Unaccountable Midnight Rulemaking? Evidence from Public Comments

Jed Stiglitz*

Stanford University–Department of Political Science
Stanford Law School

January 23, 2012

Draft: Comments Welcome

*Contact: jeds@stanford.edu. I thank Tino Cuellar, Barbara Fried, Dan Ho, Anne Joseph O'Connell, Mitch Polinsky, Connor Raso, Matthew Spitzer, Matthew Stephenson, Alan Sykes, Chris Warshaw, Barry Weingast, and the Hon. Stephen F. Williams for helpful comments or discussions. The paper also benefited from the comments of participants at a Stanford University seminar and a CELS panel. I use colors to highlight patterns in several figures and textual displays. However, the document is entirely grey-scale compatible (for example, red will appear as dark grey in a printout and this is sufficient for interpretation of the figure).

ABSTRACT

Pundits and scholars have long criticized so-called midnight regulations on a variety of grounds. Although scholars have demonstrated that rulemaking activity increases during the midnight period, no study has addressed the question of whether midnight rules in fact systematically differ in any material way from rules passed in other times of a president's administration. In this essay, I follow a common critique of midnight rules by asking whether the president passes relatively unpopular rules during the midnight period. I address this question by examining a new data source: tens of thousands of comments submitted by the public in response to NPRMs issued by an agency, OSHA, over a 27 year period. Although I find some evidence of politics in rulemaking—for example, first term presidents, hoping to win reelection, appear to issue fewer unpopular rules at the end of their terms—the data do not appear to support the idea that the president uses the midnight period to systematically smuggle unpopular rules into law.

1 Introduction

On February 4, 2009 the House Committee on the Judiciary held a hearing entitled "Midnight Rulemaking: Shedding Some Light." During this hearing, witness after witness appeared before Congress to criticize regulations passed in the final months of a president's administration. Although participants expressed a variety of concerns about these "midnight" regulations, accountability was the dominant theme. Because agencies pass midnight regulations after an election, at a time when the administration is not electorally accountable for its actions, the president is free to disregard the preferences of the electorate. This accountability failure, participants reasoned, allowed the president to pass unpopular regulations with minimal political fallout.

This hearing was not the first challenge to midnight rules. Indeed, scholars have long expressed concern over midnight rulemaking, and many have observed that rulemaking activity appears to increase in the final year of a presidential administration. In the most comprehensive and detailed empirical study of the topic, O'Connell, for example, finds that "agencies complete more rulemaking actions in the final three months of a President's administration than in any other year's final quarter" (2008, 957; see also 2011). Using different data, Cochran (2001) and Dudley (2001) support O'Connell by finding that the number of pages published in the Federal Register dramatically increases in the final quarter of a president's administration relative to the same period in non-election years.

By these measures, ample evidence suggests that presidents tend to increase rulemaking activity as they exit office. It is not clear, however, how to interpret this increase in rulemaking activity. Are the midnight rules qualitatively different than the rules that a president passes at other times in his administration? Following the suggestion from the House hearing, does the midnight period of unaccountability allow the president to smuggle unpopular rules into law? A count of pages published in the Federal Register or of the num-

ber of final rules passed in a period is informative about the volume of rulemakings. But this counts-based methodology is not informative about the nature of rulemakings in general, much less about any difference in the nature of midnight and non-midnight (daylight?) rulemakings.

After all, the increase in rulemakings in the midnight period may be benign. Perhaps, for example, midnight regulations represent principally instances of agency procrastination. Or perhaps the end-of-term spikes in final rules simply reflect a spike in well-considered mid-term rule-making initiations that have run their course. O'Connell (2008), for example, finds that presidents tend to increase the number of Notice of Proposed Rulemakings (NPRMs) in the second and third years of their administrations. Newspaper accounts, of course, detail instances of unpopular midnight rulemakings. Under Clinton, for example, OSHA passed a rule in the midnight period imposing costly workplace ergonomics standards. The United Parcel Service estimated that compliance with the rule would cost the company $20 billion; estimates of cost to the economy as a whole ran into the hundreds of billions.[1] George W. Bush, for his part, passed a rule in the midnight period that allowed surface coal mining within 100 feet of streams, enraging environmental groups.[2]

But these anecdotes cannot tell us whether midnight rules in fact represent the scourge to our democracy that many suggest. It is, for example, possible that the type of midnight rules identified by journalists as problematic occur throughout a president's administration. Perhaps, for example, Clinton passed rules similar to the ergonomics standard not just in the midnight period, but also earlier in his administration. The midnight rulemaking phenomenon, that is, may be largely a product of the journalist news cycle: we notice ideological rulemakings toward the end of the administration because journalists seek out

[1] http://www.nytimes.com/2000/11/18/business/18SAFE.html. Of course, companies have an incentive to over-state compliance costs, but the fact that Congress later overturned the rule—in the first and only exercise of the Congressional Review Act—suggests that the regulation was a reach for the administration.

[2] See, for example, Earth Justice (2009).

stories about the transition in administrations in this period; but the rules themselves do not differ qualitatively from rules passed in earlier periods. If so, midnight regulations, as such, hardly seem to "represent a growing threat to the rule of law and the rights of individuals" that some scholars and journalists suggest (Morriss et al 2003).

Of course, it is also possible that journalists and legal observers have it right—that presidents systematically use the midnight period of unaccountability to pass rules that have little popular support. Certainly, we have anecdotal evidence that presidents exploit the period for this purpose. And the evidence on the volume of rulemakings is consistent with this view. But, as suggested above, the interpretation of these two sources of evidence is unclear: the anecdotes may be misleading and the volume of rulemakings may increase for a variety of reasons.

In this essay, I examine a new data source: the comments submitted in response to proposed rulemakings. Unlike simple counts of rules, these comments allow us to study systematically the *nature* of rulemakings, both inside and outside the midnight period. Most importantly, it is possible to determine whether agencies push through regulations that have little popular support during the period of unaccountability. Suppose, for example, that the rules agencies pass in the midnight period have substantially higher levels of adverse comments than the rules that the agencies pass in earlier periods. Such a pattern indicates that presidents employ the midnight period to smuggle in unpopular rules. If on the other hand, the level of adverse comments is not higher for rules passed in the midnight period, this indicates that the midnight regulations do not pose the grave threat to our democracy that many observers suggest.

This essay proceeds as follows. First, I discuss the normative concerns that midnight rulemaking raises and I emphasize that the prevailing counts-based methodology interfaces with these concerns in only a limited way. Second, I introduce the comments data that I employ in this essay and discuss the methods used to analyze the data. Third, I discuss the

results from this study. My conclusions follow.

2 Normative Concerns and Empirical Regularities

Playing off "midnight judges," the term "midnight regulation" appears to have emerged in the early 1980s as a reference to the Carter Administration's efforts to pass last-minute regulations ahead of Reagan's inauguration. In fact, Carter was neither the first nor the last president to increase rulemaking activity toward the end of his term. Using the number of pages contained in the Federal Register as a measure of rulemaking activity, for example, Cochran (2001) finds a pattern of increased rulemaking in post-election quarters dating to 1948.[3] Using counts of final rules as reported in the Unified Agenda, O'Connell (2008, 2011) similarly finds a significant increase in rulemaking activity in post-election quarters from the Reagan Administration through the Clinton Administration. President GW Bush, likewise, appears to have increased the tempo of rulemaking toward the close of his administration, prompting a House hearing on the subject and inspiring countless op-eds.

Although scholars and pundits almost universally adopt a dim view of midnight rules, it is difficult to identify precisely the motivating concern. Midnight regulations, after all, must still conform to the strictures of the relevant enabling statutes, of the Administrative Procedures Act (APA), and of the constitution. That is, they represent lawful policymaking actions. And the president who passes the rules continues to operate under the legitimate democratic mandate of his elected term in office. Acknowledging the difficulty of identifying the normative basis of concern, Beerman remarks that "the arguments of principle [regarding midnight rulemaking] are somewhat difficult to grasp" and that both he and previous observers "have found it difficult to articulate exactly what is wrong" with midnight regulations (2003, 951 and note 8).

[3] In particular, he finds a 17 percent increase in rulemaking in the quarter following an election relative to the same quarter in non-election years.

Scholars have proposed various reasons to disapprove of midnight rules. In his pioneering study, Beerman, for example, writes that the outgoing president "should not order significant regulatory actions that the new President or his newly appointed officials are likely to oppose" (2003, 948). This position plausibly reflects a sense that the outgoing president has lost a measure of democratic legitimacy during the midnight period. He also suggests a more pragmatic explanation: "The new administration should not be forced to spend its first several months in office digging out from under a large pile of activity that occurred immediately before the transition" (2003, 948). Thus, midnight rules plausibly throw sand into the gears of the new administration, hindering its effort to carry out the program of policies most recently endorsed by the electorate. Mendelson, likewise, regards midnight regulations as "unseemly," and notes that "the agency's choice in the last few weeks to proceed regardless of the new President's views suggests an unsatisfied craving for power" (2003, 563).[4] Other scholars worry that agencies, fearing the impending midnight deadline, craft hasty and ill-considered rules (Brito and McLaughlin 2008; Rhinehart 2009). Others focus on the possibility that last-minute regulatory actions short-circuit public participation (e.g., Rhinehart 2009).[5]

Although many of these rationales appeal to commonly held normative intuitions and rest on plausible empirical grounds, my focus on this essay is on a different normative concern. Democratic accountability is perhaps the dominant concern in most critiques of midnight rulemaking. During the period between the election and the inauguration, a president who loses the election is not accountable to the electorate. He knows he has lost the election so the preferences of the electorate no longer discipline his behavior. Unchecked by elections, the president is free to pursue a policy agenda that is unresponsive to usual demands of

[4]Consider also Cochran, who expresses a similar concern, arguing that "a midnight deluge undermines our abilities as ordinary citizens to understand and follow the law. Respect for law erodes when it changes for no other apparent reason than the fact that an administration's drop-dead date draws near" (2001).

[5]See Seidenfeld (1991) for an account emphasizing the importance of public participation in rulemaking for civic republican values.

politics.

The typical view is that this gap in accountability carries negative implications. The lack of electoral accountability, that is, allows the president to pursue an ideological and unpopular policy agenda. Or, in a similar view, he may exploit the accountability failure to issue rules as favors to special interest groups. Clinton's ergonomics standards and G.W. Bush's coal mining regulations both exhibit traces of this account. This approach also explains our sharply discontinuous concern over regulations passed before and after an adverse election. Before an election, the president is disciplined by the upcoming election and the rules his administration passes raise few concerns. The president may, of course, disregard the electorate's preferences before an election, but the electorate then enjoys the capacity to punish the president for his behavior. After an adverse election, however, the electorate's ability to respond to presidential mis-behavior is curtailed; this permits the president to pursue policies without regard to the public's views.

Throughout the recent House hearing on midnight rulemaking, participants focused on the problem of accountability in the midnight period. Sub-Committee Chairman Cohen, for example, opened the proceeding by noting that "this midnight regulation period is a time without political accountability, where controversial actions will not cost the Administration's party votes" (Cohen, 1). Representative Nadler, similarly, stressed the notion of accountability: "lack of accountability in its waning weeks enabled the [G.W. Bush] Administration to adopt highly controversial rules on family planning, endangered species and global warming, that may not have passed muster in the more public debate. But since there was no more public accountability, no election to look forward to, they could do what they wanted..." (23). Another witness, Veronique de Rugy, of the Mercatus Center, echoed this sentiment: "[the] lack of accountability frees the president and his administration to enact regulations that previously had been politically impossible" (209). David Mason, of the Heritage Foundation, noted that, "there is a danger that an [outgoing] administration

... might make politically unpopular decisions which it was reluctant to make before the election" (269).

My focus in this essay is on precisely this issue of accountability. In particular, I ask: does the president employ the midnight period of unaccountability to issue regulations that have little popular support? A contention to this effect motivates much—though not all—of the concern over midnight rules. Yet we have virtually no evidence that the president, in fact, uses the midnight period for this purpose. My objective is to assess this question using a new data source: public comments.

3 Data and Methods

The Unified Agenda (UA) is the backbone of this analysis. The Unified Agenda is published on a semi-annual basis in the Federal Register and contains a nearly comprehensive view of an agency's rulemaking activities over the last six months.[6] The UA reports, for instance, when an agency issues a NPRM; when it finalizes a rule; and if the agency withdraws a proposed rule. The UA includes both independent and executive agencies. I have assembled UA reports from 1983 through 2010.[7] In total, the UA reports indicate that agencies initiated over 24,000 rules between 1983 and 2010. The IRS is the most prolific agency over the series, initiating over 1400 rules, or roughly one rule per week for the last 27 years; a number of less active agencies, such as the Federal Council on the Arts and the Humanities, part of the National Endowment for the Arts, initiated only one rule over the series. OSHA, the central focus of this essay, initiated 118 rules over the series; of these, 95 became final rules.

I focus on OSHA for a number of reasons. First, the agency issued one of the most

[6]In a notable omission, the Unified Agenda does not appear to report variances admitted by agencies' to their rules. The EPA, for example, occasionally provides waivers to various regulations; such waivers do not appear in the Unified Agenda, even if the agency takes public comments on the action.

[7]In particular, I obtained XML files of the Unified Agendas from the Regulatory Information Service Center, as division of the Government Accountability Office. Although the UA has existed since 1978, RISC did not have XML files for years before 1983.

high-profile midnight regulations: President Clinton's ergonomics standards, which sought to protect workers from repetitive motion injuries. Although Congress later invalidated the rule in a rare exercise of the Congressional Review Act, critics of midnight rules often point to the ergonomics standards as an example of ill-considered, ideological policymaking. Second, pragmatically, the agency kept good records of the comments it received. Unlike most of the agencies covered by the UA, OSHA retained digital copies of public comments submitted in response to NPRMs. Third, most of the policy issues under OSHA's jurisdiction shift rights and resources between employers and employees. This employee-employer tradeoff maps nicely to the dominant left-right dimension in politics, increasing the odds that we observe ideological policymaking in the midnight period, if a systematic relationship in fact exists. Fourth, OSHA issues a reasonable number of rules over the series. The UA tracks rules using Regulation Identification Numbers (RIN) whereas the comments data follow docket identification numbers.[8] The same RIN can have multiple docket numbers, and each docket must be manually matched to each RIN. In many cases, I completed this matching process only after reading the texts of the rules and comments for their substantive content, cross-checking for dates and other corroborating information such as the title of the regulatory action.

An initial question is whether OSHA engages in midnight rulemaking. Although a number of studies indicate a general increase in rulemaking during the midnight period, scholars have yet to assess the pattern at the agency level. My objective is to interface with the existing literature on this topic and I adopt the standard practice of classifying a final rule as a "midnight rule" if it was issued in the "fourth quarter" of an election year, defined as the period between November 1 of an election year and January 19 of the following year (e.g., Cochran 2001). Notice that, so defined, the "midnight period" encompasses late-term

[8]Also problematically, the docket numbers listed on the comments do not always match with those listed by the agency during the rulemaking process. The website I recover the comments from, regulations.gov, occasionally appears to assign its own docket numbers to rulemaking proceedings.

regulations even for presidents who win re-election. The conventional definition, that is, includes *all* presidential election years, regardless of who wins the election. In this section, I adopt this definition; in the following section, I also consider separately the behavior of first and second term presidents, the latter of whom know well in advance that this is their final term in office. I also follow standard practice by evaluating an agency's proclivity for midnight rulemaking by comparing the number of fourth-quarter election-year rules to the number of fourth-quarter rules in non-election years (e.g., O'Connell 2008; Cochran 2001).

Using these definitions, we can examine an agency's tendency to issue midnight rules by considering: $\frac{3|R_e|}{|R_n|}$, where, within a presidential election cycle, $|R_e|$ is the number of fourth-quarter election-year rules and $|R_n|$ is the number of fourth-quarter rules issued in non-election years. The resulting quantity is interpreted as the relative increase in election-year fourth-quarter rulemaking over non-election-year fourth-quarter rulemaking, and provides a simple metric of which agencies rely on the midnight period to issue their rules. A quantity of 2, for example, indicates that the agency issues twice as many rules during the midnight period than it does in the average fourth quarter of a non-election year; a number below 1 indicates that the agency is inactive during the midnight period relative to the average fourth quarter of a non-election year. For this exercise, I pool data for all election cycles between Reagan's second term and GW Bush's second term (i.e., 1985-2009).[9] The figure I present below thus provide a general sense of whether, over the series, a given agency tends to engage in midnight rulemaking.

In figure 1, I plot the relative increase in election-year fourth-quarter rulemaking for the agencies in the UA dataset.[10] The x-axis of the figure displays the "excess" election-

[9] An "election cycle" is defined, for example, as the period between January 20, 1985 and January 19, 1989.

[10] The UA generally lists both a "parent" and a sub-level agency. For example, OSHA is a sub-level agency and the Department of Labor is a parent agency. I attempt to keep the analysis at the lowest informative level of administration. Generally, this means that I consider sub-level agencies. However, in a number of cases, I collapse sub-level agencies into their parent codes. The EPA, for example, lists regional offices as sub-level agencies; I collapse distinctions between regional offices. Similarly, most agencies list the "Office of

year rulemaking, as defined above; the y-axis displays the total number of fourth-quarter rulemakings for the agency over the series. As is evident from the figure, most agencies have relative increases above unity, indicating that they ramp up rulemaking activity in the midnight period. Sixty percent of agencies increase their rulemaking activity during the midnight period (112 of 185); over 25 percent increase their activity by a factor of two or more (50 of 185); and about 15 percent increase their activity by a factor of three or more (27 of 185). It is notable, though, that nearly forty percent of agencies *decrease* their rulemaking activity in the midnight period. Although most of these agencies either (a) have excess rulemaking quotients near unity or (b) issue few fourth-quarter rules overall, aggregate analyses obscure the fact that many agencies decrease activity in the midnight period.[11] The general pattern, however, is consistent with the findings from earlier research: agencies increase activity in the midnight period.

[Figure 1 About Here]

Importantly, OSHA, the focus of this essay, denoted by a textual point in the figure, increases its rulemaking activity by a factor of 3 during the midnight period. In this sense, OSHA is a relevant agency to consider for midnight rulemaking activity. Consider OSHA's behavior in further detail. Over the series, OSHA issues 12 rules during the midnight period, or approximately 13 percent of the total number of final rules over the 24 years I examine. On the average fourth quarter of an election year, OSHA thus issues 2 rules. By comparison, the agency also issues 12 rules during the fourth quarter of *all* non-election years. Thus, on the average fourth quarter of a non-election year, OSHA issues only .67 rules.

the Secretary" and the "General Counsel's Office" as sub-level agencies. Here, too, I collapse sub-levels into their parent codes.

[11] If we conduct a chi-square test of the frequency of rulemaking activity in normal and election-year fourth quarters, only two of these agencies reach even a rough sense of statistical significance (i.e., $p \leq .2$). Notice that this test is sensitive to the general rulemaking activity of an agency. An agency that publishes many rules is likely to appear in the figure, even if it only modestly changes rulemaking activity in the midnight period. It is also sensitive to the magnitude of the differences between the frequencies. The two agencies that pass this modest test are the U.S. Coast Guard and the Department of State.

It is not clear, however, how to view this increase in OSHA rulemaking activity in the midnight period. Perhaps these midnight rules differ little from the rules passed in other periods of a president's administration. Perhaps, that is, midnight rules represent wholly unexceptional examples of policymaking. Under this view, the surge in rules reflects nothing more pernicious than agency procrastination.[12] Or perhaps the president, in fact, uses the midnight period to smuggle into law unpopular regulations. This scenario speaks directly to the accountability failure that motivates concern about midnight regulations. The simple counts-based methodology, however, cannot adjudicate between these possibilities.

To address the question of whether the president passes unpopular regulations during the midnight period, I consider the comments submitted in response to NPRMs. As part of an e-governance campaign, President GW Bush created a website, regulations.gov, in 2003 as a public portal to the rulemaking process. On this website, the public can view proposed regulations for over 300 agencies and provide comments in response to NPRMs. Citizens can also view comments filed by other members of the public. Fortunately, a number of agencies also uploaded comments from rulemaking proceedings that transpired before 2003. I identified and downloaded all available comments relating to OSHA rulemakings. In total, I downloaded approximately 195,000 comments covering 312 dockets. I saved the comments as PDFs. After discarding "phantom" comments, or false records of comments, I had a database of roughly 180,000 comments. It is unclear whether these phantom comments represent errors on the part of the website—falsely indicating a comment where none ever existed—or on the part of the agency—neglecting to upload the comment.[13]

I then extracted the text from all comments with renderable text. The agency saved approximately 44,000 comments, however, as images, and I OCR'ed these files. I then

[12]This procrastination plausibly leads to rushed rules, discussed above, but it is not related in a direct way to the accountability failure.

[13]The vast majority of these phantom comments relate to proposed rules that would be dropped from the analysis in any event because (a) they relate to rules finalized outside the 1983-2010 period or (b) the proposed rule was never finalized.

extracted the text from these OCR'ed files. I found that many of the original files without renderable text represented hand-written comments, largely in response to a single rule-making, Clinton's effort to enhance indoor air quality standards.[14] It is not possible to OCR hand-written text, and I exclude these comments, on the order of 28,000, from the main analysis below. At the end of this exercise, I am left with the text from roughly 150,000 comments.[15]

Next, I connect the comments data with the backbone of this analysis, the UA data. The challenge to this exercise is that the comments come indexed by docket numbers and the UA data is indexed by RINs, as noted above. The matching process revealed that I had recovered comments for 107 of the 118 RINs created by OSHA over the period between 1983 and 2010; I recover comments for 88 of the 95 RINs associated with a final rule.[16] Thus, the comments data covers over 90 percent of the rules issued by OSHA over the series. The missing comments do not appear to load on the fourth quarter. Of the 7 final rules for which I cannot recover comments, 2 were issued in the fourth quarter, 2 in the third quarter, and 3 in the first quarter.[17] In total, I identify approximately 26,000 comments for rules finalized in the period between 1983 and 2010.

The objective of the analysis below is to determine whether the president smuggles unpopular rules into law during the midnight period. It is possible to partially respond to this query by simply examining the counts of comments received for rules in and out of the midnight period. Presumably more unpopular or politically controversial rules attract more

[14] As discussed below, this proposed rule was never finalized and is therefore dropped from the analysis.

[15] As I discuss below, another approximately 15 percent of these comments have unusable text. Most commonly, the text is unusable because the OCR engine attempted, unsuccessfully, to read hand written text. In addition, as also discussed below, some of these 150,000 comments relate to rules outside the period covered by the UA data.

[16] For roughly 20 percent of the RINs, the agency proposes and then withdraws a rule.

[17] The fourth quarter is defined by the dates above. The "first quarter" runs from January 20 to April 24, the"second quarter" runs from April 25 to July 29, and the "third quarter" runs from July 30 to October 31. The first through third quarters, that is, equally partition the days not covered by the fourth quarter, defined as it is for conceptual reasons. Note that the fourth quarter is modestly shorter than the other quarters.

comments. A more complete assessment, however, requires us to consider the comments themselves. It would be useful, in particular, to know whether the comments submitted for a particular rule tend to support or oppose the regulatory action.

I adopt the standard Naive Bayes classification algorithm to assess whether a given comment supports the regulatory action in question. I discuss the mathematical details of the algorithm at length in the appendix; here, I address the intuition. The fundamental idea behind the algorithm is that the probability of observing a given word depends on the type of the document in question. Supporting comments, for example, may tend to emphasize the importance of the proposed regulation, or urge the agency to expedite the rulemaking process. In the context of OSHA, supporting comments may stress the positive consequences of the rule, using words such as "health," "safety," or "worker." By contrast, disapproving comments may tend to emphasize a set of concepts expressed by words such as "cost," "expense," "unnecessary," and "excessive."

It is possible to estimate the probability that each class generates each word by hand-coding a sample of comments as supporting or opposing the OSHA action.[18] For this hand-coded sample, that is, we know the class of each document. Using the labels, we can estimate the probability of observing each word conditional the class of a document using a standard maximum likelihood approach. This step is commonly referred to as "training" the algorithm. In the next step of the algorithm, these word parameters help us to determine the hidden or "latent" classification of documents that we have not hand-coded.

The intuition behind this next step is plain: if a comment employs many "supporting" words and few "opposing" words, the comment likely supports the proposed rule; by contrast, if a comment employs many "opposing" words and few "supporting" words, the comment likely opposes the proposed rule. As suggested above, for example, if a comment refers to

[18] As I discuss in the appendix, I use different training samples for Democratic and Republican rules under the theory that the same word can have different meanings in the context of regulatory and deregulatory actions.

"workers" repeatedly, the comment most likely supports the OSHA regulation. By contrast, a comment that uses the word "costs" is likely to oppose the regulation. I refer the reader to the appendix for the details of the algorithm and its application to the data.

I employ this algorithm to classify the roughly 26,000 comments for finalized rules in the dataset. As explained at length in the appendix, I classify a comment as "supporting" the rule if it is more likely than not, based on the algorithm, that the comment supports the rule.[19] Before addressing the results from this exercise, it is useful to illustrate the application of this algorithm to two comments, one positive and one negative. Both comments relate to Clinton's ergonomics standard.

Below, I have transcribed the two comments using a shading scheme to illustrate the words that drive the algorithmic results. Words in typewriter font underlined in red associate with an opposing classification; words in typewriter font without underlining associate with a supporting classification. The darker the shading for the word, the stronger the "polarity" of the word (e.g., a dark word with red underlining indicates a strongly negative word).[20] Words colored in black in Times New Roman font do not enter the algorithm: as described in the appendix, such words either occur very frequently, and thus tend to provide little information about the nature of the comment, or occur very infrequently, and thus threaten to create a problem of over-fitting.

Now consider the two comments:

Example 1: Opposing Comment. As president and general manager of Eagle Dist. Inc. in Freeport, Illinois, an independent beer distributorship employing 11 people, I am writing to urge you to withdraw the proposed rule on ergonomics Docket S-777. I am familiar with the positions of Anheuser-Busch, the U.S. Chamber of Commerce the

[19] In addition to the classes of supporting and opposing rules, I also estimate the probability that the comment is purely procedural in nature. A number of comments, that is, do nothing more than ask for more time to respond to the NPRM, without expressing any view on the regulation. I attempt to remove these comments from the main analysis.

[20] I use, in particular, the log relative risk to color each word. That is, I calculate, $log[\frac{p(w_i|\text{"}supporting\text{"})}{p(w_i|\text{"}opposing\text{"})}]$, and use the resulting scores to shade the words. See the appendix for details on the mechanics of the algorithm.

National Beer Wholesalers Association and the National Coalition on Ergonomics. I would like to join in those detailed comments as well as the comments of others that have been raised in opposing the proposed rule. The ergonomics rule is ill founded and should be withdrawn. The legal deficiencies of the proposal have been well addressed by others. I particularly want to comment on the practical and economic deficiencies of the proposal. The rule is based on theory and speculation and not sound medical science. We are convinced this rule will be understood as a declaration that hard work in and of itself causes injury. The rule fails to provide any definition of when any given physical activity will be considered injurious. The rule also fails to tell employers either what specific hazards must be eliminated or what specific controls must be implemented. Apparently an employers only guides in that regard will be employee complaints and then after the fact the judgment of OSHA inspectors who will undoubtedly know an ergonomic hazard when they see one. This kind of uncertainty and arbitrariness makes such a rule completely impractical. The rule would require employers to train employees about OSHA's theories which we believe medical science has shown to be false. Such training will improperly mislead people to believe that vigorous physical activity which is essential to good health is likely to cause injury or disease. The rule improperly defines mere symptoms of discomfort as injury or disease and requires employers to respond to complaints of such symptoms by investigating and installing feasible controls. The rule also requires employers to engage in costly experimentation with unproven controls that may well present a greater hazard than the original activity. All of this is to be done in the clearly misguided belief that employers can somehow eliminate employee complaints about the aches and pains that all of us experience as an aspect of daily life. For these reasons, and for those presented by the organizations listed above, we urge you to withdraw this misguided and fatally flawed rule

Example 2: Supporting Comment. I am writing to express my support for OSHA's proposed ergonomics standard. It has been nearly 10 years since then-secretary of labor Elizabeth Dole first announced the need for ergonomics standard. Since then more than 6 million workers have suffered disabling ergonomic injuries while numerous studies have proven there is a relationship between musculoskeletal disorders and ergonomic hazards in the workplace. Each year these injuries account for about one-third of all serious injuries recorded by the bureau of labor statistics. According to these data the main causes of msds — overexertion and repetitive motion — result in nearly 600 000 lost-time injuries each year. These injuries are estimated to cost the u s economy more than billion annually. After the senate failed last year to pass legislation that would have prohibited OSHA from moving forward, and after nine years of previous delay, OSHA finally issued a proposed ergonomics rule. The opposition still is trying to stop the proposal from becoming a final standard. They aim to delay the rule making until the current administration, which supports ergonomics standard, leaves. We know that msds are costly both physically and economically and

we know they can be prevented. Employers must take action to fix hazardous jobs. OSHA must issue a final ergonomics standard. Attempts to further delay a final ergonomics rule will result in additional preventable injuries that are costly and destroy workers lives. Now is the time to move forward with, not delay, an ergonomics standard.

The most obvious point to draw from these examples is that, sensibly, the algorithm finds many more negative words in the opposing comment than in the supporting comment. Approximately 65 percent of the words in the supporting comment were classified as supporting words; by contrast, roughly 75 percent of the words in the opposing comment were classified as opposing words.

It is also worth noting several features of the words that drive the algorithmic results. As expected, the opposing comment employs words that, a priori, an observer might expect to convey a negative attitude toward a regulation: "deficiencies," "fails," "impractical," "misguided," for example. And, similarly, the supporting comment employs words that, a priori, an observer might expect to convey a positive attitude toward a regulation: "injuries," "preventable," and so on. Thus, the opposing comment spends its time discussing the problems created by regulation, whereas the supporting comment spends its time discussing the problems addressed by the regulation.

The algorithm also identifies informative words that, a priori, an observer would not expect to convey a positive or negative attitude toward regulations. For example, "constitutional" is not clearly a word that one would expect to reflect opposing or supporting sentiment. Plausibly, both those opposed to regulation and in favor of it could refer to the constitution. Yet, in the discourse of OSHA regulations, "constitutional" happens to associate with opposition to proposed rules. In the discourse of workplace regulations, that is, the word "constitutional" refers to certain rights, including the ability to be free from supposedly onerous regulations. One could imagine the word carrying just the opposite polarity—with "constitutional" reflecting a sentiment supporting instead of opposing proposed regulations–

but, empirically, in the language of comments submitted to OSHA, the concept is clearly associated with opposition to proposed rules. Although neither of the examples uses the word "constitutional," much the same can be said of words such as "employer" or "organization."

Other words appear to owe their polarity to the style of writing adopted by commenters rather than to the concept expressed by the word. The examples above, for example, show that "ergonomics" has a negative polarity whereas "ergonomic" has a positive polarity. Obviously, little difference exists between the concept implied by "ergonomic" and the concept implied by "ergonomics." Yet, due to differences in writing style, the two words carry different statistical meanings. Supporters of the ergonomics standard, as above, were more likely to refer to "ergonomic hazards" or "ergonomic dangers," whereas opponents of the standard tended to write, simply, "ergonomics." In this sense, the words "ergonomic" and "ergonomics" contain information about the class of comment, despite the fact that the concepts underlying these two words do not differ.

Cross validation is a standard test of this type of algorithm, a topic discussed in the appendix. The central idea is that we want to evaluate how the model performs on comments that did not contribute to the "training sample." One way to conduct such an evaluation is to divide the full training sample of N comments into N different samples, each consisting of $N-1$ documents, where the Nth document is held out as a test case, and the $N-1$ documents generate the parameter estimates. The performance of the algorithm on the test cases allows us to assess how the algorithm will perform on the vast majority of comments for which we do not have hand-coded labels. If the algorithm correctly classifies the comments in these hold-out samples, we should be fairly confident about its performance in the broader sample of comments. As I discuss in the appendix, the algorithm's overall classification performance is between between 70 and 85 percent, a reasonable level for machine learning applications. Of perhaps more importance, the validation procedure suggests minimal bias.

Using these data and tools, we can now address the question—does the president use

the midnight period the smuggle unpopular rules into law?

4 Results

I examine both the volume and the content of comments submitted in response to OSHA rulemaking efforts. Both metrics plausibly contain information about the popularity of a given rulemaking. Plausibly, the public responds to a highly unpopular rule by submitting many comments. It is also possible, though perhaps less likely, that the public rewards a highly popular rule with many comments lavishing praise on the agency. I also examine the content of the public comments. A highly unpopular rule, it is fair to suppose, tends to attract negative comments.

4.1 Comment Volume

To provide an initial assessment of the motivating question in this essay, I consider the counts of comments submitted to OSHA for each rule. Not surprisingly, the distribution of comments is highly skewed: a few proposed rules receive an enormous number of comments; most proposed rules receive relatively few comments.

Among finalized rules, Clinton's ergonomics standards received far and away the highest number of comments.[21] This rule, famously, required employers to provide ergonomically safe workplaces that would reduce the incidence of repetitive motion injuries. The ergonomics standard received almost 12,500 comments, over 30 times the average number of comments

[21] OSHA never finalized two proposed rules with the highest counts of comments. The proposed rule with the largest number of comments, by far, is OSHA's proposed Indoor Air Quality in the Workplace Standards, NPRM 59 Fed. Reg. 15968. This proposed rule, initiated by President Clinton in April 1994, sought to protect workers from hazardous indoor air conditions by, among other requirements, forcing employers to develop sophisticated ventilation systems. Another unfinalized rule with a large number of comments stemmed from a 2007 effort by President GW Bush to enhance protections afforded to employees working with explosives and firearms. This proposed rule received about 10,200 comments, or nearly 28 times the number of comments received for the average finalized rule. See NPRM 72 Fed. Reg. 18792. In another paper under construction, I address the question of responsiveness to comments.

for a finalized rule. Employers appear to have written the majority of comments for this rule, with almost all opposing the regulation and citing the high compliance costs. Employees also submitted comments, and they generally favored the new standard. OSHA's bloodborne pathogen standard, issued in 1991 amid mounting concern over AIDS, also received a large number of comments: about 3,300, or roughly 9 times the average number of comments for a finalized rule. This rule sought to reduce the transmission of bloodborne diseases, centrally Hepatitis B and HIV, by encouraging safe workplace practices and the use of protective clothing and equipment.[22] The proposed rule attracted widespread attention from the medical community. Dentists represented a particularly vocal group, with most appearing to oppose the new regulation.

By comparison, over half of OSHA's finalized rules received fewer than 100 comments. These less-noted rules cover a wide range of mostly limited policy objectives, from modestly revising permissible exposure levels to cotton dust in textile plants,[23] to revising the language used to regulate exit routes.[24]

In figure 2, I plot the distribution of comments to finalized OSHA rules. The figure makes plain the fact that the majority of rules receive relatively few comments, at least by comparison to the ergonomics standard. Nevertheless, fully 18 final rules—or between 2 and 3 rules per presidential administration—receive at least 250 comments.

[Figure 2 About Here]

A central question is whether midnight rules receive a greater number of comments than rules passed in other periods of a president's administration. I investigate this question

[22] See Final Rule, 52 Fed. Reg. 45438.

[23] 48 Fed. Reg. 26962.

[24] 61 Fed. Reg. 47712. The primary purpose of this modest rule is to exchange references to "means of egress" for "exit routes."

by estimating the following equation:

$$E[log(comments_i+1)] = \alpha+\beta_1 Last_i+\beta_2 PostElection_i+\beta_3 PostElection_i X Last_i+\Gamma Q_i, \quad (1)$$

where $comments_i$ is the number of comments submitted for rule i, $Last_i$ is an indicator for whether the rule is issued in a president's last term in office, $Election_i$ is an indicator for whether the rule is issued in an election quarter, and Q_i represents quarter fixed effects. Notice that $Last_i$ takes a 1 if the president is in his second term *or* if he loses the election at the end of his first term. Two coefficients hold central interest. First, the coefficient on election year informs us of whether OSHA issues relatively comment-inducing rules during election quarters that result in re-election. Notice that I include quarter fixed effects, so this election coefficient is adjusted for any general fourth-quarter rulemaking effect. Second, the interaction between election quarter and last term informs us of whether, above and beyond the election quarter effect identified in β_2, OSHA issues relatively comment-inducing rules during election quarters that produce a new presidential administration. If presidents issue highly unpopular rules during the midnight period due to their unaccountability to the electorate, one might expect positive coefficients on the interact term.

I report the results from this regression in the first column of table 1. There, we see that OSHA rules issued in election quarters that result in re-election do not appear to attract a significantly greater number of comments than rules issued in non-election fourth quarters. The coefficient on election quarter is far from statistically significant. As is the coefficient on the interaction term. This pattern suggests that OSHA rules issued in election quarters that produce a new presidential administration do not, on average, attract more comments than OSHA rules issued in election quarters that do not produce a new administration. If the president is on the way out the door, in other words, OSHA's rules do not attract a greater number of comments.

[Table 1 About Here]

The rulemaking process is often lengthy. A plausible concern with this first equation, therefore, is that the coefficient on *Last* understates the accountability problem with midnight rulemaking. By the time a first term president can reasonably forecast that he will lose the election, it is too late to initiate unpopular rulemakings. The coefficient on *Last* thus reflects the behavior of presidents effectively disciplined by elections, in addition to the behavior of presidents unencumbered by concern over re-election. A straightforward way to resolve this issue is to focus on second term presidents, who know with certainty that they will not face re-election. For these presidents, the accountability problem is potentially acute.

I estimate the same regression, but replace *Last* with an indicator for whether the president is in his second term in office. The results from this regression, reported in the second column of table 1, indicate that OSHA rules issued in the post-election quarter of a president's second term may attract a somewhat greater number of comments than rules issued in a first term or in non-election quarters. The coefficient on the interaction between second term and election quarter is modestly larger than the corresponding interaction term from the first model. And although it remains statistically insignificant, it begins to approach the conventional threshold ($p = .26$). As indicated by the coefficient on post-election quarter, rules issued in the election quarter of a first term president's administration do not attract more comments than rules in non-election year fourth quarters. Notice also that the coefficient on second term is not significant, suggesting that second term presidents do not, across quarters, issue rules that attract a relatively large number of comments.

These first two regressions thus provide modest support for the notion that president's use the midnight period to issue unpopular rules, as judged by the number of comments received during election quarters in which they plausibly do not fear the electorate. The coefficients on the interactions between election quarter and last term and election quarter

and second term, that is, are positive, even if not significant at conventional levels. However, even though the number of comments is log-transformed, a natural concern is that the ergonomics standard, a clear outlier in figure 2, is wholly responsible for the (modest) apparent increase in comments for last- and second-term election-year quarter rules. To account for this possibility, I estimate the same regressions, though this time removing comments related to the ergonomics standard from the data. I report the results from these regressions in columns three (last term results) and four (second term results). Outside of the interaction terms, all of the coefficients in these regressions remain essentially unchanged from the first two regressions. The coefficients on the interaction terms, however, reduce considerably in magnitude. The relevant coefficients now do not remotely approach conventional levels of significance. This exercise indicates that, to the extent midnight rules on average attract a greater number of comments, the pattern is largely due to a single, highly unusual rule—Clinton's ergonomics standard.

4.2 Comment Content

A more complete analysis, of course, requires us to examine the comments themselves. Even if midnight rules do not attract a greater number of comments, it remains quite possible that the nature of comments submitted to midnight rules differs materially from the nature of comments submitted to other rules. It is quite possible, in particular, that midnight rules attract a disproportionate number of negative comments.

Using the procedure described above and in the appendix, I estimate whether each of the roughly 26,000 comments submitted for finalized OSHA rules supports or opposes the regulatory action. Based on these estimates, approximately 70 percent of the public comments oppose the rule in question, a proportion consistent with the hand-coded sets of data. However, this aggregate proportion masks the fact that many regulations, in fact, appear to receive substantial public support.

OSHA's 1983 amendment to its hearing conservation standard is instructive.[25] In that regulatory action, the Reagan administration effectively relaxed a midnight rule issued by Carter on January 16, 1981.[26] Carter's hearing amendment sought to reduce workplace-induced hearing impairment by setting forth a highly detailed "hearing conservation program" for loud workplaces involving noise testing, noise monitoring, employee training, and employee protective equipment. Early in the Reagan administration, OSHA softened this regulation by removing the specific requirements of the Carter rule and replacing them with a "performance approach," which "generally allows the employer to chose his own method of complying with the [regulatory] obligations."[27] The 1983 amendment, for example, allowed employers to monitor noise levels in the "area" of employees, whereas Carter's rule required employers to sample noise levels close to employees' ears. Although several labor groups generally opposed the Reagan amendments, the vast majority of commenters supported OSHA's 1983 action. Based on the machine coding results, approximately 92 percent of the comments submitted for the Reagan amendment support the rule.

Now return to the main focus of this essay—an evaluation of midnight regulations. To approach the data, consider initially figure 3. There, I plot the proportion of supporting comments for each final rule in the dataset. I display rules adopted outside the midnight period with unfilled blue circles; rules passed in the midnight period, defined as the fourth quarter of an election year, appear as solid red dots. The size of the dot or circle reflects the number of comments submitted for the rule. I use arrows to point to several significant OSHA rules. I denote the initiation of a new presidential administration with a dashed grey line, and mark the name of the president who is *starting* his term below the grey line.

[Figure 3 About Here]

[25]Final Rule, 48 Fed. Reg. 9738.

[26]Final Rule, 46 Fed. Reg. 4078. Notice that this is an example of a "crack-of-dawn" response to a midnight regulation. See O'Connell (2011) for an insightful and more comprehensive analysis of rulemaking after political transitions.

[27]Final Rule, 48 Fed. Reg. 9738.

Several conclusions follow from this figure. Most obviously, the ergonomics standard was a true outlier in terms of the number of comments submitted. This is evident, too, in figure 2, but figure 3 provides visually striking evidence of how unusual the ergonomics standard was. A second conclusion, however, is that commenters have quite varied views of OSHA regulations. Some, such as Reagan's hearing conservation amendment, receive fairly universal support from commenters. For other rules, such as HW Bush's bloodborne pathogens standard, commenters almost universally disapproved of the rule. Most rules fall somewhere between these extremes.

It is also evident from the figure that, even within the midnight period, commenters' support for rules is varied. Consider the midnight period in which Clinton passed the ergonomics standard. There, we see that, in addition to the famous ergonomics standard, Clinton also passed (a) a rule enhancing protections for construction workers who erect steel,[28] and (b) a rule simplifying the reporting requirements for occupational injuries and illnesses.[29] By and large, commenters disapproved of the steel erection standard. A large number of commenters, for example, opposed the rule, in part, because it imposed requirements on general contractors who supervise the project; any regulatory requirements, they thought, should apply exclusively to the subcontractor hired to perform the steel work. The final rule maintained general contractor liability, despite the opposing comments. By contrast, the majority of commenters favored the rule simplifying reporting requirements. Most of the comments opposing the rule objected to OSHA dropping workplace violence injuries from the reporting log. The final rule deletes this proposed exemption.[30] Thus, considerable heterogeneity exists in commenters' support for OSHA's rules, both inside and outside the midnight period.

Having examined the data visually, I now turn to a regression framework to assess the

[28]Final Rule, 66 Fed. Reg. 5196.
[29]Final Rule, 66 Fed. Reg. 5916.
[30]Final Rule, 66 Fed. Reg. 5945.

data systematically. In particular, I estimate,

$$P(c_i = 1) = F(\alpha + \beta_1 Last_i + \beta_2 PostElection_i + \beta_3 PostElection_i X Last_i + \Gamma Q_i), \qquad (2)$$

where F is the logistic CDF, c_i is an indicator for whether comment i supports the rule, and I otherwise maintain the definitions of the variables identified above. Notice, however, that the unit of observation in this regression is the comment and not the rule.[31] As before, β_3 is the central—though not the only—coefficient of interest. If the coefficient is negative, this indicates that rules passed in the fourth quarter of the election year of a president's last term tend to attract relatively adverse comments; if positive, just the opposite.[32]

I report the results from this regression in the first column of table 2. The coefficients from this regression suggest the following story. In general, last term presidents do not issue more unpopular rules than presidents who continue in office. This is suggested by the non-significant coefficients on both last term and the interaction of last term with election quarter. Regardless of whether a president continues in office or is ousted, however, he appears to issue *more* popular rules in fourth quarters that follow an election (post-election quarter) than in fourth quarters that do not follow an election.

[Table 2 About Here]

The substantive conclusion suggested by this regression is that an accountability problem does not afflict midnight rules. The probability of receiving an opposing comment to a rule passed between the election and inauguration day does not appear to much matter on whether the president is on his way out or whether he is continuing in office. Indeed, comments submitted for rules passed in quarters following elections tend to be *more* favorable than comments submitted to other fourth-quarter rules. From the perspective of the

[31] I cluster standard errors on rules.

[32] Note that I drop procedural comments asking for additional time.

accountability thesis, this is a puzzling pattern. A plausible explanation, however, is that presidents do not want to discuss controversial rules in the months leading up to an election. As a result, the rules finalized shortly after an election tend to be relatively popular or at least bland.

A concern with this specification is, as above, that first term presidents do not have time to roll out controversial regulations in the period between the loss of the election and the inauguration of a new president. Rather than relying on the time consuming notice and comment process, a president who is unexpectedly leaving office is more likely to rely on executive orders or other policy instruments that can be quickly initiated and completed. I address this possibility by again focusing on second term presidents, who obviously anticipate their departure, and first term presidents, who hope to remain in office for another term. Second term presidents, unlike first term presidents, have time to plan a midnight deluge of unpopular rules.

As reported in the second column of table 2, the results from this approach suggest a possible accountability issue associated with midnight rulemaking. In particular, the results suggest that comments submitted for rules passed by second-term presidents in the quarter following an election are considerably more likely to be opposing than comments submitted for rules passed by first-term presidents in the corresponding quarter. That is, suppose the baseline is the likelihood of an opposing comment to an election-year fourth-quarter rule passed by a first term president. Relative to this baseline, a comment submitted for a rule passed by a second-term president in the corresponding quarter is significantly more likely to oppose the regulatory action. This suggests that presidents may reserve highly unpopular rules for the midnight period of their second term, a period in which they know well in advance that they will be unaccountable to the electorate. The president may avoid finalizing unpopular rules until after the election to increase the odds that their vice presidents win

the election.[33] Or, as discussed below, avoiding unpopular rules until the midnight period may be a response to the lame duck status of his administration.

The coefficient on the interaction is partially offset by the fact that second term presidents, in general, pass relatively popular rules. The coefficient on second term is positive and nearly statistically significant ($p = .12$). A plausible explanation for this result is that second term presidents, facing a lame duck dilemma, must generally maintain positive public profiles to accomplish anything of note during their second term in office. As the lights begin to fade on his second term, however, he passes unpopular regulations, as indicated by the coefficient on the interaction of second term and post-election quarter. In this sense, the president treats midnight regulations much like a controversial pardon. Nevertheless, despite the general popularity of second term presidents' rules, the probability of observing a positive comment in the post-election quarter of a president's second term is about 31 percent; the corresponding probability during a president's first term is about 53 percent.

This is, to be sure, a dramatic decrease in the popularity of rules. And so one interpretation of the results is that second term presidents use the midnight period to smuggle unpopular rules past the electorate. This conclusion, however, is undermined by examining the coefficient on post-election quarter. This coefficient indicates that, on average, rules passed in the post-election quarter attract positive comments. The magnitude of this positive coefficient is, indeed, somewhat greater than the magnitude of the negative coefficient on the interaction term. Taken together, the main effect for post-election quarter and the interaction term with second-term presidents basically cancel out. This indicates that the rules passed by second term presidents in the midnight period attract roughly the same level of support as ordinary, non-election year fourth-quarter rules passed by second term

[33] Anecdotally at least, Reagan, for example, appears to have delayed unpopular administrative actions ahead of HW Bush's election. In an article written on October 28, 1989, the Washington Post noted that "the Reagan administration appears to be accelerating the implementation of popular measures and postponing unpopular ones" (Havemann, 1988).

presidents. The election simply has little influence over the type of comments received by second term presidents. The spread noted above—between 53 and 31 percent—appears to be driven largely by first term presidents issuing relatively *popular* rules during the post-election quarter; not by second term presidents issuing relatively *unpopular* rules during the post-election quarter.

As before, a natural concern with the first two regressions is that the ergonomics standard drives the results. Accordingly, I re-estimate the equations after dropping comments submitted for the ergonomics standard. I report the results in the third (last term) and fourth (second term) columns of table two. Little changes as a result of this exercise.

The results, in sum, suggest the following. Midnight rules do not generally attract significantly more comments than rules passed outside the midnight period. Moreover, second term presidents do not generally use the midnight period to pass highly unpopular rules. The rules passed by second term presidents in this period resemble the rules passed by second term presidents in other, non-election year fourth quarters. First term presidents, however, appear to pass relatively popular rules during their midnight periods. Taken together, this pattern indicates that much of the concern about the nature of rules passed in the midnight period is over-stated. Electoral accountability appears to drive up the popularity of late-term rules passed by first term presidents; but the lack of accountability for second-term presidents does not appear to drive down the popularity of their late-term rules.

5 Discussion

This essay charts a new approach to studying midnight rules. It is, in a number of ways, only a first step. Most obviously, this essay focuses on a single agency, OSHA. Over the 27 year series, this agency issues only 95 rules. The fact that I consider a single agency precludes any sweeping conclusions about the role of midnight rules in the broader administrative state.

And, even within OSHA, it is difficult to render solid conclusions due to the small number of rules passed by the agency over the period. So the scope of the data render the conclusions preliminary.

It is also natural to be skeptical of the algorithmic results. I use a standard algorithm, however, and the sample of comments used to train the classifier has the virtue of being representative of the larger population of comments.[34] Measurement error afflicts the algorithmic results, but there is no obvious reason to suspect a form of bias that would induce the observed pattern of results. The evidence from the validation procedure, reported in the appendix, supports the view that the algorithm provides an approximately unbiased assessment comments' general viewpoints.

As an initial effort to study the connection between the content of comments and the content of rules this essay naturally opens several questions. It is, initially, clear that the population of commenters is not representative of the population generally. It is costly to submit a comment, and people who comment on proposed rules tend to have a stake in the regulation. So it is clear that a level of support in the comments of X percent does not imply that X percent of the general population support the rule. It is mainly for this reason that I have attempted to avoid extended discussions over the levels of support. I have, instead, attempted to focus on changes in support for rules passed in different periods of an administration. Are rules passed in the midnight period unpopular relative to rules passed in other times of an administration? If the process that generates comments remains the same in the relevant periods of an administration, we can at least address this more modest question. This inference, though, relies on a kind of stationarity assumption, one worth further scrutiny.[35]

[34] In many applications this is not true. For example, if attempting to classify blog entries, the training sample will generally not be representative of future blog entries. As time passes, the content of blog entries shifts. Hand (2006) discusses problems that can arise from this type of sample selectivity bias. Hopkins and King (2010) follow Hand and suggest a possible correction to the resulting bias.

[35] As a simple illustration of the possibility of stationarity, consider the following example. Suppose the

Another question involves the difference between proposed and finalized rules. Ultimately, I wish to draw conclusions regarding finalized rules. The comments, however, respond to proposed rules. The design of this paper assumes that the differences between proposed and final rules is modest. Yet, of course, agencies sometimes adjust rules in response to comments. Clinton's revised reporting standards, discussed above, stands as an example of one such OSHA regulation. Determining the relationship between each proposed and final rule is outside the scope of this project.[36] Were measures of the differences between proposed and final rules available, it would be of interest to consider whether rules finalized in the midnight period exhibit less sensitivity to commenters' concerns.

It is also reasonable to question whether, following the canonical Priest-Klein model (1984), equilibrium adjustments in agency and commenter behavior confound our ability to draw inferences about *rules* on the basis of observed *commenting* behavior. Two suppositions must hold to motivate this concern. First, the election must differentially influence the president's supporters and opponents. For instance, opponents must become less likely to comment in response to proposed rules, while supporters remain unaffected by the election; or, similarly, supporters become more likely to comment, while opponents remain unaffected by the election. Second, the agency must adjust the content of rules in response to the differing behavior of supporters and opponents so that the mix of supporting and opposing comments remains unchanged by the election. In this sort of scenario, the underlying

distribution of citizen preferences is given by F. Further suppose that the cutpoint dividing those who prefer the rule to the status quo is represented by k, with all citizens to the right of k preferring the rule and all those to the left preferring the status quo, and that the cost of writing a comment discourages citizens in the interval $[k-t, k+t]$ from submitting a comment. In this case, the proportion of submitted comments that support the rule is $\frac{1-F(k+t)}{1-F(k+t)+F(k-t)}$. Now suppose that the agency instead proposes a rule with an implied cutpoint of $k+e$, such that $F(k+e) \geq F(k)$ (i.e., the proportion of citizens who support the rule decreases, at least weakly), and that t remains the same. In this case, the proportion of submitted comments that support the rule is $\frac{1-F(k+t+e)}{1-F(k+t+e)+F(k-t+e)}$. Because $F(k+t+e) \geq F(k+t)$ and $F(k-t+e) \geq F(k-t)$, the proportion of comments supporting this second rule is weakly less than the proportion of comments supporting the first rule. In other words, the proportion of submitted comments that support a rule is weakly increasing the the proportion of the population that support a rule.

[36] For an excellent study along these lines, see Cuellar (2005), who provides a in-depth study of a small number of rules.

popularity of rules plausibly differs in the midnight period, despite no observed changes in commenting behavior. Systematically assessing this possibility is beyond the scope of this essay. The literature casts little light on the plausibility of either of these necessary conditions, and I leave the matter to future research.

Despite these limitations, tentative conclusions emerge from this study. Most importantly, I find little evidence that midnight rules represent the blight to our democracy suggested by many observers. It is true that a few midnight rules attract many opposing comments. But much the same can be said for many other rules passed during a president's administration. Presidents do not appear to use the midnight period of unaccountability to smuggle unpopular rules into law. Instead, the results suggest that first term presidents, perhaps fearing the electorate, tend to pass relatively popular rules during the end of their administrations.

6 Technical Appendix

The objective of this methodology is to classify comments according to whether they express positive or negative views toward the rule in question. The approach I develop is based on the common naive Bayes classification algorithm. This algorithm posits the following two-step generative process: first, a document is assigned a class, c_k for $k \in \{p, n\}$; second, for each word position in the document, we pick a word independently according to a multinomial distribution parameterized by θ_k. The "naive" part of the algorithm is that each word is sampled independently. Words, of course, often exhibit strong dependence, so the assumption is unrealistic. However, the assumption of conditional independence greatly reduces the number of parameters that must be estimated, and the algorithm has shown strong classification performance across countless applications (Pang and Lee 2008).

6.1 Statistical Model

In most applications, researchers want to estimate,

$$p(c_k|\mathbf{w}) \tag{3}$$

where c_k is a candidate classification, and $\mathbf{w}$ is the vector of words in the document. Researchers, that is, want to estimate the probability that a document is of class c_k given that it contains words $\mathbf{w}$. It is possible to estimate this quantity using Bayes rule. In particular,

$$p(c_k|\mathbf{w}) = \frac{p(c_k)p(\mathbf{w}|c_k)}{p(\mathbf{w})} \tag{4}$$

Furthermore, under the naive but useful assumption of conditional independence,

$$p(\mathbf{w}|c_k) = \prod_{i=1}^{W} p(w_i|c_k), \tag{5}$$

where W is the number of words appearing in the document. Using this, it is possible to determine which class is most likely. In particular,

$$p(c_k|\mathbf{w}) \propto p(c_k) \prod_{i=1}^{W} p(w_i|c_k). \tag{6}$$

The algorithm simply assigns to each document the most likely class.

6.2 Naive Bayes Application

In practice, the naive Bayes algorithm involves two central steps. First, the model described above is "trained" on a small set of documents with hand-coded classification assignments. The estimates $\hat{p}(w_i|c_k)$ and $\hat{p}(c_k)$ derive from this training phase of the analysis. When possible, researchers randomly sample from the population of documents. They then estimate the parameters under the standard maximum likelihood approach, often smoothing the estimates to avoid estimating zero probability words.[37] Second, using $\hat{p}(w_i|c_k)$ and $\hat{p}(c_k)$, we calculate the log relative risk for *unlabeled* documents, as described above.

Consider the implementation of the training phase in detail. I first draw a random sample of 200 comments from rules passed by Democratic presidents. After sampling the comments, I discard any comment that (a) contains garbled or unusable text or (b) does not express a clear view on the proposed regulation—a total of 18 out of 200 in the sample. I then code each comment as (a) supporting the rule, (b) opposing the rule, or (c) requesting an extension of the comment period. I add this third, procedural category because a non-trivial minority of "comments" in fact do little more than beseech the agency for additional time. I want to exclude these comments from the central analysis.

I then process the text of these comments for analysis. I first tokenize the text of

[37] Because of spareness inherent to the natural language context, the unlabeled documents will always contain some words not present in the training sample. Such words will have an estimated zero probability using the maximum likelihood approach. Zero probabilities create serious problems for the algorithm because $p(\mathbf{w}|c_k)$ goes to zero and becomes uninformative.

each comment, breaking down a comment into a series of individual words. I then reverse the polarity of any word that follows within three words of "not" or a contraction involving "not" by appending "NOT" to the word in question. Thus, the word "fair" in the phrase "this regulation is not fair," becomes "NOTfair."[38] After removing stop words such as "the" as well as any word that appears in less than one percent or more than 75 percent of the documents, I assemble a word-document matrix. I drop infrequent words to guard against over-fitting; I drop highly frequent words because they likely contain little information about the class of the documents. Element a_{wd} of this matrix represents the number of times that word w appears in document d. This word-document matrix consists of 173 documents and 5,776 words. I use this matrix to estimate $\hat{p}(w_i|c_k)$ and $\hat{p}(c_k)$ for rules passed by Democratic presidents for each of the three categories.

I estimate the parameters using the leave-one-out cross validation technique.[39] This procedure involves first removing one document from the word-document matrix; then estimating the word parameters with the N-1 documents; then "testing" the success of the estimated parameters of predicting the class of the document which was earlier removed. Notice that the words from this N-1th document did not contribute to the parameters used to predict its class. This is then repeated N times. This technique produces an out of sample prediction for each of the hand-coded documents, and also produces N different parameter estimates. I later use the mean of these parameter estimates to evaluate the class of the documents that I did not hand code.

I report the errors in predictions from this validation exercise for Democrats in table 3. The overall classification success rate is 84 percent. The model, moreover, appears to do well at predicting each class: it correctly predicts 85 percent of the comments hand coded as opposing the regulation; 73 percent of the comments hand coded as supporting the

[38] Although this is an extremely primitive approach to handling negation, most sophisticated approaches yield only modest improvements in performance (Pang and Lee 2008).

[39] This is a degenerate form of K-Fold cross validation.

regulation; and 96 percent of the comments hand coded as procedural in nature. In addition, the proportions predicted classes approximate the proportions of hand-coded classes: based on hand-coding of the random sample, 64 percent of comments oppose the regulation; 22 percent support; and 14 percent comment on a procedural issue. The corresponding figures for the predicted classes follow: 59 percent oppose the regulation; 19 percent support; and 22 percent comment on a procedural issue.

[Table 3 About Here]

I repeat this exercise for rules passed by Republican presidents, drawing 200 comments randomly, hand coding the comments, processing the text, and estimating the parameters of interest through a leave-one-out approach. I estimate separate parameters for Republican and Democratic presidents to acknowledge the possibility that, in the view of many, Republicans pass principally deregulatory measures whereas Democrats pass principally regulatory measures. If so, the word "safety," for instance, may have a different polarity in Republican and Democratic regimes. Plausibly, the word is used to support OSHA's actions under Democratic regimes and to oppose OSHA's actions under Republican regimes.

The validation results for Republicans appear in table 4. There, we see that the model performs somewhat less well for Republicans. Overall, the classification success rate is 73 percent. Within each class, the classification success rate is also lower: 83 percent for opposing comments; 48 percent for supporting comments; and 0 percent for procedural comments. It is clear that measurement error afflicts these results. The level of bias in the estimates, however, appears to remain minimal. The predicted proportions of classes closely follow the hand-coded proportions of classes. The randomly selected hand-coded comments indicate that 76 percent of comments oppose the regulation; 20 percent support; and 4 percent comment on a procedural issue. Based on the predictions, 77 percent of comments oppose the regulation; 23 percent support; and less than 1 percent comment on a procedural

issue.

[Table 4 About Here]

After estimating the relevant parameters for Republican and Democratic rules, I estimate $p(c_k|\mathbf{w})$ for each comment in the dataset. I follow the statistical model identified above in a straightforward way.

References

Beerman, Jack M. 2003. "Presidential Power in Transitions." Boston University Law Review, 83.

Brito, Jerry and Pattrick McLaughlin. 2008. "OIRA at Midnight." Regulation (Fall).

Cochran, Jay III. 2001. "The Cinderella Constraint: Why Regulations Increase Significantly During Post-Election Quarters." Typescript, George Mason University.

Cohen, Steven. 2009. "Prepared Statement: Midnight Rulemaking Shedding Some Light." Hearing Before the Subcommittee on Commercial and Administrative Law of the Committee on the Judiciary, House of Representatives, No. 111-02.

Cuellar, Mariano-Florentino. 2005. "Rethinking Regulatory Democracy," Administrative Law Review, 57: 411.

de Rugy, Veronique. 2009. "Prepared Statement: Midnight Rulemaking Shedding Some Light." Hearing Before the Subcommittee on Commercial and Administrative Law of the Committee on the Judiciary, House of Representatives, No. 111-02.

Dudley, Susan E. 2001. "Reversing Midnight Regulations." Regulation (Spring).

Earth Justice. 2009. "Prepared Statement: Midnight Rulemaking Shedding Some Light." Hearing Before the Subcommittee on Commercial and Administrative Law of the Committee on the Judiciary, House of Representatives, No. 111-02.

Hand, David J. 2006. "Classifier Technology and the Illusion of Progress." *Statistical Science* 21(1): 114.

Havemann, Judith. 1989. "Popular Policies Bloom As Election Approaches; Less Palatable Initiatives Appear Dormant." *Washington Post*, October 28, 1988, A6.

Hopkins, Daniel and Gary King. 2010. "Extracting Systematic Social Science Meaning from Text." *American Journal of Political Science* 54(1):229-247.

Mason, David. 2009. "Prepared Statement: Midnight Rulemaking Shedding Some Light." Hearing Before the Subcommittee on Commercial and Administrative Law of the Committee on the Judiciary, House of Representatives, No. 111-02.

Mendelson, Nina A. 2003. "Agency Burrowing: Entrenching Policies and Personnel Before a New President Arrives." New York University Law Review, 78.

Morriss, Andrew P. Roger E. Meiners and Andrew Dorchak. 2003. "Between a Hard Rock and a Hard Place: Politics, Midnight Regulations and Mining." Administrative Law Review (55).

O'Connell, Anne Joseph. 2008. "Political Cycles of Rulemaking: An Empirical Portrait of the Modern Administrative State." Virginia Law Review, 94 (889).

O'Connell, Anne Joseph. 2011. "Agency Rulemaking and Political Transitions." Northwestern Law Review, 105(2).

Pang, Bo and Lillian Lee. 2008. *Opinion Mining and Sentiment Analysis.* Foundation Trends in Information Retrieval 2(1-2).

Rhinehart, Lynn. 2009. "Prepared Statement: Midnight Rulemaking Shedding Some Light." Hearing Before the Subcommittee on Commercial and Administrative Law of the Committee on the Judiciary, House of Representatives, No. 111-02.

Seidenfeld, Mark. 1991. "A Civic Republican Justification for the Bureaucratic State." Harvard Law Review, 105:1511.

Figure 1: Midnight Rulemaking: An Agency-Level Assessment

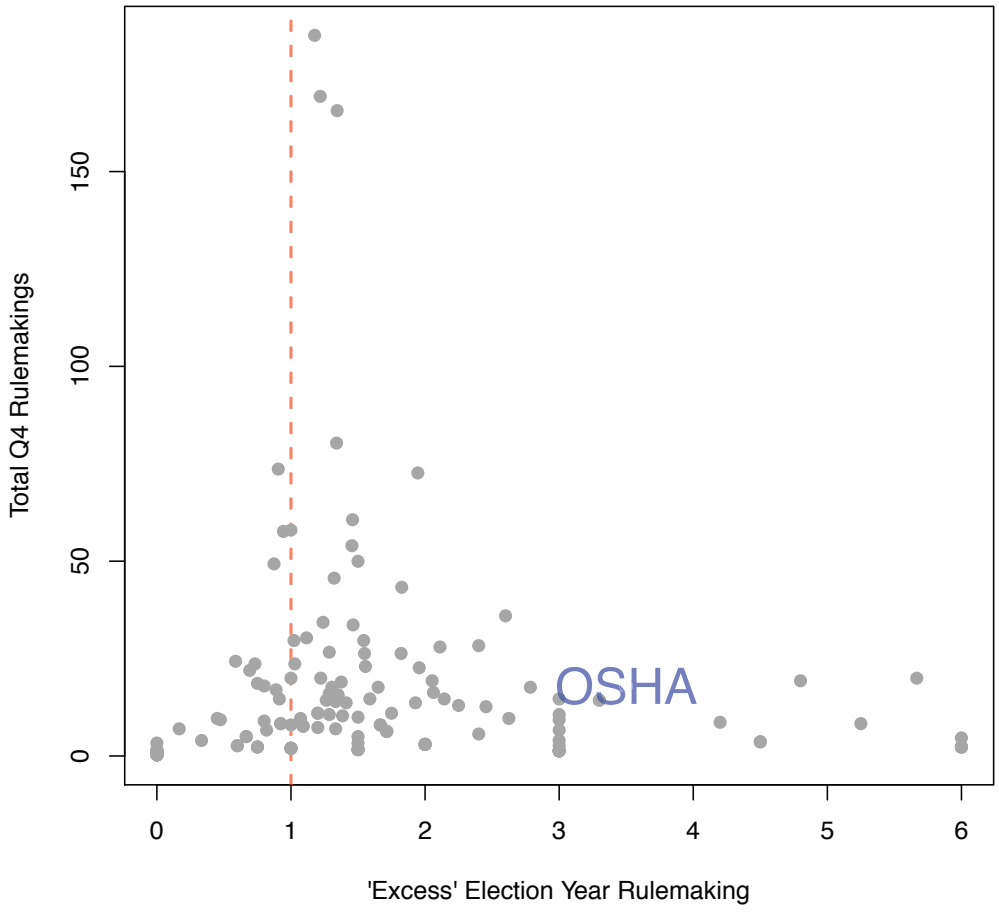

Figure 2: Distribution of Comments to OSHA NPRMs: 1983-2010

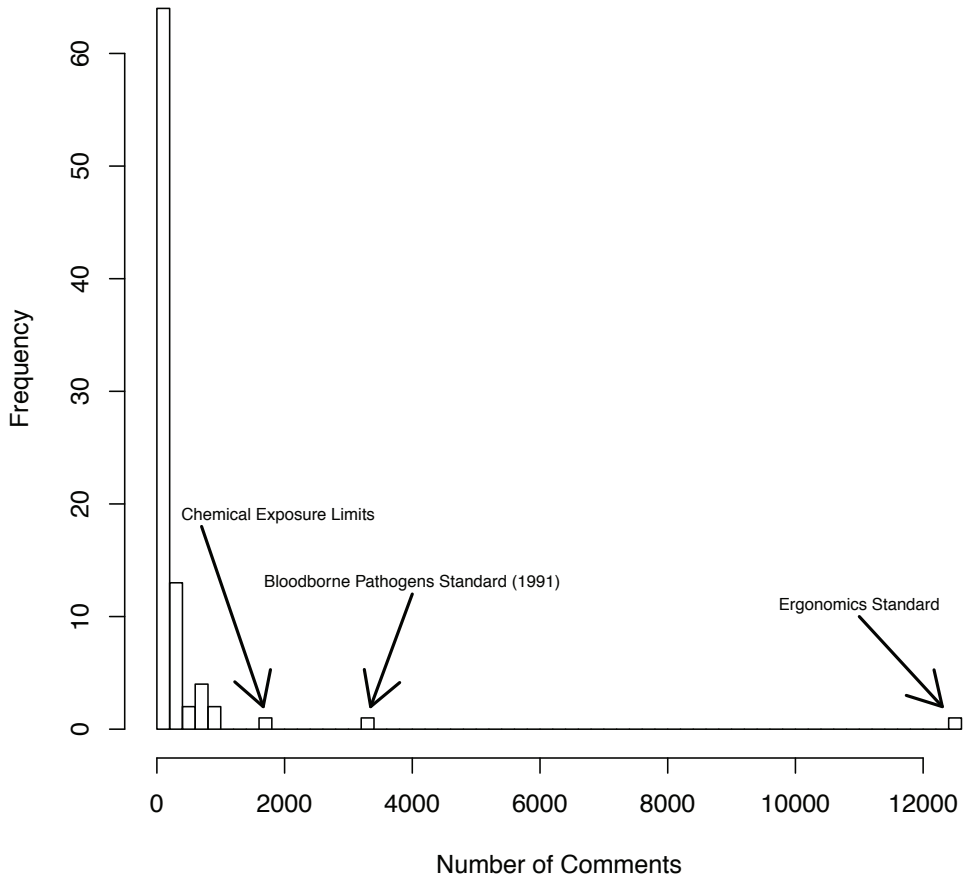

Figure 3: Supporting Comments by Presidential Administration: A Rule-Level Assessment

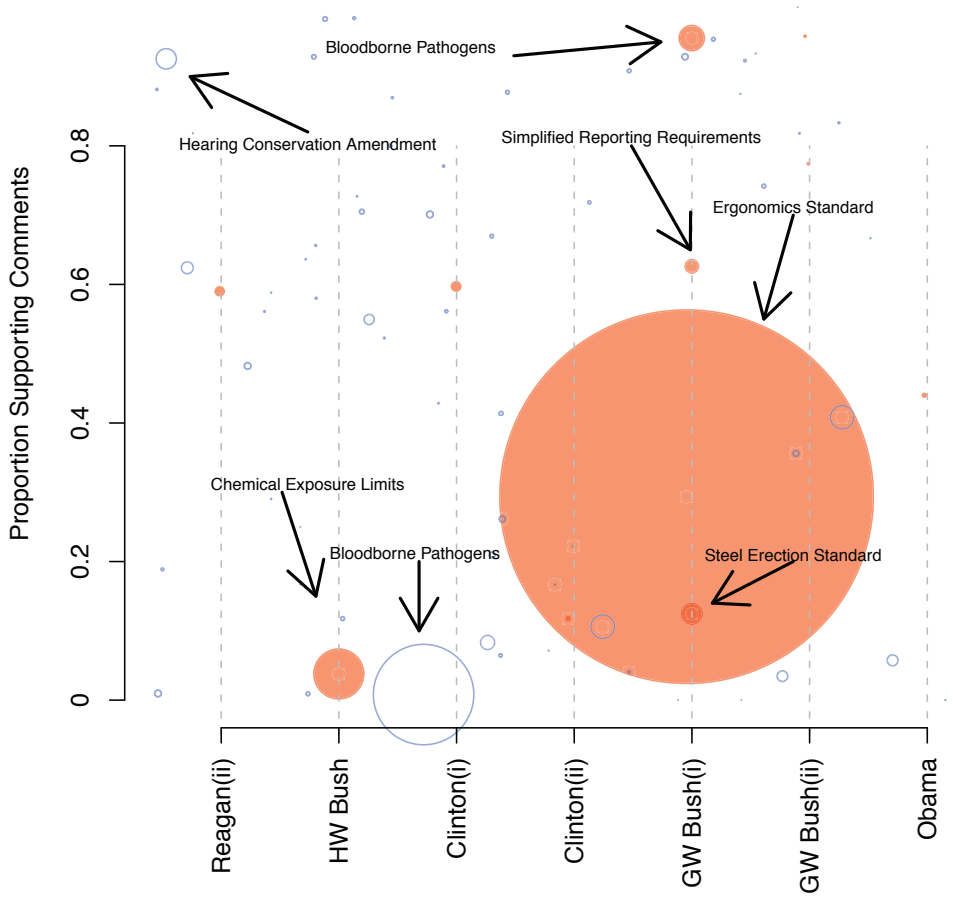

Table 1: Number of Comments Submitted for OSHA Rules

	Model 1	Model 2	Model 3	Model 4
Constant	4.58	4.81	4.58	4.81
	(0.42)	(0.39)	(0.4)	(0.38)
Last Term	0.21	–	0.21	–
	(0.38)		(0.37)	
Second Term	–	-0.29	–	-0.29
		(0.39)		(0.38)
Post-Election Quarter	-0.1	-0.2	-0.1	-0.2
	(0.86)	(0.8)	(0.83)	(0.77)
Post-Election Quarter X Last Term	0.8	–	0.32	–
	(0.97)		(0.97)	
Post-Election Quarter X Second Term	–	1.06	–	0.5
		(0.95)		(0.94)
Quarter Fixed Effects	Yes	Yes	Yes	Yes
R^2	0.12	0.12	0.1	0.1
N	88	88	87	87

Table 2: Commenters' Support for Midnight Rules

	Model 1	Model 2	Model 3	Model 4
Constant	0.14	0.01	0.14	0.01
	(0.7)	(0.56)	(0.7)	(0.56)
Last Term	-0.08	–	-0.08	–
	(0.63)		(0.63)	
Second Term	–	1	–	1
		(0.65)		(0.65)
Post-Election Quarter	2.03	2.64	2.03	2.64
	(1.02)	(0.82)	(1.02)	(0.82)
Quarter (2)	-0.13	-0.27	-0.13	-0.27
	(0.73)	(0.72)	(0.73)	(0.72)
Quarter (3)	0.56	0.49	0.56	0.49
	(0.61)	(0.64)	(0.61)	(0.64)
Quarter (4)	-2.21	-2.51	-2.21	-2.51
	(0.82)	(0.88)	(0.82)	(0.88)
Last Term X Post-Election Quarter	-0.65	–	-0.39	–
	(0.82)		(1.12)	
Second Term X Post-Election Quarter	–	-1.94	–	-1.74
		(0.75)		(1.13)
$PseudoR^2$	0.15	0.17	0.24	0.27
N	20688	20688	12402	12402

Table 3: Validation: Democrats

		Hand-Coding		
		Opposing	Supporting	Procedural
Prediction	Opposing	99	7	1
	Supporting	8	27	0
	Procedural	10	3	27

Table 4: Validation: Republicans

		Hand-Coding		
		Opposing	Supporting	Procedural
Prediction	Opposing	105	17	5
	Supporting	21	16	1
	Procedural	1	0	0