Congress as Manager: Oversight Hearings and Agency Morale[*]

John D. Marvel
&
Robert J. McGrath[†]

School of Policy, Government, and International Affairs
George Mason University

October 26, 2015

[†]Corresponding author: `rmcgrat2@gmu.edu`

Abstract

"Congress as Manager: Oversight Hearings and Agency Morale"

Federal agencies perform many important tasks, from guarding against terrorist plots to mailing social security checks. A key question is whether Congress can effectively manage such a large and influential bureaucracy. We argue that Congress, in using oversight to ensure agency responsiveness to legislative preferences, risks harming agency morale, which could have negative long-run effects on performance and the implementation of public policy. More specifically, we argue that oversight's effects on agency morale are conditional on whether oversight is adversarial or friendly. We assess our claims using a novel dataset of the frequency and tone of hearings in which federal agencies are called to testify before Congress from 1999-2011 and merge it with data on agency autonomy and job satisfaction. Our findings suggest that agency morale is sensitive to congressional oversight attention and thus speak to questions regarding democratic accountability, congressional policymaking, and the implementation of public policy.

The agencies that comprise the federal government's executive branch do many things: they protect the environment, guard against foreign and domestic threats to national security, mail millions of social security checks each month, and perform many additional functions, some more glamorous than others. Given that the federal bureaucracy is a non-electoral institution, Congress is charged with overseeing its execution of these tasks. In particular, congressional committees monitor the bureaucracy through oversight hearings, often attempting to increase agency responsiveness to congressional policy preferences. Yet, scholarship has paid scant attention to the possibility that such oversight may have significant managerial consequences. In particular, theory suggests that oversight may, at least conditionally, negatively affect agency morale, particularly as reflected in agencies' collective senses of autonomy and job satisfaction. That is, an empirical, as well as a theoretical, trade-off may exist between political responsiveness and agency autonomy. We assess this possibility, examining the link between oversight and survey-based measures of morale, and find that congressional oversight, when it is adversarial in tone, can indeed have negative consequences for the functioning of bureaucracy. Yet, we also find that more "friendly" congressional attention can actually improve agency morale.

Our research speaks to persistent questions concerning the correct balance between politics and administration. The impact of politics on policy implementation has been the subject of longstanding scholarly debate, particularly within public administration (Waldo, 1948). Echoing arguments from the Progressive Era (Wilson, 1887), contemporary government reform movements such as the New Public Management advance the argument that politics interferes with agencies' fulfillment of their duties (see, e.g., Light, 2006). On the other hand, some have argued that politics and administration are inextricably intertwined, and that attempts to neatly separate them are hopeless and naïve (Rosenbloom, 1993; Waldo, 1948). By examining whether and under which conditions congressional oversight is related to agency morale, we aim to make an empirical contribution to this debate. We also contribute to the burgeoning public management literature on organizational performance. While the notion that political actors influence agencies is central to this literature's prominent theories (e.g., O'Toole and Meier, 1999; Rainey and Steinbauer, 1999), very little empirical research addresses whether there is in fact a relationship between the activities of these actors and agency morale, a variable that

theory and empirical evidence suggest will affect performance.[1]

We seek to synthesize and contribute to two distinct, but related, fields of research. Studies in political science have traditionally been concerned with questions of congressional monitoring and control of bureaucratic outputs and policy (McCubbins and Schwartz, 1984; Bendor, Taylor and Van Gaalen, 1985; Moe, 1989; Ferejohn and Shipan, 1990; Wood and Bohte, 2004; Balla, 1998). This is a crucially important issue for a democratic system of government. If duly elected political actors must rely on unelected bureaucrats to implement policy programs, control and responsiveness are normative imperatives (see, e.g., Finer, 1941). Research in public administration, on the other hand, has focused on the roles of professional norms and ethics in constraining bureaucratic policy implementation. In this view, public agencies should be subject to internal constraints (the so-called "inner check"), yet minimally encumbered by the intrusion of political actors (Friedrich, 1940). These two perspectives—democratic control of public agencies enforced by external political actors vs. professional democratic norms developed through internal discipline[2]—are often seen as driving contemporary normative debates across political science and public administration. We seek to test the implicit claim of the latter perspective: that political intervention can serve to limit agency discretion in deleterious and counterproductive ways (Behn, 1995). In particular, we utilize novel data on oversight hearings directed at particular agencies from 1999 through 2011, and assess whether increases in oversight attention affect agencies' collective feelings of autonomy and job satisfaction.

We begin by discussing agency morale and its importance. Since theory suggests that morale is positively associated with agency performance, as well as with work attitudes and work behaviors that feed into performance, we see it as particularly worthy of empirical attention. We then argue that congressional oversight is likely to affect agency morale, but that the direction of these effects should depend on the content and tone of the oversight. Next, we describe our data and empirical strategy, paying particular attention to our measures of oversight and agency morale. After presenting our results, we close with a discussion of our findings' practical and theoretical implications. The main takeaway is that oversight seems to negatively affect morale, but only when the oversight is adversarial and negative in tone. In fact, we provide evidence that so-called "advocacy" oversight, on the other hand, can actually bolster agency

morale (Aberbach, 1990).

The Importance of Agency Morale

Part of the job of any manager, in any organizational setting, is to motivate employees. Doing so involves cultivating employee work attitudes (e.g, job satisfaction, organizational commitment) and behaviors (e.g., arriving to work on time, aiding coworkers) that are thought to be associated with individual- and organizational-level performance. In exercising its oversight function, however, Congress is not necessarily interested in doing these things.[3] Instead, it is primarily interested in ensuring federal agencies' responsiveness to legislative preferences.[4] But in pursuing responsiveness, Congress can unwittingly harm agency morale. Before fully developing this argument below, we define the empirical focus of our study—agency morale—and discuss its importance for agency performance.

We use the term "agency morale" to denote agency employees' collective feelings of autonomy and job satisfaction. Theory and evidence from the organizational behavior literature suggest that, at the individual level, both of these traits are positively related to job performance. In a meta-analysis of 312 independent samples, Judge et al. (2001) find a correlation between job satisfaction and job performance of 0.30. Similarly, in a meta-analysis of 101 independent samples, Spector (1986) finds a correlation between autonomy and job performance of 0.26. In fact, these correlations likely underestimate the total impact of job satisfaction and autonomy on performance, given that both are associated with numerous other work attitudes and behaviors that are themselves related to performance. These include, for instance, organizational commitment, role conflict, role ambiguity, emotional distress, absenteeism, turnover intention, and actual turnover (Mathieu and Zajac, 1990; Meyer et al., 2002; Riketta, 2002; Tett and Meyer, 1993; Spector, 1986).

Theories of public sector organizational effectiveness and political control pay special attention to autonomy. The former typically emphasize autonomy's salutary operational qualities: It allows agencies to use their expertise to solve pressing implementation problems, make and execute decisions quickly, and pursue their missions in an administratively rational manner (see, e.g., Wilson, 1989; Wolf, 1993; Meier, 1997; Rainey and Steinbauer, 1999; Brewer and

Selden, 2000). These theories also assume that autonomy has motivational benefits at the employee level. Individuals—particularly individuals with high levels of formal education and professional training—value autonomy and work harder when it is given to them (see, e.g., Gagné and Deci, 2005). By contrast, theories of political control tend to view autonomy as necessary—bureaucracies have expertise that political actors lack, and so delegations of authority are sometimes unavoidable—but potentially problematic, given that bureaucracies are non-electoral institutions. Yet, even political theories note the importance of autonomy for organizational performance. Gailmard and Patty (2007, 2012), for example, argue that congressional principals, who generally prefer informed to uninformed policymaking, proactively grant autonomy and policymaking discretion to bureaucratic agents in order to incentivize investments in expertise. Whatever their differences, both schools tend to agree that autonomy is systematically associated with organizational performance and the development of policy expertise. Consequently, we believe it is important to examine whether congressional oversight is associated with agency autonomy.

Congressional Oversight and its Managerial Consequences

We expect that congressional oversight will be negatively associated with autonomy and job satisfaction when such oversight is primarily meant to monitor and control the bureaucracy for political reasons, rather than to aid it in the performance of agency duties (Weingast and Moran, 1983; Ferejohn and Shipan, 1990; Shipan, 2004). Congress is often unlike the manager or firm owner described in standard economic accounts of principal-agent theory. In these accounts, it is usually assumed that the principal is concerned with securing some outcome and is, moreover, happy to let the agent choose whatever means or behaviors best serve that end (for a review, see Eisenhardt (1989)). The congressional impulse to control, however, often seeks to dictate the bureaucracy's choice of means. This impulse is intensified in our separation of powers system, where Congress often competes with the president for agency influence (Shapiro, 1994; Whitford, 2005). Below, we identify three particular mechanisms through which congressional oversight can harm agency morale and conclude by arguing that oversight's relationship with morale is ultimately conditional on whether it is adversarial or friendly.

Mechanism I: Micromanagement

Consistent with the predilection of Congress to be interested in control rather than performance, scholars have long noted that its oversight relationship with the federal bureaucracy has been characterized by micromanagement, or "intervention by Congress in administrative details" (Gilmour and Halley, 1994, p. 10). As early as 1885, Woodrow Wilson complained that Congress "has entered more and more into the details of the administration, until it has virtually taken into its own hands all the substantial powers of government" (Wilson 1896, cited in Beermann 2006). Similarly, James Q. Wilson (1989) writes that "Congress is commonly criticized for 'micromanaging' government agencies; it does and it always has" (p. 241). More recently, Behn (1995) identifies political micromanagement as one of public administration's most pressing problems and elucidates how it hampers agency performance: "The legislative branch is, for some reason, unhappy with the way an executive-branch agency is performing; so the legislators impose some rules on the agency... These new rules prevent, or at least constrain, the agency from doing what the legislature dislikes. Unfortunately, these rules also constrain the agency from producing the results for which it is responsible" (p. 316).

There is reason to believe that oversight has become increasingly driven by this impulse to micromanage and constrain bureaucratic discretion. Summarizing a series of ten case studies on oversight, Gilmour and Halley (1994) conclude,

> The cases show a 'congressional co-manager' intervening directly in the details of policy development and management rather than enacting vague, wide-ranging, sweeping statutes to change fundamental policy directions...
>
> Gone almost without a trace is the post-New Deal Congress that optimistically delegated broad-scale public problems and policy questions for solution and resolution by the executive branch. Much diminished as well is an executive branch relied upon by Congress for neutral competence and specialized expertise. Instead, the story... is one of the retrieval of executive discretion and the highly specific redefinition—by Congress—of prior delegations of authority (p. 335-336).

In the same vein, Aberbach (1990) shows that the average number of pages per statute enacted by Congress rose sharply between the 80th (1947-1948) and 103rd (1993-1994) sessions of

Congress, indicating an increased command-and-control orientation in legislative-bureaucratic relations. More recently, Balla and Deering (2013) code a sample of all Congressional hearings that occurred during the 96th (1979-1981), 100th (1987-1989), 104th (1995-1997), and 108th (2003-2004) sessions of Congress. They find that most hearings—over 80% in each session—are police patrols, as opposed to fire alarms, indicating that Congress has an abiding interest in monitoring what the federal bureaucracy is doing and in how it is doing it.

As a recent illustration of this mechanism, scholarly research and witness testimony from administrators from the Centers for Medicare and Medicaid Services (CMS) attest that members of Congress are keen to micromanage policies governing provider payment (Pham, Ginsburg and Verdier, 2009). The data that we compile below support these claims, indicating that there were no less than 377 oversight hearings from 1999-2013 where members of Congress expressed their views on this issue, often disagreeing with CMS policies. Representative of these interactions is a May 15, 2007 hearing of the House Committee on Ways and Means's Subcommittee on Health, under the direction of subcommittee chairman Pete Stark (D-CA). In this hearing, titled "Payments to Certain Medicaid Fee-for-Service Providers," Stark belies his intent to intervene in CMS regulations, upon "hearing from industry that many of these regulations, particularly the inpatient hospital regulations, are nothing but backdoor attempts to circumvent Congress and cut spending." And, despite being "loathe to intervene in the nuts and bolts of regulations," and generally thinking "that level of detail is best left to the experts like Mr. Kuhn [Herb Kuhn, then Acting Deputy Administrator, Centers for Medicare and Medicaid Services]," Congressman Stark felt impelled to give pages of suggestions on how CMS should direct fee-for-service payments to providers. Such intricate congressional involvement in agency decisions is common in our hearings data and an indication that more oversight often means more direct congressional involvement in policy implementation.

Micromanagement is fundamentally a psychological mechanism. It is harmful to agency morale because it politicizes employees' work and, in doing so, undermines employees' ability to *experience meaning* while performing their jobs (Hackman and Oldham, 1976; Ryan and Deci, 2000; Barrick, Mount and Li, 2013). A large body of research on "public service motivation" suggests that for many individuals who are employed in the public sector, the experience of

meaning flows from doing work that is thought to advance the public good (see, e.g., Perry and Wise, 1990; Houston, 2009). At its core, public service motivation is an "other-regarding" orientation; it entails a broad-based concern for the wellbeing of one's fellow citizens, as opposed to a more narrow concern for particularistic interests (Ward, 2014). Micromanagement can hurt agency morale by appropriating an agency's collective work effort for partisan purposes and, in doing so, stripping that effort of its politically-neutral public service meaning. Just as a generic manager's use of monetary rewards to incentivize employee effort can "crowd out," or displace, an employee's intrinsic motivation for doing a job well (Frey and Oberholzer-Gee, 1997), congressional micromanagement can crowd out agency employees' public service motivation by signaling to employees that their work is ultimately partisan in nature.

We view congressional micromanagement as a variable that shapes an agency's shared understandings of, and collective beliefs about, the purpose of its core work. In other words, micromanagement affects agency morale via its influence an agency *culture*. In this view, an employee need not be directly exposed to congressional oversight for the micromanagement mechanism to be operative; the employee need only be exposed to the agency's prevailing cultural beliefs. In agencies that are subject to a significant amount of politically-motivated oversight, we would expect a "politicized" culture to obtain. In these agencies, employees would understand their work to be primarily partisan, and would be demoralized by this understanding. By contrast, in agencies subject to little political oversight, we would expect a relatively "apolitical" culture to obtain. In these agencies, employees would understand their work to be primarily in service of the public good, and would be heartened by this understanding.

Mechanism II: Short-Term, Recurring Opportunity Costs

Besides this micromanagement mechanism, there are at least two more possible avenues by which oversight may harm agency morale. First, preparing for and participating in oversight hearings, especially high profile ones, levies opportunity costs on agency employees. Rather than focusing on, say, fulfilling their missions, or competently implementing legislative policy, agency employees must respond to the priorities of a committee holding an oversight hearing. We call these opportunity costs short-term to differentiate them from the more fundamental

(and psychological) crowding-out of experienced meaning that congressional micromanagement entails.

Short-term opportunity costs likely fall most squarely on agency managers, especially those who are called to testify in an oversight hearing. These employees must, quite literally, put down whatever they are working on to prepare for and attend a hearing. A recent journalistic account of declining morale among high-level agency managers at the Department of Homeland Security supports this line of reasoning. As the article notes, "Many former and current officials said the most burdensome part of working for DHS is the demands of congressional oversight. More than 90 committees and subcommittees have some jurisdiction over DHS, nearly three times the number that oversee the Defense Department. Preparing for the blizzard of hearings and briefings, officials say, leaves them less time to do their jobs" (Markon, Nakashima and Crites, 2014).

While we assume that oversight hearings will produce higher opportunity costs for managerial than non-managerial employees, it is plausible that at least some of these costs will impinge on the daily work routines of an agency's middle and lower-level employees. Managers will likely need help preparing for and responding to hearings, and it is reasonable to expect that they will delegate some of their hearings-related work to non-managers. Still, in terms of their impact on the felt autonomy and job satisfaction of non-managerial employees, we view short-term opportunity costs as secondary to micromanagement. Whereas micromanagement undermines the very meaning of work done by agency employees, opportunity costs are merely temporary (albeit perhaps frequent) disruptions to an employee's work routine.[5]

Mechanism III: Public Shaming

Finally, it is reasonable to assume that negative congressional attention whose aim is to publicly embarrass high-level agency managers would be demoralizing to the agency as a whole. A recent example of this involves the General Services Administration and the attention it received in 2012, after stories of wasteful spending at its Western Regions Conference surfaced in the media. The aftermath included many high profile oversight hearings and numerous internal reports that sought to assign responsibility for the agency's actions. Since "fraud, waste, and

abuse" are anathema to both parties, Democrats as well as Republicans relentlessly attacked the GSA in hearings. In this instance, Congress can be seen to have had a genuine interest in improving GSA performance into the future. In other words, this was an ideal opportunity for Congress to act as a genuine performance manager; that is, to take a sincere interest in remedying whatever underlying organizational problems (e.g., issues with organizational culture, ineffective internal accountability structures, etc.) may have contributed to the GSA scandal. Instead, Congress appeared to be more interested in obtaining whatever political mileage it could by publicly scolding top-level GSA employees.

Of course, agency managers *should* be called to account for agency misbehavior. Nevertheless, it is important to emphasize that public shaming is not viewed as a constructive managerial practice in the organizational behavior and public management literatures. In fact, recent research suggests that "abusive supervision," which includes "nonphysical actions such as angry outbursts, public ridiculing, taking credit for subordinates' successes, and scapegoating subordinates," is negatively associated with job satisfaction, turnover intention, and additional markers of employee morale (Tepper, 2007, 2000; Aryee et al., 2007).[6] Importantly, research in this vein also indicates that the abusive supervision endured by an organization's higher-level employees "trickles down" to its lower-level employees (Aryee et al., 2007). In this view, the supervisory treatment that high-level employees receive influences the manner in which they treat their own subordinates. Notwithstanding these potential trickle-down effects, we assume that public shaming is more strongly associated with managerial employees' morale than non-managerial employees' morale. Whereas the high-level managerial employees who attend hearings will endure any shaming attempts firsthand, non-managerial employees' exposure will be indirect.

"Advocacy" and the Conditional Effects of Oversight

Thus far we have discussed three mechanisms via which oversight hearings may negatively affect agency morale. These mechanisms would seem to operate across qualitatively different types of oversight hearing. Police patrol oversight, for example, is most likely to reflect Congress's desire to micromanage (Balla and Deering, 2013). These hearings also require dili-

gent agency preparation and are likely to command persistent short-term opportunity costs. Fire-alarm hearings (McCubbins and Schwartz, 1984) also require agency preparation, often on short notice, so we expect agencies to be burdened by high opportunity costs here as well. In addition, fire-alarms are more likely to trigger particularly adversarial hearings, thus activating the public shaming mechanism.[7] In fact, all of these mechanisms rely on the assumption that oversight hearings are contentious affairs.

Yet, existing work (Aberbach, 1990) cautions us against making the assumption that all hearings serve the same purpose. Aberbach (1990), drawing on survey evidence from committee members and their staffs, shows that much congressional oversight activity takes place in what he calls an "advocacy context." Aberbach stresses that there are two general types of committee oversight: adversarial hearings meant to score political points or forcibly change agency policy (through micromanagement, as discussed above); and advocacy hearings, where members of Congress defend "their" preferred programs and agencies by holding hearings and officially voicing praise and approval. This type of oversight is qualitatively different from that assumed in our theoretical discussion regarding the negative effects of hearings on agency morale. There is little reason to expect any of the three proposed mechanisms to drive down morale when committees are friendly towards agencies in hearings. In fact, we might even expect that advocacy hearings *increase* agency morale, as they publicly demonstrate agency accomplishments, and can serve to justify increased appropriations (Aberbach, 1990, chap. 8). In addition, when Congress's and the bureaucracy's goals are aligned and oversight is positive and advocacy-driven, it is conceivable that Congress might assume the salutary managerial role that is exalted in theories of public sector organizational effectiveness (Rainey and Steinbauer, 1999; O'Toole and Meier, 1999; Fernandez, 2005; Lee and Whitford, 2013).

We ultimately argue that the relationship between congressional oversight activity and agency morale is a conditional one. When oversight is politically-driven and adversarial, we expect it to harm agency morale, for the reasons discussed above. Yet, when oversight is more "friendly," agencies can benefit, both tangibly and intangibly, from congressional attention. Although agencies still have to prepare for these hearings, the outcomes of these preparations (potential praise and material rewards) can often outweigh the short-term opportunity costs of

hearing involvement. Thus, to the extent that oversight hearings are positive toward the target agency, we expect them to increase agency morale.

Data, Variables, and Methods

In order to assess the conditional relationship between congressional oversight and agency morale, we first create empirical measures of each. We focus exclusively on formal oversight hearings as, of the myriad forms of oversight,[8] these are the most straightforward to quantify and have been the focus of many empirical studies (Dodd and Schott, 1979; Aberbach, 1990; Ogul and Rockman, 1990; Smith, 2003; Balla and Deering, 2013; McGrath, 2013; MacDonald and McGrath, N.d. (forthcoming). Nevertheless, existing studies have not considered oversight as an agency-level demand-side variable, and have instead focused almost entirely on the supply-side of oversight. The few studies that consider oversight from an agency perspective have focused on small samples of agencies or hearings and have not documented the overall extent to which agencies are called to appear before Congress (see, e.g., Parnell, 1980; May, Workman and Jones, 2008; May, Sapotichne and Workman, 2009; May, Jochim and Sapotichne, 2011). Therefore, we develop a unique measure of oversight hearings directed at federal agencies as our primary independent variable.

Oversight Hearings Data

We collected data on oversight hearings from the Government Printing Office's Federal Digital System (`http://www.gpo.gov/fdsys/search/advanced/advsearchpage.action`).[9] The GPO began publishing a sizable number of hearing transcripts in 1997, so we begin our collection there.[10] The description of the GPO's hearings data indicates that committees sometimes take up to two years to publish hearings, so we attenuate our dataset to conclude at the end of 2011.[11] We collected the universe of hearings by searching the "Congressional Hearings" database with an empty keyword field and saved each full text transcript. Each transcript contains a list of witnesses called before Congress for the hearing, including their affiliation with federal agencies, when applicable. All told, we identified 17,572 hearings in these data.

We parsed the text of each individual hearing transcript to create witness data and then

narrowed the witnesses by whether or not they represented an agency. We consider a hearing to be directed at a particular agency only if the committee or subcommittee holding the hearing called a witness from that agency. There are often cases where there are no agency-affiliated witnesses for a given hearing and still others where an individual hearing applies to multiple, and sometimes many, agencies. Next, we attempted to identify hearings that were meant to conduct oversight and separate them from legislative hearings. As described in appendix A, we followed recent research (McGrath, 2013; MacDonald and McGrath, N.d. (forthcoming) and filtered oversight hearings by searching the full text transcripts for keywords that might indicate oversight.[12] After filtering, we identified a total of 11,407 oversight hearings in our data.

Figure 1 goes here.

Once we identified agency witnesses and separated oversight from non-oversight hearings, we grouped hearings by agency and year. The agency-year dataset then has 1,053 observations—13 full years of data for 80 agencies and 2 agencies with fewer than 13 observations due to being created after 1999.[13] The agencies were grouped by the coding scheme for the 2012 Federal Human Capital Survey so as to allow us to match the hearings data to the agency morale data described below. Generally speaking, the data are grouped at the department level, including independent agencies and the Office of Management and Budget (part of the Executive Office of the President), with some departmental subunits included.[14]

Table A1 (appendix A) indicates each agency for which we have collected hearings data and gives descriptive statistics for such oversight activity. Figure 1 displays how the total number of oversight hearings committees held across the 82 coded agencies varies over time. The data cover a time period which was characterized by the full diversity of institutional and partisan configurations. Namely, we have been through unified government, divided government with a unified Congress, divided government with a divided Congress, Republican presidents, Democratic presidents, and changes in the partisan control of each chamber during this period. Figure 2 displays temporal changes in oversight hearings across the 15 cabinet-level departments, further demonstrating the variation that exists in these data. Additionally, figure 3 shows, via box plots, the distributions of oversight hearings for each department. While obvi-

12

ously crucial for testing how how oversight can affect agency morale, these data are inherently interesting in demonstrating the significant variation that exists in how often certain agencies are called to appear before Congress, and future research should model this variation as an outcome, as well as a determinant of agency characteristics (MacDonald and McGrath, N.d. (forthcoming).

Figure 2 goes here.

Figure 3 goes here.

Measuring Hearing Sentiment

As we argue above, the effects of oversight on morale should depend on the fundamental *tone* and purpose of the hearings. As such, we additionally analyze the content of each hearing to categorize it as either adversarial or advocacy-driven. Adversarial hearings reflect what most observers think of when they consider oversight. Here, members of congressional committees call agencies to task for poor performance, or simply for implementing policy inimical to the wishes of a committee. These hearings are often acerbic affairs, and are unpleasant experiences for agency employees called to testify. They additionally require agencies to prepare extensive reports and testimony to avoid public embarrassment. These are the hearings that we expect to negatively affect agency morale.

On the other hand, Aberbach (1990, p. 118) describes an alternative to adversarial hearings: "While one's first reaction to the word 'oversight' is that Congress is at odds with an agency or program targeted, committees sometimes use oversight because they want to defend 'their' program or agency against others who would do it harm." This brand of advocacy oversight has been largely overlooked by empirical studies, though there is evidence that this makes up a good part of Congress's oversight agenda, especially during unified government (see Aberbach (1990, chap.8) and MacDonald and McGrath (N.d. (forthcoming)). We do not expect such hearings to negatively affect agency morale; rather, we expect that when hearings are positive in tone, they will actually improve agency morale.

We thus seek to categorize congressional oversight as either adversarial or friendly, and we do so by measuring hearing *sentiment*. Specifically, we undertake computer-assisted sentiment

analyses of each hearing, following standard practice in the computer science literature and a growing trend in the social sciences.[15] Hearing transcripts follow a fairly standard format. They open with metadata about the hearing (those in attendance, the time and location of the meeting, a list of witnesses, etc.), and then invariably commence with the opening statements of the committee or subcommittee chair and other interested members of Congress. These opening statements are the primary source of our sentiment data, as they provide many instances where a member of Congress expresses sentiment towards an agency.

For each observation in the agency-hearing dataset described above and in appendix A, we calculated a *Targeted Sentiment* score which we use to measure how positive (positive values to 1) or negative (negative values to -1) each hearing is with respect to the agency at hand.[16] There is a good deal of variation in sentiment scores across the data, with a mean score of 0.068 and a standard deviation of 0.278 (empirical range: -0.901 to 0.925). As our data are organized at the agency-year level, we aggregate from individual hearings by taking the mean sentiment for each agency and year (*Hearings Sentiment*). We assess our conditional hypotheses below by interacting this overall measure of oversight sentiment with the total volume of oversight hearings conducted involving each agency in each year.

Measuring Agency Morale

Viewing agency morale as a set of characteristics best discerned from individual responses to surveys of federal employees, we adopt the approach of Bertelli et al. (2015) of measuring agency-level characteristics by aggregating these individual responses. This approach builds on earlier attempts to use individual employee attitudes to approximate unobservable agency characteristics,[17] and seeks to overcome some of the limitations of these types of data. In particular, Bertelli et al. (2015) provide a framework for aggregating survey responses in such a way as to put agency-level summaries on a common scale for cross-agency and over-time comparisons. Such an approach is key for our endeavor to test the effects of oversight activity on agency morale in a panel data setup. Having consecutive years of data on oversight and agency morale across agencies thus allows us to use a fixed effects design, isolating the within-agency effects of changes in oversight activity on self-reported agency characteristics.

Bertelli et al. (2015) begin by identifying the agency characteristics they wish to measure: autonomy, job satisfaction, and intrinsic motivation. They consider these characteristics to be latent attributes and use individual responses to particular questions from federal personnel surveys to measure these constructs using a dynamic Bayesian item-response model similar to the approach in Martin and Quinn (2002) (see also, Clinton, Jackman and Rivers, 2004; Bertelli and Grose, 2011; Clinton et al., 2012).[18]

Of these measured agency-level characteristics, we focus particularly on agency autonomy and job satisfaction as constructs which relate to agency "morale" as a meta-characteristic of interest. Bertelli et al. (2015), among other studies, do not necessarily equate autonomy with the possession of objectively large amounts of statutory administrative discretion (Epstein and O'Halloran, 1999; Huber and Shipan, 2002). Instead, autonomy refers to the extent to which bureaucrats *feel* in control of their own surroundings in performing their duties: a more subjective sense of discretion. The job satisfaction variable is what organizational behavior researchers typically call a "global" measure; that is, a measure of *overall* job satisfaction. Each of the three survey items that together constitute this measure encourage respondents to think in very broad terms about their jobs. One of the items asks, for instance, "Considering everything, how satisfied are you with your job?"[19]

Figure 4 goes here.

Figure 5 goes here.

Figure 4 displays the autonomy measures and the variation that exists in each across the cabinet departments, as figure 5 does for the measure of job satisfaction.[20]

Empirical Strategy

Having collected panel data[21] on levels of oversight and agency morale characteristics, with each measure varying considerably over time (again, see figures 2, 4, and 5), we turn now to identifying the most appropriate empirical design by which to assess the relationship between oversight and morale. We are primarily interested in the effects that changes in oversight might

have on agency morale over time. Ideally, we would like to tease out temporally causal relationships from confounded, spurious, or endogenous correlations and have chosen a design and model specifications that we believe will help us get there. In particular, we take advantage of our data structure to estimate fixed effects models, thus accounting for unobserved agency heterogeneity and isolating the effects of time-varying covariates on time-varying agency characteristics.

Yet, this design does not erase the potential for biased estimates; nor does it guarantee casual interpretations of these estimates. In particular, we are careful to measure and account for factors that might simultaneously cause increases in oversight activity *and* changes in autonomy and job satisfaction, respectively. Our primary explanatory variable, *Oversight Hearings*, varies both across and within agencies over time, and our research is designed to isolate the effects of within-agency across-time changes in oversight on expressed agency traits. Therefore, we limit our attention to control variables that similarly vary within agencies over time, as the fixed effects eliminate all sources of time-invariant agency heterogeneity, observed and unobservable.

News Sentiment and Other Controls

Perhaps most importantly, we control for the possibility that something—like an agency scandal of the sort described above with respect to the GSA— contributes both to the variation in *Oversight Hearings* and in the measures of agency morale. Agency scandals and aggregations of smaller issues related to poor agency performance invariably lead to "fire-alarm" oversight by congressional committees eager to show constituents how they can fix agency problems (McCubbins and Schwartz, 1984). Scandals and poor performance also generate negative media attention that presumably has deleterious effects on agency morale, independent of the potential effects of the hearings themselves. It is thus necessary to disentangle the effects of negative media attention from the effects of congressional oversight.[22]

To this end, we created a measure of media attention by collecting all stories published in the *Washington Post* that mention each agency in our dataset.[23] We grouped the stories by agency and year and calculated the total number of stories and pages of coverage. This approach is similar to recent attempts to measure mass media attention to federal agencies

(Lee, Rainey and Chun, 2009; Lee and Whitford, 2013), but we must also take into account the sentiment that these aggregated stories reflect towards agencies. Therefore, exactly as we did with the hearing transcripts, we measured the targeted sentiment of each news article in these data to create *News Sentiment* scores reflecting how positive (positive values to 1) or negative (negative values to -1) each piece of coverage is with respect to the agency at hand. We then calculated the sum of *News Sentiment* scores for each agency-year and use this as our measure, *Total Washington Post Sentiment*, capturing both the amount and direction of news coverage of the agencies in our data.

We also account for political attention to agencies, apart from the attention that oversight hearings themselves indicate. First, we separately include the volume of *Non-oversight Hearings* for each agency-year into our models. These are the hearings that we collected from the GPO that did *not* include the keywords we consider to indicate oversight.[24] Likewise, we recognize that agencies may be the recipients of other kinds of political attention that may affect employees' responses to survey questions. As in Lee and Whitford (2013), we operationalize a *Presidential Attention* variable, using the GPO's FDsys to search for mentions of each agency in the Public Papers of the Presidents of the United States.[25] Lee and Whitford (2013) argue specifically that presidential attention might signal that political resources (time and money) are available for agency policy priorities.

In addition to these measures of media and political attention, we include indicators for various regimes of political control. While we are mostly agnostic about the potential effects of these variables on changes in agency morale, we know that they are important determinants of congressional delegation to agencies in the first place (see, e.g., Kiewiet and McCubbins, 1991; Epstein and O'Halloran, 1996; Huber and Shipan, 2002; Volden, 2002) and of congressional incentives to hold hearings with or investigate agencies (see, e.g., Mayhew, 2005; Kriner and Schwartz, 2008; Parker and Dull, 2009; McGrath, 2013). These variables include an indicator for *Divided Government*, and one each for *Republican Control of Congress*, *Democratic President*, and *Presidential Transition Year*.

Notably, we do not include any time-invariant agency characteristics, as they would present identification issues in a fixed effects setup. This ultimately means that we cannot directly as-

sess which specific mechanisms are at play in generating the relationships that we find. While these mechanisms have distinct observable implications, these are found in agency-level characteristics and unmeasured characteristics of the hearings. For example, we argued above that public shaming can cascade from those managers who were involved in an oversight hearing to agency careerists. This mechanism might imply that such cascades should have larger impacts on agency morale in small, tight-knit agencies. Yet, agency size is largely time-invariant and is thus collinear with agency fixed effects.[26] Indeed, these agency fixed effects are crucial for us to make reliable estimates of the relationships between oversight and morale, as agency characteristics (e.g., size, budget, political insulation) are so often correlated with each other and with congressional attention. We thus limit our current attention to uncovering reliable estimates, net of the effects of agency-level characteristics, and leave the subtle task of mechanism assessment to future research.[27]

We should also note that we have some *ex ante* concerns regarding endogeneity. Specifically, it might be the case that instead of oversight activity affecting agency morale, the relationship is the inverse, with congressional committees choosing to hold hearings with agencies with particular latent characteristics, such as high or low autonomy. We take a number of steps to ameliorate this inferential pitfall. First, we lag the hearings covariates one year. There is little reason to expect a contemporaneous and swift reaction in the autonomy or job satisfaction dependent variables to a change in hearings activity. Instead, by lagging each of the hearings variables, we can assess what we see as a more realistic temporal ordering, where the effects of hearings in period $t-1$ take until the survey in period t to be reflected in the measured agency traits.[28] Next, we have specified each dependent variable as the one time period *change* in agency autonomy and job satisfaction from time $t-1$ to time t. As plausible as it is to consider oversight and morale being endogenously related, it is less worrisome to consider the unlikely scenario that Congress oversees agencies with especially high (or low) changes from year to year in autonomy (or job satisfaction). For these reasons, we have both lagged the primarily important hearings independent variables and created differenced *change in* autonomy and job satisfaction dependent variables.

In addition, we have modeled remaining endogeneity directly with an instrumental variables

18

approach (see e.g., Wooldridge, 2010; Angrist and Krueger, 2001). Generally, for instrumental variables regression to solve endogeneity problems, one must find an IV that is strongly correlated with the endogenous regressor (*Oversight Hearings*), but not directly related to the outcome variable (*Agency Autonomy/Job Satisfaction*). We have identified two such instruments, *Second Session of a Congress* and *Presidential Election Year*, both of which drive down congressional oversight, but show no direct correlation with our dependent variables. Inclusion of these instruments and estimation of two-stage least squares regression does not change any of our substantive interpretations, lessening our concerns regarding endogeneity.[29]

Results

Table 1 displays results for both dependent variables. For each column, we include all of the control variables described above, as well as agency fixed effects, and additional fixed effects for each year in the time series to account for systematic heterogeneity across time.[30]

Table 1 goes here.

In columns 1 and 2, we purposefully begin with a naïve model specification. In these columns, we exclude information regarding hearing sentiment and assess the *unconditional* relationships between *Oversight Hearings* and the *Change in Autonomy* and *Change in Job Satisfaction* dependent variables. Estimating this unconditional relationship serves to highlight the importance of the models found in columns 3 and 4, where we empirically distinguish between adversarial and more friendly oversight. These unconditional results demonstrate that increases in lagged *Oversight Hearings* are associated with decreases in both autonomy and job satisfaction. Both of these effects are statistically distinguishable from zero and are relatively substantial in their magnitude. In contrast, only one of the control variables across these first two models is statistically significant (*Non-oversight Hearings* in column 2).

Columns 3 and 4 introduce our operationalization of the conditionality implied by theory. While the results from columns 1 and 2 indicate that increased oversight activity leads to decreased agency autonomy and job satisfaction, we suspect that this is the case due to the distributions of adversarial and advocacy oversight hearings, with the former more likely to occur

19

than the latter in the time period being studied. To assess this explanation, and to evaluate how oversight's effect on agency morale depends on the content of the oversight attention it receives, we include our measure of *Hearings Sentiment*. As described above and in appendix B, we measured a sentiment score (ranging from most negative [-1] to most positive [+1]) for each agency hearing in the data. We then take the mean values of all of the hearings involving an agency as a global approximation of how negatively or positively Congress has interacted with each agency in each year (the mean of this variable for the estimation sample is .03, with a standard deviation of .15 and an empirical range from -.52 to .84). We then interact the lagged values of this hearings sentiment measure with the lagged number of *Oversight Hearings* involving each agency in each year to capture the intensity, as well as the direction, of agency-congressional interactions.

Column 3 presents results for *Change in Autonomy* when we add the interaction of *Oversight Hearings* and *Hearings Sentiment* to the specification from column 1. Here, the constitutive term for *Oversight Hearings* tells us that the effect of additional oversight hearings when the mean sentiment of hearings towards an agency are neutral (sentiment score of zero) is negative and statistically significant. Alternatively, we can approximately interpret this as meaning that the marginal effect of additional *neutral* oversight hearings is significantly negative, indicating that at least one of the mechanisms discussed above is at work even when hearings are not expressly negative in tone. The interaction term, on the other hand, indicates that as hearings become more positive, the effect of oversight on autonomy reverses and becomes statistically significantly positive at a *Hearings Sentiment* score around .50. These very positive hearings likely constitute what Aberbach calls "advocacy" oversight and when agencies see more of this type of oversight, it tends to increase feelings of autonomy. Since such extremely positive hearings are relatively rare in the data, this conditional relationship is obscured when we look at the results from columns 1 and 2. On the other hand, the results demonstrate that extremely negative hearings are even more likely to reduce agency autonomy than neutral hearings. To illustrate, the marginal effect of increases in hearing activity for neutral hearings (sentiment score of 0) is -.002, which more than triples for more negative hearings (sentiment score of -.25 has a marginal effect of -.007) and increases all the way to -0.011 for the most negative hearings

in the data (sentiment score of -.52). Thus, we have evidence that feelings of agency autonomy respond not only to the volume of activity, but also to the degree of negativity (or positivity) they express.

These results are substantively meaningful. Consider the distribution of the *Change in Autonomy* dependent variable — mean: -.0085, standard deviation: .34, range: -1.05 to .925. When hearings are commonly negative (say, a standard deviation below the mean of *Hearings Sentiment*: a sentiment score of -.12), it would take about 80 such hearings to lead to a standard deviation decrease in agency autonomy. On the other hand, if these hearings each carried a strongly positive sentiment (say, a sentiment score of .50), these 80 hearings would lead to an increase in agency autonomy of .045, which is significantly larger than the variable's standard deviation. While large increases in oversight are relatively rare (see figure 2 and table A1 for more information on the distribution of the variable across agencies and over time), certain agencies *do* see relatively large changes in oversight over time. Consider the Department of Defense. Figure 2 shows that hearings involving the DOD increased from a minimum of 23 to a maximum of 129. In addition, focusing solely on the coefficient estimates and their marginal effects alone may obscure the importance of oversight. A change in oversight may lead to only a small change in autonomy, but that shifts the baseline for the next period, where more oversight can further decrease (or increase, if the tone of the hearings are positive) autonomy. The dynamics of the oversight-autonomy relationship thus allows us to treat the one period effect as a floor for the true substantive impact of oversight activity.

Table 1, column 4 displays results for the same specification just described, but this time for the *Change in Job Satisfaction* dependent variable. Here, we see the same pattern of results as in column 4. Specifically, neutral and adversarial hearings tend to decrease aggregate (overall) job satisfaction within an agency, while more friendly hearings engender increases in such job satisfaction. Despite the statistical significance of the coefficient on the interaction term, the marginal effect for increases in friendly oversight is only marginally statistically significant, and only for the most positive hearings (sentiment scores of .65 or greater; compared to a .50 threshold for the *Change in Autonomy* dependent variable). Despite the smaller coefficients and effect magnitudes, we can make similar substantive interpretations of these results, as *Change*

in Job Satisfaction has a smaller standard deviation (.25) than does *Change in Autonomy* (.45). In addition, across columns 3 and 4, the oversight and sentiment variables are the only factors that consistently affect agency measures of morale, suggesting that future studies of the determinants of morale, especially those using the Bertelli et al. (2015) approach, should at least control for oversight in their empirical models.

Conclusion

As a manager of the federal bureaucracy, Congress gets mixed reviews. On one hand, when it engages in friendly oversight, it bolsters agency morale. On the other hand, when it engages in adversarial oversight, it undermines agency morale. Some of the time, then, it appears to assume the salutary managerial role that is exalted in theories of public sector organizational effectiveness (Rainey and Steinbauer, 1999; O'Toole and Meier, 1999; Fernandez, 2005; Lee and Whitford, 2013). At other times, it appears to be more interested in micromanaging and publicly shaming agencies than in abetting their performance. While it is of course Congress's prerogative to oversee the federal bureaucracy in the manner of its choosing, our results suggest that its interactions with agencies have concrete consequences for employee motivation. It strikes us reasonable that Congress should at least consider these consequences as it exercises its oversight function.

Quite simply, there is a balancing act that Congress should perform when considering oversight, and to truly understand it, scholars need to assess the managerial consequences of oversight, as well as its causes. Oversight may indeed be an effective mechanism for ensuring that agencies are responsive to the policy preferences of committee majorities (Kriner and Schickler, 2013; McGrath, 2013; MacDonald and McGrath, N.d. (forthcoming), but the congressional desire to monitor and control the bureaucracy should be balanced against adversarial oversight's likely detrimental effects on agency morale and, ultimately, agency performance. Our results suggest that "micromanagement" is more than a mere theoretical possibility. Apart from losing the benefits of delegation (expertise, insulation, etc.), Congress risks harming agency morale when it too vigorously monitors its agents. This should especially be concerning for a particular flavor of "show-horse" oversight that lacks policy content and is instead motivated by

the desire to embarrass political opponents. Yet, it is also problematic in policy areas where technical expertise is required and political incentives align to meddle with policy details, as in the Medicare example above.[31]

Ours is the first study to examine the relationship between oversight activity and latent agency characteristics, but it should not be considered the last word on the topic. We admit to a number of specific drawbacks of this study, as currently constructed. First, we do not directly measure agency performance. Instead, we focus on publicly available data on agency autonomy and job satisfaction as precursors to performance. Second, though we have proposed three theoretical mechanisms via which adversarial oversight negatively affects agency morale, our analyses cannot distinguish between these mechanisms. We envision progress on this front occurring as existing approaches to textual analysis are refined. Ultimately, we hope to be able to distinguish adversarial oversight hearings in which Congress is micromanaging from adversarial hearings in which Congress is simply shaming an agency. At the same time, we hope to be able to distinguish friendly oversight hearings in which Congress is genuinely engaged in the role of a performance manager from friendly hearings in which Congress is simply patting an agency on the back. When genuinely engaged, we would expect Congress to express commitment to a clear mission, to be attentive to agency exigencies, to allocate resources when necessary, and to buffer agencies from the demands of the external environment (e.g., from the demands of particularistic interest groups). Knowing with a greater degree of precision what sort of oversight is actually happening during a hearing will allow scholars to pin down the theoretical mechanism (or mechanisms) via which oversight operates on agency attitudes and behavior.

Up to now, empirical research has largely ignored the potential managerial consequences, both positive and negative, of congressional oversight. In particular, oversight's negative managerial consequences have long been a cause for concern in the public administration and management literatures. At the same time, the political science literature evinces a deep concern for democratic accountability and its theoretical guarantor—political control. We have sought to synthesize these two perspectives and feel that we have identified an area where more research could lead to better agency performance on the ground. Our research speaks to classic

debates concerning the politics-administration dichotomy and identifies a tangible consequence of the increase in oversight activity that has recently attracted much attention. Yet, a great deal remains to be done regarding empirical assessments of the consequences of congressional oversight.

Notes

[1]Although we argue that morale is related to organizational performance, we are careful not to conflate the two concepts. Indeed, we expect to see future research make more direct assessments of the relationship between congressional oversight and agency performance.

[2]Or selection of a "representative bureaucracy" (e.g., Meier, 1975; Meier and Nigro, 1976).

[3]We recognize that Congress is not, strictly speaking, the "boss" of the federal bureaucracy. Yet, we use workplace terminology as a metaphor for the principal-agent relationship that is said to exist between Congress and the bureaucracy (Miller, 2005). In addition, Congress is hardly an agency's only boss (Whitford, 2005; Gailmard, 2009), yet we would argue that the existence of multiple principals actually attenuates the empirical results we find below.

[4]That congressional oversight is primarily determined by political and policy motivations is well-established in the political science literature (see, e.g., Dodd and Schott, 1979; Aberbach, 1990; Kriner and Schwartz, 2008; Parker and Dull, 2009; McGrath, 2013; Kriner and Schickler, 2013).

[5]We do not separate managers from non-managers in our empirical analyses below, as this would limit the number of agencies for which we had enough managers to calculate aggregate morale. If managers are more negatively influenced by congressional oversight than non-managers, the large presence of non-managers in our data would bias our coefficients downward, making our estimates conservative.

[6]In their account of declining morale at the Department of Homeland Security, Markon, Nakashima and Crites (2014) identify "relentless congressional carping" as a source of employee unhappiness.

[7]In the empirical analyses that follow, we are largely agnostic as to the mechanism driving the findings, and suspect that all three are at work across the heterogeneous sample of hearings and agencies in our dataset.

[8]There are many ways in which legislatures can review, monitor, and supervise executive action. Committee members may engage in personal communication (even when this communication is technically illegal as "ex parte" communication) with bureaucratic staff or agency heads. Committee staff may also engage in such casework on behalf of their constituents. Besides committees, inspectors general reports (Light, 1993), General Accounting Office reports, and resolutions of inquiry (Oleszek, 2001) can serve to supplement the formal oversight work that committees engage in through hearings

[9]Smith (2003), McGrath (2013), and MacDonald and McGrath (N.d. (forthcoming) use hearings data from the Policy Agendas Project's (`www.policyagendas.org`) Congressional Hearings database (`http://www.utexas.edu/cola/_webservices/policyagendas/ch/instances.csv?from=1945&to=2012`) to construct summaries of oversight activity. Designed to capture congressional behavior, this data source fails to indicate any agency information for the identified hearings. That is, although one can measure how often each committee or subcommittee of Congress met with agency personnel in a formal hearing, the Policy Agendas Project does not allow us to recover which agency is being scrutinized in each hearing.

[10]We found hearings from 1993-1996 in the FDsys, but these constitute far less than a universe of committee hearings in those years. In addition, we limit our sample by ignoring data from 1997-1999, as the number of hearings identified in the GPO data for those years is far fewer than the number recorded in the Policy Agendas Project data.

[11]The GPO further reports (`{http://www.gpo.gov/fdsys/browse/collection.action?collectionCode=CHRG}`)that

"Not all congressional hearings are available on FDsys. Whether or not a hearing is disseminated on FDsys depends on the committee. GPO continues to add hearings as they become available during each session of Congress. If a congressional hearing is not listed in FDsys, it is not available electronically via GPO at this time. NOTE: If a committee has not made a hearing available electronically via GPO for a specific Congress, the committee's name will not appear in the browse list until a hearing for that committee is made available in FDsys." While this is a worrying disclosure, each standing committee with oversight jurisdiction published hearings through the GPO in all years of the data. In addition, as mentioned above, the hearings published via the GPO closely track those identified in the Policy Agendas Project from 1999-2004 (the end year of complete data in that dataset). In short, missing hearings data may be a problem, but there is no way to confirm the extent to which it is, or to correct for such missing data. We are confident that we have collected the universe of publicly available hearings data from 1999-2011.

[12]These keywords are: "oversight," "investigation", and "budget request."

[13]For this paper, we focus on subsets of these agencies where data on agency morale are currently available, as described below and further in appendix A.

[14]For example, the U.S. Air Force, Army, and Navy are treated discretely apart from their parent Department of Defense.

[15]See, e.g., Hopkins and King (2010) and Grimmer and Stewart (2013) for overviews and applications for political science. The basic idea is that articles, tweets, posts, hearings, etc. express positive, negative, or neutral sentiments and that we can uncover and estimate these sentiments using statistical models.

[16]As opposed to calculating the sentiment of the entire hearing, our scores measure sentiment directed toward a particular agency, denoted by the presence of the agency's name in each transcript. We use a general algorithm to calculate sentiment scores and describe our process and the algorithm in more detail in appendix B.

[17]See Bertelli et al. (2015) for a brief review of these studies.

[18]See appendix C for a description of these data sources and a list of the questions and the surveys from which they were drawn. For more information on the aggregation and estimation procedures, see Bertelli et al. (2015).

[19]By contrast, "facet" measures of job satisfaction are comprised of survey items that refer to specific aspects of an individual's job, such as satisfaction with one's coworkers, satisfaction with one's opportunities for career advancement, or satisfaction with one's pay. We use a global measure of job satisfaction because meta-analytic evidence suggests that global measures are stronger predictors of job performance than are facet measures (see Judge et al., 2001).

[20]These are the summary measures found at `http://agencydata.wordpress.com`. They are bounded at -5 and 5.

[21]The panel is unbalanced, as some agencies are missing data on key variables in some years. See appendices A and C for more information regarding missingness in the data.

[22]Fire-alarms may also spur court action, as federal courts have vast jurisdiction over federal agency policymaking and have the power to overturn agency decisions. Such court action might simultaneously affect agency morale and drive oversight activity, and while it would be best if future research could directly measure and incorporate judicial attention, we rely on the likely correlation of such attention with the media attention variable we create below to assuage our concern that is a threat to inference.

[23]We accessed the stories using the *Lexis Nexis Academic* database. We chose to explore *Washington Post* stories, in

particular, because this newspaper dedicates more of its coverage to the federal bureaucracy than other national news organizations, such as the *New York Times*. See appendix D for more detailed coding information.

[24]The bulk of these non-oversight hearings concern prospective legislation, where agency testimony is used by a congressional committee to inform their policy decisions.

[25]`http://www.gpo.gov/fdsys/search/advanced/advsearchpage.action`. We did the same with the Congressional Record for a measure of *Congressional Attention*, apart from attention through hearings. Yet, with agency fixed effects, we are limited in the number of covariates that we can include in a single model and this variable proved highly correlated (Pearson's r of 0.67) with our measure of *Oversight Hearings*.

[26]We have included some slowly changing agency-level variables, such as agency size, budget, and percentage of employees who are appointed, but they do not add to model fit or change the substantive interpretations we present.

[27]Our data are not particularly amenable to testing mechanisms, so future research should focus on establishing the micro-foundations of our theory by examining individual survey data, rather than aggregating to the agency level.

[28]We lag the *Total Washington Post Sentiment* and *Hearings Sentiment* variables for the same reasons.

[29]For ease of interpretation, and since the results are largely identical across specifications, we present the standard regression results below, but present the second stage instrumental variables results in appendix table E1.

[30]We also cluster all standard errors by agency to allow for agency-specific trends in the error term. This has the effect of increasing the standard errors and makes finding statistical significance more difficult.

[31]In fact, Congress itself has recognized the problem with this type of micromanagement and has sought to remedy its own proclivities. To wit, the 2010 Patient Protection and Affordable Care Act created an Independent Payment Advisory Board that has the ability to change payment schedules without prior congressional approval (although these decisions are subject to a supermajoritarian congressional veto).

References

Aberbach, Joel D. 1990. *Keeping a Watchful Eye: The Politics of Congressional Oversight*. Washington, D.C.: Brookings Institution Press.

Angrist, Joshua D. and Alan B. Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *The Journal of Economic Perspectives* 15(4):69–85.

Aryee, Samuel, Zhen Xiong Chen, Li-Yun Sun and Yaw A Debrah. 2007. "Antecedents and outcomes of abusive supervision: test of a trickle-down model." *Journal of Applied Psychology* 92(1):191.

Balla, Steven J. 1998. "Administrative Procedures and Political Control of the Bureaucracy." *The American Political Science Review* 92(3):663–673.

Balla, Steven J. and Christopher J. Deering. 2013. "Police Patrol and Fire Alarms: An Empirical Examination of the Legislative Preference for Oversight." *Congress & the Presidency* 40(1):27–40.

Barrick, Murray R, Michael K Mount and Ning Li. 2013. "The Theory of Purposeful Work Behavior: The Role of Personality, Higher-order Goals, and Job Characteristics." *Academy of Management Review* 38(1):132–153.

Behn, Robert D. 1995. "The Big Questions of Public Management." *Public Administration Review* pp. 313–324.

Bendor, Jonathan, Serge Taylor and Roland Van Gaalen. 1985. "Bureaucratic Expertise versus Legislative Authority: A Model of Deception and Monitoring in Budgeting." *American Political Science Review* 79(4):1041.

Bertelli, Anthony M. and Christian R. Grose. 2011. "The Lengthened Shadow of Another Institution: Ideal Point Estimates for the Executive Branch and Congress." *American Journal of Political Science* 55:767–781.

Bertelli, Anthony M., Dyana P. Mason, Jennifer M. Connolly and David A. Gastwirth. 2015. "Measuring Agency Attributes with Attitudes Across Time: A Method and Examples Using Large-Scale Federal Surveys." *Journal of Public Administration Research and Theory* 25(2):513–544.

Brewer, Gene A. and Sally Coleman Selden. 2000. "Why Elephants Gallop: Assessing and Predicting Organizational Performance in Federal Agencies." *Journal of Public Administration Research and Theory* 10(4):685–712.

Clinton, Joshua D., Anthony Bertelli, Christian Grose, David E. Lewis and David C. Nixon. 2012. "Separated Powers in the United States: The Ideology of Agencies, Presidents, and Congress." *American Journal of Political Science* 56(2):341–354.

Clinton, Joshua D., Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *The American Political Science Review* 98(2):355–370.

Dodd, Lawrence C. and Richard L. Schott. 1979. *Congress and the Administrative State*. New York: John Wiley & Sons.

Eisenhardt, Kathleen M. 1989. "Agency Theory: An Assessment and Review." *Academy of Management Review* 14(1):57–74.

Epstein, David and Sharyn O'Halloran. 1996. "Divided Government and the Design of Administrative Procedures: A Formal Model and Empirical Test." *The Journal of Politics* 58(2):373–397.

Epstein, David and Sharyn O'Halloran. 1999. *Delegating Powers: A Transaction Cost Politics Approach to Policymaking Under Separate Powers.* Cambridge: Cambridge University Press.

Ferejohn, John and Charles Shipan. 1990. "Congressional Influence on Bureaucracy." *Journal of Law, Economics, & Organization* 6:1–20.

Fernandez, Sergio. 2005. "Developing and Testing an Integrative Framework of Public Sector Leadership: Evidence from the Public Education Arena." *Journal of Public Administration Research and Theory* 15(2):197–217.

Finer, Herman. 1941. "Administrative Responsibility in Democratic Government." *Public Administration Review* 1:335–350.

Frey, Bruno S and Felix Oberholzer-Gee. 1997. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-out." *The American Economic Review* pp. 746–755.

Friedrich, Carl J. 1940. "Public Policy and the Nature of Administrative Responsibility". In *Public Policy*, ed. Carl J. Friedrich and Edward S. Mason. Cambridge, MA: Harvard University Press.

Gagné, Marylène and Edward L. Deci. 2005. "Self-Determination Theory and Work Motivation." *Journal of Organizational Behavior* 26(4):331–362.

Gailmard, Sean. 2009. "Multiple Principals and Oversight of Bureaucratic Policy-Making." *Journal of Theoretical Politics* 21:161.

Gailmard, Sean and John W. Patty. 2007. "Slackers and Zealots: Civil Service, Policy Discretion, and Bureaucratic Expertise." *American Journal of Political Science* 51(4):873–889.

Gailmard, Sean and John W. Patty. 2012. *Learning While Governing: Expertise and Accountability in the Executive Branch.* University of Chicago Press.

Gilmour, Robert Scott and Alexis A. Halley, eds. 1994. *Who Makes Public Policy?: The Struggle for Control Between Congress and the Executive.* Chatham House Publishers.

Hackman, J Richard and Greg R Oldham. 1976. "Motivation Through the Design of Work: Test of a Theory." *Organizational behavior and human performance* 16(2):250–279.

Houston, David J. 2009. "Motivating Knights or Knaves? Moving beyond Performance-related Pay for the Public Sector." *Public Administration Review* 69(1):43–57.

Huber, John D. and Charles R. Shipan. 2002. *Deliberate Discretion? The Institutional Foundations of Bureaucratic Autonomy.* Cambridge: Cambridge University Press.

Judge, Timothy A., Carl J. Thoresen, Joyce E. Bono and Gregory K. Patton. 2001. "The Job Satisfaction–Job Performance Relationship: A Qualitative and Quantitative Review." *Psychological Bulletin* 127(3):376.

Kiewiet, Roderick and Mathew McCubbins. 1991. *The Logic of Delegation: Congressional Parties and the Appropriation Process.* Chicago: University of Chicago Press.

Kriner, Douglas and Eric Schickler. 2013. "The Policy Consequences of Congressional Investigations." Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago, IL, April 2013.

Kriner, Douglas and Liam Schwartz. 2008. "Divided Government and Congressional Investigations." *Legislative Studies Quarterly* 33:295–322.

Lee, Jungwook, Hal G. Rainey and Younghan Chun. 2009. "Of Politics and Purpose: Political Salience and Goal Ambiguity of U.S. Federal Agencies." *Public Administration* 87:457–484.

Lee, Soo-Young and Andrew B. Whitford. 2013. "Assessing the Effects of Organizational Resources on Public Agency Performance: Evidence from the US Federal Government." *Journal of Public Administration Research and Theory* 23(3):687–712.

Light, Paul C. 1993. *Monitoring Government: Inspectors General and the Search for Accountability.* Washington, D.C.: The Brookings Institution.

Light, Paul C. 2006. "The Tides of Reform Revisited: Patterns in Making Government Work, 1945–2002." *Public Administration Review* 66(1):6–19.

MacDonald, Jason A. and Robert J. McGrath. N.d. (forthcoming). "Retrospective Congressional Oversight and the Dynamics of Legislative Influence over the Bureaucracy." *Legislative Studies Quarterly* .

Markon, Jerry, Ellen Nakashima and Alice Crites. 2014. "Top-level turnover makes it harder for DHS to stay on top of evolving threats.".
URL: *http://www.washingtonpost.com/politics/top-level-turnover-makes-it-harder-for-dhs-to-stay-on-top-of-evolving-threats/2014/09/21/ca7919a6-39d7-11e4-9c9f-ebb47272e40e_story.html*

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10(2):134–153.

Mathieu, John E and Dennis M Zajac. 1990. "A Review and Meta-Analysis of the Antecedents, Correlates, and Consequences of Organizational Commitment." *Psychological Bulletin* 108(2):171.

May, Peter J., Ashley E. Jochim and Joshua Sapotichne. 2011. "Constructing Homeland Security: An Anemic Policy Regime." *Policy Studies Journal* 39(2):285–307.

May, Peter J., Joshua Sapotichne and Samuel Workman. 2009. "Widespread Policy Disruption: Terrorism, Public Risks, and Homeland Security." *Policy Studies Journal* 37(2):171–194.

May, Peter J., Samuel Workman and Bryan D. Jones. 2008. "Organizing Attention: Responses of the Bureaucracy to Agenda Disruption." *Journal of Public Administration Research and Theory* 18(4):517–541.

Mayhew, David R. 2005. *Divided We Govern: Party Control, Lawmaking, and Investigations, 1946-1990.* 2nd ed. New Haven, CT: Yale University Press.

McCubbins, Mathew D. and Thomas Schwartz. 1984. "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms." *American Journal of Political Science* 28(1):165–179.

McGrath, Robert J. 2013. "Congressional Oversight Hearings and Policy Control." *Legislative Studies Quarterly* 38(3):353–380.

Meier, Kenneth J. 1975. "Representative Bureaucracy: An Empirical Analysis." *American Political Science Review* 69:526–542.

Meier, Kenneth J. 1997. "Bureaucracy and Democracy: The Case for More Bureaucracy and Less Democracy." *Public Administration Review* 57(3):193–199.

Meier, Kenneth J. and Lloyd G. Nigro. 1976. "Representative Bureaucracy and Policy Preferences." *Public Administration Review* pp. 458–69.

Meyer, John P, David J Stanley, Lynne Herscovitch and Laryssa Topolnytsky. 2002. "Affective, Continuance, and Normative Commitment to the Organization: A Meta-Analysis of Antecedents, Correlates, and Consequences." *Journal of Vocational Behavior* 61(1):20–52.

Miller, Gary J. 2005. "The Political Evolution of Principal-Agent Models." *Annual Review of Political Science* 8:203–25.

Moe, Terry M. 1989. The Politics of Bureaucratic Structure. In *Can the Government Govern?*, ed. John Chubb and Paul Peterson. Washington, DC: Brookings Institution pp. 267–329.

Ogul, Morris S. and Bert A. Rockman. 1990. "Overseeing Oversight: New Departures and Old Problems." *Legislative Studies Quarterly* 15:5–24.

Oleszek, Walter J. 2001. *Congressional Procedures and the Policy Process.* 5th ed. Washington, D.C.: CQ Press.

O'Toole, Laurence J. and Kenneth J. Meier. 1999. "Modeling the Impact of Public Management: Implications of Structural Context." *Journal of Public Administration Research and Theory* 9(4):505–526.

Parker, David C.W. and Matthew Dull. 2009. "Divided We Quarrel: The Politics of Congressional Investigations." *Legislative Studies Quarterly* 34:319–345.

Parnell, Arche. 1980. "Congressional Interference in Agency Enforcement: The IRS Experience." *Yale Law Journal* pp. 1360–1394.

Perry, James L and Lois Recascino Wise. 1990. "The Motivational Bases of Public Service." *Public Administration Review* pp. 367–373.

Pham, Hoangmai H, Paul B Ginsburg and James M Verdier. 2009. "Medicare Governance and Provider Payment Policy." *Health Affairs* 28(5):1382–1394.

Rainey, Hal G. and Paula Steinbauer. 1999. "Galloping Elephants: Developing Elements of a Theory of Effective Government Organizations." *Journal of Public Administration Research and Theory* 9(1):1–32.

Riketta, Michael. 2002. "Attitudinal Organizational Commitment and Job Performance: A Meta-Analysis." *Journal of Organizational Behavior* 23(3):257–266.

Rosenbloom, David H. 1993. "Editorial: Have an Administrative Rx? Don't Forget the Politics!" *Public Administration Review* 53(6):503–507.

Ryan, Richard M and Edward L Deci. 2000. "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being." *American psychologist* 55(1):68–78.

Shapiro, Sidney A. 1994. "Political Oversight and the Deterioration of Regulatory Policy." *Administrative Law Review* pp. 1–40.

Shipan, Charles R. 2004. "Regulatory Regimes, Agency Actions, and the Conditional Nature of Congressional Influence." *American Political Science Review* 98(3):467–480.

Smith, Keith. 2003. "The Growth of Congressional Oversight". Paper presented at the 2003 meeting of the American Political Science Association, Philadelphia, PA.

Spector, Paul E. 1986. "Perceived Control by Employees: A Meta-Analysis of Studies Concerning Autonomy and Participation at Work." *Human Relations* 39(11):1005–1016.

Tepper, Bennett J. 2000. "Consequences of Abusive Supervision." *Academy of Management Journal* 43(2):178–190.

Tepper, Bennett J. 2007. "Abusive supervision in work organizations: Review, synthesis, and research agenda." *Journal of Management* 33(3):261–289.

Tett, Robert P and John P Meyer. 1993. "Job Satisfaction, Organizational Commitment, Turnover Intention, and Turnover: Path Analyses Based on Meta-Analytic Findings." *Personnel Psychology* 46(2):259–293.

Volden, Craig. 2002. "Delegating Power to Bureaucracies: Evidence from the States." *Journal of Law, Economics, & Organization* 18(1):187–220.

Waldo, Dwight. 1948. *The Administrative State*. Holmes & Meier Publishers.

Ward, Kevin D. 2014. "Cultivating public service motivation through AmeriCorps Service: A longitudinal study." *Public Administration Review* 74(1):114–125.

Weingast, Barry R. and Mark J. Moran. 1983. "Bureaucratic Discretion or Congressional Control? Regulatory Policymaking by the Federal Trade Commission." *The Journal of Political Economy* 91(5):765–800.

Whitford, Andrew B. 2005. "The Pursuit of Political Control by Multiple Principals." *Journal of Politics* 67(1):28–49.

Wilson, James Q. 1989. *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.

Wilson, Woodrow. 1887. "The Study of Administration". In *Classics of Public Administration, 6th edition*, ed. Jay M. Shafritz and Albert C. Hyde. Boston: Thomson Wadsworth.

Wolf, Patrick J. 1993. "A Case Survey of Bureaucratic Effectiveness in US Cabinet Agencies: Preliminary Results." *Journal of Public Administration Research and Theory* 3(2):161–181.

Wood, B. Dan and John Bohte. 2004. "Political Transaction Costs and the Politics of Administrative Design." *Journal of Politics* 66:176–202.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Author Contact Information

John D. Marvel
School of Policy, Government, and International Affairs
George Mason University
4400 University Drive, MSN 3F4
Fairfax, VA 22030
jmarvel@gmu.edu

Robert J. McGrath – *corresponding author*
School of Policy, Government, and International Affairs
George Mason University
4400 University Drive, MSN 3F4
Fairfax, VA 22030
rmcgrat2@gmu.edu

Tables

Table 1: OLS Models of Agency Autonomy and Job Satisfaction, 1999-2011

	(1) Autonomy Δ	(2) Satisfaction Δ	(3) Autonomy Δ	(4) Satisfaction Δ
Oversight Hearings (Lag)	-.00260*	-.00195**	-.00252*	-.00191**
	(.00146)	(.00091)	(.00145)	(.00091)
Hearings Sentiment (Lagged mean)			-.53547*	-.20647
			(.32041)	(.20086)
Oversight Hearings (Lag)× Hearings Sentiment (Lagged mean)			.01634***	.00877***
			(.00579)	(.00363)
Divided Government	.05127	.01691	.03571	.01032
	(.04974)	(.03100)	(.04960)	(.03109)
Republican Control of Congress	.06938	.01750	.07672	.02118
	(.05048)	(.03145)	(.05008)	(.03140)
Democratic President	-.06628	-.02864	-.05459	-.02286
	(.04879)	(.03040)	(.04856)	(.03044)
Presidential Transition Year	.01378	-.01149	.03486	-.00108
	(.07490)	(.04667)	(.07470)	(.04683)
Presidential Attention	-.00032	.00071	-.00025	.00075
	(.00122)	(.00076)	(.00120)	(.00076)
Non-oversight Hearings (Lag)	.01320**	-.00062	.01294**	-.00074
	(.00541)	(.00337)	(.00535)	(.00336)
Total *Washington Post* Sentiment (Lag)	-.00001	.00005	-.00011	.00001
	(.00028)	(.00017)	(.00028)	(.00017)
Constant	.02891	.06925	-.06920	.01794
	(.20208)	(.12592)	(.20863)	(.13079)
Agency FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	470	470	470	470
R^2	.543	.676	.555	.680
AIC	126.855	-317.777	118.732	-320.221
BIC	475.685	31.052	475.867	36.914

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$

Note: Entries are linear regression coefficient estimates and standard errors, clustered by agency. The dependent variables are created by calculating the change in the Bertelli et al. (2015) measures of autonomy and job satisfaction (excluding compensation questions) from time $t-1$ to time t. Agency and year fixed effects are included in all models but not reported. See appendix A for further description of the oversight data, appendix C for more information on the hearings sentiment scores, and appendix D for a description of the *Washington Post* sentiment scores.

Figures

Figure 1: Oversight Hearing Days (1999-2011)
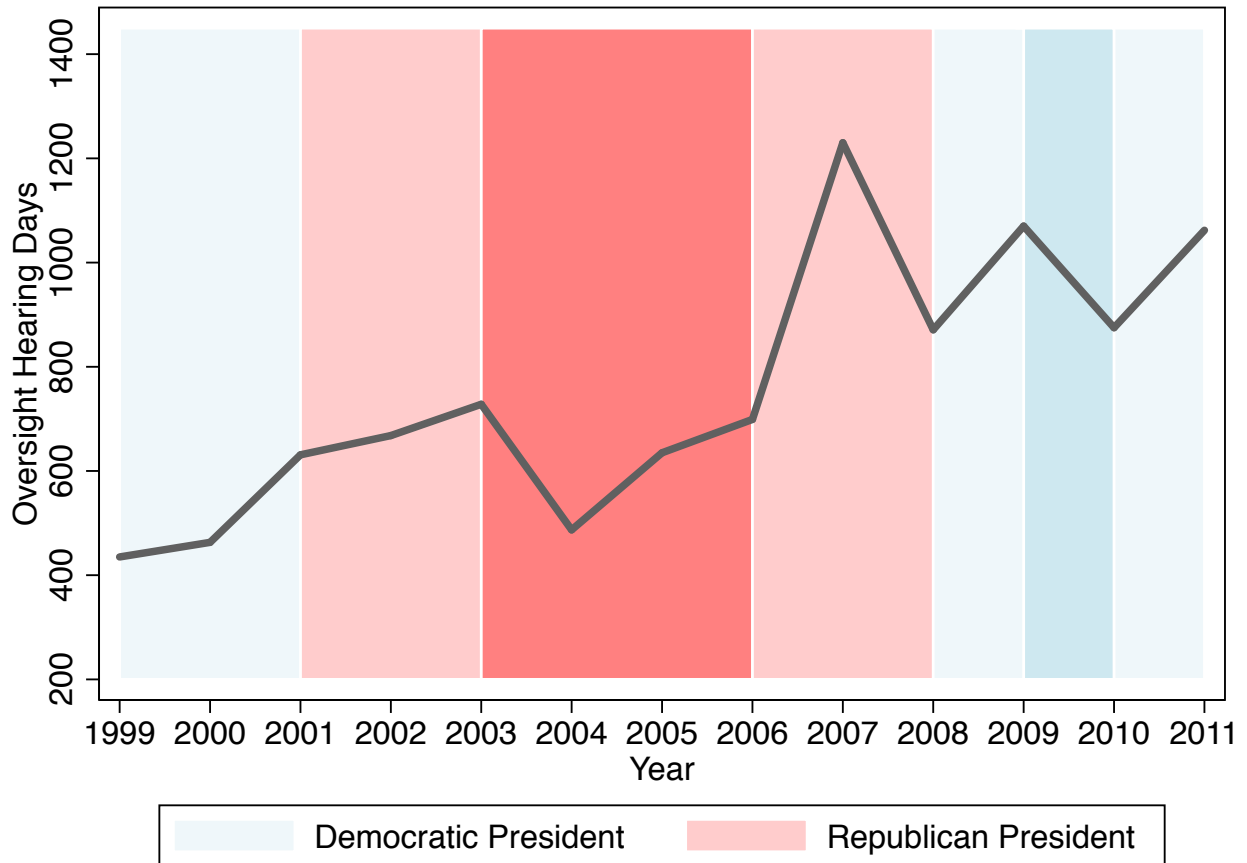
Note: Figure created by summing all oversight hearing days for all agencies in a given year. Hearings can involve more than one agency at a time, so this can result in double counting hearings. We show this double-counted measure of hearings activity as it more accurately captures total agency attention to congressional priorities. More intense colors indicate unified Democratic or Republican control.

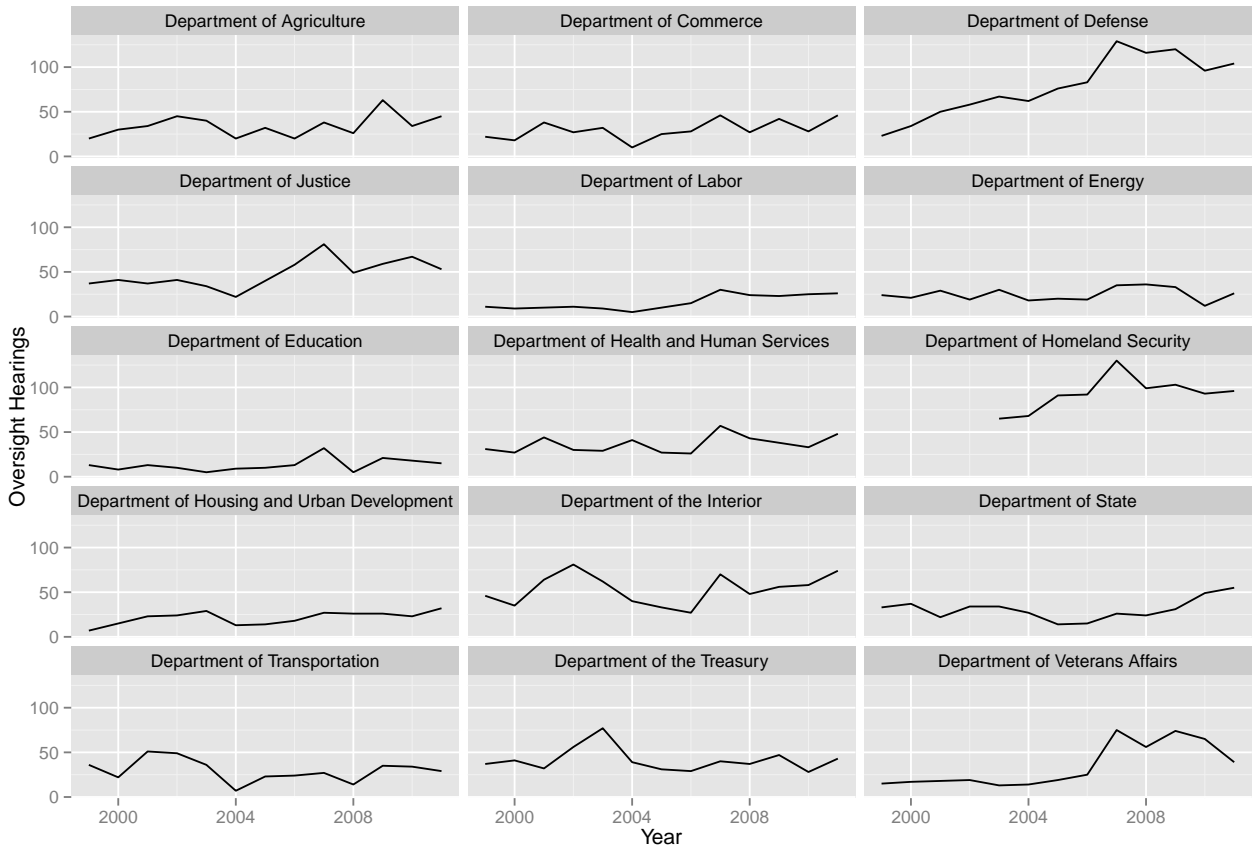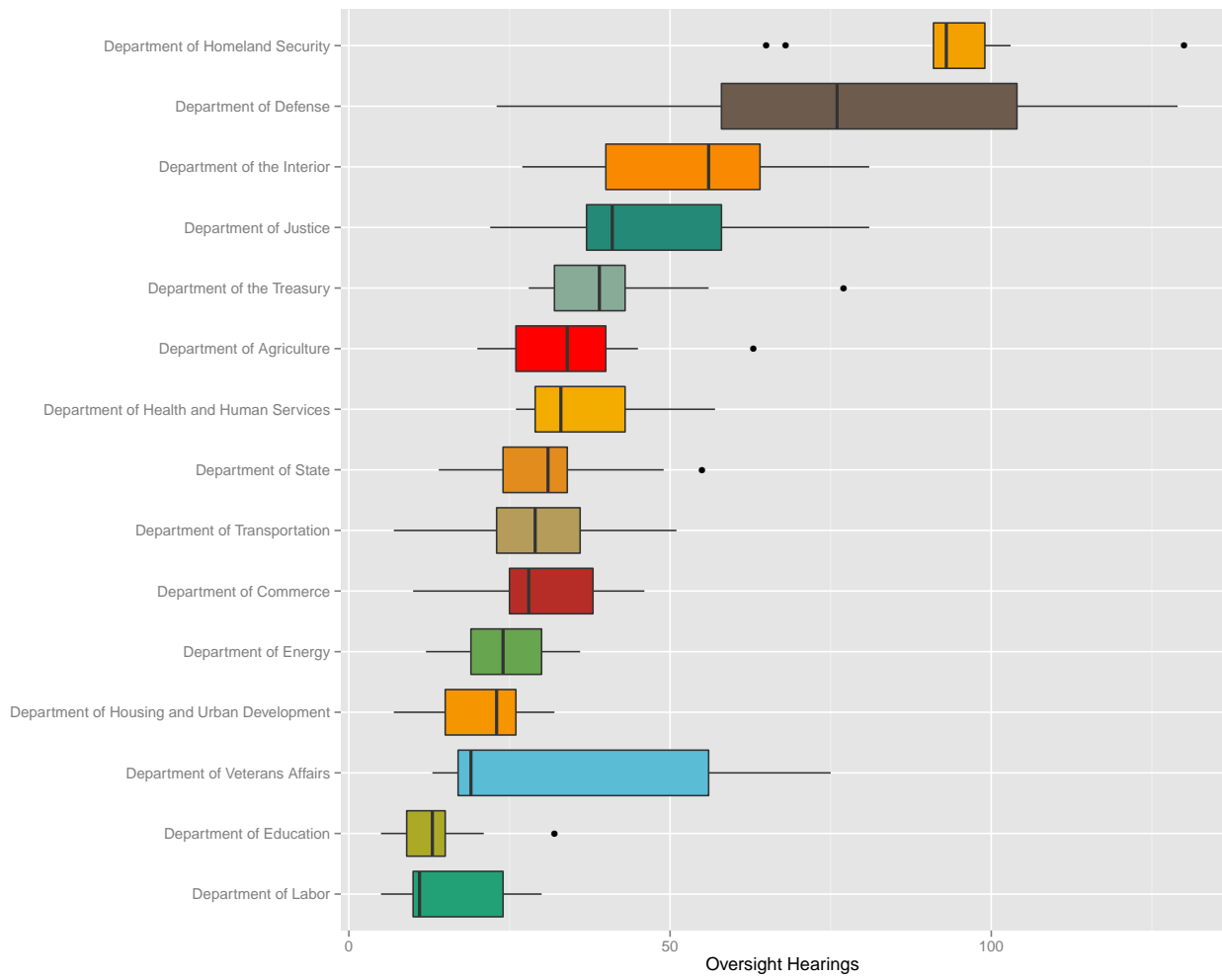Figure 2: Oversight Hearings over Time, by Department

Figure 3: Distribution of Oversight Hearings, by Department

Note: Horizontal lines give the median for each agency, boxes give the bounds of the interquartile range, and dots show outliers.
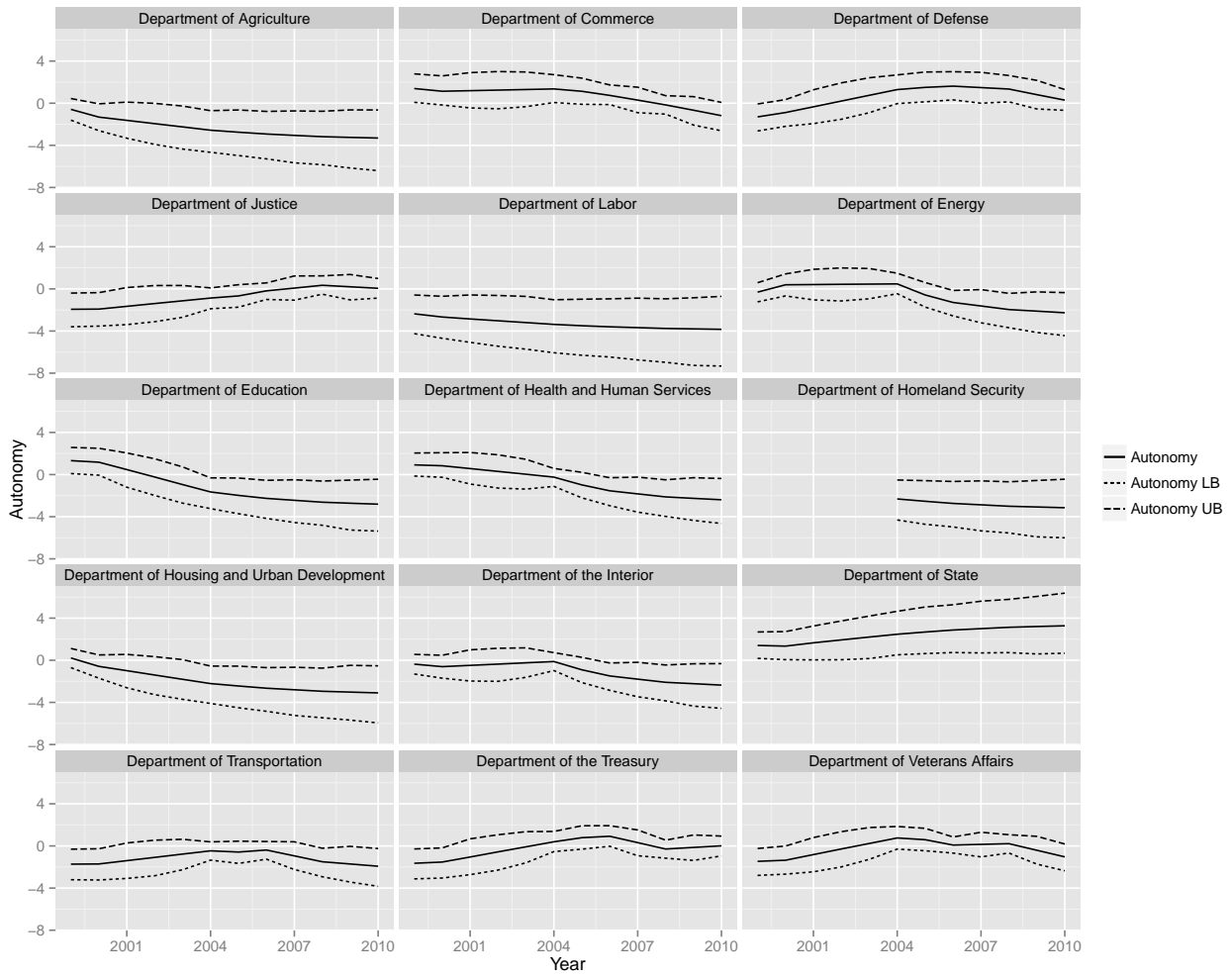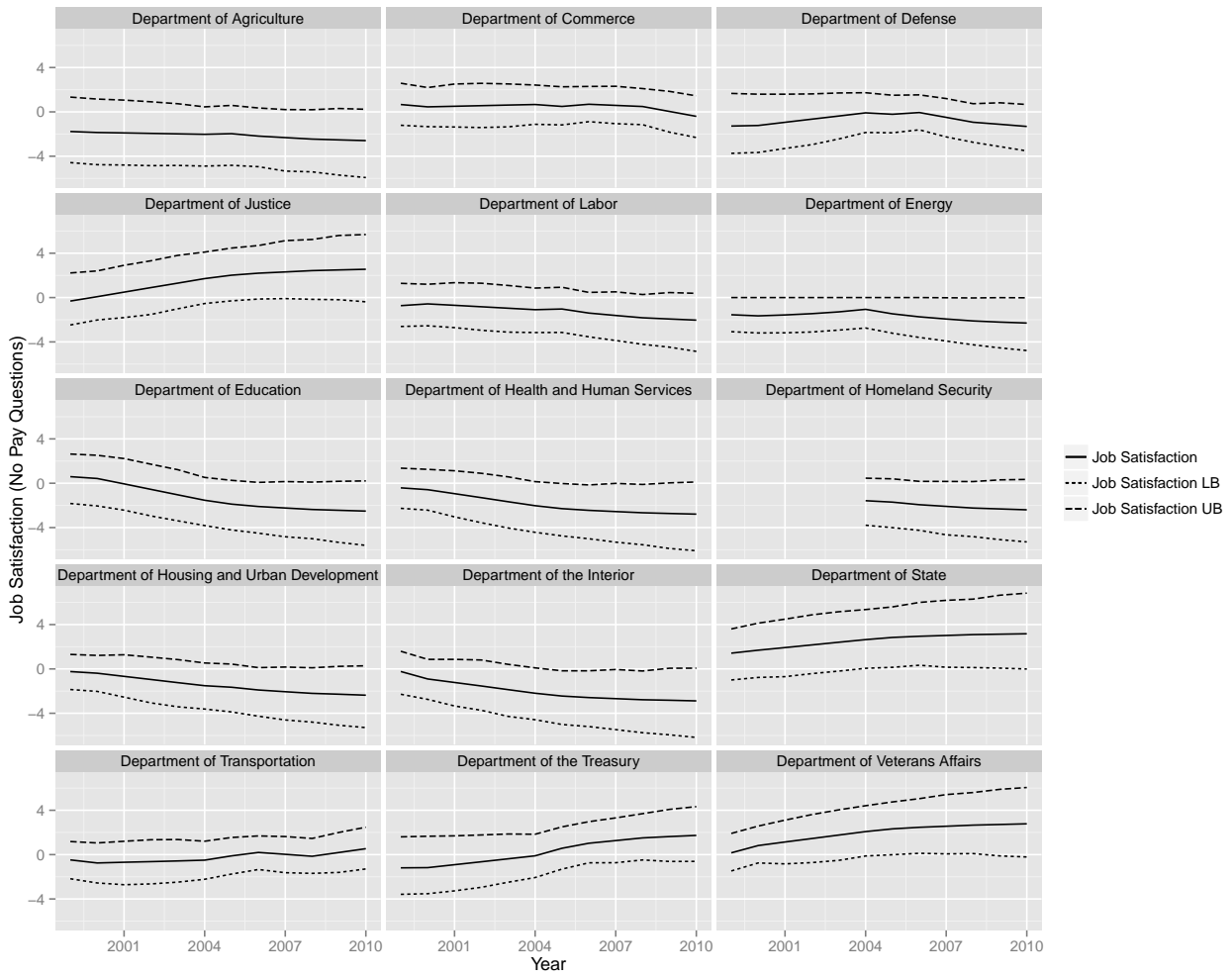
Figure 4: Agency Autonomy, by Department

Figure 5: Job Satisfaction, by Department

Appendix A - Description of Oversight Data

As described in the text, we collected all available data on hearings from the Government Printing Office's Federal Digital System (`http://www.gpo.gov/fdsys/search/advanced/advsearchpage.action`). We batch downloaded metadata on each hearing (N: 17,572) in XML format and converted these individual files to a CSV-formatted spreadsheet of hearings metadata, with each XML file becoming a row in the data. The XML tags within each file became columns in the aggregate metadata and included such information as "Title," "Held Date," "Committee," "Subcommittee," etc., and then, importantly, information on the identity of witnesses. This information included the full names and titles of witnesses in each hearing, including institutional affiliation(s). Some hearings do not have witnesses and the maximum number of witnesses in our data is 157. Using the scheme of the 2012 Federal Human Capital Survey, we matched each agency witness to their respective agency and used this information to create an agency-year dataset of hearings.

Next, we needed to make sure that we did something to separate oversight hearings from hearings having to do with proposed legislation or appropriations. We collected the full text transcripts of each hearing and searched for the following keywords meant to indicate oversight activity: "oversight," "investigation," and "budget request." This is a subset of keywords used in recent research to identify oversight (McGrath, 2013; MacDonald and McGrath, N.d. (forthcoming), but our study differs from these in that we have access to the full text of each hearing, rather than just the abstracts provided by the Policy Agendas Project. We found that nearly 65% of the total number of hearings included at least one of these keywords and was thus coded as an oversight hearing (N: 11,407).

After identifying oversight hearings from the universe of available hearings and identifying agency witnesses, we were able to compile the full agency-year data on hearings, oversight and non-oversight. The agency-year dataset has 1,053 observations—13 full years of data (1999-2011) for 80 agencies and 2 agencies with fewer than 13 observations due to being created after 1999.

Table A1 below displays agency-level summary statistics for oversight hearings in which an agency employee was a witness.

Table A1: Descriptive Statistics for Yearly Hearings, by Agency (**Departments in bold**, *agencies included in models in italics*)

Agency	MeanSD	Min-Max	Agency	MeanSD	Min-Max
United States Air Force	15.08 14.38	1-39	**Department of Homeland Security**[2]	93 19.13	65-130
Department of Agriculture	34.38 12.26	20-63	**Department of HUD**	21.31 7.31	7-32
National Endowment for the Arts	0.07 0.27	0-1	*Broadcasting Board of Governors*	0.23 0.43	0-1
National Endowment for the Humanities	0.15 0.37	0-1	*Inter-American Foundation*	0.08 0.27	0-1
Institute of Museum and Library Services	0.23 0.59	0-2	*National Indian Gaming Commission*	0.62 0.87	0-2
Agency for International Development	4.31 4.31	0-13	**Department of the Interior**	53.38 16.81	27-81
United States Army	29.07 17.11	7-66	*Kennedy Center*	0.07 0.27	0-1
Federal Labor Relations Authority	0.38 0.65	0-2	*Corporation for National and Community Service*	0.15 0.37	0-1
Merit Systems Protection Board	0.62 0.87	0-3	*Federal Election Commission*	0.23 0.83	0-3
Defense Nuclear Facilities Safety Board	0.15 0.37	0-1	*Federal Maritime Commission*	0.62 0.77	0-2
Pension Benefit Guaranty Corporation	0.92 1.49	0-4	*National Science Foundation*	3.62 3.09	0-10
Office of Management and Budget	12.92 5.95	3-24	*National Labor Relations Board*	0.15 0.38	0-1
U.S. Access Board	0.08 0.28	0-1	*National Mediation Board*	0 0	0-0
Department of Commerce	29.92 10.74	10-46	*National Aeronautics and Space Administration*	5.07 3.3	0-12
Commodity Futures Trading Commission	4.84 5.68	0-19	*National Capital Planning Commission*	0.07 0.27	0-1
National Credit Union Administration	2.46 1.89	0-6	*National Archives and Records Administration*	0.84 0.98	0-3
Department of Defense	78.31 33.36	23-129	*Nuclear Regulatory Commission*	3.69 3.06	0-11
Department of Justice	47.62 15.73	22-81	*United States Navy*	13.07 11.96	1-33
Department of Labor	16 8.33	5-30	*Office of Personnel Management*	6.07 4.01	0-11
Department of Energy	24.76 7.41	12-36	*Occupational Safety and Health Review Comm.*	0 0	0-0
Federal Energy Regulatory Commission	2.31 2.06	0-5	*Postal Regulatory Commission*	0.69 0.94	0-2
Export-Import Bank of the United States	1.31 1.6	0-5	*Office of Navajo and Hopi Indian Relocation*	0 0	0-0
Department of Education	13.23 7.31	5-32	*Federal Retirement Thrift Investment Board*	0.53 0.87	0-3
Equal Employment Opportunity Commission	0.31 0.48	0-1	*Railroad Retirement Board*	0 0	0-0
Environmental Protection Agency	19.46 12.64	8-55	*Small Business Administration*	10.15 5.58	3-26
Trade and Development Agency	0.46 0.52	0-1	*Securities and Exchange Commission*	16.31 10.01	3-34
Federal Communications Commission	6.84 3.72	3-15	*Consumer Product Safety Commission*	1.92 2.21	0-7
Chemical Safety and Hazard Investigation Board	1 1.29	0-3	*National Gallery of Art*	0 0	0-0
Federal Mediation and Conciliation Service	0 0	0-0	*Selective Service System*	0 0	0-0
Court Services and Offender Supervision Agency	0.15 0.37	0-1	**Department of State**	30.85 11.81	14-55
Federal Trade Commission	9.23 4.89	2-18	*Woodrow Wilson International Center for Scholars*[2]	2 4.02	0-15
U.S. Office of Special Counsel	0.69 0.94	0-3	*Social Security Administration*	5.85 1.72	3-9
Overseas Private Investment Corporation	0.54 0.66	0-2	*National Transportation Safety Board*	2.53 2.06	0-7
U.S. Office of Government Ethics	1.85 1.95	0-6	*U.S. International Trade Commission*	0.54 0.97	0-3
General Services Administration	4.38 2.78	1-10	**Department of Transportation**	29.76 12.44	7-51
International Boundary and Water Commission	0.76 0.27	0-1	*Office of the U.S. Trade Representative*	1.46 1.81	0-5
Department of Health and Human Services	36.46 9.59	26-57	**Department of the Treasury**	41.31 13.2	28-77
Federal Housing Finance Agency[1]	5 2.31	3-7	*Surface Transportation Board*	1.15 0.89	0-3
Advisory Council on Historic Preservation	0.15 0.38	0-1	**Department of Veterans Affairs**	34.53 24.19	13-75
American Battle Monuments Commission	0.23 0.44	0-1	*Administration for Children and Families*	0.69 1.18	0-4
Committee for Purchase…Severely Disabled	0 0	0-0	*Commission on Civil Rights*	1.46 2.29	0-6

[a] 4 observations. Agency created in 2008
[b] 9 observations. Agency created in November 2002.

Appendix B: Measuring Hearing Sentiment

We measured targeted hearing sentiment for each of the oversight hearings we identified in the GPO data (N= 11,407, see Appendix A). The first step in our process involved identifying the agencies involved in each hearing and expanding the dataset so that each agency-hearing pair is an observation, resulting in 55,831 unique observations (29,268 observations for the House of Representatives and 26,563 for the Senate). Once we identified these agency-hearings, we prepared each transcript for processing by removing special characters and metadata.

There are a number of different approaches to sentiment analysis and we have tried two particular methodologies. In particular, we have used Alchemy API, an unsupervised deep-learning algorithm for the analyses in the main body of the paper, but we have also used a lexicon-based approach for validation.[1] Alchemy API is available as a commercial data analysis product: `http://www.alchemyapi.com`. Conveniently, the Alchemy API algorithm has a built-in procedure for weighting sentiments that are directed at user-supplied strings (in our case, the names of agencies), using both distance from the target string, and syntactic and semantic context.

Using a custom Python script, we ran each hearing through the Alchemy API algorithm to recover estimates of targeted sentiment (targeted at the agency) for each agency-hearing in the data.[2] These scores theoretically range from -1 to 1. Positive scores denote that a hearing reflected largely positively on an agency, negative scores mean the opposite. Scores at or near zero reflect hearings that did not express strong sentiment in either direction towards the agency. As a rough validation of these scores, we took a sample of very high and very low sentiments and referenced the hearing transcripts to verify that the scores had a basis in common sense. This exercise did much to ensure us of the face validity of the scores. For example, the following statement by House Committee on Government Reform Chairman Rep. Henry Waxman (D-CA) regarding the Department of Homeland Security in a 2005 hearing received a strongly negative score (-.79), for good reason:

> I'm going to be blunt in my remarks. This administration is squandering literally billions of dollars on wasteful Federal contracts. Private contractors are reaping a bonanza while taxpayers are being gouged. Whether the explanation is gross incompetence or deliberate malfeasance, the result is the same: Taxpayers are being vastly overcharged. . . . Nearly every week the papers are full of stories of contract abuse. The Department of Homeland Security has wasted hundreds of millions of dollars on security contracts that have produced virtually no result.

Besides directly expressing positive or negative sentiment towards agencies in their members' comments, committees carefully choose witness testimony to serve a particular purpose in their hearings. For example, in a 2009 hearing held by a House Select Committee on Energy Independence and Global Warming, Chairman Rep. Edward J. Markey (D-MA) called as a witness Mr. Stephen Seidel, Vice President for Policy Analysis, Pew Center on Global Climate Change. In his testimony, Mr. Seidel

---

[1]Lexicon-based approaches use "dictionaries" containing lists of positive (e.g., "fantastic," "excellent") and negative words (e.g., "failure," "abysmal"), assigning an arbitrary positive value to positive words found in the text and an arbitrary negative value to negative words, weighting sentiment words that appear closer to the target phrase more than sentiment words that are further away. While useful in many domains, the lexicon-based approach is arguably not ideal for our purposes. For one, available lexicons do not include every possible positive or negative word, especially across different domains. Since each target agency deals with its own substantive focus, from environmental protection, to banking regulation, to personnel management, the extent to which the generic lexicon applies to hearings involving each agency would be expected to vary across agencies. We could develop custom lexicon dictionaries for each agency, but it would be difficult to ensure that these apply the same standards of sentiment polarity across agencies. As our goal is explicitly to measure the tenor of congressional attention across agencies, we sought a more general approach.

[2]That is, we analyzed each of the 11,407 hearings, repeating the analysis for each specific agency mentioned, giving us 55,831 agency-hearing scores.

praises the efforts of a number of federal agencies in their strategic plans regarding curbing global warming. He is particularly effusive about the Department of the Interior ("And I would say the Department of the Interior is a great example of moving forward and looking through each of their program areas and coming up with what needs to be done to deal with the types of changes that have been discussed this morning."), thus leading to a very positive sentiment score for that agency (0.89). These statements were not counteracted by negative congressional response, but had they been, the algorithm would have yielded a lower score. In fact, a number of other witnesses chime in to agree with Mr. Seidel's assessment. We mean this as an example of positive language leading to a high sentiment score, but it also is an indication that committees express sentiment towards agency policy and performance by strategically calling sympathetic witnesses. One potential drawback of the sentiment method we use is that it is unable to weight sentiment by speaker and assumes a statement by a witness is of equal importance to a statement by a member of Congress.

Table B1 below shows the considerable cross-sectional variation by giving summary statistics for the sentiment directed at each agency in our data. Figure B1 shows that there is also a fair bit of variation in hearing sentiment over time, especially when the sentiment scores are aggregated by how often department agencies are called before Congress. We see here that most agencies get positive hearings and negative hearings each year and that, on the whole, congressional attention is rather neutral. We also see that some agencies get more positive attention than not (e.g., Department of Justice, Department of Health and Human Services, Department of Energy), but others regularly receive negative congressional attention (e.g., Department of Veterans Affairs).

Table B1: Targeted Sentiment, by Agency

Agency	Mean Sentiment	SD	Min-Max	Agency	Mean Sentiment	SD	Min-Max
United States Air Force	.043	.262	-.866-.899	Department of HUD	-.110	.172	-.716-.492
Department of Agriculture	.025	.213	-.783-.892	Broadcasting Board of Governors	-.022	.200	-.558-.580
Agency for International Development	.078	.186	-.550-.767	Inter-American Foundation	.136	.261	0-.801
United States Army	.017	.218	-.813-.784	Department of the Interior	.088	.252	-.855-.924
Federal Labor Relations Authority	-.008	.108	-.426-.367	Corporation for National and Community Service	.137	.177	0-.577
Merit Systems Protection Board	.537	.155	-.309-.805	Federal Election Commission	-.011	.206	-.610-.574
Defense Nuclear Facilities Safety Board	.030	.162	-.521-.409	Federal Maritime Commission	.026	.202	-.607-.728
Pension Benefit Guaranty Corporation	.269	.201	-.648-.706	National Science Foundation	.087	.198	-.801-.749
Office of Management and Budget	-.099	.251	-.821-.845	National Labor Relations Board	-.020	.206	-.719-.791
Department of Commerce	.064	.228	-.753-.867	National Mediation Board	-.060	.215	-.394-.626
Commodity Futures Trading Commission	-.175	.142	-.527-.434	National Aeronautics and Space Administration	.074	.243	-.820-.858
National Credit Union Administration	-.013	.141	-.500-.690	National Capital Planning Commission	.146	.206	-.282-.662
Department of Defense	-.003	.253	-.769-.917	National Archives and Records Administration	.011	.089	-.285-.561
Department of Justice	.293	.195	-.726-.881	Nuclear Regulatory Commission	-.014	.208	-.758-.856
Department of Labor	-.077	.235	-.777-.782	United States Navy	-.060	.216	-.788-.793
Department of Energy	.154	.247	-.791-.925	Office of Personnel Management	.037	.215	-.795-.856
Export-Import Bank of the United States	-.0008	.107	-.259-.393	Occupational Safety and Health Review Comm.	.418	.065	.230-.566
Department of Education	.060	.237	-.819-.861	Federal Retirement Thrift Investment Board	-.079	.274	-.734-.847
Equal Employment Opportunity Commission	-.031	.256	-.790-.733	Railroad Retirement Board	-.201	.192	-.542-.447
Environmental Protection Agency	.403	.121	-.594-.904	Small Business Administration	.069	.249	-.740-.868
Trade and Development Agency	.164	.167	0-.699	Securities and Exchange Commission	.043	.213	-.743-.771
Federal Communications Commission	.005	.220	-.742-.819	Consumer Product Safety Commission	.212	.213	-.900-.780
Federal Mediation and Conciliation Service	0	.	0-0	Department of State	.045	.250	-.870-.847
Federal Trade Commission	-.059	.245	-.834-.799	Social Security Administration	-.094	.230	-.884-.817
U.S. Office of Special Counsel	.176	.166	0-.413	National Transportation Safety Board	.078	.217	-.738-.779
Overseas Private Investment Corporation	.056	.131	0-.476	U.S. International Trade Commission	.002	.197	-.681-.541
U.S. Office of Government Ethics	-.226	.081	-.451-0	Department of Transportation	.104	.224	-.822-.849
General Services Administration	.024	.188	-.764-.774	Office of the U.S. Trade Representative	-.0006	.158	-.761-.563
Department of HHS	.596	.113	-.481-.886	Department of the Treasury	-.043	.206	-.822-.819
Federal Housing Finance Agency	-.110	.170	-.610-.276	Department of Veterans Affairs	-.398	.110	-.889-.488
Department of Homeland Security	.074	.256	-.789-.838	Commission on Civil Rights	-.065	.153	-.516-.523

Figure B1: Total Sentiment (Sum of Sentiment Scores * Number of Hearings), by Department
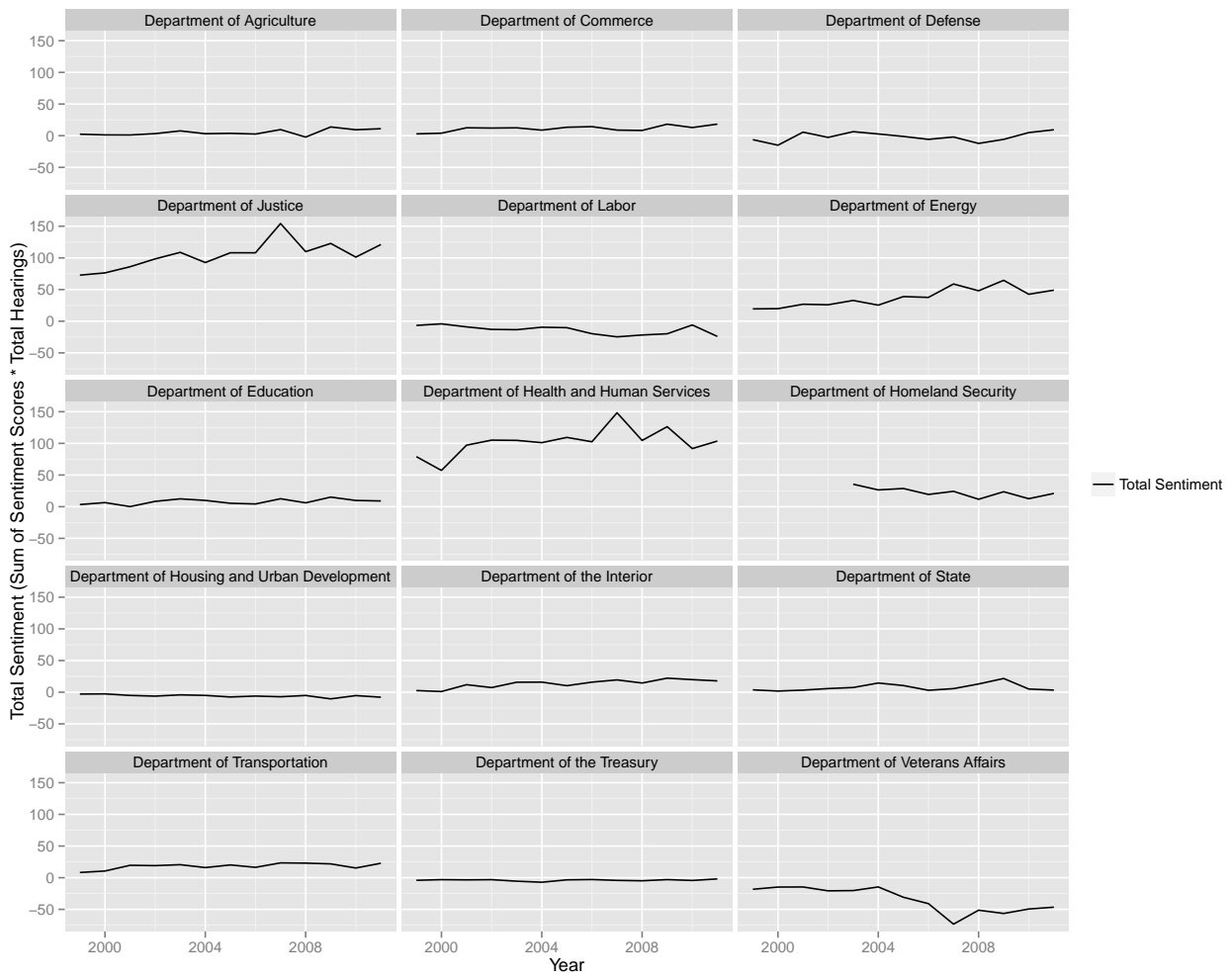
Appendix C - Description of Agency Characteristics Data

Agency characteristic data are available at `http://agencydata.wordpress.com`. Bertelli et al. (2015) used the following surveys to construct their measures:

- Merit Principles Surveys (administered by the Merit Systems Protection Board) from 2000 and 2005;

- Federal Human Capital Surveys (now known as Federal Employee Viewpoint Surveys; administered by the Office of Personnel Management) from 2004 2006, 2008, and 2010

- Reinventing Government Surveys (established by the National Partnership for Reinventing Government) from 1998, 1999, and 2000

The following questions were used to measure the agency autonomy and job satisfaction characteristics used in this paper. From Bertelli et al. (2015), Table 1:

Autonomy (see Figure 4 for how this varies over time in cabinet agencies)

- "I feel encouraged to come up with new and better ways of doing things."
- "Employees have a feeling of personal empowerment with respect to work processes."
- "Creativity and innovation are rewarded."
- "How satisfied are you with decisions that affect your work?"
- "How satisfied are you with your involvement in decisions that affect your work?"
- "I have been given more flexibility in how I accomplish my work."
- "Creativity and innovation are important."
- "In the past two years, I have been given more flexibility in how I accomplish my work."

Job Satisfaction (Overall) (see Figure 5 for how this varies over time in cabinet agencies)

- "Considering everything, how satisfied are you with your job?"
- "In general, I am satisfied with my job."
- "I would recommend the government as a good place to work."

Job (Compensation) Satisfaction

- "Considering everything, how satisfied are you with your pay?"
- "Overall, I am satisfied with my current pay."
- "Overall, I am satisfied with my pay."

Not all agencies are represented in all surveys, so the agency-year dataset of agency characteristics has a total number of 573 observations. See Table A1 (appendix A) for an indication of which agencies have enough survey responses to be included in our analyses.

Appendix D - Description of Media Attention Data

We collected *Washington Post* news stories using the *Lexis Nexis Academic* database, employing keyword searches for agency names for each year from 1999 through 2011 to match up with the oversight and morale data. This was mostly a straightforward process, with agencies clearly and unambiguously identified by their names. Yet, there were times where we searched a common acronym, as well as the official agency name, taking care to remove duplicate articles from the dataset. For example, the National Aeronautics and Space Administration is more commonly known as NASA, so we obtained many of this agency's media coverage by searching for the acronym. Once we downloaded these data[1], we prepared each file (each file contains one news story) for the computer-assisted text analysis, by removing quotation marks and special characters.

Using the same approach as above with the hearing transcripts (appendix B), we measured the *targeted* sentiment of articles, discounting positive/negative words that appear far from the target phrases, e.g., "Department of Agriculture," "Office of Management and Budget," "Securities and Exchange Commission," etc. As mentioned in the text, to capture both sentiment and volume of *Washington Post* coverage, we simply sum the sentiment scores by agency and year. Figure D1 below displays this agency-year aggregate score for each department agency over time. The correlation (Pearson's r) between news sentiment and hearings sentiment scores is 0.56.

---

[1]We collected a total of 106,554 stories, totaling 286,094 pages.

Figure D1: Total *Washington Post* Sentiment, by Department
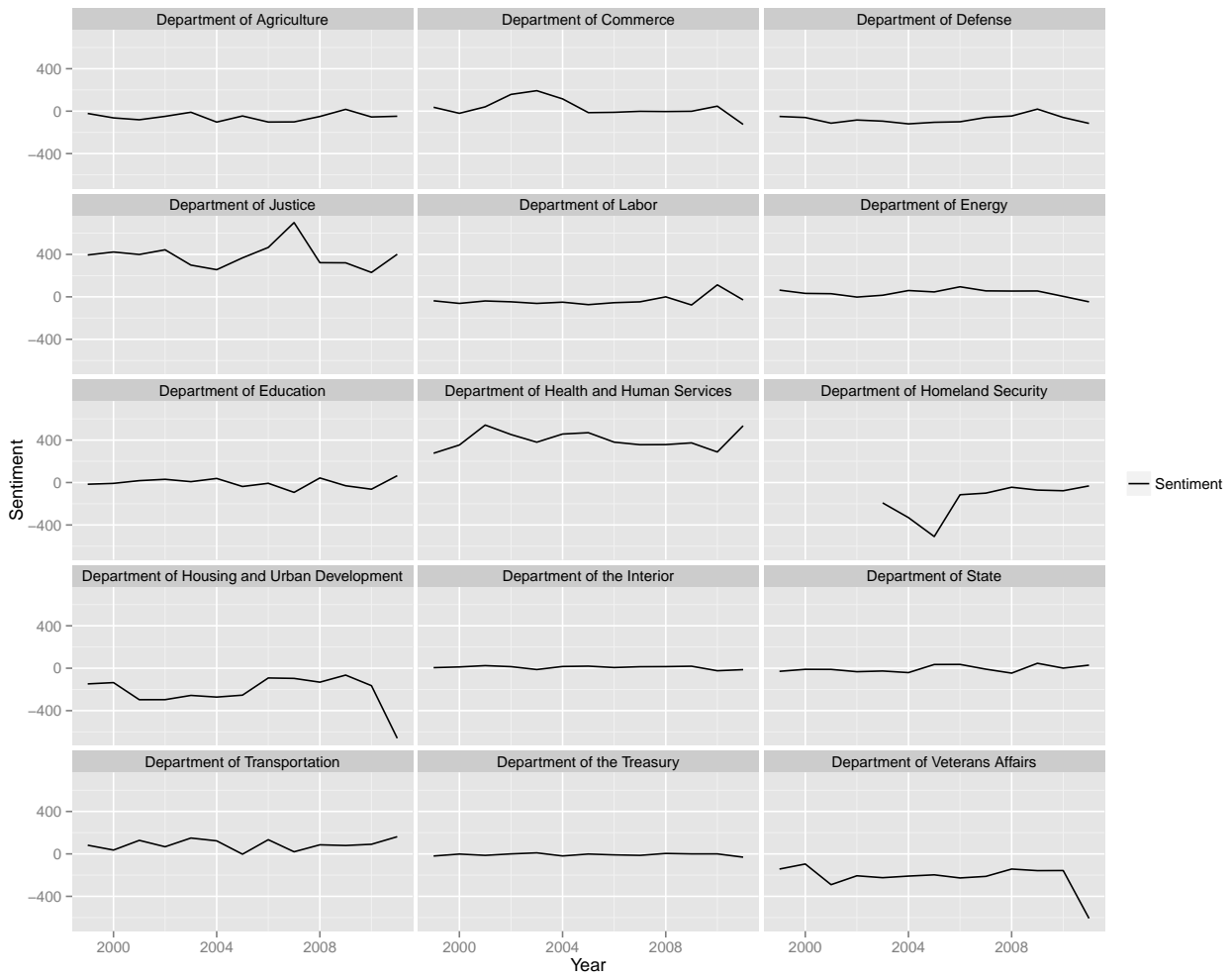
Appendix E - Instrumental Variables Results

Table E1. 2SLS Models of Agency Autonomy and Job Satisfaction (Overall), 1999-2011

	(1) Autonomy Δ	(2) Satisfaction Δ	(3) (1) with Interaction	(4) (2) with Interaction
Oversight Hearings (Lag)	-.00273*	-.00312**	-.00203**	-.00191**
	(.00140)	(.00140)	(.00086)	(.0086)
Hearings Sentiment (Lagged mean)			-.52879	-.20510
			(.32124)	(.20108)
Oversight Hearings (Lag)× Hearings Sentiment (Lagged mean)			.01531***	.00777***
			(.00542)	(.00333)
Divided Government	.03392	.02897	-.00427	-.00645
	(.03408)	(.03390)	(.02098)	(.02095)
Republican Control of Congress	.07748*	.08344**	.03712	.03974
	(.04032)	(.04011)	(.02481)	(.02478)
Democratic President	.00631	.00484	-.01583	-.01648
	(.03313)	(.03291)	(.02039)	(.02034)
Presidential Transition Year	.03951	.04087	.00349	.00409
	(.03594)	(.03570)	(.02212)	(.02206)
Presidential Attention	-.00004	-.00001	.00069	.00071
	(.00121)	(.00121)	(.00075)	(.00075)
Non-oversight Hearings (Lag)	.01426***	.01529***	-.00026	.00019
	(.00541)	(.00539)	(.00333)	(.00333)
Total *Washington Post* Sentiment (Lag)	-.00001	.00005	-.00011	.00001
	(.00028)	(.00017)	(.00028)	(.00017)
Constant	-.04961	-.04340	.02883	.03155
	(.04600)	(.04575)	(.02831)	(.02827)
Agency FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	470	470	470	470
R^2	.545	.678	.552	.681

$*p < 0.10$, $**p < 0.05$, $***p < 0.05$

Note: Entries are two-stage least squares regression coefficient estimates and standard errors, clustered by agency. First stage results are available on request. The dependent variables are created by calculating the change in the Bertelli et al. (2015) measures of autonomy and job satisfaction (excluding compensation questions) from time $t-1$ to time t. Agency and year fixed effects are included in all models but not reported. See appendix A for further description of the oversight data.