

Controlling for Oversight in Studies of Oversight

Jenn S. Kim*

Benjamin Kinnard[†]

John Patty[‡]

May 16, 2025

Abstract

In honor of the creation of the Congressional Oversight Research Database (CORD), we explore some of the selection effects that the data in this resource may be subject to. We present two simple, related principal-agent models and discuss how they offer different comparative statics, with a focus on the occurrence of oversight of the agent by the principal. Within these models, we also explore the issue of timing of oversight, which raises classic political economy questions, such as credibility of commitments, dynamic consistency, and the origins of accountability. We then discuss various forms of selection effects that may affect empirical studies of congressional oversight and empirically test the effect of policy disagreement on oversight using the data contained in CORD, drawing implications about our models for the study of congressional oversight.

1 Introduction

Accountability is a central concept in the study of democratic politics. From theoretical questions such as how and what kind of institutions can incentivize politicians and civil servants to make policy choices in line with the “will of the people,” to empirical questions such as whether and how individuals have sufficient information and/or can make “rational” decisions that maintain these incentives, the concept of accountability is nearly ubiquitous in the study of political science.

It is well known that achieving accountability can, if not almost always, require **oversight**. This has been acknowledged in the United States since the nation’s founding. Indeed, the historical evolution of oversight in the US Federal Government¹ reflects the social, economic, and

*Ph.D. Candidate, Department of Political Science, Emory University. seoyeon.kim@emory.edu

[†]Ph.D. Candidate, Department of Political Science, Emory University. benjamin.kinnard@emory.edu

[‡]Professor of Political Science and Quantitative Theory & Methods, Emory University. jwpatty@gmail.com

¹For a small sample of works on this topic, see Ogul (1976), Arnold (1979), Oleszek (2010), and Gailmard and Patty (2012).

political evolution of the United States. Accordingly, the newly developed and emerging Congressional Oversight Research Database (CORD) presents a new opportunity for the qualitative and quantitative study of the evolution of the United States itself.

Our goal in this article is to explore some simple theoretical models of oversight with the goal of deriving and presenting empirical implications of these models. One central theme of the models is that oversight activities are inherently subject to several **selection effects**. Some simple examples are useful at this point. Consider measuring the efficacy of your city’s fire department. What is the “most effective” fire department? Consider the following two possible definitions:

A. Best Fire Fighters: Put fires out as quickly and safely as possible.

B. Best Fire Preventers: Minimize the number of fires that occur.

It can be difficult (or impossible) to measure the effectiveness of a fire department, under notion *A*, if it scores highly under notion *B*. While this may seem like a silly example, it is useful as a thought experiment. To make it tangible, suppose that the citizens of your town collectively decide to provide incentives to the fire department to make decisions in line with notion *A*. At least at the margins (again, this is simply a thought experiment), providing such an incentive can easily incentivize the department to actively undermine its effectiveness under notion *B*. Similar thought experiments can be constructed with application to policing, medical care, safety regulation, and many other practical realms, particularly when the notion of a fire — the “bad” to be prevented — is the product of strategic individual behavior.

Matters become even more complicated when one recognizes that oversight in practice usually consists of multiple layers, and often circular. Citizens seeking to prevent crime, for example, must hold their politicians accountable for holding law enforcement agencies and their employees accountable for holding individual citizens accountable for their individual behaviors. In this article, we “try to crawl before we try to run,” and mostly sidestep this level of complication.² Instead,

²For recent theoretical contributions exploring this vein, see Ashworth, Bueno de Mesquita and Friedenberg (2018), Gailmard and Patty (2019), and Alexander (2024).

we will consider a few classic principal-agent settings as our central accountability problem. They share a common structure, to which we now turn.

2 The Theoretical Setup

Our framework involves two players: a **principal** and an **agent**. The agent possesses private information and chooses an action, and the principal's preferences over the agent's choice of action depends on the agent's private information. This, indeed, is the basis of the *agency problem* faced by the two players: the agent has his or her own preferences over the actions, which may also depend on the agent's private information.³ The two models we consider in this article differ with respect to the nature of the “carrots and sticks” that the principal can choose from to alter the agent's *induced preferences* over the actions. This will become clearer when we describe the two models individually below in this section.

Who is the Principal, and Who is the Agent? As alluded to above, oversight is often a multi-layered concept in practice. Fortunately, the scope of CORD and the framing of this issue help narrow our focus in useful ways. CORD contains reports about Congress's investigations into the executive branch. In many cases, these reports are explicitly about the functioning and effectiveness of federal agencies. Some reports detail the committee in question's investigations and findings about larger phenomena, such as the 2007-08 financial crisis,⁴ but many of these can be thought of as oversight of multiple agencies. Accordingly, for the purposes of discussion, it is reasonable to conceive of the agent as representing a bureaucratic agency.

³For simplicity of presentation, we consider only settings in which the agent's preferences over actions are independent of the agent's private information, but our analysis can relax this.

⁴For example, the October 6th, 2008 report issued by the minority (then GOP) staff of the House Committee on Oversight and Government Reform, “Examining the Causes of the Credit Crisis of 2008.”

2.1 Auditing to Deter

Many policy decisions have significant irreversible effects. We refer to oversight designed to induce the agent to not make such choices as the **deterrence model**. We assume that agent and principal's policy preferences are always in conflict, *but only the agent will know when this is the case*. In other words, the agent's best response is possibly what the principal would have the agent "do" in that situation, but not always, which leads to a *credibility problem* between the agent and the principal.

In the auditing to deter model, the agent is assumed to dislike oversight (or, **auditing**) *per se*. Without this (fairly uncontroversial) assumption, it is nearly impossible for the prospect of the principal auditing the agent to affect the agent's incentives. The deterrence model is wide-ranging and flexible. We suppose that the principal observes a noisy, but informative, signal about the policy that the agent actually chose and then chooses whether to audit the agent. Because we are focusing on auditing solely to *deter* bad choices by the agent, we assume that the principal cannot actually reverse the policy decision after an audit. Rather, the differential conditional probabilities of the two signals that the principal might observe being audited offer a direct incentive to the agent to choose the "correct policy" in equilibrium.⁵

The imprecision of the principal's information about the agent's policy choice can emerge through a wide array of different mechanisms. For example, an agent might issue a formal "directive" that can be directly observed by the principal, but the agent's actual policy choice (*e.g.*, in terms of implementation of the directive) might be less precisely observed.⁶ On the other end of the *de jure/de facto* spectrum of principal-agent problems, the principal might be uncertain about the agent's policy choice because there was no formal "directive" from the agent and the principal instead must infer the policy choice through alternative mechanisms (*e.g.*, media and/or lobbying ((*e.g.*, Kollman (1997), Gordon and Hafer (2005, 2007) ?)).⁷

⁵Throughout, our notion of equilibrium is Perfect Bayesian.

⁶For example, consider the classic study of the United States Forest Service offered in Kaufman (2006).

⁷There is an interesting link between these two ends of the spectrum and the seminal arguments in Aghion and Tirole (1997). Unfortunately, it lies outside of our main focus in this article, so we leave this as a possible direction for future work.

2.2 Auditing to Correct

The second model of oversight that we consider is similar to the auditing to deter model. The two differences are:

1. The agent pays no direct cost from being audited, and
2. The principal “corrects” the agent’s policy choice if necessary when the principal audits.

The agent’s optimal strategy in this setting is quite simple: he or she should simply set his or her most-preferred policy. This induces what is sometimes referred to as a “bailout problem” for the principal. This type of effect emerges in many oversight models, but we have chosen the simplest one in which it emerges.

2.3 Oversight & Credible Commitments

In addition to comparing two related models of oversight, we also consider two different timings in each of the models. The first timing, which we refer to as **Stackelberg timing**, involves the principal publicly committing to an auditing strategy, after which the agent makes his or her policy choice. Under this timing, the principal chooses his or her auditing strategy as a function of her prior belief about the agency’s willingness to set the “correct” policy. The agency perfectly knows whether he or she will face an oversight after his or her policy choice. In the second timing, which we refer to as **Cournot timing**, the order is reversed: the agent makes his or her policy choice, after which the principal observes a signal about the chosen policy and chooses whether to audit the agent. In this setting, the principal cannot commit to an auditing strategy; the agency knows that the principal will make his or her auditing decision *after* observing some signal about the agency’s policy choice.

The two timings often lead to different equilibrium predictions, particularly in the auditing to deter model. Specifically, in the auditing to deter model, the principal has nothing to gain from auditing the agent in the Cournot timing and, because the agent realizes this, oversight is completely ineffective. In both models, the principal can do no worse in equilibrium within the

Stackelberg timing than he or she would do in equilibrium in the Cournot timing. This is a classic and well-known result—it highlights the importance of *credible commitments*. By construction, any commitment that the principal makes in the Stackelberg timing *is* credible. Thus, the principal can in principle commit to the strategy that he or she would use in the Cournot timing in the Stackelberg case as well, while the converse does not hold, because the Stackelberg timing does not impose a sequential rationality constraint on the principal’s choice of strategy in equilibrium.⁸

In substantive terms, each of the two timings can be reasonable. For example, if the oversight is being done by Congress, *per se*, the Cournot timing is arguably more realistic because Congress is sovereign and is not “required” to audit any particular agency or program. On the other hand, if the audit is being conducted by a court to ensure that the agency in question followed the law, the Stackelberg timing is more realistic. Indeed, the benefits we find for the overseer in the Stackelberg timing are essentially a common argument for why Congress benefits from an independent judiciary (*e.g.*, Salzberger (1993), Stephenson (2003), Helmke and Rosenbluth (2009)). We now turn to discuss the equilibria for each timing in the two oversight models.

3 Four Oversight Equilibria

We now discuss the oversight models in order, first considering the Stackelberg timing and then the Cournot timing for each. Regardless of the model or timing, we denote P ’s strategy by $\alpha(y) \equiv \Pr[a = 1 \mid y]$ for each signal $y \in \{0, 1, \varphi\}$, and denote the strategy used by A by $\sigma(\omega) \equiv \Pr[x = \omega \mid \omega]$ for each state $\omega \in \{0, 1\}$. An overview of the two timings is provided in Table 1.

With the timings in hand, we now describe the common components of the oversight models.

Players, Actions, Information, and Sequence of Play. There are two players in each model: an **agent**, A , and a **principal**, P . The agent will observe a **state of nature**, $\omega \in \Omega = \{0, 1\}$. The state

⁸This is analogous to the Bayesian persuasion framework, where the principal can commit to any information generation mechanism, including more specific forms of communication such as cheap talk. As in our model, the principal can do no worse with commitment than he or she would do without commitment.

| Stackelberg Timing | Cournot Timing |
|---|--|
| <ol style="list-style-type: none"> 1. P commits to an auditing strategy, α, 2. A observes state of nature, ω, 3. A chooses policy, x, 4. P observes signal, y, and 5. Audit decision, a, generated by $\alpha(y)$. | <ol style="list-style-type: none"> 1. A observes state of nature, ω, 2. A chooses policy, x, 3. P observes signal, y, and 4. P chooses audit decision, a. |

Table 1: Two Oversight Timings: Stackelberg and Cournot

of nature, ω , is drawn according to a commonly known probability mass function, F :

$$F(0) = p,$$

$$F(1) = 1 - p,$$

for an exogenous and commonly known probability $p \in (0, 1)$. After observing ω , A then chooses a **policy**, $x \in X = \{0, 1\}$. After the agent chooses the policy, x , the principal observes a **signal**, $y \in Y = \{0, 1, \varphi\}$, which is potentially informative about x . The signal, y , is drawn from a pair of commonly known conditional distributions, $G_0, G_1) : X \times \Omega \rightarrow \Delta(Y)$:⁹

$$G(\varphi|x) = \nu \quad \text{for both } x \in \{0, 1\},$$

$$G(0|x = 0) = (1 - \nu) \cdot q,$$

$$G(1|x = 0) = (1 - \nu) \cdot (1 - q),$$

$$G(0|x = 1) = (1 - \nu) \cdot (1 - q),$$

$$G(1|x = 1) = (1 - \nu) \cdot q,$$

where $\nu \in [0, 1]$ and $q \in [1/2, 1]$ are each exogenous and commonly known parameters. The signal $y = \varphi$ represents “no information” being revealed to P , and this occurs with probability ν ,

⁹The notation $\Delta(Y)$ represents the set of all lotteries over Y .

regardless of A 's choice of policy, x . Similarly, $q \in (1/2, 1]$ is the conditional probability of an “informative signal” being correct. Accordingly, P perfectly observes A 's choice with certainty if and only if $\nu = 0$ and $q = 1$.¹⁰ Finally, after observing y generated by G_x , the agent is either *audited* ($a = 1$) or not ($a = 0$). Referring to Table 1, how this auditing decision is “made” is the distinction between the Stackelberg and Cournot timings.

Policy Payoffs. We denote player $i \in \{A, P\}$'s **policy payoff** from policy choice x , given state of nature ω , by $v_i(x \mid \omega)$, and consider the following family of policy payoff functions:

$$\begin{aligned} v_A(x \mid \omega) &= \eta \cdot x, \\ v_P(x \mid \omega) &= -(x - \omega)^2, \end{aligned}$$

where $\eta > 0$ is a commonly known exogenous parameter. The parameter η captures the degree to which A 's preferences differ from those of P . These policy payoffs and information structure describe a classic **moral hazard** situation: the agent always prefers choosing $x = 1$, but the principal wants $x = \omega$. In other words, the agent does not have any *policy* incentive to act in an accountable fashion.

We now describe our two models of oversight.

3.1 Auditing to Deter

In our version of the auditing to deter model, the players' payoff functions are as follows:

$$\begin{aligned} u_A(x, a \mid \omega) &= v_A(x \mid \omega) - \pi \cdot a, \\ u_P(x, a \mid \omega) &= v_P(x \mid \omega) - c \cdot a, \end{aligned}$$

¹⁰We restrict attention to $q > 1/2$ for simplicity: because q is common knowledge, the case where $q < 1/2$ is symmetric to this setting.

where $\pi > 0$ represents A 's **penalty** from being audited and $c > 0$ represents P 's **cost** of auditing A . Throughout, we assume that π and c are common knowledge. Substantively, the agent, A , always wants to avoid being audited, and always wants to choose $x = 1$. The principal always prefers that A choose $x = \omega$, and would like to audit as rarely as possible.

The derivations of the formal model are presented in Appendix A. In all cases, we refer to the equilibrium as achieving **perfect accountability** if the agent's equilibrium strategy is to always choose the principal's preferred policy (in other words, when P has "solved" the moral hazard problem in equilibrium).

Proposition 1 *In the Stackelberg timing of the auditing to deter model, P audits only after observing $y = 1$, and audits $y = 1$ with positive probability if and only if*

$$\begin{aligned} \eta &\leq \pi \cdot (1 - \nu) \cdot (2q - 1), \text{ and} \\ p &\geq \frac{c \cdot (1 - \nu)q}{c \cdot (1 - \nu)(2q - 1) + 1}. \end{aligned}$$

Perfect accountability is achieved if and only if P audits with positive probability in equilibrium, and the probability that A is audited after $y = 1$ is equal to

$$\alpha^*(1) = \frac{\eta}{\pi \cdot (1 - \nu) \cdot (2q - 1)}. \quad (1)$$

Proof: All proofs of numbered results in the body of the article are contained in Appendix A. ■

In this setting, P will audit in equilibrium only when it achieves perfect accountability. P is more likely to audit the agent in equilibrium for small costs of auditing (*i.e.*, c), high precision of the signal s (*i.e.*, q), or high probability that the agent's interim policy preferences differ from the principal's (*i.e.*, $p \equiv \Pr[\omega = 0]$). The following proposition summarizes the comparative statics of auditing in equilibrium in the Stackelberg timing of the auditing to deter model.

Proposition 2 *In the Stackelberg timing of the auditing to deter model, auditing is more likely in equilibrium for*

1. *Higher values of $p \in (0, 1)$,*
2. *Lower values of $c > 0$,*
3. *Higher values of $q \in (1/2, 1)$,*
4. *Lower values of $\eta > 0$,*
5. *Lower values of $\nu \in (0, 1)$, and*
6. *Higher values of $\pi > 0$.*

These comparative statics, some of which are arguably counterintuitive, can be explained as follows. The first comparative static (with respect to p) is intuitive: p is the probability that the agent will prefer to make a choice different from the policy that the principal would choose. The second comparative static (with respect to c) is similarly intuitive, because c is P 's marginal cost of auditing. Figure 1a illustrates these effects of p and c . The third comparative static (with respect to q) is slightly less obvious. It reflects the fact that higher values of $q > 1/2$ imply that the probability that the probability that an audit will occur in equilibrium looms larger in A 's strategic calculations when considering what policy to choose *when a moral hazard problem arises*, because $q \cdot (1 - \nu)$ is the probability that a choice of $x = 1$ will generate a signal of $y = 1$ increases.

The fourth comparative static (with respect to η) describes the effect of the alignment of the agent's policy preferences with those of the principal. When η is close to zero, P can achieve perfect accountability with a very small probability of auditing after $y = 1$. As η grows, the minimal auditing probability after $y = 1$ required to sustain perfect accountability increases as well, eventually requiring a "probability" greater than 1, at which point only baseline accountability can be achieved in equilibrium. The fifth comparative static (with respect to ν) follows from the fact that P will never find it profitable to audit when y is uninformative ($y = \varphi$): as ν increases, A will eventually no longer have sufficient incentive to choose $y = 0$ when $\omega = 0$, and only baseline accountability will be achievable in equilibrium. The final comparative static (with respect to π) is intuitive: as the penalty incurred by the agent from being audited (π) increases, the principal

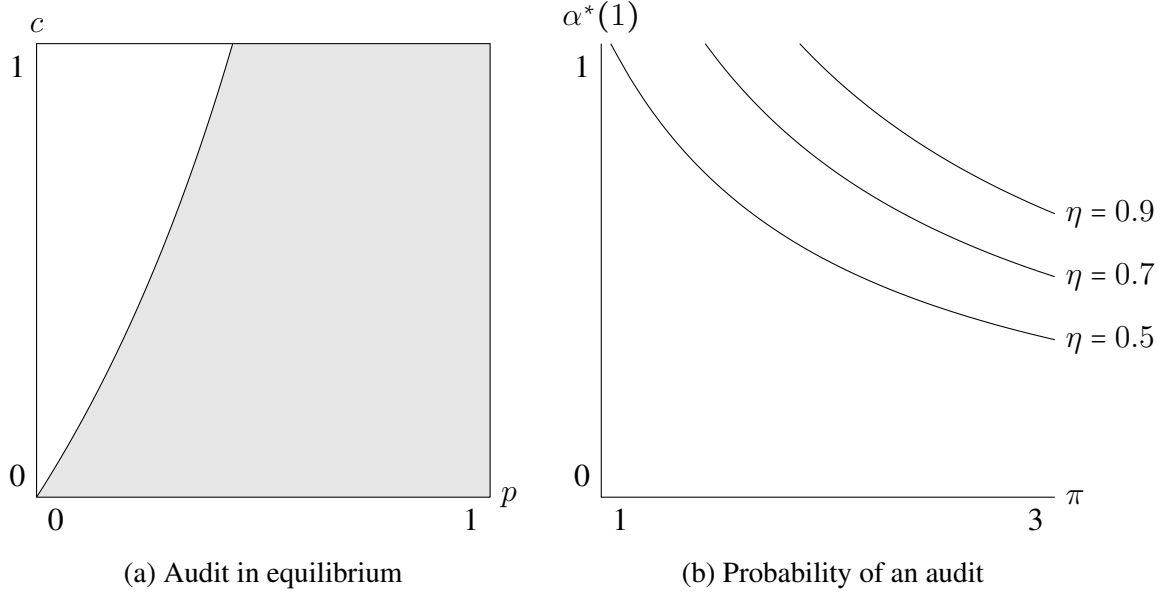


Figure 1: Auditing to Deter, Stackelberg Timing

can sustain perfect accountability with a lower probability of auditing $y = 1$. Figure 1b further illustrates the effects of η and π on the probability that the principal audits in the equilibrium.

3.1.1 Cournot Timing

In the Cournot timing, the principal's auditing strategy, $\alpha : \{0, 1, \varphi\} \rightarrow [0, 1]$ must satisfy sequential rationality for P in equilibrium, which implies that

$$\alpha^*(0) = \alpha^*(1) = \alpha^*(\varphi) = 0,$$

in equilibrium, because auditing is strictly dominated for P in the Cournot timing of the auditing to deter setting. Intuitively, and unfortunately from P 's perspective, this *never* achieves perfect accountability. Accordingly (and intuitively), P *always* weakly benefits from being able to commit to a strategy (*i.e.*, the Stackelberg timing).

Proposition 3 *In the Cournot timing of the auditing to deter model, neither auditing nor perfect accountability occur in equilibrium.*

We will see below that this is not true under either timing when P is auditing to correct rather than to deter. We now turn to the auditing to correct model, beginning with the Stackelberg timing.

3.2 Auditing to Correct

In the auditing to correct model, we suppose the players' payoff functions are as follows:

$$\begin{aligned} u_A(x, a \mid \omega) &= (1 - a) \cdot v_A(x \mid \omega) + a \cdot v_A(x_P^*(\omega) \mid \omega), \\ u_P(x, a \mid \omega) &= (1 - a) \cdot v_P(x \mid \omega) - c \cdot a, \end{aligned}$$

where $c \geq 0$ again represents P 's cost of auditing A (which is again assumed to be common knowledge). The change in the payoffs represents a change in “what audits result in.” In the auditing to deter model, the only role oversight plays is to punish A when P believes A might have chosen $x = 1$ when $\omega = 0$. In the auditing to correct model, oversight does not “cost” A anything, but it results in the final policy being equal to P 's ideal policy (conditional on the true state, ω).

3.2.1 Stackelberg Timing

The agent's strategic calculations are simple in this model: A should *always* choose $x = 1$, because $x = 0$ is weakly dominated for A after either state of nature, ω . The principal's strategic calculations are also simple: P should audit *all* signals $y \in \{0, \varphi, 1\}$ if and only if $p \geq c$. This is stated formally in the next proposition.

Proposition 4 *In the subgame perfect Nash equilibrium of the Stackelberg timing of the auditing to correct model, A always chooses $x = 1$, regardless of ω :*

$$\sigma^*(\omega) \equiv \Pr[x = 1 \mid \omega] = 1 \quad \forall \omega \in \{0, 1\}.$$

and P audits all signals $y \in \{0, \varphi, 1\}$ with equal probability:

$$\alpha^* = \begin{cases} (0, 0, 0) & \text{if } p < c, \\ (1, 1, 1) & \text{if } p \geq c. \end{cases} \quad (2)$$

In the Stackelberg timing of the auditing to correct model, the principal's optimal auditing strategy is always a corner solution and, more importantly, *the optimal auditing decision is independent of the signal the principal observes*. Note that equilibrium auditing in this model is qualitatively different from in the auditing to deter model. In the Stackelberg timing of that model, P *never* audits $y = 0$ or $y = \varphi$ in equilibrium. Furthermore, when P audits $y = 1$, the auditing probability is (generically) less than 1. As in the model of auditing to deter, Figure 2 shows that the audit is more likely in equilibrium for sufficiently large p and sufficiently small c . We now turn to the Cournot timing of the auditing to correct model.

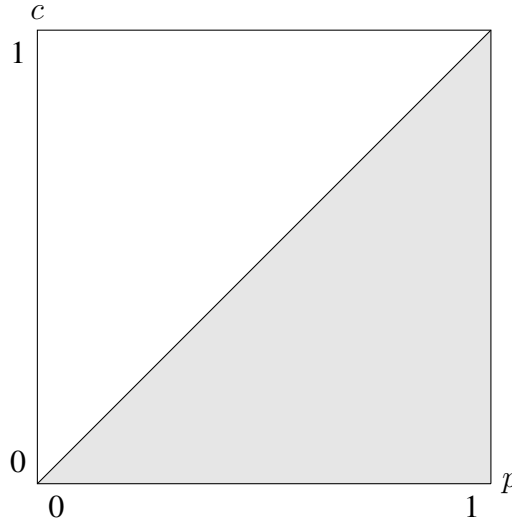


Figure 2: Audit in Equilibrium (Auditing to correct, Stackelberg Timing)

3.2.2 Cournot Timing

In the Cournot timing, the agent should still always choose $x = 1$, regardless of ω . Accordingly, P 's equilibrium strategy in this case is identical to P 's equilibrium strategy in the Stackelberg

timing.

Proposition 5 *Perfect accountability is achieved in the Cournot timing of the auditing to correct model if and only if it is achieved in the Stackelberg timing (i.e., if $p > c$), and the principal's equilibrium auditing strategy is identical (Equation (2)).*

In other words, a principal facing the challenge to “audit to correct” as opposed to “auditing to deter” does not gain anything (at least in instrumental terms) from having the ability to commit to an auditing strategy. This is an important difference between the two forms of auditing. The following proposition summarizes the comparative statics in both timings of the auditing to correct model.

Proposition 6 *In both timings of the the auditing to correct model, auditing is more likely to occur in equilibrium for*

1. *Higher values of $p \in (0, 1)$, and*
2. *Lower values of $c > 0$.*

Furthermore, the equilibrium auditing probabilities are insensitive to q and ν .

Comparing Proposition 6 with Proposition 2, it is clear that auditing to deter has “more” comparative statics: equilibrium behavior in the auditing to correct model depends entirely on the relationship between p and c . Similarly, comparing Propositions 5 and 3 demonstrates another distinction between the two models with respect to the impact of timing. In the Cournot timing of the auditing to deter model, auditing never occurs in equilibrium (unlike the Stackelberg timing of that model), whereas auditing can occur (and under identical conditions) in the two timings of the auditing to correct model. We now turn to the broader empirical implications of these models and timings for studying oversight.

3.3 Empirical Implications

Proposition 2 provides six potentially testable implications for auditing to deter under the Stackelberg timing, and Proposition 6 provides two such implications for both timings of the auditing to correct model. The two implications of Proposition 6 are contained in Proposition 2, so they are useful within this space of four models/timings only to the degree that they might rule out the Cournot timing of the auditing to deter model. The four implications of Proposition 2 not included in 6 can be used to rule out the auditing to correct model.

However, leveraging any of our comparative statics results in empirical work requires choices about measurement. Using Proposition 2 as the backdrop because it contains the maximal set of testable implications, a cursory list of possible examples is presented in Table 2. We have broken out the examples with respect to p by the “party in control,” because this reflects something about “spatial ideology.” Similarly, these are the only ones to which we have assigned “agencies,” *per se*, because agencies are commonly viewed as having a ‘secular ideology’ independent of the partisanship of the President.

| Parameter | Low Example | High Example |
|-----------|--|-------------------------------|
| p | Dem: Dept of Labor GOP: Dept of Defense | Dem: SEC GOP: EPA |
| c | Taxes/Unpopular Programs | Entitlements/Popular Programs |
| q | Consumer Issues | Eminent Domain, Anti-Trust |
| η | National Defense | Public Interest Regulation |
| ν | Appropriations/Budgeting | Intelligence Gathering/Spying |
| π | Courts/States & Cities | Discretionary Programs |

Table 2: Some Examples for Comparison through Comparative Statics

4 Oversight & Selection Effects

The series of oversight models above showcase the importance of credible commitment and the ‘type’ of auditing being conducted on the observable outcomes of oversight, such as committee hearings, records of private communications between members of Congress and agencies, and

oversight reports. In the following section, we discuss some of the various selection effects that affect empirical research on congressional oversight and how they connect with our models of oversight. Though a full test of all of the comparative statics of the oversight models falls outside of the scope of this paper, we also present a series of empirical tests of p , which can be substantively interpreted as the likelihood of the agent wanting to select a different policy than the principal, i.e. policy preference divergence. The predictive power of preference divergence, particularly ideological divergence, on oversight intensity is an unsettled debate in the congressional oversight literature, a debate that we discuss more in the following section preceding the empirical analysis. We also present some descriptive statistics of CORD which showcase the diversity of the congressional oversight agenda and its potential in future empirical research on congressional oversight.

4.1 Selection Effects in the Empirical Oversight Literature

Much of the empirical literature on congressional oversight focuses on the decision of Congress to opt-in to oversight, but data and endogeneity challenges make analysis of agency responsiveness to congressional policy preferences difficult.¹¹ To observe the principal (i.e. Congress) correcting the policy choices of the agent (i.e. an agency), the agent has to make an observable policy decision that is incongruent with the preferences of the principal and the principal has to make an observable change to the policy. Although this does indicate that there is some accountability happening, this does not satisfy perfect accountability as it is described in the oversight models from earlier, where the agent's equilibrium strategy is to always select the principal's preferred policy. The issue with perfect accountability is that it is often observationally equivalent to a situation where the principal chooses not to conduct any oversight because the principal cannot correct the agent's policy choice, cannot induce accountability out of the agent in the future, or simply views the cost of oversight as exceeding the potential benefits gained from it. This is not a novel contribution, of course—many

¹¹For some of the recent advances in the oversight literature in linking congressional oversight to policy outcomes, see Mills, Kalaf-Hughes and MacDonald (2016), Neiheisel and Brady (2017), Ritchie and You (2019), and Ban and Hill (2025).

scholars have already examined this observational equivalence in a variety of ways, including with formal models and qualitative case studies (Weingast and Moran (1983), McCubbins and Schwartz (1984), Aberbach (1990), Lupia and McCubbins (1994), among others).

Libgober’s model of notice-and-comment rulemaking (Libgober (2020))¹² neatly presents the dilemma these selection effects present for empirical research—in his model, commenting on an agency’s policy proposal (which can be viewed as an attempt to correct a policy choice, potentially by a principal such as Congress) can only be effective if the agency is neither totally independent from the commenter (i.e. no accountability in equilibrium) nor completely captured by the commenter (i.e. perfect accountability in equilibrium). Our correction model also captures a similar effect through auditing costs instead of agent independence—if the cost c of an audit to the principal outweighs the potential benefits gained from correcting the agent’s policy choice, the audit does not occur, despite there being potentially substantively high disagreement between the agent and the principal, thus presenting the possibility of observationally equivalent perfect accountability and no accountability. Empirical research on congressional oversight and agency responsiveness can thus often be limited in scope to situations where there are moderate levels of independence of the agency from Congress, with the extreme cases of high or low independence being difficult to empirically distinguish, and situations where the cost of oversight to Congress is sufficiently low relative to potential policy benefits from oversight.

Another form of selection effects that can affect the empirical study of congressional oversight is how the different modes of congressional oversight (hearings, reports, private communications, etc.) connect with one another. Selin and Moore describe congressional oversight as “a series of actions designed both to gather information on the executive and to facilitate legislative credit claiming and position taking” (Selin and Moore (2023), 187). Oversight is not just one action taken at one period of time—it is a process that involves multiple observable actions that occur at different periods of time, all with different implications for what factors may affect the intensity of oversight. Figure 3 showcases a rough and non-exhaustive ‘timeline’ of congressional oversight,

¹²For recent closely related work, see Gailmard and Patty (2017) and Bills, Carroll and Rothenberg (2024).

with later phases of oversight typically being more public, though this should not be taken as a strict ordering or a strict one-directional process—some of these steps can occur in different orderings, such as a private communication between a member of Congress and an agency spurring a member of Congress to request an investigation to be conducted by either the Government Accountability Office (GAO) or the agency’s Inspector General (IG).

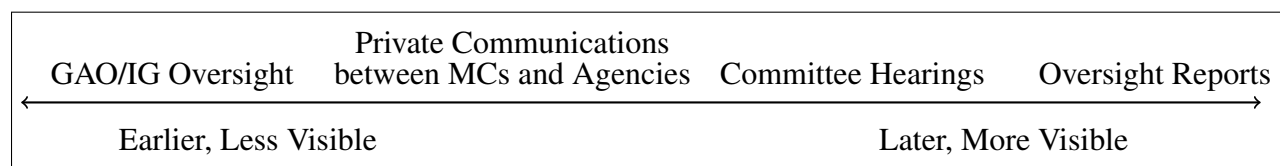


Figure 3: Congressional Oversight as a Process

Members of Congress select into conducting more oversight at each step of this ‘process’, meaning that the observable production of oversight reports is conditioned on there being a committee hearing, which is itself conditioned on members of Congress deciding to conduct a hearing with the information gathered from themselves, their staff, and oversight support organizations such as the GAO and Offices of Inspector General. There are a few reasons why members of Congress may want to either stop at an earlier phase or continue on with oversight, the most notable of which is whether the agency responds to Congress’s preferences at an earlier stage of oversight before, say, a costly committee hearing is held that may damage the reputation of an agency. This ties back in with the previous selection effect of the difficulty in empirically distinguishing accountability and no accountability—if an agency corrects a policy (like in our correction model) in response to a private email from a member of Congress, accountability is achieved, but this accountability cannot be distinguished from no accountability if a scholar examines a mode of oversight that typically occurs later in the process, such as committee hearings.

Oversight scholars have been making significant advances in empirically analyzing these earlier, more private forms of oversight, however, uncovering what drives selection into the later stages of oversight in the process. Some of the new modes of congressional oversight that have been studied include non-statutory directives in appropriations reports (Bolton (2022)), strategic

staffing and appointments (Patty and Moore (2023)), limitation riders in appropriations bills (Macdonald (2010)), the aforementioned private communications between members of Congress and federal agencies (Lowande (2018), Ritchie (2018)), comments on federal rules (Lowande and Potter (2021)), legislative structuring of judicial review (Patty and Turner (2021)), and more. One common debate in the literature is whether ideological disagreement between Congress and agencies predicts oversight intensity and, depending on the mode of oversight being looked at, the results are quite different, with public-facing forms of oversight typically exhibiting an increase in intensity and frequency when ideological disagreement increases and private forms of oversight seeing either no effect or substantively different effects. This is congruent with an emerging theory, first posited by Lowande (2018), that legislators respond to electoral incentives to focus their oversight efforts on ideologically distant agencies in conducting public-facing oversight, such as through committee hearings, and that the lack of these electoral incentives in private oversight drive them to conduct oversight focused on correcting policy valence.¹³ If one envisions oversight as a process, this theory implies that legislators consider electoral incentives when deciding which private oversight investigations to move into the public realm, producing yet another form of selection effects that affect empirical studies of congressional oversight.

Finally, an important qualitative difference between some of these forms of oversight is the degree to which they are ‘routine’. Our models have two timings—Stackelberg and Cournot timing. The distinction between the two is that Stackelberg timing involves the principal credibly committing to an auditing strategy before the agent makes a policy decision, thus the agent has perfect information about whether they will be audited after making their policy decision. On the other hand, Cournot timing does not have the commitment to an auditing strategy, therefore the agent has uncertainty about whether they will be audited or not after making their policy choice. In the Cournot timing of the deterrence model, neither auditing nor perfect accountability occur

¹³Ritchie (2018) examines private communications between members of Congress and the Department of Labor and does find an effect of ideology on oversight intensity, but the picture is much more nuanced compared to studies of public-facing oversight—she finds that senators who are cross-pressured between their constituents and their party pursue potentially electorally unpopular positions through their private communications with agencies, which is congruent with the importance of the public-and-private distinction that we discuss in this section.

in equilibrium. The distinction between these two timings is important, as one can conceive of the different modes of congressional oversight as adhering more to a ‘routine’ structure, where agency policymakers know whether or not they’ll be overseen upon making a policy choice akin to Stackelberg timing, or an *ad hoc* structure, where agency policymakers are unsure about the likelihood of an audit upon making a policy choice akin to Cournot timing. For example, Inspectors General (IGs) report to Congress on a biennial basis with a report that outlines which IG policy recommendations have been implemented by their respective agency and which ones have not (Wilhelm (2023)).¹⁴ Agencies, knowing that they will be audited every 6 months on their progress towards implementing these policy recommendations, may be more likely to implement these recommendations pre-emptively, satisfying perfect accountability, than if audits were irregular and at the whim of Congress. Another example of a Stackelberg-esque oversight mechanism are reporting requirements in legislation that authorizes agencies to ‘fill in’ the details of a law passed by Congress with their policymaking powers. Finally, though this example steps out of the congressional oversight literature a bit, the Office of Information and Regulatory Affairs (OIRA) in the Executive Office of the President mandates review of ‘economically significant’ regulations above \$100 million, thus agencies have relatively clear information about their likelihood of being audited by OIRA (Acs and Cameron (2013), Benn (2023)).¹⁵ The extent to which agencies have perfect knowledge about their probability of being overseen by Congress upon making their policy decision can not only affect the observable outcomes of oversight, it is in part determined by the level of oversight a scholar looks at, though it is difficult to neatly categorize different modes of oversight into the Stackelberg or Cournot categories because of congressional oversight often resembling a ‘process’ with multiple steps, as mentioned earlier.

These selection effects are not ‘bad’, per se—in fact, they create great opportunities for theo-

¹⁴Though IGs are not a part of Congress and are housed in the executive branch, they have a dual reporting role to both the head of their agency and Congress, which is unique among executive branch actors, and Congress has also implemented significant statutory protections for IGs from executive branch interference. Members of Congress also often request IGs to conduct investigations, so it is reasonable to view IGs as a part of Congress’s oversight ‘toolkit’, even though they are formally part of the executive branch.

¹⁵This example is far from perfect, as OIRA is known to review some federal rules below \$100 million and not review some above \$100 million, but the cutoff still gives agencies some degree of information of the likelihood of an audit in the form of an auditing ‘strategy’ from the principal, i.e. the President.

retical development and empirical research. However, it is still important to keep them in mind, particularly for empirical scholars of congressional oversight, as many of them determine the observable outcomes of oversight that are used for empirical studies of oversight.

4.2 The Congressional Oversight Research Database

The Congressional Oversight Research Database (CORD) presents a great opportunity for further empirical investigation of these oversight selection effects. Oversight is a process with multiple steps, often starting with an outside actor notifying Congress of an issue or Congress proactively conducting its own informal oversight, such as through private communications, and then evolving into more formalized oversight, such as committee hearings. Oversight reports are typically the product of multiple committee hearings, thus representing an ‘end-product’ of oversight, as shown earlier in Figure 3.

In contrast to much excellent recent empirical work on congressional oversight, which has examined more informal and private means of oversight to obtain a broader cross-section of congressional oversight, this dataset goes in the opposite direction in some respects, with the production of committee oversight reports being one of the last steps in the oversight process. These reports are highly public and are often the product of several committee hearings, which has been the standard empirical unit of analysis for congressional oversight before the advent of new data sources in the 2010s (Kriner and Schwartz (2008), McGrath (2013)), thus giving an opportunity to examine the end-product of the selection effects that drive congressional oversight and understand how the observable outcomes of oversight change throughout the process of oversight. This dataset is also useful for descriptively understanding committees’ oversight jurisdictions and agendas.

At the time of this article’s writing (early 2025), CORD contains a selection of reports issued by seven Congressional Committees¹⁶ since 1992, as well as a selection of reports issued by individual

¹⁶Specifically, the House Oversight Committee, the Senate Homeland Security and Governmental Affairs Committee, the House’s and Senate’s Foreign Affairs Committees, the House’s Committee on Homeland Security, the House’s Committee on Education and Labor, and the Senate Committee on Health, Education, Labor and Pensions

members of Congress. The CORD Codebook¹⁷ provides a clear discussion of the methodology used to determine whether a report is an *oversight* report. The details of this methodology are slightly beyond our scope in this article, but the length, and care taken in the development, of the methodology illustrates the importance of selection effects in this realm.

Though CORD is still under development, the database already displays the wide variety of topics (as structured by the Policy Agenda Project (PAP) Major and Minor Codes) examined by the committees contained in the database. Table 3 illustrates the breakdown of the PAP Major Topic Codes by each committee in the dataset, alongside the topic codes for reports authored by multiple committees and individual members of Congress.

¹⁷Available here: <https://hdl.handle.net/20.500.14300/1577>.

Table 3: Number of Reports in CORD by Committee & by Most Common PAP Major Topic Codes

| House Committees | 1st | 2nd | 3rd |
|---|--|---|---|
| <i>Education & Labor</i> (15, 100%) | Gov't Operations (15, 100%) | | |
| <i>Foreign Affairs</i> (36, 100%) | N/A (21, 58.3%) | Gov't Operations (14, 38.9%) | Defense (1, 2.8%) |
| <i>Oversight & Reform</i> (596, 100%) | Health (167, 28%) | Gov't Operations (123, 20.6%) | N/A (99, 16.6%) |
| <i>Homeland Security</i> (195, 100%) | Defense (57, 29.2%) | Immigration (48, 24.6%) | N/A (44, 22.6%) |
| Senate Committees | 1st | 2nd | 3rd |
| <i>Health, Education, Labor and Pensions</i> (43, 100%) | N/A (36, 83.7%) | Gov't Operations (3, 8.3%) | Health (2, 5.6%) |
| <i>Foreign Relations</i> (108, 100%) | N/A (50, 46.3%) | Int'l Affairs & Foreign Aid (44, 40.7%) | Gov't Operations (12, 11.1%) |
| <i>Homeland Security & Gov't Affairs</i> (277, 100%) | Gov't Operations (153, 55.2%) | Law, Crime, & Family Issues (18, 6.5%) | Defense/Health (tie) (17, 6.1% each) |
| Other Sources | 1st | 2nd | 3rd |
| <i>Multiple Committees</i> (54, 100%) | N/A/Gov't Operations (tie) (15, 27.8% each) | Defense/Law, Crime, & Family Issues (tie) (6, 11.1% each) | Transportation (3, 5.6%) |
| <i>Independent MC Report</i> (32, 100%) | N/A (17, 53.1%) | Gov't Operations (4, 12.5%) | Energy/Health (tie) (3, 9.4% each) |
| Total (1356, 100%) | Gov't Operations (357, 26.3%) | N/A (283, 20.9%) | Health (199, 14.7%) |

The diversity in topics within committees is a sign that the oversight boundaries of committees are more fluid than the titles of the committees may imply. It is notable that the House Education and Labor committee produces far more reports on Government Operations than its Senate counterpart, and vice versa for the Senate Homeland Security and Government Affairs Committee and the House Oversight and Reform committee, despite the latter having an ostensibly smaller 'oversight' workload by there being a separate House Homeland Security committee. It is also notable

that the plurality of these Government Operations reports, roughly 46%, have the Government Efficiency, General Regulatory Policy and Bureaucratic Oversight minor topic code. This, perhaps, points to committees having different incentives to select into producing these oversight reports.

Another indicator of the jurisdictional boundaries of committees is the federal agencies being overseen in these oversight reports. Not all oversight reports are focused on oversight of federal agencies— many focus on private sector actors or state and local governments, for example— but many of them do, and the agencies that receive the most attention from these reports can tell us a lot about the jurisdictions of each committee and the political variables that induce selection into oversight. We examined the titles and abstracts of all oversight reports in CORD and were able to match 421 of 1356 reports with at least one federal agency, either through the explicit mention of an agency or the mention of a specific policy/program that can be matched directly to an agency.¹⁸ Table 4 shows the breakdown of agencies being overseen by each committee in CORD.

¹⁸A small number of reports oversaw more than one federal agency and are thus counted more than once in Table 4 and in our empirical analysis. More information on our coding process is in Appendix B.

Table 4: Number of Reports in CORD by Committee & by Most Common Agencies Overseen

| House Committees | 1st | 2nd | 3rd |
|---|---------------------|-------------------------------|---|
| <i>Education & Labor</i> (15, 100%) | N/A (15, 100%) | | |
| <i>Foreign Affairs</i> (36, 100%) | N/A (31, 86.1%) | State Department (2, 5.6%) | DOD, USAID, & USAGM (tie) (1, 2.8% each) |
| <i>Oversight & Reform</i> (616, 100%) | N/A (386, 62.7%) | HHS (83, 13.5%) | DOD (17, 2.8%) |
| <i>Homeland Security</i> (195, 100%) | N/A (163, 83.6%) | DHS (32, 16.4%) | |
| Senate Committees | 1st | 2nd | 3rd |
| <i>Health, Education, Labor and Pensions</i> (43, 100%) | N/A (36, 83.7%) | HHS (5, 11.6%) | DOL & EEOC (tie) (1, 2.3% each) |
| <i>Foreign Relations</i> (110, 100%) | N/A (99, 90%) | State Department (4, 3.6%) | DHS (3, 2.7%) |
| <i>Homeland Security & Gov't Affairs</i> (296, 100%) | N/A (148, 50%) | DHS (23, 7.8%) | HHS (16, 5.4%) |
| Other Sources | 1st | 2nd | 3rd |
| <i>Multiple Committees</i> (54, 100%) | N/A (31, 57.4%) | DOJ (6, 11.1%) | HHS (5, 9.3%) |
| <i>Independent MC Report</i> (35, 100%) | N/A (26, 74.3%) | DOD (3, 8.6%) | Multiple Agencies (1, 2.9% each) |
| Total (1400, 100%) | N/A (935, 66.8%) | HHS (110, 7.9%) | DHS (76, 5.4%) |

A few patterns emerge — first is that the substantive focus of the committees matches up with the federal agencies that they conduct the most oversight of. Another notable pattern is that both of the oversight committees — the House Oversight and Reform committee and the Senate Homeland Security and Government Affairs committee — have a higher percentage of reports that can be matched with a specific federal agency than the other committees in the CORD database, perhaps indicative of these committees' greater focus on specifically conducting federal government

oversight. We recognize the shortcomings of this measure and do not argue that we are capturing all of the agencies being overseen in these oversight reports — an analysis of the full text of these reports, rather than just the titles and abstracts, would be more complete — but we do believe that this measure is still helpful in the sense that the measure captures the agencies that receive the most frequent congressional oversight.¹⁹ Despite our measure being an undercount of agencies receiving oversight, we identify 44 federal agencies and independent commissions receiving oversight by at least one report in CORD, with 33 being overseen at least once by the House Oversight and Reform committee and 30 being overseen at least once by its Senate counterpart.²⁰ Even with the limited scope of our data, it is clear that the oversight agenda, particularly from these oversight committees, is diverse and spans the entire gamut of federal government operations.

The general topic of jurisdictional diversity and overlap between committees is not new—much research has been conducted on shifting jurisdictional boundaries between committees (King (1997)), including how these shifts can affect the decision to conduct oversight. Overlapping committee jurisdictions can create incentives for committees to compete or free-ride off of each other’s efforts based on their level of agreement in how the agency should be corrected, leading to inefficiencies in oversight (Clinton, Lewis and Selin (2014), King, Gailmard and Wood (2023)). They can also present opportunities for empirical inference on the determinants of oversight, which we will examine more in the next section.

4.3 Empirical Analysis

Oversight of “big issues,” such as the 2007-08 financial crisis, is not uncommon and frequently involves multiple federal agencies. A related type of occurrence is when multiple committees exercise oversight of the same agency and/or for a common reason, or simply have the jurisdiction to exercise oversight over the same agencies. Such “turf wars” have attracted sustained scholarly

¹⁹We argue that an agency that receives a direct mention in the title or abstract of a report is most likely the target of the most intense scrutiny in the report as compared to an agency only mentioned in the body of the report.

²⁰We also aggregate upwards to the department level, so some sub-agencies, such as FEMA and the Centers for Medicare and Medicaid Services, are grouped together under their parent agency. Therefore, there is even more diversity in the oversight agenda not being captured by this dataset.

attention (*e.g.*, King (1997), Gailmard (2009), King, Gailmard and Wood (2023)).

The overlapping jurisdictional boundaries between House and Senate committees present an opportunity for empirical insight into aforementioned questions about when Congress selects into conducting oversight while controlling for time-correlated unobservable factors. Oversight can be conducted in response to agency errors, such as FEMA’s response to Hurricane Katrina.²¹ Many scholars have also found evidence that oversight is motivated by ideological and partisan concerns (McGrath (2013), Bolton (2022), Kriner and Schwartz (2008)). One issue with disentangling these competing motives that may explain selection into oversight is that the ‘universe’ of issues that Congress can select into investigating differs from year to year— one year, allegations of corruption in the Securities and Exchange Commission may be a salient news story, whereas next year the salient news story may be improper payments to Section 8 housing voucher recipients. Both of these news stories focus on policy valence rather than ideological policy preferences, meaning investigations would target correction of policy valence, but there are also incentives for Congress to differentially investigate these issues based on partisan alignment with the President and the ideological leanings of the agency in question (*e.g.*, Clinton, Bertelli, Grose, Lewis and Nixon (2012), Bonica (2014)), making it unclear the degree to which ideology, partisanship, and/or valence are explaining variation in oversight.

CORD, with report data from House and Senate committees with overlapping jurisdictions, presents an opportunity to disentangle these motivations by holding fixed the baseline ‘universe’ of issues/agencies that can be investigated by Congress at any given moment in time. In a world where congressional oversight activity is strictly motivated by policy valence, there should be no difference in oversight intensity between two co-jurisdictional House and Senate committees at the same moment in time that correlates with partisan (mis)alignment with the President or federal agencies. On the other hand, if members of Congress are also motivated by ideological/partisan concerns, there should be differences in oversight intensity when the House and Senate are con-

²¹CORD helps illustrate this: in the year following Hurricane Katrina’s landfall in Mississippi and Louisiana on August 29th, 2005, there were 7 oversight reports published that explicitly mention Katrina or FEMA in the title of the report out of 76 total. These 7 oversight reports sum to 851 pages out of 3395 pages total, accounting for 9.2% of all reports and 25.1% of all pages published in that time period within CORD.

trolled by different parties (*e.g.*, Gailmard (2009), Clinton, Lewis and Selin (2014), King, Gailmard and Wood (2023)).

To analyze selection into oversight and test p from the deterrence and correction models of oversight, we estimate a series of fixed-effects regression models examining reports from the House Committee on Oversight & Reform, the House Committee on Homeland Security, and the Senate Committee on Homeland Security & Government Affairs from 2003 to 2022. The focus of these oversight committees on conducting oversight of federal agencies fits the substantive focus of the principal-agent models in the paper, and these committees have jurisdiction over executive branch oversight broadly, thus not restricting our analysis to any one agency or policy area.²² To ensure the greatest jurisdictional overlap, we aggregate reports from both the House Oversight Committee and the House Homeland Security Committee to match the jurisdiction of the Senate Committee on Homeland Security & Government Affairs.

One advantage of CORD, relative to committee hearings data, is that some reports are coded regarding whether they were issued by the majority or minority party on a given committee. Both parties in the committee have the right to issue reports, which allows for the analysis of the difference in oversight strategies between majority and minority parties in Congress, rather than just varying the party in control of the chamber of Congress at a given time. To analyze this, we subset to reports that can be linked to a specific party in Congress.²³ To control for the possibility that committees may be producing a large number of duplicate reports or short reports without much substantive oversight,²⁴ we also estimate our empirical models with the number of pages produced in these oversight reports as an outcome variable.

One common finding in the oversight literature is the importance of partisan misalignment with

²²As mentioned earlier, 33 federal agencies were overseen at least once by the House Committee on Oversight & Reform and 30 federal agencies were overseen at least once by its Senate counterpart.

²³Of the 1356 reports in CORD, we were able to link 1047 of them to a specific party. There are instances where reports are issued by both the minority party in one chamber and the majority party in another chamber, but these are usually the same party in an instance of divided government. For the purposes of analysis, we drop reports authored by multiple committees.

²⁴There are a number of examples of largely duplicated reports with low page counts in CORD. For example, the House Democrats' staff on the House Oversight Committee produced 51 reports entitled 'Trump Administration Actions Threaten People with Pre-Existing Conditions', with 50 of the reports dedicated to focusing on one state in particular.

the President in predicting variation in oversight, especially with committee hearings (Kriner and Schwartz (2008), McGrath (2013)). The series of oversight models also point to the importance of partisanship and ideology in explaining oversight intensity. The auditing to deter model with Stackelberg timing and both timings of the auditing to correct model predict an increase in oversight as p approaches 1, and p can be substantively interpreted as the likelihood in which the agent will prefer to select a different policy from what the principal wants. If one makes the assumption that partisan opposition between the President and Congress means a higher likelihood that the agent, such as a federal agency, will select a policy that Congress does not want, we can predict that partisan opposition will lead to an increased production of oversight reports, which is in line with findings from the empirical literature. Therefore, we expect a positive coefficient of partisan misalignment on the production of reports.

Another way that we operationalize p is the ideological divergence between agencies and Congress. Agencies are often conceptualized as having a ‘secular’ ideology independent of the party of the agency head or the President (Richardson, Clinton and Lewis (2018)). For example, the EPA is commonly viewed as a liberal agency, even if the EPA is headed by a Republican. This secular ideology can influence the degree of oversight that Congress, the principal, may want to conduct on the basis that preference divergence is highest among non-ideologically aligned agencies. As discussed earlier in the CORD section of the paper, we assess if an agency is the target of oversight by examining the title and abstract of each oversight report in CORD and connect the agency to its corresponding ideology score from Richardson, Clinton and Lewis (2018).²⁵ For simplicity of measurement, we collapse agency ideology into ‘liberal’ and ‘conservative’ (either below or above 0) and create a dummy variable for ideological alignment on the basis of the party of the report issuer and the ideology of the agency (a value of 1 indicates ideological misalignment). Therefore, we expect a positive coefficient in the regression model.

To control for differential capacity for oversight between these committees, we control for the average committee staff numbers for each year in a given chamber-Congress pair, with House

²⁵More information about this measure is in Appendix B.

Oversight Committee & House Homeland Security Committee staffers summed together. Finally, we control for whether the chamber is controlled by Republicans to assess whether Republicans have a secular desire to conduct more government oversight than Democrats due to their fiscally conservative political views, as well as a Senate indicator variable to control for any secular characteristics of each chamber that may explain oversight report production.

Our unit of analysis is aggregated up from the report level to the congress-chamber-party-alignment level, with alignment being the dummy variable assessing the ideological alignment between the agency being overseen in the report and the party of the issuer of the report. This produces a unique observation for each unique combination of these four variables, producing 80 observations. To model the within-period House-Senate variation that controls for time-varying factors that correlate with oversight intensity, each model is specified with Congress fixed effects. As mentioned earlier, this holds fixed the state of the world that may determine the ‘universe’ of issues that can be overseen by Congress, helping tease out the effect of preference divergence on oversight. The results are presented below in Table 5.

Table 5: Effect of agency ideological misalignment and partisan opposition to the President on Oversight and Homeland Security Committee report outcomes (2003-2022)

| | <i>Dependent variable:</i> | | | | | |
|---------------------------------|----------------------------|---------------------|---------------------|------------------|--------------------|--------------------|
| | Report Count | | | log(Page Count) | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Agency Ideological Misalignment | 0.800 (1.355) | | 0.800 (1.245) | 0.431 (0.563) | | 0.431 (0.534) |
| Opposite Party of President | | 4.450*** (1.248) | 4.942*** (1.306) | | 1.400** (0.540) | 1.477** (0.560) |
| Committee Staff | | | 0.133 (0.103) | | | 0.021 (0.044) |
| Majority Party | | | 1.639 (1.313) | | | 0.257 (0.563) |
| Republican Issuer | | | -0.064 (1.252) | | | 0.395 (0.537) |
| Senate | | | 4.659 (4.978) | | | -0.191 (2.136) |
| Observations | 80 | 80 | 80 | 80 | 80 | 80 |
| R ² | 0.005 | 0.156 | 0.220 | 0.008 | 0.089 | 0.171 |
| Congress F.E. | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |

Note:

*p<0.1; **p<0.05; ***p<0.01
Standard errors are clustered by Congress.

As shown in Table 5, there is a significant and correctly signed effect of partisan misalignment on both the number of oversight reports issued and the number of pages of oversight reports issued. On average, presidential out-party members of Congress produce 4-5 more oversight reports and 338% more pages of reports than members of Congress aligned with the President. However, there is no significant effect of ideological misalignment between the agency and the member of Congress on the production of reports, though the coefficients are correctly signed. This can mean a few things—partisan control of an agency may be the primary factor in dictating preference misalignment between agencies and members of Congress, and members of Congress may be selecting into producing oversight reports more so based on the electoral benefits of embarrassing the President than the policy benefits of conducting oversight of a misaligned agency. This illustrates the impact of selection effects on drawing a clear conclusion from an empirical study of oversight, particularly a mode of oversight that comes after so many layers of selection effects.

That being said, when combined with previous research assessing the impact of ideological disagreement on oversight, particularly in public-facing modes of oversight, the evidence is consistent with ideological disagreement increasing congressional oversight, even if the scope of this conclusion is limited to public-facing oversight and may be driven by members of Congress selecting into oversight on the basis of electoral motives. This evidence is consistent with p in the Stackelberg timing of the auditing to deter model and both timings of the auditing to correct model.

One potential issue with this empirical analysis is that limiting the scope of my analysis to the subset of reports that can be matched to an agency can induce some bias in our results by focusing on the most salient cases of oversight. In Appendix C, we re-run the analysis from Table 5 without the agency ideological alignment variable and find largely similar and correctly signed results, though not as consistently statistically significant.

5 Conclusion

This paper presents the importance of selection effects in understanding congressional oversight, both from a theoretical and empirical perspective. To illustrate this, we presented a series of oversight models that capture different perspectives of oversight that Congress may be engaging in. The ability for the principal to induce accountability through oversight depends significantly on the ability for the principal to correct an auditing decision made by the agent and, importantly, on whether or not the principal can credibly commit to an oversight strategy before the start of the game.

Using insights from our models and the theoretical and empirical literature on congressional oversight, we then discussed some forms of selection effects that may affect the observable outcomes of congressional oversight. Worth noting is that oversight is often a process with multiple steps and the continuation of oversight at each of these steps is itself driven by selection effects, such as the willingness of the agency to comply with congressional preferences and the desire to capitalize on credit-claiming and position-taking in public venues (Selin and Moore (2023)). These selection effects are not ‘bad’, per se, but present opportunities for theoretical development and new empirical research on congressional oversight.

We then analyzed our models of oversight using CORD, finding some evidence that ideology influences selection into oversight which is consistent with the Stackelberg timing of the deterrence model and both timings of the correction model. Oversight reports, due to their public nature and their coming at the end of the oversight process, can be considered the output of congressional oversight with the strongest bias induced from selection effects. Despite this, CORD is nonetheless a very helpful addition to the discipline’s repertoire of oversight datasets, allowing for quantitative and qualitative analysis of the biggest questions in the congressional oversight literature.

A Technical Derivations (For Online Publication)

In this appendix, we first analyze the two timings of the auditing to deter model, and then proceed to the two timings of the auditing to correct model.

A.1 The Auditing To Deter Model

A.1.1 Stackelberg Timing

We first consider a setting in which P can commit to a strategy, $\alpha(y) \equiv \Pr[a = 1 \mid y]$ for each signal $y \in \{0, 1, \varphi\}$, and denote the strategy used by A by $\sigma(\omega) \equiv \Pr[x = \omega \mid \omega]$ for each state $\omega \in \{0, 1\}$. When $\sigma \equiv (\sigma(0), \sigma(1)) = (1, 1)$, A sets P 's most preferred policy with probability 1. We will refer to this profile as satisfying **perfect accountability**.

Incentive Compatibility. For any α , A 's incentive compatibility (IC) conditions are derived as follows. When $\omega = 0$, A 's expected payoffs from $x \in \{0, 1\}$ are:

$$\begin{aligned} EU_A(x = 0 \mid \omega = 0) &= -\pi \cdot ((1 - \nu) \cdot (q \cdot \alpha(0) + (1 - q) \cdot \alpha(1)) + \nu \cdot \alpha(\varphi)), \\ EU_A(x = 1 \mid \omega = 0) &= \eta - \pi \cdot ((1 - \nu) \cdot (q \cdot \alpha(1) + (1 - q) \cdot \alpha(0)) + \nu \cdot \alpha(\varphi)), \end{aligned}$$

yielding the following IC condition for A when $\omega = 0$:

$$\begin{aligned} EU_A(x = 0 \mid \omega = 0) &\geq EU_A(x = 1 \mid \omega = 0), \\ -\pi \cdot ((1 - \nu) \cdot (q \cdot \alpha(0) + (1 - q) \cdot \alpha(1))) &\geq \eta - \pi \cdot ((1 - \nu) \cdot (q \cdot \alpha(1) + (1 - q) \cdot \alpha(0))), \\ \pi \cdot (1 - \nu) \cdot (2q - 1) \cdot (\alpha(1) - \alpha(0)) &\geq \eta, \\ \alpha(1) - \alpha(0) &\geq \frac{\eta}{\pi \cdot (1 - \nu) \cdot (2q - 1)}. \end{aligned}$$

Similarly, the IC conditions for A when $\omega = 1$ are derived from the following:

$$\begin{aligned} EU_A(x = 0 \mid \omega = 1) &= -\pi \cdot ((1 - \nu) \cdot (q \cdot \alpha(0) + (1 - q) \cdot \alpha(1)) + \nu \cdot \alpha(\varphi)), \\ EU_A(x = 1 \mid \omega = 1) &= \eta - \pi \cdot ((1 - \nu) \cdot (q \cdot \alpha(1) + (1 - q) \cdot \alpha(0)) + \nu \cdot \alpha(\varphi)), \end{aligned}$$

yielding the following IC condition for A when $\omega = 1$:

$$\begin{aligned} EU_A(x = 1 \mid \omega = 1) &\geq EU_A(x = 0 \mid \omega = 1), \\ \eta - \pi \cdot ((1 - \nu)(q\alpha(1) + (1 - q)\alpha(0)) + \nu\alpha(\varphi)) &\geq -\pi \cdot ((1 - \nu)(q\alpha(0) + (1 - q)\alpha(1)) + \nu\alpha(\varphi)), \\ \pi \cdot (1 - \nu) \cdot (2q - 1)(\alpha(1) - \alpha(0)) &\leq \eta, \\ \alpha(1) - \alpha(0) &\leq \frac{\eta}{\pi \cdot (1 - \nu) \cdot (2q - 1)}, \end{aligned}$$

Putting these together and recalling that $\eta \in [0, 1]$ and $\alpha(1) - \alpha(0) \in [-1, 1]$, P can achieve perfect accountability in equilibrium if and only if:

$$\pi \cdot (1 - \nu) \cdot (2q - 1) \geq \eta. \quad (3)$$

When Inequality (3) holds, the optimal strategy for P achieves perfect accountability and is equal to the following:

$$\begin{aligned} \alpha^*(0) &= 0, \\ \alpha^*(1) &= \frac{\eta}{\pi \cdot (1 - \nu) \cdot (2q - 1)}, \\ \alpha^*(\varphi) &= 0. \end{aligned}$$

We now consider whether it is optimal for P to achieve perfect accountability, the cost of which involves the cost of auditing, $c > 0$. Supposing that Inequality (3) is satisfied (so that P can achieve perfect accountability in equilibrium but only with $\alpha(1) > 0$), the expected payoff that P will

receive in this equilibrium is equal to P 's direct cost of oversight:

$$\begin{aligned} EU_P &= -c \cdot (1 - \nu)((1 - p) \cdot q + p \cdot (1 - q)), \\ &= c \cdot (1 - \nu)(p(2q - 1) - q), \end{aligned}$$

P 's expected payoff from committing to $\alpha(0) = \alpha(1) = 0$ (i.e., “committing to no review”) is

$$EU_P = -p,$$

because A would always choose $x = 1$ regardless of ω in this case. Thus, achieving perfect accountability is optimal for P if and only if

$$\begin{aligned} c \cdot (1 - \nu)(p(2q - 1) - q) &\geq -p, \\ p &\geq \frac{c \cdot (1 - \nu)q}{c \cdot (1 - \nu)(2q - 1) + 1}. \end{aligned} \tag{4}$$

Before describing this equilibrium (Proposition 1), we state a simple lemma that is very helpful and interesting in its own right.

Lemma 1 *All equilibria of the auditing to deter model support either baseline accountability or perfect accountability.*

Proof: The result follows from the fact that if P can induce (and would benefit from inducing) partial accountability in equilibrium, then P strictly benefits from inducing perfect accountability. A 's IC condition for inducing partial accountability is identical to the IC condition for achieving perfect accountability (Inequality (3)). ■

We now summarize the derivations above as a proposition.

Proposition 1 *In the Stackelberg timing of the auditing to deter model, P audits only after ob-*

serving $y = 1$, and audits $y = 1$ with positive probability if and only if

$$\begin{aligned}\eta &\leq \pi \cdot (1 - \nu) \cdot (2q - 1), \text{ and} \\ p &\geq \frac{c \cdot (1 - \nu)q}{c \cdot (1 - \nu)(2q - 1) + 1}.\end{aligned}$$

Perfect accountability is achieved if and only if P audits with positive probability in equilibrium, and the probability that A is audited after $y = 1$ is equal to

$$\alpha^*(1) = \frac{\eta}{\pi \cdot (1 - \nu) \cdot (2q - 1)}. \quad (5)$$

Proof: Both claims follow immediately from the derivations above (specifically, Inequalities (3) and (4)). ■

Proposition 2 *In the Stackelberg timing of the auditing to deter model, auditing is more likely in equilibrium for*

1. *Higher values of $p \in (0, 1)$,*
2. *Lower values of $c > 0$,*
3. *Higher values of $q \in (1/2, 1)$,*
4. *Lower values of $\eta > 0$,*
5. *Lower values of $\nu \in (0, 1)$, and*
6. *Higher values of $\pi > 0$.*

Proof: Claim 1 follows from the fact that perfect accountability is achieved in any equilibrium in which auditing occurs with positive probability, and this implies that $x = 1$ if and only if $\omega = 1$. The conditional auditing probability after $\omega = 1$ (Equation 5) is independent of p , so that the overall probability of observing auditing in equilibrium is weakly increasing in p (but strictly so if auditing is occurring in equilibrium). Claim 2 follows directly from Inequality (4) (the right hand side of

which is increasing in c). Finally, Claims 3–6 each follow immediately from partial differentiation of Equation (5), above. ■

A.1.2 Cournot Timing

Proposition 7 *In the equilibrium of the Cournot timing of the auditing to deter model, P never audits and A chooses $x = 1$ with certainty.*

Proof: Auditing after any signal y is strictly dominated for P , implying that $\alpha^* = (0, 0, 0)$ in any subgame perfect Nash equilibrium. Accordingly, A choosing $x = 1$ regardless of ω is A 's strict best response, establishing that P never audits and A chooses $x = 1$ with certainty in equilibrium, as was to be shown. ■

Proposition 3 *In the Cournot timing of the auditing to deter model, neither auditing nor perfect accountability occur in equilibrium.*

Proof: Follows immediately from Proposition 7. ■

A.2 The Auditing to Correct Model

Turning to the auditing correct model, we suppose now that the players' payoff functions are as follows:

$$\begin{aligned} u_A(x, a \mid \omega) &= (1 - a) \cdot v_A(x \mid \omega) + a \cdot v_A(x_P^*(\omega) \mid \omega), \\ u_P(x, a \mid \omega) &= (1 - a) \cdot v_P(x \mid \omega) - c \cdot a, \end{aligned}$$

where $c \geq 0$ again represents P 's cost of auditing A (which is again assumed to be common knowledge). The change in the payoffs represents a change in “what audits result in.” In the auditing to deter model, the only role oversight plays is to punish A when P believes A might have chosen $x = 1$ when $\omega = 0$. In the auditing to correct model, oversight does not “cost” A anything, but it results in the final policy being equal to P 's ideal policy (conditional on the true state, ω).

A.2.1 Stackelberg Timing

The agent's strategic calculations are simple in this model: A should always choose $x = 1$:

$$\sigma^*(\omega) \equiv \Pr[x = 1|\omega] = 1 \quad \forall \omega \in \{0, 1\}.$$

P 's expected payoff from a strategy $\alpha = (\alpha(0), \alpha(\varphi), \alpha(1)) \equiv (\alpha_0, \alpha_\varphi, \alpha_1)$ is

$$EU_P^*(\alpha|\sigma^*) = \underbrace{(1-\nu) \cdot (1-q) \cdot (p-c) \cdot \alpha_0}_{EU_P(1|y=0)} + \underbrace{\nu \cdot (p-c) \cdot \alpha_\varphi}_{EU_P(1|y=\varphi)} + \underbrace{(1-\nu) \cdot q \cdot (p-c) \cdot \alpha_1}_{EU_P(1|y=1)} - p. \quad (6)$$

which is linear in α_0 , α_φ , and α_1 , and yields a very simple optimal strategy for the principal, as stated in the following proposition.

Proposition 4 *In the subgame perfect Nash equilibrium of the Stackelberg timing of the auditing to correct model, A always chooses $x = 1$, regardless of ω :*

$$\sigma^*(\omega) \equiv \Pr[x = 1|\omega] = 1 \quad \forall \omega \in \{0, 1\}.$$

and P audits all signals $y \in \{0, \varphi, 1\}$ with equal probability:

$$\alpha^* = \begin{cases} (0, 0, 0) & \text{if } p < c, \\ (1, 1, 1) & \text{if } p \geq c. \end{cases}$$

Proof: The result follows immediately from the derivations above. ■

A.2.2 Cournot Timing

In the Cournot timing, the agent should still always choose $x = 1$, regardless of ω . Accordingly, inspection of each of the first three (under-bracketed) terms in Equation (6) (which correspond to each of the principal's three possible information sets after A chooses x) reveals that P 's equilibrium strategy in this case is identical to P 's equilibrium strategy in the Stackelberg timing (Propo-

sition 5 in the article).

The principal's equilibrium expected payoff from a strategy $\alpha = (\alpha_0, \alpha_\varphi, \alpha_1)$ is then

$$EU_P^*(\alpha|\sigma^*) = \underbrace{(1-\nu) \cdot (1-q) \cdot (p-c) \cdot \alpha_0}_{EU_P(a=1|y=0)} + \underbrace{\nu \cdot (p-c)}_{EU_P(a=1|y=\varphi)} \cdot \alpha_\varphi + \underbrace{(1-\nu) \cdot q \cdot (p-c) \cdot \alpha_1}_{EU_P(a=1|y=1)} - p. \quad (7)$$

which is linear in α_0 , α_φ , and α_1 , and yields a very simple optimal strategy for the principal (Proposition ?? in the article).

Proposition 5 *Perfect accountability is achieved in the Cournot timing of the auditing to correct model if and only if it is achieved in the Stackelberg timing (i.e., if $p > c$), and the principal's equilibrium auditing strategy is identical (Equation (2)).*

Proof: Follows immediately from the derivations above. ■

Proposition 6 *In both timings of the the auditing to correct model, auditing is more likely to occur in equilibrium for*

1. *Higher values of $p \in (0, 1)$, and*
2. *Lower values of $c > 0$.*

Proof: Follows from Propositions 4 and 5. ■

B Matching Reports with Agencies

To match oversight reports with agencies, we examined the title and abstract variables for each report in CORD, which are labeled as ‘dc.title[]’ and ‘dc.description.abstract[]’ in the dataset respectively. The basic criteria we utilized is whether the title and abstract mention the following:

1. The name of an agency or sub-agency
2. A program that can be directly connected to an agency or sub-agency
3. A person directly employed by an agency or sub-agency

The other criteria we used is whether the title or abstract suggest that the agency in question is the target of oversight in the report. In most cases, this is true, but there are cases where an agency is named in the title or abstract of a oversight report but does not appear to be the substantive target of oversight in a report. Below is an example which illustrates this coding rule:

“The House Committee on Oversight and Government Reform investigated hospital-associated infections, particularly central-line-associated bloodstream infections (CLAB-SIs), which contribute to a significant number of deaths and high costs. The Centers for Disease Control and Prevention (CDC) estimates that in 2002, around 1.7 million hospital-associated infections caused about 99,000 deaths. This report highlights the efficacy of preventative measures, specifically in reducing CLABSIs, and reveals that only 14 state hospital associations have adopted or planned to adopt programs to reduce CLABSIs.” (House Committee on Oversight and Government Reform, 2008)

This abstract, taken from a report focusing on hospitals’ lack of measures taken to prevent certain types of hospital-associated infections, does mention the Center for Disease Control and Prevention (CDC). However, they mention the CDC only to provide a data point for the severity of the problem being investigated, which the abstract suggests could be solved by the efforts of state hospital associations. Therefore, this report would not be coded as an oversight report for the

Department of Health and Human Services, the parent agency of the CDC.²⁶

If multiple agencies were coded as being the subject of oversight in a report, an observation was generated for each unique agency-report pair. This did not increase the number of reports that much, however—the number of reports matched with an agency is 421 and the number of report-agency pairs in the dataset is 462. This means that roughly one third of all reports were matched with at least one agency, with higher percentages of matched reports in both of the oversight committees as illustrated in Table 3 of the main text.

We recognize that this measure is likely an undercount of the true number of agencies being overseen in these reports, but looking solely at the titles and abstracts of reports does give a measure of the most ‘intense’ cases of oversight—ones where the agency received a significant amount of the oversight, if not the most, that went into the creation of the report. This precludes any potential over-inclusiveness that might dilute the measure. Future improvements to this measure would benefit from some form of automated text analysis of the body of the report, rather than just the title and abstract.

²⁶Sub-agencies were aggregated up to their parent agencies for ease of analysis. Despite the aggregation, there were still 44 unique federal agencies and independent commissions matched to a report, indicating a high level of diversity in the oversight agenda, particularly for the dedicated oversight committees in the House and Senate.

C Empirical Analysis of Oversight without Agency Alignment

Variable

Table 6: Effect of partisan opposition to the President on Oversight and Homeland Security Committee report outcomes (2003-2022)

| Variable | <i>All Reports</i> | | | |
|-----------------------------|----------------------|----------------------|-------------------|-------------------|
| | Report Count | | log(Report Pages) | |
| | (1) | (2) | (3) | (4) |
| Opposite Party of President | 21.400*** (7.035) | 22.743*** (7.049) | 0.877 (0.627) | 0.823 (0.628) |
| Committee Staff | | 0.433 (0.553) | | 0.013 (0.049) |
| Majority Party | | 4.478 (7.084) | | -0.181 (0.631) |
| Republican Issuer | | -1.548 (6.758) | | 0.160 (0.602) |
| Senate | | 4.236 (26.867) | | -0.912 (2.392) |
| Observations | 40 | 40 | 40 | 40 |
| R ² | 0.242 | 0.404 | 0.063 | 0.264 |
| Congress F.E. | ✓ | ✓ | ✓ | ✓ |

Note:

*p<0.1; **p<0.05; ***p<0.01
Standard errors are clustered by Congress.

References

- Aberbach, Joel D. 1990. *Keeping a Watchful Eye: The Politics of Congressional Oversight*. Washington, DC: Brookings Institution Press.
- Acs, Alex and Charles M. Cameron. 2013. “Does White House Regulatory Review Produce a Chilling Effect and “OIRA Avoidance” in the Agencies?” *Presidential Studies Quarterly* 43(3):443–467.
- Aghion, Philippe and Jean Tirole. 1997. “Formal and Real Authority in Organizations.” *Journal of Political Economy* 105(1):1–29.
- Alexander, Dan. 2024. “Incentives or Disincentives?” *The Journal of Politics* 86(1):257–273.
- Arnold, R. Douglas. 1979. *Congress and the Bureaucracy: A Theory of Influence*. New Haven, CT: Yale University Press.
- Ashworth, Scott, Ethan Bueno de Mesquita and Amanda Friedenberg. 2018. “Learning About Voter Rationality.” *American Journal of Political Science* 62(1):37–54.
- Ban, Pamela and Seth J. Hill. 2025. “Efficacy of Congressional Oversight.” *American Political Science Review* Forthcoming.
- Benn, Annie. 2023. “Presidential partisanship and regulatory review.” *Presidential Studies Quarterly* 53(3):316–333.
- Bils, Peter, Robert J Carroll and Lawrence S Rothenberg. 2024. “Strategic Avoidance and Rule-making Procedures.” *Journal of Theoretical Politics* 36(2):156–185.
- Bolton, Alexander. 2022. “Gridlock, Bureaucratic Control, and Nonstatutory Policymaking in Congress.” *American Journal of Political Science* 66(1):238–254.
- Bonica, Adam. 2014. “Mapping the Ideological Marketplace.” *American Journal of Political Science* 58(2):367–386.

- Clinton, Joshua D, Anthony Bertelli, Christian R 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, David E Lewis and Jennifer L Selin. 2014. "Influencing the Bureaucracy: The Irony of Congressional Oversight." *American Journal of Political Science* 58(2):387–401.
- Gailmard, Sean. 2009. "Oversight and Agency Problems in Legislative-Bureaucratic Interaction." *Journal of Theoretical Politics* 121(2):161–186.
- Gailmard, Sean and John W. Patty. 2012. *Learning While Governing: Information, Accountability, and Executive Branch Institutions*. Chicago, IL: University of Chicago Press.
- Gailmard, Sean and John W. Patty. 2017. "Participation, Process, & Policy: The Informational Value of Politicized Judicial Review." *Journal of Public Policy* 37(3):233–260.
- Gailmard, Sean and John W Patty. 2019. "Preventing Prevention." *American Journal of Political Science* 63(2):342–352.
- Gordon, Sanford C. and Catherine Hafer. 2005. "Flexing Muscle: Corporate Political Expenditures as Signals to the Bureaucracy." *American Political Science Review* 99(1):245–261.
- Gordon, Sanford C. and Catherine Hafer. 2007. "Corporate Influence and the Regulatory Mandate." *Journal of Politics* 69(2):298–317.
- Helmke, Gretchen and Frances Rosenbluth. 2009. "Regimes and the Rule of Law: Judicial Independence in Comparative Perspective." *Annual Review of Political Science* 12(1):345–366.
- Kaufman, H. 2006. *The Forest Ranger: A Study in Administrative Behavior*. Resources for the Future.
- King, David C. 1997. *Turf Wars: How Congressional Committees Claim Jurisdiction*. Chicago, IL: University of Chicago Press.

- King, Janna, Sean Gailmard and Abby Wood. 2023. "Decentralized Legislative Oversight of Bureaucratic Policy Making." *Journal of Theoretical Politics* 35(4):292–309.
- Kollman, Ken. 1997. "Inviting Friends to Lobby: Interest Groups, Ideological Bias, and Congressional Committees." *American Journal of Political Science* 41(2):519–544.
- Kriner, Douglas and Liam Schwartz. 2008. "Divided Government and Congressional Investigations." *Legislative Studies Quarterly* 33(2):295–321.
- Libgober, Brian D. 2020. "Strategic Proposals, Endogenous Comments, and Bias in Rulemaking." *The Journal of Politics* 82(2):642–656.
- Lowande, Kenneth. 2018. "Who Polices the Administrative State?" *American Political Science Review* 112(4):874–890.
- Lowande, Kenneth and Rachel Augustine Potter. 2021. "Congressional Oversight Revisited: Politics and Procedure in Agency Rulemaking." *The Journal of Politics* 83(1):401–408.
- Lupia, Arthur and Mathew D. McCubbins. 1994. "Learning from Oversight: Fire Alarms and Police Patrols Reconstructed." *Journal of Law, Economics, and Organization* 10(1):96–125.
- Macdonald, Jason A. 2010. "Limitation Riders and Congressional Influence over Bureaucratic Policy Decisions." *American Political Science Review* 104(4):766–782.
- McCubbins, Mathew and Tom Schwartz. 1984. "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms." *American Journal of Political Science* 28:165–179.
- McGrath, Robert J. 2013. "Congressional Oversight Hearings and Policy Control." *Legislative Studies Quarterly* 38(3):349–376.
- Mills, Russell W., Nicole Kalaf-Hughes and Jason A. MacDonald. 2016. "Agency policy preferences, congressional letter-marking and the allocation of distributive policy benefits." *Journal of Public Policy* 36(4):547–571.

- Neiheisel, Jacob R. and Michael C. Brady. 2017. "Congressional lettermarks, ideology, and member receipt of stimulus awards from the US Department of Labor." *Research & Politics* 4(3):1–6.
- Ogul, Morris S. 1976. *Congress Oversees the Bureaucracy: Studies in Legislative Supervision*. Pittsburgh, PA: University of Pittsburgh Press.
- Oleszek, Walter J. 2010. "Congressional Oversight: An Overview." Congressional Research Service Report R41079, February 22, 2010.
- Patty, John W and Emily H Moore. 2023. "Personnel, Politics, and Policy-Making." *Journal of Political Institutions and Political Economy* 4(1):111–158.
- Patty, John W. and Ian Turner. 2021. "Ex Post Review and Expert Policymaking: When Does Oversight Reduce Accountability." *Journal of Politics* 83(1):23–39.
- Richardson, Mark D., Joshua D. Clinton and David E. Lewis. 2018. "Elite Perceptions of Agency Ideology and Workforce Skill." *The Journal of Politics* 80(1):303–308.
- Ritchie, Melinda N. 2018. "Back-Channel Representation: A Study of the Strategic Communication of Senators with the US Department of Labor." *The Journal of Politics* 80(1):240–253.
- Ritchie, Melinda N. and Hye Young You. 2019. "Legislators as Lobbyists." *Legislative Studies Quarterly* 44(1):65–95.
- Salzberger, Eli M. 1993. "A Positive Analysis of the Doctrine of Separation of Powers, or: Why Do We Have an Independent Judiciary?" *International Review of Law and Economics* 13(4):349–379.
- Selin, Jennifer L. and Grace Moore. 2023. "Keeping Tabs on the Executive." *Presidential Studies Quarterly* 53(2):186–208.
- Stephenson, Matthew C. 2003. "When the Devil Turns: The Political Foundations of Independent Judicial Review." *Journal of Legal Studies* 32(1):59–90.

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.

Wilhelm, Ben. 2023. "Statutory Inspectors General in the Federal Government: A Primer." *Congressional Research Service* .