A Signaling Theory of Congressional Oversight*

CHARLES M. CAMERON

Department of Political Science, Columbia University, New York, New York 10027

AND

B. PETER ROSENDORFF

Department of Economics, University of Southern California, University Park, Los Angeles, California 90089-0253

Received September 26, 1991

A take-it-or-leave-it bargaining game with asymmetric information and costly signaling is used to examine Congressional supervision ("oversight") of federal agencies. Hearings signal the resoluteness of the committee—the likelihood that the committee will expend the effort to draft and pass a bill overruling an agency. Two kinds of sequential equilibria exist: a pooling equilibrium, and a set of partial-pooling equilibria in which the receiver is able to distinguish among groups of senders. When the receiver sends its utility-maximizing offer, the sender vetoes with positive probability, and if a compromise offer is sent, it is sent on the assurance of its acceptance. These results resemble patterns in oversight observed in Congress. Journal of Economic Literature Classification Numbers: 025, 026.


I. INTRODUCTION

When a federal agency announces its intention to alter a policy, a congressional oversight committee may hold hearings on the issue. However, the committee cannot overrule the agency except by incurring the costs in time and effort of drafting, debating, and passing a bill on the floor of the House or Senate. What then is the purpose of the oversight hearing? Many analysts

* We thank Phoebus Dhrymes, Susan Elmes, Richard Ericson, Andy Rutten, Peter Strauss, and two anonymous referees for helpful suggestions. Rosendorff thanks the Center for Social Sciences at Columbia University for financial support. Our interest in this topic was sparked by seminars given by John Ferejohn at Columbia University in the fall of 1990. Errors and inadequacies in the paper are the sole responsibility of the authors.

0899-8256/93 $5.00
Copyright © 1993 by Academic Press, Inc.
All rights of reproduction in any form reserved.
have suggested that hearings are simply political posturing (Mayhew, 1974); others that the hearings gather relevant information (Maass, 1983). Some have suggested that the hearings are primarily signals (Ferejohn and Shiplan, 1989; Foreman, 1988) and it is this possibility we explore here. In a bargaining game with asymmetric information and costly signaling, hearings can be used to signal the resoluteness of the committee—the likelihood that the committee will expend effort to overrule the agency. We explain why hearings take place relatively rarely, why the threat of hearings is sufficient to engender compromises, and under what circumstances hearings do in fact alter agency policy (McCubbins and Schwartz, 1984; Ogul, 1976; Foreman, 1988).

We focus on three questions: Under what conditions will a committee choose to have oversight hearings, and when will it remain silent in the face of an agency’s policy initiative? When will the threat of hearings be effective and alter agency policy? When will a committee resort to corrective legislation and go to the floor to overrule agency decisions?

The problem we consider involves a monopoly agenda setter with incomplete information. The setter makes a take-it-or-leave-it offer to a chooser; the chooser may send a signal before the offer is made. Recent work by Matthews (1989) considers this problem in the context of presidential speech-making (veto threats), a payoff-irrelevant signal, directed at Congress. In the veto-threat game all compromising types of senders—types that would accept some offer, but would reject the receiver’s most preferred policy—pool together on the same signal, limiting the amount of information the receiver can acquire. If all possible types of senders are compromisers, only uninformative equilibria can occur. The possibility of any transmission of information is predicated on the existence of accommodating types, types that will accept any offer. As noted by Banks (1991), “whenever Congress proposes its utility-maximizing outcome the President never vetoes, whereas when Congress proposes a compromise outcome, the President vetoes with positive probability.”

We resolve this somewhat paradoxical result by generating a partial-pooling equilibrium in a slightly different setting in which (a) the receiver is able to distinguish among groups of compromisers; and (b) signals are informative, even when all types are compromisers. What is more, when the receiver sends its utility-maximizing offer, the sender vetoes with positive probability, and if a compromise is sent, it is sent on the assurance of its acceptance.

Substantively we address political control of the bureaucracy, a principal-agent problem in which transactions costs play a critical role (Weingast and Moran, 1983; McCubbins and Schwartz, 1984; McCubbins et al., 1987; Spiller, 1990). Members of congressional committees try to monitor

1 On the general importance of transactions costs in politics, see North (1990).
agency behavior and compel adherence to policies that frequently fly in the face of the desires of agency personnel. Monitoring and enforcement require the expenditure of effort and time, "a House member's scarcest and most precious resource" (Fenno, 1978, p. 34). The transactions costs of oversight therefore confront congressmen with a difficult choice. As McCubbins et al. (1987) note,

> elected representatives face a tradeoff between the extent of compliance they can command and the effort that is expended to assure it, effort which has an opportunity cost because it can also be used for other politically relevant purposes. [p. 247]

This tradeoff lies at the heart of our analysis of committee behavior.

But the committee is only half of the congressional–bureaucratic nexus. Transactions costs also prevent agency personnel from acquiring relevant information about Congress, reducing their ability to estimate precisely the response of overseers to agency actions. In particular, agencies cannot know with complete precision the opportunity costs facing congressmen. Consequently, agency personnel cannot rid themselves of uncertainty about the resoluteness of the committee, its willingness to pursue a matter doggedly even at the cost of precious effort. This unavoidable uncertainty means the agency too faces a difficult tradeoff; it may try to implement its most preferred policy but only at the hazard of being overturned by the committee and forced by statute to implement a less desirable policy. Or the agency can offer a compromise policy that is less risky but also somewhat less attractive.

In this environment, a credible signal about the committee's resoluteness can be extremely valuable to the agency and may have dramatic effects on its behavior. Signaling may also be valuable to the committee if the signal elicits a more favorable proposal from the agency and saves the committee the transactions costs of actually overturning the agency. We detail below exactly the conditions under which a committee would wish to offer a credible signal of its resoluteness and suggest how oversight hearings can function as that signal.

Throughout our analysis, we focus on what Ferejohn and Shiban (1990) call *agency policy-making*, when an agency unilaterally sets policy. For example, most agencies issue regulations that have the force of law; regulatory agencies often make explicit decisions that are clearly policy decisions;

---

2 As opposed to *statutory policy making*. Statutory policy-making arises when legislators or the president attempt to alter a preexisting status quo. Given statutory authority, the bureaucracy then implements the final policy. Aspects of statutory policy-making (including bargaining over agency budgets) are analyzed in, *inter alia*, Niskanen (1971), Romer and Rosenthal (1979), Matthews (1989), and Banks (1989, 1991). Ferejohn and Shiban (1990) examine agency policy-making under complete information, focusing particularly on the later stages of the process (i.e., adding presidential vetoes and judicial review of oversight legislation).
and agencies sometimes determine effective policy through internal allocation of budgets, for example, by cutting the number of inspectors in a particular bureau. Agencies can make policy because Congress delegates this authority to the bureaucracy. However, congress always retains the right to overrule the bureaucracy's policy initiatives.

As Ferejohn and Shipan (1990), note, agency policy-making has a natural and unambiguous sequential structure of actions. This sequence of actions is determined by the basic structure of American government, the requirements of the Administrative Procedures Act, and the organizational practices of Congress. Under the provisions of the Administrative Procedures Act, an agency must announce whether it intends to issue regulations on a particular subject (typically an agency issues proposed regulations for public comment). Oversight committees may respond with hearings; if the agency subsequently commits to a policy at odds with committee preferences, the committee may resort to corrective legislation. Legislation then goes to the floor, typically under an open rule so that floor amendments are permitted. The other chamber must also act, using a similar procedure; occasionally, differences between the two chambers are resolved via bargaining in conference committee. The president may then veto legislation. Under some circumstances, the courts may review and strike down the legislation.

We focus on the early stages of the agency policy-making process. Latter parts of the process can be incorporated explicitly in a very straightforward way, as in Ferejohn and Shipan (1990) and Ferejohn (1990). For our purposes here, it is sufficient that committee members and the agency understand what the likely outcome of the process will be if the committee actually proceeds with corrective legislation.

In Section II, we review important facts about oversight hearings and justify the information structure we assume in our analysis. Section III presents a game-theoretic model of agency policy-making and a signaling model of oversight. We discuss the theory and offer some concluding remarks in the final section. All the proofs are in the Appendix.

II. CONGRESSIONAL OVERSIGHT OF THE BUREAUCRACY

Stylized Facts about Oversight Hearings

- *Formal oversight hearings are a relative rarity*. Oversight hearings appear to be held sporadically rather than systematically (Dodd and

---

1 See Spulber and Besanko (1991) for an analysis of the delegation of regulatory authority to an agency and the associated problem of commitment to a noncredible policy.

4 We assume that an agency faces a single oversight committee in a chamber. In reality, agencies often face multiple overseers (Moe, 1987). Multiple oversight committees could be added to the model, yielding modest additional insight into signaling.
Schott, 1979). Most oversight hearings are not systematic “police patrols” but instead result from irregular “fire-alarms” (McCubbins and Schwartz, 1984).

- **Agencies sometimes anticipate triggering congressional oversight and alter policy initiatives accordingly.** For example, Ogul (1976) quotes one committee chairman as saying, “I do not have to use hearings as a formal threat because the executive already knows that the threat exists. This is just understood. It seldom has to be discussed explicitly” (pp. 161–162). The capability of exercising oversight appears in some cases to deter agency actions.

- **When held, oversight hearings can have a dramatic effect on agency policy.** Foreman (1988) notes, “Congressional inquiry functions effectively to sensitize agency executives and managers to a plethora of issues, sometimes with spectacular concrete results” (pp. 34–35). Congressional oversight of the Consumer Products Safety Commission, the Federal Trade Commission, and the Occupational, Safety and Health Administration has often shown this pattern (Foreman, 1988, Chap. 2).

- **On other occasions, agencies ignore a committee’s oversight efforts and pursue policies that fly in the face of members’ expressed desires.** For example, some oversight of OSHA in the early 1980s showed this pattern (Foreman, 1988, pp. 64–65).

- **Finally, committees do occasionally introduce legislation and overturn agency decisions or alter agency policies.**

**Information Asymmetries and Oversight**

Agencies go to considerable lengths to collect information about their political overseers. For example, most agencies maintain legislative liaison officers who know the members of oversight committees and their staffs personally. These liaison officials diligently study members’ statements, voting records, and district characteristics. Not surprisingly, they usually have a very precise idea of the preferred policies of committee members.\(^3\)

As specialists in the legislative process, they also monitor general events in Congress so they can make an informed estimate of the outcome that would follow the introduction of corrective legislation—they are likely to be good “nose counters,” in congressional parlance. In addition, as long-time and close observers of the legislative process, they have a fairly precise notion of the transactions costs facing the committee if it tries to overturn agency policy. Such costs include the time and effort required to write a bill, obtain a rule from the Rules Committee, put together a coalition

\(^3\) Some agency administrators have been former committee staffers themselves; e.g., Michael Perschuck and Mary Hanford of the FTC. These officials can be expected to have a profound knowledge of their overseers.
to support the bill, manage the bill on the floor, bargain in conference committee, and so on. Given experience, legislative liaisons may be able to estimate these costs almost as well as committee members can.

Nonetheless, the cost of acquiring information about their political overseers limits what agencies can know. In particular, it is almost impossible even for legislative specialists to know precisely the opportunity costs to committee members of the time and effort spent in the laborious process of overturning agency policy initiatives. The opportunity cost of effort spent on oversight depends on the electoral returns of time and effort spent on other activities (e.g., constituency service, other policy-making activity on other committees, campaigning). These returns to effort depend on a host of factors, many of which fluctuate very rapidly and, from the perspective of an outside observer, quite unpredictably. Congressmen, however, are not outside observers of their personal political worlds. They attentively monitor and respond to such factors as an impending legislative initiative in another committee assignment, the identity of a possible new challenger in the district, the idiosyncratic problems of an influential contributor, the new concern of a local newspaper publisher, the changing economic fortunes of constituents, and a thousand other special circumstances affecting their political careers. Moreover, congressmen have generously endowed themselves with personal staffs in Washington and in their districts to help them track and respond to such developments. A legislative liaison official in an agency is not likely to understand a congressman's interests and responsibilities as well as the congressman does. Asymmetric information on members' opportunity costs of time and effort is unavoidable in agency-committee relations, given the institutional foundations of oversight.

We do not assert that the agency has no idea of the opportunity costs facing members, nor that other forms of asymmetric information are unimportant in congressional–bureaucratic relations. At any given time, legislative specialists in the agency can surely specify a reasonable range for the opportunity costs facing committee members. Also, agency personnel may be somewhat better or somewhat worse informed than committee members about the legislative process or the effects of public policies. A perfectly general model of agency policy making would allow for asymmetric information about the size of transactions costs, likely outcomes given corrective legislation, or even the committee's most preferred policy. Our analysis can be seen as a limiting case in which information asymmetries

---

6 See Ferejohn (1990) for an analysis of the case where the agency has better information about action on the floor than the committee does. Austen-Smith and Wright (1989), Banks (1989, 1990), and McCubbins et al. (1987) consider the implications of a committee facing policy specialists with superior information.
have been restricted to the *minimum* compatible with the institutional foundations of committee oversight. We show below that even this degree of asymmetric information is sufficient to alter profoundly the relationship between the agency and the committee, creating an important role for signals from the Hill.

### III. The Signaling Theory of Oversight

*The Model*

The sequence of play in this game is shown in Fig. 1. Nature chooses the committee's *type*, after which it holds a hearing. The agency offers a policy which can then be accepted or rejected by the committee.\(^7\) If accepted, the policy is implemented, and if rejected, the house’s commonly known preferred policy point is implemented. The agency must make a proposal without knowing the committee’s type \(t\), which the agency believes is a realization of a real random variable with distribution \(F\). Define \(T = [0, 1]\) as the set of all possible types,\(^8\) and let \(F\) be continuous and twice differentiable with density \(f\). We assume that \(F\) is common knowledge and exhibits the increasing hazard rate: for any prior \(F, f(t)/(1 - F(t))\) is an increasing function of \(t\) (Banks, 1991).

In the spirit of Denzau and Munger (1986), endow the committee (interpreted as a single actor) with a fixed effort constraint \(E\), an amalgam of time and resources. The committee may overrule the agency only at some cost in effort, say \(\bar{c}\). The committee may also use resources to produce other activities or services, \(S\), via a production function \(S: [0, E] \times T \to \mathbb{R}^+\). These activities are assumed to be attractive to the committee, for either electoral or policy-oriented reasons. For example, the production

\(^7\) Implicitly, the committee is made aware of the agency's intent to set policy, e.g., through a notice of proposed rule-making.

\(^8\) Any closed interval in the positive reals would suffice.
function may be simply \( S(e, t) = te \). Then \( t \) is a productivity parameter and is assumed to be private information of the committee.

The policy space \( A_1 \) is an interval in the real line. An action for the agency is just \( a_1 \in A_1 \). The agency possesses a von Neumann–Morgenstern utility function, \( U^A \), over policies which is continuous and strictly declines to either side of an ideal policy outcome \( a \); i.e., \( U^A \) is unimodal at \( a \in A_1 \).\(^9\) The committee similarly possesses a utility function, \( U^C \), over policies that is continuous and unimodal at the committee's preferred policy point \( c \in A_1 \). The parent chamber or house (which is not a player in this game) has \( h \in A_1 \) where \( h \) is to be interpreted as the ideal point of the median member of the parent chamber, and hence the policy outcome that is implemented if a bill is brought before the house under an open rule. This \( h \) need not be known with certainty but is assumed to be the commonly known expected outcome from the later parts of the policy-making process.\(^10\) We assume throughout that \( a < h \) or \( a < c \); symmetric analysis applies if \( a > h \) and \( a > c \).

The hearing is a signal or message, which is characterized by the length of time or the amount of effort \( m \) devoted to it. Holding the hearing is costly to the committee because it requires the expenditure of effort that might have been spent producing other services. Note that \( m \), the effort expended on a hearing, is quite distinct from \( \bar{e} \), the cost of actually overruling the agency by passing legislation, as discussed in Section II.

The function \( m: T \rightarrow [0, E] \) specifies the message \( m \) that a committee of type \( t \) sends. The agency now takes the message \( m \), updates its beliefs, and chooses an action according to the function \( a_1: [0, E] \rightarrow A_1 \). The committee, on receiving policy \( a_1 \), chooses its next action according to \( a_2: A_1 \times T \rightarrow \{ \text{accept, reject} \} \).

We focus on pure strategy equilibria. Since \([0, E]\) and \( A_1 \) are compact, no nondegenerate mixing can be a best response (Cho and Sobel, 1990). However, at the accept or reject stage, if a type is indifferent between the two actions given all that has occurred in the play of the game, we assume the tie is broken by playing accept.

We assume throughout that \( E \) is always large enough for the committee to send an equilibrium message and overrule if necessary. This assumption keeps the model interesting—if \( E \) is too small, the opportunity to signal is eliminated. In fact, in equilibrium, a sufficient condition for this assumption to be true is that \( E \geq 2\bar{e} \).

---

\(^9\) Single-peakedness of utility functions represents the satiation that is likely in political contexts. See Matthews (1989), Ferejohn (1990) and Ferejohn and Shipp (1989).

\(^10\) Following Ferejohn and Shipp (1990) and Ferejohn (1990) one might explore the later stages of agency policy-making by adding another chamber, conference committees, the president, and the judiciary as active players.
Definition. A pair of strategies \((a^*_1, m^*, a^*_2)\) and a set of conditional beliefs \(F(\cdot \mid m)\) is a Sequential Equilibrium (Kreps and Wilson, 1982) if

1. For all \(m^*\), \(F(\cdot \mid m^*)\) is the conditional distribution function on \(T\) obtained by updating \(F\) (the prior) using \(m^*\) in a Bayesian fashion whenever possible.

2a. For all \(t, m^*(t) \in \text{argmax } U^c(t, m, a_1^*(m), a_2^*)\).

2b. For all \(t, m, a_1, a_2 \in \text{argmax } U^c(t, m, a_1, a_2)\).

3. For all \((m, a_2), a_2^*(m) \in \text{argmax } \int U^h(t, a_1, a_2^*(a_1, t))dF(t \mid m)\).

We assume that each utility function is linear in the distance of the policy from the optimum.\(^\text{11}\)

\[
U^c(t, m, a_1, a_2) = \begin{cases} 
- tm - |c - a_1| & \text{if } a_2 = \text{accept} \\
- tm - t\bar{e} - |c - h| & \text{if } a_2 = \text{reject}
\end{cases}
\]

\[
U^h(t, m, a_1, a_2) = \begin{cases} 
- |a_1 - a| & \text{if } a_2 = \text{accept} \\
- |h - a| & \text{if } a_2 = \text{reject}.
\end{cases}
\]

Note that these utilities exhibit the Spence–Mirrless single-crossing property: \(\partial^2 U^c / \partial m \partial t < 0\). As the opportunity cost of effort increases, the signaling becomes increasingly costly.

Definition. \(p(m, t) : [0, E] \times T \rightarrow A_1\) such that \(U^c(t, m, p(m, t), a_2 = \text{accept}) = U^c(t, 0, h, a_2 = \text{reject})\).

We call the pivot the element in policy space left of \(c\) that leaves a type \(t\) committee indifferent between signaling \(m\) and accepting the pivot and signaling \(0\), rejecting the offer, and receiving \(h\). With linear utilities and \(h \leq c\), \(p(m, t) = h - t(\bar{e} - m)\).\(^\text{12}\) Figure 2 illustrates this case.

In any equilibrium each type falls conveniently into one of two groups. A committee type \(t\) is “accommodating” if \(t\) prefers accepting \(a\), the agency’s ideal point, to \(h\) (taking into account the rejection cost \(\bar{e}\)). That is, \(a > p(m^*, t)\). If \(a > p(m^*, 0)\), all types are accommodating. A committee type that is not accommodating is called “compromising.”\(^\text{13}\)

\(^{11}\) These utility functions have undergone a linear transformation to suppress \(E\), the total effort endowed to the committee.

\(^{12}\) If \(c < h\), the left pivot is \(p(m, t) = 2c - h - t(\bar{e} - m)\), a linear translation of the pivot in the text. In addition, careful readers will note that single-peakedness of \(U^c\) and costly actions for the committee imply the existence of two pivot points; by focusing on the case where \(a < h\) or \(a < c\), we need only consider the left pivot.

\(^{13}\) As in Matthews (1989), a committee type \(t\) would be “recalcitrant” if a preferred \(h\) to any outcome \(a_t\) sent by the agency. However, in this game \(p(m, t) \leq h \forall t\), \(m\) if \(\bar{e} > 0\), and so \(a_t \in [p(m, t), h]\) will not be rejected by any \(t \in T\), and so there do not exist any recalcitrant types. In contrast to Matthews (1989), the definitions of accommodating and compromising used here depend on equilibrium levels of signaling. This difference stems from the fact that signaling is costly in the oversize game but costless in Matthews’ veto threat game.
We make use of the type of the committee to refer to its resoluteness. A committee with a higher type and hence a lower pivot is more likely to agree to a compromise and is therefore less resolute. A lower type, whose value of outside services is lower, can afford to be more hard-line about the issue in question and have a higher pivot. This type gives away less in a compromise and is termed more resolute.

**Total-Pooling Equilibrium**

Let \( t^a \) satisfy \( a = p(0, t^a) \). That is, \( t^a \) is the committee type whose pivot point, when sending a zero message, is the agency’s preferred policy point. If \( h \leq c \) then \( t^a = (h - a)/\tilde{c} \); if \( h > c \), \( t^a = (2c - h - a)/\tilde{c} \). For any offer \( a_1 \in A_1 \), define \( t^{a_1} \) by \( a_1 = p(0, t^{a_1}) \). Define \( t^* = \arg\max_{t \in [0, 1]} t[1 - F(t)] \) and \( t^{**} = \min(t^*, t^n) \).

**Proposition 1.** For any \( a \in A_1 \), the following is an equilibrium: \( \forall t \in T, m^*(t) = 0 \). The agency offers \( a^*_t(m^*(t)) = p(0, t^{**}) \) and \( a^*_t(m' \neq m^*(t)) = p(0, 1) \) (i.e., out-of-equilibrium beliefs are specified to be pessimistic; any deviation \( m' \neq 0 \) is believed to be by type \( t = 1 \)). The committee plays \( a^*_t(a_1, t) = \text{accept} \) if \( t \geq t^n \); reject if \( t < t^n \).

If \( h \leq c \), \( p(0, t^{**}) = h - t^{**}\tilde{c} \); if \( h > c \), \( p(0, t^{**}) = 2c - h - t^{**}\tilde{c} \). In this equilibrium, the committee never holds oversight hearings. The agency learns nothing useful from the absence of a hearing, so the proposal it offers depends on its “best guess” about the resoluteness of the commit-
tee. If the agency believes it cannot possibly be facing an accommodating committee, or if it believes the probability of facing an accommodator is sufficiently low (i.e., if \( r^o > t^* \)), it trades off between offering less attractive proposals that have a higher probability of acceptance and more attractive proposals with lower probability of acceptance. If the agency believes it may be facing an accommodating committee (i.e., if \( 0 < r^o < t^* \)), the agency offers its own most-preferred policy. In either case, there is a positive probability that the committee will reject the agency’s offer (rejection occurs when the committee’s actual type \( t \) is greater than the agency’s estimate \( t^* \)). Finally, if \( r^o < 0 \) (i.e., \( a > h \)) then \( a > p(0, 0) \) and all types are accommodating; the agency sends up its most-preferred proposal \( a \) and there is no chance of rejection.

**Partial-Pooling Equilibria: The No-Accommodators Case**

We assume for the purposes of this section that no accommodators are present; that is, \( a < p(m, 1) \forall m \leq E \). This assumption will be relaxed in the next section. We first define the set parameterizing the equilibria; this set is shown in the Appendix to be nonempty.

**Definition.** Define the set \( T^E \) by \( T^E = \{ (r^o, r^o) \mid r^o < t^* < r^o \text{ and } r^o f(r^o) \geq F(r^o) - F(r^o) \} \).

For any equilibrium parameterized by \((r^o, r^o) \in T^E\) a signaling schedule

\[
\bar{m}(r^o, r^o) = \bar{e} \left[ \frac{r^o - r^o}{2 r^o - r^o} \right]
\]

can be specified. It is important to note that \( 0 < \bar{m}(r^o, r^o) < \bar{e} \); that is, the cost of signaling is less than the cost of enforcement (Lemma 5 in the Appendix supplies a proof). For the purposes of the next proposition for some \((r^o, r^o) \in T^E\) define \( T_1 = [0, r^o] \cup (r^o, 1] \); \( T_2 = (r^o, r^o] \) and for any offer \( a_1 \), define \( r^{m^o} \) by \( a_1 = p(\bar{m}(r^o, r^o), r^{m^o}) \).

**Proposition 2.** For any \((r^o, r^o) \in T^E\), the following is a sequential equilibrium to the signaling game:

The committee signals \( \forall t \in T_1, m^*(t) = 0 \) and \( \forall t \in T_2, m^*(t) = \bar{m}(r^o, r^o) \).

For the agency \( a_1^*(m^*(t) = 0) = p(\bar{m}, r^o), a_1^*(m^*(t) = \bar{m}) = p(\bar{m}, r^o) \), and \( a_1^*(m^o \neq 0, \bar{m}) = p(\bar{m}, 1) \) (any deviation from \( m^* \) is believed to be by type 1). For any offer \( a_1 \), the committee now plays

\[
a_1^*(a_1, t) = \begin{cases} 
\text{accept if } t \leq r^{m^o} \\
\text{reject if } t < r^{m^o}.
\end{cases}
\]
In the case of \( h \leq c \), \( p(\bar{m}, t^*) = h - t^*(\bar{c} - \bar{m}) \) and \( p(\bar{m}, t^c) = h - t^c(\bar{c} - \bar{m}) \); if \( h > c \), \( p(\bar{m}, t^*) = 2c - h - t^*(\bar{c} - \bar{m}) \) and \( p(\bar{m}, t^c) = 2c - h - t^c(\bar{c} - \bar{m}) \). As before, these differ only by a constant.

In this equilibrium, the committee will hold an oversight hearing or may remain silent depending on its resoluteness. Consider Fig. 3, drawn for the case \( h \leq c \); the top part of the figure shows the type space where the committee types fall into three groups. In the first group lying between 0 and \( t^* \), all types are very resolute ("tough"). A tough committee is not recalcitrant: it will accept some proposals the agency would be willing to offer, but only if the proposal is quite high. Signaling is worthwhile for a tough committee only if it receives a very high offer in return, for the committee—demanding to begin with—must be further compensated for its expenditure of effort in hearings. In equilibrium, tough committees know they are unlikely to receive a high enough offer and therefore do not signal. In the second group lying between \( t^* \) and 1, all types are quite irresolute ("pushovers"). A pushover committee is not an accommodator for there are some proposals it will reject. However, pushovers are easy to please. A pushover committee knows that any offer it receives will probably be high enough to accept, while holding hearings eats up very valuable time (the reason a committee is a pushover is that its alternative uses of time are so remunerative). Therefore, pushover committees do not signal either.\(^{14}\) In the third group lying between \( t^* \) and \( t^c \), committee types

\(^{14}\) A complementary way to think about this committee is the following. If the committee spent time on oversight hearings, then it could elicit a better proposal from the agency. But the proposal would not be sufficiently better to offset the losses from time spent in oversight (the level of the message in the partial-pooling equilibrium is high enough so an offsetting gain cannot occur.) Consequently, the best the committee can do is remain silent and hope the agency believes it to be a tough guy, a form of bluffing.
fall into an intermediate range between the tough types and the pushovers.
A committee in this intermediate range is willing and able to signal, and
does so in order to induce the agency to send a more attractive offer.15

In the equilibrium, if the agency observes a signal it knows it faces a
committee in the intermediate group. The agency is happy to offer the
committee a rather generous proposal \((p(\bar{m}, t^*) = h - t^*(\bar{e} - \bar{m}))\), for it
knows the committee will be satisfied with a reasonable compromise that
the agency can achieve with no risk of rejection. If the agency does not
observe a signal, it knows the committee is either a pushover or tough.
Again the agency offers a compromise proposal, but a less generous one
\((p(\bar{m}, t^*) < p(\bar{m}, t^0))\). This offer will be rejected by tough committees and
accepted by pushovers. The lower part of Fig. 3 illustrates the policies
that are offered \((A_1)\) and the committee responses \((A_2)\) for the case where
\(h \leq c\).

The message sent in equilibrium is cheaper than \(\bar{e}\). Holding hearings
extracts a compromise and is cheaper than having to overrule the agency.

A striking feature of the partial-pooling equilibrium is its capacity to
convey information. Very different policies are sent when the agency sees
an oversight hearing and when it does not. This occurs even though
committees in the high and low groups hold no hearings and all committees
in the medium range devote the same time and effort to the hearing.

**Partial-Pooling Equilibria—The Compromisers
and Accommodators Case**

Recall that any pair \((t^e, t^p) \in T^E\) parameterizes an equilibrium in which
the signal sent by the middle group is \(\bar{m}(t^e, t^p)\). Accommodators are present
in an equilibrium if \(t^e \leq t^p < 1\) where \(t^e\) satisfies \(a = p(\bar{m}(t^e, t^p), t^p)\). Then
\((t^e, t^p) \in T^E\) parameterizes a partial-pooling equilibrium where all \(t \in (t^e, 1)\)
are accommodators.16

**Corollary.** If accommodators are present (i.e., \(t^e < 1\), then

\[ T^E = \{(t^e, t^p) \mid t^e < t^* < t^p \leq t^e \text{ and } t^e f(t^p) = F(t^e) - F(t^p)\} \]

is the set of parameters of the partial-pooling equilibrium.

---

15 The sender's strategy in this equilibrium is nonmonotonic, in contrast to that in standard
signaling games (Cho and Sobel, 1990). Because the sender may choose an outside option
after the receiver responds to his signal, the oversight game is not a standard signaling game.
When the opportunity costs of rejecting are very low, taking the outside option becomes
very attractive to the sender (ex ante). In this case, the sender may essentially opt out of the
signaling portion of the game by expending no effort. Hence the nonmonotonicity.

16 In the case that \(h \leq c\), \(t^e\) satisfies \(a = h - t^e(\bar{e} - \bar{m}(t^e, t^p))\) or \(t^e = [h - a]/[\bar{e} - \bar{m}]\);
if \(h > c\), \(t^e\) satisfies \(a = 2c - h - t^e(\bar{e} - \bar{m}(t^e, t^p))\) or \(t^e = [2c - h - a]/[\bar{e} - \bar{m}]\).
Proof. Identical to the previous proposition.

The equilibria in this case are similar to those in the no-accommodators case. In fact, if the probability that the committee is an accommodator is sufficiently small, the behavior of the agency and committee is qualitatively identical to the no-accommodators case (see Fig. 4, again drawn for the case \( h \leq c \)). However, if the probability that the committee is an accommodator is higher and the committee fails to hold hearings, the agency's behavior differs in an important way from the earlier case. Consider Fig. 5 (also drawn for the case \( h \leq c \)); as before, in the absence of hearings the agency cannot know whether the committee is tough or an accommodator. But now the agency may be inclined to believe the committee is an accom-
moderator unless given evidence to the contrary. Therefore, absent a hearing, the agency may not offer a compromise proposal but instead it may offer its own most-preferred point. This behavior is similar to the agency action in the pooling case, where given a committee that is likely to be an accommodator, the agency offers its most-preferred policy. Again, absent a hearing, there is a chance the proposal will be rejected by the committee.

In both cases of the partial-pooling equilibria, the committee types fall into three groups. It is shown in Proposition 3 in the Appendix that the committee types cannot fall into two groups.

**Separating Equilibrium**

No separating equilibrium can exist (Proposition 4 in the Appendix). If it did, on receiving the message the agency could immediately infer the type of the committee. Then the agency would return that committee type’s pivot yielding the worst possible outcome for the committee. All types would defect and pretend to be the lowest possible type, or the type with the highest possible pivot.

**Equilibrium Refinement**

The total-pooling and the partial-pooling equilibria are shown in Proposition 5 to be universally divine equilibria.

**Definition.** A sequential equilibrium satisfies Universal Divinity if for all out-of-equilibrium signals s, the posterior beliefs \( \mu(t, s) > 0 \) only if there exists an action \( q^t \) which is a best response to \( s \) such that \( U^*(t) < U(t, s, q^t) \) and \( U^*(t') \geq U(t', s, q^t) \) for all \( t' \neq t \).

That is, we put positive weight only on types who are most likely to defect, or who are likely to gain the most from such a defection. For any out-of-equilibrium message \( s \), we can identify a share \( q \) making a type \( t \) sender indifferent between defecting and sticking to the equilibrium path. Any higher \( q \) would imply a strict preference for defecting. Thus by identifying the types with the lowest "indifference share" at \( s \), we have the types most likely to defect. Universal divinity requires the posterior beliefs to place positive probability only on those types.

**Proposition 5.** The partial-pooling equilibrium and the total-pooling equilibrium are universally divine equilibria.

So, when all posterior weight is placed on the type most likely to deviate, there remains no incentive for any type to deviate—hence the equilibrium path is robust to the belief restriction of universal divinity.
IV. Discussion

Explaining the Stylized Facts

Behavior in the partial-pooling equilibrium is compatible with the behavior reported in Section II. First, oversight hearings take place on occasion but are a relative rarity. This is what the theory predicts, if most committees fall into the high or low groups. Second, the threat of corrective legislation can sometimes be effective in restraining agency behavior even absent hearings. The theory predicts this outcome whenever the agency believes the committee is definitely not an accommodator. In this case, the agency offers a compromise proposal even absent a hearing. The logic of this outcome closely resembles the case of "congressional dominance" of the bureaucracy described in McCubbins and Schwartz (1984) and Weingast and Moran (1983), whose focus can be seen as this case. Third, as noted by Foreman (1988), oversight hearings can have a dramatic effect on agency behavior. If the agency initially believed the committee might be an accommodator, the model predicts a change in agency policy upon receipt of the signal, and this change could be quite dramatic. Finally, the model predicts that, absent a hearing, there is always some chance of going to the floor with corrective legislation. Indeed, this outcome seems to occur from time to time.

The model fails to predict one outcome that periodically arises: hearings yielding no perceptible change in agency behavior. This outcome occurs occasionally when the committee is unable to effectuate change through corrective legislation. For example, Republican control of the Senate effectively blocked House Democrats from overturning some OSHA policies in the early 1980s, but Democrats held oversight hearings anyway (Foreman, 1988, pp. 64–65). A pure signaling model of oversight cannot explain why a committee would waste time on hearings that a rational agency will ignore. The obvious explanation lies in the realm of political posturing or "position-taking” (Mayhew, 1974): hearings sometimes have value as political theater, not only as a signal of resoluteness. If the model were modified to allow this possibility, then a rational committee might undertake hearings that the agency would ignore. However, provided the agency knew the value of hearings as position taking, longer hearings would still function as a signal resoluteness.

In short, the signaling theory of oversight provides a coherent and plausible explanation for many observed patterns in oversight. Alternative explanations neatly complement the theory.

17 For an example, see Ogul’s discussion of congressional oversight of the Post Office (Ogul, 1976, especially at pp. 33–38).
Relationship with the Veto Threat Game

In many respects, the model presented here is quite similar to Matthew's well-known model of presidential veto threats. However, as discussed in Section I, the partial-pooling equilibria in the two models are quite different. Does this difference arise from the use of costless signals in the veto threat game and costly ones in the oversight game? The answer is no. The reason is that the information structures in the two games are quite different.

Consider the following comparative statics exercise: allow the cost of overturning the agency's proposal (i.e., $\bar{e}$) to go to zero while keeping the opportunity cost of action (i.e., $i\tilde{e}$) private information (recall that the nominal cost of overturning the agency, $\bar{e}$, is common knowledge). As the cost of overturning the agency falls, so must the length of hearings (this follows from simple inspection of the equation for equilibrium signals ($\tilde{m}$)). When the committee can overturn the agency completely costlessly (i.e., when $\bar{e}$ equals zero), the committee will stop holding hearings. More importantly, though, if the committee can costlessly overturn the agency then the opportunity cost of action must necessarily be zero as well (i.e., if $\bar{e} = 0$ then $i\tilde{e} = 0$). But then the utility function for the committee becomes simply

$$U^c(t, m, a_1, a_2) = \begin{cases} -tm - |c - a_1| & \text{if } a_2 = \text{accept} \\ -tm - |c - h| & \text{if } a_2 = \text{reject.} \end{cases}$$

Since the committee will not signal, $tm = 0$ and then every component of the committee's utility function is common knowledge. In other words, when the committee can overturn the agency's policy costlessly, the oversight game does not become a signaling game with costless signals but instead **collapses to the standard monopoly agenda setter model with complete information**. Analyzing this game is straightforward: the agency will offer its most preferred policy if the committee will accept it, and a policy utility equivalent for the committee to the floor's policy ($h$) if the committee will reject the agency's most preferred policy (see Ferejohn and Shiplan, 1990).

In short, the difference in the equilibria in the oversight game and the veto threats game stems not from the use of costly or costless signals but instead from differences in fundamental assumptions about the nature of uncertainty in political settings. In the veto threats game, uncertainty

---

18 The committee still has private information, namely, its type $t$. But only the opportunity cost of action, $i\tilde{e}$, not $t$ itself, affects the committee's payoffs. Once $\tilde{e}$ hits zero, the incomplete information remaining in the game is payoff irrelevant and cannot affect its outcome.
is about preferences: the monopoly agenda setter knows perfectly the chooser’s cost of action but is uncertain about what the chooser really wants. In the oversight game, uncertainty is about the cost of action: the monopoly agenda setter knows perfectly what the chooser wants but is uncertain about what the chooser will accept. Both assumptions capture important aspects of political uncertainty. However, for the reasons outlined in Section II, we believe that the second type of uncertainty is particularly important in the case of Congressional oversight of the bureaucracy.

V. Conclusion

In this paper, we formulate a signaling theory of congressional oversight hearings. From a modeling standpoint, we extend earlier analysis of take-it-or-leave-it bargaining games with asymmetric information by incorporating costly signals tied to transaction costs. Substantively, the theory explains many of the observed facts about oversight hearings and complements recent investigations of political control of bureaucracy.

Many unanswered questions about congressional oversight remain. To our minds, perhaps the most interesting is, why would Congress as a whole allow committees to create the institution of the oversight hearing? In other words, are oversight hearings in the interest of the average member of the whole chamber? Answering this question would require modifying the game studied here so that non-committee members become active players in an earlier stage of play.\footnote{Gilligan and Krebsiel (1990) provide an example of game-theoretic analysis of the organizational design of a legislature.}

 Appropriately modified, the model of transaction costs and signaling formulated here may be able to shed light on the performance of other political institutions; for example, the judicial hierarchy (signaling between the Supreme Court and appellate courts), parliamentary systems (signaling between cabinet ministers and the permanent civil service), and hierarchical organizations (signaling between superiors and subordinates (Milgrom and Roberts, 1990)).

Appendix

The properties of $F$, the prior distribution, are the following:

A1. $F$ is a cumulative distribution function on bounded support $T = [0, 1]$.\footnote{Gilligan and Krebsiel (1990) provide an example of game-theoretic analysis of the organizational design of a legislature.}
A2. $F$ is twice continuously differentiable with density $f(t) \neq 0$ except possibly at 0 or 1.

A3. $F$ has the increasing hazard rate property.

**Definition.** Define the function $\Psi: T \rightarrow \mathbb{R}^+$ as $\Psi(t) = t[1 - F(t)]$. Define $t^* \in \text{argmax}_{t \in T} \Psi(t)$. This maximum exists: $\Psi$ is a continuous function on a compact set.

**Lemma 1.** (1) $t^* > 0$;
(2) $t^*$ is unique;
(3) $\Psi'(t) > 0 \ \forall t \in (0, t^*)$;
(4) $\Psi'(t) < 0 \ \forall t \in (t^*, 1]$ if $t^* < 1$.

**Proof.** (1) Suppose not. That is, $t^* = 0$. Then $\Psi'(t) = 1 - F(t) - tf(t) \leq 0 \ \forall t \in T$. But $\Psi(t^*) = 1 - F(0) = 1 \neq 0$. Contradiction.
(2) Either $1 - F(t) - tf(t) > 0 \ \forall t \in T$, in which case $t^* = 1$ and is unique, or $1 - F(t^*) - t^*f(t^*) = 0$. Suppose $\exists t^* > t^{**}$ such that both are maximands. Then

$$t^* > t^{**} \Rightarrow \frac{1}{t^*} < \frac{1}{t^{**}} \Rightarrow \frac{f(t^*)}{1 - F(t^*)} < \frac{f(t^{**})}{1 - F(t^{**})} \Rightarrow \lambda(t^*) < \lambda(t^{**}) \quad \text{iff } t^* < 1$$

where $\lambda(\cdot)$ is the hazard rate. By A3 $t^* > t^{**} \Rightarrow \lambda(t^*) > \lambda(t^{**})$. The contradiction is established. If $t^* = 1$, $t^*[1 - F(t^*)] = 0 = t^{**}[1 - F(t^{**})] \Rightarrow 1 - F(t^{**}) = 0$ since $t^* \neq 0$. Then $f(t^{**}) = 0$. Then $t^{**} = 1$ by A2. Again contradiction.
(3) For all $t$ in $(0, t^*)$, $\lambda(t) < \lambda(t^*) \Rightarrow \frac{1 - F(t)}{f(t)} > \frac{1 - F(t^*)}{f(t^*)} = t^* > t$.
Then $1 - F(t) - tf(t) > 0 \ \forall t \in (0, t^*)$.
(4) Similar to 3. ■

The proofs of the propositions make use of the function $\hat{h} = h$ if $h \leq c$; $\hat{h} = 2c - h$ if $h > c$.

**Proof of Proposition 1.** I. The Agency. All types pool—the posterior distribution is the same as the prior. The expected utility of the agency can then be expressed as $U^A(t, m, a_1, a_2) = -(\hat{h} - a) + i\epsilon[1 - F(t)]$. Recall that a choice of a policy $a_1 = \hat{h} - t_i\epsilon$ corresponds to a choice of a type $(t_i)$. The agency never accepts a policy corresponding to type $t_i \in (t^*, 1)$ since there exists another $t'_i < t^*$ with equal distance from the agency’s preferred policy, but $t'_i$ has a higher probability of acceptance. So all $t > t^*$ are dominated. Therefore $t$ is chosen from $[0, t^*]$. The agency’s problem is then restated as $\max_{\epsilon \in [0, t^*]} t[1 - F(t)]$. Now $t^*$ is the maximizer over $T$. If $t^* < t^*$ then $t^*$ is the maximizer over $[0, t^*]$. If $t^* > t^*$, Lemma
I implies that $t^*$ is the minimizer. So $t^{**} = \min \{t^*, t^o\}$ is the type corresponding to the optimal policy.

II. The Committee. Consider any deviation to $m'$. Then $a_1^* = h - 1\bar{c}$ and

$$U^C(t, m', a_1, a_2) = \begin{cases} -tm' - 1\bar{c} - |c - h| & \text{if } a_2 = \text{accept} \\ -tm' - t\bar{c} - |c - h| & \text{if } a_2 = \text{reject.} \end{cases}$$

Then all types reject in deviation. Then $U^* \succeq U'$ iff $-t\bar{c} \succeq -tm' - t\bar{c}$ which is true $\forall t \in T$. Hence there is no incentive to deviate to $m'$. ■

**Definition.** $T^w = \{t^w \in [0, 1] \mid 1 - F(t^w) - 2t^w f(t^w) \geq 0\}$.

**Lemma 2.** (1) $T^w$ is nonempty;

(2) $\forall t^w \in T^w, t^w \leq t^*$.

**Proof.** (1) $t^w = 0 \in T^w$, since $1 - F(0) - 2.0f(0) = 1 \geq 0$.

(2) If $t^* = 1$ then any $t^w < 1$ has $t^w \leq t^*$. If $t^* < 1$, suppose not. That is, $\exists t^w \in T^w$ such that $t^w > t^*$. Then $\lambda(t^w) > \lambda(t^*)$ by A3. Then

$$\frac{f(t^w)}{1 - F(t^w)} > \frac{f(t^*)}{1 - F(t^*)} = \frac{1}{t^*} > \frac{1}{t^w},$$

so

$$\frac{f(t^w)}{1 - F(t^w)} > \frac{1}{t^w},$$

i.e., $1 - F(t^w) - t^w f(t^w) < 0$. Then $1 - F(t^w) - 2t^w f(t^w) < 0$ and $t^w \notin T^w$. Contradiction. ■

We introduce here a fourth assumption on the prior distribution.

A4. If $t^* < 1$, $f'(t) \geq 0 \forall t < t^*$.

A3 limits how negative $f'$ can get; A4 strengthens A3 over an interval by requiring $f'$ to be positive over that interval. A4 is equivalent to the requirement that $t^* \leq r^{\text{mode}}$, where $r^{\text{mode}}$ is the mode of $f$.

**Definition.** $T^E = \{(t^w, t^*) \in T^w \times [t^*, 1] \mid t^w f(t^w) \geq F(t^*) - F(t^w)\}$.

**Lemma 3.** If $t^* < 1$, then $T^E$ is nonempty.

**Proof.** The condition is equivalent to $1 - F(t^*) - t^w f(t^w) \leq 1 - F(t^w)$. Define $\bar{r} = \max T^w$. Then $\bar{r} \neq 0$ and $1 - F(\bar{r}) - 2\bar{r}f(\bar{r}) = 0$. So $1/2\bar{r} = \lambda(\bar{r})$. Claim $(\bar{r}, t^*) \in T^E$. Since $\bar{r} \in T^w$, $1 - F(t^*) - t^w f(t^w) = \bar{r} f(\bar{r})$ and $1 - F(t^*) = t^w f(t^*)$. Now $\bar{r} \leq t^*$ by Lemma
2 and \( f(\bar{r}^a) \leq f(t^*) \) by A4. Then \( \bar{r}^a f(\bar{r}^a) \leq t^* f(t^*) \). We require also that \( t^a \geq t^a \forall (t^a, r^a) \in T^E \). A sufficient condition for \( t^* \leq t^a \) is

\[
\frac{\hat{h} - a}{\bar{c} - \bar{m}(t^*, r^a)} \geq t^a.
\]

That is, \( \hat{h} \) and \( a \) must be sufficiently far apart or \( \bar{c} \) sufficiently small. ■

**Lemma 4.** \( 1 - F(t) - tf(t) \geq F(r^a) - F(t) \forall t \in [0, t^a], \forall (r^a, r^a) \in T^E \).

**Proof.** For any \( (t^a, r^a) \in T^E \), \( -[F(r^a) - F(t)] \approx - t^a f(t^a) \). Then \( 1 - F(r^a) - t^a f(t^a) - [F(r^a) - F(t)] \geq 1 - F(r^a) - 2t^a f(t^a) \geq 0 \). Now \( (d/dt)[1 - F(t) - t^a f(t)] = -2f(t) - tf'(t) < 0 \forall t \leq t^a \) by A4. Then \( 1 - F(t) - t^a f(t) - [F(r^a) - F(t)] \geq 1 - F(r^a) - t^a f(t^a) - [F(r^a) - F(t^a)] \geq 0 \forall t \leq t^a \). ■

**Lemma 5.** For any \( (r^a, t^a) \in T^E \), let \( \bar{m}(t^a, r^a) = \bar{m} \). Then

\[
(1) 0 < \bar{m}(t^a, r^a) < \bar{c};
\]

\[
(2) \beta = \frac{\bar{c}}{\bar{c} - \bar{m}} > 1;
\]

\[
(3) t^a < \frac{r^a}{\beta} < t^a;
\]

\[
(4) \frac{\bar{c} - \bar{m}}{\bar{m}} = \frac{r^a}{t^a - r^a}.
\]

**Proof.** (1) \( 2t^a > t^a > r^a \). Therefore \( 2t^a - t^a > t^a - t^a > 0 \). Hence \( 1 > (r^a - t^a)/(2t^a - t^a) > 0 \) as required.

(2) \( \bar{c} > \bar{m} > 0 \) from (1). Then \( \bar{c} > \bar{c} - \bar{m} > 0 \). Then \( \bar{c}/(\bar{c} - \bar{m}) > 1 \) as required. Actually \( \beta = (2t^a - t^a)/t^a = 2 - t^a/t^a \).

(3) \( \beta > 1 \Rightarrow \frac{r^a}{t^a} < \frac{t^a}{\beta} \). \( r^a < \frac{t^a}{\beta} \) iff \( r^a < t^a/(2t^a - t^a) \) iff \( 0 < \frac{t^a}{2t^a - t^a} \); that is, \( 0 < (t^a - t^a)^2 \), true \( \forall t^a \neq t^a \).

(4)

\[
\bar{c} - \bar{m} = \bar{c} - \left[ \frac{r^a - t^a}{2t^a - t^a} \right] = \bar{c} \left[ \frac{t^a}{2t^a - t^a} \right].
\]

Then

\[
\frac{\bar{c} - \bar{m}}{\bar{m}} = \bar{c} \left[ \frac{r^a}{2t^a - t^a} \right] \frac{r^a}{\bar{c} \left[ \frac{r^a - t^a}{2t^a - t^a} \right]} = \frac{r^a}{t^a - t^a}
\]

as required. ■
Proof of Proposition 2. 1. Consider the committee. For a committee of type \( t \in T_1 \),

\[
U^{c*}(t, m(t)) = 0, a_1 = \hat{h} - r^*(\hat{c} - \hat{m}), a_2
\]

\[
= \begin{cases} 
- r^*(\hat{c} - \hat{m}) - |c - h| & \text{if } t \geq t^* \text{ (accept)} \\
- t\hat{c} - |c - h| & \text{otherwise (reject)}.
\end{cases}
\]

For a committee of type \( t \in T_2 \), \( U^{c*}(t, m(t)) = \hat{m}, a_1 = \hat{h} - r^*(\hat{c} - \hat{m}), a_2 = - \hat{m} - r^*(\hat{c} - \hat{m}) - |c - h| \). Does any committee have any incentive to deviate?

1. Say a type \( T_2 \) committee that in equilibrium signals \( \hat{m} \) deviates and signals 0 instead:

\[
U'(t \in T_2, 0, a_1 = \hat{h} - r^*(\hat{c} - \hat{m}), a_2')
\]

\[
= \begin{cases} 
- r^*(\hat{c} - \hat{m}) - |c - h| & \text{if } a_2' = \text{ accept} \\
- t\hat{c} - |c - h| & \text{otherwise}.
\end{cases}
\]

So the deviator accepts if \( r^* < \beta t \), and rejects otherwise.

Case a. If the deviator accepts, \( t \in [r^*/\beta, r^*] \). Then \( U^* \geq U' \) iff

\[
t\hat{m} + r^*(\hat{c} - \hat{m}) \leq r^*(\hat{c} - \hat{m}) \ \forall t \in \left[ \frac{r^*}{\beta}, r^* \right] \Rightarrow \frac{\hat{c} - \hat{m}}{m}
\]

\[
\geq \frac{t}{r^* - r^*} \ \forall t \in \left[ \frac{r^*}{\beta}, r^* \right]
\]

which is ensured by 4 of Lemma 5.

Case b. If the deviator rejects, \( t \in [r^*, r^*/\beta] \). Then \( U^* \geq U' \) iff

\( t(\hat{c} - \hat{m}) \geq r^*(\hat{c} - \hat{m}) \ \forall t \in [r^*, r^*/\beta] \) which is always satisfied.

2. Suppose a type \( T_1 \) committee that in equilibrium signals 0 deviates and signals \( \hat{m} \) instead:

\[
U^{c'} = \begin{cases} 
- t\hat{m} - r^*(\hat{c} - \hat{m}) - |c - h| & \text{if } a_2' = \text{ accept} \\
- \hat{m} - t\hat{c} - |c - h| & \text{otherwise}
\end{cases}
\]

and \( a_2' = \text{ accept} \) if \( r^* \leq t\beta \).

Case a. If the committee accepts in equilibrium, then \( t > r^* > r^*/\beta \). So this type (\( t \in (r^*, 1) \)) accepts in deviation. Then \( U^* \geq U' \) iff

\[
- r^*(\hat{c} - \hat{m}) \geq - t\hat{m} - r^*(\hat{c} - \hat{m}) \ \forall t \in (r^*, 1] \Rightarrow \frac{\hat{c} - \hat{m}}{m}
\]
\[
\frac{t}{r^o - \bar{r}^o} \leq \frac{t}{r^o - \bar{r}^o} \quad \forall t \in (r^o, 1]
\]

which is ensured by 4 in Lemma 5.

Case b. If \( t < r^o \), the committee accepts in equilibrium. If the committee
accepts in deviation, \( t \geq \frac{r^o}{\beta} \), then \( U^* \geq U' \) iff \(-\bar{t} \geq -\bar{m} - r^o(\bar{c} - \bar{m}) \) \( \forall t \in [r^o/\beta, r^o] \Rightarrow t(\bar{c} - \bar{m}) \leq r^o(\bar{c} - \bar{m}) \) \( \forall t \in [r^o/\beta, r^o] \), which is always true.

Case c. The committee rejects in equilibrium and rejects in deviation. Then \( t \in [0, r^o/\beta] \), and then \( U^* \geq U' \) iff \(-\bar{t} \geq -\bar{m} - \bar{t} \bar{c} \) \( \forall t \in [0, r^o/\beta] \) which is always true.

3. Any other deviation, \( m' > 0 \), leads to the out-of-equilibrium belief
that type 1 has deviated. Action \( a_1 = \bar{h} - \bar{m} \) is taken. Any type bigger
than \( r^o \) can do better merely by signaling \( m = 0 \), extracting \( a_1 = \bar{h} - r^o(\bar{c} - \bar{m}) \), and accepting. Similarly with any type smaller than \( r^o \). For \( t \in [r^o, \bar{r}^o] \), we have that signaling \( m = \bar{m} \) is preferred to signaling \( m = 0 \) and
inducing \( a_1 = \bar{h} - r^o(\bar{c} - \bar{m}) \). Then there can be no incentive to signal
any \( m' > 0 \) to induce a worse response.

II. Consider the agency.

1. On receiving \( m^*(t) = 0 \), the agency infers that \( t \in T_1 \). The agency’s
expected utility on sending \( a_1 = \bar{h} - t(\bar{c} - \bar{m}) \) is \( U^A(t, \bar{m}, a_1, a_2) = -(\bar{h} - a) + (\bar{c} - \bar{m}) t (1 - F(t \mid 0)) \). We need to show that \( r^o = \arg \max_{t \in T_1} t(1 - F(t \mid 0)) \). By Bayes’ Rule,

\[
1 - F(t \mid 0) = \frac{[1 - F(t)] [I_1 + I_2] + [1 - F(r^o)] I_2 + [F(r^o) - F(r^o)] I_1}{1 + F(r^o) - F(r^o)}
\]

where \( I_1 = 1 \) if \( t \in [0, r^o] \), 0 otherwise; \( I_2 = 1 \) if \( t \in [r^o, r^o] \), 0 otherwise; and \( I_2 = 1 \) if \( t \in (r^o, 1] \), 0 otherwise; Let \( \Psi^O(t) = t(1 - F(t \mid 0)) \). Recall
that \( \Psi(t) = t(1 - F(t)) \). In the interval \([r^o, 1]\), \( 1 - F(t \mid 0) = (1 - F(t)) / (1 + F(r^o) - F(r^o)) \). Then \( \Psi^O(t) = \Psi(t) / D \), where \( D = 1 + F(r^o) - F(r^o) \), and \( \arg \max_{t \in T_1} \Psi^O(t) = \arg \max_{t \in T_1} \Psi(t) \). Now \( \Psi'(t) < 0 \) \( \forall t \in (r^o, 1) \) from Lemma 1 above. Hence argmax_{t \in (r^o, 1]} \Psi^O(t) = r^o. Hence \( r^o = \arg \max_{t \in (r^o, 1]} \Psi^O(t) \). In the low interval \( I_1 = [0, r^o] \), \( 1 - F(t \mid 0) = (1 - F(t) + F(r^o) - F(r^o)) / (1 + F(r^o) - F(r^o)) \). Then \( \Psi^O(t) = (\Psi(t) + [F(r^o) - F(r^o)]) / D \). Now \( \Psi'(t) = 1 - F(r^o) = F(r^o) \) \( \forall t \in [0, r^o] \) by Lemma 4 above. So \( \Psi^O(t) > 0 \) and so \( r^o = \arg \max_{t \in [0, r^o]} \Psi^O(t) \). Now all that remains is to compare \( \Psi^O(r^o) \) and \( \Psi^O(r^o) \):

\[
\Psi^O(r^o) = r^o \quad \frac{1 - F(r^o)}{1 - F(r^o) + F(r^o)} \quad \text{and} \quad \Psi^O(r^o) = r^o \quad \frac{1 - F(r^o)}{1 - F(r^o) + F(r^o)}.
\]
Now \( t^w > t^o \) so \( \Psi^0(t^w) > \Psi^0(t^o) \) so \( t^* = \arg\max_{t \in T} \Psi^0(t) \). Then the agency is optimizing, and there can be no incentive to deviate.

2. On receiving \( m^*(t) = \bar{m} \), the agency infers that \( t \in (t^o, t^w) \). \( U^{A*}(t, \bar{m}, a_1(t) = \bar{h} - t^o(\bar{e} - \bar{m}), a_2) = -(\bar{h} - a) + t^o(\bar{e} - \bar{m}) \) with certainty. Is there any incentive to deviate here? Suppose on receiving \( m(t) = \bar{m} \), the agency deviates to \( a_1' = \bar{h} - t'(\bar{e} - \bar{m}) < \bar{h} - t^o(\bar{e} - \bar{m}) \). That is, the agency picks a \( t' > t^o \). Then acceptance is not ensured. \( U^A(t, \bar{m}, a_1', a_2) = -(\bar{h} - a) + t'(\bar{e} - \bar{m})(1 - F(t' | \bar{m})) \). Then \( U^* \geq U' \) iff \( \bar{r}_o(\bar{e} - \bar{m}) \geq t'(\bar{e} - \bar{m})(1 - F(t' | \bar{m})) \) \( \forall t' \in (t^o, t^w) \); that is, iff

\[
\frac{t^w}{t^o} \geq \frac{F(t^o) - F(t')}{{F(t^o) - F(t^w)}} \forall t' \in (t^o, t^w).
\]

Denote \( LHS(t') = \frac{t^w}{t^o}t'; RHS(t') = \frac{[F(t^o) - F(t')][F(t^o) - F(t^w)]}{F(t^o) - F(t^w)} \). Then \( LHS(t^o) = RHS(t^o) = 1 \), and \( LHS(t^w) = \frac{t^w}{t^o} > 0 = RHS(t^w) \). All we need to check is that

\[
\left| \frac{\partial LHS(t')}{\partial t'} \right|_{t' = t^o} \leq \left| \frac{\partial RHS(t')}{\partial t'} \right|_{t' = t^o}.
\]

This is true iff \( t^w f(t^o) \geq F(t^o) - F(t^w) \) iff \( F(t^o) - F(t^w) \leq 1 - F(t^o) \) which is satisfied by Lemma 3 above.

3. Any deviation in the opposite direction does not increase the probability of acceptance (it is already 1) and moves the policy further from \( a \). This then cannot be any incentive to deviate. ■

**Proposition 3.** No two-group-equilibrium exists.

**Proof.** Properties of two-group-equilibria: \( \exists t^* \in T \) such that \( T \) is partitioned into two intervals, \( T_1 = [0, t^*] \) and \( T_2 = [t^*, 1] \). Within each group the types pool on the same signal, but the two groups have different signals.

Case 1. \( m^*(t) = 0 \ \forall t \in T_1 \) and \( m^*(t) = \bar{m} > 0 \ \forall t \in T_2 \). In equilibrium, all types that signal \( m > 0 \) must expect to receive a policy they can be sure to accept. Hence the equilibrium policy must correspond to the left edge of the messaging-type interval \( T_2 \). In this case \( a^*_t(m(t) = \bar{m}) = \bar{h} - t^o(\bar{e} - \bar{m}) \).

In the interval of zero signalers, \( T_1 \), the agency must choose a policy corresponding to a type that maximizes the expected utility with respect to the posterior distribution of types. This is equivalent to finding the type that maximizes \( \Psi^0(t) = t[1 - F(t | 0)] \) in the interval \( T_1 \). Call this type \( \bar{t} \). Then \( a^*_t(m(t) = 0) = \bar{h} - \bar{t}^o(\bar{e} - \bar{m}). \) Now \( t^o \leq t^* \), so \( a^*_t(m(t) = 0) \geq a^*_t(m(t) = \bar{m}). \)
Consider the types in $T_2$ that signal $\bar{m} > 0$, receive $a_t^*(m(t)) = \bar{m} = h - t(e - \bar{m})$ and accept. These types would be clearly better off by signaling 0 and receiving the higher $a_t^*(m(t) = 0) = h - t(\bar{e} - \bar{m})$ and accepting. Hence the positive signalers in any two-type equilibrium have an incentive to deviate in this case.

Case 2. $m^*(t) = \bar{m} \ \forall t \in T_1$ and $m^*(t) = 0 \ \forall t \in T_2$. In return, the messaging committees must receive offer $a_t^*(m(t) = \bar{m}) = h - 0(\bar{e} - \bar{m}) = h$, in order that all messengers accept. That would in turn imply that

$$0 = \arg\max_{t \in [0,t^2]} t(1 - F(t \mid \bar{m})) = \arg\max_{t \in [0,t^2]} \Psi(t).$$

Then

$$\frac{\partial \Psi(t)}{\partial t} \leq 0 \ \forall t \in [0, t^2).$$

Now

$$\frac{\partial \Psi(t)}{\partial t} = \frac{\Psi'(t) - [1 - F(t^2)]}{F(t^2)} \Rightarrow \Psi'(t) \leq 1 - F(t^2) \ \forall t \in [0, t^2).$$

But $\Psi'(0) = 1 > 1 - F(t^2)$, which is a contradiction. Hence $a_t^*(m(t) = \bar{m}) = h$ cannot be a best response for the agency. ■

**Proposition 4.** No separating equilibrium exists.

**Proof.** By contradiction. Let $m^*: T \rightarrow [0, E]$ be the equilibrium signaling function. In response to $m^*(t)$ the agency ascertains that the committee type is $t$ and responds with the pivot: $a_t^* = h - t(e - m^*(t))$. Equilibrium utility of the committee is $U^C(t, m, a_1, a_2) = -e - |c - h|$. If the committee deviates, say pretends to be type $t' < t$, and sends $m'$, then $U^C(t, m', a_1(m'), a_2) = -tm' - t'(e - m') - |c - h|$. In a separating equilibrium, $\forall t \in T, U^C \geq U^{C'}$. That is $-t\bar{e} \geq -tm' - t'(\bar{e} - \bar{m'}).$ Then $-t(\bar{e} - \bar{m'}) \geq -t'(\bar{e} - \bar{m'})$ and $t \leq t'$. The contradiction is established. ■

**Proposition 5.** The partial pooling equilibrium and the total pooling equilibrium are universally divine equilibria.

**Proof.** 1. The Partial Pooling Equilibrium. Restrict out-of-equilibrium beliefs to place weight 1 on type $t = t^*$. Let $a_t(\cdot)$ be the equilibrium offer in the second period. Let $0 < m < \bar{m}$ be an out-of-equilibrium signal (a similar analysis applies if $m > \bar{m}$), and define $a(m, i)$ as the offer that makes type $t$ indifferent between staying on the equilibrium path and
deviating to \( m \). Thus \( a(m, t) \) solves \( a(m, t) - tm = \hat{h} - t\bar{e} \) if \( t < t^* \); \( a(m, t) - tm = \hat{h} - t^t(\bar{e} - \bar{m}) - \bar{m} \) if \( t^* \leq t \leq t^* \); and \( a(m, t) - tm = \hat{h} - t^t(\bar{e} - \bar{m}) \) if \( t > t^* \). Given \( m, a(m, t) \) is minimized at \( t = t^* \); i.e., \( t^* \) is the type with the most to gain from deviation. Hence the agency must put posterior weight of 1 on type \( t^* \). Then the agency offers \( a_1 = \hat{h} - t^t(\bar{e} - \bar{m}) \) and \( U_C(t, m) = \hat{h} - t^t(\bar{e} - \bar{m}) - tm \). This utility is inferior to that obtained in equilibrium by any type: for \( t \in (t^*, 1] \), \( \hat{h} - t^t(\bar{e} - \bar{m}) > \hat{h} - t^t(\bar{e} - \bar{m}) - tm \) so \( U^*_C > U^C \). From Lemma 5, \( -\bar{m}(2t^t - t^e)/t^t > 0 \). For \( t \in [t^t, t^e] \), \( m > 0 \) implies \( t^m(t^t - t^e) > 0 \). Then \( \bar{m}(t^t - 2t^e) + \bar{e}(t^t - t^e) + t^m > 0 \) implies \( F(t^t, \bar{e} - \bar{m}) > \bar{h} - t^t(\bar{e} - \bar{m}) \) for \( t^* \in \{t^*, t^e\} \). Again \( U^*_C > U^C \). For \( t < t^* \), \( -t(\bar{e} - \bar{m}) > -t^t(\bar{e} - \bar{m}) \), so \( \hat{h} - t\bar{e} > \hat{h} - t^t(\bar{e} - \bar{m}) - tm \). Again \( U^*_C > U^C \).

2. The Total Pooling Equilibrium. If \( m > \bar{e} \) then posterior weight is put on \( t = 0 \); if \( 0 < m \leq \bar{e} \), then \( t = t^* \). In each case, these are the types most likely to have deviated by sending the signal \( m \). Moreover, in each case, each type prefers to play the equilibrium strategy rather than deviate. So this equilibrium survives refinement by universal divinity.

REFERENCES


