^e$ is

type $H$ and those below are type $L$. Which values of $S^e$ are they be able to obtain in a separating equilibrium?

As a useful point of reference, consider the aggregate sanction that arises in the first-best outcome. When $\overline{\theta}_H = \widetilde{\theta}_H$

$$\widetilde{S}_q = \int_{\theta=0}^{\widetilde{\theta}_q} g(\theta)\, dF(\theta) + \int_{\theta=\widetilde{\theta}_q}^{\theta_M} p_q(\theta)\, s dF(\theta). \tag{4}$$

Since a prosecutor of high quality wins at trial with a higher probability and takes more cases to trial, it is straightforward to verify that $\widetilde{S}_H > \widetilde{S}_L$. Thus, sentence length is a reasonable metric to use to distinguish between the types.

If the benefit to being retained is relatively small, then a low-quality prosecutor is uninterested in increasing the number of cases she takes to trial to mimic a high-quality prosecutor. This self separation occurs if the probability of being replaced is small. If a low-quality prosecutor is replaced with only a low likelihood, then she will select $S_L = \widetilde{S}_L$ and take her chances at being replaced. Therefore, there exists a threshold $\underline{z}$ defined by

$$\underline{z} \equiv \frac{1}{b} \int_{\overline{\theta}_H}^{\widetilde{\theta}_L} \left[ u(g(\theta)) - u(p_L(\theta)\, s) - c \right] dF(\theta). \tag{5}$$

where if $z < \underline{z}$, then self separation occurs.

**Lemma 1** *If $z < \underline{z}$, then the unique equilibrium has $S_L = \widetilde{S}_L$ and $S_H = \widetilde{S}_H$.*

Therefore, consider the environment where $z \geq \underline{z}$ so that both types are willing to achieve $\widetilde{S}_H$ to be retained, since the benefit to being retained is sufficiently high and the probability of being replaced is great. It follows that in separating equilibrium $S^e$ must be greater than $\widetilde{S}_H$. Additionally, since for a fixed $b$ the prosecutor's utility is proportional to welfare, a high-quality prosecutor is interested, if she remains in her job, in achieving the lowest length that achieves her retention, or rather, $S = S^e$. Thus, in a separating equilibrium a threshold $\overline{\theta}_H = \theta_H^e$, below $\widetilde{\theta}_H$, is set so that $S_H = S^e$. Consequently, an excessively aggressive aggregate sentence length is obtained by trying in court some defendants who in the first-best outcome would have received a plea offer. For this to be an equilibrium she must be willing to push for these harsher sanctions. If she is not willing, then the best choice for her is $S = \widetilde{S}_H$. Therefore,

selecting $\overline{\theta}_q = \theta_H^e$ is preferable if

$$\int_0^{\theta_q^e} u\left(g\left(\theta\right)\right) dF\left(\theta\right) + \int_{\theta_q^e}^{\theta_M} \left[u\left(p_q\left(\theta\right) s\right) - c\right] dF\left(\theta\right) + b \tag{6}$$

$$\geq \int_0^{\widetilde{\theta}_q} u\left(g\left(\theta\right)\right) dF\left(\theta\right) + \int_{\widetilde{\theta}_q}^{\theta_M} \left[u\left(p_q\left(\theta\right) s\right) - c\right] dF\left(\theta\right) + \left(1 - z\right) b$$

where $q = H$. There exists a threshold level of $\theta$, denoted $\phi_q$, where (6) holds with equality. Hence, if $\theta_H^e \geq \phi_H$, then the high-quality prosecutor is willing to take enough cases to court to remain in her position. This cutoff value corresponds to a level of $S$, denoted $S_q^\phi$ $(> \widetilde{S}_q)$.

Now consider the incentives of the prosecutor of low quality. If she does not attempt to increase $S$ to $S^e$, then she will not be retained. Since her preferences are proportional to welfare it follows that if she chooses not to achieve $S^e$ she selects $S = \widetilde{S}_L$ by adopting $\overline{\theta}_L = \widetilde{\theta}_L$. Alternatively, she may choose to misrepresent her type by taking more cases to court and plea bargaining less. Since she is less successful in court, even more cases must be tried than if she were more skilled. Thus, to obtain $S^e$ her threshold, $\overline{\theta}_L$, must be set at a value less than $\theta_H^e$. As before, denote the value of $\theta$ that generates $S_L = S^e$ as $\theta_L^e$. For $S^e$ to be supported as a separating equilibrium (6) must fail for $\overline{\theta}_q = \theta_q^e$ when $q = L$. Thus, a low-quality prosecutor is unwilling to mimic the high-quality prosecutor if $\theta_L^e < \phi_L$, or rather if $S^e > S_L^\phi$.

Combining these two results, if $S^e > \widetilde{S}_H$ a separating equilibrium exists so long as $S^e \in \left(S_L^\phi, S_H^\phi\right]$. Since $p_H\left(\theta\right) > p_L\left(\theta\right)$ $\forall\theta$ if a low-quality prosecutor is willing to achieve $S^e$ so to is a high-quality one. Therefore, $S_H^\phi > S_L^\phi$ and the interval is non-empty.

**Proposition 1** *There exists a range of aggregate sanctions,* $\left(S_L^\phi, S_H^\phi\right]$*, in which separating equilibria exist.*

As a consequence of Proposition 1, the incentives of a prosecutor encourage one of high quality to engage in an excessive number of trials. This increase has to be so great that a low-quality prosecutor is unwilling to plea bargain so few cases. Interestingly, in these equilibria it is the low-quality prosecutor that selects the welfare-maximizing number of trials.

### 3.2 Pooling Equilibria

In a pooling equilibrium both types achieve an aggregate sentence length $S^e$ and the retention agents believe that a prosecutor who achieves $S^e$ or above is type $H$ with probability $\gamma$ and type $L$ with probability $1-\gamma$. What outcomes can arise in such an equilibrium?

First, $S^e \geq \widetilde{S}_H$ since if not a high-quality prosecutor would prefer to deviate.

Second, for any outcome to be a pooling equilibrium each agent must be unwilling to deviate. Since the outcome results in a greater aggregate sanction than what the prosecutor prefers it to be, regardless of her type, then only deviations to lower aggregate sentences need be considered. If the retention agents believe that any such deviations are done by low-quality prosecutors, then they will choose not to retain one who makes the choice. Thus, a pooling equilibrium requires that deviating to $\widetilde{S}_q$, the most preferred deviation, results in a lower utility than achieving $S^e$ and being retained. Again, this requires that (6) hold when $\theta_q^e$ is the cutoff value of $\theta$ that if used obtains $S = S^e$. Hence, a prosecutor of quality $q$ is unwilling to pool with the other type if $\theta_q^e < \phi_q$. As a result, $S^e$ can be supported as a pooling equilibrium only if $S^e \leq S_q^\phi \ \forall q$.

Finally, since $p_H(\theta) > p_L(\theta) \ \forall \theta$, any aggregate sentence length a low-quality prosecutor is willing to achieve to be retained a high-quality one is also willing to obtain it to remain in the position. Thus, $S_H^\phi > S_L^\phi$. Additionally, since it is assumed that $b \geq \underline{b}$ (so that self separation is not an outcome), it must be that $S_L^\phi > \widetilde{S}_H$. As a result, the interval $\left[\widetilde{S}_H, S_L^\phi\right]$ is non-empty and describes the set of pooling equilibria.

Alternatively, if $S^e > \widetilde{S}_H$, then it may be reasonable for retention agents to believe a deviation to $\widetilde{S}_H$ is done by a high-quality prosecutor. With these beliefs no pooling equilibrium other than $S^e = \widetilde{S}_H$ exists. As is typical in signaling models, the size of the set of pooling equilibria depends on the belief regarding deviations to non-equilibrium outcomes.

**Proposition 2** *If it is believed that a deviation to a lower aggregate sentence length is done by a low-quality prosecutor, then there exists a range of aggregate sanctions, $\left[\widetilde{S}_H, S_L^\phi\right]$, in which pooling equilibria exist. If such deviations are believed to be done by a high-quality prosecutor, then the unique pooling*

*equilibrium is* $\widetilde{S}_H$.

As a consequence of Proposition 2, a prosecutor of low quality is encouraged to engage in an excessive number of trials to improve the length of her obtained sentences. Interestingly, if the retention agents believe that lower aggregate sentence lengths come from a low-quality prosecutor, then there exist equilibrium where both types of prosecutors engage in an insufficient amount of plea bargaining. Alternatively, with more reasonable beliefs high-quality prosecutors engage in the first-best amount of prosecution.

### 3.3 Testable Predictions

The theoretical model leads to a number of predictions that can be used as a test of the theory. First, consider cases that go to trial. As shown in Lemma 1 if the re-election pressures are low, then the first-best outcome is selected by the prosecutor, regardless of her type. Propositions 1 and 2 illustrate that in all separating and pooling equilibrium the number of cases taken to trial is greater than what would be selected in the first-best outcome. Thus,

**Result 1** *If re-election pressures are high, there is an increase in the number of cases taken to trial and a decrease of the number of cases plea bargained.*

As an immediate consequence of this result, the number of convictions adjusts,

**Result 2** *If re-election pressures are high, then there is an increase in the number of convictions arising from jury trials and a decrease in the number arising from plea bargaining. Consequently, the proportion of total convictions coming from trial increases.*

Additionally, this distortion affects the average sentences obtained in a year.

**Result 3** *The average sentence obtained in jury trials decreases as does the average guilty plea.*

More marginal cases with lower expected sanctions are taken to trial when the prosecutor expands her use of the courtroom. This reduces the average sentences obtained in trial even though the aggregate sentenced obtains increases.

Correspondingly, these were the better cases among those plea bargained so that the average guilty plea reduces as well. A formal proof of this result is provided in the Appendix. These three results provide the hypotheses to be tested.

# 4 Data

The theoretical model gives a number of clear, testable predictions regarding the behavior of public prosecutors who face re-election pressures (the most common form of retention in the U.S.). When election pressures are high prosecutors take more cases to trial and plea bargain less. Thus, the proportion of convictions that come from jury trials will increase. This causes the average sanction received from a case that does go to trial to be lower since more marginal cases are pursued than if retention motivations were mitigated.

Do the effects predicted by the theory arise in practice? Data on crime, convictions, and elections in North Carolina is collected. While there are one hundred counties in the state, there are only forty-three prosecutorial districts. More heavily populated counties, such as Mecklenburg county which contains the city of Charlotte and Forsyth county which has Winston-Salem, make up an entire district. More rural, less-populated counties are grouped together into a single prosecutorial district. Each district has one chief public prosecutor (known as the district attorney) who is elected by voters to serve a four-year term. Each district attorney has a staff of assistant district attorneys. The number of assistant district attorneys varies across the districts.

Across the U.S. each chief public prosecutor has a staff of on average 10.3 assistant prosecutors and 20.0 support staffers, which includes non-litigating attorneys and investigators. For offices in the U.S. serving populations under 250,000 citizens (which includes 75.2% of the districts in North Carolina) the median number of felony jury trials in a year is ten (Perry, 2006). Thus, the chief prosecutor can be expected to be either making the final decision on whether to take a case to court or to be well-informed if a member of his/her staff proceeds to trial. Therefore, the concerns and motivations of the chief prosecutor can be expected to influence the decisions of the office.

Data for felony prosecutions are collected from the North Carolina Sentenc-

ing and Policy Advisory Commission. Each year the Commission publishes the *Structured Sentencing Statistical Report for Felonies and Misdemeanors.*[11] The Report provides data on convictions obtained for the fiscal year from July 1 to June 30 of the following year. Data are collected from the 1997-98 fiscal year to the 2008-09 fiscal year. Thus, the data set covers twelve years of convictions in North Carolina. Only felony convictions are considered here. In each year a variety of information is available. First, in each district the total number of convictions obtained via guilty pleas and the total number of convictions obtained via jury trials is reported. Guilty pleas are primarily obtained through plea bargaining. The variable *jury* measures the proportion of total convictions that come from jury trials in a district for a year.

The primary testable prediction of the theoretical is that when re-election pressures are high prosecutors take more cases to trial and plea bargain less so that more convictions arise from the courtroom relative to plea bargains. Thus, the hypothesis is that *jury* is distorted by elections. Measuring jury convictions as a proportion of total convictions controls for scale effects from districts differing in the total number of crimes committed. Results with the total number of convictions obtained (for both jury trials and guilty pleas) is presented in the Appendix.

Information is given regarding the type of punishment obtained. The type of punishment is determined by the most serious sentence an offender received. Each sanction falls into one of three types. The variable *com* is the percentage of convictions that result in a community punishment as the most severe sanction. Community punishments include community service, unsupervised or supervised probation, outpatient drug or alcohol treatment, and fines. They are most frequently used in property and (non-trafficking) drug crime types. Sanctions may also be classified as having an intermediate sanction. Examples of intermediate punishments in North Carolina include intensive supervised probation, house arrest with electronic monitoring, residence in a facility (such as a substance abuse facility), or drug treatment court. Hence, *inter* is the percentage of convictions that result in an intermediate punishment as the harshest punishment. Finally, *active* is the percentage of convictions that result in an

[11] www.nccourts.org/Courts/CRS/Councils/spac/Publication/Statistical/Annual

active prison sentence. As a result, $active + inter + com = 1$. For crimes against the person most offenders receive an active punishment (64% in 2008-09), while they rarely receive community punishments (only 2% in 2008-09). For crimes against property intermediate punishments are most common (50%), while active punishments are frequent (34%). Additionally, (non-trafficking) drug crimes are rather split. While intermediate sanctions are common (50% in 2008-09), both active punishments along with community punishments commonly arise (22% and 28% respectively in 2008-09).

Third, North Carolina uses sentencing guidelines to determine minimum and maximum sanctions. District attorneys complete a worksheet to calculate a score, which measures the number and severity of past offenses. Given the offense the individual is convicted of and his/her prior record score, the guidelines provide for a range of acceptable minimum sanctions and the corresponding range of acceptable maximum ranges. Judges may, though, choose to deviate from these bounds. North Carolina is unique in that the sentencing guidelines also lay out a range of acceptable minimum and maximum sanctions for both aggravated and mitigated punishments. For example, suppose an individual is convicted of a Class D felony (e.g. armed robbery) and had a prior record score of 6 (e.g. one prior Class F and one prior Class I). The presumptive range of sanctions would be for his minimum sanction to be between 82 and 103 months, while his maximum sanction falls between 108 and 133 months. If the judge chooses to inflict an aggravated punishment, then his maximum may reach up to 164 months, while the lowest allowable mitigated punishment is 53 months. Hence, while it is mandatory for judges in North Carolina to select a sanction from the sentencing guidelines, the guidelines provide for enhanced and reduced sentences.[12] If a mitigated or aggravated sanction is selected, the judge is required to provide his/her reason for the deviation. Additionally, defendants may appeal these deviations.

Thus, $agg$ is the percentage of active punishments that exceed, $mit$ is the percentage that fall short, and $pre$ is the percentage that are within the sentencing guidelines. Hence, $agg + pre + mit = 1$.

[12] In fact, the National Center for State Courts ranks North Carolina as having the most "mandatory" of sentencing guidelines in the country.

Additionally, information on the duration of the active punishments is given. For each district for each year the average maximum, *max*, sanction received by those convicted is reported.[13]

All of the these sanction variables serve as measurements of the particular distribution of cases and crimes that arise in a district over the course of the year. The seriousness of the offenses committed in a year affects the decision to go to trial, which is not dependent on re-election motivations.

Additionally, one may expect the number of cases taken to trial to depend on the financial resources of the prosecutor's office. Unfortunately, historical, district-level budget data is not available. Data on the total annual budget, total spent on corrections, and, consequently, the percentage of the budget spent on corrections are available from 1984-85 to 2010-11. Correlations between each of these three metrics and dummy variables capturing the year before, year of, and year after a gubernatorial election (which occurs every four years) reveal no significant correlations. Also, Gubernatorial races occur in the middle of most prosecutors' terms as well so any adjustment to funding for additional prosecution of crime that does arise does not occur when prosecutors are running for re-election. Thus, the North Carolina budget does not seem to move with the election cycle and, hence, the effect of prosecutorial resources should be captured by district fixed effects.

Furthermore, socio-economic variables are created to control for unobservable differences between the districts over time. Population data is collected from the North Carolina Office of State Budget and Management.[14] Annual county-level population estimates are provided. Hence, *density* calculates the number of individuals who live in a district in each year divided by the number of square miles the district covers. The variable $\%16-24$ measures the fraction of a district's population that is between the ages of sixteen and twenty-four. Additionally, the Office provides data on the number of males and the number of whites in each county in each year. Thus, *white* is the fraction of the popula-

[13] The size of the gap between the minimum and maximum sanction is fixed by the sentencing guidelines. Thus, only *max* is considered here.

[14] www.osbm.state.nc.us/ncosbm/facts_and_figures/socioeconomic_data/population_estimates/county_estimates.shtm provides the data and a description of the estimation procedures used.

tion who are identified as white and *male* is the fraction of the district in each year that is male. Additionally, district-level labor market data is collected from the Employment Securities Commission of North Carolina. The unemployment rate, *ur*, is used as a control for economic opportunities and the opportunity cost of crime.[15]

One may be concerned that political party affiliation/ideology may affect a prosecutor's decision to go to trial as well. A dummy variable capturing the political party of the incumbent was created. It is highly correlated with most of the socio-economic variables and its inclusion in the specifications does not affect the sign or significance of the primary variables of interest. The estimations including political party affiliation, then, are not presented here.

In a number of cases county-level data is aggregated to create the district-level data. For the *active*, *inter*, and *com* variables as well as the *agg*, *pre*, and *mit* variables the total number of convictions in each category are summed for all counties within a district. The percentages are then calculated. Similarly, for *male* and *white* the total number is summed over all counties in the district and then divided by the total population of the district. The total number of people employed and unemployed across the counties in a district is aggregated to calculate *ur*. *Density* simply sums the total population in a district by the total square mileage of the district. For the district-level *max* the average maximum sanction for each county is multiplied by the total number of convictions in that country. This generates the total number of months sentenced in each county. These county-level totals are summed and then divided by the total number of convictions in the district. Similarly, the $\%16-24$ of the district is created using a population-weighted average of each county.

A few other issues are dealt with when constructing the data set. First, the demographic variables are estimated for the calendar year, while the conviction data covers the fiscal year. We simply use the estimated population for the calendar year of the second half of the fiscal year as the population. The NC Office of State Budget and Management estimated the population as of January 1 of that year. Thus, the 2009 population is the estimated population for the midpoint

[15] Labor data is obtained from www.nces.com. Labor force participation rate was statistically insignificant in the results and, hence, only the specifications with *ur* are presented.

of the fiscal year. Second, population levels, including the number of males and whites, is not available for 1998 (the first year of the data set). The 1995 data is available though. Thus, an average per year change in each district is used to estimate the 1998 population data; $POP_{98} = POP_{95} + 0.75\,(POP_{99} - POP_{95})$.

There are four cases where the composition of the district changed. While in no circumstance was a county split between multiple districts, in four cases a prosecutorial district which contained multiple districts was split into two districts. In the fiscal years 1997-98 to 2005-06 there were thirty-nine districts. District 20 initially contained four counties. For the 2006 election one county was split from the others to create districts 20a and 20b. The incumbent vacated the position and, therefore, two newly created districts originated in 2006. Also, district 29 contained five counties. Two of the counties were split off to create a new district, 29a, and an open election was held in 2006 to fill the vacancy. The remaining counties were relabeled 29b and the same DA who had been an unchallenged incumbent was again unchallenged for the position. Thus, 29b is considered a continuation of 29. Hence, for the 2006-07 and 2007-08 fiscal years there were forty-one prosecutorial districts. Initially, district 22 contained four counties. While in 2006 an individual ran unopposed to fill a vacant seat, by the 2008 election the district was divided into two districts each with two counties. The incumbent remained the DA for the district, 22b, and an election was held for the newly created district, 22a. Thus, 22b is considered a continuation of 22. Finally, one county was split from district 19b to create a brand new district, 19d, in the 2008 election. Thus, a new district was created. As a result, for the 2008-09 fiscal year there were forty-three prosecutorial districts. Consequently, there are 476 observations.

Finally, election data is collected from the North Carolina State Board of Elections.[16] A number of dummy variables are created to capture the election outcomes. The dummy variable $CI$ is equal to one if in a district in that year an incumbent runs for re-election and has a challenger either in the primary or the general election (or both).[17] $CI$ is equal to zero if there is no election,

[16] www.sboe.state.nc.us

[17] Since the fiscal year begins July 1 of the year and the general election is in November of the year the election is considered within the year if the general election is held in the fiscal year. This implies, though, that the voting for the primaries occur just before the beginning

if the incumbent is not running for re-election, or if the incumbent is running for re-election but does not have a competitor. Second, the dummy variable *reelect* is equal to one if it is the year before an incumbent runs for re-election. *Reelect* is equal to zero if it is not the year before an election or if it is the year before an election but the incumbent does not run for re-election in the following year. Third, the dummy variable *never* is equal to one if the district has not had either a contested primary election nor a contested general election in any election between 1998 and 2008. *Never* is equal to zero if in an election in any stage the voters in a district have had a choice between candidates.

These three variables are used to measure re-election pressures. It is posited that in the year before a re-election campaign an incumbent, if she does adjust her behavior to the election cycle, will respond more in that year than the previous two years. In the year of the election the presence of a challenger allows for the possibility of not being retained and, thus, compared to years without an election or without challengers the theory predicts an increase in the use of the courtroom relative to the use of plea bargaining. Therefore, *reelect* captures the motivation to remain in office, while $CI$ measures the direct pressure applied by a challenger. We assume voters focus more heavily on the prosecutor's performance in the year prior to and in the year of the election. This is a common assumption in the literature on political business cycles (Rogoff, 1990; Rogoff and Siebert, 1988). Additionally, it is predicted that the set of equilibria involve outcomes where there is less of a use of the courts when $z$, the probability of being removed from office by the voters if the incumbent underperforms, is low. It is posited here that if a district has not experienced a single contested election over the entire sample period in either primary or general elections, then the incumbent prosecutors of these districts experience a lower value of $z$ than prosecutors in districts with electoral contests. Hence, *never* captures environments where electoral concerns are minimal.

of the fiscal year in which they are recorded.

Table 1 presents the variables of interest to the econometric model.

Table 1: Variable Descriptions

| **variable** | **description** |
|---|---|
| *jury* | % of convictions obtained from a jury trial |
| *CI* | =1 if an incumbent is challenged |
| *reeleect* | =1 if in the next year the incumbent runs for re-election |
| *never* | =1 if the district had no contested election (1997-2009) |
| *agg* | % of convictions with the sanction in aggravated range |
| *mit* | % of convictions with the sanction in the mitigated range |
| *inter* | % of convictions resulting in an intermediate sanction |
| *com* | % of convictions resulting in a community punishment |
| *max* | average maximum active sanction |
| *male* | % of population that is male |
| *white* | % of population that is white |
| $\%16-24$ | % of population between the ages of 16 and 24 |
| *density* | population / mile$^2$ |
| *ur* | unemployed / (employed + unemployed) |

Table 2 presents the descriptive statistics.

Table 2: Descriptive Statistics

| **variable** | **mean** | **st. dev.** | **min** | **max** |
|---|---|---|---|---|
| *jury* | 0.0259 | 0.0147 | 0 | 0.1147 |
| *CI* | 0.0651 | 0.2470 | 0 | 1 |
| *reeleect* | 0.2668 | 0.4428 | 0 | 1 |
| *never* | 0.2815 | 0.4502 | 0 | 1 |
| *agg* | 0.0595 | 0.0477 | 0 | 0.4326 |
| *mit* | 0.1947 | 0.1025 | 0 | 0.6415 |
| *inter* | 0.4385 | 0.0721 | 0.0424 | 0.8774 |
| *com* | 0.2037 | 0.0656 | 0.0238 | 0.4425 |
| *max* | 42.466 | 9.423 | 16.39 | 77.20 |
| *male* | 0.4902 | 0.0099 | 0.4685 | 0.5302 |
| *white* | 0.7397 | 0.1560 | 0.3489 | 0.9777 |
| $\%16-24$ | 0.1294 | 0.0239 | 0.0952 | 0.2062 |
| *density* | 268.13 | 292.75 | 35.54 | 1698.5 |
| *ur* | 0.0622 | 0.0221 | 0.0127 | 0.1442 |

Just how prevalent is contested district attorney elections? As stated, each DA serves a four-year term. In North Carolina elections occur in every other year. Table 3 provides information on the elections in North Carolina.

Table 3: Prosecutor Elections in North Carolina

| | **1998** | **2000** | **2002** | **2004** | **2006** | **2008** | **total** |
|---|---|---|---|---|---|---|---|
| # of elections | 37 | 4 | 37 | 4 | 39 | 4 | 125 |
| # of contested general elections | 11 | 0 | 5 | 1 | 10 | 1 | 28 |
| # of contested primary elections | 7 | 2 | 9 | 0 | 12 | 1 | 31 |
| # uncontested elections | 19 | 2 | 25 | 3 | 21 | 2 | 72 |
| # of elections with vacancies[18] | 5 | 3 | 6 | 2 | 12 | 1 | 29 |

Contests are somewhat common. In 22.4% of the elections there was a contest in the general election, while in 24.8% of the elections there was a contest in a primary election. Considering only general election contests misses many instances of challenges to an incumbent's seat. The majority of the districts hold elections in the a different year from the U.S. Presidential election, but as of the 2008 election 11.6% districts hold elections in the year of the Presidential election.

# 5 Results

Our hypothesis is that retention motivations affect the incumbent's decision to take a case to trial. As stated, there are three variables of primary interest that capture the effect of re-election concerns (or the lack thereof) on prosecutorial decisionmaking; $CI$, *reelect*, and *never*. Table 4 presents the average number of trials that occur in a district in a year (in both absolute levels and as a proportion of total convictions) in subsamples defined by these variables.

[18] Information on whether an interim district attorney, appointed by the Governor, was running for election is not available. Thus, a seat is considered vacant if the individual, victorious in the previous election, did not run in either the primary or general election.

Table 4: Elections and Trial Convictions

| | Total Trial Convictions | % of Total Convictions |
|---|---|---|
| first two years of a term[19] | 17.269 | 0.0246 |
| year before a re-election | 18.505 | 0.0291 |
| year of re-election w/ a challenger | 19.516 | 0.0304 |
| | | |
| districts w/o a contest | 16.299 | 0.0221 |
| districts w/ at least one contest | 18.196 | 0.0273 |

This suggests that the number of convictions from jury trial is affected by election pressures. When an incumbent runs for re-election there is an increase in the total trial convictions by $(18.505 - 17.269)/17.269 = 7.2\%$ and an increase in jury trial convictions relative to total convictions by $(0.0291 - 0.0246)/0.0246 = 18.3\%$. When she faces a challenger an additional increase of 13.0% and 23.7% (for the first and second column respectively) arises. An incumbent who is safe from such pressures prosecutes less (11.6% and 23.8%). A more rigorous econometric test is needed, though, to verify that such effects actually do occur.

The variables *reelect* and $CI$ measure the existence of a re-election campaign and the presence of a challenger respectfully. Jointly, these identify the potential distortions caused by retention motivations. $CI$ is obviously collinear with *never* since for the former to equal one is required that the latter equals zero. The correlation coefficient between the two is -0.1652 and the null hypothesis of zero correlation can be rejected. Hence, due to the collinearity, separate specifications are estimated to identify the impact of the election variables on the use of jury trials. Additionally, since the bulk of the elections occur in the 1998-2002-2006 cycle $CI$ and *reelect* are highly correlated with the year fixed effects (see Table $A1$ in the Appendix). Finally, the variable *never* is a district-level control and, hence, both *never* and district fixed effects cannot be used in a specification. Consequently, the econometric models to be estimated are

$$\begin{aligned} jury_{dt} &= \alpha_0 + \alpha_1 CI_{dt} + \alpha_2 reelect_{dt} + \alpha_3 S_{dt} + \alpha_4 X_{dt} + \alpha_5 D_t + \epsilon \quad (7) \\ jury_{dt} &= \beta_0 + \beta_1 never_d + \beta_2 S_{dt} + \beta_3 X_{dt} + \beta_4 Y_t + \epsilon \end{aligned}$$

[19] This also includes the last two years of a term when the incumbent vacated the office at the end of his/her term.

where $S$ is the sanction variables: *mit*, *agg*, *com*, *inter*, and *max*,[20] $X$ is the socio-economic controls: *density*, *white*, *male*, $\%16-24$ and *ur*, and $D$ and $Y$ are district and year fixed effects respectively.

As stated, the variables $CI$ and *reelect* capture the significance of the motivation to be retained. Consequently, theory predicts $\alpha_1 > 0$ and $\alpha_2 > 0$. Also, since *never* identifies environments in which re-election concerns are minimal, theory predicts that $\beta_1 < 0$.

Table 5 presents the results.

Table 5: Results (dep. var. = *jury*, $N = 476$)

| | **I** | **II** | **III** | **IV** | **V** |
|---|---|---|---|---|---|
| *CI* | 0.0052 ** | 0.0038 ** | | | 0.0044 * |
| | (0.0026) | (0.0018) | | | (0.0026) |
| *reeleect* | 0.0034 ** | 0.0025 ** | | | 0.0031 ** |
| | (0.0015) | (0.0011) | | | (0.0015) |
| *never* | | | -0.0048 *** | -0.0028 ** | -0.0039 *** |
| | | | (0.0011) | (0.0012) | (0.0012) |
| *agg* | 0.0410 *** | 0.0341 ** | 0.01146 | 0.0048 | 0.0326 ** |
| | (0.0151) | (0.0159) | (0.0154) | (0.0136) | (0.0152) |
| *mit* | -0.0244 *** | -0.0159 ** | -0.0158 ** | -0.0070 | -0.0195 *** |
| | (0.0067) | (0.0067) | (0.0066) | (0.0074) | (0.0065) |
| *inter* | -0.0342 *** | -0.0254 ** | -0.0327 *** | -0.0334 *** | -0.0263 *** |
| | (0.0104) | (0.0104) | (0.0107) | (0.0101) | (0.0099) |
| *com* | -0.0094 | -0.0243 * | -0.0145 | -0.0178 | 0.0102 |
| | (0.0114) | (0.0142) | (0.0124) | (0.0131) | (0.0119) |
| *max* | 0.0006 *** | 0.0004 *** | 0.0006 *** | 0.0005 *** | 0.0006 *** |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| *male* | | -1.1489 *** | | -0.3050 *** | -0.1993 *** |
| | | (0.3932) | | (0.0593) | (0.0529) |
| *white* | | -0.0148 | | 0.0181 *** | -0.0032 |
| | | (0.1103) | | (0.0053) | (0.0043) |
| $\%16-24$ | | -0.1047 | | 0.0079 | -0.0461 * |
| | | (0.1291) | | (0.0273) | (0.0275) |
| *density* | | -0.000002 | | 0.00001 *** | 0.000001 |
| | | (0.000016) | | (0.00000) | (0.000001) |
| *ur* | | 0.0211 | | 0.2466 *** | 0.0506 |
| | | (0.0295) | | (0.0778) | (0.0312) |
| year controls | NO | NO | YES | YES | NO |
| district controls | NO | YES | NO | NO | NO |
| adj $R^2$ | 0.2651 | 0.5877 | 0.2941 | 0.3697 | 0.3089 |

[20] In each specification the variables *pre*, *active*, and the dummy variables for the year 1997-98 and District 1 are omitted. Also, a constant term is included in each specification, but not reported.

* 10% level; ** 5% level; *** 1% level. Heteroskedastic-robust standard errors are reported.

The variables $CI$ and *reelect* are positive and statistically significant. If an incumbent is going to run for re-election and interested in increasing her "toughness", then she increases the number of convictions that come from jury trials in the prior year. This can be done to gain notice in the news, appear hawkish to voters, obtain greater sanctions, deter challengers in the general election, and to be successful in the primary. If in the year of re-election a challenger enters, either in the primary or in the general election, she increases the number of convictions obtained from jury trials relative to the number of convictions obtained from guilty pleas. As reported in Table 2 the mean value of $jury$ is 0.0259. Thus, this represents a 13.1 to 20.1% increase in the year before the re-election and an additional 9.7 to 14.7% increase if in the re-election year a challenger enters the race. Combining the effects the re-election campaign and the presence of a challenger increases the number of convictions from jury trials (relative to the total number of convictions) by 24.3 to 33.2%. Put another way, the average number of jury convictions in a district in a year in North Carolina is 17.66 and the average number of guilty pleas is 700.92. Thus, there are 1.80 more convictions from jury trials in the year before the re-election and 2.73 more convictions in the year of the contested race, for a total of 4.53 more jury trial convictions (using II).

The specifications in columns I and II differ in whether the district-level variables, the socio-economic variables and the district fixed effects, are included in the specification. An $F$-test of the joint null hypothesis that the socio-economic variables have no effect on the use of jury trials can be rejected at the 1% level. Similarly, the hypothesis that the district fixed effects have no effect can also be rejected at the 1% level. The inclusion of the variables in Column II has no effect on the statistical significance of the primary coefficients of interest and has only a slight impact of their magnitude.

Additionally, the coefficient for *never* is negative and statistically significant. Thus, a district that has not had an election contest between 1997 and 2009 (either in a primary race or in a general election) experiences fewer jury

trials.[21] This corresponds to a 10.8 to 18.6% decrease in the number of jury trial convictions. Columns III and IV differ in whether the socio-economic variables are included in the specification. An $F$-test of the joint null hypothesis that the socio-economic variables have no effect on the use of jury trials can again be rejected at the 1% level.

The sanction control variables are, for the most part, statistically significant and have the predicted sign. A district that has more cases in which the defendant receives an aggravated sentence requires the district attorney to prosecute more cases, while if there are more cases where mitigated sanctions are handed out the prosecutor takes fewer cases to trial. Community punishments are not related to the proportion of convictions that come from jury trials, but an increase in the percentage of convictions that result in intermediate sanctions corresponds to a decrease in the number of convictions via jury trials. The magnitude of the sanction given out, as one would predict, is a significant determinant of the use of jury trials.

The socio-economic variables are valuable in explaining the use of jury trials to obtain convictions. Districts with a higher percentage of the population being male decrease the number of convictions that are obtained in court. Additional analysis reveals that *male* significantly increases the total number of guilty pleas and substantially decreases the number of convictions from jury trials (see Table *A*3 in the Appendix). There is only inconsistent evidence that the racial composition, age distribution, unemployment rate, and population density affect the use of the courtroom to decide cases.

Even though $CI$ and *never* are collinear, it is still instructive to consider their effects together. Column V presents the results. The statistical significance and the magnitude of the coefficients of primary interest are relatively unchanged. Thus, the importance of re-election pressures on prosecutorial decisionmaking is consistent throughout the specifications.

[21] Dummy variables for whether the same incumbent held the office for the sample period and another for whether the same incumbent both held the office for the entire period and had no challenger were considered. The first does not have a statistically significant impact either when used in conjunction with never or in place of it. The latter has a statistically significant and negative impact when (and only when) used to replace *never*.

While the sanction, *max*, and the level of the punishment, *agg*, *pre*, and *mit*, are set by the judge in the case the prosecutor can be expected to be able to determine or at least influence the form of the punishment. This would especially be true for negotiated guilty pleas. Hence, one may be concerned about the potential endogeneity of *com*, *inter*, and *active*. Dropping these control variables from the specifications presented in Table 5 maintains the significance of the primary variables of interest.[22] These results, though, are not presented here.

A number of other election-related variables were considered. First, a dummy variable for being in the year before an incumbent has a contested re-election, $CI_{t-1}$, was included but was not statistically significant at the 10% level. This variable would include incumbents who believed their position is secure and excludes those running for re-election who are unsure about their retention. Second, a dummy variable equal to one if an only if an incumbent is running for re-election in the year, $reelect_{t+1}$, has a positive and statistically significant impact when used to replace $CI$. Similarly, a variable capturing election contests, whether they involve an incumbent or a vacant seat, is positive and statistically significant at the 10% level. These variables do not provide new insights since they also capture both safe and vacant positions and, therefore, the additional results are not presented here.

Finally, if one replaces *reelect* with $CI_{t-1}$, year fixed effects may be added to the econometric model. The variable $CI$ maintains its sign and statistical significance in the estimation. Thus, the results in columns I and II are not sensitive to the exclusion of year effects.

While year and district fixed effects are valuable at controlling for unobservable causes of the use of the courtroom for felony cases. One may instead be interested the potential effect of district-specific error terms. Hence, a random effects model is estimated. Table 6 presents the results.

[22] One may also argue that the level of the sanctions imposed is affected by prosecutorial decisions as they select which cases to plea bargain. Removing *max* maintains the statistical significance of *reelect* in I and II and *never* in III and IV, but the p-value of $CI$ increases to 0.11.

Table 6: Random Effects (dep. var. = *jury*, $N = 476$)

| | **I** | **II** |
|---|---|---|
| *CI* | 0.0038 ** | 0.0039 ** |
| | (0.0019) | (0.0020) |
| *reelect* | 0.0033 *** | 0.0029 *** |
| | (0.0011) | (0.0011) |
| controls: | | |
| socio-economic | NO | YES |
| sanction | YES | YES |

* 10% level; ** 5% level; *** 1% level.

The significance of *CI* and *reelect* remains. A Hausman test fails to reject the null hypothesis that the estimates are consistent. Thus, the main results are not sensitive to the use of these effects.

One may also be concerned with the potential endogeneity problem of the choice of the challenger to enter the race. If, for example, a challenger enters because the incumbent is more aggressively prosecuting cases, then the results presented in Table 5 may arise not due to re-election concerns, but rather the choices made by the challenger. To address this concern a data set is constructed consisting only of the re-election contests in the 2000, 2002, 2004, 2006, 2008 elections. A dummy variable was created which equals 1 if and only if the re-election had a challenger. The objective was to assess whether the use of the courtroom explains the entrance of a challenger. Independent variables capturing the average number of jury trial convictions (in both absolute level and relative to total convictions) in the last two years of the previous term were used in both logit and probit specifications including no controls, controls for *pop*, *male*, *white*, *ur*, and *active*, and changes in these controls over the term. In every specification (sixteen in total) the use of the courtroom had no statistically significant effect on the existence of a challenger (even at the 10% level). Thus, there does not seem to be an endogeneity problem with the decision of the challenger to run being based on the decision to take cases to trial.

Also, since both the mean of the independent variable, 0.0259, and its standard deviation, 0.0147, are small one may be concerned that changes to the number of convictions from jury trials may be sensitive to the total number of

cases. Thus, to check the robustness of the results, additional variables are created to measure the prevalence of the use of jury trials. Three dummy variables are considered. The variable *above* is equal to one if the number of convictions from jury trials relative to the total number of convictions in a district in a year is greater than the mean value, or rather, if $jury > 0.0258$. It is equal to zero otherwise. The variable *abovesd* is equal to one if $jury$ is greater than one standard deviation above the mean. From Table 2 *abovesd* is, then, equal to one if and only if $jury > 0.0405$. Finally, *belowsd* is equal to one if and only if $jury$ is less than one standard deviation below the mean, or rather, $jury < 0.0112$. A logit model is used to identify the effect of the election variables on the chance of being within each of the three identified subgroups. Also, these cutoffs are used to create an ordered variable equal to zero if $jury < 0.0112$, one if $jury \in [0.0112, 0.0258)$, two if $jury \in [0.0258, 0.0405)$, and three if $jury \geq 0.0405$. An ordered logit model is estimated. Each specification in Table 7 includes both the sanction variables and the socio-economic controls.

Table 7: Logit Results ($N = 476$)

| **dep. var. =** | **ind. var. =** *CI* | *reelect* | *never* | adj $R^2$ | % correctly predicted |
|---|---|---|---|---|---|
| *belowsd* | -2.3900 ** | -0.5973 | | 0.1364 | 90.1% |
| | (1.0519) | (0.4675) | | | |
| *belowsd* | | | 0.9908 *** | 0.0994 | 91.0% |
| | | | (0.3715) | | |
| *above* | 0.6823 | 0.5127 ** | | 0.1360 | 73.1% |
| | (0.4806) | (0.2491) | | | |
| *above* | | | -0.2098 | 0.1367 | 73.7% |
| | | | (0.2638) | | |
| *abovesd* | 0.1607 | 0.2133 | | 0.1655 | 88.9% |
| | (0.5675) | (0.3333) | | | |
| *abovesd* | | | -1.2959 ** | 0.1691 | 88.9% |
| | | | (0.5072) | | |
| *ordered logit* | 0.7342 ** | 0.4261 ** | | | 53.4% |
| | (0.3396) | (0.2072) | | | |
| *ordered logit* | | | -0.6692 *** | | 56.5% |
| | | | (0.2156) | | |

* 10% level; ** 5% level; *** 1% level. QML standard errors are reported.

Having an incumbent running for re-election with a contest decreases the chance the proportion of total convictions that come from jury trials is less than one standard deviation below the mean. Being in the year before an incumbent runs for re-election increases the probability the district is above the mean in the fraction of convictions that come from jury trials. Finally, never having a contested election increases the likelihood of being substantially below the mean and decreases the chance of being substantially above the mean.

The econometric model does very well at accurately predicting the proportion of total convictions that arise from jury trials. For being below one standard deviation from the mean and being above by more than one standard deviation, the model is correct more than 88% of the time.

The specifications that include *never* also include year fixed effects. Given that some districts maintain their place in the levels over all years, district fixed effects are not included in the specifications. Matching results arise if probit models are used instead and, hence, the results are not reported here. Finally, nearly identical findings arise if the dependent variable measures whether the district is above the median (rather than the mean) and whether the district is in the top 20% along with whether it is in the bottom 20%. These results, again, are not reported here. Thus, Table 7 provides additional support for the previous results.

Furthermore, one may be interested in whether it is the primary contest or the general election contest that is motivating the prosecutor to expand the number of convictions obtained from jury trials. As shown in Table 3, a contest is just as likely to arise from a challenge within the party as coming from the other political party. Hence, $CIP$ is equal to one if the incumbent is running for re-election in that year and had a contest in the primary. $CIG$ is equal to one if the incumbent is running for re-election and has a contest in the general election. Thus, $CI = \max\{CIP, CIG\}$. Table 8 summarizes the results.

Table 8: Election Results ($N = 476$)

| | **OLS** | **OLS** | **Logit** | **Ordered** |
|---|---|---|---|---|
| dep. var. = | ***jury*** | ***jury*** | ***above*** | **Logit** |
| $CIP$ | 0.0081 ** | 0.0041 | 0.2960 | 0.7596 |
| | (0.0040) | (0.0026) | (0.7099) | (0.4932) |
| $CIG$ | 0.0014 | 0.0031 | 1.1527 ** | 0.7772 * |
| | (0.0021) | (0.0022) | (0.5769) | (0.4101) |
| *reeleect* | 0.0032 ** | 0.0025 ** | 0.5039 ** | 0.4250 ** |
| | (0.0015) | (0.0011) | (0.2501) | (0.2094) |
| controls: | | | | |
| sanction | YES | YES | YES | YES |
| socio-economic | YES | YES | YES | YES |
| district | NO | YES | NO | NO |
| adj $R^2$ | 0.2992 | 0.5870 | 0.1351 | |

* 10% level; ** 5% level; *** 1% level. Robust standard errors are reported.

Thus, it is not clear whether it is the competition within the political party or between parties that drives the results. Both seem to matter and each likely suffers from a low number of observations (the mean values of $CIP$ and $CIG$ are 0.038 and 0.029 respectively).

The theoretical model also predicts that the average sentence length a prosecutor achieves in a year should decrease when re-election pressures are greater. We propose two ways to identify this effect. First, if the prosecutor is trying to be tough and achieve longer incarcerations, then the percentage of convictions that result in community punishments as the most serious sanction should be reduced. Since community punishments are expected to arise more frequently from plea bargaining, a decrease in the percentage of convictions as community punishments is evidence that prosecutors are pursuing more severe sanctions. Second, if more cases are being taken to court and fewer cases are being plea bargained, then the marginal cases tried in court are those with weaker evidence. Such cases are expected to result in lower expected sanctions than those tried regardless of the motivation to be retained, ceteris paribus. Thus, the theory predicts that the average sanction obtained should decrease with re-election pressures.

Specifically, the following econometric models are estimated:

$$\begin{aligned} com_{dt} &= \gamma_0 + \gamma_1 CI_{dt} + \gamma_2 reelect_{dt} + \gamma_3 X_{dt} + \epsilon \\ com_{dt} &= \delta_0 + \delta_1 never_d + \delta_2 X_{dt} + \delta_3 Y_t + \epsilon \end{aligned} \tag{8}$$

and

$$\begin{aligned} max_{dt} &= \zeta_0 + \zeta_1 CI_{dt} + \zeta_2 reelect_{dt} + \zeta_3 S_{dt} + \zeta_4 X_{dt} + \zeta_5 D_t + \epsilon. \\ max_{dt} &= \eta_0 + \eta_1 never_d + \eta_2 S_{dt} + \eta_3 X_{dt} + \eta_4 Y_t + \epsilon. \end{aligned} \tag{9}$$

Additional support for incumbents, running in a contested re-election, who increase the use of jury trials to signal quality to the voting public would be provided if $\gamma_1 < 0$ & $\delta_1 > 0$ and $\zeta_1 < 0$ & $\eta_1 > 0$. Such a finding would provide evidence in support of Result 3. Table 9 presents the results.

Table 9: Effect on Sentencing ($N = 476$)

| dep. var. = | ***com*** | ***com*** | ***max*** | ***max*** |
|---|---|---|---|---|
| *CI* | -0.0213 ** | | -2.6407 * | |
| | (0.0098) | | (1.3575) | |
| *reeleect* | 0.03383 *** | | 1.4316 * | |
| | (0.0070) | | (0.7805) | |
| *never* | | 0.0248 *** | | 0.0985 |
| | | (0.0054) | | (0.9148) |
| *agg* | | | 40.418 *** | 40.142 *** |
| | | | (7.954) | (13.053) |
| *mit* | | | -2.8315 | 24.681 *** |
| | | | (4.848) | (4.608) |
| *male* | 0.5519 * | 0.4566 * | | |
| | (0.2831) | (0.2398) | | |
| *white* | 0.0844 *** | 0.1462 *** | | |
| | (0.0197) | (0.0194) | | |
| $\%16-24$ | -0.0045 | 0.3897 *** | | |
| | (0.1438) | (0.1211) | | |
| *density* | -0.00002 * | 0.00002 * | -0.0090 | 0.0041 *** |
| | (0.00001) | (0.00001) | (0.0104) | (0.0015) |
| *ur* | -0.3741 *** | 0.9470 *** | | |
| | (0.1249) | (0.2123) | | |
| district controls | NO | NO | YES | NO |
| year controls | NO | YES | NO | YES |
| adj $R^2$ | 0.1285 | 0.2899 | 0.4394 | 0.1425 |

* 10% level; ** 5% level; *** 1% level. Heteroskedastic-robust standard errors are reported.

The predictions of the theoretical model are supported by these results. If a district has an incumbent running for re-election and a challenger has entered the race, the number of convictions that result in community punishments decreases by $0.0213/0.2037 = 10.5\%$ (where 0.2037 is the mean proportion of convictions resulting in community sanctions as shown in Table 2). Similarly, the change to the average maximum sanction decreases, as would be predicted with prosecutors increasing the number of cases they take to trial by taking cases which receive lower than average sanctions. The average maximum active sanction decreases by $2.6407/42.466 = 6.2\%$ in the year an incumbent runs in a contested election.

Never having had a contested election between the 1998 and 2008 elections increases the proportion of convictions that have community punishments, while

it does not have a statistically significant effect on the sanctions for active punishments. Interestingly, in the year before a re-election campaign there is an increased use of community punishments and there is an increase in the average maximum active sanction. This could be explained, for example, by a forward-looking incumbent dispensing with weaker cases in the year prior to the re-election by encouraging pleas to community punishments so that in the re-election year resources can be devoted to the achievement of tougher sentences through trials.

As before, the sanction variables and the socio-economic controls are valuable. The use of community punishments is related to the gender and racial composition of the districts. Other than the population density, the socio-economic variables do not affect the average maximum sanction. An $F$-test of the joint null hypothesis that they have no effect on $max$ fails to be rejected. The length of the incarceration is influenced by whether the judge assigns an aggravated or mitigated punishment. Obviously, the use of the aggravated sentences increases the size of the maximum punishment. Interestingly, the use of mitigated sentences also increases the size of the maximum punishment. This could be explained by judges applying mitigated sentences to those with greater prior record scores or those who commit more serious offenses.

Thus, it seems that the contested incumbent influences which cases go to trial. Community punishments are used less and, with more marginal cases being taken to trial, the average sanction reduces.

# 6 Conclusion

In the U.S. prosecutors exercise a significant amount of discretion. They decide whether to file charges, which charges to file, and whether to proceed to trial or plea bargain. How they exercise this discretion affects the people they serve. Understanding how the institutions used affect prosecutor's choices and the outcomes of the criminal justice system is crucial to developing the best mechanisms to determine guilt and innocence and protect society. There is a serious void in the formal analysis of the decisions of prosecutors. Here a particular decision is analyzed, whether to take a case to trial or engage in plea

bargaining. We investigate how retention motivations, specifically the desire to win re-election and remain in office, affects this decision. A theoretical model of asymmetric information is developed to explore how trial outcomes can be used as a signal to the voting public to make the retention decision. The theory predicts that when re-election pressures are high prosecutors increase the number of cases taken to trial and plea bargain less. Data from all forty-three districts in North Carolina over twelve years provides empirical verification.

The theoretical model developed assumes the voters use aggregate sanctions as the metric to base the retention decision on. Recent theoretical work has considered conviction rates as a potential measurement of prosecutorial performance (Bandyopadhyay and McCannon, 2010; Rasmusen, Raghav, and Ramseyer, 2009). As shown in Bandyopadhyay and McCannon (2010), the use of such a metric would cause distortions that decrease the number of trials. Thus, the empirical results falsify this theory's application to prosecutors in North Carolina.

The results stongly suggest that election pressures affect the decision to take a case to trial. Whether or not this in fact suboptimal, as predicted by the theory, depends on whether alternative explanations can be ruled out. One may, instead, argue that the retention concerns induce additional effort. If prosecutors are motivated to work harder, then one may expect more cases to be taken to trial since significantly more time, effort, and resources must be expended. It seems reasonable to presume, though, that the average sanction achieved in trial would increase with the additional expended resources on investigation and preparation for the cases. Hence, the result that the average maximum sanction obtained decreases with re-election pressures can be argued to work against an explanation of incentives to exert effort and provides additional support for the asymmetric information explanation provided. Furthermore, the North Carolina Sentencing and Policy Advisory Commission also publishes data on the number of pending felony cases in each district between 1999-00 and 2008-09. Our preliminary analysis of this data reveals that the average number of pending cases in the year before a re-election campaign is 1377.7, while in all other years it is only 1260.9. This is a 9.3% *increase* in the number of unresolved cases.[23]

[23] The correlation coefficient between the two variables is positive. The null hypothesis that

The number of pending cases coming up to the re-election represents a 5.6% increase from the number of pending cases from the previous year. Again, this is suggestive of re-election concerns not inducing an improved amount of effort, but rather incumbents strategically taking more cases to trial to be retained, which increases the number of convictions from jury trials and (potentially) increases the number of unresolved cases. Our future work, though, will focus on the issues of case backlogs and the incentives provided by elections.

# 7 References


Bandyopadhyay, Siddhartha and Bryan C. McCannon (2010), Prosecutorial Retention: Signaling By Trial, Working Paper, University of Birmingham.

Besley, Timothy (2006), *Principled Agents? The Political Economy of Good Government*, Oxford: Oxford University Press.

Boylan, Richard T. (2005), What Do Prosecutors Maximize? Evidence from Careers of U.S. Attorneys, *American Law and Economics Review* 7(2), 379-402.

Burke, Alafair (2007), Prosecutorial Passion, Cognitive Bias, and Plea Bargaining, *Marquette Law Review* 91(1), 183-211.

Forst, Brian and Kathleen B. Brosi (1977), A Theoretical and Empirical Analysis of the Prosecutor, *Journal of Legal Studies* 6(1), 177-191.

Garoupa, Nuno and Frank Stephen (2006), Law and Economics of Plea Bargaining, Working Paper.

Glaesar, Edward L., Daniel P. Kessler, and Anne Morrison Piehl (2000), What Do Prosecutors Maximize? An Analysis of the Federalization of Drug Crimes, *American Law and Economics Review* 2(2), 259-290.

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


the total number of pending cases and the variable *reelect* has zero correlation can be rejected.

Hanssen, F. Andrew (2000), Independent Courts and Administrative Agencies: An Empirical Analysis of the States, *Journal of Law, Economics & Organization* 16(2), 534-571.

Landes, William M. (1971), An Economic Analysis of the Courts, *Journal of Law and Economics* 14(1), 61-107.

Leaver, Clare (2009), Bureaucratic Minimal Squawk Behavior: Theory and Evidence from Regulatory Agencies, *American Economic Review* 99(3), 572-607.

Mellon, Leonard R., Joan E. Jacoby, and Marion A. Brewer (1981), The Prosecutor Constrained by His Environment: A New Look at Discretionary Justice in the United States, *Journal of Criminal Law and Criminology* 72(1), 52-81.

Perry, Steven W. (2006), *Prosecutors in State Courts, 2005*, Bureau of Justice Statistics, U.S. Department of Justice.

Rasmusen, Eric, Manu Raghav, and Mark Ramseyer (2009), Convictions Versus Conviction Rates: The Prosecutor's Choice, *American Law and Economics Review* 11(1), 47-78.

Reinganum, Jennifer F. (1998), Plea Bargaining and Prosecutorial Discretion, *American Economic Review* 78(4), 713-728.

Reinganum, Jennifer F. (2000) Sentence Guidelines, Judicial Discretion, and Plea Bargaining, *RAND Journal of Economics* 31(1), 62-81.

Rogoff, Kenneth S. (1990), Equilibrium Political Budget Cycles, *American Economic Review* 80(1), 21-36.

Rogoff, Kenneth S. and Anne Siebert (1988), Elections and Marcoeconomic Policy Cycles, *Review of Economic Studies* 55(181), 1-16.

Simmons, Ric (2004), Election of Local Prosecutors, *ElectionLaw@Moritz*, http://moritzlaw.osu.edu/electionlaw/ebook/part7/elections_prosecutors.html.

Shepherd, Joanna M. (2009), The Influence of Retention Politics on Judges' Voting, *Journal of Legal Studies* 38(1), 169-206.

Wright, Ronald F. (2009), How Prosecutor Elections Fail Us, *Ohio State Journal of Criminal Law* 6(2), 649-660.

# 8 Appendix: Not for Publication

Table $A1$ presents the correlation coefficients between year dummy variables and the two primary variables of interest, $CI$ and $reelect$. As one would expect, there is a substantial amount of correlation between them.

Table $A1$: Correlations

| | ***CI*** | | ***reelect*** |
|---|---|---|---|
| **1998−99** | 0.2937 *** | **1997−98** | 0.3740 *** |
| **2002−03** | 0.2005 *** | **2001−02** | 0.3567 *** |
| **2006−07** | 0.2224 *** | **2005−06** | 0.2874 *** |

* 10% level; ** 5% level; *** 1% level.

As stated in the data description four of the original thirty-nine districts where divided during the sample period. This created an unbalanced panel with nine years having thirty-nine observations, two years having forty-one observations, and one year having forty-three districts. To ensure that these divisions do not affect the main message of the paper a balanced panel dataset is created eliminating these districts. Table $A2$ presents results using only the sample from the districts that remained in tact.

Table $A2$: Balanced Panel (dep. var. = $jury$, $N = 432$)

| | **I** | **II** | **III** |
|---|---|---|---|
| $CI$ | 0.0051 ** | 0.0032 * | |
| | (0.0026) | (0.0018) | |
| $reelect$ | 0.0027 * | 0.0023 ** | |
| | (0.0015) | (0.0012) | |
| $never$ | | | -0.0049 *** |
| | | | (0.0012) |
| controls: | | | |
| socio-economic | YES | YES | YES |
| sanction | YES | YES | YES |
| year | NO | NO | YES |
| district | NO | YES | NO |
| adj $R^2$ | 0.2998 | 0.5943 | 0.3682 |

* 10% level; ** 5% level; *** 1% level. Heteroskedastic-robust standard errors are reported.

The magnitude of the coefficients and their statistical significance is unaltered from those presented in Table 5.

Define *trial* as the total number of convictions that arise in a district in a year from jury trials, *total* as the total number of convictions obtained, and *plea* as the total number of convictions that arise in a district in a year from guilty pleas. Hence, $total = trial + plea$ and $jury = \frac{trial}{total}$. The total population, unemployed, and labor force are included as control variables.

Table $A3$: Results ($N = 476$)

| | ***trial*** | ***trial*** | ***plea*** | ***plea*** |
|---|---|---|---|---|
| $CI$ | 2.2837 ** | | -2.5895 ** | |
| | (1.0885) | | (1.1421) | |
| $reelect$ | 0.1869 | | -0.3951 | |
| | (0.6267) | | (0.6557) | |
| $never$ | | -0.8687 | | -0.6292 |
| | | (0.7111) | | (1.5874) |
| $total$ | 0.0041 | 0.0195 *** | 0.9911 *** | 0.9816 *** |
| | (0.0034) | (0.0017) | (0.0049) | (0.0020) |
| $population$ | 0.0002 ** | 0.00006 * | -0.0002 *** | -0.0001 ** |
| | (0.0001) | (0.00003) | (0.0001) | (0.00004) |
| $density$ | -0.0454 * | | 0.0474 ** | -0.0112 *** |
| | (0.0236) | | (0.0238) | (0.0031) |
| $unemployed$ | -0.0002 | -0.0004 | -0.0001 | 0.0006 * |
| | (0.0002) | (0.0003) | (0.0002) | (0.0003) |
| $labor\ force$ | -0.0002 * | -0.0001 * | 0.0003 ** | 0.0001 * |
| | (0.0001) | (0.0001) | (0.0002) | (0.0001) |
| controls: | | | | |
| socio-economic | YES | YES | YES | YES |
| sanction | YES | YES | YES | YES |
| year | NO | YES | NO | YES |
| district | YES | NO | YES | NO |
| adj $R^2$ | 0.8283 | 0.7194 | 0.9996 | 0.9994 |

* 10% level; ** 5% level; *** 1% level. Heteroskedastic-robust standard errors are reported.

Thus, the total number of convictions from jury trials increases by $2.2837/17.662 = 12.9\%$ when a challenger enters the race, which coincides with the previous estimation of $9.7 - 14.7\%$. Table $A3$ also more directly verifies Result 1 of the theory.

Finally, while the lemma, propositions, and the first two results follow directly from the derivations presented in the text, Result 3 may not seem entirely obvious. Thus, a formal proof is provided.

**Result 3** *The average sanction obtained in jury trials is greater than the first-best outcome, as is the average guilty plea.*

**Proof.** The average expected sanction derived from jury trials is

$$\frac{\int_{\theta_1}^{\theta_M} p(\theta)\, s(\theta)\, dF(\theta)}{1 - F(\theta_1)}$$

The derivative with respect to $\theta_1$ is

$$
\begin{aligned}
& -\left(\int_{\theta_1}^{\theta_M} p(\theta)\, s(\theta)\, f(\theta)\, d\theta\right)(1-F(\theta_1))^{-2}(-f(\theta_1)) + (1-F(\theta_1))^{-1} \\
& \text{x } (-p(\theta_1)\, s(\theta_1)\, f(\theta_1)) \\
= & \frac{\left(\int_{\theta_1}^{\theta_M} p(\theta)\, s(\theta)\, f(\theta)\, d\theta\right) f(\theta_1)}{(1-F(\theta_1))^2} - \frac{p(\theta_1)\, s(\theta_1)\, f(\theta_1)}{(1-F(\theta_1))} \\
= & \frac{f(\theta_1)}{(1-F(\theta_1))^2}\left(\int_{\theta_1}^{\theta_M} p(\theta)\, s(\theta)\, f(\theta)\, d\theta - p(\theta_1)\, s(\theta_1)\,(1-F(\theta_1))\right) \\
= & \frac{f(\theta_1)}{(1-F(\theta_1))^2}\left(\int_{\theta_1}^{\theta_M} p(\theta)\, s(\theta)\, f(\theta)\, d\theta - \int_{\theta_1}^{\theta_M} p(\theta_1)\, s(\theta_1)\, f(\theta)\, d\theta\right) \\
= & \frac{f(\theta_1)}{(1-F(\theta_1))^2}\int_{\theta_1}^{\theta_M} [p(\theta)\, s(\theta) - p(\theta_1)\, s(\theta_1)]\, f(\theta)\, d\theta
\end{aligned}
$$

Since it is assumed that $\frac{dpS}{d\theta} > 0$, $p(\theta)\, s(\theta) > p(\theta_1)\, s(\theta_1)$ for $\theta > \theta_1$. Hence, $\int_{\theta_1}^{\theta_M} [p(\theta)\, s(\theta) - p(\theta_1)\, s(\theta_1)]\, f(\theta)\, d\theta > 0$. Since $f(\theta_1)/(1-F(\theta_1))^2 > 0$ as well, if $\theta_1$ increases, then so to does the average expected sentence. Therefore, since $\phi_L$, the upper bound to the set of $\theta_H^e$ that can be supported as a separating equilibrium (Proposition 1), was shown to be less than $\widetilde{\theta}_H$, then it follows that the average expected sanction is less than in the first-best outcome for every separating equilibrium. Also, since all $\theta_H^e$ and $\theta_L^e$ that can be supported in a pooling equilibrium are less than or equal to $\widetilde{\theta}_H$ (Proposition 2), then the average expected sanction is less for every pooling equilibrium. Consequently, the average sanction obtained in jury trials is greater than the first-best outcome. The result for plea bargains follows analogously. ■