

# The Costs of Top-Down Control: Discretion and Turnover of Federal Prosecutors\*

Mitch Downey      Ben Grunwald

February 16, 2023

## Abstract

In 2003, the Bush Administration issued a policy requiring federal prosecutors to charge and seek conviction on the most severe, readily provable offense in each of their cases. To study the implementation and consequences of this controversial effort to constrain prosecutorial discretion, we exploit an understudied Senate procedure governing the confirmation of U.S. Attorneys—the managers of federal prosecutor offices. We show that having more conservative Senators in a state leads to the appointment of a U.S. Attorney who is more engaged in conservative politics. We then find that offices with more conservative U.S. Attorneys more faithfully implemented the policy, illustrating how buy-in by middle managers helps to enact organization-wide reform. Finally, we show that the offices that most faithfully implemented the memo experienced a significant increase in resignations of front-line prosecutors afterwards and that these prosecutors were particularly experienced and skilled. Our results thus highlight how politically motivated efforts to curb bureaucratic discretion can impose sizeable personnel costs.

**Keywords:** organizational economics; bureaucratic politics; worker discretion; bureaucratic discretion; prosecutorial discretion; Ashcroft memo

**JEL Classification Numbers:** D72, D73, J45, K14

---

\*Downey: Institute for International Economic Studies (IIES), Stockholm University. Grunwald: School of Law, Duke University. We thank Scott Baker, Sara Sun Beale, Stephanie Didwania, John de Figueiredo, Brandon Garrett, Lisa Griffin, Scott Guenther, Claire Lim, Aurelie Ouss, Shane Stansbury, David Strömberg, Noam Yuchtman, and seminar participants at Warwick, Queen Mary, Linnaeus University, CELS, SIOE, and the Duke Criminal Law Works-in-Progress Workshop for helpful feedback.

# 1 Introduction

In complex organizational environments, it is often impossible to specify in advance a complete set of decision rules. The standard solution is to entrust agents with discretion to exercise their own judgement. A growing literature in economics examines how expanding or constraining this discretion affects individual worker performance (Bandiera et al., 2021; Duflo et al., 2018; Hoffman et al., 2018; Rasul and Rogger, 2018; Szucs, 2020), but it does not account for the possibility that changes in discretion affect the composition of the workforce itself. Indeed, because workers value professional autonomy (Maestas et al., 2017; Nikolova and Cnossen, 2020; Stern, 2004), changing their discretion might affect whether and which workers choose to remain with the organization. In contexts where performance quality varies across workers (Best, Hjort, and Szakonyi, 2019; Fenizia, 2021; Hoffman and Tadelis, 2021), these effects on turnover might be as important as effects on individual-level performance. In this paper, we examine how 87 United States Attorney’s Offices implemented a policy to reduce federal prosecutors’ discretion and quantify the costs on the agency’s workforce.

Prior to 2003, the Department of Justice gave federal prosecutors significant discretion to pursue criminal charges based on an “individualized assessment” of whether those charges were “proportional to the seriousness of the defendant’s conduct.” But, in that year, Attorney General John Ashcroft reversed course, issuing a memo stating that prosecutors “must charge and pursue the most serious, readily provable offense.” Importantly, the memo didn’t empower or require prosecutors to do anything they couldn’t do before. Rather, it was designed to constrain discretion, and it did so in at least two ways. First, it sought to increase the average severity of case outcomes.<sup>1</sup> Second, it sought to increase uniformity in the severity of outcomes among similar cases. To many government officials, the memo represented a dangerous erosion of prosecutorial discretion (Lichtblau, 2003). Indeed, at the time, former Deputy Attorney General Philip Heymann said that “the anger [among prosecutors] is immense,” and he predicted “a huge revolt” (Liptak and Lichtblau, 2003).

Our analysis of the effects of the memo proceeds in three stages. First, because the memo applied nationwide, we need variation in the intensity of its implementation across prosecutor offices to study its effects. We exploit variation in the appointment process for US Attorneys, the leaders of these offices. Specifically, we focus on an understudied institution governing their confirmations: US Senators can veto nominees for office in their home state. We build a

---

<sup>1</sup>While our empirical evidence focuses constraints on discretion that increase sentences, anecdotally, constraints that reduce sentences can also be controversial. In early 2021, for example, state prosecutors in Los Angeles County sued the elected District Attorney to overturn a policy prohibiting them from seeking sentencing enhancements. Their union argued “Dismissals of those priors can only be based on individual circumstances, not a blanket policy” (Romero, 2021).

novel biographical database on all confirmed nominees from the Bush Administration’s first term and show that, in state’s with more conservative Senators, the President appointed US Attorneys who were more “politicized”—a shorthand we use to mean that they engaged in more partisan campaign contribution behavior and were more likely to work in politics. Consistent with senators’ veto power, these effects were driven by the ideology of a state’s most liberal Senator (the ideology of the more conservative senator is irrelevant), having at least one Democratic senator (a second Democrat has no marginal effect), and by the Senator actually in office at the time of the nomination (the ideologies of senators from earlier or subsequent terms is irrelevant).

Second, we examine whether the offices of more politicized US Attorneys implemented the Ashcroft memo with greater fidelity. To do so, we use the Justice Department’s case management database, the Legal Information Office Network System (LIONS), to develop measures of the severity and uniformity of charges convicted in each case. We then apply a difference-in-differences strategy to test whether offices in states with a conservative Senator—where more politicized US Attorneys were appointed—showed evidence of *differential* implementation, after accounting for time-invariant between-district heterogeneity and aggregate trends in case processing.<sup>2</sup> Overall, we find evidence of variation in implementation across districts. Specifically, we find that conservative-senator districts experienced a differential increase in the severity of convicted charges. We also find that these districts experienced differential reductions in line prosecutors’ discretion, as measured by the standard deviation of the severity of the most serious convicted charge and the degree to which the identity of the prosecutor assigned to a case explains the severity of its outcomes. We interpret these results as evidence that more politicized U.S. Attorneys implemented the memo with greater fidelity, which illustrates how buy-in by middle managers can be critical for enacting organization-wide reform.

Third, we test whether the memo cost the Justice Department front-line prosecutors. Observing several thousand voluntary departures, we find that conservative-senator districts suffered a markedly higher number after the memo was issued. We also find that it was precisely the prosecutors who remained in the Justice Department who changed their behavior in compliance with the memo; in contrast, those who left did not. Given that many prosecutors reportedly opposed the memo, we interpret these findings as evidence that it caused the subsequent increase in departures.

Admittedly, the Bush Administration might have quietly celebrated the departure of some

---

<sup>2</sup>To check whether our results are simply driven by background correlations between geography and ideology, we use a strategy in which senate turnover isolates idiosyncratic variation specific to the term when the Attorneys were nominated.

of these prosecutors if they were particularly low-performing. But we find little evidence they were. Instead, the prosecutors induced to leave by the memo had more experience handling complex and important cases than the average departing attorney. Reasonable minds can disagree about the value of turnover in a prosecutor’s office, but, at least from the perspective of the Bush Administration, these findings reinforce that the memo hampered its law enforcement goals.

Beyond the economics literature on the effects of discretion on worker performance, our work contributes to two additional research areas. The first studies the tensions between the ideological preferences of politicians and bureaucrats. Spenkuch, Teso, and Xu (2021), for example, show that procurement officers perform worse under administrations that conflict with their own political views. Several other papers find that a change in presidential administration drives federal bureaucrats to quit (Bolton, de Figueiredo, and Lewis, 2020; Doherty, Lewis, and Limbocker, 2019; Richardson, 2019). And recent theoretical work argues that the most motivated and skilled bureaucrats depart in the highest numbers after presidential turnover (Cameron and de Figueiredo, 2020). Our analysis contributes to this literature by showing how more modest shifts in government policy, which nonetheless constrain worker discretion, can drive up departures.

Our work contributes to a second body of scholarship on prosecutors’ charging and plea bargaining discretion (Rehavi and Starr, 2014; Berdejo, 2018; Starr and Rehavi, 2013; Didwania, 2021b; Okafor, 2022). Didwania (2022) documents that a 2013 policy prohibiting federal prosecutors from filing mandatory minimum charges for low-level, non-violent, trafficking offenses achieved its goal, but we know little about whether and how supervisors can regulate the discretion of line prosecutors, both within and across offices, more broadly. Our results suggest that the Attorney General can control prosecutorial discretion through high-profile charging and plea bargaining memos and that the size of their impact is contingent on buy-in from US Attorneys—the “middle managers” who lead each office.

In the next section, we discuss the institutional background of the US Attorney’s Office and the history of the Ashcroft memo. Section 3 designs and justifies our identification strategy and sections 4 and 5 examine the effects of the memo on case outcomes and departures, respectively. Section 6 summarizes our results that consider four alternative explanations (we confine our evidence to the appendix), and Section 7 concludes.

## 2 Institutional Background

### 2.1 Employment Context

Federal criminal cases are prosecuted by one of 93 US Attorney’s Offices, 89 of which lie within the 50 states. Each state has at least one district and populous states, like California, have up to four. Each office is run by a US Attorney who oversees its front-line prosecutors, called Assistant US Attorneys, who are themselves responsible for handling criminal cases directly. While there are only 93 US Attorneys, there are thousands of Assistants.<sup>3</sup>

For our purposes, there are three key facts about employment in US Attorney’s Offices. First, both US Attorney and Assistant US Attorney positions are prestigious. US Attorneys are the most senior federal law enforcement official in their district and are often later appointed to a federal judgeship or other appointed positions in the federal government (Boylan, 2005). Assistants gain valuable courtroom experience and exposure to federal law, and, similarly, often transition into prestigious and lucrative positions in the public or private sector (Boylan and Long, 2005).

Second, an Assistant US Attorney holds a standard, federal bureaucratic position that is meant to be apolitical. They have been afforded civil service protections since the 1980s, and, as a result, their jobs are secure and very few departures are involuntary.<sup>4</sup>

Finally, the position of a US Attorney is explicitly political. US Attorneys are appointed by the President and are expected to resign at the beginning of a new presidential administration. While they are tasked with objectively and fairly upholding the law, they are also responsible for implementing the President’s law enforcement priorities. It is therefore neither surprising nor controversial that we find evidence politics affects their behavior. Indeed, there is evidence that US Attorneys Offices increase filings for particular enforcement areas—like narcotics, immigration, and terrorism—when presidential administrations mention them more in public statements (Miller and Curry, 2019).

Our identification strategy leverages a key institutional norm in the political process for appointing US Attorneys. Like federal judges, US Attorneys are nominated by the President and must be confirmed by the Senate. Since 1913, the Senate Judiciary Committee has honored home state senators’ objections to nominees (the so-called “blue slip” procedure), giving them the ability to veto the President’s choice (for a history, see Binder (2007)). While scholars have closely studied blue slip vetoes for *judicial* nominations, few have done so for

---

<sup>3</sup>The average district has 64 Assistant US attorneys, ranging from 8 in Western Arkansas to 390 in Central California, the latter of which includes Los Angeles.

<sup>4</sup>For instance, publicly available administrative data on separations from US Attorney’s Offices shows 1704 permanent-contract attorneys leaving US Attorneys offices from FY05-FY10, of which zero were fired or dismissed.

US Attorney nominations (one exception is Nelson and Ostrander (2016)). Although these vetoes are rarely exercised formally, the threat of a veto can nonetheless have a powerful influence on the nomination process. In the words of John Suthers, the US Attorney for Colorado during the Bush Administration, “You don’t become a US Attorney without the support of your state’s senators” (Miller and Curry, 2019). More direct evidence comes from Kyle Sampson, the Chief of Staff for the Attorney General under the Bush administration. When an obscure legislative provision was enacted in 2006 that empowered the Attorney General to bypass the Senate for US Attorney nominations, Sampson proposed using it to “give far less deference to home state senators and thereby get our preferred person appointed” (Eggen and Solomon, 2007). Our analyses bear out that the ideology of the home state Senator influences who ultimately gets appointed.

## 2.2 Prosecutorial Discretion

It is now widely agreed that federal prosecutors have more influence over the outcome of individual criminal cases than any other government actors (Barkow, 2008; Bibas, 2009; Davis, 2005).<sup>5</sup> Today, 95% of convictions arise through a guilty plea to all charges filed by the prosecutor or to a subset agreed to by the prosecutor. And, due to the Federal Sentencing Guidelines, sentences are fairly predictable conditional on convicted charges. This means that prosecutors control the charges filed, the charges convicted, and much of the ultimate sentence. These decisions are also largely unmonitored (Barkow, 2008), difficult to appeal (Davis, 2005), and the prosecutor has substantial leeway to exercise judgement (Bibas, 2009).

Federal prosecutors speaking to the *Washington Post* prior to Ashcroft memo were candid about the scope of their discretion (Flaherty and Biskupic, 1996). One former US Attorney from the Southern District of New York, explained: “If you push and pull a whole lot you can reach almost any conclusion you want about what you can actually charge a person with... And that’s the whole ballgame. You can call the same act by several names, and each one brings about a different result in prison time.” Several prosecutors interviewed viewed discretionary charging as an opportunity to correct for harsh and rigid federal sentencing guidelines, which, in the words of one former US Attorney, often lead to sentences that “reasonable people would view as an unjust result” (Flaherty and Biskupic, 1996).

This discretion arises from at least three main sources. First, in some cases, the legal definition of a charge may be ambiguous. A narcotics *conspiracy*, for example, requires a certain amount of coordination among sellers, but exactly how much coordination (or what counts as coordination) is not clearly defined. In the absence of coordination, a defendant can be

---

<sup>5</sup>See Bellin (2019) for a dissenting view.

charged simply with narcotics distribution, which could carry a lighter penalty. Second, in many more cases, the evidence may be ambiguous. Depending on whether a defendant held drugs to use personally or to sell, for example, a case can be charged as simple possession or possession with intent to distribute, the latter of which comes with much stiffer penalties. Finally, even when the facts and law are clear, prosecutors still have normative discretion to decide whether to pursue a charge based on a variety of considerations—such as criminal history, age, family, and personal culpability. For example, a drug cartel member who sold 280 grams of cocaine over five distinct transactions could be charged with one count of trafficking or five; with trafficking alone, or with conspiracy and RICO charges; with trafficking 27 grams, 28 grams, or 280 grams, each of which would trigger no mandatory minimum, a five-year minimum, or a ten-year minimum, respectively.<sup>6</sup> Moreover, if the defendant possessed a gun during any one of the sales, he could be charged with an enhancement that would increase prison time by another five years. Taken together, legal, factual, and normative ambiguity can generate substantial variation across individual prosecutors. Indeed, Wright, Baughman, and Robertson (2021) showed the same criminal case to 500 prosecutors across the country and report dramatic differences in the number of charges that the respondents recommended filing. Roughly one fifth said they would file one or fewer charges, while another fifth said they would file five or more (see also Mayer (1996)).

## 2.3 Constraining Prosecutorial Discretion through the Ashcroft Memo

In recent decades, the Attorney General has issued a number of memos regulating different aspects of prosecutorial discretion. Two are most relevant here. First, from 1993 to 2003, the Reno Bluesheet gave federal prosecutors discretion to make charging and plea bargaining decisions based on “an individualized assessment of the extent to which charges... [were] consistent with the purposes of the federal criminal code, and maximize[d] the impact of federal resources on crime.” It was “appropriate” for prosecutors to consider not only whether they would be able to secure a conviction on a charge, but also whether the resulting sentence would be “proportional to the seriousness of the defendant’s conduct.”

Second, in September 2003, George W. Bush’s Attorney General, John Ashcroft, rescinded those guidelines and issued a memo instructing that federal prosecutors “must charge and pursue the most serious, readily provable offense or offenses ... supported by the facts of the case,”<sup>7</sup> and defined the most serious offense as one that would “generate the most sub-

---

<sup>6</sup>Indeed, several studies have found substantial bunching around mandatory minimum thresholds that is driven by prosecutorial discretion (Bjerk, 2017; Tuttle, 2019).

<sup>7</sup>The memo was in force until it was rescinded by the Obama Administration in 2010.

stantial sentence.” Importantly, the Ashcroft memo did not require or empower prosecutors to do anything they couldn’t do beforehand. Instead, it merely sought to constrain their discretion in two ways. First, it aimed to increase the average severity of case outcomes by instructing prosecutors to pursue the toughest charge that would reasonably result in conviction. Second, it aimed to increase the uniformity of outcomes among similar cases. Indeed, in announcing the memo, Ashcroft emphasized that prosecutors must be “uniform” and that “the charges a defendant faces should not depend upon the particular prosecutor assigned to handle the case.”

In large part because of these constraints, the memo was controversial among both prosecutors and legal commentators. Some Assistant US Attorneys characterized the new policy as “a serious erosion of their discretion in criminal cases” (Schmidt, 2003) and others observed that it would “serve only to further centralize authority in the hands of Washington policymakers” (Lichtblau, 2003). One legal commentator interpreted the memo similarly as an effort to “strip[] the discretion that federal prosecutors need to do justice” (Ely, 2004).

Yet, enforcing the memo was not easy. In the words of one anonymous federal prosecutor, while the memo “put[] significant pressure on prosecutors,” it “might be difficult for officials in Washington to enforce” (Lichtblau, 2003). One reason was that the memo carved out several exceptions, including where formal approval is granted by the local US Attorney. Therefore, as Kate Stith (2007) explains: “[The memo] depended on the incentives and attitudes of US Attorneys... There were not enough people in Main Justice to monitor and enforce ‘mandatory’ charging policies in every US Attorneys’ office.” According to Stith, the “mandatory-policy approach to controlling dispersed prosecutorial discretion [could] work (if ... at all) only by altering the practice and norms of US Attorneys’ offices.” Our empirical strategy builds on this observation to generate variation in compliance with the new policy across offices.

### 3 Designing the Identification Strategy

To justify our identification strategy for estimating the effects of the Ashcroft memo, we begin by studying how the ideology of a state’s US Senator affected the politicization of its appointed US Attorney during the Bush Administration. Given senators’ ability to veto home state nominees, we always focus on each state’s *most liberal* senator in office at the time Bush took office (2001). Below, we show empirically that this senator (rather than the other) is the key determinant of US Attorney politicization.



### 3.1 Data and Measures

*Measuring US Attorney Politicization.* We gathered data on the Bush Administration’s first nominee for US Attorney in each US Attorney’s office. Four offices were excluded because they do not have a US Senator (Guam/Northern Mariana Islands, Puerto Rico, Virgin Islands, District of Columbia) and another two were excluded because Bush did not nominate a US Attorney (District of Maine, Eastern District of Louisiana). We therefore gathered data on Bush’s first nominee in the remaining 87 offices. All 87 nominees were confirmed, 86 before Ashcroft issued the memo.<sup>8</sup> Some of these US Attorneys left during the Bush Administration, but their departures and replacements are potentially endogenous and so we examine the initial appointments.

We create two sets of variables to measure the politicization of our 87 appointed US Attorneys. The first captures political ideology based on campaign contributions from the Database of Ideology and Money in Elections (DIME; see Bonica, 2019), which includes every federal, and many state, campaign contributions from 1980 to 2014 (for more details, see Appendix A.1.1). The US Attorneys nominated by President Bush are very active donors. Roughly 85% appear in the contribution data compared to 4% of the general population and 43% of all lawyers (Bonica and Sen, 2017). The median US Attorney contributed \$5,000 from 1980 to 2014, much more than the median for all donors, \$300.<sup>9</sup> The total contributions of the median US attorney is above the 90<sup>th</sup> percentile of the all-donor distribution, and 10 of the 87 US Attorneys are in the top 1%. At the extreme, one US Attorney gave \$1.4 million over this period, putting him in the top .01% of all donors. Of the 74 US Attorneys for whom we found campaign contributions, all but three gave more to Republicans than Democrats, all but eight gave 90% or more of their contributions to Republicans, and most (53/74) gave 100% of their contributions to Republicans. These figures are particularly notable given that lawyers tend to be significantly more liberal than most donors (Bonica and Sen, 2017).

We use the DIME data to construct several outcome variables at the US-Attorney level: total campaign contributions made from 1980-2014; the share of contributions given to Republicans; and the Campaign Finance (CF) score, which assigns a higher score for contributions to relatively more conservative Republicans (Bonica, 2014).

Our second set of variables measures US Attorneys’ employment in politics based on detailed and near-complete biographical data we hand collected from many sources, including newspapers, law firm biographies, academic CVs, Wikipedia, nomination press releases,

---

<sup>8</sup>84 were confirmed by the end of 2002, well before the Ashcroft memo issued in October 2003. Robert Corrente of Rhode Island was the one US Attorney who was not confirmed until afterwards, in May 2004, although he was in office as Acting US Attorney at the time of the memo.

<sup>9</sup>We do not have information on the median contributions of all lawyers.

Congressional investigations, obituaries, and law school alumni profiles (for more details, see Appendix A.1.2). These data track the US Attorneys’ careers from the date of law school graduation to the present, with very few gaps.<sup>10</sup> We use our biographical data to construct two outcome variables at the US-Attorney level: whether the attorney ever ran for an elected office, like Congress, Mayor, Governor, President, and some others<sup>11</sup>; and ever worked directly for a politician.<sup>12</sup>

Importantly, for all five of our politicization variables, we use information from both before and after the US Attorneys were nominated. Excluding jobs after 2001 would eliminate many of the political positions because US Attorney is typically a mid-career position while holding elected office tends to occur later in life. Our preferred interpretation is that the index reflects an individual’s underlying predispositions towards partisan political activity, and the behaviors and career choices that reflect this disposition occur over their entire careers.<sup>13</sup> Nonetheless, the coefficients are directionally consistently, though somewhat smaller and no longer statistically significant, if we restrict to pre-2001 employment spells (Appendix Table B1).

Finally, after transforming our five politicization measures to have a mean of zero and a standard deviation of one, we construct a composite index by computing the mean of the five variables for each US Attorney. For the thirteen who did not make any campaign contributions, we omit Republican share and CF score from the index. We renormalize this index to have a standard deviation of one.

*Measuring Senator Political Ideology.* We measure the ideology of every senator in the 107<sup>th</sup> Congress (2001-2003) using NOMINATE scores, the most common estimates of federal legislators’ ideology (Poole and Rosenthal, 1991, 2000). NOMINATE scores assign more conservative values to legislators who vote more consistently with the Republican party.

---

<sup>10</sup>The most similar data to ours is from Boylan (2005) and Miller and Curry (2019), both of whom create biographical databases of US Attorneys’ careers only *after* their tenure as US Attorneys.

<sup>11</sup>We omit elected judges from this classification, which we see as primarily a legal job that, for idiosyncratic state-by-state reasons, sometimes requires participation in an election.

<sup>12</sup>Positions directly serving politicians include Chair of the Missouri Republican Party, Chief Legal Counsel for Mike Huckabee (former Republican Governor and Presidential Candidate), and Chief of Staff for Congressman Ron Lewis (R-KY).

<sup>13</sup>An alternative interpretation, however, is that our findings reflect the treatment effect of serving as a US Attorney in a district with a liberal/conservative senator on subsequent career and contribution decisions. We find this interpretation unlikely. We think that the reason James Comey (USA for Southern District of New York) never ran for elected office, worked for a politician, or contributed more than a trivial amount of money to political campaigns is because he is not an inherently political person. While the reason Thomas Marino (USA for Central District of Pennsylvania) went on to run for Congress as part of the 2010 Tea Party wave and later served on Donald Trump’s transition team is because he is a more politically active person with a deeper engagement with partisan politics. We think it is unlikely that Comey’s political career was held back as a result of the senators in New York being liberal, or that Marino’s career is the result of having conservative senators in Pennsylvania at the time.

Given our interest in US Attorneys, we use Poole and Rosenthal’s statistical software (Poole et al., 2020) to estimate “justice-specific” NOMINATE scores based only on votes that they code as related to “Public Safety,” “Narcotics,” “Firearms,” and the “Judiciary” (for more details, see Appendix A.2.1).<sup>14</sup> The correlation between justice-specific and all-issue scores is very high, roughly 0.97, so this has little effect on our results. In Appendix A.2.2, we illustrate the relationship between these scores and document and explain some instances where influential Senators’ justice-specific scores differs meaningfully from their all-issue scores.

## 3.2 Results

Panel (a) of Figure 1 maps the justice-specific NOMINATE score for each state’s most liberal senator in office in 2001. Panel (b) then presents a binned scatterplot showing that states with more liberal senators tended to have less politicized US Attorneys. This is entirely driven by states with at least one Democratic senator (represented by blue diamonds), where the slope is quite strong. Among the remaining states, all of which have two Republicans (represented by red squares), there is no discernible relationship between ideology and politicization.

[Figure 1 about here.]

Table 1 shows the results of a series of regressions at the US Attorney level in which the dependent variables are different measures of US Attorney politicization and the independent variable measures the ideology of the most liberal Senator in the state (normalized to range from zero to one, with one being the most conservative). The constant represents the expected value in the state with the most liberal Senator, and the coefficient for the “Senator ideology” variable represents the difference between that state and the state with the most conservative senator. Column 1 shows that the total dollars contributed to political campaigns by US Attorneys does not vary based on the NOMINATE score of the most liberal senator.<sup>15</sup> But the recipients of the donations do. Column 2 shows that, in the most liberal-senator state, roughly 90% of US Attorneys’ contributions go to Republican

---

<sup>14</sup>The standard, all-issues NOMINATE scores are valuable for a wide variety of applications, but we are sympathetic to arguments that other measures of ideology can perform better in the specific contexts for which they were developed (e.g., certain issues and certain periods of time). Caughey and Schickler (2016) emphasize “the feasibility and value of tailoring one’s model and data to one’s research goals rather than relying on off-the-shelf NOMINATE scores.”

<sup>15</sup>We measure the contribution amount as percentiles within the US Attorney distribution because there are extreme outliers and zeros. The results are the same, however, when we add an offset and take the log of contribution or when we use the inverse hyperbolic sine.

candidates, while in the most conservative-senator state, that number increases to 100% ( $p < .05$ ). Column 3 similarly shows that, in the most liberal-senator state, US Attorneys' CF scores are roughly .75, while in the most conservative-senator state, that number rises to 1 ( $p < .05$ )—a difference roughly equivalent to shifting from the contribution profile of the Illinois or Massachusetts Republican Party to that of the Kentucky or Alabama Republican Party.

[Table 1 about here.]

Columns 4 and 5 show that shifting from the most conservative-senator state to the most liberal more than doubles the probability that a US Attorney has or will work for a politician or run for elected office, respectively, but neither result is statistically significant. When we combine the two outcomes together, column 6 shows that the probability of holding a political job nearly triples—from 17pp in the most liberal districts to 44pp in the most conservative. While this coefficient is also not statistically significant at conventional levels ( $p = .113$ ), the corresponding logit regression—which is more efficient but harder to interpret—is statistically significant ( $p = .069$ ). Combining all of our measures, column 7 shows that shifting from the most conservative-senator state to the most liberal increases the index by a statistically significant 0.67 standard deviations.<sup>16</sup>

### 3.3 Alternative explanations

One compelling explanation for the correlation between senator ideology and US Attorney politicization is that senators wield veto power over nominees in their states. However, it is also possible that our results simply reflect broader geographic variation in ideology—that is, that more conservative US Attorneys tend to come from states with more conservative politicians. Since we use this variation throughout the paper, it is important to understand the source of this cross-sectional correlation.

We apply three tests to separate veto power from broad geographic variation. First, if the correlation between senator ideology and US Attorney politicization is driven by the home state senators' veto power, then the ideology of the most liberal senator in the state—who is more likely to exercise a veto—should matter more than the ideology of the other senator. In Table 2, column 1 replicates our preferred specification from Table 1, and column 2 adds the ideology of the other senator as another independent variable. Predictably, the ideologies of the two senators are correlated, so the standard errors increase considerably and neither

---

<sup>16</sup>As reported in Appendix Table B1, using the first principal component produces very similar results ( $\hat{\beta} = .68, p < .05$ ), but it gives little weight to the employment variables. We prefer our approach because it gives equal weight to all variables.

coefficient is statistically significant. However, the coefficient for the most liberal senator's ideology remains large, increasing by 20%, while that of the other senator is negative and roughly 90% smaller.

[Table 2 about here.]

Second, if veto power is driving our results, we would expect that having one Democratic senator matters more than having a second. In column 3, we regress our politicization index on dummy variables for whether the home state has one Democrat or two Democrats (with no Democrats as the omitted category). The magnitudes of both coefficients are both large and similar in size, roughly  $.4\sigma$ ; though, only the coefficient for two democrats is statistically significant. When we collapse the one- and two-Democrat dummies together in column 4, we obtain a similar coefficient that is statistically significant. Taken together, these results suggest there is a large difference between the politicization of US Attorneys when at least one Democratic senator represents the state but little difference between having one Democrat and having two. This is consistent with veto power, but difficult to explain based on geographic variation in ideology.

Finally, if veto power is driving our results, we would expect US Attorney politicization to be correlated with the ideology of the sitting senators at the time of nomination but not with that of senators in previous or subsequent Congresses. In column 5, we add controls for the ideology of the most liberal senator in each state six years (or one term) before Bush took office (the 104<sup>th</sup> Congress, 1995-1996) and six years after (the 110<sup>th</sup>, 2007-2008). The coefficient for the ideology of senators sitting at the time of the US Attorney nominations grows larger and more statistically significant, while the coefficients for other terms are smaller, not statistically significant, and not consistently signed. Thus, our results appear to reflect more politicized US Attorneys coming from states with conservative senators at the time of the confirmation hearings, not simply coming from states that have more conservative senators in general. Below, we use this result (that only variation in senator ideology idiosyncratic to the 2001-2002 term is correlated with US Attorney politicization) to motivate a robustness check that our estimates of the effects of the Ashcroft memo do not simply reflect differential trends in more or less conservative areas of the country.

To verify these tests help separate the veto power from geographic variation in ideology, in Appendix Table B2 we regress the state-level vote shares of President George W. Bush in the 2000 election on the same specifications in Table 2. We find that the ideology of the most conservative senator is significantly correlated with Bush vote share and that including it as a control eliminates one-third of the correlation between vote shares and the ideology of the most liberal senator (in sharp contrast to our results in column 2 of Table 2 where

the coefficient grew relative to column 1); that a second Democrat has the same “effect” on vote shares as the first Democrat (in contrast to column 3 of Table 2); and that Bush vote shares are even more strongly correlated with the ideology of the most liberal senator from the 110<sup>th</sup> Congress than they are with that from the 107<sup>th</sup> Congress (in contrast to column 5 of Table 2).

Thus, interpreting Table 2 alongside Table B2 suggests that veto power is an important channel by which senator ideology influences the characteristics of nominated US Attorneys. As we describe in greater detail below, our identification strategy to estimate the effects of the Ashcroft memo follows directly from our results documented here on US Attorney politicization.

### 3.4 US Attorneys’ experience

While US Attorneys appointed in states with more conservative senators are more politicized, we find no evidence that they are less qualified. Appendix Table B3 applies our regression specifications from Tables 1 and 2 to dependent variables measuring US Attorneys’ experience and qualifications at the time of nomination. In our sample, the attorneys nominated in states where the most liberal senator was more conservative were more qualified: they had more years since law school, they were more likely to be an Assistant US Attorney being promoted from within, and they had spent a greater share of their career as prosecutors and a smaller share in private practice. None of these differences, however, are statistically significant (though, a composite index based on all five variables is,  $p < .10$ ). Separately, we tested for differences in whether the US Attorney held a clerkship after law school—a proxy for legal skill—and also find no statistically significant differences. Overall, these results provide little evidence that US Attorneys nominated under more conservative senators are less qualified than those nominated under liberal senators; they are simply more ideologically conservative and politically engaged.

## 4 Differential Implementation

With our identification strategy in hand, we next turn to our analysis of the effects of the Ashcroft memo. Legal scholar and former federal prosecutor Kate Stith observes that DOJ had little capacity to enforce the memo and therefore had to rely on “the incentives and attitudes of US Attorneys” (Stith, 2007). Given that the US Attorneys who were appointed in states with more conservative senators were significantly more politicized, we next examine whether their offices implemented the Ashcroft memo with greater fidelity.

## 4.1 Data and Measures

Our core data source is LIONS,<sup>17</sup> an administrative database used by all US Attorneys Offices since late 1998 to track case processing.<sup>18</sup> For our purposes, the most important table is a charge-level dataset containing a row for each charge filed against a criminal defendant (for more details, see Appendices A.3.1 and A.3.2). For every charge, we observe a statute string and disposition (e.g., conviction, dismissal, or acquittal). The dataset also indicates the federal district in which charges were filed and a case number assigned by the corresponding US Attorney’s Office. Because many case numbers are associated with multiple defendants, we define a “case” as any unique combination of district, case number, and defendant identifier. We use a separate district-case number-level dataset that contains a “program category” variable describing, with reasonable precision, the type of crime underlying each case.<sup>19</sup> Finally, we also merge into the charge-level dataset anonymized identifiers denoting the prosecutors assigned to each case and the sentence ultimately awarded by the judge.

We exclude all cases initiated before 1999 and after 2008. We also exclude cases with rarely used or vague program categories, and, due to a large increase in case volume during the study period, all immigration cases.<sup>20</sup> Our final sample consists of roughly 530,000 cases. Our data appendix (Appendix A.3.3) provides more detail on the LIONS data and how we define our key variables.

In announcing the memo, Ashcroft explained that prosecutors should be “fair, uniform and tough.” We view “toughness” and “uniformity” as the memo’s two core goals. Moreover, the memo admonished: “Once filed, the most serious readily provable charges may not be dismissed except [under four specified conditions].” While likely unenforceable, the memo also instructed that “charges should not be filed simply to exert leverage to induce a plea.”

*Measuring Toughness.* The memo instructed that “prosecutors must charge and pursue

---

<sup>17</sup>A redacted version of LIONS is publicly available due to FOIA lawsuits. There is relatively little academic research using LIONS (with Didwania (2020b,a, 2021a) being notable exceptions), in part because it is a complex operational database not designed for research purposes.

<sup>18</sup>Syracuse University’s “Transactional Records Access Clearinghouse” provides excerpts of the 1992 and 1998 US Attorneys’ Manual, which describe the initiation of universal LIONS usage (<https://trac.syr.edu/data/jus/eousaDataHistorical.html>.)

<sup>19</sup>There are 109 categories. As one example, there are 23 different codes for “white collar crime” (including federal procurement fraud, securities fraud, commodities fraud, health care fraud, etc.), 6 different codes for organized crime, 6 different codes for terrorism crimes, etc. For more details, see Appendix A.3.3.

<sup>20</sup>From 2001 to 2008, the number of immigration cases increased from 21% of all cases to 62%. The immigration docket also differs substantially from other federal criminal cases. Our results are generally stronger when we include immigration cases (available upon request), though we can show that this is partially due to the change in case composition. In Appendix C, we consider whether even after dropping immigration cases, the large increase in their volume can explain our main findings, and we conclude that it cannot.

the most serious, readily provable offense ... supported by the facts of the case.” It defined the “most serious offense” as the charge that would “generate the most substantial sentence.” To test whether the memo achieved its goal, we examine whether it led to an increase in the severity of the most serious charge filed and convicted in each case.<sup>21</sup> To do so, we develop a charge severity index based on United States Sentencing Commission (USSC) data on all sentences imposed for convictions of felonies and of class A misdemeanors from FY1999-FY2016 (United States Sentencing Commission, 2018). For every case in the USSC data, we observe the statute—sometimes noisily recorded<sup>22</sup>—for each convicted charge. We regress the length of incarceration sentences on roughly 3,000 dummy variables, one for each statute, and we also include a rich set of controls. The coefficients on the statute dummies reflect the marginal increase in the expected sentence as a result of securing conviction on the given charge. We normalize the coefficients so they have a standard deviation of one within each program category. We then use the resulting normalized coefficients as a charge severity index—which is far more granular than typical measures of crime severity—to test whether the Ashcroft memo led to a differential increase in the severity of the most serious charge filed and convicted. In constructing our measure of the most serious charge convicted in cases where no charges resulted in conviction, we assign the severity value of the least severe charge within the same program category to avoid excluding cases on the basis of this potentially endogenous outcome. For more details on our index, see Appendix A.4.

To examine whether any changes in prosecutorial behavioral translated into changes in the severity of sentences, we also examine whether the memo led to a differential increase in the probability of a prison sentence or the length of incarceration. Though, we note that sentencing decisions are ultimately up to the judge rather than the prosecutor.

*Measuring Uniformity.* As noted, the Ashcroft memo sought to increase uniformity in outcomes in similar cases. Relatedly, it also sought to reduce the influence of the individual

---

<sup>21</sup>Whether charges result in conviction or dismissal can be decided by the jury rather than the prosecution, but in practice, only about 2% of federal criminal cases go to trial.

<sup>22</sup>Statutes can be noisily recorded in both LIONS and in the USSC data. Our “statutes” are therefore better thought of as “statute strings.” For example, 21 U.S.C. § 960(a) prohibits several acts of drug trafficking. For violations of that provision, 21 U.S.C. § 960(b) describes the range of punishment available, depending primarily on drug type and quantity but also other case characteristics, such as whether the defendant has a prior conviction for a serious drug or violent offense. § 960(b)(2) generally imposes a minimum sentence of five years in a variety of cases, including, as described in § 960(b)(2)(A), those involving 100 grams or more of heroin. § 960(b)(1) generally imposes a minimum sentence of ten years in a variety of cases, including, as described in § 960(b)(1)(A), those involving 1,000 grams or more of heroin. While we observe convictions on § 960(b)(1)(A) and § 960(b)(2)(A), we also observe convictions on § 960(a) and § 960(b), which do not distinguish between the 100 gram and 1 kilogram quantities (nor do they distinguish between other sub-categories described within them). Because we observe these “vague” § 960(a) and (b) charges in both LIONS and the Sentencing Commission data, we estimate a severity index value for them and the index value is best thought of as a weighted average of the different potential charges that fall within each.



prosecutor on case outcomes. According to the memo’s introduction, for example, “the charges a defendant faces should not depend upon the particular prosecutor assigned to handle the case.”

To measure the uniformity of outcomes in similar cases, we first compute the standard deviation in the charge severity of the most serious convicted charge (according to our index described above) across cases within the same district, year, and program category. Second, we estimate the importance of the prosecutor assigned to a case by regressing the severity of the most serious charge resulting in conviction on district and program category fixed effects, and then computing the increase in  $R^2$  that results from including fixed effects for the lead attorney assigned to the case (see Appendix A.3.3 for details). We use these two measures to test whether the Ashcroft memo disproportionately reduced the variability in conviction outcomes and the importance of the individual prosecutor for explaining conviction outcomes.<sup>23</sup>

## 4.2 Summary Statistics

Table 3 presents summary statistics on the characteristics of cases filed before 2003, the year the memo was issued. Column 1 presents the mean of each case characteristic. Because our results are identified from within program category Column 2 presents an adjusted standard deviation of each characteristic: we regress each variable on program category fixed effects and then compute the standard deviation of the residuals. Column 3 then summarizes the overall change in the mean of each characteristic from the first two years of the Bush Administration to the last two years, again, adjusted for program category fixed effects.

[Table 3 about here.]

Column 3 provides suggestive evidence the Ashcroft memo was successful in achieving its goals. While prosecutors filed fewer total charges at the end of the Bush Administration compared to the beginning (9% of a standard deviation; henceforth  $.09\sigma$ ), there was no change in the number of charges convicted. Instead, the decrease in charges filed was accompanied by a similar decrease in the number of charges dismissed ( $.10\sigma$ ). There was also a 4 percentage point decline in the share of cases with any charges dismissed (from a

---

<sup>23</sup>One formula for  $R^2$  is  $1 - \text{var}(\varepsilon)/\text{var}(y)$  where  $y$  is the dependent variable and  $\varepsilon$  is the regression residuals. When we calculate the added  $R^2$  at the district-year level, we calculate  $\text{var}(\varepsilon)$  using only the residuals from within that district-year, but we calculate  $\text{var}(y)$  using the full sample. This is because, as we show below, variation in outcomes *decreases* as a result of treatment. By calculating  $\text{var}(y)$  using the full sample, we can avoid the case where the importance of lead attorney falls, but the  $R^2$  goes up because they are explaining a larger share of a smaller amount of variation.

base of 52%). Moreover, although the severity of the most serious charge *filed* in each case did not change, the severity of the most serious charge *convicted* rose by  $.09\sigma$ . There was also a modest 1 percentage-point increase in the likelihood a case results in incarceration and, conditional on incarceration, a 5.5 ( $.10\sigma$ ) month increase in the length of sentence. These results are consistent with the goals of the Ashcroft memo, and our analyses below test whether there is evidence for *differential* implementation depending on the politicization of the nominated US Attorney.

### 4.3 Methods

Our identification strategy to estimate the effects of the Ashcroft memo follows directly from our results on US Attorney politicization above. Specifically, we estimate the *differential* change in outcomes across districts, depending on the ideology of the most liberal senator in the district’s state. Our most straightforward estimating equation is given by:

$$Y_{p,d,t} = \theta_p + \alpha_d + \delta_t + \sum_{\tau \neq 2002} \beta_\tau (NOM_{s(d)} \times D(t = \tau)) + \varepsilon_{p,d,t} \quad (1)$$

where  $Y_{p,d,t}$  is some outcome among cases in program category  $p$  in prosecutorial district  $d$  during year  $t$ . The key cross-sectional heterogeneity comes from  $NOM_{s(d)}$ —the NOMINATE score of the most liberal senator in office in 2001-2002, which we normalize to range from zero to one, with higher numbers indicating a more conservative senator. We control for time-invariant cross-district heterogeneity with the district fixed effects ( $\alpha_d$ ) and nationwide aggregate trends with the year fixed effects ( $\delta_t$ ), and so this specification is a standard difference-in-difference approach in which we estimate differential exposure to a nationwide policy change as a function of predetermined characteristics. The core coefficients of interest ( $\beta_\tau$ ) are identified from differential post-memo changes in our outcomes of interest in states where the most liberal senator is more conservative (after accounting for differences in time invariant levels and for time trends affecting all districts equally). We also control for program category fixed effects ( $\theta_p$ ) to ensure that our results are not affected by changes in the composition of cases over time.

Below, we also report the results of several robustness tests to address potential concerns about causal inference and our measure of senator ideology.

### 4.4 Results

We begin by examining whether the Ashcroft memo increased the severity of case outcomes pursued by line prosecutors. Figure 2 shows, after the memo, a 0.11 differential

increase in the severity of the most serious convicted charge in districts where the most-liberal Senator was more conservative. The magnitude of this estimate is equal to just over half of the aggregate change documented in the summary statistics in Table 3.

[Figure 2 about here.]

We next check whether the differential increase in the severity of convicted charges translates into more severe sentences. Figure 3 shows that districts where the most-liberal senator is more conservative experienced a differential increase in the share of cases resulting in incarceration (Panel A) and in the average length of sentence imposed (Panel B).<sup>24</sup> While there is some evidence of a pre-2003 differential trend, below, we report several robustness tests that show no pre-trends and yield similar magnitudes.

[Figure 3 about here.]

It is natural to wonder whether this increase in the severity of case outcomes was driven by a change in charging or plea bargaining. In Appendix Table B5, we find a statistically significant but very small differential increase in the severity of the most serious charge filed. However, the magnitude is just a tenth the size of the increase in the severity of the most serious convicted charge and, below, we report two robustness checks in which the corresponding estimate is not statistically significant. We, therefore, find little evidence that the differential increase in the severity of the most serious charge convicted was predominantly driven by changes in charge filing. Turning to plea bargaining, unfortunately, we cannot examine differential changes in the severity of the most severe dismissed charge because no charges are dismissed in roughly half of all cases and the results are sensitive to different reasonable solutions to this problem. We can, however, examine whether there was a differential change in the probabilities that the most severe charge in a case was dismissed and a that a less severe charge was dismissed. As a drawback, we note that the structure of these variables renders them unusually noisy.<sup>25</sup> In Panel (a) of Appendix Figure B6, we find a differential reduction in the probability the most serious charge was dismissed, which could

---

<sup>24</sup>We assume a sentence of zero when no incarceration is imposed.

<sup>25</sup>Suppose, for example, there are two charges filed in a case: the more severe charge has a severity level of 0.75 and is ultimately dismissed, while the lesser charge has a severity level of 0.70 and results in conviction. Suppose further that, due to noise in our severity index, we erroneously conclude the second charge is more severe. Under these circumstances, our measure of the severity of the most serious charge will still be relatively accurate. The error is much larger, however, if we estimate whether the most serious charge was dismissed. In that case, our estimate would be 0 because the second charge wasn't dismissed but the true value is 1.

explain the increase in severity of case outcomes observed above.<sup>26</sup> Though, the estimate is not statistically significant, in part because of the noise in the outcome variable. Taken together, these results provide suggestive evidence, consistent with the text of the memo, that changes in plea bargaining, rather than charging, better explain the differential increase in the severity of case outcomes after the memo.

Next, we consider whether the memo differentially increased the uniformity of case outcomes. Figure 4 shows a differential decrease in the standard deviation of the severity of the most serious charge resulting in conviction and a differential decrease in the explanatory power of prosecutor fixed effects.

[Figure 4 about here.]

Taken together, our results provide evidence that districts in states where the most-liberal senator was relatively conservative experienced a differential increase in the severity of the most serious charge convicted and in the severity of sentences. They also indicate a differential decrease in the variation in outcomes across similar cases and the importance of prosecutorial assignments for explaining case outcomes. These findings mirror the aggregate summary statistics shown above, and suggest that the Ashcroft memo had a greater impact in the districts where senators' veto power led to a more politicized US Attorney.

## 4.5 Robustness

We report the results of two primary robustness checks. First, liberal and conservative states differ on a variety of dimensions, and some of those dimensions may have changed around the time of the Ashcroft memo, which our district fixed effects would not account for. To address this concern, we control for the interaction between year effects and the NOMINATE score of senators *in office one term before and after the start of the Bush Administration*. This approach isolates variation in NOMINATE scores that is idiosyncratic to the 2001-2002 term and is uncorrelated with broader political geography. As shown in the map in Appendix Figure B1, this idiosyncratic variation is substantially different than the full variation shown in Figure 1. For instance, states like California and Texas no longer contribute to identification because they experienced no change in their most liberal senator over this period. Instead, identification is now driven by comparing states like Ohio and Colorado (both of which happened to be in a brief period of unified Republican control of

---

<sup>26</sup>In Panel (b) we also find a statistically insignificant reduction in the probability that a less severe charge was dismissed. One possible but counter-intuitive explanation is that, before the memo, some prosecutors may have filed lesser charges in anticipation of dismissing the most severe charge but ceased doing so after the memo discouraged such dismissals.

their Senate seats at the time Bush nominated US Attorneys) with states like North Carolina and Georgia (both of which are frequently Republican-held states, but had a Democratic Senator at the time of nominations). It is important to note, however, that restricting to variation idiosyncratic to 2001-2002 eliminates about 65% of the total variation in senator ideology. Unsurprisingly, then, these results tend to be much less precise than our main specification.

Second, because our estimated NOMINATE scores are somewhat noisy, we also present regressions in which the NOMINATE score is replaced with a dummy variable indicating whether the state had two Republican senators in 2001-2002 and therefore no inter-party oversight.

Table 4 presents our implied constant treatment effects, and our dynamic event study plots are in the appendix (see Figures B2-B4). Panel A reports the results of our primary specifications. Unlike in our primary specification, the independent variables in our robustness checks lack sufficient variation to interact program category fixed effects with district fixed effects.<sup>27</sup> As a baseline comparison, therefore, Panel B shows the results of our primary specification without interacting program category and district. Panels C and D show the results of the idiosyncratic-variation specification and the party-only specification, respectively. To facilitate interpretation of magnitudes, we note that one can compare effect sizes between Panels A through C but not with D.<sup>28</sup> For all outcome variables, the results are reasonably robust across specifications. While most specifications do not yield statistically significant effects across *all* outcomes, the signs of our estimated effects never change and the magnitudes are quite stable.

[Table 4 about here.]

Interestingly, when we use these same specifications for an instrumental variables strategy—instrumenting for our index of politicization for the appointed US Attorney of the district—the differences in magnitudes across specifications all shrink (see Table B4).<sup>29</sup>

---

<sup>27</sup>The reasons is that the robustness checks already ignore a large share of the cross-sectional variation in senator ideology. In the first case, the variation idiosyncratic to the 107<sup>th</sup> Congress is only 35% of the total variation in ideology. In the second case, as shown in column 4 of Table 2, the party-only variation has only half as much influence over US Attorney politicization as the ideology variation has.

<sup>28</sup>In Panels A through C, the key independent variable measures the ideology of the most-liberal senator in a state. For Panel D, the key variable is a dummy indicating whether there were two Republican senators in a given state at the time.

<sup>29</sup>In other words, the reason that our party-only estimates (Panel D) tend to be smallest estimates and our idiosyncratic variation estimates (Panel C) tend to be the largest could be that the party-only variation has the smallest effects on US Attorney politicization while the idiosyncratic variation has the largest effects (as shown in Table 2). The IV estimates that account for this produce very consistent and stable magnitudes across different specifications, though they are not always statistically significant.

Finally, common pre-trends are more compelling when the treatment and control group series move together. For each outcome that we consider, we first residualize out district-by-program-category fixed effects, and then calculate the average level of the residual among districts covered by at least one Democrat compared to those represented only by Republican senators. Figure B5 shows the over-time patterns of our outcomes for these two types of districts. The figure shows that even when pre-trends are not perfectly parallel, the two series tend to move together in fairly consistent ways.<sup>30</sup>

## 5 Differential prosecutorial turnover

We next seek to quantify the organizational costs of the Ashcroft memo by examining whether it caused more Assistant US Attorneys to leave the Justice Department. As noted, both the *New York Times* and the *Washington Post* reported that many federal prosecutors were frustrated by the memo. At the time, one former Deputy Attorney General said their “anger is immense” and predicted “a huge revolt.”

These sentiments are broadly consistent with the findings of academic research in other employment contexts. Both surveys (Maestas et al., 2017; Nikolova and Cnossen, 2020) and real-world decisions (Stern, 2004) of workers, and especially educated workers, show they value autonomy and discretion in deciding whether to take a job. Relatedly, influential models in occupational psychology posit that losing discretion, or “job decision latitude,” is stressful—particularly in demanding positions (Karasek, 1979). Standard measures of management quality thus emphasize worker autonomy (Bloom and Van Reenen, 2007; Rasul and Rogger, 2018), and, more generally, higher management quality improves retention of the highest skilled employees (Bender et al., 2018; Lazear et al., 2015). For all these reasons, we suspect that some prosecutors left the Justice Department in response to efforts to curb their discretion.

Of course, other contextual factors may have interacted with the memo to magnify its effects. For example, Ashcroft’s relationship with career prosecutors was, perhaps, more strained than that of other Attorney Generals before him. One journalistic account reported that “the scorn that Ashcroft’s aides feel for ‘the bureaucracy’—the career attorneys and employees—... helped to create a tense environment within the Justice building” (Toobin, 2002). It also reported that some prosecutors perceived Ashcroft was inappropriately politicizing the Justice Department, a view expressed by then-Chair of the Senate Judiciary Committee Senator Patrick Leahy: “In many ways, Ashcroft has created a far more politi-

---

<sup>30</sup>Note that because we residualize out district-by-program-category fixed effects, both series are mechanically mean zero over time.

cal department than any Justice Department since I’ve been here [28 years].” Anecdotally, prosecutors were also frustrated with Ashcroft’s distinctively religious views on social and policy issues. This rich context is not a threat to our identification strategy, but it may bear on external validity. It is possible that efforts to constrain the discretion of government bureaucrats would have smaller effects on departures if enacted by a principal who is more trusted and respected by those bureaucrats.

## 5.1 Data and Measures

LIONS contains two datasets needed to measure departures. First, the staff table assigns an anonymous identifier to each staff member in the US Attorney’s Office and also indicates their title, including a code for Assistant US Attorneys.<sup>31</sup> Second, the assignment table identifies staff members assigned to each case. We code the last year an attorney receives an assignment as their final year in the US Attorney’s Office. Because new assignments are frequent, a median of 48 cases per year for each attorney, we detect the end of employment with reasonable precision for the vast majority of departures. Based on our estimates, roughly 300 prosecutors leave the US Attorney’s Office each year, which is just shy of the average of 340 annual permanent-contract attorney departures per year, from FY2005-FY2010, reported by the Office of Personnel Management (OPM).

We assume all these departures are voluntary. Indeed, Assistant US Attorneys are afforded the full protections of the federal civil service and it is “very hard to fire” them (Miller and Curry, 2019). According to aggregate data from the OPM, of the 1,700 permanent-contract attorneys who left the US Attorney’s Office from FY2005-FY2010, none were fired or dismissed.

## 5.2 Effects on Departures

To estimate the effect of the memo on departures, we use a Poisson pseudolikelihood regression where the dependent variable is the log of the number of Assistant US Attorneys who departed in a given district-year from 1999 to 2008, and the independent variables follow the same difference-in-difference specification above.<sup>32</sup> We also present estimates using only the idiosyncratic variation (comparable to Panel C in Table 4 above) and the

---

<sup>31</sup>Only 70% of staff members assigned to cases are Assistant US Attorneys. The remaining assignments are for paralegals, administrative assistants, inter-agency liaisons, etc.

<sup>32</sup>We prefer a Poisson pseudolikelihood regression over simply analyzing the log number of departures because roughly 17% of district-year observations show no departures. One alternative is to calculate the *fraction* of Assistant US Attorneys who leave. The results are similar. For details on the Poisson pseudolikelihood regression, see Correia, Guimarães, and Zylkin (2019, 2020).

party-only variation (comparable to Panel D). Because US Attorney’s Offices have vastly different numbers of Assistant US Attorneys, we always weight each office by the average number of attorneys employed annually during the study period.

Figure 5 presents the estimated effects on departures. It shows reasonably flat pre-trends before departures begin differentially increasing in conservative-senator districts in 2005.<sup>33</sup> Given our finding that the memo began affecting case outcomes in 2004, we suspect that Assistant US Attorneys spent about a year deciding whether to leave and looking for other jobs, a reasonable timeline for a highly skilled labor force. Though, it is also possible some Assistant US Attorneys were waiting to see whether Bush would be reelected in 2004.

[Figure 5 about here.]

The magnitude of our estimates are quite large, peaking in 2007 at around 40 log points. The average US Attorney’s Office has 64 Assistant US Attorneys and roughly five departures per year. Comparing the most conservative-senator district with the most liberal, our estimates imply that departures increased to roughly 8 per year, nearly a 50% difference. Although this trend would not cripple an office, it could impair its functioning.<sup>34</sup>

### 5.3 Differential Implementation by Leavers and Stayers

If the memo caused prosecutors to leave the US Attorney’s Office, we would expect that the prosecutors who left were less compliant with the memo’s guidance than those who stayed. One challenge in testing this hypothesis is that prosecutors receive different case assignments as they accumulate experience over time. Because departing attorneys tend to be more experienced than non-departing ones, we must be careful to disentangle differences in prosecutorial behavior due to caseload from differences in prosecutorial behavior between those who left and those who stayed.

---

<sup>33</sup>Above, our fully flexible difference-in-difference plots presented implied constant treatment effects and their significance. That was because the Ashcroft memo meant a permanent shift from one policy regime to another. Here, we study departures, and we view these as short-run responses to the *transition* from one policy regime to another. We do not present implied constant treatment effects because we do not expect an increase in departures to be sustained as a new status quo.

<sup>34</sup>Our back-of-the-envelope estimate is that the memo drove over a hundred prosecutors to leave. Roughly 300 prosecutors leave each year. We estimate that in the most conservative district ( $NOM = 1$ ), relative to the most liberal ( $NOM = 0$ ), the memo increased departures by between 10 and 35 log points per year over four years (or roughly 20 log points per year for four years). Since the average NOMINATE score is roughly 0.5, relative to a minimal-implementation counterfactual ( $NOM = 0$ ), this is 10 log points per year, or an increase of roughly 30 departures per year over the baseline of 300. Thus, this rough calculation implies roughly 120 departures due to the Ashcroft memo, relative to a counterfactual where all districts showed minimum ( $NOM = 0$ ) implementation.



As a partial solution, we focus our analysis on the most senior Assistant US Attorneys we can isolate in the data. Specifically, we identify those who were working in the US Attorney’s Office in 1998—when LIONS was adopted—and who continued to work through at least 2006 so we can observe at least three years of their behavior after the memo. We separate these attorneys into two groups: “leavers,” who left in 2007 or 2008, and “stayers,” who remained until the end of 2008 when Bush left office. For each group, we separately estimate our earlier regression on the severity of the top charge resulting in conviction and, this time, include  $\eta_a$  fixed effects for lead Assistant US Attorney to isolate changes in an individual prosecutors’ behavior:

$$Y_{c,d,t} = \alpha_{d,p} + \delta_t + \eta_a + \sum_{\tau \neq 2002} \beta_{\tau} (NOM_{s(d)} \times D(t = \tau)) + \varepsilon_{c,d,t} \quad (2)$$

where  $\alpha_{d,p}$  denotes district-by-program-category fixed effects,  $\delta_t$  denotes year fixed effects, and  $\eta_a$  denotes lead attorney fixed effects.

Figure 6 presents the results. Panel (a) shows that, among cases assigned to “stayers,” the severity of the most serious convicted charge differentially increased by a statistically significant 0.14 standard deviations, somewhat more than the overall effect estimated above. Panel (b), however, shows no such change for leavers; though, the estimates are less precise because there are a fifth as many leavers as stayers. Taken together, the results suggest that it was the stayers, and not the leavers, who implemented the memo. If anything, during their final years in office, the leavers in conservative-Senator districts—where we estimate departures were concentrated—obtained *less* severe convictions than their counterparts in liberal-Senator districts, which could reflect a backlash to the controversial policy or poor morale.

[Figure 6 about here.]

## 5.4 Who Leaves?

We next examine who left as a result of the memo, an important question for at least two reasons. First, the number of departures caused by the memo may understate its true costs if the attorneys who left were particularly experienced or skilled.<sup>35</sup> Second, at least to the Bush administration, increased departures may have represented a policy success if

---

<sup>35</sup>We note that turnover is a major challenge faced by many organizations, particularly those relying on highly skilled and specialized workers (Hoffman and Tadelis, 2021). Thus, even if the marginal departures are not systematically different from regular departures, an increase in the rate of departures can still have important organizational consequences.

certain prosecutors—specifically, those who were low-performing—were more likely to leave. To advance these goals, we developed a series of performance measures at the attorney level.

To create measures of experience and skill, we spoke with three former Assistant US Attorneys who proposed eleven potential measures, eight of which we could reliably implement in LIONS (for more details, see Appendix A.6.1). We tie our hands and analyze all eight. The first is a proxy for years of experience in the US Attorney’s Office. We measure whether someone was an Assistant US Attorney in 1999 (64% of departing attorneys), the first full year when LIONS was used and, therefore, the point at which our measure of experience is censored. Our remaining seven variables focus on the number and kinds of cases assigned to prosecutors. All are constructed based on cases filed between 2000 and 2002 (i.e., pre-memo) to ensure they are uncontaminated by the effects of the memo. First, we measure the total number of cases assigned to each prosecutor as lead attorney (mean of 144). Second, we measure the number of cases each prosecutor took to trial, a predictor of whether prosecutors later secure positions in large firms (Boylan and Long, 2005). Because 95% of convictions are secured through a guilty plea, trials are scarce opportunities to gain critical litigation skills (Wright and Levine, 2014). Third, we measure the number of times prosecutors were assigned cases with multiple defendants, an indicator of case complexity (12% of cases). Fourth, we measure the number of times prosecutors were assigned a case with a priority of “Main Justice,” DOJ’s central office (0.5% of cases).<sup>36</sup> Fifth, we measure the number of times each prosecutor was assigned a case with multiple investigative agents, another indicator of case complexity (5% of cases). Sixth, we observe the number of cases in which the prosecutor filed a motion for a “5K1.1 Credit” on behalf of the defendant, indicating that the defendant provided substantial assistance in another investigation (5% of cases). This indicates not only that the case was sufficiently complicated to warrant gathering evidence across different defendants but also that the prosecutor was able to obtain this information from the defendant.<sup>37</sup> Finally, we measure the number of times a prosecutor is assigned a case in the Public Corruption or Organized Crime program categories, which are typically more prestigious and complex (7% of cases).

With our attorney-level outcome measures in hand, we next develop a regression to compare the characteristics of the attorneys who left because of implementation of the memo to the characteristics of those who would have left regardless. To do so, we construct a

---

<sup>36</sup>To do so, we observe whether DOJ reported monitoring and tracking the case or whether Main Justice attorneys joined in the prosecution.

<sup>37</sup>One former Assistant US Attorney we spoke to doubted whether this measure was informative because 5K1.1s are recorded in the case against the cooperating witness rather than the case in which that witness is testifying. Our composite index has a correlation of .988 with the same index excluding the 5K1.1 variable. The variable thus makes little difference for our results.

person-level dataset of every Assistant US Attorney who was employed in 2002—the year before the memo issued—and also departed between 2002 and 2008. Similar to our previous analyses, we regress our outcomes variables on the interaction of Senator ideology and a dummy indicating whether the prosecutor departed after the memo (rather than before). We include district and year of departure fixed effects. In our preferred specification, we define the pre-period as 2002 to 2004 and the post-period as 2005 to 2008, reflecting the timing of the memo’s effects on departures observed in Figure 5. The results are substantively similar, however, when we define the pre-period as 2001 to 2003. Our estimating equation is as follows:

$$X_{a,d,t} = \alpha_d + \delta_t + \beta[NOM_{s(d)} \times D(2005 \leq t \leq 2008)] + \varepsilon_{a,d,t} \quad (3)$$

where  $X_{a,d,t}$  is some characteristic of attorney  $a$  who left district  $d$  during year  $t$ .<sup>38</sup> As in our analysis of differential implementation, the  $\alpha_d$  fixed effects account for the fact that the attorneys leaving different districts might differ from one another for reasons unrelated to policy implementation (e.g., the most skilled attorneys might work in populous areas, which have both more lawyers and senators who are more liberal). The  $\delta_t$  fixed effects account for the fact that nationwide departure patterns may differ across years, for instance because of high levels of overall turnover after the change in Presidential administration.

Our core coefficient of interest is  $\beta$ , which captures how much a given characteristic of departing attorneys differentially changes, from before to after the memo, in conservative-senator districts relative to liberal ones. If a given characteristic differentially changes—after accounting for pre-memo cross-district differences and nationwide post-memo patterns—then we conclude that the marginal attorney who quit because of the memo differed from the average attorney who would have left regardless.

Figure 7 presents estimates of the  $\beta$  coefficients of interest for all of our outcome variables.

---

<sup>38</sup>Standard methods for characterizing the compliers in an instrumental variables regression (e.g., Angrist and Pischke (2008)) are developed for a randomly assigned binary instrument and a binary treatment. For reasons discussed in Abadie (2003) and Marbach and Hangartner (2020), they cannot be used with a continuous instrument, like Senator ideology. In the canonical case of a randomly assigned binary instrument, our approach is equivalent to regressing the characteristics of program participants on whether they were assigned the instrument. In Appendix A.6.2, we show that the  $\beta$  coefficient from that regression (i.e., the coefficient on the instrument) is proportional to the difference in average characteristics between the compliers and the “always takers,” where the non-negative multiplier is the share of program participants who are compliers (a monotonic transformation of the traditional “complier share”). Because this coefficient is positive under monotonicity, the sign of the  $\beta$  coefficient is the same as from the traditional calculation, and as long as the first stage is sufficiently strong, the statistical significance is roughly equal. Thus, our approach closely replicates the intuition of the traditional approach to characterizing compliers. We recover the correct sign of the difference and an approximation of the correct statistical significance. However, our magnitudes will be systematically under-estimated by an unknown amount that depends on the complier share and the “always taker” share, neither of which can be calculated in the presence of a continuous instrument.

With the exception of the binary measure of experience, each variable has been normalized to a standard deviation of one. Of the eight characteristics we consider, seven generate positive point estimates (all but the number of multi-defendant cases), and five of those point estimates are greater than 0.10 standard deviations (three greater than 0.20), although only one is statistically significant. When we create an index of these characteristics, however, we estimate that marginal departures induced by the memo were 0.22 standard deviations more skilled and experienced than regular departures ( $p < .05$ ). Put differently, the Ashcroft memo’s reduction of prosecutorial discretion appears to have pushed out particularly well-qualified US Attorneys.

[Figure 7 about here.]

## 6 Alternative Explanations

We have shown that, after the Ashcroft memo, conservative-senator districts experienced a differential increase in the severity of the most serious convicted charge, a differential reduction in the variation of case outcomes across similar cases and across prosecutors, and a differential increase in prosecutorial departures. Given federal prosecutors’ vociferous objections to the memo when it was issued, we interpret these results as the causal effects of the memo.

Nonetheless, other important events happened during the Bush Administration, and they plausibly had differential effects in different places. In Appendix C we take seriously four plausible alternative explanations for our findings: the increase in immigration cases, the *United States v. Booker* Supreme Court decision increasing judicial discretion, the controversial firing of seven US Attorneys, and the Thompson Memo guiding criminal charges filed against corporations. In all four cases, we develop a series of tests, and present a range of evidence suggesting that none of these explanations can explain our empirical results.

## 7 Conclusion

Our analysis provides evidence that a policy designed to constrain the discretion of federal prosecutors achieved its goal of increasing the severity and uniformity of case outcomes. From the perspective of the administration that issued the memo, however, this success did not come without costs. Indeed, we find evidence that the memo also had another, more subtle, effect on the organization: it encouraged a substantial number of line prosecutors to quit—particularly those who specialized in important and complex cases.

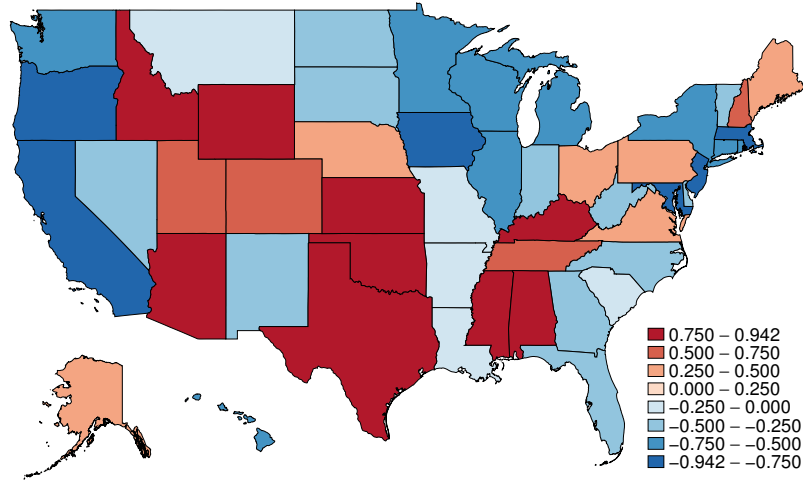
These results have important implications. First, a growing body of evidence in economics suggests that the skills and preferences of bureaucrats and the extent of their professional discretion are important predictors of organizational success. Best et al. (2019), for example, find large variation in performance quality across government bureaucrats in procurement. Other papers, on both the public and private sector, provide evidence that giving workers more discretion is associated with better performance (Bandiera et al., 2021; Bloom and Van Reenen, 2007). But much of this literature holds fixed the stock of workers. Our results highlight how a policy can affect agency outcomes, not only by changing worker performance, but also by changing the workforce itself—by encouraging workers to quit. Public sector agencies should take these effects into account when designing a policy to constrain discretion, particularly when it is politically motivated and targeted at bureaucrats with unaligned ideological backgrounds.

A second strand of work has also found that large-scale shifts in policy—specifically, when the White House transfers from one political party to another—drive federal bureaucrats to quit in large numbers (Bolton, de Figueiredo, and Lewis, 2020; Doherty, Lewis, and Limbocker, 2019; Richardson, 2019). Our results show how far more modest shifts in policy that constrain worker discretion can increase departures, too.

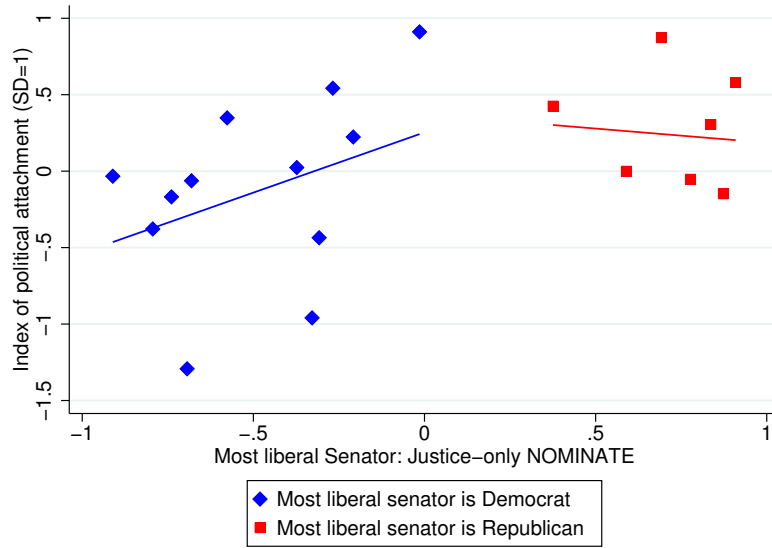
Finally, prosecutors are known to have wide discretion in their case work, and we have relatively little systematic empirical evidence about whether and how this discretion can be controlled by supervisors. Our results suggest that the Attorney General’s charging and plea bargaining memo—perhaps the most publicly visible effort to constrain prosecutorial discretion—can change the behavior of line prosecutors only if the US Attorney in their office is committed to enforcing it.

Figure 1: Senator Ideology and US Attorney Politicization

(a) Variation in most liberal senators' NOMINATE scores

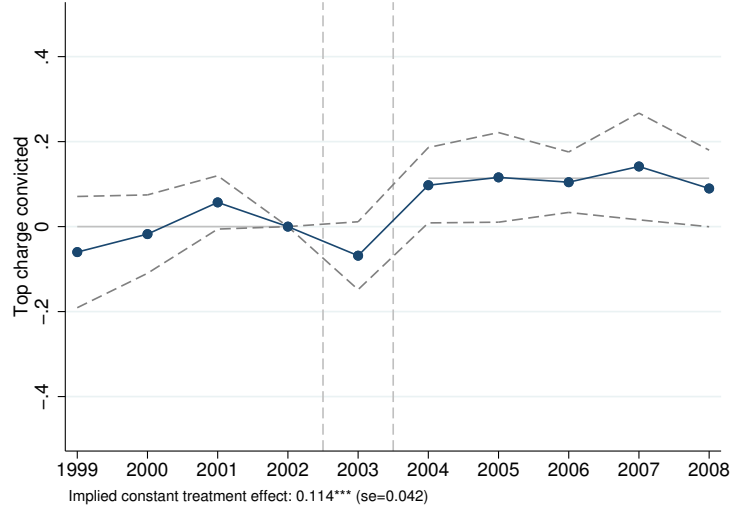


(b) Ideology and Politicization



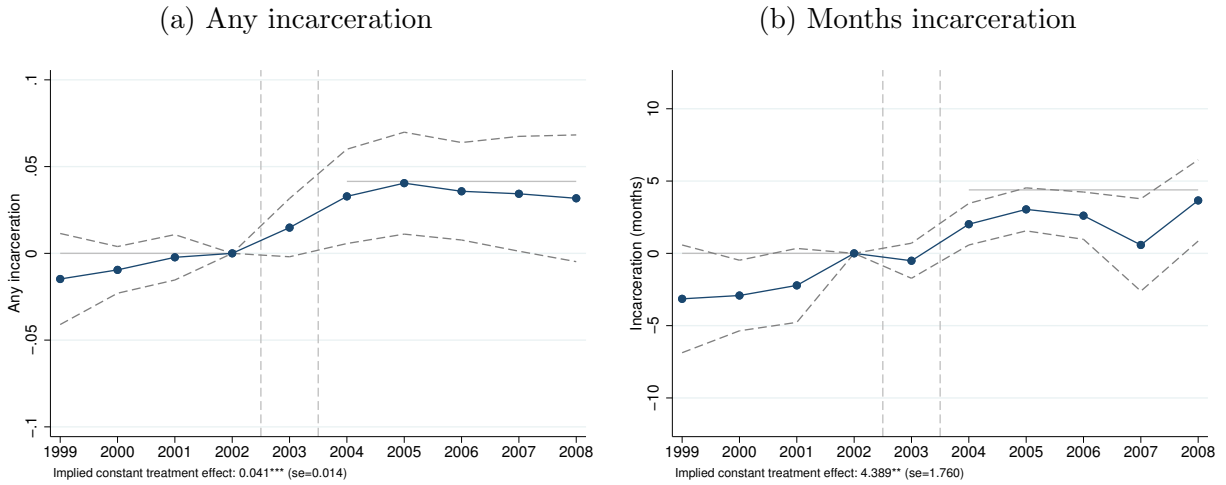
*Notes:* Panel (a) displays our justice-specific NOMINATE scores, which measure ideology from roll call votes on bills related to justice issues, for the most liberal Senator in the state during the 107<sup>th</sup> Congress (2001-2002). Panel (b) presents a binned scatterplot showing the correlation between the ideology of the most liberal senator in the state, as measured by NOMINATE scores, and an index of US Attorney politicization that combines five variables measuring campaign contributions and career histories (shown in Table 1).

Figure 2: Effects on most serious charge resulting in conviction



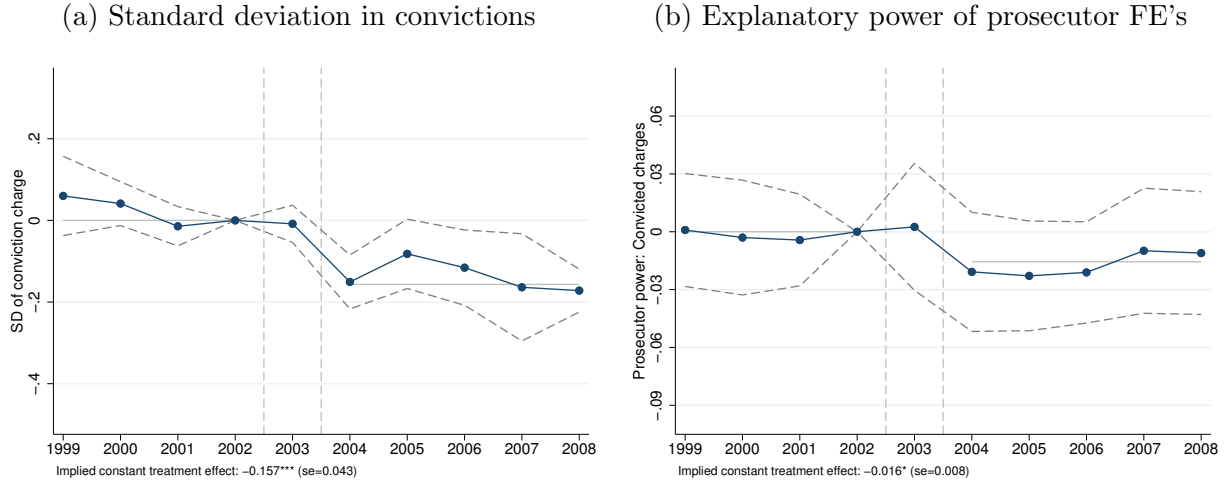
*Notes:* Unit of observation is district-by-year-by-program-category. Observations are weighted by the number of cases. Regressions include year and district-by-program-category fixed effects. Standard errors are clustered at the state level and the program category level. Plotted coefficients are based on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator.

Figure 3: Effects on sentencing outcomes



*Notes:* Unit of observation is district-by-year-by-program-category. Observations are weighted by the number of cases. Regressions include year and district-by-program-category fixed effects. Standard errors are clustered at the state level and the program category level. Plotted coefficients are based on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator.

Figure 4: Effects on charging variation across cases

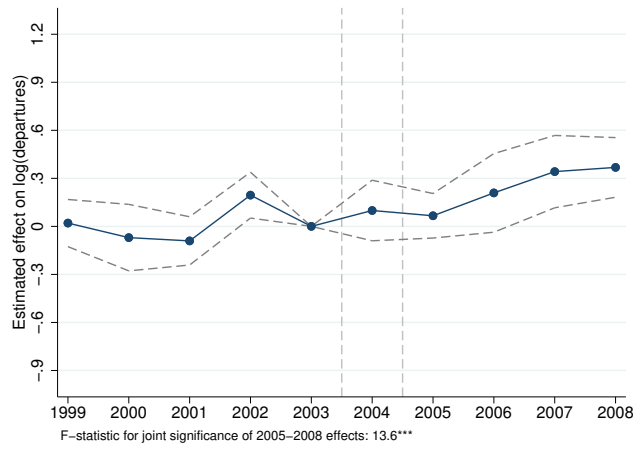


*Notes:* Panel (a): Unit of observation is district-by-year-by-program-category. Regressions include year and district-by-program-category fixed effects. Standard errors are clustered at the state level and the program category level. Panel (b): Unit of observation is district-by-year. Regressions include year and district fixed effects. Standard errors are clustered at the state level. Both: Observations are weighted by the number of cases. Plotted coefficients are based on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator.

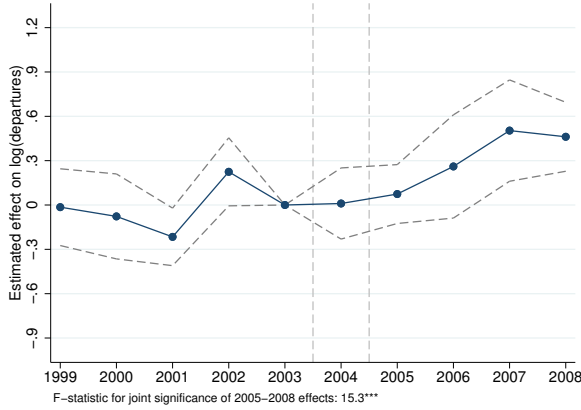


Figure 5: Effects on attorney departures

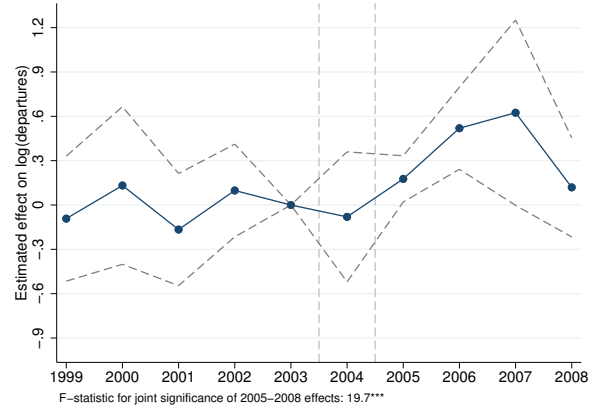
(a) Full variation



(b) Party-only variation

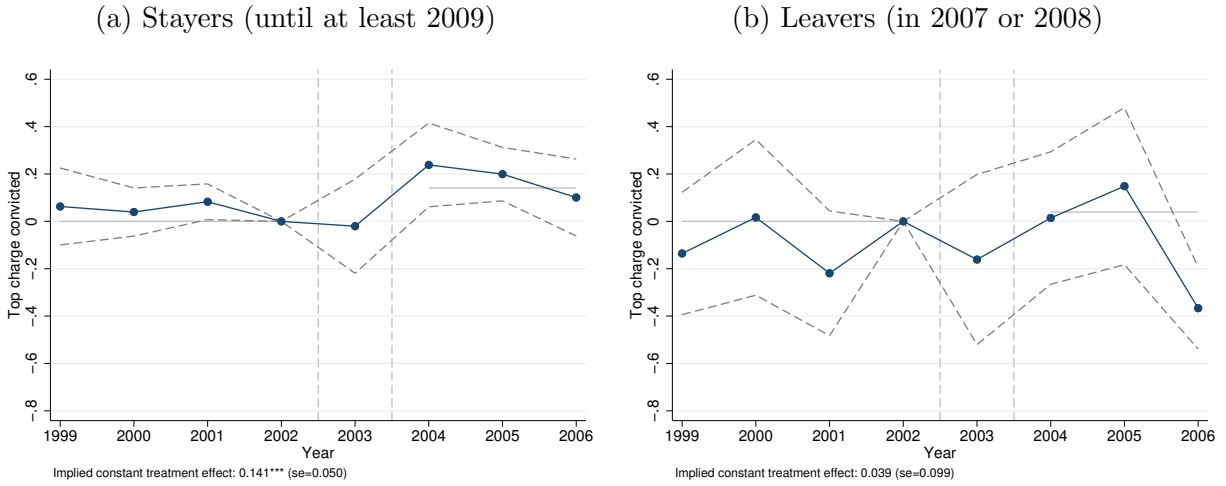


(c) Idiosyncratic variation only



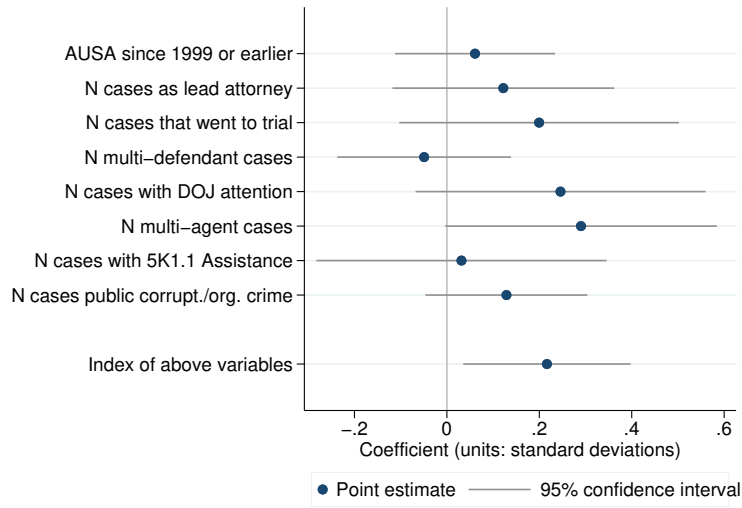
*Notes:* Unit of observation is district-year. Observations are weighted by the average number of Assistant US Attorneys in the district (averaged across years). Estimates come from a Poisson pseudolikelihood regression in which the dependent variable is the number of attorneys departing the district in a given year. All regressions include district and year fixed effects. Standard errors are clustered at the state level. Plotted coefficients are on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator. “Full variation” estimates use raw NOMINATE scores. “Party-only” estimates use a dummy for whether the state had two Republican Senators. “Idiosyncratic variation” estimates also control for NOMINATE scores from 1995 and 2007 interacted with year effects (see Figure B1).

Figure 6: Differential responses of leavers and stayers



*Notes:* Unit of observation is attorney-by-year-by-program-category. Observations are weighted by the number of cases. Regressions include year, attorney, and district-by-program-category fixed effects. Standard errors are clustered at the state level and the program category level. Plotted coefficients are based on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator.

Figure 7: Differential specialization of induced departures compared to average departures



*Notes:* Unit of observation is the departing Assistant US Attorney. Figure plots coefficients on  $NOM_{s(d)} \times Dep\{ '05 - '08 \}_i$  from 9 different regressions in which departing attorneys' characteristics are regressed on district and year-of-departure fixed effects and the interaction of NOMINATE scores and a dummy variable indicating that a given attorney left between 2005 and 2008. The coefficients represent the differential specialization of attorneys induced to leave in response to the memo compared to the average of attorneys leaving from 2002-2008. All variables except "AUSA since 1998 or earlier" are normalized to have a standard deviation of one across departing assistants. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator and one in the state with the most-conservative-most-liberal senator.

Table 1: States with more liberal senators are appointed less politicized US Attorneys

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
DV:	Campaign contributions			Political jobs			Index
	Amount (pct.)	Repub. share	CF score	Work for politician	Elected office	Either	
Senator ideology	0.045 (0.094)	0.096** (0.036)	0.275** (0.118)	0.087 (0.088)	0.157 (0.155)	0.271 (0.168)	0.671** (0.261)
Constant	0.479*** (0.063)	0.902*** (0.033)	0.758*** (0.082)	0.086* (0.050)	0.133* (0.074)	0.172** (0.079)	-0.314* (0.184)
$R^2$	0.003	0.043	0.065	0.008	0.017	0.039	0.050
N	87	74	74	87	87	87	87

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . Unit of observation is the US Attorney. Standard errors are clustered at the state level and shown in parentheses. NOMINATE variables have been normalized to have minimum zero (most liberal) and maximum one (most conservative) within the sample. Index is based on the first five variables and normalized to a mean of zero and a standard deviation of 1. Contribution amount is normalized as percentiles within the US Attorney distribution. “CF Score” is a continuous measure of how conservative a profile of campaign contributions is (Bonica, 2014).

Table 2: Separating geography and veto power

DV: Political index	(1)	(2)	(3)	(4)	(5)
Most liberal (107th)	0.671** (0.261)	0.801 (0.491)			1.054*** (0.297)
Most conserv. Senator		-0.095 (0.315)			
2 Democrats			-0.353* (0.208)		
1 Dem./1 Rep.			-0.432 (0.309)		
2 Dem. or 1D/1R				-0.383** (0.190)	
Most liberal (104th)					0.004 (0.530)
Most liberal (110th)					-0.460 (0.435)
$R^2$	0.050	0.052	0.035	0.034	0.061
N	87	87	87	87	87

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . Unit of observation is the US Attorney. Standard errors are clustered at the state level and shown in parentheses. Dependent variable is the index of the five variables discussed in Table 1.

Table 3: Case-level summary statistics

	(1)	(2)	(3)
Variable	Pre-memo mean (1999-2002)	Pre-memo standard deviation	Change in mean during Bush Administration
Number of charges filed	2.04	1.46	-0.13**
Number of charges convicted	1.16	0.82	-0.01
Number of charges dismissed	0.86	1.21	-0.12***
Any charge dismissed	0.52	0.49	-0.04**
Severity index: Top charge filed	6.24	0.83	0.03
Severity index: Top charge convicted	5.25	2.27	0.20***
Any incarceration	0.69	0.43	0.01*
Months of incarceration (if any)	59.4	55.4	5.5***

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . Unit of observation is the case.  $p$ -values are based on two-way clustered standard errors, clustered at the district and program category levels. Columns 1 and 2 are based on 226,000 pre-memo cases (1999-2002). Columns 2 and 3 are computed within program category (i.e., crime type): we first regress each variable on program-category fixed effects and then compute the standard deviation and change in mean based on the resulting residuals. Column 3 is based on 225,000 cases and represents the change in the average level of the residualized variable from 2001/2002 to 2007/2008.

Table 4: Differential implementation: Robustness to specification changes

	(1)	(2)	(3)	(4)	(5)
	Conviction severity	Any incar.	Months incar.	SD of conv. severity	$R^2$ from pros. FE's
<b>Panel A:</b> District-by-program-category fixed effects (preferred specification)					
$Post_t \times NOM_s$	0.114*** (0.042)	0.041*** (0.014)	4.389** (1.760)	-0.157*** (0.043)	-0.016* (0.008)
$R^2$	0.951	0.796	0.888	0.747	0.491
<b>Panel B:</b> Basic specification (district, year, and program category fixed effects)					
$Post_t \times NOM_s$	0.078 (0.053)	0.030* (0.017)	3.399* (1.755)	-0.150*** (0.051)	-0.018* (0.010)
$R^2$	0.924	0.688	0.770	0.638	0.497
<b>Panel C:</b> Only senate ideology variation idiosyncratic to 2001-2002					
$Post_t \times NOM_s$	0.113*** (0.042)	0.021* (0.012)	4.432*** (0.526)	-0.110 (0.075)	-0.020 (0.016)
$R^2$	0.924	0.688	0.771	0.639	0.503
<b>Panel D:</b> Only senate ideology variation due to party					
$Post_t \times TwoRep_s$	0.053** (0.025)	0.017* (0.010)	2.094** (1.026)	-0.092*** (0.034)	-0.008 (0.006)
$R^2$	0.924	0.688	0.770	0.638	0.496

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . Columns 1-4: Unit of observation is district-by-year-by-program-category. Standard errors are clustered at the state level and the program category level. Column 5: Unit of observation is district-by-year. Standard errors are clustered at the state level. All columns: Observations are weighted by the number of cases. Plotted coefficients are based on justice-specific NOMINATE scores interacted with a “post-memo” dummy (2004-2008). NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator. Panel B (basic specification) includes district, program category, and year fixed effects. Panel A modifies panel B by using district-by-program-category fixed effects. Panel C modifies panel B by controlling for senator ideology from other terms (1995-1996 and 2007-2008) interacted with year effects to isolate variation idiosyncratic to the Senators initially in office when Bush took office. Panel D modifies panel B by using a dummy for whether the state had two Republican senators, rather than using NOMINATE scores. For panel A, corresponding event study plots are in the text. For panels B, C, and D, corresponding plots are in Figures B2-B4 of the appendix.

## References

- Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of econometrics* 113(2), 231–263.
- Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics*. Princeton university press.
- Bandiera, O., M. C. Best, A. Q. Khan, and A. Prat (2021). The allocation of authority in organizations: A field experiment with bureaucrats. *The Quarterly Journal of Economics* 136(4), 2195–2242.
- Barkow, R. E. (2008). Institutional design and the policing of prosecutors: Lessons from administrative law. *Stanford Law Review* 61, 869.
- Bellin, J. (2019). The power of prosecutors. *NYUL Rev.* 94, 171.
- Bender, S., N. Bloom, D. Card, J. Van Reenen, and S. Wolter (2018). Management practices, workforce selection, and productivity. *Journal of Labor Economics* 36(S1), S371–S409.
- Berdejo, C. (2018). Criminalizing race: Racial disparities in plea-bargaining. *Boston College Law Review* 59, 1187–1249.
- Best, M. C., J. Hjort, and D. Szakonyi (2019). Individuals and organizations as sources of state effectiveness. *NBER Working Paper* 23350.
- Bibas, S. (2009). The need for prosecutorial discretion. *Temple Political and Civil Rights Law Review* 19, 369.
- Binder, S. A. (2007). Where do institutions come from? exploring the origins of the senate blue slip. *Studies in American Political Development* 21(1), 1.
- Bjerk, D. (2017). Mandatory minimums and the sentencing of federal drug crimes. *The Journal of Legal Studies* 46(1), 93–128.
- Bloom, N. and J. Van Reenen (2007). Measuring and explaining management practices across firms and countries. *The quarterly journal of Economics* 122(4), 1351–1408.
- Bolton, A., J. M. de Figueiredo, and D. E. Lewis (2020). Elections, ideology, and turnover in the us federal government. *Journal of Public Administration Research and Theory*.
- Bonica, A. (2014). Mapping the ideological marketplace. *American Journal of Political Science* 58(2), 367–386.
- Bonica, A. (2019). Database on ideology, money in politics, and elections (dime).
- Bonica, A. and M. Sen (2017). The politics of selecting the bench from the bar: the legal profession and partisan incentives to introduce ideology into judicial selection. *The Journal of Law and Economics* 60(4), 559–595.
- Boylan, R. T. (2005). What do prosecutors maximize? evidence from the careers of u.s. attorneys. *American Law Economics Review* 7, 379–402.
- Boylan, R. T. and C. Long (2005). Salaries, plea rates, and the career objectives of federal prosecutors. *Journal of Law Economics* 48, 651–627.
- Cameron, C. M. and J. M. de Figueiredo (2020). Quitting in protest: Presidential policy-making and civil service response. *Quarterly Journal of Political Science* 15(4), 507–538.
- Caughey, D. and E. Schickler (2016). Substance and change in congressional ideology: Nominate and its alternatives. *Studies in American Political Development* 30(2), 128–146.
- Correia, S., P. Guimarães, and T. Zylkin (2019). Verifying the existence of maximum likelihood estimates for generalized linear models. *arXiv preprint arXiv:1903.01633*.



- Correia, S., P. Guimarães, and T. Zylkin (2020). Fast poisson estimation with high-dimensional fixed effects. *The Stata Journal* 20(1), 95–115.
- Davis, A. J. (2005). The power and discretion of the american prosecutor. *Droit et cultures. Revue internationale interdisciplinaire* (49), 55–66.
- Didwania, S. H. (2020a). The immediate consequences of federal pretrial detention. *American Law and Economics Review* 22(1), 24–74.
- Didwania, S. H. (2020b). Mandatory minimum entrenchment and the controlled substances act. *Ohio State Journal of Criminal Law* forthcoming.
- Didwania, S. H. (2021a). Discretion and disparity in federal detention. *Northwestern University Law Review* forthcoming.
- Didwania, S. H. (2021b). Gender favoritism among criminal prosecutors. *Journal of Law Economics* forthcoming.
- Didwania, S. H. (2022). Charging leniency and federal sentences. *University of Wisconsin Legal Studies Research Paper Series Paper No. 1746*.
- Doherty, K. M., D. E. Lewis, and S. Limbocker (2019). Executive control and turnover in the senior executive service. *Journal of Public Administration Research and Theory* 29(2), 159–174.
- Duflo, E., M. Greenstone, R. Pande, and N. Ryan (2018). The value of regulatory discretion: Estimates from environmental inspections in india. *Econometrica* 86(6), 2123–2160.
- Eggen, D. and J. Solomon (2007). Firings had genesis in white house. *The Washington Post* March 13 2007.
- Ely, A. N. (2004). Prosecutorial discretion as an ethical necessity: The ashcroft memorandum’s curtailment of the prosecutor’s duty to seek justice. *Cornell Law Review* 90, 237.
- Fenizia, A. (2021). Managers and productivity in the public sector. *Econometrica* forthcoming.
- Flaherty, M. P. and J. Biskupic (1996). Prosecutors can stack the deck. *The Washington Post* October 7 1996.
- Hoffman, M., L. B. Kahn, and D. Li (2018). Discretion in hiring. *The Quarterly Journal of Economics* 133(2), 765–800.
- Hoffman, M. and S. Tadelis (2021). People management skills, employee attrition, and manager rewards: An empirical analysis. *Journal of Political Economy* 129(1), 243–285.
- Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative science quarterly*, 285–308.
- Lazear, E. P., K. L. Shaw, and C. T. Stanton (2015). The value of bosses. *Journal of Labor Economics* 33(4), 823–861.
- Lewis, J. B., K. Poole, H. Rosenthal, A. Boche, A. Rudkin, and L. Sonnet (2021). Voteview: Congressional roll-call votes database.
- Lichtblau, E. (2003). Ashcroft limiting prosecutors’ use of plea bargains. *The New York Times* September 23 2003.
- Liptak, A. and E. Lichtblau (2003). New plea bargain limits could swamp courts, experts say. *The New York Times* September 24 2003.
- Maestas, N., K. J. Mullen, D. Powell, T. Von Wachter, and J. B. Wenger (2017). Working conditions in the united states: Results of the 2015 american working conditions survey. *RAND Corporation*.

- Marbach, M. and D. Hangartner (2020). Profiling compliers and noncompliers for instrumental-variable analysis. *Political Analysis* 28(3), 435–444.
- Mayer, K. B. (1996). Applying open records policy to wisconsin district attorneys: can charging guidelines promote public awareness. *Wis. L. Rev.*, 295.
- Miller, B. P. and B. W. Curry (2019). *US Attorneys, Political Control, and Career Ambition*. Oxford University Press, USA.
- Nelson, M. J. and I. Ostrander (2016). Keeping appointments: The politics of confirming united states attorneys. *Justice System Journal* 37(3), 211–231.
- Nikolova, M. and F. Cnossen (2020). What makes work meaningful and why economists should care about it. *Labour economics* 65, 101847.
- Okafor, C. O. (2022). Prosecutor politics: The impact of election cycles on criminal sentencing in the era of rising incarceration. *Working Paper*.
- Poole, K. T., J. B. Lewis, J. Lo, and R. Carroll (2020). Scaling roll call votes with w-nominate in r. *Working Paper*.
- Poole, K. T. and H. Rosenthal (1991). Patterns of congressional voting. *American journal of political science*, 228–278.
- Poole, K. T. and H. Rosenthal (2000). *Congress: A political-economic history of roll call voting*. Oxford University Press on Demand.
- Rasul, I. and D. Rogger (2018). Management of bureaucrats and public service delivery: Evidence from the nigerian civil service. *The Economic Journal* 128(608), 413–446.
- Rehavi, M. M. and S. B. Starr (2014). Racial disparity in federal criminal sentences. *Journal of Political Economy* 122(6), 1320–1354.
- Richardson, M. D. (2019). Politicization and expertise: Exit, effort, and investment. *The Journal of Politics* 81(3), 878–891.
- Romero, D. (2021). Los angeles county prosecutors file lawsuit against new district attorney over justice reform efforts. *NBC News January 2 2021*.
- Schmidt, S. (2003). Ashcroft issues tougher prosecutorial guidelines. *The Washington Post September 23 2003*.
- Silveira, B. S. (2017). Bargaining with asymmetric information: An empirical study of plea negotiations. *Econometrica* 85(2), 419–452.
- Spenkuch, J. L., E. Teso, and G. Xu (2021). Ideology and performance in public organizations. *NBER Working Paper 28673*.
- Starr, S. B. and M. M. Rehavi (2013). Mandatory sentencing and racial disparity: Assessing the role of prosecutors and the effects of booker. *Yale Law Journal* 123, 2–80.
- Stern, S. (2004). Do scientists pay to be scientists? *Management science* 50(6), 835–853.
- Stith, K. (2007). The arc of the pendulum: Judges, prosecutors, and the exercise of discretion. *Yale Law Journal* 117, 1420.
- Szucs, F. (2020). Discretion and corruption in public procurement. *Job Market Paper*.
- Toobin, J. (2002). Ashcroft’s ascent: How far will the attorney general go? *The New Yorker April 7, 2002*.
- Tuttle, C. (2019). Racial disparities in federal sentencing: Evidence from drug mandatory minimums. *Available at SSRN 3080463*.
- United States Sentencing Commission (2018). Monitoring of federal criminal sentences, 2015-2016.

- Wright, M. S., S. B. Baughman, and C. T. Robertson (2021). Inside the black box of prosecutor discretion. *UC Davis Law Review*, *Forthcoming*.
- Wright, R. F. and K. L. Levine (2014). The cure for young prosecutors' syndrome. *Ariz. L. Rev.* 56, 1065.
- Yang, C. S. (2014). Have interjudge sentencing disparities increased in an advisory guidelines regime-evidence from booker. *NYUL Rev.* 89, 1268.

Downey & Grunwald

## Appendix

# **A Data and Methods Appendix**

## **A.1 Measuring US Attorney politicization**

We use data from the US Senate to identify all confirmed US Attorneys during the Bush Administration. For the first US Attorney appointed by Bush in each district, we collect data on campaign contributions and professional biography.

### **A.1.1 Campaign contributions**

To gather US Attorneys’ campaign contribution history, we searched manually in DIME Version 2 (Bonica, 2019) for their names in the state in which they held office. We sought to capture all instances where US Attorneys used different names over time (e.g., James McMahon vs. Jim McMahon). We used a variety of informal rules to ensure that we did not match different individuals who lived in the same state (e.g., occupation must be attorney, contribution years cannot overlap, check spouse’s name for joint contributions from spouses, use middle names and initials, use the city of residence we observe in the biographical data). A majority of US Attorneys appear with more than one unique Bonica-generated contributor ID. None appear with more than four.

### **A.1.2 Professional Biographies**

We obtained the biographies through extensive searches conducted in Spring 2019. For each US Attorney, we determine the year they completed their JD. For each calendar year since then, we collect information to classify their employment into fairly detailed categories. We use this to determine whether they were working as prosecutors (and specifically as Assistant US Attorneys), were working in private practice, were working directly for a politician, and the other categories we use in Tables 1 and B3. We also recorded any documented instance of running for elections, even if they lost (and therefore never held the relevant position).

Out of 19,246 possible attorney-years (i.e., 87 US Attorneys with an average of 35 years from JD to death or 2018), we are able to determine the position held during 19,183 (99.7%) of them. For 68 of our US Attorneys, we have no gaps at all. For the remaining 19, the median length of the gap is only three years, and it never exceeds six.

## A.2 Measuring senator ideology NOMINATE scores

### A.2.1 Estimating justice-specific scores

We use data on roll call votes from voteview (Lewis et al., 2021). Each vote through the 113<sup>th</sup> Congress (2013-2014) has one of 109 distinct “issue codes.”<sup>39</sup> We include a vote in our justice-specific measure of ideology if its issue code is “Judiciary,” “Public safety,” “Narcotics,” “Firearms,” “Death penalty,” and “Legal Services Corporation.” We focus on the 87<sup>th</sup> (1961-1962) through the 113<sup>th</sup> Congresses, covering 1,078 justice-specific roll call votes (5% of all roll calls in Congress during the period, but a larger percent in recent years), 75% of which were labeled “judiciary” (including confirmations of US Attorneys and federal judges). The Senate did not vote on any “Death penalty” or “Legal Services Corporation” roll calls during this period. We estimate the NOMINATE scores using the R package from Poole et al. (2020).

### A.2.2 Relationship between all-issue scores and justice-specific scores

The correlation between a state’s most-liberal-senator using the justice-only scores and the all-issue scores is very high, 0.97, so all results are similar with the all-issue NOMINATE, but the all-issue NOMINATE results tend to have slightly smaller point estimates and slightly larger standard errors. Figure A1 shows the relationship between a state’s most-liberal-senator based on both measures. Clearly, the relationship is strong, but there is some meaningful movement.

For example, the most liberal senator across all issues is Paul Wellstone (MN), a well-known leader of the party’s progressive wing. On criminal justice issues, however, Wellstone was best known for promoting victim’s rights and combatting domestic violence. This position often put him in alliance with “tough on crime” Republicans, and so his justice-specific NOMINATE score is notably more moderate than his all-issue score. As another example, Ron Wyden (Oregon), a fairly moderate Democrat, holds fairly centrist views on most issues but is the most influential federal legislator pushing for cannabis legalization and decriminalization (co-sponsoring the Marijuana Justice Act, the Marijuana Opportunity Reinvestment and Expungement Act, and SB 420). It is not surprising, then, that he appears to be one of the most liberal senators on justice issues. This is useful variation that could plausibly affect the confirmation process for US Attorneys.

[Figure A1 about here.]

---

<sup>39</sup>Issue codes can be found at [https://voteview.com/articles/issue\\_codes](https://voteview.com/articles/issue_codes).

## A.3 Measuring prosecutorial behavior and case outcomes

### A.3.1 Charge-level LIONS data

To construct the dataset for our analysis of differential implementation in Section 4, we began with a charge-level dataset and dropped rows based on a series of restrictions. First, we merged in the participant dataset on district, case number, and participant number and then dropped all rows in which participant role was not a criminal defendant. Second, we merged in the instrument dataset and dropped all rows associated with post-disposition action (e.g., probation violations). We further dropped rows we could identify as post-disposition based on disposition and variables. Third, we dropped all rows in which instrument type indicates the case concerned contempt charges. Fourth, we dropped rows associated with forfeiture. Most also have a participant role of "PR," which means they were already dropped above. But we drop additional rows if the disposition variables or charge string indicate they relate to forfeiture. Fifth, we dropped all rows associated with cases not primarily handled by the US Attorney's Office (e.g., by Main Justice). Sixth, we drop all rows associated with courts other than the District or Magistrate Court (e.g., U.S. Court of Appeals). Finally, we drop all rows associated with an "exempted case," for which we lack charge information, likely because the case has been sealed.

### A.3.2 Measuring disposition

We use two variables in the charge-level dataset—called "disposition" and "disp\_reason"—to interpret the disposition and method of disposition for filed charges. The most common values for disposition indicate that a charge resulted in conviction ("GT-Guilty"), non-conviction ("NG-Not Guilty"), dismissal ("DJ-Dismissed with Prejudice"; "DM-Dismissed without Prejudice"; "DP-Dism prejudice (code retired)"), a new filing ("NW-New Filing"); removal ("RE-Removal (Rule 40)"), or transfer ("TR-Transfer from District (Rule 20, 21)"). The second variable, disp\_reason, contains over 100 codes. Some of the most frequent indicate that a charged was disposed by guilty plea (e.g., "PLED-Plea (District Court)"; "PLOD-Plea Other Charge(s) (District Court)"), went to trial (e.g., "BTRD-Bench Trial Verdict (District Court)"; "JTRD-Jury Trial Verdict (District Court)"), or was replaced by a superceding information or indictment (e.g., "SIFD-Superceding Information (from District Court)"; "SIND-Superseding Indictment (from District Court)").

We use the interaction of these two variables to create a new variable indicating whether each charge resulted in a "guilty plea," "dismissal," "trial," or is "not disposed." We define a charge as disposed by guilty plea if the disposition variable equals "GT-Guilty", "NC-Nolo Contendere", or "GD-Adjudged Juvenile Delinquent" and the disp\_reason variable

equals “PLED-Plea (District Court)”, “PLEM-Plea (Magistrate Court)”, “PLE”, “PLEA”, “GWDD-By Government from District Court w/DOJ Authorization”, “GWDM-By Government from Magistrate Court w/ DOJ Authorization”, “DWOD-By Government from District Court no DOJ Authorization”, or “DWOM-By Government from Magistrate Court no DOJ Authorization”.

We define a charge as dismissed if it meets one of the following three requirements. First, we define a charge as dismissed if the disposition variable equals “DM-Dismissed without Prejudice”, “DJ-Dismissed with Prejudice”, “DP-Dism prejudice (code retired)”, “NW-New Filing”, and “SU-Superseded (code retired)”. Second, we define a charge as dismissed where disposition equals “NG-Not Guilty” or “ND-Adjudged not Juvenile Delinquent” and disp\_reason equals “GWDD-By Government from District Court w/ DOJ Authorization”, “GWDM-By Government from Magistrate Court w/DOJ Authorization”, “GWOD-By Government from District Court no DOJ Authorization”, “GWOM-By Government from Magistrate Court no DOJ Authorization”, “PLED-PLEA (District Court),” “PLEM-Plea (Magistrate Court)”, “DEMD-By Defense Motion (District Court)”, or “DEMM-By Defense Motion (Magistrate Court)”. Third, we define a charge as dismissed where disp\_reason equals “PLOD-Plea Other Charge(s) (District Court)” or “PLOM-Plea Other Charge(s) (Magistrate Court)”.

Before applying the sample restrictions (described in A.3.1), roughly 4.5% of all rows in the charge-level file lack a disposition value. After applying sample restrictions, that number falls dramatically. We drop all charges for cases in which we are missing disposition information for one charge or more. We also drop rows in which the disposition variables indicate the charge was removed or transferred to another district, was filed in another case, was declined by the US Attorney’s Office, was created in LIONS in error, or was rejected by the grand jury.

### **A.3.3 Program categories**

LIONS contains 117 program categories describing the primary cause of action in each criminal case. Most categories are quite granular. For instance, the LIONS appendix notes there are 7 different categories within “public corruption,” 5 within “organized crime,” 23 within “white collar crime/fraud,” and 6 within “anti-trust violations.” Of these 117, 109 are used during our 11-year study period.

We drop “fugitive crimes” and “probation revocation” because there is no new offense committed; “assimilated crimes” where the federal government prosecutes a state criminal charge because, for example, the offense was committed on a US military base; “all other violent crimes” (any violent crime other than firearms, bank robbery, or domestic violence)



because it is too broad; and the category defined as “Including in this category everything which does not fall into one of the above categories.” These codes account for roughly 10% of all cases, with the majority being “fugitive crimes.” Dropping “all other violent crimes” only drops 0.5% of cases.

We also drop program categories that appear fewer than 500 times during our 11-year study period. This represents about a third of all program categories but only 0.4% of cases. Our final analytic dataset contains 59 program categories, which give a relatively clear picture of the criminal activity alleged in the case.

### **A.3.4 Thompson Memo**

We identify districts that specialize in corporate criminal prosecution by calculating the share of cases during our sample period that fall within one of the following program categories: tax fraud, securities fraud, commodities fraud, other investment fraud, intellectual property violations, corporate fraud, or antitrust violations. We drop the 10 districts where these categories made up the largest share of cases: NYS (6.8% of cases), CT, NYE, MN, CAC, MA, OHN, NJ, CAN, and ILN (3.8% of cases). TXW had the lowest share across all districts (.1% of all cases).

## **A.4 Severity index**

### **A.4.1 Estimation**

We measure the severity of charges based on average severity in similar cases. To do so, we use data from the United States Sentencing Commission (USSC). The Sentencing Commission collects data on sentences imposed for convictions of felonies and of Class A misdemeanors, which are punishable by over 6 months of prison time (United States Sentencing Commission, 2018). We use data from FY1999-FY2016 on 1.1 million cases (omitting 2012 and 2013 due to inconsistency in data formatting). For every case in the USSC data, we observe the statute—sometimes noisily recorded—for each convicted charge. We also observe the total sentence imposed.

Statutes can be noisily recorded in both LIONS and in the USSC data. Our “statutes” are therefore better thought of as “statute strings.” For example, 21 U.S.C. § 960(a) prohibits several acts of drug trafficking. For violations of that provision, 21 U.S.C. § 960(b) describes the range of punishment available, depending primarily on drug type and quantity but also other case characteristics, such as whether the defendant has a prior conviction for a serious drug or violent offense. § 960(b)(2) generally imposes a minimum sentence of five years in a variety of cases, including, as described in § 960(b)(2)(A), those involving 100 grams or

more of heroin. § 960(b)(1) generally imposes a minimum sentence of ten years in a variety of cases, including, as described in § 960(b)(1)(A), those involving 1,000 grams or more of heroin. While we observe convictions on § 960(b)(1)(A) and § 960(b)(2)(A), we also observe convictions on § 960(a) and § 960(b), which do not distinguish between the 100 gram and 1 kilogram quantities (nor do they distinguish between other sub-categories described within them). Because we observe these “vague” § 960(a) and (b) charges in both LIONS and the Sentencing Commission data, we estimate a severity index value for them and the index value is best thought of as a weighted average of the different potential charges that fall within each. Some of this vagueness may not simply represent noise or measurement error. Instead, it may reflect deliberate and strategic prosecutorial behavior. A former US Attorney for the Eastern District of New York has said that his office intentionally files charges that do not disclose the amount of drugs in cases where prosecutors’ truthful disclosure would trigger mandatory minimum sentences that are excessive (Flaherty and Biskupic, 1996).

Because sentences are recorded at the case level, not the charge level, we need to apportion the resulting sentence across convicted charges in the same case.<sup>40</sup> We do so by regressing the case’s sentence on dummy variables for each unique convicted charge. The coefficients on these “charge dummies” represent the marginal impact on sentence length of one conviction of a given charge, which we call a “charge severity index”.

Specifically, we regress the inverse hyperbolic sine of the length of incarceration from the Sentencing Commission data on 3,000 dummy variables, as well as a series of controls. We always work with the inverse hyperbolic sine of sentences because *i*) sentences with zero incarceration are common (precluding logarithms), and *ii*) many statutory sentencing adjustments are explicitly proportional and are thus better captured by a non-linear specification that estimates proportional effects (like the logarithm or the inverse hyperbolic sine). Our estimating equation is:

$$\sinh^{-1} Y_{d,t} = \delta_t + \sum_c \beta_c D_c + X'_{d,t} \gamma + \varepsilon_{d,t}$$

where the indices  $d$  and  $t$  refer to defendant and year, respectively, and  $D_c$  is a dummy variable denoting that charge  $c$  was one of the statutes resulting in conviction. Note that up to 24 different statutes can be recorded for a single conviction in the Sentencing Commission data.

The controls include defendant race; sex; citizenship; the type of disposition (including guilty plea vs. trial conviction); dummies for whether an adjustment was made for armed career criminal status, career offender status, or acceptance of responsibility; a dummy for

---

<sup>40</sup>Roughly 60% of cases involve two or more different statutes/charges (18%: three or more).

whether the defendant has any criminal history; a quadratic polynomial number of criminal history points; and a cubic polynomial in the number of counts.

The  $\beta_c$  coefficients represent how much longer, proportionally, the defendant’s sentence would be if one more conviction on charge  $c$  were added, holding fixed the other convicted charges and the defendant’s other characteristics. Note that this is not a simple fixed effects specification because multiple charge dummies can be equal to one at once.

In total, we measure the average marginal severity for over 3,000 different charges in LIONS. In our full sample of 530,000 cases from LIONS, there are 1.3 million charges filed, 6,000 of which are unique. The 3,000 charges for which we developed severity indices represent 99% of total filings. Many of the charges for which we are unable to generate an index are minor, such as providing alcohol to a minor within a National Park, and thus are not collected for the USSC data.<sup>41</sup> Our measure of severity is, therefore, far more detailed than those available in other common criminal justice datasets.<sup>42</sup>

As noted, we normalize our index to have standard deviation of one for each program category. To do so, we first winsorize the sentencing index at its 99.9<sup>th</sup> and 0.1<sup>st</sup> percentiles, editing 0.2% of the sentencing index values (3.7 standard deviations below or 3.2 standard deviations above the mean). Next, we determine which program categories are most/least serious by averaging over the maximum severity for each cases ever assigned to that program category (noting that cases can be assigned to multiple program categories). We then assign each case to only one program category: That which is most serious. Finally, we normalize all cases within each program category so that the standard deviation of most serious charge (across cases) is equal to one.

---

<sup>41</sup>Rehavi and Starr (2014) also study prosecutorial discretion over charging. They study the importance of charging behavior for explaining racial inequality in sentences using Federal Justice Statistics Program (FJSP) data. Some advantages of their data are that they can observe the arrest offense and defendants’ race, while we cannot, and their sentencing information is more detailed and reliable than ours. Some disadvantages are that they only observe a subset of charges filed (up to five at the initial filing and up to five at case termination) while we observe all charges ever filed, and their data includes no information on the prosecutor handling the case. Methodologically, they code charge seriousness based on the statutory maximum and minimum sentence (i.e., the bounds written into the federal code defining the longest and shortest sentences that can legally be applied in a case). As they note, these bounds are extremely wide (for instance, the statutory range for 5g of methamphetamine is 5-40 years). Thus, an advantage of our approach is that it better reflects the typical sentences actually awarded for a particular charge. Though an advantage of their approach is that it explicitly identifies charges with mandatory minimums, which are substantively important.

<sup>42</sup>For example, the most commonly used data source on crime, the Uniform Crime Reports (UCR), has only 29 crime types. The most commonly used data on federal prosecutions, from the Federal Justice Statistics Program, has 284. Our highly granularized measure of charge severity allows us to observe changes in the charges filed even among substantively similar charges (i.e., without switching across different program categories).

### A.4.2 Validation

To validate our severity index, we compare it to the sentences imposed in cases where a single charged statute resulted in conviction—roughly 43% of all cases in the Sentencing Commission data (the median case has 2 statutes; only 5% have 3 or more). In these cases, there is no ambiguity about the effects of each charge on the ultimate sentence.

Roughly a thousand charges never appear in a single-statute case. Thus, if we only used single-statute cases, we would be unable to calculate a severity index for roughly one-third of the relevant charges.

For the other two-thirds of charges, however, the average sentence matches up very well with our calculated index value. We first residualize the inverse hyperbolic sine of the prison sentence in single-statute cases from the controls we use in estimating the sentencing index (e.g., criminal history variables) to ensure we do not classify a charge as severe simply because it is disproportionately present in cases where, for example, the defendant has a longer criminal history. We then take the average residual sentence length across all single-statute cases with a conviction for the given statute. Figure A2 presents the relationship between this average residualized-sentence length and our severity index. The correlation between the two variables is 0.78.

[Figure A2 about here.]

Figure A3 presents the same exercise using the inverse hyperbolic sine of the midpoint of the range suggested by the sentencing guidelines. Note that this is based on guideline recommendations, not actual sentences. Nonetheless, we still find a high correlation between our index and the guideline recommendations, particularly when we restrict to charges with a moderate number of single-statute cases for the calculation of average guidelines midpoint (Panel (b)).

[Figure A3 about here.]

## A.5 Prosecutor fixed effects

### A.5.1 Prosecutor assignment

A case can be assigned to more than one prosecutor, either because multiple prosecutors are assigned concurrently or because the prosecutor on a case is reassigned over time. To address concurrent assignments, we only attribute cases to a prosecutor if they are denoted the lead on the case. To address (potentially endogenous) changes in assignments over time, we assign cases to the first lead prosecutor.

### A.5.2 Calculating the fixed effects

To calculate prosecutor fixed effects, we first regress outcomes (e.g., severity of most serious convicted charge) on district-by-program-category and year fixed effects:

$$Y_{c,p,d,t} = \alpha_{p,d} + \delta_t + \varepsilon_{c,p,d,t} \quad (4)$$

where indices  $c$ ,  $p$ ,  $d$ , and  $t$  denote case, program category, district, and year, respectively.

We then regress the residuals on AUSA fixed effects, separately for each year:

$$\hat{\varepsilon}_{c,p,d,t} = \gamma_{a,t} + \nu_{c,p,d,t} \quad (5)$$

Note that our “prosecutor fixed effects” are actually prosecutor-by-year fixed effects (i.e., we estimate  $\gamma_{a,t}$  and not  $\gamma_a$  fixed effects). If we did not allow the values of the prosecutor fixed effects to change over time, then it would be mechanically impossible for their importance to decline.

We calculate the contribution of prosecutor fixed effects to the  $R^2$ , at the district-by-year level, as:

$$P_{d,t} = \left( 1 - \frac{\frac{1}{|\mathcal{C}_{d,t}|} \sum_{c \in \mathcal{C}_{d,t}} \hat{\nu}_{c,p,d,t}^2}{\sigma_Y^2} \right) - \left( 1 - \frac{\frac{1}{|\mathcal{C}_{d,t}|} \sum_{c \in \mathcal{C}_{d,t}} \hat{\varepsilon}_{c,p,d,t}^2}{\sigma_Y^2} \right) \quad (6)$$

where  $\sigma_Y^2$  is the variance of  $Y$  across all cases (i.e., all districts, all program-categories, all years),  $\mathcal{C}_{d,t}$  denotes the set of cases within the district-year, and  $\varepsilon$  and  $\nu$  are as defined above in equations (4) and (5), respectively. The second term represents the  $R^2$  for a regression without prosecutor fixed effects (but with district-by-program-category and year fixed effects) and the first term represents the  $R^2$  for a regression with AUSA fixed effects (as well as the fixed district-by-program-category and year fixed effects).

This is a non-standard calculation of the  $R^2$ . Conventionally, the residual sum of squares ( $\sum_{c \in \mathcal{C}_{d,t}} \hat{\nu}_{c,p,d,t}^2$  and  $\sum_{c \in \mathcal{C}_{d,t}} \hat{\varepsilon}_{c,p,d,t}^2$ ) and the total sum of squares would be taken from the same sample. Instead, we take the residual sum of squares separately for each district-year, and the total sum of squares for the full sample.<sup>43</sup> This is because we wish to hold  $\sigma_Y^2$  fixed for the full sample so that the denominator of the  $R^2$  calculation does not change when the variation across cases falls in response to the memo.

---

<sup>43</sup>Obviously, there are more cases in total than there are for any district-year. Thus, we always work with the mean of squares, not the sum of squares.

## A.6 AUSA Heterogeneity

### A.6.1 Measuring Assistant US Attorney performance

We consulted several former AUSAs about how to measure attorney skill, ability, or experience with important or complex cases. We proposed four potential measures, denoted in column 3 of Table A1: experience as an AUSA, number of cases handled, the share of cases that were criminal rather than civil, and the share of cases that resulted in appeal. The former AUSAs rejected the third and fourth measures, which we do not use, and proposed nine more, which are also listed in Table A1. We check whether each variable is actually used in each district (e.g., not all districts record 5K1.1 credits). Our heterogeneity analysis only uses variables in districts that actually used the variable.

[Table A1 about here.]

### A.6.2 Deriving our “complier variant” estimator

In Section 5.4 we propose an approach to determine whether the AUSAs induced to leave by the memo were different than those leaving under normal circumstances. In the canonical case of a randomly assigned binary instrument, our approach is equivalent to regressing the characteristics of program participants on whether they were assigned the instrument. Here, we show that the  $\beta$  coefficient from that regression (i.e., the coefficient on the instrument) is proportional to the difference in average characteristics between the compliers and the ‘always takers.

Let  $\bar{x}_A$ ,  $\bar{x}_C$ , and  $\bar{x}_N$  be the average level of characteristic  $x$  among the population of “always-takers” (who depart regardless of the memo), “compliers” (who depart because of the memo), and “never-takers” (who do not depart), respectively, where our terminology follows standard treatments of complier analysis when random assignment is used as an instrument for program participation (e.g., Angrist and Pischke (2008) or Marbach and Hangartner (2020)), as well as the standard monotonicity assumption (in our case, no one who would have otherwise departed is induced to stay because of the memo). Let  $p_A$ ,  $p_C$ , and  $p_N$  be the share of the population who are always-takers, compliers, and never-takers, respectively. Note that  $\bar{x}_J$  and  $p_J$  are not observed. Our interest is in  $\Delta \equiv \bar{x}_C - \bar{x}_A$ .

Let  $Z$  be a binary, random instrument for departures. Then one could focus on the sample of departures, and regress  $x$  on  $Z$ , thereby estimating:

$$x = \beta_0 + \beta_1 Z + \varepsilon$$

Because OLS estimates the conditional mean, this provides an estimate for  $\bar{x}$  when

the instrument is zero ( $\hat{\beta}_0$ ). By definition, this sample includes only the always-takers, and so, in expectation,  $\beta_0 = \bar{x}_A$ .

This regression also provides an estimate for  $\bar{x}$  when  $Z = 1$ , which is where  $\bar{x}$  is a weighted average of  $\bar{x}_A$  and  $\bar{x}_C$  weighted by their relative frequency:

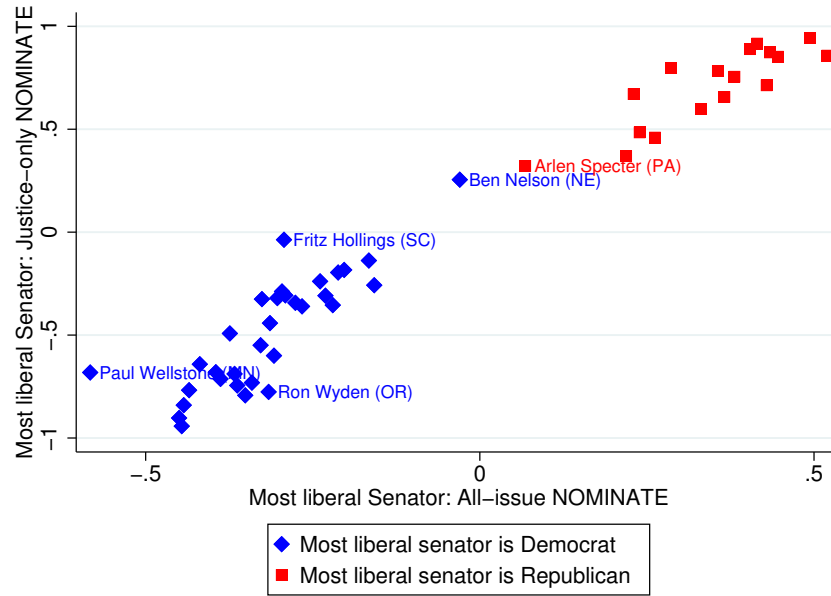
$$\beta_0 + \beta_1 = \frac{p_A}{p_A + p_C} \bar{x}_A + \frac{p_C}{p_A + p_C} \bar{x}_C$$

Note that  $\beta_0$  can be written as  $\frac{p_A + p_C}{p_A + p_C} \bar{x}_A$ , so  $\beta_1$  can be written as:

$$\begin{aligned} \beta_1 &= \frac{p_A}{p_A + p_C} \bar{x}_A + \frac{p_C}{p_A + p_C} \bar{x}_C - \frac{p_A + p_C}{p_A + p_C} \bar{x}_A \\ &= \frac{p_A \bar{x}_A}{p_A + p_C} + \frac{p_C \bar{x}_C}{p_A + p_C} - \frac{p_A \bar{x}_A + p_C \bar{x}_A}{p_A + p_C} \\ &= \frac{p_A \bar{x}_A}{p_A + p_C} + \frac{p_C \bar{x}_C}{p_A + p_C} - \frac{p_A \bar{x}_A}{p_A + p_C} - \frac{p_C \bar{x}_A}{p_A + p_C} \\ &= \frac{p_C \bar{x}_C}{p_A + p_C} - \frac{p_C \bar{x}_A}{p_A + p_C} \\ &= \frac{p_C}{p_A + p_C} (\bar{x}_C - \bar{x}_A) = \frac{p_C}{p_A + p_C} \Delta \end{aligned}$$

As noted in the text, this means that our regression estimate is proportional to the difference between compliers and always takers, where “the (non-negative) multiplier is the share of program participants who are compliers.”

Figure A1: Most liberal senator: All-issue NOMINATE vs. Justice-specific NOMINATE



*Notes:* Figure plots an all-issue NOMINATE score for the most-liberal Senator in each state based on roll call votes for all bills against a justice-specific NOMINATE score based on the subset of roll calls for justice-related bills.

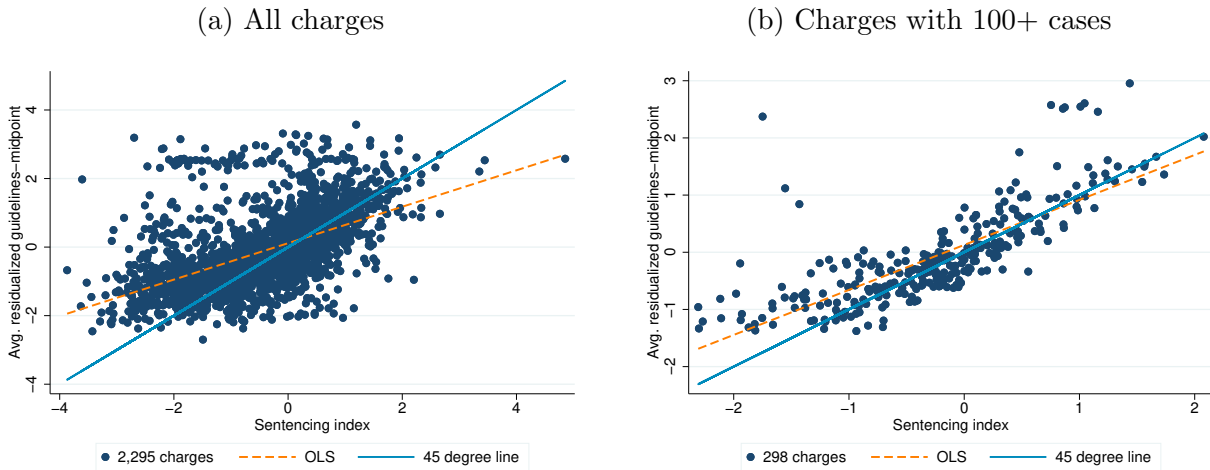


Figure A2: Sentencing index and sentences in single-statute cases



*Notes:* Sample is based on charges in cases in the Sentencing Commission dataset that involve a single criminal statute. The inverse hyperbolic sine of sentence length is residualized on the controls mentioned in Section A.4.1 and then averaged across all cases with the same statute.

Figure A3: Sentencing index and guidelines in single-statute cases



*Notes:* Sample is based on charges in cases in the Sentencing Commission dataset that involve a single criminal statute. For each case, the midpoint of the guidelines range is residualized on the controls mentioned in Section A.4.1 and then averaged across all cases with the same statute. Panel (b) restricts to charges for which we observe at least 100 single-statute cases.



Table A1: Measures suggested for Assistant US Attorney heterogeneity

Measure	Observable	Ours	Comments
Experience as an AUSA	Yes, with censoring when our data starts in 1998	Yes	AUSA-1 pointed out that this measure is imperfect because many AUSAs have substantial litigation experience prior to becoming AUSAs
Number of cases	Yes	Yes	
Criminal share of cases	Yes	Yes	Multiple AUSAs felt this was a bad measure because civil cases are often very important. <b>We do not use this measure.</b>
Cases resulting in appeal	Yes	Yes	Multiple AUSAs felt this was a bad measure because appeals are common even when there's little chance of overturning the original conviction and little reason to believe the prosecutor made a mistake. <b>We do not use this measure.</b>
Cases that went to trial	Yes		
Multi-defendant cases	Yes		Note: We define cases based on internal LIONS case numbers, not court docket numbers.
Cases with Main Justice attention	Yes		Either litigation responsibility is shared between DOJ and the US Attorneys Office, or DOJ delegated directly to the USAO or reports monitoring the USAO.
Number of agents working the case	Yes		
Cases with informants	No		
Cases with 5K1.1 credits	Yes		This is recorded during the pre-memo period, but they stopped reporting it during later years for safety reasons. AUSA-2 was concerned that this variable measures which defendants testified in other cases (earning a 5K1.1 credit), but not which case they testified in (which would be the better measure of skill or complexity).
Cases with advanced investigation techniques (e.g., Title III wiretaps)			
Complexity weight	Yes		There is a LIONS variable where prosecutors themselves can assign a complexity weight to the case, but this variable is very rarely used and we determined it was not reliable. <b>We do not use this measure.</b>
Public corruption or organized crime cases	Yes		We identify these using the program categories.

## **B Additional Results and Robustness**

### **B.1 Ideology and US Attorney Politicization**

[Table B1 about here.]

[Table B2 about here.]

[Table B3 about here.]

[Figure B1 about here.]

### **B.2 Differential implementation of the memo**

[Table B4 about here.]

[Figure B2 about here.]

[Figure B3 about here.]

[Figure B4 about here.]

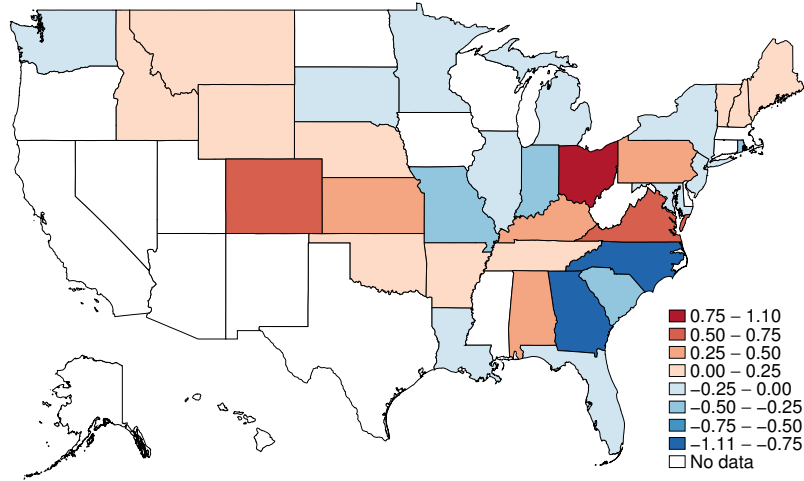
[Figure B5 about here.]

[Figure B6 about here.]

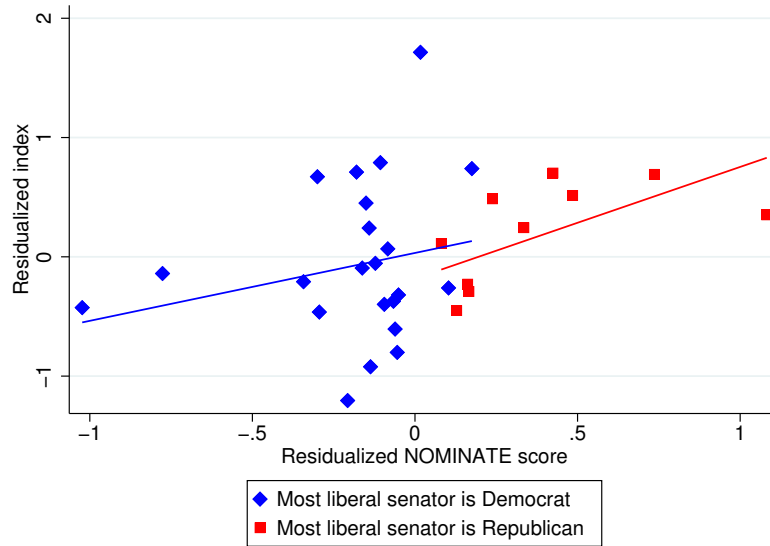
[Table B5 about here.]

Figure B1: Senator Ideology and US Attorney Politicization

(a) Variation in NOMINATE scores idiosyncratic to 2001-2002

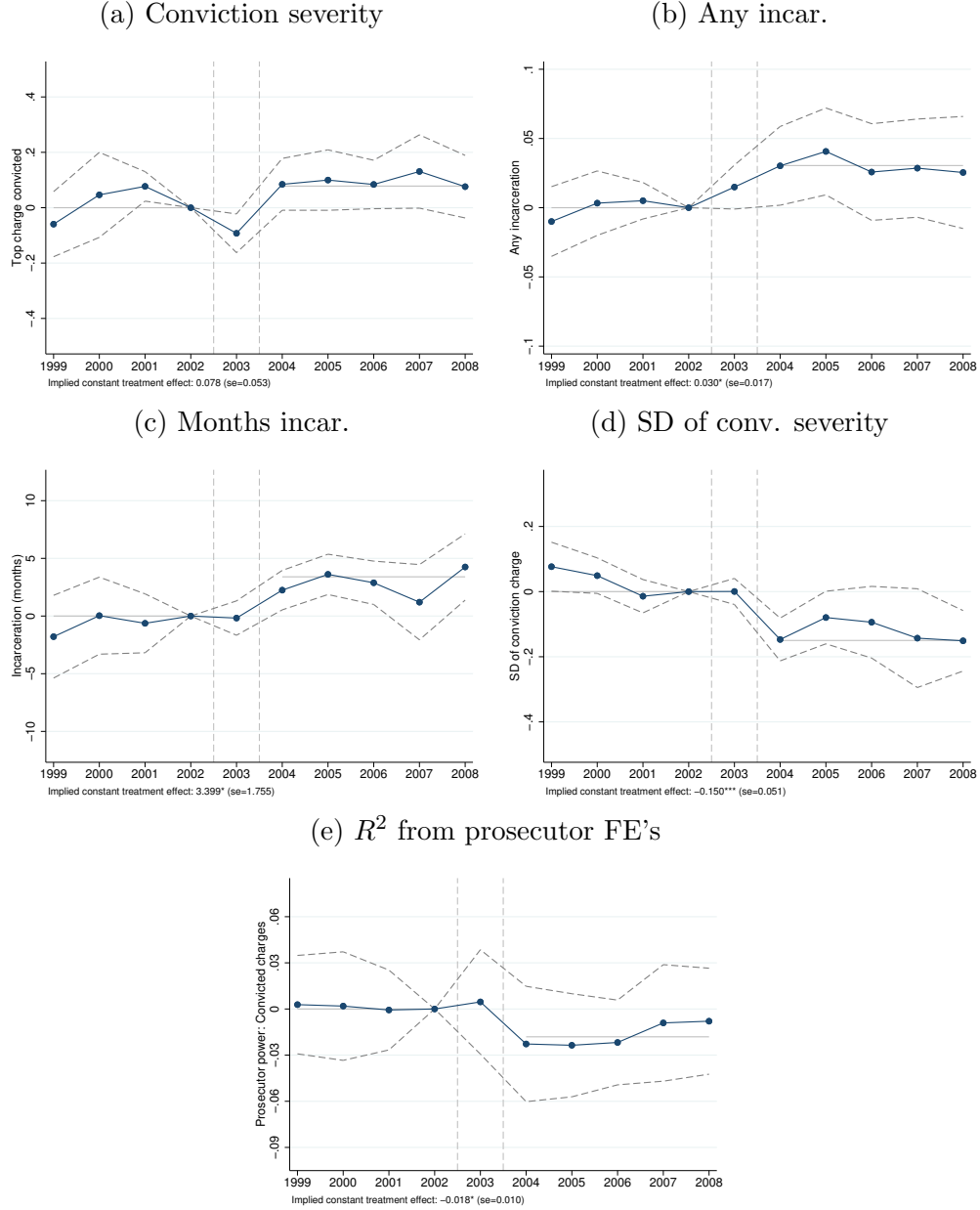


(b) Ideology and Politicization



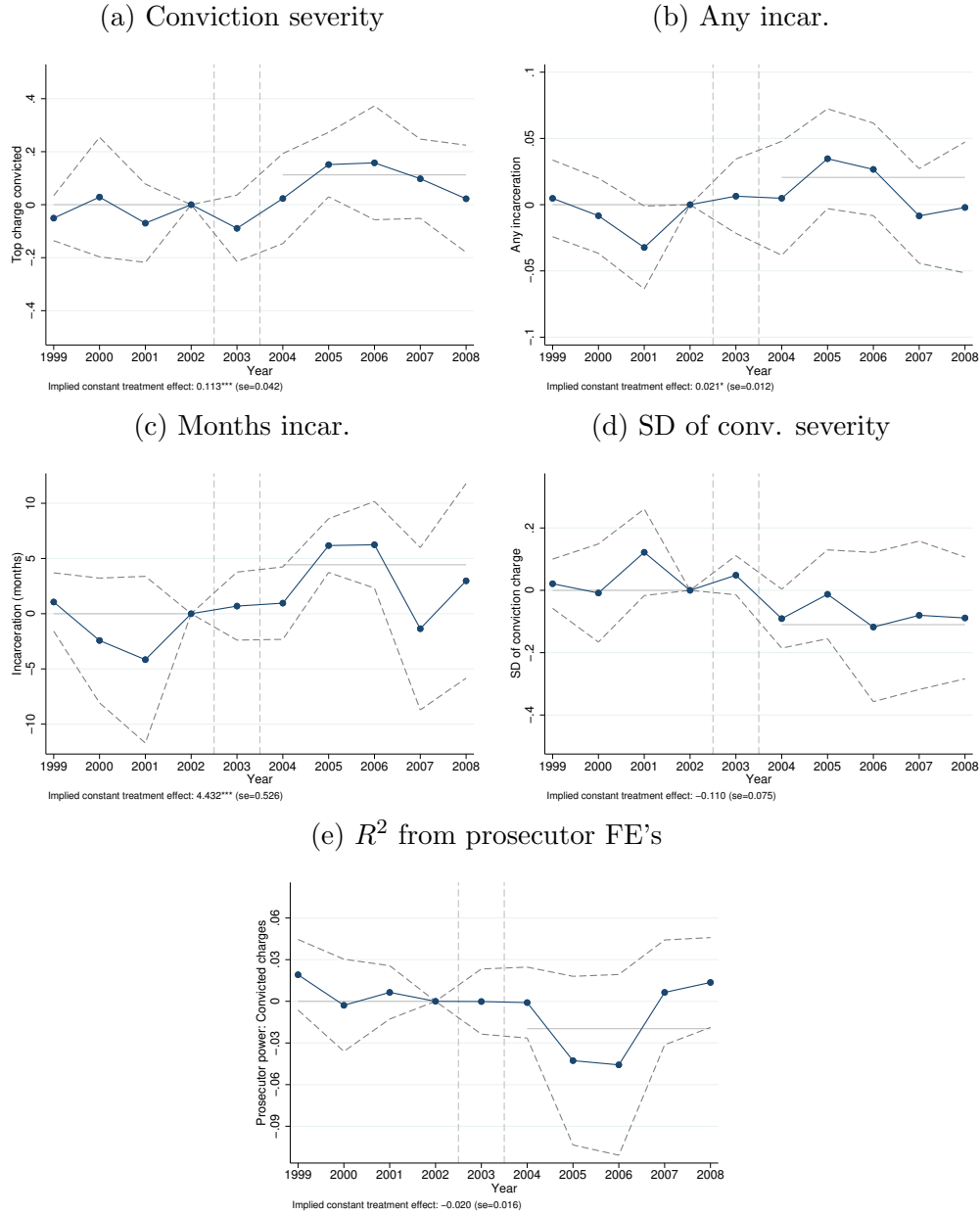
*Notes:* Panel (a) displays the variation in states' most-liberal senators' justice-specific NOMINATE scores which is idiosyncratic to 2001-2002 (the 107<sup>th</sup> Congress). To do so, we first regress these NOMINATE scores from the 107<sup>th</sup> Congress on the corresponding scores from the 104<sup>th</sup> (1995-1996) and 110<sup>th</sup> (2007-2008) Congresses. We then plot the residuals, only for states that experienced a change in the identity of their most-liberal senator. Panel (b) shows the correlation between these residualized NOMINATE scores and our index of US Attorney politicization (also residualized).

Figure B2: Differential implementation: Basic specification



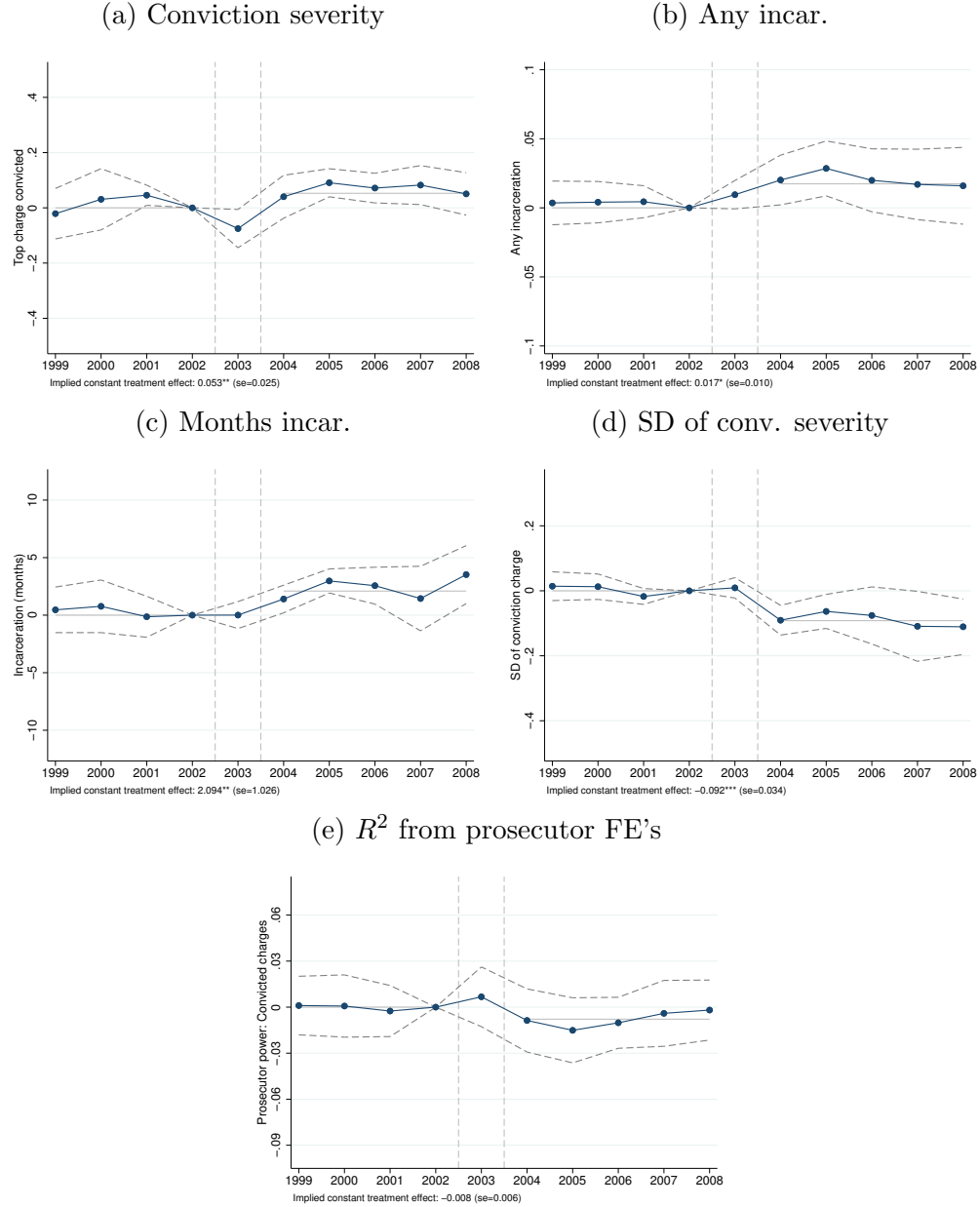
*Notes:* Unit of observation is district-by-year-by-program-category. Observations are weighted by the number of cases. Regressions include year, district, and program category fixed effects. Standard errors are clustered at the state level and the program category level. Plotted coefficients are based on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator.

Figure B3: Differential implementation: Idiosyncratic variation only



*Notes:* Unit of observation is district-by-year-by-program-category. Observations are weighted by the number of cases. Regressions include year, district, and program category fixed effects, as well as ideology from previous and subsequent senate terms (1995-1996 and 2007-2008) interacted with year effects. Standard errors are clustered at the state level and the program category level. Plotted coefficients are based on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator.

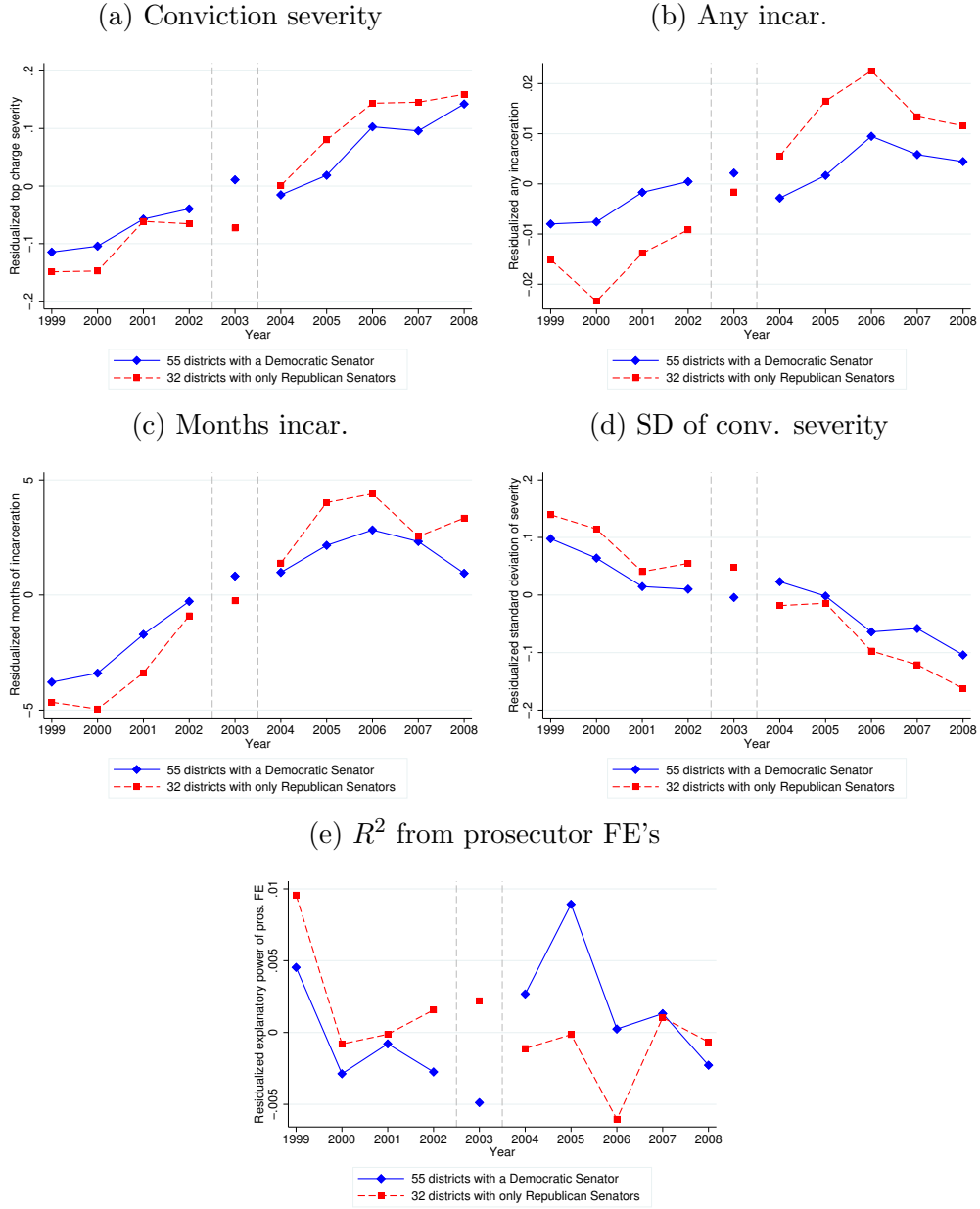
Figure B4: Differential implementation: Between-party only



*Notes:* Unit of observation is district-by-year-by-program-category. Observations are weighted by the number of cases. Regressions include year, district, and program category fixed effects. Standard errors are clustered at the state level and the program category level. Plotted coefficients are based on a dummy for whether the state had two Republican senators, interacted with year effects.

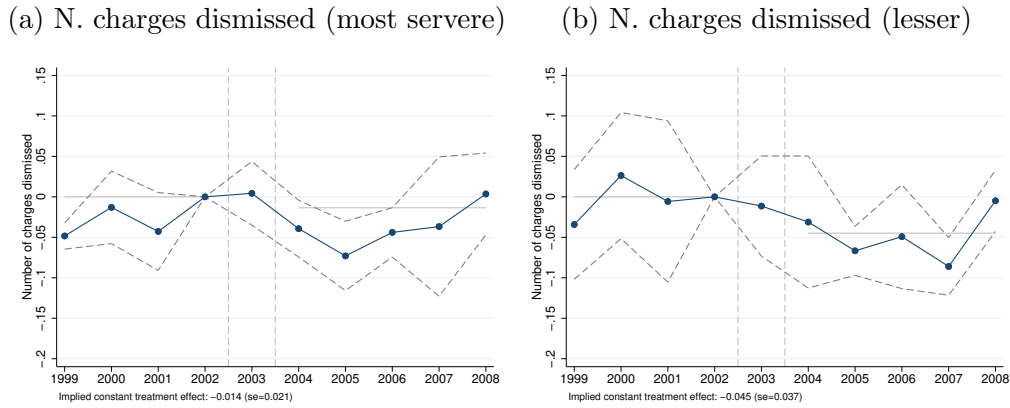


Figure B5: Changes in residualized outcomes by partisan control



*Notes:* Figure constructed by residualizing district-by-program-category fixed effects from all outcomes, then aggregating depending on composition of districts' senate delegation.

Figure B6: Changes in dismissed charges relative to top convicted charge



*Notes:* Unit of observation is district-by-year-by-program-category. Observations are weighted by the number of cases. Regressions include year and district-by-program-category fixed effects. Standard errors are clustered at the state level and the program category level. Plotted coefficients are based on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator.

Table B1: Robustness: Senate ideology and US Attorney politicization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
					Pre-USA work histories only		
DV:	Index for political attachment				Work for politician	Elected office	Either
Justice-only NOMINATE	0.680** (0.253)		0.943 (0.753)	1.279 (0.873)	0.093 (0.080)	0.016 (0.113)	0.113 (0.131)
All-issue NOMINATE		0.673** (0.272)					
2 Republicans			-0.197 (0.552)	0.865 (1.014)			
NOMINATE $\times$ 2 Rep.				-1.466 (1.402)			
Constant	-0.433** (0.188)	-0.426** (0.198)	-0.484** (0.214)	-0.562** (0.235)	0.049 (0.042)	0.096 (0.062)	0.131* (0.072)
$R^2$	0.052	0.042	0.053	0.059	0.011	0.000	0.009
N	87	87	87	87	87	87	87

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . Unit of observation is US Attorney. Standard errors clustered at the state level are reported in parentheses. NOMINATE variables have been normalized to have a minimum of zero (most liberal) and maximum one (most conservative). Index is based on the first five variables in Table 1 and normalized to a mean of zero and a standard deviation of 1.

Table B2: Understanding variation in Bush 2000 vote share

DV: Bush vote share ('00)	(1)	(2)	(3)	(4)	(5)
Most liberal (107th)	0.190*** (0.028)	0.133*** (0.037)			0.092*** (0.025)
Most conserv. Senator		0.040** (0.020)			
2 Democrats			-0.116*** (0.026)		
1 Dem./1 Rep.			-0.062** (0.029)		
2 Dem. or 1D/1R				-0.094*** (0.023)	
Most liberal (104th)					0.016 (0.044)
Most liberal (110th)					0.101*** (0.035)
$R^2$	0.468	0.498	0.299	0.244	0.527
N	49	49	49	49	49

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . Unit of observation is the state. Robust standard errors in parentheses. Dependent variable is the share of the two-party vote going to George W. Bush in the 2000 General Election.

Table B3: Senators and US Attorney experience

	(1)	(2)	(3)	(4)	(5)	(6)
	Years since JD	Internal promo.	Share of career spent in Prosec.	DOJ	Private	Index
Most liberal senator	1.928 (1.802)	0.181 (0.119)	0.095 (0.117)	0.020 (0.021)	-0.060 (0.131)	0.540* (0.303)
Constant	16.985*** (1.020)	0.157** (0.077)	0.468*** (0.076)	0.007 (0.011)	0.355*** (0.086)	-0.253 (0.194)
$R^2$	0.012	0.020	0.008	0.011	0.004	0.033
N	84	87	87	87	87	87

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . Unit of observation is US Attorney. Standard errors clustered at the state level are reported in parentheses. NOMINATE variables have been normalized to have a minimum of zero (most liberal) and maximum one (most conservative). Index is based on the first five variables and normalized to a mean of zero and a standard deviation of 1.

Table B4: Differential implementation: IV estimates of US Attorney politicization

	(1)	(2)	(3)	(4)	(5)
	Conviction severity	Any incar.	Months incar.	SD of conv. severity	$R^2$ from pros. FE's
<b>Panel A:</b> District-by-program-category fixed effects (preferred specification)					
$Post_t \times PolIndex_d$	0.130** (0.055)	0.048*** (0.015)	5.042** (2.207)	-0.179** (0.072)	-0.017 (0.011)
<b>Panel B:</b> Basic specification					
$Post_t \times PolIndex_d$	0.090 (0.063)	0.035* (0.019)	3.945* (2.164)	-0.174** (0.077)	-0.019 (0.012)
<b>Panel C:</b> Only senate ideology variation idiosyncratic to 2001-2002					
$Post_t \times PolIndex_d$	0.118* (0.061)	0.024* (0.014)	4.496*** (1.579)	-0.125 (0.085)	-0.019 (0.015)
<b>Panel D:</b> Only senate ideology variation due to party					
$Post_t \times PolIndex_d$	0.090* (0.051)	0.030* (0.017)	3.566* (2.061)	-0.156** (0.074)	-0.013 (0.011)

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ . All specifications instrument for our index of politicization, based on campaign contributions and political employment (see Table 1). Columns 1-4: Unit of observation is district-by-year-by-program-category. Standard errors are clustered at the state level and the program category level. Column 5: Unit of observation is district-by-year. Standard errors are clustered at the state level. All columns: Observations are weighted by the number of cases. Plotted coefficients are based on justice-specific NOMINATE scores interacted with a “post-memo” dummy (2004-2008). NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator. Panel B (basic specification) includes district, program category, and year fixed effects. Panel A modifies panel B by using district-by-program-category fixed effects. Panel C modifies panel B by controlling for senator ideology from other terms (1995-1996 and 2007-2008) interacted with year effects to isolate variation idiosyncratic to the Senators initially in office when Bush took office. Panel D modifies panel B by using a dummy for whether the state had two Republican senators, rather than using NOMINATE scores.

## C Alternative explanations

### C.1 Increase in Immigration Cases

During the study period, there was a dramatic rise in immigration cases that is correlated with senator ideology because Southern border states often have more conservative senators. Immigration cases are politically contentious and are viewed by many as a distraction from

Table B5: Effects on filing, conviction, dismissal outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	Number of charges			Most serious charge		
	filed	convicted	dismissed	filed	convicted	dismissed
$Post_t \times NOM_s$	-0.086*** (0.003)	-0.017 (0.018)	-0.058*** (0.000)	-0.014** (0.006)	0.114*** (0.042)	0.200 (0.144)
$R^2$	0.766	0.767	0.645	0.984	0.951	0.857

\*  $p < .10$ , \*\*  $p < .05$ , \*\*\*  $p < .01$ .

the federal government’s role in addressing crime.<sup>44</sup> It is possible that frustration with this trend drove prosecutors to leave the US Attorney’s Office.

The rise in immigration cases, however, cannot explain our results. For the analysis of differential implementation, our preferred specifications exclude all immigration cases (though the results are generally stronger when we include them). The immigration docket cannot explain our results on departures either. The entire correlation between the volume of immigration cases and senator ideology is driven by three states: Arizona, Texas, and New Mexico. While many states had high *levels* of immigration cases, only Arizona, New Mexico, and Texas experienced substantial *growth*, and it is these *changes* that pose a threat to our difference-in-difference analysis.

[Figure C1 about here.]

Because immigration cases are very common, these states have substantial influence in case-level regressions (especially Arizona and Texas, which account for 54% of all cases in 2008), but little influence on the district-level departure regressions (they account for only 11% of assistant US attorneys). Thus, ex ante, it is unlikely these states can drive our results. Figure C2 replicates our departures analysis dropping these three states. The results are unchanged. Likewise, controlling for lagged immigration cases has no effect on our estimates.

[Figure C2 about here.]

<sup>44</sup>The immigration docket was less polarized then than it is today. There was a much smaller difference between the immigration platforms of George W. Bush and John McCain vs. John Kerry and Barack Obama than between the post-2016 Democratic and Republican parties.

## C.2 United States v. Booker

In January 2005, the Supreme Court ruled in *United States v. Booker* that the U.S. Sentencing Guidelines are not binding on federal sentencing decisions and are instead merely advisory. *Booker* caused judges to more frequently impose sentences below the recommendations of the Guidelines (Yang, 2014), which shifted some control over punishment from prosecutors to judges.

Because we always include year fixed effects, *Booker* is only a threat to identification if it had differential effects in states where the more liberal senator was more conservative than in other states. For several reasons, the Ashcroft memo better explains our results than *Booker*. First, regarding differential implementation, many of the effects we estimate in Section 4 begin in 2004, the year before *Booker*.

Second, our idiosyncratic-variation specifications reassure that our results are not simply driven by unobserved correlations between the ideology of US Attorneys and judges in the same federal districts. The identifying variation in these specifications arises from “purple” swing states (e.g., Ohio, Virginia, Colorado, North Carolina, Georgia) and depends on whether those states happened to have more liberal senators around 2001—when all the US Attorneys we study were nominated. Compared to US Attorneys, turnover for federal judges is much slower and comparatively few federal judges were nominated in that same period. Thus, the idiosyncratic variation has little expected relationship with judicial ideology.

Third, our results focus on the specific charges convicted, not on the sentence imposed by the judge. While prosecutors’ charging decisions may anticipate judicial changes in sentencing (Silveira, 2017), any influence of *Booker* on charging or plea bargaining would be only indirect. Moreover, the best evidence we have suggests *Booker* had little lasting impact on charging and plea bargaining. One common prediction, for example, was that *Booker* could encourage prosecutors to file and secure conviction more frequently on mandatory minimums to undercut the expanded judicial discretion bestowed by *Booker*. But, as Starr and Rehavi (2013) carefully document, aside from a few blips in the month or two after the opinion issued, there’s little evidence the opinion had any long-term effects on the rate at which prosecutors filed and obtained convictions on mandatory minimum charges. As the authors explain, “*Booker* is only what federal judges make of it, and, so far, that appears not to have been much (Starr and Rehavi, 2013, p. 71).”

Fourth, any “indirect effects” of *Booker* on charging or plea bargaining would more likely produce the opposite of our results. In practice, *Booker* substantially increased the frequency with which judges sentenced below the guidelines but not above (Yang, 2014). The surge in departures after *Booker* was, therefore, likely concentrated in districts with more liberal judges, who were seeking to undo the perceived harshness of the Federal Sentencing

Guidelines. Any indirect effects of *Booker* on charging or plea bargaining likely would have concentrated in those same liberal districts, driving up the severity of their convicted charges more than in conservative districts—the opposite of our results.

Although LIONS lacks direct measures of guideline departures, we can test whether downward departures differentially increased in conservative districts after *Booker* by fitting our main regression specification on sentences and controlling for the most serious charge resulting in conviction. If downward departures differentially increased in conservative-senator districts, then we would expect that, after controlling for the most serious convicted charge, the average sentence imposed would differentially fall in those districts. Figure C3 shows that, instead, the average sentence continues to differentially rise in conservative-senator districts (likely due to the Ashcroft memo’s guidance on prosecutors behavior during sentencing). For that reason, the results suggest downward departures did not differentially increase in conservative districts after *Booker*.

[Figure C3 about here.]

Like our results for differential implementation, *Booker* probably does not explain our results for differential departures. If *Booker* increased the frequency with which liberal judges sentenced below the guidelines, it is plausible that frustrated prosecutors might have left in response. However, this likely would have differentially increased departures in liberal-senator districts (where the liberal judges are concentrated), the opposite of what we find.

### C.3 US Attorney Firing Controversy

In December 2006, the Bush Administration fired seven United States Attorneys. When this story was publicized in the middle of 2007, it became one of the most publicized scandals of the second Bush Administration. As a result, the top three officials at the Justice Department and the Director of the Executive Office for US Attorneys all resigned. Some Assistant US Attorneys may have left, too, in protest.

While the scandal may have driven some departures, we do not believe it can explain our results. First, while numerous reasons were given for firing the US Attorneys (and the reasons for each differed), neither critics nor the Bush Administration attributed the firings to implementation of the memo. Second, our estimated effects on departures begin both before the firings occurred and before they were publicized.<sup>45</sup> Third, the correlation between the ideology of the most liberal senator and whether the state had any fired US Attorneys

---

<sup>45</sup>Several of the departing US Attorneys were instructed to make public statements that they were resigning to pursue other opportunities. The public and the news media were not aware that these were firings until mid-2007. It is unclear whether Assistant US Attorneys would have known before then.



is low (-.14), so these pressures were likely orthogonal to our identifying variation. Finally, when we drop the seven districts with a fired USA, we obtain very similar departure results (see Figure C4).

[Figure C4 about here.]

## C.4 Thompson Memo

In 2003, the Department of Justice issued the Thompson Memo, which signalled a new federal enforcement priority in criminal corporate cases. In anticipation of an increase in corporate prosecutions, private sector demand may have risen for lawyers with DOJ experience, especially current or former Assistant US Attorneys and even those who had no white collar experience. The Thompson Memo, therefore, may have encouraged departures at the same time and in the same districts as the Ashcroft memo.

We find little evidence to support this hypothesis. First, when we drop 10 federal districts with the most cases involving charges that corporations may worry about (e.g., tax fraud, securities fraud, antitrust violations), the magnitudes of our estimates remain nearly identical; though, they become statistically insignificant in the idiosyncratic variation specification (see Figure C5).

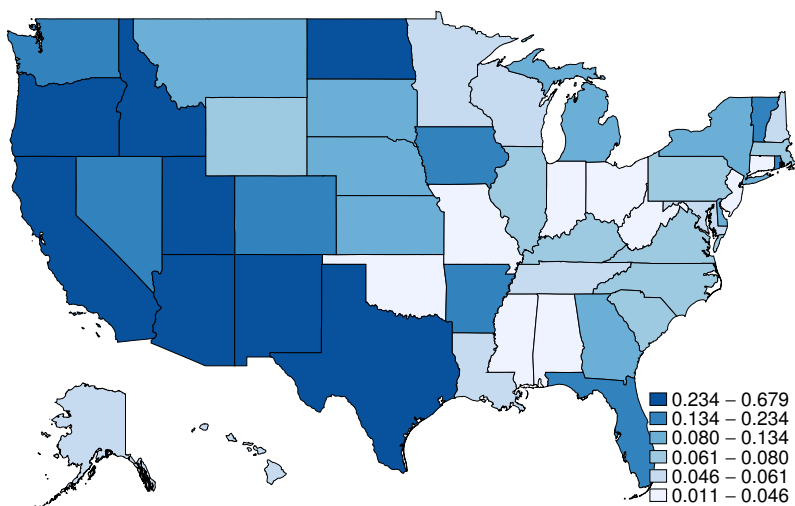
[Figure C5 about here.]

Second, we examine whether our identifying variation is correlated with private sector options for attorneys. Specifically, we use American Community Survey data to construct state-year-level measures of the share of private sector workers who are attorneys and the log of average earnings for private sector attorneys. We then estimate our usual regressions at the state-year level with those two measures as dependent variables (Figure C6). We find no change in private sector attorneys' employment or earnings. It is therefore unlikely that our departure results are driven by a surge in attractive outside options for Assistant US Attorneys in conservative-senator districts where our departure effects are concentrated.

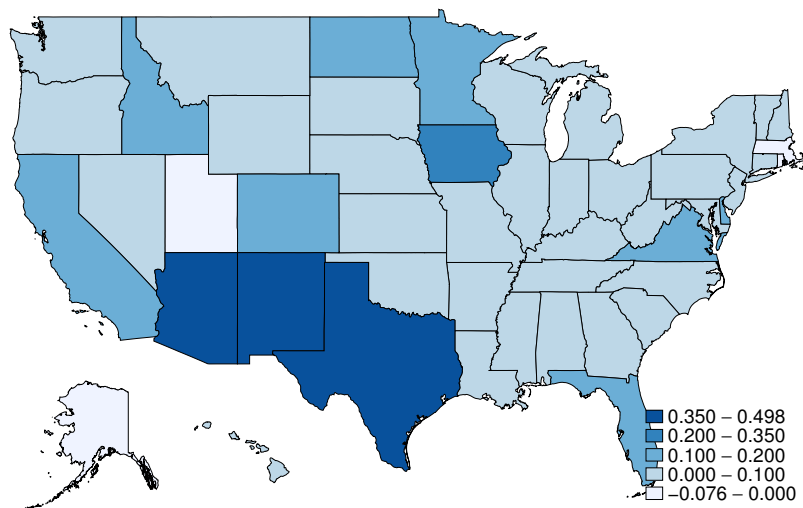
[Figure C6 about here.]

Figure C1: Immigration caseloads

(a) Immigration cases as share of all cases (average of 2001-2008 annual values)



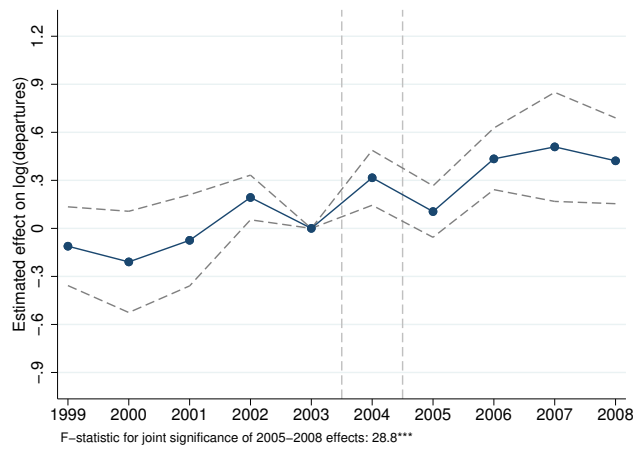
(b) Change in immigration case share (2001-2008)



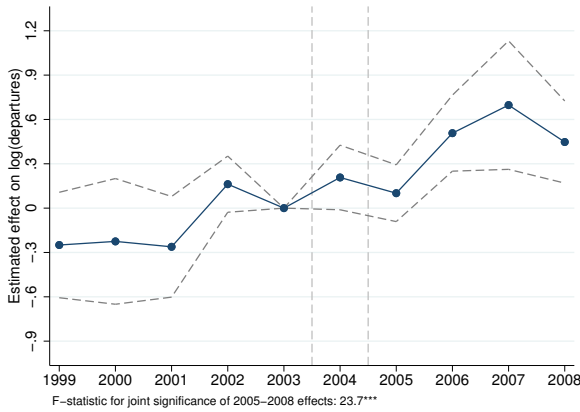
*Notes:* Panel (a) displays the average of annual values for the share of cases in the state involving immigration charges. Panel (b) displays the change, from 2001-2018, in the share of cases that involve immigration charges.

Figure C2: Robustness of departure results to immigration cases

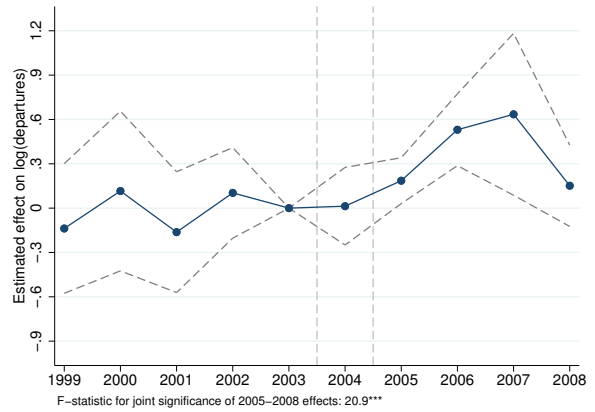
(a) Drop AZ, NM, TX  
Full variation



(b) Drop AZ, NM, TX  
Party-only variation

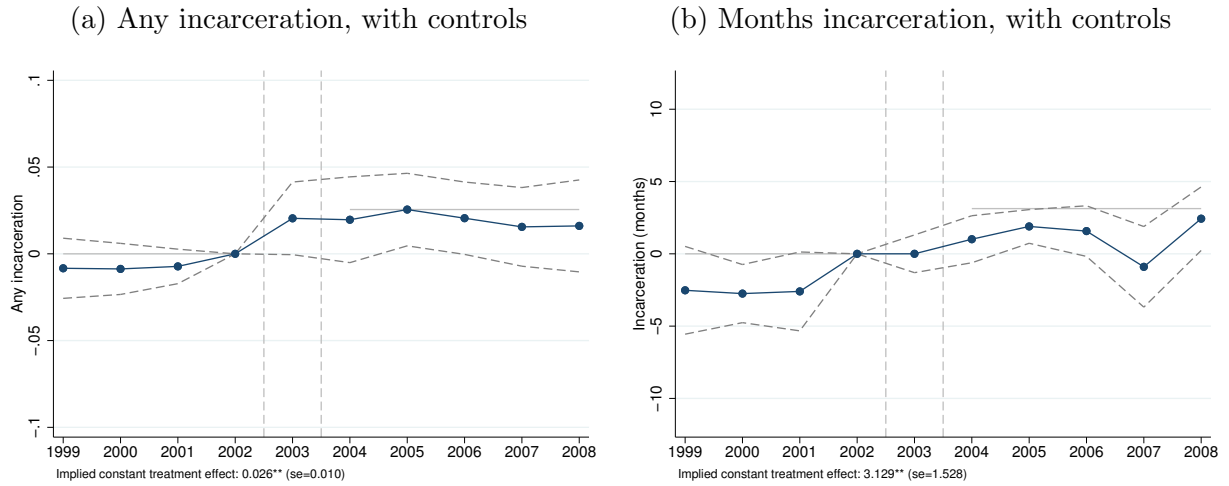


(c) Drop AZ, NM, TX  
Idiosyncratic variation



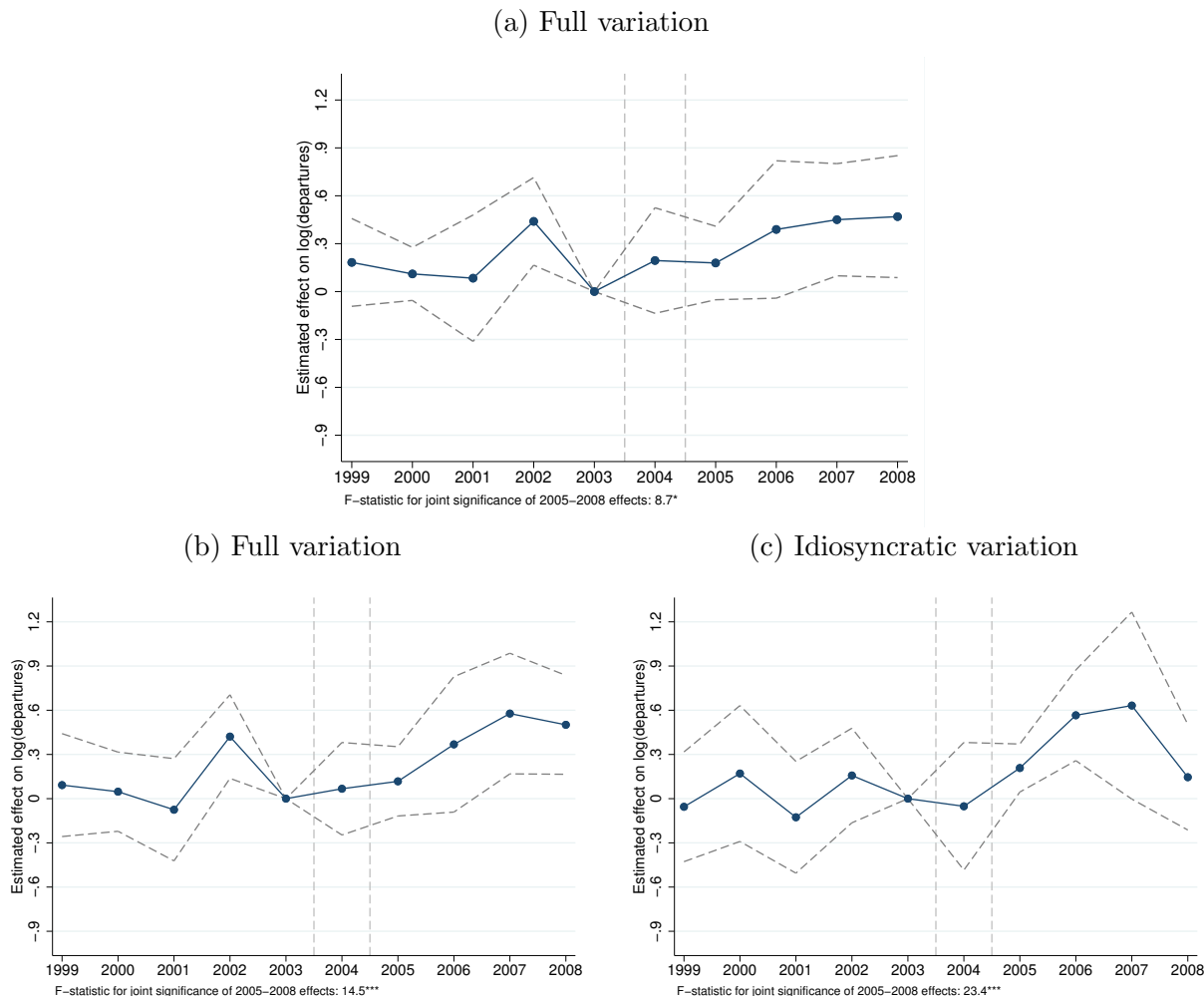
*Notes:* Unit of observation is district-year. Observations are weighted by the average number of Assistant US Attorneys in the district (averaged across years). Estimates come from a Poisson pseudolikelihood regression in which the dependent variable is the number of attorneys departing each district in a given year. All regressions include district and year fixed effects. Standard errors are clustered at the state level. Plotted coefficients are on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator and one in the state with the most-conservative-most-liberal senator. “Full variation” estimates use raw NOMINATE scores. “Party-only” estimates use a dummy for whether the state had two Republican Senators. “Idiosyncratic variation” estimates also control for 1995 and 2007 NOMINATE scores interacted with year effects (see Figure B1).

Figure C3: Effects on sentences with and without conviction charge controls



*Notes:* Unit of observation is district-by-year-by-program-category. Observations are weighted by the number of cases. Regressions include year and district-by-program-category fixed effects. Unlike Figure 3 in the text, regressions also control for the severity of the most serious charge resulting in conviction. Standard errors are clustered at the state level and the program category level. Plotted coefficients are based on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator, and one in the state with the most-conservative-most-liberal senator.

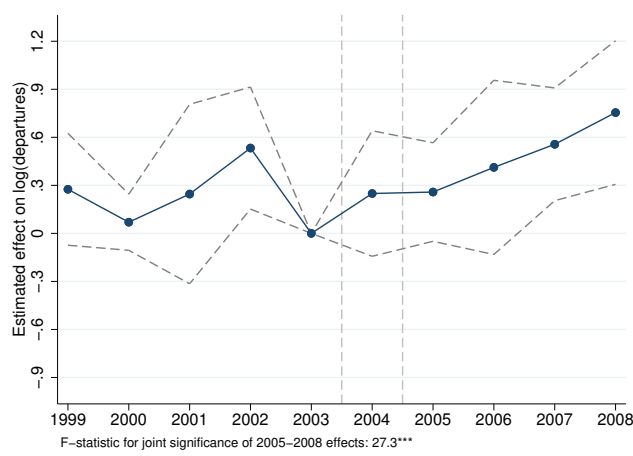
Figure C4: Effects on attorney departures, excluding seven districts with a fired US Attorney



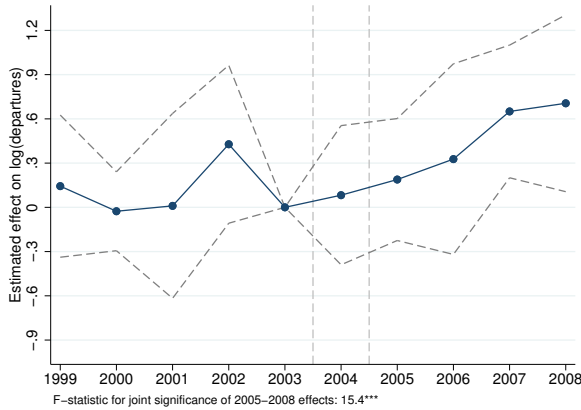
*Notes:* Unit of observation is district-year. Observations are weighted by the average number of Assistant US Attorneys in the district (averaged across years). Estimates come from a Poisson pseudolikelihood regression in which the dependent variable is the number of attorneys departing each district in a given year. All regressions include district and year fixed effects. Standard errors are clustered at the state level. Plotted coefficients are on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator and one in the state with the most-conservative-most-liberal senator. “Full variation” estimates use raw NOMINATE scores. “Party-only” estimates use a dummy for whether the state had two Republican Senators. “Idiosyncratic variation” estimates also control for 1995 and 2007 NOMINATE scores interacted with year effects (see Figure B1). Relative to Figure 5, these analyses exclude the seven districts that experienced a controversial firing of a US Attorney in December 2007 (NM, CA-Central, WA-Western, CA-Northern, AZ, NV, WI-Western).

Figure C5: Dropping main White Collar districts

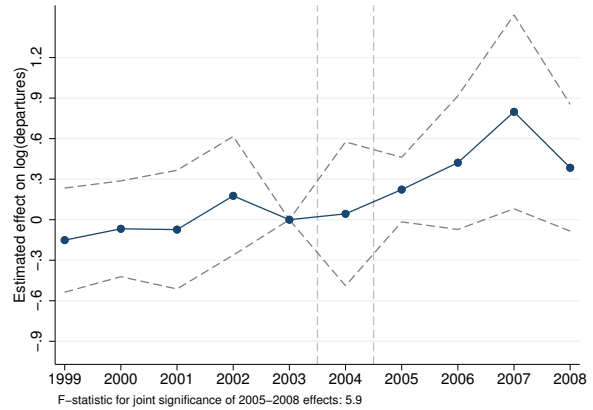
(a) Full variation



(b) Full variation

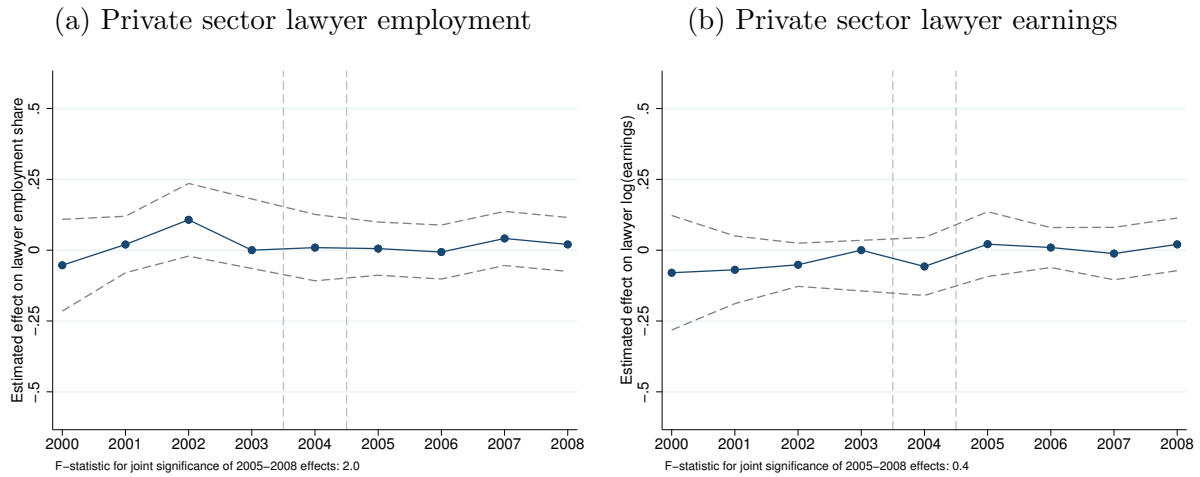


(c) Idiosyncratic variation



*Notes:* Unit of observation is district-year. Observations are weighted by the average number of Assistant US Attorneys in the district (averaged across years). Estimates come from a Poisson pseudolikelihood regression in which the dependent variable is the number of attorneys departing each district in a given year. All regressions include district and year fixed effects. Standard errors are clustered at the state level. Plotted coefficients are on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator and one in the state with the most-conservative-most-liberal senator. “Full variation” estimates use raw NOMINATE scores. “Party-only” estimates use a dummy for whether the state had two Republican Senators. “Idiosyncratic variation” estimates also control for 1995 and 2007 NOMINATE scores interacted with year effects (see Figure B1). Relative to Figure 5, these analyses exclude the districts most specialized in White Collar crime.

Figure C6: Correlation between identifying variation and private sector opportunities



*Notes:* Unit of observation is state-year. Observations are weighted by the average number of Assistant US Attorneys in the district (averaged across years). All regressions include state and year fixed effects. Standard errors are clustered at the state level. Plotted coefficients are on justice-specific NOMINATE scores interacted with year effects. NOMINATE scores have been normalized to equal zero in the state with the most-liberal-most-liberal senator and one in the state with the most-conservative-most-liberal senator.