

A Jamming Theory of Politics

William Minozzi¹
Department of Political Science
The Ohio State University
minozzi.1@osu.edu

April 21, 2010

Abstract

Competitive political elites frequently offer conflicting, irreconcilable accounts of policy-relevant information. This presents a problem for members of the public who lack the skill, time, and attention to become experts on every complicated policy question that might arise. To analyze problems like these, this article presents a formal theory of political communication with competitive senders who have privately known preferences. In equilibrium, senders can *jam* messages from their opponents; that is, they can send messages designed to leave receivers uncertain about who has sent a truthful message. The article identifies differences between jamming and existing theories, reports empirical predictions, and discusses substantive implications for the politics of representation, the judiciary, and expertise.

Forthcoming, *Journal of Politics*

¹I would like to thank David Austen-Smith, David Baron, Jon Bendor, Brandice Canes-Wrone, Cliff Carrubba, Jamie Druckman, Kevin Esterling, Keith Krehbiel, Skip Lupia, Mat McCubbins, Michael Neblo, Ken Shotts, David Siegel, Craig Volden, Alan Wiseman, and Jack Wright; seminar participants at Emory University, Florida State University, Northwestern University's Kellogg Graduate School of Management, New York University, the University of California, University of Michigan, and the University of Toronto; and three anonymous reviewers.

Clashing viewpoints are ubiquitous in democratic politics. Voters are bombarded by competing, sometimes conflicting messages from candidates during elections; legislators are lobbied by multiple interests on important bills; courts hear arguments from both sides and issue conflicting opinions; and regulators receive comments for and against proposed rules.

But competition is not the only quality common to democracies. Democratic politics is also beset by uncertainty about the motives of political actors. Whether politicians intend to poach from the public purse, limit civil liberties, concentrate wealth among a few, and so forth is unclear to voters. Citizens cannot easily distinguish honest politicians from scoundrels by their manner of dress, speaking, education, prior occupation, or other characteristics. There is no screening process that allows only honest politicians to take the stage. Many argue that competition resolves this conflict. But might not competition merely muddy the waters, leaving citizens as confused as ever?

For example, as in debates on most complicated political issues, conflicting evidentiary claims dominate the politics of climate change and carbon policy.² Environmental activists claim that climate change is caused by human behavior, while skeptics maintain that this claim is not grounded in scientific fact. Both sides boast that scientific experts are on their side. Pollster Frank Luntz propounded this communications strategy in an infamous 2002 memo: “Voters believe that there is *no consensus* about global warming within the scientific community. Should the public come to believe the scientific issues are settled, their views about global warming will change accordingly. ... The scientific debate is closing (against us) but [is] not yet closed. There is still a window of opportunity to challenge the science.

²The theory developed here applies to many, many strategic situations in democratic politics. However, for the sake of exposition, I use the language of elite discourse, public opinion, and climate change to animate the analysis that follows. In a later section, I consider how jamming extends to other political environments.

You [the skeptic] need to be even more active in recruiting experts who are sympathetic to your view, and much more active in making them part of your message” (Luntz Research Companies, 2001; emphasis in original).

It is difficult to know exactly why skeptics like Luntz take the position they do. Is it because they believe that the weight of scientific evidence gives good reason to doubt that humans can and do cause climate change, or because, regardless of the evidence, they believe that the costs of policies like cap-and-trade outweigh the benefits? On any complicated and controversial issue, non-expert citizens, including almost all of the public, face a challenge in discerning fact from fiction.

Uncertainty about motivations, and especially policy preferences, lies at the root of this problem. To see why, suppose (for the sake of argument) that climate change is a threat recognized privately by both activists and skeptics, but that the two disagree about the circumstances under which a program like cap-and-trade should be implemented. Policy preferences might concern the tradeoff between climate change mitigation and economic loss due to the program. The activist prefers to implement the program even if economic losses far outstrip the benefits of climate change reduction, although there is a point at which he³ would abandon the program. Similarly, the skeptic prefers not to implement the program even if this cost-benefit ratio is low, but there is a point at which the ratio is so low that he would embrace the program. Each point is the private information of that elite.

The elite whose side is aided by a piece of evidence can benefit from credibly revealing this information (if possible). Less clear is whether the opposing actor prefers to confirm this message. As an extreme counterfactual, suppose that reliable, yet dense and difficult to

³I use male pronouns for activists and skeptics (senders) and female pronouns for members of the public (receivers).

understand analysis yielded the prediction that sea levels will rise 100 feet within a decade due to human causes but that, with severe changes in policy, this disaster can be averted. It is reasonable to conjecture that, in this case, most skeptics would argue in favor of such a policy change. But what messages would these skeptics send if sea levels were predicted to rise 25 feet, 5 feet, 6 inches, or $1\frac{1}{2}$ inches over the same amount of time? How does this uncertainty about preferences affect the credibility of messages from activists? How would members of the public update their beliefs given these messages?

Many features of this environment are captured by a sender-receiver game, in which the elites are senders and a key member of the public is a receiver. But existing theories cannot address the above questions because they fail to consider either competitive senders or uncertainty about preferences. Although many works consider one or the other, none consider both simultaneously. Table 1 illustrates this pattern. To fill this gap, this article develops a new two-sender cheap talk game. In the game, there is uncertainty about the senders' preferences over the outcome of a policy process. This process is affected both by some feature of the world that only the senders see perfectly (e.g., the extent of human-driven change in sea levels) and an action taken by the receiver (e.g., forming an opinion about cap-and-trade legislation).

In equilibrium, one sender can tell a lie to prevent the receiver from learning information the other sender would like to reveal. I call this phenomenon *jamming*. For jamming to occur, the receiver must be uncertain who has lied and who has told the truth. To ensure this, it must be that, if a jamming message were actually true, the other sender might want to jam it and could do so using the truthful statement the current lie is intended to jam. Faced with jamming messages, the receiver cannot tell who has told the truth and who has lied. If allowed to say so, either sender could credibly claim that the other is lying.

Table 1 Formal Models of Political Communication

	Public Preferences	Private Preferences
Single Sender	Crawford and Sobel (1982) Gilligan and Krehbiel (1987) many others	Sobel (1985) Austen-Smith (1995) Lupia and McCubbins (1998) Morris (2001)
Multiple Senders	Gilligan and Krehbiel (1989) Austen-Smith and Wright (1992) Austen-Smith (1993) Krishna and Morgan (2001 <i>a,b</i>) Battaglini (2002, 2004)	This article

Jamming messages differ from the separating and pooling phenomena familiar from existing formal models (e.g., Gilligan and Krehbiel, 1989). In a separating equilibrium, the receiver observes identical messages from both senders and learns the underlying information perfectly. In pooling, both senders offer vague messages that prevent precise information from being revealed, and the receiver learns nothing precise about the underlying information. Jamming is different. As in separating equilibria, jamming messages are responsive to changes in underlying information, although they leave the receiver uncertain as to which of the two senders has been truthful. As in pooling equilibria, jamming messages prevent the receiver from becoming fully informed, although they do so by offering precisely counterprogrammed alternatives. Moreover, jamming messages differ from both other kinds because they directly contradict each other.

The Model and an Example

Consider a simplified model of elite discourse, in which two opposing elites are aware of some piece of policy-relevant information that is difficult to convey to the mass public. The

elites also privately know whether they want to reveal this information to the public, given the potential political impact of rendering it uncontroversial. Both sides craft messages for consumption by a representative member of the public who draws what information she can from the messages and updates her opinion about the policy at hand. Of particular interest is how elites craft these messages and how the public responds.

To put more structure on this idea, label the elites as senders 1 and 2, and the member of the public as the receiver. Policy-relevant evidence is represented by w , the realization of a random variable that Nature randomly chooses according to F_w , an atomless, strictly increasing distribution with support on the interval $[0, 1]$. Each sender $i \in \{1, 2\}$ observes w and sends a message $m_i \in [0, 1]$ to the receiver. Messages are simultaneous and have no direct effect on the payoffs. Upon receipt of messages m_1 and m_2 , the receiver updates her beliefs about w and forms an opinion $y \in [0, 1]$ about which policy is best.

To determine how beliefs and opinions depend on messages, we need assumptions about the relationship between policies and outcomes, and about preferences. The first assumption is familiar from many incomplete information models of politics.

Linear Outcomes *Given w and y , the outcome $x = w - y$.*⁴

Each player has single-peaked preferences and a unique most-preferred outcome, or type, x_i . The receiver's type is normalized to 0. Assume players' preferences are Euclidean, so that a player with type x_i prefers x to x' if and only if x is closer to x_i than is x' . Preferences that

⁴Other theories that assume Linear Outcomes use the functional form $x = w + y$. I prefer $x = w - y$ because (as we will see) it ensures that senders with higher types prefer to jam with higher messages. In substantive terms, w might be the adverse impact of climate change, while y is the costly abatement of this change. Both functional forms satisfy the important property that x changes monotonically in w and y , and yield logically equivalent results.

satisfy these properties can be represented by a linear-loss utility function.

Single-Peaked Preferences *A player with type x_i has utility $u(x_i, x) = -|x_i - x|$.*⁵

These two assumptions imply that for any w and y , a player with type x_i has utility $u(w, x_i, y) = -|x_i - (w - y)|$. While the receiver may be thought of as preferring to learn the truth about w , senders may be thought of as preferring to influence public opinion.

Senders' types are common knowledge in most sender-receiver games. In contrast, assume that x_i , the type of sender i , is the privately observed realization of a random variable with the distribution F_i . Assume that F_i is increasing on support X_i with a hazard rate that increases in distance from 0.⁶ The next assumption focuses attention on senders from opposing sides.

Opposing Senders *For $\bar{x} \in (0, \frac{1}{4}]$, $X_1 = [-\bar{x}, 0]$ and $X_2 = [0, \bar{x}]$.*⁷

Essentially, this assumption means that the receiver knows who prefers lower outcomes (sender 1) and who prefers higher outcomes (sender 2), but not who has preferences closer to her own. This assumption sets weak limits on the situations covered by this theory; for example, it excludes cases in which the senders may or may not be in opposition. But the range of political situations that can be described by these three assumptions remains vast.

Each sender i uses a message strategy $m_i(w, x_i)$. A message is essentially equivalent to both, "The value of w is m_i ," and, "You should prefer the policy m_i ." The receiver uses an opinion formation strategy $y(m_1, m_2)$ and has beliefs given by a probability function $h(\cdot|m_1, m_2)$. In words, when the receiver observes messages m_1 and m_2 , she believes that w is w_0 with probability $h(w_0|m_1, m_2)$. The equilibrium concept is perfect Bayesian, which

⁵The results hold for all symmetric, single-peaked functions, including quadratic-loss.

⁶A distribution satisfies this property if the conditional density of x_i given $|x_i| > x$ is always increasing in $x > 0$. Many distributions (e.g., uniform, truncated normal) have this property.

⁷Since F_i is monotone increasing in distance from 0, $\frac{f_1(x)}{F_1(x)}$ is decreasing and $\frac{f_2(x)}{1-F_2(x)}$ is increasing. The upper limit $\frac{1}{4}$ ensures that almost all types prefer to reveal at least one state.

requires (1) posterior beliefs to be calculated via Bayes' rule when possible and (2) strategies to maximize expected utilities given beliefs and others' strategies. The following example illustrates when and why jamming messages are possible and optimal.

An Example. Suppose x_1 is uniformly distributed on $[-\bar{x}, 0]$, x_2 is uniformly distributed on $[0, \bar{x}]$, and w is uniformly distributed on $[0, 1]$. There is an equilibrium with the following strategies and beliefs. Sender 1 sends $m_1 = 1 - w$ for $w \in (\frac{1}{2} - 2|x_1|, \frac{1}{2})$ and $m_1 = w$ otherwise, and sender 2 sends $m_2 = 1 - w$ for $w \in (\frac{1}{2}, \frac{1}{2} + 2|x_2|)$ and $m_2 = w$ otherwise. If the receiver observes confirmatory messages ($m_1 = m_2$), she chooses $y = m_1 = m_2$. If she observes conflicting messages ($m_1 \neq m_2$), she chooses $y = \frac{1}{2}$. If $m_1 = m_2$, the receiver believes that $w = m_1 = m_2$ with probability 1. If $m_1 = 1 - m_2$, the receiver believes that $w = m_1$ with probability $h(m_1) = \frac{2\bar{x} - (m_1 - \frac{1}{2})}{4\bar{x} + m_2 - m_1}$ and that $w = m_2$ with probability $h(m_2) = 1 - h(m_1)$.⁸

In this example, a sender jams when he sends $1 - w$ rather than the truthful message w . For a pair of messages to emerge in equilibrium, senders must correctly evaluate the consequences their messages, and the receiver must behave as the senders conjecture she will. In particular, the receiver must form the opinion $y = \frac{1}{2}$ if she observes conflicting messages. Since at least one sender always sends the truthful message, a jamming message must offset the impact of the truthful message on the receiver's beliefs. To see why the strategies and beliefs in the example constitute an equilibrium, consider each player's incentives.

First, the message strategies of senders reflect incentives generated by the opinion formation strategy of the receiver. Sender 1, whose type is $x_1 \leq 0$, prefers to induce the receiver to choose $y = \frac{1}{2}$ rather than $y = w$ if and only if $w \in (\frac{1}{2} - 2|x_1|, \frac{1}{2})$.⁹ Since sender 1 has private

⁸Many off-the-equilibrium-path beliefs sustain this equilibrium. One example is, if $m_1 \neq m_2$ and $m_1 \neq 1 - m_2$, the receiver believes w is uniformly distributed on $[\frac{1}{2} - 2\bar{x}, \frac{1}{2} + 2\bar{x}]$.

⁹To see why this is true, compare $u(w, x_1, y = w) = -|x_1|$ and $u(w, x_1, y = \frac{1}{2}) = -|x_1 - (w - \frac{1}{2})|$. Note that

information about x_1 , only he knows the exact interval on which he jams. To an outsider, it is possible that sender 1 jams the entire interval $(\frac{1}{2} - 2\bar{x}, \frac{1}{2})$ in the case that $x_1 = -\bar{x}$, or no states at all in the case that $x_1 = 0$. Similarly, sender 2, whose type is $x_2 \geq 0$, prefers $y = \frac{1}{2}$ to $y = w$ if and only if $w \in (\frac{1}{2}, \frac{1}{2} + 2|x_2|)$. Thus, both senders prefer to reveal w (and they know they both prefer to reveal w , etc.) for any $w \leq \frac{1}{2} - 2\bar{x}$ and for any $w \geq \frac{1}{2} + 2\bar{x}$. For these extreme values of w , it is a best response for each sender to issue a truthful message. For values of w near the middle, $w \in (\frac{1}{2} - 2\bar{x}, \frac{1}{2} + 2\bar{x})$, one sender always prefers $y = w$ to $y = \frac{1}{2}$. If $w < \frac{1}{2}$, then sender 2 prefers to reveal w , and if $w > \frac{1}{2}$, sender 1 prefers to reveal w . Moreover, for each $w \in (\frac{1}{2} - 2\bar{x}, \frac{1}{2} + 2\bar{x})$, it is possible that the other sender prefers to reveal w as well. As a consequence, it is always optimal for at least one sender to try to reveal w by sending a truthful message. The only potential downside of sending a truthful message is that the receiver might choose $y = \frac{1}{2}$ if the opposing sender jams. But since $y = \frac{1}{2}$ is the only opinion the receiver forms except for $y = w$, this downside is really no downside at all.

The consequence of these considerations is that the message strategies in the example are optimal, conditional on the receiver using the strategy specified for her. Because the receiver has symmetric, single-peaked preferences, she maximizes her utility by choosing y equal to $E(w|m_1, m_2)$, her expected value of w given the messages she received and her posterior beliefs given those messages. To understand why it is optimal for the receiver to use her specified strategy, it is necessary to understand why she forms the beliefs she does.

It is straightforward to see why the receiver chooses $y = m_1 = m_2$ whenever she observes confirmatory messages. Because the senders oppose each other, the only values of w on which they can agree are those that are sufficiently extreme for both to prefer $y = w$ to $y = \frac{1}{2}$. $-|x_1 - (w - \frac{1}{2})| = -(x_1 - (w - \frac{1}{2}))$ iff $\frac{1}{2} - |x_1| > w$. For $\frac{1}{2} - |x_1| > w$, $u(w, x_1, y = \frac{1}{2}) > u(w, x_1, y = w)$ iff $w > \frac{1}{2} - 2|x_1|$. Similarly, for $w > \frac{1}{2} - |x_1|$, $u(w, x_1, y = \frac{1}{2}) > u(w, x_1, y = w)$ iff $\frac{1}{2} > w$.

This is even possible for values of w near the middle, $w \in (\frac{1}{2} - 2\bar{x}, \frac{1}{2} + 2\bar{x})$. For example, if w is only slightly greater than $\frac{1}{2} - 2\bar{x}$ and sender 1 has a type x_1 very close to 0, then both senders prefer to send truthful messages.

In contrast, if the receiver observes messages m_1 and $m_2 = 1 - m_1$, she cannot infer who has been truthful and who has jammed. Instead, she forms beliefs using her information about the preferences of the senders and her conjectures about message strategies. From her perspective, the probability that, say, sender 1 has been truthful corresponds to the probability that sender 2 would want to jam m_1 . As determined above, this is the probability m_1 is in the interval $(\frac{1}{2}, \frac{1}{2} + 2x_2)$. Since x_2 is uniformly distributed, this probability is $1 - \frac{1}{2\bar{x}}(m_1 - \frac{1}{2})$. Similarly, the receiver believes that sender 2's message is truthful with the probability that m_2 is in $(\frac{1}{2} - 2|x_1|, \frac{1}{2})$, which is $1 - \frac{1}{2\bar{x}}(\frac{1}{2} - m_2)$.¹⁰ Therefore, given m_1 and m_2 , the receiver believes that the w is m_1 with the probability

$$h(m_1) = \frac{1 - \frac{1}{2\bar{x}}(m_1 - \frac{1}{2})}{1 - \frac{1}{2\bar{x}}(m_1 - \frac{1}{2}) + 1 - \frac{1}{2\bar{x}}(\frac{1}{2} - m_2)} = \frac{2\bar{x} - (m_1 - \frac{1}{2})}{4\bar{x} - m_1 + m_2}.$$

In addition to being the probability with which the receiver believes the w is m_1 , $h(m_1)$ is also the probability the receiver assigns to the event that sender 2 lied. Given $h(m_1)$, the receiver also believes that w is m_2 with probability $h(m_2) = 1 - h(m_1) = \frac{2\bar{x} - (\frac{1}{2} - m_2)}{4\bar{x} - m_1 + m_2}$.

These beliefs motivate the receiver to use the strategy specified in the example. If the receiver observes messages m_1 and $m_2 = 1 - m_1$, she assigns probability $h(m_1)$ to the event that w is m_1 and probability $h(m_2) = 1 - h(m_1)$ to the event that w is m_2 . Substituting in

¹⁰To be more clear, since the receiver knows m_1 but is uncertain about x_2 , she believes that $\Pr(m_1 \leq \frac{1}{2} + 2x_2) = \Pr(\frac{1}{2}(m_1 - \frac{1}{2}) \leq x_2)$. Because x_2 is uniformly distributed on $[0, \bar{x}]$, $\Pr(\frac{1}{2}(m_1 - \frac{1}{2}) \leq x_2) = 1 - \frac{1}{2\bar{x}}(m_1 - \frac{1}{2})$. Similarly, since the receiver knows m_2 but not x_1 , she believes $\Pr(\frac{1}{2} - 2|x_1| \leq m_2) = \Pr(x_1 \leq -\frac{1}{2}(\frac{1}{2} - m_2)) = 1 - \frac{1}{2\bar{x}}(\frac{1}{2} - m_2)$ because x_1 is uniformly distributed on $[-\bar{x}, 0]$.

$1 - m_1$ for m_2 and $1 - h(m_1)$ for $h(m_2)$, and doing the algebra yields the expectation

$$E(w|m_1, m_2) = m_1 h(m_1) + (1 - m_1)(1 - h(m_1)) = \frac{1}{2}, \quad (1)$$

which is precisely what is prescribed by her original strategy. This completes the explanation of why the strategies and beliefs in the example constitute an equilibrium.

The jamming message in the example may seem arbitrary, but it is actually tailored for this situation. A jamming message must be chosen so that, if it were truthful, the first sender might want to jam it, and would do so by using the truthful message that it is intended to jam. In the example, $m_J = 1 - w$ has exactly that effect. To see why, suppose sender 1 prefers to jam some $w < \frac{1}{2}$ and sends the message $m_J = 1 - w > \frac{1}{2}$. If w actually were m_J , sender 2 could jam m_J with the message $w = 1 - m_J$, which 2 sent in the first place.

In Figure 1, which illustrates this example, the middle, “jamming” interval is indicated by the heavy black line on the horizontal axis. Truthful messages are represented by the lines with positive slope, while jamming messages are represented by the lines with negative slope. All possible jamming messages that can be sustained in equilibrium are illustrated by the light gray line, although all jamming messages are not necessarily used in any given game. Jamming is possible in the range corresponding to the two light gray triangles. The height of these triangles above any particular value of w indicates the amount of disagreement between jamming messages for that value of w . Which jamming messages are used depends on the type of each sender, which is the private information of that sender. This is illustrated in the figure for $\bar{x} = \frac{1}{8}$, $x_1 = -\frac{1}{20}$, and $x_2 = \frac{1}{10}$. Sender 1, whose messages appear as a solid line, sends jamming messages for w between $\frac{1}{2} - 2\left|-\frac{1}{20}\right|$ and $\frac{1}{2}$, and sends truthful messages for all other w . Similarly, sender 2, whose messages are dashed, sends jamming messages for w between $\frac{1}{2} - 2\left|\frac{1}{10}\right|$ and $\frac{1}{2}$, and sends truthful messages for all other w .

Before moving to more complete analysis of the model, several features of jamming are worth emphasizing. First, in related formal theories, truthful messages appear in separating equilibria, in which senders use different messages to distinguish between different values of policy-relevant information, w , and thereby reveal precise information to the receiver (e.g., Gilligan and Krehbiel, 1989). In the complementary pooling equilibria, senders select a single, constant, pooling message to send in each of a large range of values of w , thus preventing the receiver from learning which particular value in the pool is true. Jamming messages share similarities with each of these other sorts of messages. Like truthful messages, jamming messages are not a constant function of w ; instead, they change continuously. Like messages in pooling equilibria, jamming messages prevent the receiver from learning what is true, namely the information corresponding to the messages that jam each other.

Second, the receiver uses each sender's message to validate the other's. Information can only be credibly transmitted to the receiver if each sender confirms the other's message. To see why, suppose there is a w in which sender 1 can credibly convey the value of w by himself, without confirmation. If sender 1 does so, the receiver infers w and chooses $y = w$. But then sender 1 has the incentive to induce $y = w$ by sending that message even when it is not. For example, if sender 1 has type x_1 , he can send the message when the true value is $w + x_1$. The receiver would still choose $y = w$, which is sender 1's ideal most preferred outcome. Therefore, when sender 1 sends this message, the receiver has reason to suspect that w is false. Hence, sender 1 cannot credibly convey the value of w by himself; both senders must prefer to reveal the information if it is to be revealed at all. This sets the stage for jamming, since only one sender must object for the policy-relevant information to remain hidden.

Third, activities like elite discourse are seldom limited to two individuals, and the model should not be interpreted as requiring only one pair of identical messages among many more

conflicting to yield policy change. Instead, the model sets limits on the ability of senders to mislead the receiver. Only consensus among senders can convince the receiver that a message is true. Consequently, a sender can cast doubt on a truthful message sent by his opponent by refusing to send a confirmatory message.

Fourth, senders make dual allegations, one explicitly focused on policy-relevant information and the other implicitly asserting that anyone who says otherwise is lying. Jamming requires that, when senders' message conflict, the receiver does not know who sent the jamming message. Thus, sending a jamming message is akin to alleging that the other sender is jamming and therefore has an extreme type. The dual nature of messages allows the receiver to use information about senders to inform her beliefs about policies.

Equilibrium Characterization

Although the above example illustrates many properties of jamming, a richer model allows for a more general equilibrium characterization and the derivation of empirical implications. This section provides such an analysis in three steps. First, I formally define what it means for a message to be jamming. Next, I present general strategies and beliefs that use jamming messages. Finally, I discuss the existence result that undergirds jamming equilibria.

The first task is to formally define a jamming message. It is useful to contrast jamming messages with truthful messages, which convey information without distortion.

Definition. *For any w , the **truthful message** is simply w .*

The truthful message is uniquely determined by w . The same is not obviously true for jamming messages. The latter sort is constrained by its principal use: jamming messages inhibit acceptance of truthful messages by leaving the receiver uncertain who has been truthful.

Therefore, each jamming message must satisfy a reciprocity condition. If the jamming message for w is m_J , then w must be a jamming message in the counterfactual case that m_J is actually truthful. This requirement is formalized in the following definition:

Definition. A **jamming function** $m_J : W \rightarrow W$ is *self-invertible*, i.e., $m_J(m_J(w)) = w$ for $w \in W$. For any $w \in W$, a **jamming message** is $m_J(w)$.

The jamming messages from the example are generated by a jamming function that satisfies this definition. In the example, the jamming function is $m_J(w) = 1 - w$. To verify that this is a jamming function, observe that $m_J(m_J(w)) = 1 - (1 - w) = w$.

For it to be possible for a sender to use a jamming message to deceive the receiver in equilibrium, that message must have consequences that the sender prefers to those of revealing the truth. Suppose (as in the example) that there is a default y_d , which might represent opinion stasis or the status quo policy. In the example, $y_d = \frac{1}{2}$, but in the general model y_d can take a range of values. Given y_d and a jamming function $m_J(\cdot)$, the following set of strategies generalizes those in the example.

$$\begin{aligned}
m_1(w, x_1) &= \begin{cases} m_J(w) \\ w \end{cases} \text{ if } w \in \begin{cases} (y_d - 2|x_1|, y_d) \\ \text{otherwise} \end{cases}, \\
m_2(w, x_2) &= \begin{cases} m_J(w) \\ w \end{cases} \text{ if } w \in \begin{cases} (y_d, y_d + 2|x_2|) \\ \text{otherwise} \end{cases}, \text{ and} \\
y(m_1, m_2) &= \begin{cases} y_d \\ m_1 \end{cases} \text{ if } m_1 \begin{cases} \neq \\ = \end{cases} m_2. \tag{2}
\end{aligned}$$

According to these strategies, jamming messages are sent only if w lies in the jamming interval $W = (y_d - 2\bar{x}, y_d + 2\bar{x})$, but not necessarily for every $w \in W$ in any given game.

The incentives for senders to use jam depend on the receiver's response to jamming

messages. This response is based on beliefs that in turn reflect senders' message strategies. Because jamming messages leave her uncertain who has been honest, she countenances two possibilities. First, 1's message may be truthful, in which case w is m_1 . Because 1 and 2 sent conflicting messages, this means that 2 must have preferred to jam. The indifference cutpoint between truthtelling and jamming is $\frac{1}{2}(m_1 - y_d)$. Thus, the receiver assigns the (conditional) probability that $x_2 > \frac{1}{2}(m_1 - y_d)$, which is $1 - F_2\left(\frac{1}{2}(m_1 - y_d)\right)$, to the event that 2 jammed. Using similar reasoning, the receiver assigns the (conditional) probability $F_1\left(-\frac{1}{2}(y_d - m_2)\right)$ to the event that 1 jammed. Bayes' rule yields the following beliefs:

$$h(w|m_1, m_2) = \begin{cases} \frac{1 - F_2\left(\frac{1}{2}(m_1 - y_d)\right)}{1 - F_2\left(\frac{1}{2}(m_1 - y_d)\right) + F_1\left(-\frac{1}{2}(y_d - m_2)\right)} & \text{if } w = m_1. \\ 1 - h(m_1|m_1, m_2) & \text{if } w = m_2. \\ 0 & \text{if } w \notin \{m_1, m_2\}. \end{cases} \quad (3)$$

The content of a jamming message—the value of $m_J(w)$ given w —knits together these strategies and beliefs into an equilibrium. That is, the strategies and beliefs in (2) and (3) form an equilibrium if and only if the jamming messages senders use justify those strategies and beliefs. The crucial requirement is that y_d must be the expected value of w given the posterior beliefs that follow jamming messages.

Equilibrium Proposition.¹¹ *Let $W = (y_d - 2\bar{x}, y_d + 2\bar{x})$. The strategies in (2) and the beliefs in (3) constitute an equilibrium if and only if there exists a jamming function $m_J(\cdot) : W \rightarrow W$ such that for all $w \in W$,*

$$y_d = wh(w|w, m_J(w)) + m_J(w)(1 - h(w|w, m_J(w))). \quad (4)$$

Note the similarity between equation (4) and equation (1) from the example. Given the

¹¹Proofs are relegated to the appendix, where off-equilibrium-path beliefs also appear. The proposition describes the more straightforward of two cases.

existence of a jamming function that satisfies (4), it is optimal for the receiver to choose y_d after observing a pair of messages that jam each other. The existence of such a jamming function is reported by the following lemma.

Jamming Function Existence Lemma. *There exists a jamming function $m_J(\cdot) : W \rightarrow W$ such that for any $w \in W$, w and $m_J(w)$ satisfy equation (4) given the beliefs in (3).*

Under general assumptions, a sender can find a message that will jam his opponent's. Thus, jamming messages in the above example are more than mere curiosity. Rather, in a very wide range of cases the receiver's ability to learn what she needs to know in order to become informed can be stymied by an informed, interested party.

Empirical Implications

Several empirical predictions can be drawn from this framework. A first question is when is jamming most likely, as a function of the policy-relevant information w senders observe. As w gets closer to the default opinion y_d , more types of senders prefer to jam, and jamming therefore becomes more likely.

Moderate Information is More Likely to Be Jammed. *If $w < w' < y_d$ or $w > w' > y_d$, the probability of jamming is (weakly) larger for w' than for w .*

There is almost always a difference between y_d and the true value of w . When the two are close, at least one sender is more likely to prefer to jam because the gains from biasing the receiver's beliefs in his favored direction outweigh the costs of preventing policy from responding to the state of the world. In the context of climate change, the skeptic tolerates small man-made changes in the environment and jams information that would cause unwelcome changes in opinion and might lead to policy changes. When calamitous changes are

forecast, the skeptic is less apt to jam.

A second question is how jamming messages change as the policy-relevant information w changes. Returning to the example and to Figure 1, as w increases away from $y_d = \frac{1}{2}$, the jamming messages sent by sender 2 (represented by the dashed line) become more extreme in the opposite direction. This is a general feature of jamming.

Countervailing Jamming Messages. *More extreme policy-relevant information must be jammed by more extreme jamming messages. That is, $m_J(w)$ decreases as w increases.*

In the context of climate change, this would correspond to a skeptic claiming that sea levels rises are negligible when there are small increases and claiming that the sea levels are diminishing when there are larger increases. As policy-relevant information compels larger and larger shifts away from y_d , jamming messages become more and more extreme in the opposite direction so as to offset the effect of truthful messages. More dire warnings must be complemented with more strident defenses of their efficacy because, in equilibrium, the receiver's expectation cannot change.

While the previous implications depend on objectively observing policy-relevant information, the following empirical implications consider messages and public opinion. Not all senders react in the same way to the same information. More extreme information leads senders with relatively moderate preferences to change their messages abruptly, depending on their types. That is, for each type of sender, there is a tipping point, a value of w at which he switches from sending a jamming message to a truthful one.

Senders Have Tipping Points. *Each sender has a point at which more extreme information leads him to send truthful messages and less extreme information leads him to jam.*

The prediction emphasizes the connection between information about policy and information

about the preferences of senders. When there is disagreement in equilibrium, it is impossible for the receiver to extricate one from the other. A sender’s tipping point is essentially his indifference cutpoint between revealing the policy-relevant information and jamming. In Figure 1, two tipping points are shown. For sender 1, the tipping point is $\frac{1}{2} - 2 \left| -\frac{1}{20} \right|$. Thus 1 will jam any w between this point and $y_d = \frac{1}{2}$, and will reveal all w less than this point. Tipping points are empirically observable without observing policy-relevant information. For example, by 2007, pollster Luntz, who earlier so clearly articulated the jamming strategy used by climate change skeptics, had himself switched sides. He explained this change by saying, “[O]ver the last 10 years the science has been much clearer. The results have been much more comprehensive and I, like millions of Americans, have changed my point of view.”¹²

The next results focus on messages when one sender is more likely to be moderate than the other. Moderation is defined with respect to the receiver, so a more moderate sender is more likely than a more extreme sender to have preferences close to those of the receiver. Such a sender might be seen as more moderate because of past actions or associations. The following definition formalizes this intuition:

Definition. *Sender i is more moderate (equivalently, less extremist) than sender j if, for any $z > 0$, $\Pr(|x_i| < z) > \Pr(|x_j| < z)$. That is, i is more moderate if the probability that i ’s type is closer than z from the receiver’s type, 0 , is greater than the probability that j ’s type is closer than z from 0 .*

More moderate senders are simply more likely to have preferences like the receiver than more extreme senders. Because the term is set in probability terms, it is still possible that a sender who is “more moderate,” to actually have a privately known type that is further from the

¹²“Tony Jones speaks with US political consultant Frank Luntz.” *Lateline*, Australian Broadcasting Corporation, May 2, 2007.

receiver. The definition is a variation on first order stochastic dominance from the literature on choice under uncertainty.

The more moderate sender always sends messages that appear to be more extreme, in the sense that they are further from the default y_d .

Moderate Senders Send More Extreme Messages. *If i is more moderate than j , i sends more extreme messages than j , and the receiver places more weight on j 's messages.*

This prediction reflects the fact that the moderate sender is always less likely to jam than the extremist. Suppose sender 2 is more extreme and that he observes some $w > y_d$ that he prefers to jam. He knows his more moderate opponent, sender 1, will attempt to reveal w with $m_1 = w$. Consider what would happen if the 2 used the symmetric jamming message from above, $m_2 = 2y_d - w$, which is just as far from y_d as w is. The probability the receiver assigns to a message from a sender is the probability that the other sender would want to jam it. Because 2 is more extreme, he is always more likely to jam, and so the receiver would assign a lower probability to m_2 than to m_1 . The receiver's optimal choice for y is her expectation of w , a weighted average of the messages she receives. Therefore, she would choose to shade y toward m_1 and away from m_2 and y_d . But this cannot happen in equilibrium. Instead, sender 2 chooses a more moderate jamming message m_J , one between the symmetric message m_2 and y_d . Sending m_J rather than m_2 pulls the receiver's opinion back to y_d because the more moderate message is more likely to be jammed, which increases the weight the receiver places on 2's message.

Figure 2 illustrates this prediction. In the figure, sender 1 is more moderate than sender 2. The full extent of jamming messages is displayed, as though both senders were located at their most extreme types. Sender 1 always sends messages that appear to be more extreme than those sent by 2, regardless of who is telling the truth. Two examples are displayed.

First, if $w < y_d$, sender 2 sends a truthful message (represented by the dashed diagonal line), and sender 1 sends a jamming message (represented by the solid, curvy line). The jamming message sent by 1 is further from y_d than is the truthful message from 2. The corresponding $w' > y_d$ represents the counterfactual case in which 1's message is truthful and 2's is jamming. But these messages are the same as in the first case. Even though 1 is now telling the truth, she appears to be sending a more extreme message than 2.

A corollary is that as the preferences of the receiver change, the messaging behavior of senders also changes. Think of the receiver as the median member of the public. If the median's preferences change, for example, if she were to become more liberal, then she moves toward the sender on the left. This is akin to making the sender on the left more moderate relative to the public, and, consequently, leading that sender to issue more extreme messages, according to the previous result. The same logic can be applied to comparisons between public opinion in different constituencies or countries with the following prediction.

Audience Corollary. *If the median member of audience A has an ideal point to the left of the median member of audience B, the sender on the left (respectively, right) sends more extreme (moderate) messages to A than to B.*

Thus, the composition of the audience has an effect on the strategic calculus of senders. This allows us to apply the logic of jamming to different states, countries, demographic groups, or issue publics—in short, any group that constitutes an audience to some set of senders. Upon reflection, the Audience Corollary makes clear that, according to the jamming theory, elites and the public influence each other's behavior.

Discussion

The kernel of jamming is the simple idea that competing senders can exploit uncertainty about their preferences to manipulate receivers. This has deep implications for our understanding of public opinion, representation, and the potential for democratic legitimacy. For example, jamming provides microfoundations for Zaller’s (1992) important theory of public opinion. This is notable because Zaller’s theory is intended to be an empirical tool, not a model of rational public opinion. Zaller states that, “[he] makes no allowance for citizens to think, reason, or deliberate about politics” (p. 45). In fact, Zaller develops no microfoundations for the behavior of citizens or elites in his model; the former are assumed to be automatons, the latter reduced to information flows. Jamming provides strategic underpinnings for this theory, substantiating Zaller’s axioms using thinking, reasoning actors.

At least three similarities emerge between Zaller’s theory and jamming. First, equilibrium opinions in the jamming theory resemble Zaller’s Response Axiom (p. 49), which states that citizens answer survey questions by averaging across available considerations. In jamming, the opinion of the receiver is a weighted average of elite messages. Furthermore, the two types of message pairs that appear in jamming equilibria (confirming and conflicting) resemble Zaller’s one-message and two-message models, in which elites are exogenously assumed to send the relevant number of messages. Moreover, predictions about opinion in each case match equilibrium predictions of jamming. The “mainstream effect” (p. 98) from the one-message model resembles the full information opinion under confirming messages, and the “polarization effect” (p. 100) from his two-message model resembles the default opinion under conflicting messages. Unlike Zaller’s theory, however, these features are not assumed; they are equilibrium implications of the micro-level assumptions of the jamming theory.

But jamming also complicates our understanding of public opinion and representation. Consider whether elite actors persuade their constituents, or conform their actions to public opinion. Zaller (1992) and Lupia and McCubbins (1998) offer theories in which elites are leaders rather than followers. While the difference between the assumptions of the jamming theory and those of Lupia and McCubbins seem slight (two senders vs. one sender), the change implies that, while elites do lead, their messages change according to their audience's composition (the Audience Corollary). This distinction is meaningful. For example, in his examination of public opinion on European Union integration, Carrubba (2001) draws on Zaller (1992) and Lupia and McCubbins (1998) to develop the hypothesis that elite policy positions shape public opinion on integration. Instead, he finds strong support for the countervailing claim that elite messages follow public opinion. In the context of the jamming theory, given that different countries have audiences with differing distributions of public opinion, the Audience Corollary implies that elites adapt their leadership strategies to the preferences of their audience. This more nuanced understanding of the elite-mass public relationship can also be subject to empirical examination. Moreover, these predictions and others that are less amenable to empirical test (like Countervailing Jamming Messages) can be tested experimentally. In the lab, subjects would play the roles of sender and receiver, and messages chosen for a variety of underlying information could be compared.

Outside of public opinion research, the theory adds to our understanding of the politics of expertise and lobbying. In Austen-Smith and Wright's (1992) theory, multiple interest groups are necessary to reveal the information policymakers need. These theories do not account for lobbyists who offer conflicting advice in order to stymie policy changes. In contrast, Esterling (2004) argues that complex issues lend themselves to arguments in which, "[b]oth sides of the debate make instrumental arguments (often, pro and con partial truths), so

debate simply muddies the waters..., an uninformative pooling equilibrium” (p. 234), which is the goal of one side. His argument focuses on policy-relevant uncertainty, whereas the jamming theory highlights the role of preference uncertainty, and the conclusions differ in a subtle way. In the pooling equilibria Esterling has in mind, the receiver learns the same thing from messages sent in a large number of states. In contrast, when jamming occurs, the receiver learns that the state takes one of two values. Jamming is more informative than pooling, but not informative enough to reveal the true state of the world.

The jamming theory applies to institutions that divide informational labor among political actors, like the judiciary. In arguments before the Supreme Court, there may be considerable variation in how well the justices know the preferences of different litigants, providing fertile ground for jamming. Bailey, Kamoie and Maltzman (2007) apply a single-sender model to this two-sided, adversarial environment, and it would be fruitful to test their predictions against those of jamming. In the literature on judicial legitimacy, Vanberg (2001) and Staton (2006) formalize the role of the public in enforcing judicial decisions, and Carrubba (2009) provides a formal model in which a court builds authority endogenously through sanctions by a monitoring public. These theories assume that the public monitors elite behavior probabilistically, in a non-competitive setting. Applying the jamming theory, opposed elites would be able to jam public monitoring, complicating the legitimacy problem faced by the judiciary. Because this would require additional assumptions, it would be best to begin such analysis with a simplified example like the one presented above.

As it is, this theory is broadly applicable. But this basic model may be enriched to encompass even more political environments. First, both senders in this theory observe the policy-relevant information without error. Relaxing this assumption would admit a new dimension of strategic interaction, in which senders might seek to make correct predictions to

prove their competence to the receiver. It would be valuable to analyze the tradeoffs between this impulse and the incentive to affect the receiver's action considered here. Second, the model presents a reduced-form account of how the public shapes policy. So long as elites care solely about aggregate public opinion, or so long as the median voter theorem applies (in which case the member of the public represents that voter), the model is sufficiently general to be applied without amendment. Nevertheless, for many other purposes it would be valuable to embed the jamming model in more complicated models of politics.

While the logic of jamming applies to many political environments, it is important to identify when it does not apply. Jamming requires at least two things to be true. First, receivers cannot become experts themselves, perhaps because it is too expensive or difficult or because it requires an outlay of attention that is not rewarded sufficiently. If instead receivers could become informed, the potential for jamming would be limited. Second, receivers cannot participate in reasoned dialogue with senders. This means they cannot ask for rationales, or explore reasons for the contradictory messages senders offer. The combined force of these two "boundary conditions" makes clear that jamming is an unavoidable consequence of dividing informational labor in a democracy. Consequently, jamming can be mitigated by bridging this divide with institutions that help and encourage citizens to understand the reasons for policies, rather than just to rely on the opinions of experts.

Regardless of its applications, extensions, and potential for empirical tests, the final take-away point from this research is that jamming provides counterintuition to the idea that competition leads to a more fully informed society (e.g., Page and Shapiro 1992). According to the jamming theory, an increasingly competitive informational environment often provides more chances for jamming, rather than ideal conditions under which the truth will win out. Additional informational providers are empowered to cast an effective veto by disputing any

information they prefer not to have communicated. In practice, the jamming mechanism is at work when voters are unable to screen signal from noise in elite messages, or when legislators, executives, justices, and regulators are confronted by contradictory information from opposing organized interests.

Appendix

I first state five lemmas, proofs of which can be found in the web appendix.

Lemma 1. *In equilibrium, the receiver believes that the state is w with probability 1 only if $m_1 = m_2 = w$. If Opposing Senders is satisfied, then it is possible in equilibrium for such confirmatory messages to be sufficient for the receiver to maintain such a belief.*

Lemma 2. *If players use the strategies in (2), equation (4) must hold whenever $m_1 \neq m_2$.*

Lemma 3. *Assume w and $m_J(w)$ satisfy equation (4), and that the receiver plays the strategy in (2). If sender $i \in \{1, 2\}$ has type $x_i = 0$, he prefers to reveal w . If sender 1 has type $x_1 < 0$, he prefers to jam w if and only if $w \in (y_d - 2|x_1|, y_d)$, and if sender 2 has type $x_2 > 0$, he prefers to jam w if and only if $w \in (y_d, y_d + 2|x_2|)$. ■*

Lemma 4. *Assume w and $m_J(w)$ satisfy equation (4), and that the receiver plays the strategy in (2). If $w > y_d$, sender 1 prefers to send $m_1 = w$ and sender 2 prefers to send*

$$m_2 = \begin{cases} w \\ m_J(w) \end{cases} \text{ if } x_2 \begin{cases} < \\ > \end{cases} \frac{1}{2}(m_1 - y_d). \text{ If } w < y_d, \text{ sender 2 prefers to send } m_2 = w$$

$$\text{and sender 1 prefers to send } m_1 = \begin{cases} w \\ m_J(w) \end{cases} \text{ if } x_1 \begin{cases} > \\ < \end{cases} -\frac{1}{2}(y_d - m_2).$$

Lemma 5. *If senders use the strategies in (2) and $m_1 = m_J(m_2)$, the receiver's beliefs are described by (3).*

Jamming Function Existence Lemma. *If Opposing Senders is satisfied, there exists a default opinion y_d , a subset $W \subseteq (y_d - 2\bar{x}, y_d + 2\bar{x})$, and a jamming function $m_J : W \rightarrow W$, such that, (i) for any $w \in W$, w and $m_J(w)$ satisfy (4) and (3), and*

$$(ii) \ y_d = \frac{\int_{w \in [2\bar{x}, 1 - 2\bar{x}] \setminus W} w dF_w}{\int_{w \in [2\bar{x}, 1 - 2\bar{x}] \setminus W} dF_w}.$$

Proof. Let $y_d \in [2\bar{x}, 1 - 2\bar{x}]$. Use of $m_J(\cdot)$ in equilibrium requires that w and $m_J(w)$ satisfy

equations (4) and (3). That is,

$$y_d = wh(w|w, m_J(w)) + m_J(w) h(m_J(w)|w, m_J(w)), \quad (\text{A1})$$

where $h(\cdot|\cdot, \cdot)$ represents the jamming beliefs given by (3). Suppose $m_1 = m_J(w)$, so that equation (3) implies

$$h(w|w, m_J(w)) = \frac{F_1\left(-\frac{1}{2}(y_d-w)\right)}{1-F_2\left(\frac{1}{2}(m_J(w)-y_d)\right)+F_1\left(-\frac{1}{2}(y_d-w)\right)}, \text{ and} \quad (\text{A2})$$

$$h(m_J(w)|w, m_J(w)) = \frac{1-F_2\left(\frac{1}{2}(m_J(w)-y_d)\right)}{1-F_2\left(\frac{1}{2}(m_J(w)-y_d)\right)+F_1\left(-\frac{1}{2}(y_d-w)\right)}. \quad (\text{A3})$$

Algebraic manipulation of equation (A1) and the fact that $h(m_J(w)|w, m_J(w)) = 1 - h(w|w, m_J(w))$ yields

$$(y_d - w) h(w|w, m_J(w)) = (m_J(w) - y_d) h(m_J(w)|w, m_J(w)). \quad (\text{A4})$$

Substitute (A2) and (A3) into (A4) and multiply by the common denominator to obtain

$$\frac{1}{2}(y_d - w) F_1\left(-\frac{1}{2}(y_d - w)\right) = \frac{1}{2}(m_J(w) - y_d) \left(1 - F_2\left(\frac{1}{2}(m_J(w) - y_d)\right)\right). \quad (\text{A5})$$

All jamming functions that are used in equilibrium must satisfy (A5). The next part of the proof constructs a jamming function $m_J(\cdot)$ that satisfies equations (4) and (3) from the text, given $y_d \in [2\bar{x}, 1 - 2\bar{x}]$.

Define the functions $\delta_1(m_1) : (y_d, y_d + 2\bar{x}) \rightarrow (0, \bar{x})$ as the right side of (A5) and $\delta_2(m_2) : (y_d - 2\bar{x}, y_d) \rightarrow (0, \bar{x})$ as the left side of (A5), and consider two properties of $\delta_1(\cdot)$ and $\delta_2(\cdot)$.

(P1) Because $\frac{f_1(x)}{F_1(x)}$ is increasing in x and $\frac{f_2(x)}{1-F_2(x)}$ is decreasing in x , both $\delta_1(\cdot)$ and $\delta_2(\cdot)$ are concave and have unique maximizers, m_1^* and m_2^* respectively. (These maximizers solve $\frac{1}{2}(m_1^* - y_d) - \frac{1-F_2\left(\frac{1}{2}(m_1^*-y_d)\right)}{f_2\left(\frac{1}{2}(m_1^*-y_d)\right)} = 0$ and $\frac{1}{2}(y_d - m_2^*) - \frac{F_1\left(-\frac{1}{2}(y_d-m_2^*)\right)}{f_1\left(-\frac{1}{2}(y_d-m_2^*)\right)} = 0$, which are the virtual valuations familiar from auction theory.)

(P2) $\delta_1(y_d) = \delta_1(y_d + 2\bar{x}) = \delta_2(y_d) = \delta_2(y_d - 2\bar{x}) = 0$.

There are two cases to consider.

Case 1: $\delta_1(m_1^*) = \delta_2(m_2^*)$. Properties (P1) and (P2) imply that $\delta_1(\cdot)$ and $\delta_2(\cdot)$ are u-shaped. This has implications for the existence of inverses for (portions of) each. First, $\delta_1(\cdot)$ is strictly decreasing on the interval $(m_1^*, y_d + 2\bar{x})$, from its maximum $\delta_1(m_1^*)$ to $\delta_1(y_d + 2\bar{x}) = 0$, and $\delta_2(\cdot)$ is strictly increasing on $(y_d - 2\bar{x}, m_2^*)$, from $\delta_2(y_d - 2\bar{x}) = 0$ to its maximum $\delta_2(m_2^*)$. Thus, there exist inverses $\eta_1(\cdot) : (0, \delta_1(m_1^*)) \rightarrow (m_1^*, y_d + 2\bar{x})$ and $\eta_2(\cdot) : (0, \delta_2(m_2^*)) \rightarrow (y_d - 2\bar{x}, m_2^*)$ such that $\eta_i(\delta_i(m)) = m$, $i = 1, 2$. Note that $\eta_1(\cdot)$ is strictly decreasing and $\eta_2(\cdot)$ is strictly increasing. Using similar reasoning, (P1) and (P2) imply the existence of (different) inverses $\zeta_1(\cdot) : (0, \delta_1(m_1^*)) \rightarrow (y_d, m_1^*)$ and $\zeta_2(\cdot) : (0, \delta_2(m_2^*)) \rightarrow (m_2^*, y_d)$, such that $\zeta_i(\delta_i(m)) = m$, $i = 1, 2$. Note that $\zeta_1(\cdot)$ is strictly increasing and $\zeta_2(\cdot)$ is strictly decreasing.

Define $m_J(w) : W_1 \rightarrow W_1$ to be a function with domain and range $W_1 = (y_d - 2\bar{x}, m_2^*) \times (m_1^*, y_d + 2\bar{x})$ such that

$$m_J(w) = \begin{cases} \eta_1(\delta_2(w)) \\ \zeta_1(\delta_2(w)) \\ \zeta_2(\delta_1(w)) \\ \eta_2(\delta_1(w)) \end{cases} \text{ for } w \in \begin{cases} (y_d - 2\bar{x}, m_2^*) \\ (m_2^*, y_d) \\ (y_d, m_1^*) \\ (m_1^*, y_d + 2\bar{x}) \end{cases}. \quad (\text{A6})$$

To verify that $m_J(w)$ is a jamming function, note that it is self-invertible. For any $w \in (y_d - 2\bar{x}, m_2^*)$, $m_J(m_J(w)) = \eta_1(\delta_2(\eta_2(\delta_1(w)))) = w$. A similar argument applies for other w 's.

The left side of (A5) is equal to $\delta_2(w)$, and the right side is $\delta_1(m_J(w))$. This means that equation (A5) is solved by $m_J(w)$ and w if and only if $\delta_2(w) = \delta_1(m_J(w))$. But this is exactly what is required of $m_J(\cdot)$ by construction, since

$$\delta_1(m_J(w)) = \delta_1(\eta_1(\delta_2(w))) = \delta_2(w). \quad (\text{A7})$$

A similar argument shows that $m_1 = w'$ and $m_2 = m_J(w')$ solve (A5) for $w' \in (y_d - 2\bar{x}, m_2^*)$.

A substantially similar argument to that of the last three paragraphs extends these jamming messages to the entire space $(y_d - 2\bar{x}, y_d + 2\bar{x})$ using the inverses $\zeta_1(\cdot)$ and $\zeta_2(\cdot)$. These inverses are used to define $m_J(\cdot)$ on (m_2^*, m_1^*) as $m_J(w) = \eta_1(\delta_2(w))$ for $w \in (y_d - 2\bar{x}, m_2^*)$, and $m_J(w) = \eta_1(\delta_2(w))$ for $w \in (y_d - 2\bar{x}, m_2^*)$. The resulting jamming function $m_J(\cdot)$ along with the stipulation $m_J(y_d) = y_d$ completes the construction of m_J .

Case 2: $\delta_1(m_1^*) \neq \delta_2(m_2^*)$. Let $\delta_1(m_1^*) > \delta_2(m_2^*)$ without loss of generality. The construction method from Case 1 may be applied to $w \in (y_d - 2\bar{x}, m_2^*) \cup (\bar{m}_1, y_d + 2\bar{x})$, where $\bar{m}_1 \in (m_1^*, y_d + 2\bar{x})$ solves $\delta_1(\bar{m}_1) = \delta_2(m_2^*)$. Such a \bar{m}_1 must exist by the intermediate value theorem. Thus there exists a jamming function on the set $W_2 = (y_d - 2\bar{x}, m_2^*) \cup (\bar{m}_1, y_d + 2\bar{x}) \subseteq (2\bar{x}, 1 - 2\bar{x})$.

This leaves $w \in (m_2^*, \bar{m}_1) = (y_d - 2\bar{x}, y_d + 2\bar{x}) \setminus W_2$. To provide the receiver with the incentive to choose y_d for pooling messages that indicate $w \in (m_2^*, \bar{m}_1)$, it must be that

$$y_d = \frac{\int_{w \in [2\bar{x}, 1 - 2\bar{x}] \setminus W} w dF_w}{\int_{w \in [2\bar{x}, 1 - 2\bar{x}] \setminus W} dF_w}. \quad (\text{A8})$$

Applying an appropriate fixed point theorem (e.g., the Fan-Browder Theorem; Border, 1985) implies existence of y_d satisfying (A8) and a jamming function on a domain/range $W_2 \subseteq (y_d - 2\bar{x}, m_2^*) \cup (m_1^*, y_d + 2\bar{x})$. ■

*Proof of the **Equilibrium Proposition.*** The Jamming Function Existence Lemma implies there is a default opinion $y_d \in [2\bar{x}, 1 - 2\bar{x}]$, a subset $W \subseteq (y_d - 2\bar{x}, y_d + 2\bar{x})$, and a jamming

function $m_J : W \rightarrow W$. In Case 1, $W = (y_d - 2\bar{x}, y_d + 2\bar{x})$. Define the strategies

$$\begin{aligned}
m_1(w, x_1) &= \left\{ \begin{array}{c} w \\ w \\ m_J(w) \\ \tilde{m} \end{array} \right\} \text{ if } w \in \left\{ \begin{array}{l} [0, y_d - 2|x_1|] \cup [y_d + 2\bar{x}, 1] \\ (y_d, y_d + 2\bar{x}) \cap W \\ (y_d - 2|x_1|, y_d) \cap W \\ (y_d - 2\bar{x}, y_d + 2\bar{x}) \setminus W \end{array} \right\}, \\
m_2(w, x_2) &= \left\{ \begin{array}{c} w \\ w \\ m_J(w) \\ \tilde{m} \end{array} \right\} \text{ if } w \in \left\{ \begin{array}{l} [0, y_d - 2|x_1|] \cup [y_d + 2\bar{x}, 1] \\ (y_d - 2|x_1|, y_d) \cap W \\ (y_d, y_d + 2x_2) \cap W \\ (y_d - 2\bar{x}, y_d + 2\bar{x}) \setminus W \end{array} \right\}, \text{ and} \\
y(m_1, m_2) &= \left\{ \begin{array}{c} m_1 \\ y_d \end{array} \right\} \text{ if } m_1 \left\{ \begin{array}{c} = \\ \neq \end{array} \right\} m_2.
\end{aligned}$$

Note that if $W = (y_d - 2\bar{x}, y_d + 2\bar{x})$, then these strategies reduce to the candidate strategies in (2). Therefore, it is sufficient to prove that these strategies, together with on-the-path beliefs from the Lemmas 1 and 5 and suitable off-the-path beliefs, satisfies equilibrium requirements.

Let off-the-path beliefs be as follows. If neither $m_1 = m_2$ nor $m_1 = m_J(m_2)$, let $h(w|m_1, m_2) = \left\{ \begin{array}{c} \frac{1}{4\bar{x}} \\ 0 \end{array} \right\}$ if $w \left\{ \begin{array}{l} \in [y_d - 2\bar{x}, y_d + 2\bar{x}] \\ \notin [y_d - 2\bar{x}, y_d + 2\bar{x}] \end{array} \right.$. Thus, following any off-the-path message pairs (m_1, m_2) , the receiver prefers to choose $y(m_1, m_2) = y_d$.¹³

It remains to make certain there are no profitable deviations from the proposed strategies.

First note that, if senders use the specified message strategies, the receiver's beliefs are

¹³Note that these beliefs match those used in Gilligan and Krehbiel (1989), rather than the beliefs presented by Krishna and Morgan (2001a) or in the one-dimensional model of Battaglini (2002), that Battaglini and others argue are implausible.

formed using Bayes' Rule wherever possible, according to Lemmas 1 and 5. Her optimal opinion strategy is to set $y(m_1, m_2)$ equal to her posterior expectation of w , which is either $y(m_1, m_2) = m_1$ if messages are confirmatory or, by construction, $y(m_1, m_2) = y_d$ if they are conflicting.

Now consider possible deviations by sender 1. Since the receiver only forms opinions y_w and y_d , senders are effectively choosing between these two options. Lemma 3 indicates that senders' optimal message choices are exactly those specified in (2). ■

The following results assume the existence of an equilibrium with a jamming function as given in the previous result. This jamming function has construction $m_J(w) = \eta_i(\delta_i(w))$, where $i \neq j$, and $i, j \in \{1, 2\}$.

Countervailing Jamming Messages. *Let $m_J : W \rightarrow W$ be a jamming function constructed via the method in the equilibrium proposition. Then $m_J(\cdot)$ is decreasing in w .*

Proof. Consider the case $w \in (y_d - 2\bar{x}, m_2^*)$ with the jamming function is $m_J(w) = \eta_1(\delta_2(w))$. The chain rule yields

$$\frac{d}{dw}m_J(w) = \frac{d}{d\delta}\eta_1(\delta_2(w)) \frac{d}{dw}\delta_2(w).$$

Since $\delta_1(\cdot)$ is strictly decreasing on the interval $(m_1^*, y_d + 2\bar{x})$, $\eta_1(\cdot)$ is strictly decreasing; hence, $\frac{d}{d\delta}\eta_1(\delta_2(w)) < 0$. Since $\delta_2(\cdot)$ is strictly increasing on $(y_d - 2\bar{x}, m_2^*)$, $\frac{d}{dw}\delta_2(w) > 0$. Thus $\frac{d}{dw}m_J(w) < 0$ for $w \in (y_d - 2\bar{x}, m_2^*)$. The proof for $w \in (y_d - 2\bar{x}, m_2^*)$ is similar. ■

Moderate Senders Send Extreme Messages. *For any F_1 , choose F_2 such that $1 - F_2(x) > F_1(-x), \forall x \in (0, \bar{x})$. Let y_d be the default opinion, $m_J(\cdot) : W \rightarrow W$ the jamming function, and $m_i(\cdot)$, $i = 1, 2$, the message strategies. $\forall (w, x_1, x_2) \in [0, 1] \times X_1 \times X_2$, if $m_1(w, x_1) \neq m_2(w, x_2)$, then:*

$$(i) |y_d - m_1(w, x_1)| > |y_d - m_2(w, x_2)|, \text{ and}$$

(ii) $h(m_1(w, x_1) | m_1(w, x_1), m_2(w, x_2)) < \frac{1}{2}$.

Proof. Applying the above method yields $\delta_1(m_1) = \frac{1}{2}(m_1 - y_d) \left(1 - F_2\left(\frac{1}{2}(m_1 - y_d)\right)\right)$ and $\delta_2(m_2) = \frac{1}{2}(y_d - m_2) F_1\left(-\frac{1}{2}(y_d - m_2)\right)$. Because $1 - F_2(x) > F_1(-x), \forall x \in (0, \bar{x})$, the maximum of $\delta_1(\cdot)$ exceeds that of $\delta_2(\cdot)$. Thus this claim must fall under Case 2 of the Jamming Function Existence Lemma, in which there is a jamming function with the domain range $W \subseteq (y_d - 2\bar{x}, m_2^*) \cup (m_1^*, y_d + 2\bar{x})$. The jamming function $m_J(\cdot)$ is defined on the decreasing part of $\delta_1(\cdot)$ and the increasing part of $\delta_2(\cdot)$. Furthermore, $\forall x \in [0, \bar{x}], 1 - F_2(x) > F_1(-x)$ implies $x(1 - F_2(x)) > xF_1(-x)$.

Let $w \in W \cap (y_d - 2\bar{x}, y_d)$, and let $x^\circ = \frac{1}{2}(y_d - w)$. Lemma 3 indicates that $m_2 = w, \forall x_2$. If $x_1 \in [-x^\circ, 0]$, Lemma 4 implies $m_1 = w = m_2$. If instead $x_1 \in [-\bar{x}, -x^\circ)$, then $m_1 = m_J(w)$. By hypothesis,

$$x^\circ(1 - F_2(x^\circ)) > x^\circ F_1(-x^\circ). \quad (\text{A9})$$

It must be that $\frac{1}{2}(m_1 - y_d) > \frac{1}{2}(y_d - m_2)$, since w is in the decreasing part of $\delta_1(\cdot)$. Hence, $|y_d - m_1| > |y_d - m_2|$. Moreover, this requires $1 - F_2\left(\frac{1}{2}(m_1 - y_d)\right) < F_1\left(-\frac{1}{2}(y_d - m_2)\right)$, so Lemma 5 implies $h(m_1 | m_1, m_2) < \frac{1}{2}$. ■

References

- Austen-Smith, David. 1993. "Interested Experts and Policy Advice: Multiple Referrals under Open Rule." *Games and Economic Behavior* 5(1):3–43.
- Austen-Smith, David. 1995. "Campaign Contributions and Access." *American Political Science Review* 89(3):566–582.
- Austen-Smith, David and John R. Wright. 1992. "Competitive Lobbying for a Legislator's Vote." *Social Choice and Welfare* 9(3):229–257.
- Bailey, Michael, Brian Kamoie and Forrest Maltzman. 2005. "Signals from the Tenth Justice: The Role of the Solicitor General at the Merits Stage." *American Journal of Political Science* 49(1):72–85.
- Battaglini, Marco. 2002. "Multiple Referrals and Multidimensional Cheap Talk." *Econometrica* 70(4):1379–1401.
- Battaglini, Marco. 2004. "Policy Advice with Imperfectly Informed Experts." *Advances in Theoretical Economics* 4(1):1–32.
- Border, Kim C. 1985. *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge: Cambridge University Press.
- Carrubba, Clifford J. 2001. "The Electoral Connection in European Union Politics." *The Journal of Politics* 63(1):141–158.
- Carrubba, Clifford J. 2009. "A Model of the Endogenous Development of Judicial Institutions in Federal and International Systems." *The Journal of Politics* 71(1):55–69.
- Crawford, Vincent and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(6):1431–1451.
- Esterling, Kevin M. 2004. *The Political Economy of Expertise*. Ann Arbor: University of Michigan Press.

- Gilligan, Thomas W. and Keith Krehbiel. 1987. "Collective Decisionmaking and Standing Committees: An Informational Rationale for Restrictive Amendment Procedures." *Journal of Law, Economics, and Organization* 3(2):287–335.
- Gilligan, Thomas W. and Keith Krehbiel. 1989. "Asymmetric Information and Legislative Rules with a Heterogeneous." *American Journal of Political Science* 33(2):459–490.
- Krishna, Vijay and John Morgan. 2001a. "Asymmetric Information and Legislative Rules: Some Amendments." *American Political Science Review* 95(2):435–452.
- Krishna, Vijay and John Morgan. 2001b. "A Model of Expertise." *Quarterly Journal of Economics* 116(2):747–775.
- Luntz Research Companies. 2001. "The Environment: A Cleaner, Safer, Healthier America." http://motherjones.com/files/LuntzResearch_environment.pdf. (February 4, 2010).
- Lupia, Arthur and Matthew D. McCubbins. 1998. *The Democratic Dilemma*. Cambridge: Cambridge University Press.
- Morris, Stephen. 2001. "Political Correctness." *Journal of Political Economy* 109(2):231–265.
- Page, Benjamin I. and Robert Y. Shapiro. 1992. *The Rational Public*. Chicago: University of Chicago Press.
- Sobel, Joel. 1985. "A Theory of Credibility." *Review of Economic Studies* 52(4):557–573.
- Staton, Jeffrey K. 2006. "Constitutional Review and the Selective Promotion of Case Results." *American Journal of Political Science* 50(1):98–112.
- Vanberg, Georg. 2001. "Legislative-Judicial Relations: A Game-Theoretic Approach to Constitutional Review." *American Journal of Political Science* 45(2):346–361.