doi:10.1017/S0003055409990207

Assessing Partisan Bias in Federal Public Corruption Prosecutions

SANFORD C. GORDON *New York University*

The 2007 U.S. Attorney firing scandal raised the specter of political bias in the prosecution of officials under federal corruption laws. Has prosecutorial discretion been employed to persecute enemies or shield allies? To answer this question, I develop a model of the interaction between officials contemplating corruption and a prosecutor deciding whether to pursue cases against them. Biased prosecutors will be willing to file weaker cases against political opponents than against allies. Consequently, the model anticipates that in the presence of partisan bias, sentences of prosecuted opponents will tend to be lower *than those of co-partisans. Employing newly collected data on public corruption prosecutions, I find evidence of partisan bias under both Bush (II) and Clinton Justice Departments. However, additional evidence suggests that these results may understate the extent of bias under Bush, while overstating it under Clinton.*

In the United States, it is not only corrupt conduct by elected officials, but also its detection and prosecution that may have significant political consequences. On the one hand, officials may undertake illegal activities to further their political fortunes. For example, an incumbent might seek to further his or her political career by accepting a bribe, rigging a competitive bidding process to secure the support of an influential private citizen, or diverting public resources into his or her reelection campaign. At the same time, the prosecution of corrupt practices—whether real or merely alleged—can also produce political fallout. An investigation or indictment can irreversibly harm the future career prospects of the official under scrutiny, even if that individual is eventually exonerated.

Discretion over whether to bring forward a criminal case lies largely with the prosecutor.[1] In the enforcement of federal public corruption laws, this authority rests almost entirely in the hands of U.S. Attorneys in the Department of Justice (DOJ). U.S. Attorneys are political appointees, raising the possibility that this discretion might be deployed in public corruption cases on the basis of partisan political motivations. If U.S. Attorneys are motivated to embarrass or discredit political opponents through an indictment, they might decide to bring relatively weak cases they would otherwise decline. Alternatively, those officials might decline to prosecute co-partisans except in the most egregious cases. On the other hand, federal prosecutors might be limited in the degree to which they can act on partisan motivations even to the extent that they harbor them, whether by formal institutional constraints or by informal norms of professional ethical conduct.

The issue of partisan bias in federal corruption prosecutions became the subject of public attention in 2007, when newly empowered Democratic congressional majorities convened hearings to investigate the dismissal of a number of U.S. Attorneys by the administration of George W. Bush in the previous two years. Critics have contended that some of these officials may have been dismissed either because they did not pursue corruption investigations against prominent Democrats with sufficient vigor or because they did pursue investigations against prominent Republicans (Goldstein and Eggen 2007; Lipton 2007a). Others have pointed to high-profile prosecutions of Democrats, such as Alabama Governor Donald Siegelman, as evidence of politicized prosecution (Lipton 2007b).

Moving beyond these high-profile cases, have Bush-appointed U.S. Attorneys systematically targeted Democratic public officials or avoided targeting Republicans for public corruption prosecutions? To the extent that partisan motivations have influenced such prosecutions, has the phenomenon been unique to the Bush administration? This article proposes an approach to address these questions systematically. First, I evaluate the conclusions of a recent, widely cited study suggesting that Democrats were up to seven times more likely to be investigated by the Bush Justice Department than were Republicans (Shields and Cragan 2007). I argue that our ability to infer partisan bias by the Justice Department from the study's data on publicized investigations is sharply limited. Because U.S. Attorneys nearly always prosecute cases that have been referred by other investigative agencies such as the Federal Bureau of Investigation (FBI) or state and local law enforcement, the absence of information (owing to privacy concerns) on the partisan breakdown of referrals *not* taken up by prosecutors threatens to invalidate the authors' claims of selective prosecution.

Second, I develop a simple model of prosecutorial discretion in the presence of (possible) partisan bias. Employing an approach from economic models of discrimination, I treat partisan bias as a "taste" or

Sanford C. Gordon is Associate Professor, Wilf Family Department of Politics, New York University, 19 W. 4th Street, New York, NY 10012 (sanford.gordon@nyu.edu).

The author gratefully acknowledges the helpful suggestions of James Alt, Nathaniel Beck, Catherine Hafer, Gregory Huber, Dimitri Landa, Andrew Gelman, James Jacobs, Nicola Persico, Anne Morrison Piehl, Howard Rosenthal, Alastair Smith, Andrew Whitford, three anonymous reviewers, and the editors. Earlier versions of this article were presented at the 2007 Conference on Empirical Legal Studies at NYU Law School and in seminars at Harvard's Center for American Political Studies, the NYU Department of Politics, and the Harris School of Public Policy at The University of Chicago, where the author received valuable comments. Angela Zhu and Leslie Huang provided excellent research assistance.

[1] For two recent theoretical examinations of prosecutorial discretion and its implications for political accountability, see Gordon and Huber (2002) and Shotts and Wiseman (2008). For a review of the literature on prosecutorial discretion in general, see Gordon and Huber (2009).

preference for prosecuting one's political opponents (or for not prosecuting allies). This approach, pioneered by Becker (1957), has been employed recently to study discrimination against minorities in setting bail (Ayres 2001; Ayres and Waldfogel 1994), racial profiling (Knowles, Persico, and Todd 2001), and discrimination against female candidates in congressional elections (Anzia and Berry 2007). In the model, the prosecutor receives a signal of the official's culpability (the referral) and decides whether to pursue the case. A biased prosecutor's relative enthusiasm for pursuing cases against political opponents will lead her to prosecute weaker cases than she would against her allies. The empirical implication of this behavioral effect is that in the presence of partisan bias, observed sentences for the prosecutor's partisan opponents should be *lower* on average than those of the prosecutor's co-partisans. Critically, because the empirical consequences of partisan bias concern the outcomes of cases *actually prosecuted*, the approach taken here does not require knowledge of the partisan breakdown of all referrals, and thus sidesteps the selection problem noted previously.

The validity of this approach to detecting partisan bias among prosecutors *within a given administration*, however, is threatened by the existence of confounding influences on sentencing. In particular, underlying interparty differences in the severity of the typical corruption case might lead to differences in observed average sentences even in the absence of prosecutorial bias. To circumvent this obstacle to inference, my identification strategy exploits additional information emerging from the comparison of sentences handed down *during adjacent administrations of different parties*. First, patterns of disproportionate prosecution of defendants from one party that persist across administrations may permit the observer to determine the direction of the confounding effect of underlying partisan differences in case severity. For example, disproportionate prosecution of Democratic defendants under both Democratic and Republican administrations would be consistent with more severe corruption on the part of Democrats than Republicans; in that case, evidence of partisan bias based on differences in average sentences would tend to understate the true extent of bias under a Republican administration and overstate it under a Democratic one. Second, differences-in-differences estimation can recover a measure of total partisan bias across those administrations free from the effect of these confounding influences, although it does not permit one to apportion the bias between the administrations.

To implement this strategy, I employ newly collected data on the outcomes of a sample of national priority public corruption prosecutions by U.S. Attorneys in the Bush Justice Department from 2004 to 2006 and the Clinton Justice Department from 1998 to 2000. I find evidence of bias in *both* samples: under Bush, a Republican president, Republican-affiliated defendants tended to receive sterner sentences than Democrats, while under Clinton, a Democratic president, it was Democratic-affiliated defendants who tended to receive harsher punishments. These patterns persist even after controlling for the region in which the prosecution took place, the partisanship of the president who appointed the sentencing judge, and other features of defendants and cases that might affect the incentives of prosecutors. I also find evidence that defendants with identifiable partisan affiliations are disproportionately Democratic under both administrations. This suggests that the findings concerning disparities in sentencing might understate the actual level of partisan bias under Bush (i.e., the true level of bias under Bush was higher than measured), while overstating partisan bias under Clinton (i.e., the true level of bias under Clinton was lower than measured). I conclude by considering the implications of these findings for our understanding of the autonomy and politicization of government agencies.

BACKGROUND

Official Corruption Abroad and in the United States

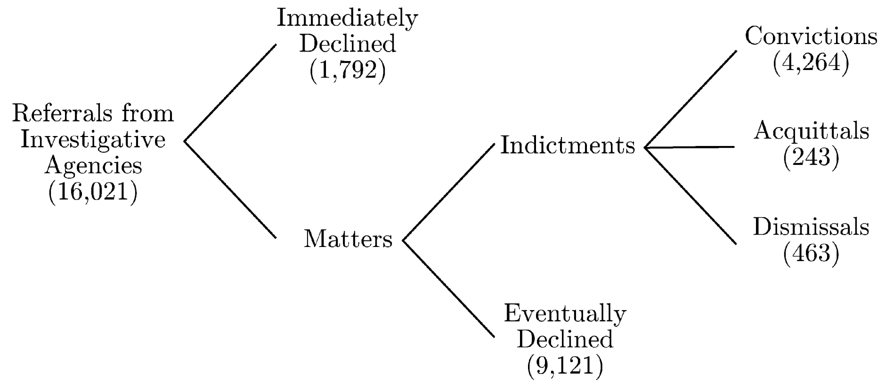
Official corruption, conventionally defined as the use of public office—elected or otherwise—for private gain (e.g., Nye 1967), is the subject of a fertile research agenda in empirical and theoretical political economy. Scholars working in the comparative context have examined the causal antecedents of corruption, examining its sociodemographic (Lipset, Seong, and Torres 1993), cultural (Rose-Ackerman 1999; Treisman 2000), and institutional (Brown, Touchton, and Whitford 2006; La Porta et al. 1999; Persson, Tabellini, and Trebbi 2003) origins.

Corruption has also been the subject of scrutiny for numerous observers of American politics. In Volume II of *The American Commonwealth*, Lord Bryce endeavors to assess the prevalence of bribery, kickbacks, and misappropriation of public funds in different branches of federal, state, and local governments (1995 [1888], 832–3). More recently, scholars of domestic governance have sought to explain variation among the states in perceptions of corruption (Boylan and Long 2003) and in federal corruption prosecutions and conviction rates. Goel and Nelson (1998) find higher corruption conviction rates in states with larger government expenditures. Meier and Holbrook (1992) detect higher conviction rates in more urban states with less educated citizenries, larger fractions of citizens with Irish or Italian ancestries, lower voter turnout, and higher total government employment. Alt and Lassen (2002) find higher perceived levels of corruption in states with closed primaries, restrictive ballot access for voter initiatives, and loose restrictions on campaign finance.

Prosecution of Public Corruption in the United States

The Justice Department prosecuted nearly 5,000 defendants classified as state or local "corrupt government officials" from 1986 to 2008 (Transactional Records Access Clearinghouse [TRAC], Syracuse University). The

FIGURE 1. Disposition of State and Local Corruption Referrals from Investigative Agencies by U.S. Attorneys' Offices, 1986–2008



Note: Outcomes not depicted include referrals in which charges were eventually folded into other cases, or in which charges were dropped and then refiled.

lead charge in 25% of the indictments was violation of 18 U.S.C. 1951, the Hobbs Act, which prohibits robbery or extortion affecting interstate commerce. The next most common offense was under 18 U.S.C. 666 (19% of prosecutions), which outlaws theft and bribery in entities receiving more than $10,000 in federal funds. Other commonly employed corruption statutes included the mail fraud statute (18 U.S.C. 1341, 12% of cases), the statute governing conspiracies to defraud the federal government (18 U.S.C. 371, 9% of cases), and the Racketeer Influenced and Corrupt Organizations statute (18 U.S.C. 1962, 4% of cases).

Although the 93 U.S. Attorneys and their assistants retain significant discretion over whether to bring a case (cf. Richman 1999), they do not operate in a vacuum.[2] Corruption cases nearly always begin with investigations by other agencies, which may or may not end in referrals to a U.S. Attorney's Office. Of the approximately 16,000 referrals to the Justice Department concerning state and local corruption cases from 1986 to 2008, a large majority (82%) came from the FBI. By comparison, the next largest source, the Internal Revenue Service (IRS), accounted for only 3% of referrals. Figure 1 outlines the typical progression of a case once it is referred to a U.S. Attorney's Office, along with the number of referrals by outcome. The office can drop the referral outright. If it does not, the case becomes a "matter," typically developed by an Assistant U.S. Attorney and the referring agency. The office can subsequently decline the case or bring it before a grand jury to bring an indictment. (It is exceedingly rare for a grand jury to refuse to indict.) Roughly one third of referrals (4,970 cases) eventually resulted in indictments. Of the indictments, 86% (4,264 cases) resulted in convictions (roughly four fifths of which were guilty pleas, and the remainder guilty verdicts), 5% (243 cases) in jury verdicts of not guilty, and 9% (463 cases) in dismissals.

The use of law enforcement generally, and federal law enforcement in particular, as a means to reduce public corruption at the state and local levels is largely an innovation of the second half of the 20th century (Anechiarico and Jacobs 1994). Not surprisingly, this development has been accompanied by concerns that the discretion to investigate and prosecute officials for violating anticorruption statutes could be abused for partisan gain (Abrams and Beale 2006).

Politicization and Bias

Political scientists and public administration scholars have long recognized efforts by presidents to influence bureaucratic policy making and performance, especially through the placement of loyalist appointees in key positions of authority (e.g., Lewis 2008; Moe 1985; Nathan 1975). At the same time, other scholars of bureaucracy have pointed to the success of bureaucratic entrepreneurs in cultivating a degree of policy-making autonomy from their political principals (Carpenter 2001) or advancing "strategically neutral" (i.e., self-consciously apolitical) procedures (Huber 2007).

The possibility that law enforcement might be politicized to serve the president's political aims, and the institutional resilience of the Justice Department against such efforts, has been a recurrent theme throughout the department's history. Throughout much of the 20th century, "Main Justice" (the department's Washington, DC, headquarters) sought to centralize control over the far-flung U.S. Attorneys, who had been largely politically autonomous since the office's creation in 1789 (Eisenstein 1978). Whitney North Seymour, Jr., himself a former U.S. Attorney, describes the electoral motivations behind the Nixon Justice Department's prioritization of narcotics prosecution (1975, 119–35,

[2] Occasionally, if a case is especially sensitive or involves defendants from multiple districts, the Public Integrity Section at the DOJ headquarters will prosecute a case in the federal courts.

226). More recently, Whitford (2002) finds evidence that U.S. Attorneys' aggregate caseloads respond to national political trends. Meier and Holbrook (1992) find some evidence that public officials from Democratic localities were disproportionately targeted under Reagan, whereas Heidenheimer (1989, 583–4) discusses anecdotal evidence of disproportionate targeting under Nixon and Ford.

The issue of politicization of law enforcement achieved most recent salience in April 2007, when newly elected Democratic majorities in Congress convened hearings to investigate the dismissal of seven U.S. Attorneys on December 7, 2006, as well as several others in the preceding two years. Critics of the Bush administration contended that a failure to employ public corruption prosecutions for partisan ends lay in part behind the firings, mustering anecdotal evidence to support the claim. One fired attorney, Carol Lam (U.S. Attorney for Southern California), had indicted San Diego Republican Congressman Randy "Duke" Cunningham. A January *New York Times* editorial posited that the Cunningham investigation had also led her to investigate Republican Representative Jerry Lewis, to the consternation of Republican officials. Another attorney, David Iglesias (New Mexico), testified before the Senate Judiciary Committee that he had received telephone calls from Senator Pete Domeneci and Representative Heather Wilson (both New Mexico Republicans) inquiring about progress in corruption investigations of prominent New Mexico Democrats. A third, John McKay (Western Washington), testified that the Chief of Staff of Representative Doc Hastings (R-WA) had called to inquire about federal investigations into voter fraud following the 2004 Washington gubernatorial election. A fourth, Thomas M. DiBiagio (Maryland), in an interview with the *New York Times*, claimed his dismissal was the result of his investigation of Republican Governor Robert L. Ehrlich, Jr. (Lichtblau 2007). Most recently, a joint report of the DOJ's Office of Professional Responsibility and Inspector General references allegations that the Attorney General's Chief of Staff, Kyle Sampson, removed U.S. Attorney for Eastern Wisconsin Steven Biskupic from a list of potential targets for termination following his prosecution of a Democratic public official (USDOJ 2008, 19).

EVIDENCE OF DISPROPORTIONALITY AS EVIDENCE OF BIAS

To move beyond anecdotal claims of partisan bias in corruption investigations and prosecutions requires a more systematic approach. One such approach would be simply to compare the number of prosecutions of Democratic officials to that of Republican officials. At the time the U.S. Attorney scandal was beginning to break, two communications professors, Donald C. Shields and John F. Cragan, had been compiling data on Justice Department corruption investigations from 2001 through 2006. They published their findings in February in the online journal *E Pluribus Media*, garnering substantial coverage in the media, including a mention in Paul Krugman's *New York Times* column and an appearance by Shields on the satirical *Colbert Report*. Their main finding was that U.S. Attorneys "investigate seven (7) times as many Democratic officials as they investigate Republican officials, a number that exceeds even the racial profiling of African Americans in traffic stops." They refer to the disparity as "political profiling." The study does not describe the methodology employed to obtain the sample of investigations; however, in an interview with *E Pluribus Media*, Shields defined the unit of observation as an instance in which "the popular press (newspapers or television stations) ran a story naming a particular elected official or candidate as being under investigation by the Department of Justice (DOJ) under the direction of a U.S. Attorney-led grand jury or U.S. Attorney's Office."[3] Shields updated the study for testimony before the House Judiciary Committee on October 23, 2007.

To evaluate this study and its claim of disproportionate targeting, one must first define the quantity of interest to be measured. Most immediately relevant is the difference between the probability a Democrat is targeted for investigation and the probability a Republican is so targeted:

$$\Pr(target \mid Dem) - \Pr(target \mid Rep). \qquad \textbf{(1)}$$

This difference is logically bounded between −1 (all Republicans and no Democrats targeted) and 1 (all Democrats and no Republicans targeted). Unfortunately, the approach of Shields and Cragan has significant drawbacks and does not permit recovering a reliable estimate of this difference.[4] First, use of media accounts of grand jury investigations can easily generate selection bias in the authors' sample. These investigations are typically secret, and the *U.S. Attorneys' Manual* directs that no DOJ personnel may respond to questions about an ongoing investigation except in unusual circumstances (1 U.S.A.M. 7.530). In the interview, Shields maintains that the data could imply selective leaking by the Justice Department, but we cannot be certain that any investigation that becomes public does so through a leak by the Justice Department.

Second, the Shields/Cragan study is implicitly premised on the notion that U.S. Attorneys initiate investigations of public officials *de novo*. Federal prosecutors, however, are limited by the referrals they receive from investigating agencies such as the FBI, the IRS, and state and local law enforcement. To estimate the quantity of interest in Equation (1), it is necessary to know the distribution of Democrats and Republicans in the pool of referrals—including those *not* pursued by the U.S. Attorneys' Offices. Unfortunately, the identity of the subjects of declined referrals is strictly confidential, so it would be extremely challenging to determine their partisanship even if one were inclined

[3] www.epluribusmedia.org/features/2007/20070425_donald_shields_interview.html (Accessed July 14, 2009).

[4] Mosedale (2007) also takes issue with coding problems in the data not discussed here.

to do so. It is, further, erroneous to assume that the partisan distribution of referrals mirrors the partisan distribution in the general population, insofar as underlying differences may exist in the rates at which members of each party engage in corrupt activity.[5]

In light of these issues, what inferences can one make from the study's data? Incorporating some additional information not employed in the study, and with no additional assumptions about the partisan distribution of all referrals to the Justice Department, we can calculate logical bounds (Manski 1995) on the difference in Equation (1). Based on figures compiled by the Executive Office for U.S. Attorneys and summarized by TRAC, we know that U.S. Attorneys disposed of 10,043 corruption referrals (federal, state, and local) from investigative agencies from 2001 to 2006. The Shields/Cragan data identify 298 investigated Democrats, 67 Republicans, and 10 Independents.[6] We therefore know that the total number of referred Democrats could range from 298, which would imply a 100% rate of Justice Department investigation of Democrats, to $10{,}043 - 67 - 10 = 9{,}966$, which would imply an investigation rate of 3%. Likewise, the number of Republicans referred could range from 67 (a 100% rate of investigations of Republicans) to 9,735 (a rate of 0.7%). Thus, all we can learn from these data is that $\Pr(target \mid Dem) - \Pr(target \mid Rep)$ must fall somewhere between -0.97 and 0.993.[7,8] In other words, the authors' data do not, by themselves, warrant conclusions about disproportionate targeting, to say nothing of partisan motivation. This indeterminacy suggests the value of an alternative approach.

A MODEL OF DISCRETION IN PUBLIC CORRUPTION PROSECUTIONS

The challenge of inferring bias from aggregate descriptive statistics has parallels in the study of racial and gender discrimination. Accordingly, I adopt a strategy employed in the literature on the economics of discrimination. Echoing an approach first advanced by Becker (1957), I treat partisan bias on the part of prosecutors as a "taste for discrimination," which may take the form either of an additional benefit associated with prosecuting opponents or an additional cost associated with prosecuting co-partisans. In addition to any partisan motivations, I assume that U.S. Attorneys prefer prosecuting "big" cases (i.e., crimes publicly perceived as more severe). Career concerns may motivate them to pursue high-profile cases (see, e.g., Boylan 2005; Glaeser, Kessler, and Piehl 2000); alternatively, they may simply be motivated to root out the most egregious offenders.

The simple model described in this section has two immediate benefits. First, it yields empirical implications concerning the *outcomes* of public corruption prosecutions, and thus bypasses the problem of uncertainty regarding the partisan composition of the full set of referrals by other agencies to U.S. Attorneys' Offices. A focus on outcomes echoes several recent studies of racial bias in law enforcement. Ayres and Waldfogel (1994), for example, argue that if judges act from racial bias when setting bail for defendants, then African-American defendants will be systematically "over-deterred," and consequently pose less of a flight risk, than their white counterparts. Similarly, Knowles, Persico, and Todd (2001) assess racial bias in traffic stops and searches, arguing that if police officers have an incentive to detect drugs, racial bias will be in evidence if African-American motorists are less likely, conditional on a search, to be carrying drugs.

Second, the model permits a precise examination of a fundamental identification problem: because partisan affiliation of defendants is not randomly assigned, differences in observed case outcomes may be attributable to underlying differences between the parties in the distribution of case characteristics, rather than to prosecutorial partisan bias. The model provides insights concerning what can be learned about bias in the presence of these confounding effects and establishes conditions under which it is possible to assign a direction to the bias associated with them.

Primitives and Equilibrium

There is a prosecuting attorney a and N public officials, indexed by $i = 1, \ldots, N$.[9] The prosecuting attorney is characterized by her partisan affiliation $p_a \in \{D, R\}$.[10] A public official is described by his partisan affiliation $p_i \in \{D, R\}$ and his type $m_i \in \mathbb{R}$.[11] The partisan affiliations of official i and prosecuting attorney a are common knowledge. Type is initially the official's private information and represents the marginal benefit to the official of engaging in (unprosecuted) corruption (described in greater detail later). Types are drawn from party-specific distributions with probability density functions $g_D(\cdot)$ and $g_R(\cdot)$. These have continuous

5 In his March 13, 2007 interview of Professor Shields, Steven Colbert slyly asked whether the study's figures were not "just overwhelming proof that Democrats are corrupt."

6 The 7-to-1 ratio touted by Shields and Cragan was obtained by restricting attention to investigations of local officials in their sample.

7 For a recent application of Manski's approach in political science, see Ashworth et al. 2008. Employing the updated data from Shields' testimony and information on referrals in 2007, the revised interval is $[-0.944, 0.987]$.

8 Shields and Cragan compare the partisan distribution of investigation targets with "the available normative data," suggesting that 50% of elected officials are Democrats and 41% are Republicans. (In the original paper, no citation for these figures is given. In his Congressional testimony, Shields cites a survey conducted by the Eagleton Institute at Rutgers University [Mandel and Kleeman 2004], which sampled elected officials younger than 35 years.) Assuming these figures accurately reflect the partisan breakdown of referrals, the estimate for Equation (1) is 0.046.

9 Throughout this section, I employ the term "official" when referring to potential defendants in public corruption cases. In reality, the set of individuals actually prosecuted for public corruption includes elected officials, appointees, civil servants, and private citizens. I consider this heterogeneity in my following empirical analysis.

10 I employ female pronouns for prosecutors and male pronouns for public officials.

11 Adding a third category of officials with no known partisan affiliation would not alter the intuition.

support over the real line, and their shape is common knowledge to the players.

The game unfolds in three steps. First, officials observe their types and choose levels of corruption $c_i \geq 0$. Second, if an official does break the law ($c_i > 0$), with probability q the infraction is detected by law enforcement and referred to the prosecutor, and with probability $(1 - q)$ it is undetected (or unreferred). Finally, the prosecutor observes the referrals and associated values of c_i, and decides whether to prosecute them.[12]

The payoffs to the prosecutor stem from three sources. First, there is an opportunity cost k associated with pursuing a case.[13] Second, there is a career benefit associated with prosecuting high-profile cases (i.e., those with larger c_i). Finally, there is a political benefit $b_{p_i}^{p_a}$ to a prosecutor from party p_a associated with prosecuting a public official from party p_i. It is this parameter that captures the notion of partisan bias.

Definition. *A prosecutor from party p maintains* ***partisan bias*** *if* $b_{p'}^{p} > b_{p}^{p}$ *for* $p \neq p'$.

In other words, the prosecutor is biased if the political benefit to her of pursuing a case against an official from the opposite party is greater than the benefit of pursuing a case against a co-partisan. To simplify exposition, I will assume throughout that k always exceeds b for all configurations of partisanship; this rules out prosecutions against defendants engaging in *de minimis* levels of corruption. I will further assume that $b_{p'}^{p} \geq b_{p}^{p}$ for $p \neq p'$ (note the weak inequality). This implies that a prosecutor can be biased against her political opponents or unbiased, but not biased against her allies. (I return to this assumption in the Discussion section.)

Rather than fully model procedures following a formal indictment, I adopt the following reduced form: if an official is indicted, he will receive an expected sentence (brought about by a plea bargain or the certainty equivalent of the lottery of a full trial) equal to $s(c_i)$, which is continuously differentiable and increasing at an increasing rate for $c_i \geq 0$, with $s(0) = 0$ and $s'(0) = 0$.

Let Q_i be an indicator equal to one if a case is referred and zero otherwise, A_i an indicator equal to one if the prosecuting attorney brings a case against official i and zero otherwise, and $\mathbf{c}$ an $N \times 1$ vector of corruption levels. Utility functions are given by

$$u_a\big(b_D^{p_a}, b_R^{p_a}, k, \mathbf{c}\big) = \sum_{i=1}^{N} A_i\big(Q_i c_i + b_{p_i}^{p_a} - k\big) \text{ and}$$

$$u_i(c_i; m_i, s(\cdot)) = (1 - Q_i A_i) m_i c_i - Q_i A_i s(c_i) \quad \text{for all } i.$$

To simplify notation, I henceforth assume without loss of generality that the prosecuting attorney is from party R and label the political benefit from prosecuting a party p official b_p.

The subgame perfect Nash equilibria to this game are straightforward, so I defer formal characterization to Lemma 1 in Appendix A. In equilibrium, prosecutors never indict officials against whom no referral is brought. Conditional on observing a referral, the prosecuting attorney will pursue a formal charge against an official if and only the observed level of corruption of an official from party p exceeds a threshold equal to $k - b_p$. Officials fall into one of three relevant categories (a fourth category, described in Lemma 1, is knife's edge): the first consists of those who experience negative marginal utility from corruption ($m_i < 0$) and who therefore do not engage in it ($c_i^* = 0$). The second group consists of individuals whose optimal level of corruption given the threat of prosecution falls below the prosecutor's threshold for indictment. These officials will engage in corruption at the threshold level ($c_i^* = k - b_p$), knowing that even if they are detected, they will not be prosecuted. The remaining officials engage in levels of corruption greater than the prosecutor's threshold; those levels are strictly increasing in the official's type, m_i, and decreasing in the probability of referral q.[14]

Empirical Implications and Identification

In this section, I explore the empirical implications of the model, consider potential threats to inferring partisan bias with observational data on corruption prosecutions, and derive a strategy for overcoming these threats. Formal propositions and proofs appear in Appendix A.

The foregoing suggests that a partisan-biased prosecutor will be willing to file weaker cases against political opponents than allies, which will in turn produce lower sentences. Suppose the underlying distribution of official types for party D, $g_D(\cdot)$, is identical to that of party R, $g_R(\cdot)$. Partisan bias on the part of a prosecutor from party R should then lead to (1) *lower* average sentences for defendants from party D than defendants from party R, and (2) a *higher* proportion of defendants from party D in the set of prosecutions than in the general population.

The assumption that the distribution of benefits to corrupt behavior is the same across the two parties is, of course, unlikely to hold in reality. In the U.S., for example, Democrats tend to be disproportionately concentrated in urban areas, where opportunities for (and rewards to) corruption may be greater. Differences in these distributions threaten to confound efforts to assess bias using cross-sectional data on corruption prosecutions from a given administration.

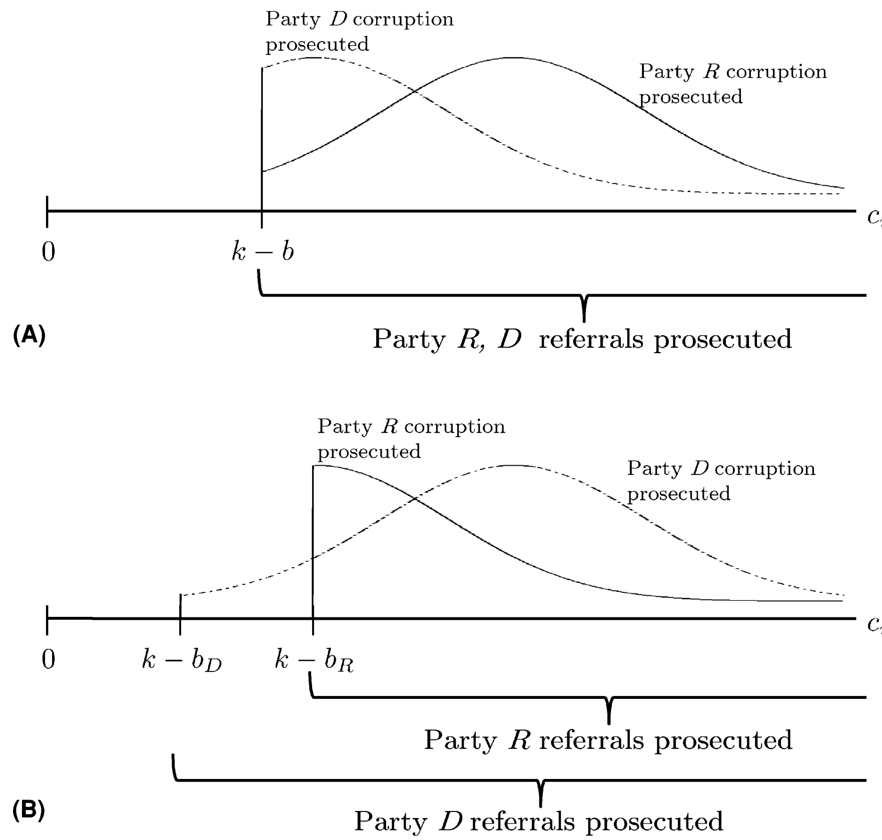
To make this issue concrete, assume that the party-specific distributions of types are not necessarily equal, but may be ordered by monotone likelihood ratio

[12] To economize on notation, the model allows the prosecutor to indict an official against whom no referral was made. As will become clear, this has no effect on the results.

[13] Public corruption cases represent only a fraction of the total caseload of a U.S. Attorney's Office. Therefore, one can think of k as capturing the foregone benefit of prosecuting a noncorruption case.

[14] A nonstrategic version of the model in which officials were characterized by exogenously given corruption levels would yield similar empirical implications to the ones described later.

FIGURE 2. Potential Threats to Inferring Bias Stemming from Underlying Interparty Differences in Returns to Corrupt Activity: (A) Confounding Yields Overestimate of Partisan Bias, (B) Confounding Yields Underestimate of Partisan Bias



Note: In the situation depicted in (A), the equilibrium distribution of (prosecuted) corruption for party *R* dominates that of party *D*, leading to a lower average sentence for party *D* even in the absence of prosecutorial partisan bias against party *D*. In the situation depicted in (B), the equilibrium distribution of corruption for party *D* dominates that of party *R*, pushing the average sentence for party *D* up relative to that of party *R*; if the average sentence for party *D* is still lower than that of party *R*, the source of the disparity must be prosecutorial bias.

(MLR) dominance.[15] Informally, MLR dominance is a common way of characterizing one distribution as "larger" than another; an example would be two normal distributions with the same variance, one of which has a higher mean than the other.[16] First, suppose $g_R(\cdot)$ strictly MLR dominates $g_D(\cdot)$; this would imply officials from party *R* tend to enjoy higher returns from corruption than officials from party *D* and will, ceteris paribus, engage in more corrupt activity as a consequence. Given a prosecutor from party *R*, the expected sentence for officials from party *D* could be less than the expected sentence for officials from party *R in the presence or absence* of partisan bias. Figure 2A displays the intuition graphically: the posited difference in distributions will lead to a lower expected level of corruption—and thus lower sentences—from party *D* than from party *R*, even when there is no partisan bias on the part of the prosecutor ($b_D = b_R = b$). In the presence of partisan bias, it would be impossible to distinguish the effect of the bias from the effect of the difference in the underlying distributions.

Next, suppose that $g_D(\cdot)$ weakly MLR dominates $g_R(\cdot)$. Given a prosecutor from party *R*, the expected sentence for officials from party *D* would be lower than the expected sentence for officials from party *R only in the presence of partisan bias.* As Figure 2B suggests, dominance of the distribution of party *D* types pushes the expected sentence for defendants from party *D* up relative to the expected sentence for defendants from party *R*. If, in spite of this, we nonetheless observe a significantly lower average sentence for defendants from party *D* than party *R*, the source of the difference must be partisan bias.

In terms of estimation, a difference in the underlying party-specific distributions is akin to omitted variables bias (OVB). In the first case, the OVB is away from zero, whereas in the second, the OVB is toward zero.

[15] $g_a(m)$ MLR dominates $g_b(m)$ if and only if $g_a(m)/g_b(m)$ is increasing over the entire domain of *m*. MLR dominance implies first-order stochastic dominance.

[16] A property of the MLR dominance relation useful in the current context is its persistence under truncation—a feature that does not hold generally for a ranking of distributions by their means.

Note that if one is assessing partisan bias of a prosecutor from party D, the potential threats to inference will be reversed: if $g_R(\cdot)$ dominates $g_D(\cdot)$, the OVB will be toward zero, whereas if $g_D(\cdot)$ dominates $g_R(\cdot)$, the bias will be away from zero.

Rather than assume, implausibly, that the underlying distributions are equal across parties, my identification strategy instead relies on a comparison of corruption case outcomes under *both* party D and party R prosecutors, under the less stringent assumption that $g_D(\cdot)$ and $g_R(\cdot)$ do not change drastically over adjacent administrations.[17] First, I consider conditions under which one can infer an ordering of the party-specific distributions from the data, which would in turn allow one to sign the OVB. It turns out that patterns in the data would permit this. In particular, if we assume that $g_D(\cdot)$ and $g_R(\cdot)$ can be ranked by MLR dominance, then disproportionate targeting of defendants from party D (relative to their proportion in the general population—*not* the set of referrals) by a prosecutor from party D can occur only if $g_D(\cdot)$ strictly MLR dominates $g_R(\cdot)$.

The intuition is simple: the presence of a party D prosecutor will push down the proportion of partisan-affiliated defendants from party D if she is biased, and have no effect on that proportion otherwise. If, in spite of this, we observe that proportion exceed its analog in the general population, then the only possible source of the disproportionality is a discrepancy in the distributions. For similar reasons, disproportionate targeting of officials from party R by prosecutors from party R can occur only if $g_R(\cdot)$ strictly MLR dominates $g_D(\cdot)$.[18] In other words, if we observe disproportionate prosecution of *co-partisans* by a prosecutor, then differences in observed sentences between the parties under that prosecutor will tend to *overstate* the extent of that prosecutor's partisan bias (OVB away from zero), whereas differences in observed sentences under a prosecutor from the opposite party will tend to *understate* the extent of that prosecutor's bias (OVB toward zero).

Second, let $E[s \mid p_i, p_a]$ represent the expected sentence of officials from party p_i under an administration from party p_a. In Appendix A, I demonstrate that the sample analog of the quantity

$$(E[s \mid p_i = R, p_a = R] - E[s \mid p_i = D, p_a = R]) \\ - (E[s \mid p_i = R, p_a = D] - E[s \mid p_i = D, p_a = D])$$

provides an unconfounded measure of the *total* extent of partisan bias summed across two adjacent administrations of different parties. Readers may recognize this quantity as a differences-in-differences (DiD). The DiD estimator recovers an unbiased measure of the aggregate effect of partisanship by controlling for ("differencing out") time-invariant differences between the party-specific type distributions and administration-specific effects (e.g., how much the administration prioritizes public corruption prosecutions generally). Critically, however, the estimator does not permit one to apportion the partisan bias between the two administrations.

Discussion of the Model

Before proceeding to the empirical analysis, I discuss the reasoning behind several important features of the model. First, the model does not distinguish between cases that go to trial and those that result in plea bargains. A simple permutation in which the prosecutor offered the defendant a "take it or leave it" plea bargain to avoid the gamble of a trial and with an exogenously given bargaining breakdown probability would yield results substantively similar to those described previously. Note that in the model, the function $s(\cdot)$ represents sentences brought about by plea bargains as well as the expected sentence given the lottery of a full trial; consequently, proper empirical tests for differences in sentencing will include plea bargains, acquittals, dismissals, and guilty verdicts in the calculation of the average observed sentence.

Second, to keep the exposition simple, the model assumes the referral process and the sentence function $s(\cdot)$ are not affected by the partisanship of the defendant. With respect to the latter, the fact that federal judges are themselves appointees with political affiliations may call this assumption into question. Later in the article, I discuss my approach for handling this challenge in the empirical analysis. With respect to the former, a plausible case could be made that pressure brought to bear on federal investigators works in the same direction as the partisan bias of the U.S. Attorneys themselves (and, indeed, might be initiated by them). If this is the case, then evidence of partisan bias may be attributed to the Justice Department as a whole, rather than the U.S. Attorneys in isolation.

Third, the model assumes that the referral accurately reflects the level of corruption in which an official engaged. In practice, the referral may contain noise that is (partially or fully) resolved during the discovery period between indictment and a trial. A simple version of the model with uncertainty resolution would have the prosecutor learn the true value of c_i conditional on paying the cost k. Such a model would share many features of the model described previously: most important, the prosecutor's strategy would be characterized by a threshold governing whether to indict given the quality of the initial referral, and that threshold would be lower for opponents given partisan bias.

The chief difference between such a model and the one described previously is that these thresholds would incorporate, in addition to any partisan bias, the prosecutor's prior beliefs about a defendant's type. So, for example, suppose the distribution of types from party D dominated the distribution from party R. The prosecutor would respond to this information by lowering the threshold for indictment for party D below that of party R, even in the absence of partisan bias. At the margin—that is, given a referral that makes the

[17] Clearly, this time invariance assumption is more justifiable at the state and local levels than at the national level. Accordingly, I confine my empirical analysis to state and local prosecutions.

[18] Note that this is a weak test, because there are configurations in the data that would not permit us to order the distributions even if the sample size grew very large (cf. Anwar and Fang 2006).

prosecutor exactly indifferent between indicting and not indicting—the equilibrium thresholds would exactly offset the difference in the distributions. As an empirical matter, however, we observe the average sentence; uncovering the marginal sentence is considerably more difficult.[19] To the extent that the model with uncertainty resolution better captures reality, the DiD test for aggregate partisan bias across administrations will be more informative than the tests for determining bias within an administration.

EMPIRICAL RESULTS

Description of Data

To test for systematic partisan bias at the Justice Department, I gathered two samples of state and local corruption prosecutions. The first consists of cases concluded during the Bush administration from 2004 to 2006, during which John Ashcroft and Alberto Gonzales each served as Attorney General. The second consists of cases concluded during the Clinton administration from 1998 to 2000, toward the end of Janet Reno's term as Attorney General. The method employed to compile the data was somewhat involved and is described in detail in Appendix B. These time frames were chosen for two reasons. First, the quality of electronic court records prior to 1998 (particularly in some judicial districts) is dramatically lower than more recent data. Second, using these periods eliminates cases that began under one president and concluded under another. At first glance, such cases may seem particularly attractive, pointing in the direction of a regression discontinuity-type research design. However, my theoretical model offers no coherent predictions for those cases.

For the 2004–2006 sample, I obtained information on the prosecution outcomes of 222 defendants. Before discussing their partisan composition, I note several other interesting features of the data. First, notwithstanding the DOJ's classification of these cases as "official corruption," 87 of the individuals (39%) are not government officials, but private citizens who may or may not be involved with corrupt officials. These include contractors accused of bribery, party operatives accused of buying votes, and a minister accused of establishing a bogus welfare-to-work program. Second, slightly more than half of the observations (115) are individuals who cannot be readily classified as belonging to one or the other party. In addition to some of the private citizens just mentioned, there are also public officials engaged in routine government operations, such as law enforcement officials (e.g., sheriffs' deputies and probation officers accused of extortion), school superintendents, liquor commission investigators, and water district managers. Officials who can be classified as partisan include mayors and town supervisors, city council members and treasurers, judges, and aides to elected officials. The modal lead charge for defendants is the bribery statute (48 defendants), followed by the Hobbs Act (39), conspiracy to defraud the United States (37), and mail fraud (20). Of the 222 cases, 204 resulted in guilty verdicts: 177 of those from plea bargains and the remainder from jury trials. Approximately two thirds of cases in which the defendant was found guilty (130) resulted in at least some prison time for the defendant. The median length of incarceration given a prison sentence was 21 months. The maximum sentence was 188 months. Of the remaining 18 cases, six were dismissed and 12 concluded with the defendant's acquittal at trial.

There are 223 cases in the 1998–2000 sample. Sixty-six of these were private citizens, and 77 could be readily classified as having a partisan affiliation. Violations of the Hobbs Act constituted the modal lead charge (72 defendants), followed by violations of the bribery statute (56). Of the 210 cases resulting in guilty verdicts, 175 emerged from plea bargains. Seven cases ended with not guilty verdicts, and six with dismissals. Interestingly, a slightly larger proportion of guilty verdicts in the Clinton sample ended in incarceration: 157, or 75%. Of those cases, the median sentence was 24 months, and the maximum 488 months.[20]

Partisan Composition of Defendants

Table 1 displays information on the partisan breakdown of defendants in the two sampling periods. Of the 222 cases for which I have collected data in the 2004–2006 (Bush administration) period, 84 of these individuals could be identified either as Democrats or as affiliated with a Democratic politician (e.g., in a fraudulent transaction). In contrast, only 23 defendants could be classified as either Republican or Republican affiliated. The ratio of 3.65 Democrats to 1 Republican prosecuted echoes the discrepancy observed in the Shields/Cragan study cited previously. The discrepancy becomes even more pronounced if one restricts attention to public officials: over six times as many Democratic public officials are observed in the data as Republican.

How do these figures differ from the partisan breakdown of defendants prosecuted under Clinton? Of the 223 cases coded, 77 could be identified with one of the two major parties: 49 Democrats and 28 Republicans. The data indicate that Democrats were disproportionately represented in the sample of defendants relative to Republicans *in a Democratic administration.* Among all defendants, there were 1.75 times as many Democratic-affiliated defendants as Republican-affiliated ones. Restricting attention to public officials, this ratio climbs to 2.6 to 1.

While I lack precise data on the ratio of Democrats to Republicans in the general population, it is not controversial to assume, based on extant surveys during the

[19] If all defendants pled out, then the best estimate of the marginal sentence for party p would be the smallest observed sentence for a defendant from that party. The possibility always exists, however, that plea bargaining will break down, leading to a trial with some possibility of an acquittal. Consequently, the smallest observed sentence (i.e., zero) is not truly the marginal sentence.

[20] This extreme outlier sentence was for a police officer with no identifiable partisan leanings; thus, it has no effect on the following interparty comparisons.

TABLE 1. Partisan Affiliation of Public Corruption Prosecution Defendants, Bush and Clinton Samples

	Bush Sample 2004–2006			Clinton Sample 1998–2000		
	Public	Private	All	Public	Private	All
Dem affiliated	49	35	84	36	13	49
Rep affiliated	8	15	23	14	14	28
Other	78	36	114	107	39	146
Ratio (Dem/Rep)	6.13	2.33	3.65	2.57	0.93	1.75
Ratio (Dem/non-Dem)	0.57	0.69	0.61	0.3	0.25	0.28
Ratio (Rep/non-Rep)	0.06	0.21	0.12	0.1	0.27	0.14

period, that the ratio was less than 1.8:1.[21] The overrepresentation of Democrats in the set of prosecutions under Clinton, although less pronounced than under Bush, suggests that disproportionality cannot *by itself* be taken to indicate anti-Democratic bias or "profiling" by the Bush administration in corruption prosecutions.

More important for current purposes, the fact that Democrats appear to be disproportionately prosecuted by the Clinton administration, provides evidence that the benefit to engaging in corruption typically tends to be larger at the state and local levels for Democrats than for Republicans. Employing the criteria discussed previously, this pattern in the data allows me to assess the direction in which the following results might be confounded: in particular, estimates of sentencing disparities are likely to overstate the true extent of partisan bias in the Clinton Justice Department, whereas those estimates are likely to understate the true extent of partisan bias in the Bush Justice Department.

Descriptive Results

Having assessed an important issue of potential confounding, I next turn to the specific prediction that in the presence of prosecutorial partisan bias, the average sentence of prosecuted partisan opponents will tend to be lower than that of prosecuted allies. Results appear in Table 2. Because sentences involving probation are not scalable on the same metric as those involving incarceration, I employ two different measures of total punishment. Columns (1) through (3) consider prison sentences in months, treating defendants who received probationary sentences as equivalent to those who were acquitted. In contrast, columns (4) through (6) incorporate information on the length of probationary sentences, under the uncontroversial assumption that the most severe probationary sentence (60 months) is more lenient than the least severe incarcerative sentence (one month). How the "pain" associated with an extra month of probation compares with that of an extra month of prison is unknown, so I present an intermediate case in which an extra month of prison is assumed to be five times as painful. So, for example, five years (60 months) probation would be coded as 12 units of punishment, and one month of incarceration as 13. Results employing different "exchange rates" between prison and probation are qualitatively similar to those reported here. Columns (1) and (4) in Table 2 display average sentences for different categories of defendants with identifiable partisan affiliations prosecuted under the two administrations, using the two measures of punishment. Interestingly, the data suggest that the average prison sentence for partisan-affiliated public corruption defendants under the Bush DOJ (22 months) was slightly longer than under Clinton (16 months), although this difference is not statistically significant.

Next, I turn to tests of administration-specific sentencing differences, reported in columns (2) and (5) of Table 2. The numbers in parentheses are one-tailed p values for corresponding t tests assuming unequal variances. (One-tailed tests are appropriate given the directional nature of the hypotheses tested.) For differences between Republican and Democratic average sentences during the Bush administration, a positive difference is evidence of partisan (pro-Republican) bias, as it is consistent with prosecution of weaker cases against Democrats. Turning to the results for the Bush administration, we find a positive difference pooling across all defendants, consistent with anti-Democratic or pro-Republican bias. However, the pooled results miss statistical significance at conventional levels. Confining attention to the sentences of public employees, however, the difference in means is statistically significant. This difference is substantively large as well: for example, the average imprisonment under Bush for Republican public employees is nearly twice that of the Democratic average.[22] Confining attention to private citizens with partisan affiliations, differences are quite small and nowhere close to significant. Results employing the alternative punishment measure incorporating probation are broadly similar to those using the imprisonment measure.

[21] The 2000 ANES reports a ratio of 1.37:1 for respondents. Using that ratio as a baseline, the binomial tail probability of observing 49 or more Democrats among the 77 defendants with identifiable partisanship is .18. Restricting attention to public officials, the tail probability is .03.

[22] The difference in medians is stark as well: the median prison sentence for Republican public employees under Bush was 23 months, compared to 12 months for Democratic public employees.

TABLE 2. Descriptive Indicators of Partisan Bias in Sentencing: Administration-specific Differences Between Republican and Democratic Defendants and Differences-in-Differences

	Prison Only			Incorporating Probation		
Defendant Party/ Prosecuting Administration	Mean (1)	Dif. in Means (2)	Dif. in Difs. (3)	Mean (4)	Dif. in Means (5)	Dif. in Difs. (6)
			All Partisan-Affiliated Defendants			
Republicans under Bush	28.39			38.06		
Democrats under Bush	20.07	8.32 (.17)		29.59	8.48 (.17)	
Republicans under Clinton	12.39		14.85 (.06)	22.42		15.14 (.06)
Democrats under Clinton	18.92	−6.53 (.04)		29.08	−6.66 (.06)	
			Public Employee Defendants			
Republicans under Bush	39.25			51.25		
Democrats under Bush	19.9	19.35 (.07)		29.55	21.7 (.05)	
Republicans under Clinton	10.79		30.32 (.01)	21.41		32.34 (.01)
Democrats under Clinton	21.75	−10.96 (.003)		32.05	−10.64 (.01)	
			Private Citizen Defendants			
Republicans under Bush	22.6			31.03		
Democrats under Bush	20.31	2.29 (.42)		29.64	1.39 (.45)	
Republicans under Clinton	14		−.64 (.52)	23.43		−1.18 (.53)
Democrats under Clinton	11.08	2.92 (.68)		20.86	2.57 (.64)	

Notes: Sentences and differences in columns (1) through (3) denote months of incarceration placing zero value on probation; sentences in columns (4) through (6) are calculated as 0.2 times the number of months of probation in the sentence if the sentence is solely probationary, and 12 plus the number of months of incarceration plus 0.2 times the number of months of probation if the sentence includes imprisonment (see text for additional explanation). One-tailed p values in parentheses.

For differences between Republican and Democratic average sentences during the Clinton administration, a negative difference is evidence of partisan (pro-Democratic) bias. Pooling over all defendants with identifiable partisan leanings, the data indicate that Democrats received prison sentences approximately six months more severe than Republicans. This difference is statistically significant for both punishment measures reported in Table 2. Confining attention to public employees only, differences are highly significant and of the anticipated sign. Under Clinton, Democratic officials received sentences 10 to 11 months more severe than Republican officials on average.

The data are thus consistent with partisan bias in the prosecution of public officials in both the Clinton and Bush administrations; however, they do not indicate partisan bias in the prosecution of private citizens. Taken together with the results in the previous section concerning confounding, the evidence suggests that observed differences under Bush are attributable to partisan bias and not to underlying differences between defendants from the two parties. The observed differences under Clinton may, however, be wholly or party attributable to such differences.

I report DiD estimates in columns (3) and (6). A positive, significant DiD reflects high total partisan bias summed across the Clinton and Bush administrations. Pooling across all defendants with identifiable partisan affiliations, I detect large, positive effects with both measures of punishment. Confining attention exclusively to public officials prosecuted under public corruption statutes, one observes effects even more pronounced, both in their magnitude and their statistical significance. Across the two administrations, the DiD estimates imply between 30- and 40-month aggregate differences in sentences attributable to partisan bias. As the p values in Table 2 indicate, these effects are very precisely estimated. Last, I detect no significant aggregate partisan effects when confining attention to private citizens.

Conditioning on Covariates

The foregoing analysis assumes that conditional on the partisan affiliation of the defendant, individual cases are exchangeable. In fact, there may be other features of the legal and political environment that affect sentencing. For example, prosecutors may get greater recognition from prosecuting elected officials than private citizens or administrative functionaries. Likewise, prosecutors may obtain scale economies from going forward with cases involving multiple defendants. Finally, judges, who must approve all sentences whether they follow plea bargains or guilty verdicts at trial, are themselves political appointees and may differ in their attitudes toward sentencing.

To account for this potential heterogeneity in cases, I assess the robustness of the descriptive results via three different statistical approaches: least squares regression, exact matching, and genetic matching.[23] These approaches permit me to estimate the average treatment effect of being a Republican defendant as compared to a Democratic defendant conditional on the previously

[23] I implement nearest-neighbor, one-to-one matching with replacement. Genetic matching is a weighted generalization of Mahalanobis distance matching with weights determined via a genetic optimization algorithm; see Diamond and Sekhon 2005.

TABLE 3. Regression and Matching Estimates of Administration-specific Differences in Sentences Between Republican and Democratic Defendants and Differences-in-Differences, Excluding Crime-level Covariates

		Prison Only			Incorporating Probation			
Method	Admin.	Est. (1)	S.E. (2)	Pr($T > t$) (3)	Est. (4)	S.E. (5)	Pr($T > t$) (6)	N (7)
				All Partisan-Affiliated Defendants				
Regression	Bush	9.2	7.74	0.12	9.13	8.12	0.13	107
	Clinton	−6.37	4.94	0.1	−6.53	5.49	0.12	77
	Dif. in difs.	15.57	9.6	0.05	15.66	10.24	0.06	184
Exact matching	Bush	27.16	5.5	0.000	28.66	5.57	0.000	59
	Clinton	−5.91	2.63	0.01	−6.62	3.05	0.01	46
	Dif. in difs.	33.07	6.09	0.000	35.28	6.35	0.000	105
Genetic matching	Bush	14.59	10.06	0.07	16.06	10.34	0.06	107
	Clinton	−8.63	4.38	0.02	−9.36	4.94	0.03	77
	Dif. in difs.	23.21	10.97	0.02	25.42	11.46	0.01	184
				Public Employee Defendants				
Regression	Bush	18.94	9.86	0.03	21.15	9.78	0.02	57
	Clinton	−12.46	5.08	0.01	−12.49	5.71	0.02	50
	Dif. in difs.	31.39	11.42	0.004	33.64	11.67	0.002	107
Exact matching	Bush	24.58	6.04	0.000	26.69	6	0.000	30
	Clinton	−10.07	3.32	0.000	−10.92	3.92	0.003	27
	Dif. in difs.	34.64	6.89	0.000	37.61	7.17	0.000	57
Genetic matching	Bush	16.36	12.67	0.1	18.68	13	0.08	57
	Clinton	−12.48	5.42	0.01	−13.47	6.21	0.02	50
	Dif. in difs.	28.84	13.78	0.02	32.15	14.41	0.01	107
				Private Citizen Defendants				
Regression	Bush	7.32	12.54	0.28	5.2	13.46	0.35	50
	Clinton	−5.84	9.19	0.26	−5.7	10.16	0.29	27
	Dif. in difs.	13.16	18.52	0.24	10.9	19.93	0.29	77
Exact matching	Bush	29.83	9.55	0.000	30.7	9.81	0.000	29
	Clinton	0.00	4.03	0.5	−0.51	4.6	0.456	19
	Dif. in difs.	29.83	10.37	0.000	31.21	10.84	0.000	48
Genetic matching	Bush	17.52	13.11	0.09	17.25	13.76	0.1	50
	Clinton	−1.48	6.47	0.41	−1.75	7.18	0.4	27
	Dif. in difs.	19	14.62	0.1	19	15.52	0.11	77

Notes: See notes in Table 2 and text for description of measure incorporating probation. Heteroscedasticity-consistent standard errors reported for regression estimates; Abadie and Imbens (2006) standard errors employed for matching estimates, with standard errors for differences-in-differences matching estimates calculated assuming independence of Bush and Clinton samples.

mentioned features of the environment. The included covariates are whether the defendant was a *private citizen*, whether he or she was an *elected official*, whether the sentence was approved by a *judge appointed by a Democratic president*, and whether the sentence was part of a case with *multiple defendants*. Because case priorities may vary by the region of the U.S. Attorney's Office (Eistenstein 1978), I also match on regional indicators: *Midwest*, *Northeast*, and *South* (the excluded category is the West). Heteroscedasticity-consistent standard errors are reported for the regression estimates, and bias-corrected standard errors (Abadie and Imbens 2006) for the matching estimates.

In the first four columns of Table C.1 in Appendix C, I report pre- and postmatching balance statistics as a means of diagnosing the success of the genetic matching procedure. (By definition, exact matching entails perfect balance.) Specifically, I employ standardized mean differences and p values for a t test that the covariate means are equal across defendant partisan affiliation (Imai, King, and Stuart 2008; Sekhon n.d.). Balance is indicated by smaller mean differences and larger p values. The diagnostics suggest that excellent balance is obtained on these covariates in both the Bush and Clinton samples. For example, the lowest p value after matching is .7 in the Bush sample and .37 in the Clinton sample.

Table 3 displays estimates for the average effect of being a Republican defendant compared to a Democrat using regression and matching approaches.[24] Estimates from all three approaches confirm the descriptive results described previously. Pooling over public officials and private citizens, we continue to observe higher sentences for Republicans than Democrats under Bush and lower sentences for Republicans under Clinton, as well as large and highly significant DiD

[24] The full statistical output is rather voluminous and is available from the author on request.

TABLE 4. Accounting for Interdistrict Heterogeneity: Random Coefficient Model Estimates of Administration-specific Differences in Sentences Between Republican and Democratic Defendants and Differences-in-Differences, Excluding Crime-level Covariates

	Prison Only				Incorporating Probation			
Admin.	Est. (1)	S.E. (2)	Pr($T > t$) (3)	S.D. (Partisan) (4)	Est. (5)	S.E. (6)	Pr($T > t$) (7)	S.D. (Partisan) (8)
				All Partisan-Affiliated Defendants				
Bush	10.1	11.79	0.2	15.36	10.13	12.68	0.21	16.84
Clinton	−8.96	6.27	0.08	10.22	−10.16	6.9	0.07	10.85
Dif. in difs.	19.06	13.35	0.08	—	20.29	14.43	0.08	—
				Public Employee Defendants				
Bush	20.06	12.96	0.06	14.19	22.64	12.99	0.04	12.71
Clinton	−17.24	4.95	0.000	0.014	−18.98	5.2	0.000	0.015
Dif. in difs.	37.3	13.88	0.004	—	41.62	13.99	0.002	—
				Private Citizen Defendants				
Bush	18.49	23.6	0.22	33.45	16.34	24.99	0.26	35.29
Clinton	1.54	9.02	0.57	20.84	1.1	9.88	0.54	22.37
Dif. in difs.	16.95	25.26	0.25	—	15.23	26.87	0.29	—

Notes: See notes in Table 2 and text for description of measure incorporating probation. Standard errors for differences-in-differences estimates calculated assuming independence of Bush and Clinton samples.

estimates. The most substantial departures from the descriptive results occur in the exact matching estimates, which boost the estimated treatment effect of being a Republican under Bush to more than 27 months. One is hesitant to make too much of this difference, however, because obtaining exact matches requires discarding more than 40% of the cases in the sample.

Note that I do not include controls for observable features of the underlying crime in the list of covariates. This omission is made to avoid posttreatment bias, which can emerge if the analyst controls for variables that are causally dependent on the treatment variable of interest (in this case, the partisanship of the defendant). The theoretical model implies that the distribution of criminal activity itself is, in the presence of bias, affected in equilibrium by the partisanship of the defendant. The sentence is simply a noisy measure of that criminal activity. Other features of crimes aside from the sentence are also dependent on the defendant's partisanship; thus, controlling for them may bias the estimated effect of partisanship.

Nonetheless, it is still useful to see whether matching on these features of the cases changes anything. I therefore reran the analysis from Table 3, including three indicator variables for whether the top charge in the indictment was *Racketeering*, *Fraud*, or *Bribery* (the excluded category is *Obstruction of Justice*). Balance statistics for the genetic matching appear in the fifth through eighth columns of Table C.1 in Appendix C. Results for regression and matching analysis incorporating crime-level covariates are displayed in Table C.2 in Appendix C, and are quite similar to the estimates that do not include them in the list of covariates. Note that exactly matching on the type of crime drastically reduces the sample size by more than 75%; in fact, in the analysis of sentences of private citizen defendants under Clinton, only four exact matches could be found.

Heterogeneity Across U.S. Attorneys' Offices

The foregoing analysis is premised on the assumption that it is possible to assess the extent of partisan bias in the Justice Department taken *in toto*. The decision to prosecute a public corruption cases, however, lies largely with the individual U.S. Attorneys and their assistants, who may exercise autonomy in setting office priorities. Accordingly, any heterogeneity in the extent of partisan bias among different U.S. Attorneys may manifest itself in variation across districts in differences in the observed sentences of Republican and Democratic public corruption defendants.

Ideally, one would estimate sentencing differences (and DiDs) for every U.S. Attorney's Office. However, many of the districts have only one or a handful of cases associated with them. An alternative approach is to model heterogeneity in the effects of partisan bias parametrically, in the context of a random coefficient model (see, e.g., Gelman and Hill 2007, ch. 11), in which district-level differences in Republican and Democratic corruption sentences are assumed to be draws from a normal distribution whose estimated mean and standard deviation reflect, respectively, the typical effect and the extent of heterogeneity.

Table 4 displays estimates from a series of such models. In addition to the estimate of the mean partisan effect, its standard error, and the one-tailed p value, I also report, in columns (4) and (8) the estimated standard deviation of the distribution of partisan differences. Estimates of the mean effect of defendant party and the uncertainty surrounding them are quite similar to the models estimated assuming homogeneous effects across districts. Two features of these estimates are particularly noteworthy, however. First, the magnitude of estimated standard deviation of the effects across districts in the pooled (public and private)

TABLE 5. Regression and Matching Estimates of Administration-specific Differences in Sentences Between Republican and Democratic Defendants and Differences-in-Differences, Northern District of Illinois Only

		Prison Only			Incorporating Probation			
Method	Admin.	Est. (1)	S.E. (2)	Pr($T > t$) (3)	Est. (4)	S.E. (5)	Pr($T > t$) (6)	N (7)
Regression	Bush	33.37	16.21	0.02	34.65	16.23	0.02	23
	Clinton	−17.46	10.91	0.06	−18.75	11.28	0.05	26
	Dif. in difs.	50.83	22.71	0.02	53.4	22.97	0.01	49
Exact matching	Bush	51.24	13.21	0.000	52.92	13.15	0.000	15
	Clinton	−16.67	4.08	0.000	−18.07	4.46	0.000	12
	Dif. in difs.	67.91	13.83	0.000	70.99	13.88	0.000	27
Genetic matching	Bush	34.45	16.76	0.02	35.61	16.77	0.02	23
	Clinton	−16.88	5.15	0.001	−17.53	5.5	0.000	26
	Dif. in difs.	51.33	17.53	0.002	53.14	17.65	0.000	49

Notes: See notes in Table 2 and text for description of measure incorporating probation. Heteroscedasticity-consistent standard errors reported for regression estimates; Abadie and Imbens (2006) standard errors employed for matching estimates, with standard errors for differences-in-differences matching estimates calculated assuming independence of Bush and Clinton samples.

sample suggests considerable variation in the differences in sentences for Republican and Democratic defendants. To see which districts departed most significantly from the mean effect, I calculated best linear unbiased predictors for district-level party effects. In the Bush data, the districts of Connecticut, Western North Carolina, and Western Tennessee showed unusually small (but still positive) differences between Republican and Democratic defendants, whereas Eastern California and Northern Illinois featured the largest differences. Interestingly, none of the dismissed U.S. Attorneys served in those districts. In the Clinton data, Southern Florida and Eastern New York displayed unusually small (but still negative) differences, and the difference for Western Oklahoma was the wrong (i.e., positive) sign. Differences for Northern Alabama, Northern Illinois, and Southern Ohio were atypically large in magnitude.

Second, restricting attention to public employee defendants only, the variance of effects across districts remains large for the Bush administration, but is quite small for the Clinton administration. In fact, a likelihood ratio test comparing the random coefficient model to simple linear regression for the Clinton/public employee sample is right at the threshold of rejecting the latter ($\chi^2 = 4.48, p = .11$). The apparently greater homogeneity in effects across districts under Clinton relative to Bush is intriguing, and suggests an avenue for future research on the extent of administrative centralization under each.

Although limitations in the data preclude estimating separate sentencing differences for most districts, an important exception to this rule is the Northern District of Illinois, which handles corruption cases from Chicago and its surrounding suburbs. I have data from this district on Democratic and Republican defendants under both the Bush and Clinton administrations. Northern Illinois is also a "hard case" for demonstrating partisan bias because of the reputation of Bush-appointed U.S. Attorney Patrick Fitzgerald for nonpartisanship. The U.S. Attorney for Northern Illinois from 1998 to 2000 was Clinton appointee Scott Lassar.

Table 5 displays the results of regression and matching analyses employing cases only from the Northern District of Illinois. Because of the small sample size, I report only results pooling across both public and private defendants. (Balance statistics for the genetic matching appear in Table C.1 in Appendix C. The procedure failed to produce good balance on the party of the president who appointed the presiding judge in a case.) The results for the single district mirror those for the national sample. We observe Republican defendants receiving substantially and statistically significantly higher sentences under Fitzgerald, and Democrats receiving higher sentences under Lassar. One challenge associated with focusing attention on this particular district is the fact that Democratic-affiliated defendants tend to work in Chicago government, whereas Republican-affiliated defendants tend to work for the state or come from smaller towns such as Cicero. Of course, these differences tend to persist across administrations, affording validity to the DiD estimates as evidence of partisan bias aggregated across the two administrations.

DISCUSSION

Some Remaining Issues Addressed

Here, I address some remaining challenges to the validity of the model and empirical findings. First, although the model allows for prosecutorial discretion in case selection, it does not consider discretion over what sentence to push or bargain for—presumably, a biased prosecutor would push for more severe sentences against opponents than allies. It is immediate that this behavior would work *against* my empirical strategy for detecting partisan bias. For example, the model anticipates that under a biased Republican prosecutor, the typical sentence should be lower for Democrats than Republicans, *ceteris paribus*. Discretion of the sort considered by this objection would push the average sentence for Democrats up and the average sentence for Republicans down. If, in spite of this, we still observe Democrats receiving lower sentences under a Republican prosecutor, the source of this difference is likely to be the bias in case selection considered explicitly in the model (accounting for the possibility of confounding effects discussed previously).

The second issue concerns professional norms of neutrality in their application of the law. Suppose U.S. Attorneys, in an effort to appear politically neutral, devote their limited time toward building a few strong, high-profile cases against officials from their own party. The empirical implications would be similar to those anticipated by my model in the presence of partisan bias. There are two responses to this concern. To begin, it is not clear whether a prosecutor trying to appear neutral would actually benefit from adopting this strategy rather than prosecuting many smaller cases against co-partisans. The posited strategy would open the prosecutor to the very charge of bias (supported by evidence of disproportionality, which is easier to observe than the prosecutor's allocation of costly effort) that she had sought to avoid. Moreover, although prosecutors may be sensitive to professional norms of neutrality, it is unlikely that their political superiors are. Consequently, if prosecutors truly behaved in the posited way, we would expect to see Republican administrations exhibiting a systematic preference for appointing Democratic U.S. Attorneys, and vice versa. The fact that we do not observe this further mitigates against this alternative account.

Third, my analysis examines cases concluded relatively late in the Clinton and Bush administrations. Suppose that U.S. Attorneys tend to prioritize cases against partisan opponents early in their administrations. It could then be that the differences I observe are simply a consequence of the Justice Department having already completed the strongest cases against its opponents. In fact, this is unlikely to be the case. To begin with, it is not the case that each administration starts with a finite supply of corruption cases that it depletes over time; the set of matters under consideration by U.S. Attorneys is constantly replenished by a steady stream of new referrals from investigative agencies. Moreover, the most high-profile cases tend to be those that stretch out over years; thus, for example, some convictions or plea agreements reached in 2006 were initially filed in 2001. Nonetheless, to make sure that these dynamic effects do not drive my results, I divided the data into "early" (1994–1996 and 2001–2003) and "late" (1998–2000 and 2004–2006) strata and reran my analysis for each. If the objection is correct, we would expect to see the partisan differences anticipated by the theory muted for the early strata relative to the late ones. In fact, we observe precisely the opposite.

Implications for the Study of Politicization and Autonomy

Contemporary theories of bureaucratic politics emphasize the tension between the partisan loyalties and expertise of officials. Lewis (2008), for example, argues that a president will be more apt to politicize agencies (foregoing the opportunity to capitalize on careerist expertise) when careerists do not share his or her underlying policy preferences. Consequently, whether an administration will politicize an agency may depend as much on *which* administration as on *which* agency. To the extent that an agency fosters professional norms devoted to the neutral application of the law (e.g., Huber 2007), however, presidents from both parties may benefit from seeking to politicize. The evidence presented in this article is consistent with politicization attempts by both Bush and Clinton, although it is not possible to definitively assess *relative* politicization between the two administrations.

At the same time, accounts of policy making at both the Justice Department (Seymour 1975) and elsewhere (Carpenter 2001) point to agency culture when accounting for the success of some agencies in some periods at resisting efforts at politicization. To be sure, a strong culture is not an insurmountable obstacle: several U.S. Attorneys have cited their defense of department culture as an immediate cause of their dismissals. More generally, the causal antecedents of strong bureaucratic culture are complex. Carol Lam, one of the fired U.S. Attorneys, suggested that the terminations "so transgressed the unwritten understanding and traditions of the department that, ironically, I think it has now reinforced them" (Fisher 2007). Another fired attorney, David Iglesias, was less sanguine about the department's resilience, writing that "it will take time for the damage done to the Justice Department to be completely remedied" (Iglesias 2008, 232).

Finally, to the extent that politicization efforts, once uncovered, can damage an administration politically, it is worth considering the potential role of oversight by either Congress or the media. The U.S. Attorney firing scandal would likely not have received as much attention if the majority control of Congress had not changed in 2006. In this vein, it is worth noting that the sample of corruption prosecutions from the Clinton administration employed in this article was drawn from a period of divided government and intense congressional scrutiny of the Reno Justice Department. It is possible that the mere potential for such scrutiny mutes the incentive to politicize the administration of justice for partisan ends. Likewise, the Bush sample was drawn from a period of unified government. A valuable extension, albeit one beyond the scope of the current article, would be to examine Clinton administration corruption prosecutions during unified Democratic control of government and Bush administration prosecutions during divided government, to better understand this effect.

Conclusion

The 2007 U.S. Attorney firing scandal raised the possibility that federal corruption laws could be deployed for partisan ends. In this article, I have sought to move beyond anecdotes to construct a systematic test of partisan bias in corruption prosecutions. My simple theoretical model anticipates differences in the distribution of observed sentences between defendants from the two major parties in the presence of partisan bias. Based on estimates of those differences, I find strong evidence of aggregate partisan bias across the Justice Departments of Bill Clinton and George W. Bush. Apportioning the bias between the administrations is slightly more challenging. Although evidence of bias is detected under both administrations, these estimates are susceptible to confounding because we cannot directly observe the actual magnitude of the benefit to an individual of engaging in corruption apart from the (potentially incomplete) evidence on which the prosecutor relies. However, under a plausible set of assumptions, evidence pointing to the disproportionate prosecution of Democrats under both Bush and Clinton speaks to the direction of this confounding effect, implying that the effect of partisanship is most likely stronger than implied by my analysis of Bush-era prosecutions, and weaker than implied by my analysis of Clinton-era prosecutions.

I conclude by noting that these results may be interpreted as both good news and bad news for Democrats and Republicans alike. For Democrats, the evidence suggests that the Bush Department of Justice may indeed have discriminated against them. For Republicans, the fact that Democrats were disproportionately targeted under both Clinton and Bush is evidence that Democrats at the state and local level may have access to more substantial opportunities for corruption than Republicans. This is perhaps not surprising given the concentration of Democratic public officials in urban areas. It is important to keep in mind, however, that the presence of opportunities for corruption for some members of a party is not synonymous with a tendency of the typical official from that party to violate the law. Indeed, both the model and data are fully consistent with compliance by the vast majority of officials and private citizens from both parties.

APPENDIX A: FORMAL RESULTS

I begin with a formal characterization of equilibrium to the game described in the text:

LEMMA 1 (Equilibrium Characterization). *In any subgame perfect Nash equilibrium to the game:*

(a) The prosecutor prosecutes official i if and only if i is referred and $c_i > k - b_p$; and
(b) For any official $i \in 1, \ldots, N$:

(i) $c_i = 0$ if $m_i < 0$,
(ii) $c_i \in [0, k - b_p]$ if $m_i = 0$,
(iii) $c_i = k - b_p$ if $m_i \in (0, \frac{q}{1-q}s'(k - b_p)]$, and
(iv) $c_i = s'(\frac{(1-q)}{q}m_i)^{-1}$ if $m_i \in (\frac{q}{1-q}s'(k - b_p), \infty)$.

Proof. The game is solved via backward induction. To establish (a), note that if i is not referred, under the assumption that $k > b_p$, the marginal utility to the prosecutor associated with bringing a case against i is $b_p - k < 0$, so not prosecuting strictly dominates prosecuting. From the prosecutor's best response correspondence, if i is referred, the prosecutor strictly prefers prosecuting if $c_i > k - b_p$, strictly prefers not prosecuting if $c_i < k - b_p$, and is indifferent if $c_i = k - b_p$.

To establish part (b)(i), note that if $m_i < 0$, official i's expected utility is strictly decreasing in c_i, making $c_i = 0$ a dominant strategy. Next, consider officials for whom $m_i \geq 0$. Let $\pi(c_i)$ be the probability the prosecutor prosecutes an official engaged in corruption c_i given a referral, and suppose the prosecutor does not prosecute if $c_i = k - b_p$ (i.e., $\pi(k - b_p) = 0$). The expected utility to official i of engaging in corruption c_i is

$$E[u_i(c_i \mid m_i, q, s(\cdot))] = (1 - \pi(c_i)q)m_i c_i - \pi(c_i)qs(c_i).$$

The official's first-order condition is given by

$$-\pi'(c_i)q(m_i + s(c_i)) + (1 - \pi(c_i)q)m_i - \pi(c_i)qs'(c_i) = 0. \quad \textbf{(A.1)}$$

Suppose a candidate solution to Equation (A.1), $\hat{c}_i(m_i; q, s(\cdot))$, exceeds $k - b_p$, so, via part (a), $\pi(\hat{c}_i(\cdot)) = 1$ and $\pi'(\hat{c}_i(\cdot)) = 0$. Substituting into Equation (A.1) and solving for $\hat{c}_i(\cdot)$ yields the optimal value of c_i given in part (b)(iv). (Second-order conditions guarantee a unique maximum.) By assumption, $s'(c_i) > 0$, $s''(c_i) > 0$, and $s'(0) = 0$, so $s'(\cdot)^{-1}$ exists and is strictly increasing, with $s'(0)^{-1} = 0$. This implies $\hat{c}_i(m_i; q, s(\cdot))$ is increasing in m_i, with $\hat{c}_i(0; q, s(\cdot)) = 0$. Consequently, there exists an m_i' such that for all $m_i > m_i'$, $\hat{c}_i(m_i; q, s(\cdot)) > k - b_p$. Substituting the value obtained for $\hat{c}_i(m_i; q, s(\cdot))$ into this inequality and solving yields the condition in part (b)(iv).

For values of $m_i \in (0, \frac{q}{1-q}s'(k - b_p)]$, $\hat{c}_i(m_i; q, s(\cdot)) < k - b_p$, in which case $\pi(\hat{c}_i(\cdot)) = 0$. In such cases, the official can always benefit from increasing his level of corruption up to $c_i = k - b_p$ given $\pi(k - b_p) = 0$, establishing part (b)(iii). Officials for whom $m_i = 0$ are indifferent between not engaging in corruption and engaging in any level of corruption up to $k - b_p$ given $\pi(c_i) = 0$ for any $c_i \leq k - b_p$, establishing part (b)(ii).

To see why any strategy profile that includes $\pi(k - b_p) \neq 0$ cannot be an equilibrium, suppose otherwise. Then, for any official with type $m_i \in (0, \frac{q}{1-q}s'(k - b_p))$ playing strategy $c_i' < k - b_p$, there exists an $\varepsilon > 0$ such that such an official would benefit from playing $c_i' + \varepsilon$. ■

Let N_p be the total number of party p officials in the population, and $\hat{N}_{p_i}^{p_a}$ the number of officials from party p_i prosecuted by a prosecutor from party p_a. Let $h_D(c)$ be the density of prosecuted *corruption levels* for party D, and $h_R(c)$ the density for party R. The next result summarizes empirical implications of the model under the assumption that party D and party R types are drawn from the same distribution:

Proposition 1 (Empirical Implications Given Equal Type Distributions)**.** *Suppose $g_D(m) = g_R(m)$ for all m. In the presence of a biased prosecutor from party R, (a) the expected sentence for a defendant from party D will be less than the expected sentence for a defendant from party R; and (b) in expectation, the proportion of prosecuted officials from party D will exceed its analog in the population.*

Proof. (a) From part (a) of Lemma 1, the distribution of prosecuted corruption levels will be left truncated at $k - b_p$ for officials from party p. From part (b)(iv), corruption levels prosecuted with positive probability are a strictly increasing function of an official's type m_i and are independent of b_p. Consequently, if $g_D(\cdot) = g_R(\cdot)$, $h_D(c)$ and $h_R(c)$ will both be left-truncated manifestations of the same underlying density, and the expected level of corruption will be lowest for the party with the lower truncation point—party D in the presence of partisan bias ($b_D > b_R$). Because $s(\cdot)$ is a strictly increasing function of c, this in turn implies that the expected sentence for prosecuted officials from party D will be less than the expected sentence from party R.

(b) From Lemma 1, the probability an official from party p engages in a prosecutable level of corruption given a prosecutor from party R is given by

$$\theta_p^R \equiv \int_{\frac{q}{1-q}s'(k-b_p)}^{\infty} g_p(m)\, dm. \tag{A.2}$$

If $g_D(\cdot) = g_R(\cdot)$, then $\theta_D^R > \theta_R^R$ if and only if the prosecutor is biased ($b_D > b_R$), in which case

$$\frac{E\left[\hat{N}_D^R\right]}{\hat{N}} = \frac{N_D\theta_D^R}{N_D\theta_D^R + N_R\theta_R^R} > \frac{N_D}{N_D + N_R},$$

where $\hat{N}$ is the observed number of prosecuted officials from parties D and R. ■

Let $\succeq_{MLR}$ and $\succ_{MLR}$ denote, respectively, weak and strict monotone likelihood ratio orderings. The next result characterizes threats to inference stemming from differences in $g_D(\cdot)$ and $g_R(\cdot)$:

Proposition 2 (Potential Threats to Inference)**.** *Given a prosecutor from party R:*

(a) If $g_D(\cdot) \succeq_{MLR} g_R(\cdot)$, then the expected observed sentence for officials from party D will be smaller than the expected observed sentence for officials from party R only in the presence of partisan bias ($b_D > b_R$); and
(b) If $g_R(\cdot) \succ_{MLR} g_D(\cdot)$, then the expected observed sentence for officials from party D could be smaller than the expected observed sentence for officials from party R in the presence or absence of partisan bias ($b_D \geq b_R$).

Proof. Note two useful properties of the monotone likelihood ratio order: (1), it is preserved under truncation; and (2), if the density of a random variable X MLR dominates that of random variable Y, then the density of $\psi(X)$ MLR dominates that of $\psi(Y)$ for any increasing function $\psi(\cdot)$ (see Theorems 1.C.6 and 1.C.8 of Shaked and Shanthikumar 2007).

The conjunction of these two properties and part (a) of Lemma 1 imply that in the absence of partisan bias ($b_D = b_R$), if $g_D(\cdot) \succeq_{MLR} g_R(\cdot)$, then the distribution of party D sentences will weakly MLR dominate the distribution of party R sentences. But then the expected observed sentence for party D will be weakly larger than that of party R. The only means for this inequality to be reversed is for the lower bound of the support of prosecuted corruption levels (and thus sentences) to be lower for party D than party R, which can only occur given partisan bias ($b_D > b_R$). In contrast, if $g_R(\cdot) \succ_{MLR} g_D(\cdot)$, then in the absence of partisan bias the expected observed sentence for party R will exceed that of party D, and the presence of partisan bias will reenforce this difference. ■

The following result gives conditions under which an ordering of $g_D(\cdot)$ and $g_R(\cdot)$ may be inferred from disproportionatility in the data:

Proposition 3 (Sufficient Conditions for Ordering Type Distributions)**.** *Suppose $g_D(\cdot)$ and $g_R(\cdot)$ can be ordered by monotone likelihood ratio dominance. Then disproportionate targeting of party R by prosecutors from party R will occur in expectation only if $g_R(\cdot) \succ_{MLR} g_D(\cdot)$. Likewise, disproportionate targeting of party D by prosecutors from party D will occur in expectation only if $g_D(\cdot) \succ_{MLR} g_R(\cdot)$.*

Proof. Assume a prosecutor from party R. There are two cases to consider: first, suppose the prosecutor is unbiased ($b_D = b_R$). If $g_D(\cdot) \succeq_{MLR} g_R(\cdot)$, then via Equation (A.2), party D officials will be prosecuted with weakly higher probability than party R officials. If $g_R(\cdot) \succ_{MLR} g_D(\cdot)$, party R officials will be prosecuted with strictly higher probability than party D officials. Second, suppose the prosecutor is biased ($b_D > b_R$). Then the lower threshold for officials from party D (see Lemma 1 and Proposition 1) will entail an increase in the probability party D officials are prosecuted relative to the probability party R officials are prosecuted. $g_D(\cdot) \succeq_{MLR} g_R(\cdot)$ will further increase the probability for party D officials relative to party R officials. In contrast, if $g_R(\cdot) \succ_{MLR} g_D(\cdot)$, the effect of the ordering of the densities will be to to increase the probability party R officials are prosecuted relative to party D officials. The conjunction of these effects could lead to disproportionate prosecution of party R, party D, or neither. Consequently, disproportionate prosecution of party D officials is consistent with any ordering of $g_D(\cdot)$ and $g_R(\cdot)$, whereas disproportionate targeting of party R officials is consistent only with $g_R(\cdot) \succ_{MLR} g_D(\cdot)$. An analogous proof establishes the proposition for a party D prosecutor. ■

The final result motivates the DiD estimator described in the text:

Proposition 4 (Aggregate Effect of Partisan Bias)**.** *Suppose the party-specific distributions of official types, $g_D(\cdot)$ and $g_R(\cdot)$, remain stable across two adjacent administrations of different parties. Then, partisan bias exists within at least one of those administrations if and only if*

$$\begin{aligned}&(E[s\,|\,p_i = R, p_a = R] - E[s\,|\,p_i = D, p_a = R])\\&- (E[s\,|\,p_i = R, p_a = D] - E[s\,|\,p_i = D, p_a = D])\\&> 0.\end{aligned} \tag{A.3}$$

Proof. Consider the following model of the conditional expected sentence for defendant i:

$$\begin{aligned}E[s\,|\,p_i, p_a] = \gamma_{p_i} + \alpha_{p_a} &- \beta_D^R I(p_i = D) \times I(p_a = R)\\&- \beta_R^D I(p_i = R) \times I(p_a = D),\end{aligned}$$

where $I(\cdot)$ is an indicator function, γ_{p_i} captures the effect on the expected sentence of the underlying distribution of official types in party p_i, α_{p_a} captures the effect of prosecutorial priorities under a prosecutor from party p_a, β_D^R is the *diminution* of sentences attributable to partisan bias by a prosecutor from party R against officials from party D, and β_R^D is the diminution attributable to bias by a party D prosecutor against party R defendants. Then, $E[s\,|\,p_i = R, p_a = R] = \gamma_R + \alpha_R$, $E[s\,|\,p_i = D, p_a = R] = \gamma_D + \alpha_R - \beta_D^R$, $E[s\,|\,p_i = R, p_a = D] = \gamma_R + \alpha_D - \beta_R^D$, and $E[s\,|\,p_i = D, p_a = D] = \gamma_D + \alpha_D$. Substituting into inequality Equation (A.3) and simplifying yields $\beta_D^R + \beta_R^D$, which is the sum of the effects of partisan bias across the party D and party R administrations. ■

APPENDIX B: DATA COLLECTION

The Executive Office of U.S. Attorneys (EOUSA) releases data on prosecutions to two forums: TRAC, cited previously, and the Federal Justice Statistics Resource Center, a subdivision of the Bureau of Justice Statistics (BJS), co-sponsored by the Urban Institute. Although each database contains information on the same universe of cases, each has advantages. For example, TRAC contains DOJ program categories, whereas BJS data contain readily accessible information on case priority: District (D), National (N), or Both National and District (B). I restrict attention to categories N and B (priority D cases are typically clerical workers). Neither data source has the defendant's name. The data universe for this article consists of the intersection of BJS records of the relevant priority levels and TRAC records with DOJ categorized as state and local corruption for the years under consideration.

To locate defendant names, I relied on the federal district courts' Case Management/Electronic Case Files (CM/ECF) system and, when necessary, Nexis. This process was fairly involved. The first step was to enter the filing date from the EOUSA data into a district's CM/ECF system. This typically yielded a list of cases. The second step was to restrict attention to cases in which the United States was a plaintiff and in which the CM/ECF closing date approximated the TRAC closing date. The third was to examine the electronic docket to find the case whose details matched those summarized in TRAC. From there, I obtained the defendant name(s). If this approach did not work, I searched Nexis for articles in state news sources that included the judge's name around the time of sentencing, then searched for the name of the defendant listed in the news story in CM/ECF.

Finally, using the defendant's name and common variants thereof, I obtained information on his or her position and party using Nexis, searching U.S. Attorney press releases, election returns, telephone calls to local city halls, or the Internet more broadly. (A list of sources is available on request.) This approach identified approximately 83% of the defendants.

APPENDIX C: ADDITIONAL TABLES

TABLE C1. Balance Tests for Covariates Used in Genetic Matching Analyses on Pooled Samples of Public and Private Defendants

	Basic Matching				Controls for Crime				Northern Illinois Only			
	Before Matching		After Matching		Before Matching		After Matching		Before Matching		After Matching	
Covariate	Std. Mean Diff.	*t* test *p* value	Std. Mean Diff.	*t* test *p* value	Std. Mean Diff.	*t* test *p* value	Std. Mean Diff.	*t* test *p* value	Std. Mean Diff.	*t* test *p* value	Std. Mean Diff.	*t* test *p* value
					Bush Sample, 2004–2006							
Private citizen	48.36	0.05	0	1	48.36	0.05	−3.75	0.48	61.55	0.16	8.51	0.32
Elected official	−44.2	0.1	0	1	−44.2	0.1	0	1	28.87	0.34	20.85	0.32
Multiple defendants	26.4	0.28	1.9	0.84	26.4	0.28	−1.87	0.88	3.89	0.93	0	1
Democrat-appointed judge	46.98	0.04	0	1	46.98	0.04	8.84	0.26	113.28	0.002	64.69	0.004
Midwest	41.55	0.08	0	1	41.55	0.08	9.57	0.3	—	—	—	—
Northeast	3.56	0.88	2.26	0.71	3.56	0.88	0	1	—	—	—	—
South	−16.2	0.5	2.07	0.78	−16.2	0.5	11.73	0.22	—	—	—	—
Racketeering	—	—	—	—	−76.21	0.01	−2.54	0.32	—	—	—	—
Bribery	—	—	—	—	−0.31	0.99	−3.83	0.41	—	—	—	—
Fraud	—	—	—	—	27.06	0.25	−6.15	0.51	—	—	—	—
					Clinton Sample, 1998–2000							
Private citizen	46.09	0.05	0	1	46.09	0.05	0	1	65.28	0.02	47.85	0.02
Elected official	−115.01	0	−8.05	0.37	−115.01	0	−8.05	0.37	−295	0	−7.93	0.32
Multiple defendants	52.57	0.03	5.21	0.68	52.57	0.03	5.24	0.73	139.07	0.02	−15.5	0.15
Democrat-appointed judge	−48.1	0.05	0	1	−48.1	0.05	0	1	−70.75	0.06	−47.57	0.05
Midwest	48.1	0.05	0	1	48.1	0.05	0	1	—	—	—	—
Northeast	−24.3	0.39	0	1	−24.3	0.39	0	1	—	—	—	—
South	−31.11	0.2	5.2	0.48	−31.11	0.2	5.2	0.48	—	—	—	—
Racketeering	—	—	—	—	24.3	0.31	−13.41	0.35	—	—	—	—
Bribery	—	—	—	—	−6.54	0.79	0	1	—	—	—	—
Fraud	—	—	—	—	−26.86	0.28	5.57	0.68	—	—	—	—

TABLE C2. Regression and Matching Estimates of Administration-specific Differences in Sentences Between Republican and Democratic Defendants and Differences-in-Differences, Including Crime-Level Covariates

		Prison Only			Incorporating Probation			
Method	Admin.	Est. (1)	S.E. (2)	Pr($T > t$) (3)	Est. (4)	S.E. (5)	Pr($T > t$) (6)	N (7)
				All Partisan-Affiliated Defendants				
Regression	Bush	9.79	7.5	0.1	10.04	7.87	0.1	107
	Clinton	−5.43	5.09	0.14	−5.44	5.68	0.17	77
	Dif. in difs.	15.22	9.41	0.05	15.48	10.09	0.06	184
Exact matching	Bush	26.21	4.38	0.000	28.24	4.44	0.000	30
	Clinton	−5.08	1.4	0.000	−6.37	1.56	0.000	13
	Dif. in difs.	31.28	4.6	0.000	34.61	4.7	0.000	43
Genetic matching	Bush	6.3	9.43	0.25	6.99	9.9	0.24	107
	Clinton	−7.95	4.46	0.04	−8.65	5.02	0.04	77
	Dif. in difs.	14.25	10.43	0.09	15.64	11.1	0.08	184
				Public Employee Defendants				
Regression	Bush	19.43	9.58	0.02	21.79	9.6	0.01	57
	Clinton	−11.06	4.95	0.01	−11.06	5.74	0.03	50
	Dif. in difs.	30.49	10.55	0.002	32.85	10.99	0.002	107
Exact matching	Bush	19.87	3.92	0.000	22.19	4.02	0.000	15
	Clinton	−8.67	2.27	0.000	−10.53	2.52	0.000	9
	Dif. in difs.	28.53	4.53	0.000	32.72	4.75	0.000	24
Genetic matching	Bush	14.91	11.08	0.09	17.28	11.32	0.06	57
	Clinton	−9.79	6.27	0.06	−10.15	7.34	0.08	50
	Dif. in difs.	24.7	12.73	0.03	27.43	13.49	0.02	107
				Private Citizen Defendants				
Regression	Bush	13.01	12.68	0.15	11.36	13.74	0.21	50
	Clinton	−4.78	8.38	0.29	−4.54	9	0.31	27
	Dif. in difs.	17.78	17.85	0.16	15.9	19.18	0.21	77
Exact matching	Bush	32.55	8.61	0.000	34.29	8.71	0.000	15
	Clinton	3	0.62	0.000	3	0.62	0.000	4
	Dif. in difs.	29.55	8.63	0.000	31.29	8.73	0.000	19
Genetic matching	Bush	10.35	15.98	0.26	9.54	16.86	0.29	50
	Clinton	2.5	8.53	0.62	2.46	9.81	0.6	27
	Dif. in difs.	7.85	18.11	0.33	7.08	19.5	0.36	77

Notes: See notes in Table 2 and text for description of measure incorporating probation. Heteroscedasticity-consistent standard errors reported for regression estimates; Abadie and Imbens (2006) standard errors employed for matching estimates, with standard errors for differences-in-differences matching estimates calculated assuming independence of Bush and Clinton samples.

REFERENCES

Abadie, Alberto, and Guido Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74 (1): 235–67.

Abrams, Norman, and Sara Sun Beale. 2006. *Federal Criminal Law and Its Enforcement*. 4th ed. Saint Paul, MN: West Group.

Alt, James E., and David Dryer Lassen. 2002. "The Political Economy of Institutions and Corruption in American States." *Journal of Theoretical Politics* 15 (3): 341–65.

Anechiarico, Frank, and James B. Jacobs. 1994. "Review: Visions of Corruption Control and the Evolution of American Public Administration." *Public Administration Review* 54 (5): 465–73.

Anwar, Shamena, and Hanming Fang. 2006. "An Alternative Test of Racial Prejudice in Motor Vehicle Searches: Theory and Evidence." *American Economic Review* 96 (1): 127–51.

Anzia, Sarah, and Christopher R. Berry. 2007. "The Jackie (and Jill) Robinson Effect: Congresswomen and the Distribution of Federal Spending." Working paper 07.16. The University of Chicago Harris School.

Ashworth, Scott, Joshua D. Clinton, Adam Meirowitz, and Kristopher W. Ramsay. 2008. "Design, Inference, and the Strategic Logic of Suicide Terrorism." *American Political Science Review* 102 (2): 269–73.

Ayres, Ian. 2001. *Pervasive Prejudice? Unconventional Evidence of Race and Gender Discrimination.* Chicago: The University of Chicago Press.

Ayres, Ian, and Joel Waldfogel. 1994. "A Market Test for Race Discrimination in Bail Setting." *Stanford Law Review* 46 (5): 987–1046.

Becker, Gary. 1957. *The Economics of Discrimination*. Chicago: The University of Chicago Press.

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

Boylan, Richard, and Cheryl X. Long. 2003. "A Survey of State House Reporters' Perception of Public Corruption." *State Politics and Policy Quarterly* 3 (4): 420–38.

Brown, David S., Michael Touchton, and Andrew Whitford. 2006. "Political Polarization as a Constraint on Government: Evidence from Corruption." Working paper. University of Colorado at Boulder.

Bryce, James. 1995 [1888]. *The American Commonwealth*. Indianapolis, IN: Liberty Fund.

Carpenter, Daniel P. 2001. *The Forging of Bureaucratic Autonomy.* Princeton, NJ: Princeton University Press.

Diamond, Alexis, and Jasjeet Sekhon. 2005. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." Working paper. University of California, Berkeley.

Eisenstein, James. 1978. *Counsel for the United States: U.S. Attorneys in the Political and Legal Systems.* Baltimore: The Johns Hopkins University Press.

Fisher, George. 2007. "Legal Matters with Carol Lam." *Stanford Lawyer* 77 (Fall): 24–8.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University Press.

Glaeser, 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): 259–90.

Goel, Rajeev K., and Michael A. Nelson. 1998. "Corruption and Government Size: A Disaggregated Analysis." *Public Choice* 97 (1–2): 107–20.

Goldstein, Amy, and Dan Eggen. 2007. "Renzi Aide Called U.S. Attorney to Ask About Probe; Chief of Staff Inquired About Land Deal Investigation; Prosecutor Amount Eight Who Were Fired." *The Washington Post*, April 26, p. A04.

Gordon, Sanford C., and Gregory A. Huber. 2002. "Citizen Oversight and the Electoral Incentives of Criminal Prosecutors." *American Journal of Political Science* 46 (2): 334–51.

Gordon, Sanford C., and Gregory A. Huber. 2009. "The Political Economy of Prosecution." *Annual Review of Law and Social Science* 5: 135–56.

Heidenheimer, Arnold J. 1989. "Problems of Comparing American Political Corruption." In *Political Corruption: A Handbook*, eds. A. J. Heidenheimer, M. Johnston, and V. T. LeVine. New Brunswick, NJ: Transaction Books, 573–85.

Huber, Gregory. 2007. *The Craft of Bureaucratic Neutrality.* New York: Cambridge University Press.

Iglesias, David, with David Seay. 2008. *In Justice: Inside the Scandal That Rocked the Bush Administration.* Hoboken, NJ: John Wiley and Sons.

Imai, Kosuke, Gary King, and Elizabeth A. Stuart. 2008. "Misunderstandings Among Experimentalists and Observationalists About Causal Inference." *Journal of the Royal Statistical Society, Series A (Statistics in Society)* 171 (2): 481–502.

Knowles, John, Nicola Persico, and Petra Todd. 2001. "Racial Bias in Motor Vehicle Searches: Theory and Evidence." *Journal of Political Economy* 109 (1): 203–29.

La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, and Robert Vishny. 1999. "The Quality of Government." *Journal of Law, Economics, and Organization* 15 (1): 222–79.

Lewis, David E. 2008. *The Politics of Presidential Appointments.* Princeton, NJ: Princeton University Press.

Lichtblau, Eric. 2007. "Ex-Prosecutor Says Departure Was Pressured." *The New York Times*, March 6, p. 1.

Lipset, Seymour Martin, Kyoung-Ryung Seong, and John Charles Torres. 1993. "A Comparative Analysis of the Social Requisites of Democracy." *International Social Science Journal* 136 (2): 155–75.

Lipton, Eric. 2007a. "One Prosecutor's Ouster Central to Inquiry." *The New York Times*, April 19, p. 24.

Lipton, Eric. 2007b. "Some Ask if U.S. Attorney Dismissals Point to Pattern of Investigating Democrats." *The New York Times*, April 30, p. 20.

Mandel, Ruth B., and Katherine E. Kleeman. 2004. "Political Generation Next: America's Young Elected Leaders." Eagleton Institute of Politics Report. Rutgers University.

Manski, Charles F. 1995. *Identification Problems in the Social Sciences.* Cambridge, MA: Harvard University Press.

Meier, Kenneth J., and Thomas M. Holbrook. 1992. "'I Seen My Opportunities and I Took 'Em:' Political Corruption in the United States." *Journal of Politics* 54 (1): 135–55.

Moe, Terry M. 1985. "Control and Feedback in Economic Regulation: The Case of the NLRB." *American Political Science Review* 79 (4): 1094–116.

Mosedale, Mike. 2007. "Not Quite Ready for His Close-Up." *City Pages*, March 21. http://citypages.com/2007-03-21/news/not-quite-ready-for-his-close-up (Accessed July 14, 2009).

Nathan, Richard P. 1975. *The Plot That Failed: Nixon and the Administrative Presidency.* New York: John Wiley and Sons.

Nye, J. S. 1967. "Corruption and Political Development: A Cost Benefit Analysis." *American Political Science Review* 61 (2): 417–27.

Persson, Torsten, Guido Tabellini, and Francesco Trebbi. 2003. "Electoral Rules and Corruption." *Journal of the European Economic Association* 1 (4): 958–89.

Richman, Daniel C. 1999. "Federal Criminal Law, Congressional Delegation, and Enforcement Discretion." *UCLA Law Review* 46 (February): 757–814.

Rose-Ackerman, Susan. 1999. *Corruption and Government.* Cambridge: Cambridge University Press.

Sekhon, Jasjeet. n.d. "Alternative Balance Metrics for Bias Reduction in Matching Methods for Causal Inference." Working paper. University of California, Berkeley.

Seymour, Whitney North, Jr. 1975. *United States Attorney.* New York: William Morrow.

Shaked, Moshe, and J. George Shanthikumar. 2007. *Stochastic Orders.* New York: Springer.

Shields, Donald C., and John F. Cragan. 2007. "The Political Profiling of Elected Democratic Officials: When Rhetorical Vision Participation Runs Amok." *E Pluribus Media: A Collaborative Journal for New Media*. www.epluribusmedia.org/columns/2007/20070212_political_profiling.html (Accessed July 14, 2009).

Shotts, Kenneth, and Alan Wiseman. 2008. "Information, Accountability and the Politics of Investigations." Presented at the Annual Meeting of the Midwest Political Science Association, Chicago.

Treisman, Daniel. 2000. "The Causes of Corruption: A Cross-National Study." *Journal of Public Economics* 76 (3): 399–457.

U.S. Department of Justice (US DOJ). 2008. "An Investigation into the Removal of Nine U.S. Attorneys in 2006." Washington, DC: USDOJ Office of Professional Responsibility and Office of the Inspector General. www.usdoj.gov/opr/us-att-firings-rpt092308.pdf (Accessed April 17, 2009).

Whitford, Andrew. 2002. "Bureaucratic Discretion, Agency Structure, and Democratic Responsiveness: The Case of the United States Attorneys." *Journal of Public Administration Research and Theory* 12 (1): 3–27.