CONTENT AND CONTEXT: THREE ESSAYS ON INFORMATION IN POLITICS

by

Ian Palmer Cook

Submitted to the Graduate Faculty of

the Kenneth P. Dietrich School of Arts and Sciences

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

KENNETH P. DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Ian Palmer Cook

It was defended on

March 31, 2016

and approved by

Jonathan Woon, Political Science

Kristin Kanthak, Political Science

George Krause, Political Science

Kenneth Shotts, Graduate School of Business, Stanford University

Dissertation Director: Jonathan Woon, Political Science

ABSTRACT

CONTENT AND CONTEXT:
THREE ESSAYS ON INFORMATION IN POLITICS

Ian Palmer Cook, PhD

University of Pittsburgh, 2016

This dissertation explores the implications of information asymmetries in three specific political environments: primary campaign speeches; negotiating behavior; and testimony delivered in a congressional hearing. First, dog whistling can dramatically affect the outcome of elections, despite observers never being sure it actually occurred. I build a model that addresses how a whistle operates, and explore implications on candidate competition. I find that whistling lets candidates distinguish themselves from competitors in the minds of voters. Second, political negotiation frequently looks like two sides staring each other down, where neither side wishes to concede, claiming that doing so would incur the wrath of voters. Little theory or evidence exists to explain how voters allocate blame for different outcomes. We conduct a laboratory experiment to investigate how anticipation of blame drives negotiating behavior, and how observers allocate blame. We find that the presence of an observer has little effect on standoff outcomes but appears to shorten the duration of standoffs. Third, while congressional hearings give legislators a national stage on which to score political points by publicly chastising high-level bureaucrats, and gives lobbyists a forum to demonstrate their access and importance to policymakers, less clear is how well hearings serve the purposes of oversight. I address this question through automated text analysis of hearings in the $105^{th} - 112^{th}$ Congresses. I show that the oversight function of hearings is only effective when it is least likely to be used: when the congressional committee and the bureaucrat agree on policy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

For Marion:

Papa done.

PREFACE

Number Six: Where am I?
Number Two: In the Village.
Number Six: What do you want?
Number Two: Information.
Number Six: Whose side are you on?
Number Two: That would be telling. We want information... information... *information.*
Number Six: You won't get it.
Number Two: By hook or by crook, we will.

The Prisoner

1.0 INTRODUCTION

Information is a protean concept in political science research. Depending on the purpose, it can be made more or less explicit, contrasting it sharply with traditional entities such as floor votes, regulatory actions, survey answers, campaign contributions, budgetary allocations, committee assignments, or Supreme Court rulings. Information is a broad term, more akin to concepts that require the "political" modifier to first winnow the options: ideology, will, violence. Indeed, one portion of the rational choice tradition does not require any more definition than to simply state some information exists, one or both sides may know it (and perhaps know the other knows it...), then concerns itself with the impact of differing levels of information – anywhere on the spectrum from complete to incomplete – on the strategic choices actors make in a vast array of situations (Calvert 2013; Gates and Humes 1997). Information, in that work, is an abstraction, and often operates on an epistemological level where actors have beliefs about their own and others' knowledge independent of whatever they are supposed to have knowledge *of*. However, while each essay in this dissertation draws directly from this rational choice perspective, explicitly using the insights from formal models to gain insights about politics, I choose to be far less abstract. I make the subject of concern very clear in each case: campaign messages and coded language; testimony delivered in a congressional hearing; and negotiating behavior in a costly waiting game. As a result, I trade generalization of information for applicability to specific political environments.

My work continues a long history of work positing information as a tangible good to explore. The part of the vast literature most relevant to my work includes studies of voter knowledge, and the information asymmetries that exist between principal and agent. Voters have long been assessed, queried, and tested about how well they can identify political actors, events, or materials (the content of propositions, bills, trade agreements, etc.) (L. M. Bartels 1986; Campbell 1964; Campbell et al. 1960; Lewis-Beck, Jacoby, et al. 2008; Lupia, Mathew D McCubbins, and S. Popkin 2000; Lupia and Mathew Daniel McCubbins 1998). Perhaps unsurprisingly, they frequently fail to measure up to an ideal, be it one of rationality or normative desirability (L. Bartels 1996; Carmines et al. 1980; Erikson and Wright 1993; Kuklinski and Quirk 1998; Lau 1997). Similarly, information on policy choices in government agencies is at the heart of literature on delegation from the legislature to the bureaucracy. Foundational work on delegation take a very general view of information, considering broadly some principal who must decide whether to take an action herself or delegate the action to an agent who is better informed about the relevant topic (Bendor and Meirowitz 2004; Holmstrom 1984; Huber and Shipan 2002; Vickers 1985). Here, too, there is extensive work that considers more tactile forms of information. Scholars have examined both the sources (Carpenter, Esterling, and Lazer 2004; Esterling 2004) and subjects (Carpenter 2010; Olson 1999; K. T. Poole and Rosenthal 1994; Splawn 1939) of information used in decision making, but rarely address the specific form it takes. Part of the contribution of this dissertation is examining the particular content of messages between actors and showing how it contributes to achieving political ends.

My first essay, "Dog Whistling" addresses the use of coded language in campaign speeches. I ask whether messages that can be heard in different ways by different members of the same audience allow candidates to capture vote share by signaling two possible types. The title of the essay takes its name from the common term from this practice, often described as a tool for delivering information to a subset of voters that would upset the rest, if only that group had heard the whistle. The presence of these coded messages privileges one group over another, possibly without the difference being known. This undermines the ability of voters to develop accurate information about politicians, and thus represents a threat to basic representative democracy. The traditional media depiction of dog whistling paints the user as a biased villain (racist, sexist, homophobic), appealing to the worse natures of intolerant voters Haney-López 2014. I develop a formal model to explore when candidates might engage in coded speech, generalizing away from the particular moral content, focusing instead on shared group affiliations between the politician and the voters.

The model addresses the motives of two candidates campaigning during a primary election, confronting audiences who may or may not share some group affiliation with the candidate. There are two notable aspects of the model that serve as contributions independent of the model's results. First, I allow the two groups in the audience to differ in their ability to hear the coded message, making it a function of their group identity. This introduces behavioral differences in the model that stem from political psychology, a still uncommon feature of formal modeling in political science. Second, I develop a probabilistic model of message reception that makes the individual differences in hearing a dog whistle a strategic tool for candidates to exploit. Dog whistles allow politicians to split voters such that a single audience may hold disparate, but well-founded (given the message they hear), beliefs. The model, then, represents a useful integration of two fields in political science that should operate together more often than they do.

The results from my model undermine the traditional view of dog whistling. Coded speech is not solely the tactic of a crypto-aggressor towards marginal groups attempting to reveal a true, nefarious agenda. Put succinctly, dog whistles introduce an uncertainty in the mind of a voter about the true identify of the candidates. This uncertainty can sway some voters – who need not share the candidate's group affiliation – to cast their lot for the whistling candidate because even a slim possibility of electing the more-preferred kind of candidate is better than definitely electing the less preferred. Moreover, the strategic benefits of using coded speech often inspire its use by *both* candidates in an attempt to mitigate competitor advantages. That is, given the opportunity, candidates may whistle regardless of their true group identity, making it even harder for voters – and the media – to correctly identify the candidate's true intentions.

My second essay, "Time and Punishment: Blame and Concession in Political Standoffs," is joint work with Jonathan Woon, and addresses the interaction between representative behavior in negotiations and voter assessment of performance. Political negotiation frequently looks like two sides staring each other down, waiting for the other to blink. In these showdowns, neither side wishes to concede, claiming that doing so would incur the wrath of voters. Whether this consideration of potential punishment influences behavior during stalemates is not well understood, and little theory or evidence exists to explain how voters allocate blame for different outcomes. Taken together, we investigate two interrelated questions: how does anticipation of blame drive behavior, and how do observers with a stake in the outcome allocate blame? Answering the first question sheds light on how politicians internalize, and act on, information about their constituents' preferences. Answering the second question provides insight into how voters with the same information differ on how to apply it to condition the behavior of their representatives.

We conduct a laboratory experiment that implements a dynamic war of attrition model for negotiation situations. In the experiment, concession time is the key choice variable for the negotiators. Our design compares versions of the game with and without an observer (whose payoffs depend on the outcome and who can punish the players). Without an observer the game is a baseline measure of individual risk assessments about holding out for the higher payoff. Adding the observer captures the salient parts of negotiations that appear to be a growing feature of U.S. politics: standoffs between the two political parties over legislation, with the continued funding of the government hanging in the balance. The two sides appear to have incentives to out-wait the other, increasing the possibility of total collapse of negotiations, but have to consider the real chance of costly punishment if that happens. Our expectations for behavior during this game are based on a simple model that extends the literature on blame by making the mechanism explicit and the level of blame endogenous. The use of a laboratory experiment lets us control the conditions to isolate the particular incentives we wish to test, such as the level of affiliation between one negotiator and the observer (in terms of payoffs from the game).

We find that the presence of an observer has little effect on standoff outcomes but appears to shorten the duration of standoffs. We also find that observers tend to punish the winning player, which is qualitatively consistent with a rational or instrumental form of punishment, but the amount of punishment is less than optimal. Negotiators, like politicians, may not adequately anticipate when, or how much, they will be blamed, preventing them from altering their choices and achieving better outcomes.

The third essay, "Information Transmission in Congressional Hearings" returns to theoretical examination of single messages, but empirically implements the theory in assessing a large set of interactions during U.S. Congressional hearings. Using the Crawford and Sobel (1982) model as inspiration, I explore the informativeness of a bureaucratic agent's message as a function of the committee chair receiving it.

Congress' ability to extract information from the bureaucracy is central to performing its oversight function. Hearings give legislators a national stage on which to score political points by publicly chastising high-level bureaucrats, and gives lobbyists a forum to demonstrate their access and importance to policymakers. But how well they serve their nominal function, enabling oversight by revealing useful policy information, remains in doubt. I address this question by directly examining the content of hearings from the $105^{th} - 112^{th}$ Congresses. Through automated text analysis methods, I measure the *language specificity* of the actors in the hearing in order to quantify the transmission of information between legislators and agency witnesses. This process marks two contributions. First, this method dramatically enhances the size of the corpus of documents on which to base assessments of hearings. Indeed, this process can be used to quantify crucial information in any large set of texts pertinent to important research questions. Second, as of this writing I believe this is the first direct measurement of the central quantity in the Crawford and Sobel (1982) model, which has informed a great portion of work on signaling in political science since its publication. The value of a theory does not lie solely in testing it against empirical reality, but basic empirical work on a rigorous theoretical model can dramatically improve interpretation of the results. This method, then, provides a bridge between theory and empirics for a wide range of important questions.

The empirical work in the third essay is a an empirical model of hearings; the results indicate confirm my (and the theorists') expectations, showing that increasing ideological distance between congressional committee chairs and executive branch witnesses decreases the amount of information revealed. The oversight function of hearings is thus only effective when it is least likely to be used: when the congressional committee and the bureaucrat (as representative of the president) agree on policy. In addition to providing insight into a longstanding question about whether hearings serve as effective oversight, this paper demonstrates the value of a new, but widely applicable technique for quantifying the informativeness of communication, a central concept in the study of political information and knowledge.

The sharing of information impacts representation, accountability, oversight, and more. Voters that cannot even identify when a candidate is delivering certain messages cannot adequately hold them accountable during an election. On the other hand, a voter rendering judgment of a result she is only partially qualified to evaluate may force delegation-style representation that prevents compromise and policy improvement. Political control is called into question if Congress cannot evaluate the information it receives from the bureaucracy. The essays in this dissertation attempt to expand our understanding of the competing issues that arise in how information is shared among actors.

2.0 DOG WHISTLING

> For of the three elements in speech-making — speaker, subject, and person addressed — it is the last one, the hearer, that determines the speech's end and object.
>
> *Rhetoric, Book I*
> ARISTOTLE

Senator Claire McCaskill (D-MO) credits the use of dog whistling with her successful reelection in 2012. During Missouri's *Republican* primary for the Senate her team manipulated voter behavior by running advertising that, she suggests, convinced a portion of voters that Todd Akin was their kind of conservative:

> "Using the guidance of my campaign staff and consultants, we came up with the idea for a "dog whistle" ad, a message that was pitched in such a way that it would be heard only by a certain group of people. [...] And we needed to run the hell out of that ad."[1]

McCaskill chose language to run in a *single* television advertisement that *simultaneously* motivated extreme Republicans to support Akin, while cautioning moderate voters against him. Of course, the McCaskill ad was observable by any viewer; the important facet is the intent to communicate a particular message to one audience subset that bypasses the complementary set. This changed the competitive landscape, she argues, by shifting voter support away from more moderate, and thus harder to beat, Republican challengers. McCaskill went on to decisively win her re-election bid. This raises questions about the mechanism that allows multiple voters to have different interpretations of the same message, and whether that mechanism can actually produce the shift in voting behavior McCaskill attributes to it. In this paper, I develop a signaling game played among

[1] Senator Claire McCaskill, "How I Helped Todd Akin Win — So I Could Beat Him Later." *Politico Magazine*, 11 August, 2015

candidates and voters that includes just such a mechanism, and demonstrate that yes, this purposeful manipulation of voter information can alter a candidate's support.

Dog whistling is frequently an epithet, used by media commentators to suggest politicians are attempting to hide their true, and, according to the commentator, distasteful views. If all the claims about whistles and the whistlers were true, the list of offenders would be long, notable, and multinational.[2] Prominent members of both U.S. political parties – Ronald Regan,[3] Newt Gingrich,[4] George W. Bush (Grossman and Matthews 2009, p. 39), both Bill[5] and Hillary Clinton (Haney-López 2014, p. 111) – willfully exacerbated economic inequality through the use of racially-based dog whistling. Current President Barack Obama revealed his antisemitism during his push to gather support for an agreement with Iran.[6] Member of the British Parliament and Labour Party leader Ed Miliband sent a dog whistle to opposition loyalists that his normally stout support of labor-union relations had softened.[7] Current U.K Prime Minister David Cameron revealed his xenophobia by whistling his intention to limit immigration, a particularly divisive topic in British politics during the recent economic recession.[8] And in one of the more chilling cases, a major telecommunications company in Egypt signaled tactical instructions for a terrorist attack to the Muslim Brotherhood.[9] Of course, whether their authors intended these messages to carry the secondary connotations is, and will remain, unknown. To address the phenomenon, I set aside any normative debate about the choice to use a dog whistle, and focus instead on identify a general mechanism for the whistle in any context, and how effective it might be.

[2]Indeed, the phrase "dog whistling" most likely originated in Australia, and was prominent in the United Kingdom before entering into common usage in the United States

[3]Joseph Crespino, "Did David Brooks Tell the Full Story About Reagan's Neshoba County Fair Visit?," *History News Network* 11 December, 2007.

[4]Gwen Ifill, "Dog Whistle Politics: You Talking to Me?," *PBS Newshour*, 20 May, 2011

[5]Nancy LeTourneau, "Democrats and Dog Whistles," Washington Monthly's Political Animal, 9 April, 2015.

[6]Lowi Lorenthal Marcus, "Obama 'Dog Whistles' Netanyahu (and Insults Congress, too)," *The Jewish Press*, 20 January, 2015.

[7]Patrick Wintour, "Ed Miliband's reform plans rejected as 'dog whistle' politics by union leader," The Guardian, 9 July, 2013.

[8]Daniel Knowles, "David Cameron's dog-whistle politics on immigration," The Telegraph, 10 October, 2011.

[9]Associated Press, "Egypt: Puppet Ad Draws Terror Accusations," The New York Times, 1 January, 2014.

Investigating the use and impact of dog whistling – or, "coded speech" more generally – is important because voter behavior is sensitive to messaging from politicians and their communications staff. Presidential pleas can engender public support for initiatives (Vavreck 2009), while small changes to word choice can invert a person's policy views (Hurwitz and Peffley 2005; Nelson 1999). Certainly, while any word may carry multiple meanings, some are imbued with distinct connotations that are more politically salient than their primary definitions. Some exert only indirect influence, which, I argue, makes them more powerful in certain circumstances than the direct phrasing in framing studies (Kahneman and Tversky 2000; Monroe, Colaresi, and Quinn 2009; Sniderman and Theriault 2004). Their importance, in fact, arises precisely *because* of this obliqueness, since it drives a wedge between voters' beliefs and a candidate's true intentions for governing (Goodin and Saward 2005).

This paper makes several contributions, both substantive and methodological. First, I consider dog whistles as a general phenomenon, independent of any specific political issue. Other studies have considered only a single topic at a time, such as religion (Albertson 2014) and race (Haney-López 2014). Second, I make message content a function of the voter's characteristics, explicitly connecting our understanding of political identity with formal models of information transmission. Finally, I show that the common wisdom about dog whistling is incorrect. The results of the analysis demonstrate that the value of whistling stems from inducing uncertainty in *all* potential voters, rather than, as the media suggests, acting as a klaxon for a select few while going unheard by the rest.

The rest of the paper proceeds as follows: First I define dog whistling and distinguish it from other forms of frequently studied forms of political communication. Next I discuss how the whistle impacts voter information gathering, and what this implies for voter behavior. I then describe and analyze one- and two-dimensional versions of the dog whistling model. The paper ends with a discussion of what I contribute to understanding the dog whistling phenomenon.

2.1 "YOU KNOW HOW TO WHISTLE, DON'T YOU?"

Dog whistling can be concisely defined as a single piece of communication that has at least two distinct meanings, delivered by a speaker with the intent that only a subset of the audience will hear the second meaning (Albertson 2014; Fear 2007). While I structure the model as a single speech, television ads, posters, even mailed postcards are all included within this definition.

A simple example helps to fix the concept. Ronald Reagan gave his 1979 post-nomination speech in the same Mississippi county where three civil rights activists were murdered 16 years earlier. During the speech, Reagan used the phrase "state's rights," the specific meaning of which is still debated (Smith 2013). Whether a standard trope among conservative politicians, or coded language to indicate Reagan would not interfere in race-based policy making, the effect of the phrase was important and observable. Soon after the speech, Reagan received — then publicly disowned — an endorsement from the far-right, white-supremacist, Ku Klux Klan. In the end, Reagan got to have it both ways: he received the group's support and plausibly denied that he courted them directly (Cannon 2005).

I use an electoral primary as the setting for my model of political speech in order to capture a richer strategic setting. Here, two candidates vie for votes by sending two-dimensional messages about their types. A "whistling function" divides what voters hear such that voters hear only one message or the other. Casting votes in the presence of a dog whistling strategy reflects voters knowing, perhaps via the news, that while a candidate may have sent a dog whistle, this is equivalent to seeing a candidate make the same statement with complete sincerity. Voters choose a candidate based on their beliefs about the messages, which are, according to the model, manipulable through the use or absence of a dog whistle. Several key features of the model require explanation.

First, what allows a speaker to whistle at all? Some portion of an audience must share an interpretation of images, words, or phrases of which a speaker is aware and able to use. Where code-breaking requires a key to find hidden messages inside plain text, hearing a (political) dog whistle uses shared understanding to uncover the secondary meaning.

Shared understanding derives from shared cultural identity which produces (and, to some extent, is defined by) norms of behavior and language (Geertz 1973; Hechter and Opp 2001). Social norms, in other words, generate correlated interpretations among group members. I break the candidates and voters in two, identifying them as either "in-group" or "out-group," leaving the exact group identity abstract.[10] As a brief example, consider that Bush's 2003 State of the Union address featured the phrase "wonder-working power," an encomium to American voluntarism (Bush 2003). The term references - but does not cite - a phrase found in the Christian hymn "There is Power in the Blood" (Howard 1994). Those in the audience who sung the hymn in church would know the phrase and its origin, while others would fail to recognize the allusion. This kind of common inferring of information is a sort of focal point, where each person has an expectation of what *other group members* expect the rest of the group to believe (Schelling 1960). Differences between groups — or "in-group" versus "out-group" as in the following model — imply different focal points. For modeling purposes, I assign each group a single, and different, probability of hearing a speaker's whistle. The probability value serves as a loose formalization of focal points, indicating that members of the same group share cultural knowledge allowing them to "coordinate" on a single interpretation.[11] Differences in the two probabilities reflect the difference in cultural norms between the groups.

I include multiple receivers in my models, as I am motivated by the primary election setting. To study the effects of a coded message intended for a subset of the audience, the audience must contain more than one person, each with a distinct likelihood of hearing one or the other meanings. Moreover, whistling's effect need not be outright electoral victory for it to be politically important. My interest is to closely tie the action of coded speech to the result of garnering a voter's support following a speech. Beyond the one contest, a candidate may seek to increase her vote share even if she has little chance of winning

[10] Albertson (2014) uses this in/out distinction to inform an empirical test of the use of religious code words.

[11] For more developed treatments of focal points in games, see R Sugden (1995) and Mehta, Starmer, and Robert Sugden (1994). I am less concerned than these authors with specific mechanisms for focal points in a game, though I do address a difference between "cheap talk and convention" raised in Farrell and Rabin (1996). Convention arising from group identity can be triggered in speeches, allowing for coordination on a choice. While I do not use a cheap-talk structure, no aspect of the probability-based focal points relies on costs.

the primary. A good showing in an election, even without winning, may increase the candidate's stance within a party. During the 2008 U.S. presidential primary, a number of Republican candidates, each with very small chances of garnering the nomination of the party, appeared to use dog whistles aggressively (Desmarais 2012).

When I include a second dimension in the subsequent model, candidates and voters are endowed with a political ideology. In this case, however, the ideology only ranges from moderate to extreme. This restriction both reflects a substantive condition and simplifies the analysis. Targets of whistling are correlated with ideology; conservatives do not whistle to labor unions, and liberals do not whistle to religious fundamentalists. The coded appeal reaches out to possibly skeptical portions of the politician's natural constituency (Albertson 2014), attempting to convey solidarity without giving away too much to the larger group. Restricting the ideological spectrum removes considerations of conservative speakers attempting to sway liberal sub-populations, and vice versa. The practical effect is to place candidates in competition against opponents on the same side of the political divide.

Primary elections are a natural setting for modeling contests between candidates of similar ideology. Politicians of similar bent will search for ways to differentiate themselves from their competitors. Using coded speech could achieve this differentiation by allowing some voters to believe the candidate is of a certain type, revealed (possibly) via dog whistle. Candidates and voters are members of two groups, each with binary indicators: moderate or extremist, in-group or out-group. To keep analysis as general as possible, I make no assumptions about the distribution of these types *for the voters*. There are different underlying probabilities that candidates will be of one ideology or the other, and either in- or out-group. The dichotomous structure emphasizes in-group/out-group distinctions (Winter 2006): one is either a populist or not, is Methodist or not, is in the Tea Party or not.

Dog whistling is not always cheap talk; both whistles and outright misrepresentations can be costly choices. My first, one-dimensional model imposes no costs, while my two-dimensional model incorporates penalties for both dog whistling and outright lying. These costs arise at least two possible ways: reputational losses from being accused of not speaking truthfully, or from a personal aversion to misrepresentation of any form. In the first case, consider that the media rarely lets a *potential* dog whistle go unremarked, despite their inability to make dispositive claims about their existence. Alerting voters to their existence, however, incurs a reputational cost to the candidate.[12] This media focus also appears in claims of "pandering", where a politician is chastised for making perfunctory remarks about particular group, or making insincere – in the commentator's estimation – appearances at events held by a group. Think of Republican candidates giving talks for the National Association for the Advancement of Colored People (NAACP), or Democrats addressing police violence towards minority groups. Like dog whistling, we cannot know the sincerity of the candidate, leaving the act open to audience interpretation. One important distinction, however, is that the media's usage is *not* the pandering addressed by Canes-Wrone, Herron, and Shotts (2001). There, pandering is specifically defined as politicians acting in what the public *believes* to be their best interest, but may not actually be. In whistling, and indeed in the less-defined but media-preferred version of pandering, voters cannot ever know the politician's true information. Thus, dog whistles are distinct from both forms of pandering. In terms of the second, psychic, form of cost, note that people are often averse to misrepresenting their true selves. Substantive motivation for lying aversion deviates from the common view of candidates being solely focused on achieving office (Mayhew 1974), but is consistent with recent literature that suggests individuals may have an innate aversion to lying (Gneezy 2005; Hurkens and Kartik 2008; Sánchez-Pagés and Vorsatz 2007). Indeed, such work has been put to use in formal models of political communication and as motivation for laboratory experiments (Kartik 2009; Lightle 2013; López-Pérez and Spiegelman 2012; Minozzi and Woon 2013).

Dog whistling is distinct from framing, though the result of framing and the effects of using whistles are comparable. Differences in question wording and presentation — framing the issue — produce varying, even contrary, views on a single topic from the same person (Chong and James N. Druckman 2007). Political candidates are particularly

[12] Effort required to identify and use an appropriate whistle for a particular audience or topic is another interpretation of costly messages. Political speeches are products of numerous hours of work by multiple parties; a whistle arising entirely unintentionally is unlikely. Of course, plain political speech is similarly costly, though it is often treated — usefully — as costless.

advantaged in using framing effects because of the credibility sometimes conveyed just by dint of being a candidate (James N Druckman 2001a). And framing can guide question responses through choosing how to present information in a light that favors a particular outcome (Cobb and Kuklinski 1997; Kuklinski and Quirk 1998; O'Keefe and Atwood 1981; S. L. Popkin 1991). Campaign advertising makes extensive use of this advantage. For instance, using an African-American convict's face in 1988 presidential campaign ads could, as Hurwitz and Peffley (2005) argued, have been an attempt to communicate not only a policy message, but also to incite racial tensions. Indeed, race-based intentions in coded speech have received considerable attention. Appeals that make racial divisions salient, first through implicit and then explicit language, are known to be critical to choosing a candidate (T. B. Edsall, M. D. Edsall, and D. 1991; Mendelberg 2001; Valentino, Hutchings, and I. K. White 2002). Framing may also be used in competitive fashion, with counter-frames created to shift perspectives from one frame to another (Chong and James N Druckman 2013). Whistles, by comparison, lack the pre-emptive nature of frames, since the whistle has divergent effects on the same population. While a frame may have differing resonance, it cannot separate audiences based on its content, producing differing understandings of the same message. Moreover, the whistle carries *both* the manipulation and the content, whereas a frame alters interpretation of a subsequent message.

2.2 VOTER INFORMATION AND WHISTLING

Voter preferences are heavily dependent on the information they obtain (James N Druckman 2001b), information that can easily be manipulated by careful construction of the transmitting message. To quote Edelman (1985, p. 11):

> In short, it is not "reality" in any testable or observable sense that matters in shaping political consciousness and behavior, but rather the beliefs that language helps evoke about the causes of discontents and satisfactions, about policies that will bring about a future closer to the heart's desire, and about other unobservables.

Whether that language is delivered directly by a politician or filtered through the media, it is controlled by both the subject and the source of voter interest (Zaller 1992). The way a candidate speaks thus shapes what people learn about current economic conditions, candidate positions, policy debates, and more (Gilens, Vavreck, and M. Cohen 2007; S. L. Popkin 1991). Carefully chosen coded language alters beliefs citizens have about distributional consequences of policy Nelson and Kinder (1996), and can even reverse prior support for various policies based on the how the race of beneficiaries is presented (Hurwitz and Peffley 2005). Even the course of a campaign can be influenced by the process of voters seeking, finding, and evaluating new information, resulting in a "discover, scrutinize, decline" cycle of popular support for candidates (Sides and Vavreck 2014). I build on these insights, but depart by relaxing the assumption that everyone in the audience heard exactly the same content in that language. Using a dog whistle violates this assumption, and makes it possible for two audience members to hear the same speech, yet depart with two distinct beliefs about the *substance* of what was said. Politicians can, and if the accusations are correct they frequently do, confound voters attempting to develop knowledge and preferences, calling into question the validity of elections (Goodin and Saward 2005).

Politicians chase partisan ends through rhetorical means, including careful consideration of message content. Riker famously defined heresthetics as concerning "the strategy value of sentences," noting that it "involves the use of language to accomplish some purpose" (Riker 1986, p. x). In assessing their persuasive ability, presidents are hailed for their rhetorical brilliance, or lack thereof (J. E. Cohen 1995). Presidents have a powerful tool in "going public" to deliver the specifics of policy positions, attempting to use their pulpit to shift public opinion and pressure Congress into action, though notably only when public and executive views already align (Canes-Wrone 2010). Related to dog whistling, political pandering — where the speaker parrots back what they believe the public wants to hear — is a useful tool to improve reelection chances(Canes-Wrone, Herron, and Shotts 2001; Fox and Shotts 2009). My contribution to this work is to make the selection of rhetorical content strategic, and conditional on voters' social characteristics.

Work on candidate ambiguity is a natural comparison to this study, since I embrace the notion that politicians have rational reasons to shade their true ideological or policy preferences (Alesina and Cukierman 1990; Glazer 1990; Page 1976; Shepsle 1972). This prior work shows that candidate evasiveness, and voters being less than perfectly informed, does not necessarily result in hollow elections (Chappell 1994). That said, facing less risky choices over candidates is generally preferred by voters, all else equal, while the candidates always have an incentive to make fuzzy statements (Page 1976). In game-theoretic terms, vagueness is often represented by allowing messages to represent more than one distinct state of the world. Blume and Board (2009) directly confront imprecise language by examining how a sender's single message can result in two receivers having two different interpretations. However, as the authors note, this "measure of vagueness operates in the message space as opposed to the type space" (p.4). My formulation operates in the type space: the speaker is fully informed about her own type, and sends a single, distinct message chosen specifically to trigger a (known) chance of a carefully chosen alternative interpretation that is a function of receiver types. Further, a coded message is not analogous to one privately sent (an option in Farrell, Gibbons, et al. (1989)), since the purpose of the code is to (potentially) prevent one group from identifying its true content.

Finally, mendacity has a long political tradition, with both clearly deleterious effects and even some potential upsides (Jay 2010). Audiences mis-apprehending a message's content, whether the plain or coded, cannot be the sole fault of the speaker. In a lie, however, a message has been chosen in such a way that the sole content is actually the opposite of the underlying truth. Receiving the lie occurs without regard to the audience's traits. In my model, audience traits explicitly contribute to the message they receive, making them a more active part of the process. Formal exploration of lies in politics note their potential value. Recent work (Callander 2007; Sobel 2007) addresses the use of lies by political candidates during campaigns, finding that candidates with a higher willingness to lie may actually be preferred during an election. Dog whistles differ from these lies, however, in that a lie does not contain the information being both obscured and transmitted.

2.3 MODELS AND RESULTS

The results of the forgoing discussion is distilled into the following two games between political candidates and an audience of voters.[13] All the actors in the first model have types in a single, group-membership dimension, and play a cheap-talk signaling game. The second model adds a second, ideological dimension and imposes costs for some messaging choices in the game. Keeping the first model simple and costless emphasizes the effect dog whistling has on voters. The second model develops a richer messaging environment for candidate competition, allowing communication in two dimensions, and making some of those messages costly.

2.3.1 One Dimensional Model

The first model of dog whistling restricts the actors to just two types in one dimension, in order to focus on the function of the dog whistle. The game is played between two senders – political candidates competing for votes – and a population of receivers – voters in an audience. Only one candidate is a strategic actor, though the second candidate's type is relevant for the audience's decisions. I denote the two candidates, $C_i \in \{1, 2\}$. Candidate i's type is $\gamma_i \in \{O, I\}$. C_1's type is private information, while $C_2's$ type is common knowledge. In this structure, C_1 can be thought of as a challenger attempting to sway voters away from the known incumbent, C_2. The types identify *group membership*, with O denoting the candidate is "out-group," and I denoting "in-group". Candidates are out-group with $Pr(\gamma = O) = g$ for $g \in (0, 1)$,[14] and in-group with the complementary probability.

Similarly, the two types of voters are identified by their group membership, denoted $\beta \in \{I, O\}$, matching the type space for candidates. Rather than identify probabilities of each voter being of a certain type, denote the size of each group by v_β, where $\forall \beta$, $v_\beta \geq 0$ and $\sum v_\beta = 1$.

[13]For this paper, candidates are identified as female, while voters are identified as male.

[14]I exclude the limits of this range to avoid edge cases that do not affect the analysis. If group identity is known perfectly, $g = 0$ or $g = 1$, voter choices are effectively independent of C_1's message.

The play of the game is as follows. Nature chooses each candidate's type. The first candidate's type is private information, while the second candidate — the incumbent — is a type known to all players in the game. Candidate 1 chooses a message: s_x for $x \in \{i, o, w\}$. This introduces the form and effect of the dog whistle. Candidate 1 may choose to send s_w, indicating she sends a coded message. When this occurs, Nature chooses what voters *hear* from Candidate 1: $h \in \{i, o\}$, which depends on both the candidate's message and the voter's type. Voters hear a message about the candidate's group membership according to the whistling probability function:

$$Pr(h_i \mid \beta) = \begin{cases} p & \text{if} \quad \beta = I \\ q & \text{if} \quad \beta = O \end{cases} \tag{2.1}$$

This notation is worth emphasizing Beginning with this function, I introduce a difference between what a candidate says, denoted $s_{x,y}$, and what the voter hears, denoted $h_{x,y}$. This reinforces the idea that the message spoken by a candidate is, in terms of informational content, different from what the voter may perceive.

Substantively, this function operationalizes Aristotle's insight: the hearer's type determines the final form of the message. Group membership, in this process, significantly impacts the beliefs voters form. For the duration of this paper, I set $p > q$ by assumption. This is motivated by the view that in-group members are more likely to recognize words and phrases as carrying secondary content than are people outside the group.[15] If the candidate does not use a whistle, she may either report her true group membership, or lie by claiming the opposite of her group membership. In technical terms, nature does not change either of the other messages, so when candidates send $m \neq w$, voters hear $h = \{i, o\}$ with certainty. To clarify the mapping of candidate messages to what voters may hear, a likelihood matrix of the various probabilities, given candidate and voter types, is presented in Table 1. Finally, voters hear the message determined by nature, then choose for whom to vote in the election.

Candidates' payoffs are equal to their vote share:

$$U_{C_\gamma}(v_\beta) = \sum v. \tag{2.2}$$

[15] For instance, particular religious groups are more likely than others to have recognized President George W. Bush's phrase "wonder working power" as a reference to a church hymn Albertson 2014.

Table 1: **Probabilities that Voter Types Hear Possible Message**

C_1 Message Sent	$Pr(h \mid \beta)$			
	v_I		v_O	
	h_i	h_o	h_i	h_o
s_i	1	0	1	0
s_o	0	1	0	1
s_w	p	$1-p$	q	$1-q$

A voter receives positive utility by voting for the candidate that matches his type. That makes this model one of vote-maximization for the candidate, not one of electoral victory. Candidates seek any increase in vote share, even if it does not produce an electoral victory. Similarly, voters gain utility from voting for the "correct" candidate even if the candidate does not gain a majority of votes. Voter utility is given by:

$$U_{v_\beta}(C_\gamma) = \begin{cases} 0 & \text{if} \quad C_\gamma = v_\beta \\ -1 & \text{if} \quad C_\gamma \neq v_\beta \end{cases} \tag{2.3}$$

Finally, strategies for each player are specified. For Candidate 1, strategies are mappings from types to messages: $\sigma(C_{1_\gamma}) : \{I, O\} \to \{i, o, w\}$. For voters, strategies are mappings from what they hear to candidates $\sigma(v_\beta) : \{i, o\} \to \{1, 2\}$. Note that voter strategies explicitly consider only what is heard, not the message the speaker sent. To reinforce this difference, $m = x$ is used when describing C_1's choices and utility, while $h = x$ is used when describing the voters' choices and utility. In the event that a candidate chooses to use a dog whistle, the voter only hears (and thus updates his beliefs on) an indication of group membership. Since voters never directly hear $h = w$, highlighting the voter's ignorance about whether the message *chosen by the candidate* was a dog whistle or not. The equilibrium concept is Perfect Bayesian.

2.3.2 One-Dimensional Model Results

The general finding from the model is that at least one candidate will opt to either claim group membership that differs from the truth, or will use a dog whistle. The strategic incentives ensure a campaign marked by dissembling and dishonesty. The first piece of this result is the lack of any situation where the candidates speak clearly and honestly to all voters.

Proposition 1. *Regardless of the distribution of voter types, no equilibria exist where both types of C_{1_γ} adopt a strategy of always sending signals that match their true group type: s_γ.*

I include full proofs in the appendix, but will introduce the basic logic to explain the result. Given a strategy of full separation, voters believe any messages they hear. Thus the evaluation of candidate choice is based entirely on matching group identity, which yields the highest utility. If they have any uncertainty about candidate group membership, the voters discount the utility of voting for that candidate. In this case, however, the voters can do no better than voting for the candidate whose message matches the voter's group type. Recall that there are no distributional assumptions on the voter groups; one group may be the majority, the minority, or in perfect balance. Given voter credulity, both candidates will claim the majority group-type. This occurs because the minority voter group will consider both candidates equally (un)appealing, and will choose based on a coin-flip.[16] Both candidates will, when their base is the minority, deviate from telling the truth to lying. In the knife-edge case of voters being perfectly balanced, both in-group and out-group types of C_1 would choose to use a dog whistle. Indeed, the dog whistle provides C_1 a utility gain over sending a truthful message when the voters are unbalanced, though not as much as claiming majority-group identity.

[16] Focusing here, and in the proofs, on a perfectly even chance for choosing one or the other candidate simplifies the analysis without sacrificing substantive content. Any vector of ways each type of voter might randomize a choice produces a new set of parameters that support both truth-telling and whistling equilibria. Crucially, in every set the support remains knife-edged, requiring single values (such as 0.5). If there is reason to suppose the voter chooses the incumbent with greater likelihood – such as 90% of the time – the challenger will simply choose to mimic the incumbent's message, again proving the Proposition.

The effect of the dog whistle is to make clear speech unappealing even when speaking the truth earns the candidate full support from her base. Competition for more votes makes the whistle useful when the population is balanced, otherwise, given the logic of cheap talk, mimicking the advantaged C_1 type maximizes vote share. Clear speech, as an equilibrium, cannot occur.

What happens when dog whistling is *explicitly* part of C_1's strategy? Consider the situation where, when C_1 is a member of the out-group she adopts a dog whistling strategy:

$$\sigma(C_{1_\gamma}) = \begin{cases} m = w & \text{if} \quad \gamma = O \\ m = i & \text{if} \quad \gamma = I \end{cases} \tag{2.4}$$

Whistling introduces uncertainty about candidate types since the message is probabilistic, resulting in equilibria where using the whistle requires very particular voter beliefs. Denote these beliefs by $\mu_\beta, \beta \in \{O, I\}$. Voters update with the standard Bayesian process (where possible), though now a difference arises because of the whistling portion of the strategy. The likelihood term in the updating equation is neither certainty (as it would be for a strategy positing a speaker always issuing one message) nor zero (a it would be in a strategy positing a speaker never issuing a particular message). In the case of whistling as part of C_1's strategy, the likelihood equals the voter's group-type's probability of hearing a particular message (as defined in Equation (2.1)). Since only C_{1_O} ever sends s_o, voter uncertainty only occurs when they hear $h = i$. Out-group voter beliefs about C_1's potential type, given they hear $h = i$, become:

$$\begin{aligned} \mu_O(C_{1_O} \mid h_i) &= \frac{Pr(h = i \mid C_{1_O})Pr(C_{1_O})}{Pr(h = i \mid C_{1_O})Pr(C_{1_O}) + Pr(h = i \mid C_{1_I})Pr(C_{1_I})} \\ &= \frac{qg}{gq - g + 1} \\ \mu_O(C_{1_I} \mid h_i) &= \frac{Pr(h = i \mid C_{1_I})Pr(C_{1_I})}{Pr(h = i \mid C_{1_I})Pr(C_{1_I}) + Pr(h_i \mid C_{1_O})Pr(C_{1_O})} \\ &= \frac{1 - g}{gq + 1 - g} \end{aligned} \tag{2.5}$$

The process is, of course, the same for in-group voters and produces analogous terms for beliefs (see the appendix for details). The practical effect of this change in beliefs is to shift the choice of some voters, depending on the type of incumbent C_2 is known to be. Out-group voters who are certain of C_1's type when they hear $h = i$ would receive

$U_{v_O} = -1$ by voting for C_{1_I}. Now, the uncertainty means they would expect to receive $\mu_O(C_{1_\gamma} = O \mid m = i)(U_{v_O}(C_{1_O})) + \mu_O(C_{1_I} \mid s = i)(U_{v_O}(C_{1_I})) = \dfrac{gq}{gq + 1 - g} > -1$. Voters choosing between two candidates of the same type are indifferent to the choice, and use a toss-up rule to decide. By using the dog whistle, C_1 has broken the indifference some voters feel for the two candidates, making the uncertain choice better in expectation than choosing the known quantity. The dog whistling type of C_1 draws votes from C_2, but also loses to C_{2_O} those v_O voters who hear $h = i$ (for this particular case; more generally it occurs when hearing a message that differs from their own group type) since they would, in expectation, incur some negative utility in voting for C_1. This is the bases for the following result:

Proposition 2. *There exists an equilibrium where the in-group (out-group) type of Candidate 1 will use a dog whistle while the in-group (out-group) type speaks clearly if, and only if, voter beliefs are* $p = \frac{q(1-v_I)+2v_I-1}{v_I}$ *(*$q = \frac{p(1-v_O)+2v_O-1}{v_O}$*).*

Any other value for p (or q) does not support the equilibrium, and instead results in *both* types of C_1 preferring to use a dog whistle. The precarious nature of this equilibrium, resting on knife-edge values for beliefs, is a result of both types of C_1 having identical motivations.

The probabilistic nature of the whistling function is important for the results, but also substantively contradicts a common argument for why dog whistling occurs: the message is not perfectly tuned for just one group, and some people outside the target audience *do* hear it. This aspect of whistling is often drowned out in a pundit's claims about the presence of a specific whistle. Indeed, if candidates could keep a perfect division between the targets of the whistle and the rest of the audience the ideologically opposed pundits would never hear the whistle in the first place.

Leakage between groups is an empirical regularity that my model not only allows for, but makes a specific part of the strategic calculation. A comparison of p and q from Equation (2.1) shows that as $p - q$ grows, the less the spillover into the out-group matters, since the number of in-group voters that hear the whistle gets much larger than the out-group that do. The decision to use a whistle, whether as part of the stated strategy or as an option for deviation, requires the candidate to weigh these proportions of voters. Substantively, an increasing divide between p and q reflects a greater cultural division between the in- and out-group; the language the in-group uses is increasingly opaque to the out-group. President Obama's use of the terms "bamboozled," "hoodwinked," and "okey-doke" were criticized as dog whistles, but ones that were especially hard for white citizens to understand since they (may have) come from speeches attributed to Malcolm X and Jesse Jackson.[17] While the names of these public activists are well-known, the nuances of their speech are grasped far more by African-Americans than others. This distinction between the in- and out-groups is far higher there than when President Obama (allegedly) whistled to suggest Mitt Romney was cruel to animals by promoting actual dog-related Obama campaign merchandise.[19]

Proposition 3. *Three equilibria exist where both C_{1_γ} (for $\gamma \in \{O, I\}$) pool on sending the same message, either $s = i$, $s = o$, or $s = w$.*

Given the pooling strategy, neither type of C_1 ever benefits from choosing a different message other than the one posited in the equilibrium. Utility calculations are shown in the appendix, so suffice it to say here that, given the pooling strategy, voters learn nothing from hearing any message from C_1 and so would not change their vote on hearing a different message. To see why, recall that the expected utility of voting for C_1 is now strictly in $[0, -1)$ for either voter type. C_1 thus gets all the votes from the opposite group-type of C_2 since those voters get some amount less than -1, which would be their payoff for voting for C_2. For example, v_O voters will always prefer C_{2_O} over C_1 regardless of the message heard, since there is some positive probability they would be voting for a C_{1_I}. And, in this example, v_I voters are not indifferent between the two candidates since there is some positive probability of C_{1_I}. Thus, for C_{2_O}, C_{1_γ} for $\gamma \in \{O, I\}$ attracts all v_I voters, and no v_O voters. Importantly, returning to using a clear message of group type is not beneficial for either C_{1_O} or C_{1_I}. The uncertainty about the candidate's true

[17] Jan Freeman, "Is Obama speaking in code?," *The Boston Globe*, 10 February 2008.[18]

[18] Complicating the issue further, the word choice may stem from a movie depiction of Malcolm X, and not his actual speech. Still, the audience for the movie "Malcolm X" was likely predominantly African-American.

[19] Michael Sherer, "The Obama Campaign's Humorous Dog Whistle," *Time*, 31 January, 2012.

type remains, and the voters are unmoved. The same logic explains why both types of C_1 pooling on dog whistling is an equilibrium: switching to a different messaging rule does not resolve voter uncertainty, and so does not bring back the voters lost by using the whistle.

Voters, however, frequently consider multiple aspects of candidates before making a decision. The whistling function addresses not only how this strategic tool can be represented as part of a signaling game, but has also shown the effect it has on voter information in a limited setting. I now expand the setting to include a second dimension for both candidates and voters to examine the use of dog whistling in a richer strategic setting.

2.3.3 Two Dimensional Model

I extend the basic model by adding an ideological dimension to the candidate and voter types, and introduce costs for candidates either lying or using a dog whistle. Candidates again play on voter uncertainties raised by the whistle, but do so to gain the support of voters who would otherwise be indifferent over the candidates. Extending the model to two dimensions changes candidate messaging strategies and player utility, but not the substantive process of the game.

Candidate i's type is now a pair $(\tau, \gamma) \in \{M, E\} \times \{I, O\}$, which is private information for C_1. The pair identifies ideology and group membership; τ indexes ideology, where $\tau = M$ indicates C_i is a "Moderate" type, and $\tau = E$ indicates an "Extreme" type. Candidates are moderate with $Pr(\tau = M) = t$. γ indexes group membership, where $\gamma = I$ indicates C_i is "in-group," $\gamma = O$ indicates she is "out-group". Candidates are out-group with $Pr(\gamma = O) = g$.

There are four distinct groups of voters, also identified by their ideology and group membership. Types are defined by $(\theta, \beta) \in \{M, E\} \times \{I, O\}$, matching the type space for candidates. Rather than identify probabilities of each voter being of a certain type, denote the size of each group by $v_{\theta\beta}$, where $\forall\, \theta, \beta$, $v_{\theta\beta} \geq 0$ and $\sum v_{\theta\beta} = 1$. This reflects the concern of winning portions of the electorate according to how many may hear the dog whistle.[20]

In addition, I add a weight to the voters' ideological dimension, denoted $\alpha \in (-\infty, \infty)$. This weight changes a voter's preference for matching in one dimension over another. Highly ideological voters, for instance, may prefer a candidate that is a close adherent to some ideological stance – say, anti-immigration or anti-genetically modified foods – regardless of the group identity. On the other hand, some voters may simply prefer a candidate who unswervingly supports a single group, regardless of the ideology of its members. This results in the following utility function for a voter:

$$U_{v_{\theta,\beta}} = \left\{ \begin{array}{ccc} 0 & \text{if} & v_\theta = C_{i_\tau} \\ -\alpha & \text{if} & v_\theta \neq C_{i_\tau} \end{array} \right\} + \left\{ \begin{array}{ccc} 0 & \text{if} & v_\beta = C_{i_\gamma} \\ -1 & \text{if} & v_\beta \neq C_{i_\gamma} \end{array} \right\} \tag{2.6}$$

Candidates' payoffs are a function of the vote share and whether the message matches their type:

$$U_C = \sum v_{\theta\beta} + \left\{ \begin{array}{ccc} 0 & \text{if} & x = \tau \\ -\lambda & \text{if} & x \neq \tau \end{array} \right\} + \left\{ \begin{array}{ccc} 0 & \text{if} & y = \gamma \\ -c & \text{if} & y = w \\ -\lambda & \text{if} & y \neq \gamma, w \end{array} \right\} \tag{2.7}$$

There are two parameters in the utility function that require explanation. One parameter, c (with $0 < c < 1$ by assumption), is a constant, used to capture a cost to the speaker for using the dog whistle. While vague stances on policy often go unremarked, the media is often quick to call out what they perceive to be a dog whistle, which candidates must engage, expending valuable time and resources. The λ parameters reflects an individual cost to lying outright, with higher values indicating greater aversion.

[20] I use a combined subscript to denote voter groups when two subtypes are to be considered together. For instance, v_O denotes all of the out-group voters of both moderate and extremist types, $v_{E,O} + v_{M,O}$. This both saves on notation and facilitates comparison.

2.3.4 Two Dimension Model Results

While the one-dimension model shows how whistling allows one candidate to mimic the other in order to secure votes, the two-dimensional model shows that whistling is primarily useful as a method for C_1 to distinguish herself from C_2 in the eyes (or, ears) of one group.

The first situation I consider is a single type of C_1 choosing to adopt a dog whistling strategy, while all other types clearly reveal their true type on both dimensions. As the next proposition states, the single type of C_1 will choose to use a dog whistle, given that she is identical in type to C_2. More specifically:

Proposition 4. *Given a sufficiently high cost of whistling (c), and a sufficiently low cost of lying (λ), when $C_{1_{\tau,\gamma}} = C_{2_{\tau,\gamma}}$, there exist equilibria where that type, and only that type, of C_1 uses a dog whistle and any $C_{1_{\tau,\gamma}} \neq C_{2_{\tau,\gamma}}$ reveals her true ideological and group types.*

Whistling, in this instance, is a surreptitious method for the candidate to distinguish herself from the incumbent. Voters who would have been indifferent between a type of C_1 that is identical to C_2 are swayed to vote for C_1, since the whistle makes the voter uncertain about the candidate's type. In expectation, voting for C_1 is an improvement over just choosing randomly between C_1 and C_2

This first result also demonstrates the impact costs have on the candidate's choice of message. As with any costly signaling, the higher the penalty for the signal, the less frequently it will be used. In this instance, the costs of whistling must be low enough to induce $C_{1_{\tau,\gamma}} = C_{2_{\tau,\gamma}}$ to not switch to stating her true group membership ($s_{x,y} = \{\tau, \gamma\}$). Recall that with whistling as part of the stated strategy, voters remain uncertain about C_1's type whether they hear $h_{x,y} = \{e, i\}$ or $h_{x,y} = \{e, o\}$. That is, even when $C_{1_{\tau,\gamma}} \neq C_{2_{\tau,\gamma}}$, voter beliefs are still affected by the presence of whistling. $C_{1_{\tau,\gamma}}$ sending either $s_{x,y} = \{\tau, o\}$ or $s_{x,y} = \{\tau, i\}$ means that no voter hears the other message, eliminating the cost, but ceding votes to C_2. Costs, then, are roughly akin to losing a number of votes. This could be readily observed in actual campaigns, though it does not discount the potential psychic costs the candidate faces from deceiving voters.

One might consider whistling to be a tactic employed by a certain kind of candidate; certainly Republicans are more often accused of engaging in the behavior than are Democrats.(Haney-López 2014) Or, whistling may be a machination of out-group candidates to surreptitiously indicate membership with a group out of mainstream favor. Whistling might also be used indiscriminately by all types, as a matter of routine communication during a campaign. Interestingly, however, none of these options garners $C_{1_{\tau,\gamma}}$ more votes under the dog whistling strategy. This result is a combination of three related results that differ only slightly in their analysis, but together have a larger implication for campaign rhetoric. The three facets of this result are as follows:

Proposition 5. *There are no equilibria supported by the following strategies:*

- *Ideological types pool on whistling:* $\sigma_{C_{1_{\tau,\gamma}}} = s_{\tau,w} \, \forall \, \tau \in \{M, E\}$
- *Group types pool on whistling:* $\sigma_{C_{1_{\tau,\gamma}}} = s_{\tau,w} \, \forall \, \gamma \in \{I, O\}$
- *All ideological and group types pool on whistling:* $\sigma_{C_{1_{\tau,\gamma}}} = s_{\tau,w} \, \forall \, \tau \in \{M, E\}, \gamma \in \{I, O\}$

This result says, in essence, that whistling only serves a useful purpose when it is done used by select types of challenger. Pooling, as a strategy, results in voters ignoring the message entirely, and returning to their prior beliefs about the type of candidates they are considering. The logic for each of the facets of the result follows.

Pooling messages by ideological type – both Moderate and Extremist types within either the in- or out-group – does not increase C_1's vote share. The result occurs because messages end up "clashing". When both ideological types within a single group send a whistle, the voters effectively learn no new information, and their posterior beliefs about candidate type equal their prior beliefs about the candidate's possible group-type. Note that this is different than the voters being *certain* about the candidate's type. Uncertainty remains, as the calculations in the proof show, but does not differ for each ideological type, as occurred when only one type whistled in Proposition 4. Because of the clash, the vote share C_1 obtains when sending a whistle is the same as when she reveals her true type, but incurs the cost of whistling, λ. Clearly, then, whistling is not the optimal message, and there is no equilibrium where the ideological types choose a pooling strategy.

The same process prevents group-types from pooling on whistling regardless of their ideology. Clashing messages, and the lack of information transfer to the voters, again keeps candidates from increasing their vote share via dog whistling. The flavor of this facet is identical to the previous one, except here the strategy indicates that all out-group (or in-group, but not simultaneously) types of $C_{1_{\tau,\gamma}}$ use a dog whistle, regardless of being a moderate or extremist. Finally, the logic carries over to considering whether all types of C_1 would pool on sending a dog whistle.

2.3.4.1 Empirical Implications Assuming for the moment that dog whistling can be clearly identified when it occurs, there are clear expectations that come from the results above.

First, and most generally, dog whistling appears in many campaign situations. That is, it is either the equilibrium messaging strategy for candidate one, though under very specific (and thus likely to be empirically rare) conditions. The ability to whistle does, however, make clear speech less attractive in a number of situations. The single-dimensional model rules out the possibility of clear speech entirely, since candidates can amass more votes through either whistling or outright lying. Expanding the model to two dimensions and adding costs does not entirely mitigate the endemic nature of dog whistling. Certainly, exogenous costs can be raised to the point where no prevarication occurs; the practical question is how to raise penalties for dog whistling high enough to tamp it out. One attempt was made in Kenya, during the 2013 campaign for Prime Minister. An independent organization monitored radio stations, newspapers, and television, catch the use of "retrogressive utterances" (words one ethnic group would use in an opaque, and derogatory, reference to members of a different ethnic group). A widespread media campaign, accompanied by the government threatening large fines and jail time, failed to eradicate the practice.[21] Pundits decrying potential dog whistles on popular U.S. media outlets cannot be considered sufficient penalty.

More particularly, the results suggest that dog whistling should be prevalent when the challenger and the incumbent are considered nearly identical. The 2012 Republican Primary contest is a useful case to compare against these results. The only concession to the real world I need to make is to cast Mitt Romney in the role of the incumbent, rather

[21] Staff, "Kenya cracks down on hate speech ahead of poll," *BBC News*, 26 February, 2013.

than make comparisons against President Obama. Recall that the model is restricted to the case of primary contests, and considers only a single political party. Romney, then, as the most prominent contestant for the nomination, measured by his time leading the polls (Sides and Vavreck 2014), serves the incumbent function well. Claiming a position as a fiscal and social conservative, Romney could be viewed as the extremist, out-group candidate (ignoring his Mormonism for the moment). The other extremist candidates attempting to "play to their bases" of social and fiscal conservatives in the case of the Republican Primary – included Michele Bachmann, Rick Perry, and Newt Gingrich, all of whom thus had a need to distinguish themselves from Romney through the possible use of dog whistles. John Hunstman, on the other hand, ran expressly as a moderate alternative to Romney.[22] And, as expected, all three of the extremists (possibly) resorted to sending dog whistles: Michele Bachmann opposed immigration;[23] Rick Santorum accused African Americans of being totally dependent on the federal government;[24] Rick Perry says President Obama is muslim;[25] and Newt Gingrich made multiple coded appeals.[26] The varied issues and disparate (purportedly) targeted audiences for these whistles reinforces the result above that dog whistling can be a tacit method to distinguish oneself from the competition. While this means that more than one person in the campaign used a dog whistle, the lack of any real difference between the challengers on the dimensions captured in my model means that these are simply repeated instances of the same, single, type of challenger.

A more subtle implication of my results suggests that voters who hear whistles from multiple types of challenger should ignore what they hear and rely on their initial beliefs. If differences among Bachmann, Perry, and Gingrich are sufficient to make them distinct types of candidates, my results suggest that the whistling from all three will collide, and render the content meaningless for voters. Sides and Vavreck (2014, pp. 45–52) examine Rick Perry's cycle of media coverage and support in the polls during his "boomlet". While

[22]Staff, "Not just too moderate," *The Economist, Democracy in America*, 16 January, 2012

[23]Van Le, "Michele Bachmann Blows the Dog Whistle on Immigration," *America's Voice*, 14 September, 2011.

[24]Joan Walsh, "Rick Santorum flip-flops on black people," *Salon*, 5 January, 2012

[25]Staff, "Why Rick Perry is airing an 'anti-gay' ad: 5 theories," *The Week*, 8 December, 2011.

[26]Melinda Henninberger, "Newt Gingrich: Against racial dog whistles before he was for them," *Washington Post*, 31 January, 2012.

media attention varied sharply in response to particular words and phrases (such as a dog whistle claim to being a member of the "birther" movement that claimed President Obama was not born on U.S. territory), his support in the polls was largely unmoved as a result. Whistling from the rest of the field may have rendered his coded appeals impotent.

2.4 DISCUSSION

The model and results presented here offer a single framework for an important political phenomenon that has only been approached in a piecemeal fashion (Albertson 2014; Fear 2007; Haney-López 2014). By abstracting away from the specific content of the messages, I show that dog whistling is a powerful tool for manipulating voter choice, usable regardless of candidate group identity or political affiliation. Moreover, there are several key insights produced by this work.

First, individual audience characteristics determine, in part, the effect of political speech. In this work, a candidate is able to use the different group likelihoods of hearing the whistle to gain enough votes from one audience to offset potential losses in the other. Moreover, this difference between the groups need not be pronounced, and candidates benefit by playing to a minority group even in the face of higher whistling costs. Substantively, prevarication or outright lies have entirely different effects, since they do not harness the fundamental difference in audience interpretation.

Next, while dog whistling is, by construction, a slippery tool, imperfectly identified when used, we can make clear statements about why the tool works, and in what situations we should anticipate its use. Identical candidate types, as noted in Proposition 4, make the use of a dog whistle profitable, given specific conditions on the costs of mendacity. Whistling serves as a method for a challenger to distinguish herself from an incumbent, or a single candidate from a field of similar contenders, by appealing to all voters, including those who would otherwise be indifferent in their dislike over options. When neither the challenger nor the incumbent are members of the in-group, whistling is still useful; in-group voters move from their prior beliefs to their focal point belief, earning votes from the portion that hear the whistle. Dog whistling in more general cases reaches a limit of usefulness, however. Candidates with the same ideological types cannot increase their vote share by adopting a whistling strategy, since voters simply ignore the message they hear, whistle or not, and act on their prior beliefs about types. For the same reasons, Candidate 1 will not simply opt to whistle regardless of her type in either the ideological or group dimensions. Dog whistling might be a tool with wide applicability, but it cannot be used indiscriminately if it is to be profitable. Indeed, for dog whistling to have any bite, it must be used in specific instances in relation to how similar or distinct one candidate is from another in the minds of voters. We should expect to see whistling occur, then, as an increasing function of candidate similarity, and not, as popular accounts imply, simply because of minority (or extreme) racial and religious views.

Without more statements of intent, like Senator McCaskill's, we cannot be certain when a candidate has specifically opted to use a whistle. That said, my model suggests a second empirical regularity: once one candidate is accused of whistling, other candidates should soon follow suit. Newt Gingrich and Michelle Bachmann both were accused of using whistles during the 2012 Republican Presidential Primary. Claims that Mitt Romney and Rick Perry engaged in whistling were not far behind. Introducing a dog whistle into campaign messaging makes it difficult, if not impossible, to return to unambiguous rhetoric. Candidates clearly benefit from being vague about future policy choices (Alesina and Cukierman 1990; Glazer 1990; Shepsle 1972; Tomz and Van Houweling 2009), but usually suffer at the hands of the media from outright falsehoods. Dog whistling, by its nature, shifts the source of uncertainty onto the audience, allowing candidates to credibly deny some aspect of the message. However, once a message is identified as a possible dog whistle — identifying the candidate as adopting a whistling strategy, in terms of the game — the audience will be left wondering, even if the candidate has shifted back to speaking clearly. The very existence of the dog whistle, then, makes its use viable.

Dog whistling, despite its now-clichéd name, provides a starting point for analyzing more nuanced aspects of speech, campaigning, and candidate manipulation of voter beliefs (and thus actions). Extensions to the models presented here include making the relationship between the dimensions correlated: as candidates grow more extreme, there is a corresponding increase in the belief that she will dog whistle. Expanding the game to explicitly include the media might allow voters to have correlated beliefs about the candidate's types. Issues related to deliberation and discourse might be addressed by allowing the voters to communicate their beliefs to each other, reflecting the wider use of social media in modern elections. Empirical investigations of coded speech can focus on both their active use by candidates, as well as the source of individuated beliefs in the audience. For the former, analysis of speeches and campaign material, compared against a reference corpus, may identify the number of times the source material is references without citation. For the latter, individuals could respond to prompts to see what associations arise. Any and all of these additions would build on this first step to flesh out the strategic interaction present in Aristotle's triumvirate of rhetoric.

3.0 TIME AND PUNISHMENT: BLAME AND CONCESSION IN POLITICAL STANDOFFS

Coauthor:

Jonathan Woon, Associate Professor of Political Science and Economics

Faculty Affiliate, Pittsburgh Experimental Economics Laboratory

University of Pittsburgh

> "Our message to the United States Senate is real simple: The American people don't want the government shut down, and they don't want Obamacare."
>
> Rep. John Boehner

> "I want to be absolutely crystal clear: Any bill that defunds Obamacare is dead. Dead."
>
> Sen. Harry Reid

Political negotiation often resembles a staring contest, rather than the back–and–forth offer–counteroffer process usually associated with deal-making and compromise.[1] Consider recent struggles over U.S. fiscal policy. In September 2013, congressional Republicans demanded the elimination of funding for the Affordable Care Act as a condition of passing continuing appropriations to keep the government operating. Democrats refused to back down, and a 17-day government shutdown ensued. During the shutdown, polls indicated

[1]We generally avoid the term "bargaining" throughout the paper. While standoffs occur during a bargaining process, our interest is on the impasse and potential failure of the bargaining event, rather than choices of offers during bargaining. Both our model and experiment are sufficiently different from the usual models of bargaining that we want to keep the distinction clear for the reader. We make an exception in the discussion of prior literature, where commonalities between our study and work on bargaining is clear.

that the public largely blamed Republicans and eventually the Republican leadership capitulated.[2] Four months later, having reasoned that allowing the shutdown had been a mistake, Republicans would not be so intransigent.[3] Facing a looming deadline over raising the debt limit, they acceded to Democratic demands well ahead of the potential default, and the public hardly noticed.[4] Similar events unfolded in 1995 between Newt Gingrich and Bill Clinton. These events suggest that incurring blame is costly for politicians, and that the anticipation or avoidance of blame can feature prominently in strategic calculations.

How do observers of political stalemate allocate blame for outcomes, and how does the expectation of blame affect the behavior of negotiators? Answering these questions is important for understanding how citizens can exert influence over their representatives through non-electoral means. To study these questions we conduct an experiment that contrasts standoffs with and without an audience. In our setup, subjects play a war of attrition game in which the only choice is when to back down and waiting is costly. A single-shot game is used to isolate the incentives for allocating blame for the outcome from concerns about discounting, reputation effects, and potential unraveling results. When we add an audience, we operationalize blame by allowing the observer to punish the standoff contestants. The war of attrition framework captures the essential features of standoffs outlined in the example above. It is also important for our purposes because ex ante negotiating power is symmetric, permitting us to investigate how observers ascribe blame as a function purely of outcomes, free of considerations of structural power.

This paper proceeds as follows. First, we emphasize the importance of studying blame in negotiations and standoffs, then demonstrate that our approach is both novel and substantively appropriate. Next, we describe the design and procedures of the experiment. From this we describe what payoff maximizing behavior predicts for Staring Contest outcomes and the use of blame. Briefly, if observers engage in rational punishment and

[2] Andrew Dugan, "Republican Party Favorability Sinks to Record Low." *Gallup Politics*, 9 October, 2013.
Scott Clement, "Republicans are losing the shutdown blame game", *The Fix*, 4 October, 2013.

[3] John Breshnahan, Manu Raju, Jake Sherman, and Carie Budoff Brown, "Anatomy of a shutdown", *Politico*, 18 October, 2013.

[4] Paul Kane, Robert Costa, Ed O'Keefe, "House passes 'clean' debt-ceiling bill, ending two-week showdown", *Washington Post*, 11 February, 2014.

contestants correctly anticipate this, then the presence of an observer in the Staring Contest will shorten the time players hold out before conceding. In addition, the contestant whose preferences are aligned with the observer should win more often. This occurs because the observer adopts a strategy that punishes the winning player, thereby reducing delay. The empirical results provide mixed support for the predictions of rational play; behavioral models that incorporate preferences for equity of winnings may explain some of the gap. The observer frequently uses blame to target the winning player, and contestants decrease their waiting times in the presence of an observer. Outcomes, however, change only slightly, if at all. Contestants, it seems, inadequately anticipate how they will be punished and fail to take action to improve their payoffs. We conclude with a discussion of these findings and their general implications.

3.1 THE POLITICS OF BLAME

We consider blame to be a form of punishment citizens use to register their displeasure concerning some event or action. Political decision makers seek to avoid punishment at all levels, from elites (**weaver1986**; Hood 2010; Weale 2002) to street-level bureaucrats (Brehm and Gates 1997; Lipsky 1980). Doling out punishment to sitting politicians entices some voters to the ballot box (Brown 2010; Iyengar 1989; Peffley 1984). Negative assessments of political culpability, such as opinion polls, are weak forms of blame, but arise in numerous contexts (e.g., Alcañiz and Hellwig 2011; Bengtsson 2004; De Vries and Hobolt 2012; Malhotra and Kuo 2007). We build on prior research about responsibility attribution, as well as work on strategic interactions between bargainers and an audience.

Attributing blame requires an observer to make a valid link between the causal drivers and the outcome. Citizens have weak, and malleable, understanding of these links when considering large scale failures (Malhotra and Kuo 2007). Indeed, they mostly rely on partisan identification when deciding who is most responsible for economic conditions or policies (Brown 2010). Blame is, however, more nuanced when citizens have better information (De Vries and Hobolt 2012). In contrast to our setting, these studies present large, often highly complicated problems (massive environmental disasters and macroeconomic conditions) to voters with varying internal motivations for devoting attention and thought. We isolate these competing pressures in the experiment by providing complete information and clearly delineating incentives, and making the causal link transparent.

Blame, as a logical subset of responsibility (Lagnado and Channon 2008; Shaver 1985), carries similar importance to the study of economic voting. However, while theories of prospective and retrospective voting entail the attribution of responsibility for economic conditions (Hellwig 2010; Lewis-Beck and Paldam 2000; Lewis-Beck and Stegmaier 2000; Paldam 2008), they do not address how voters *apportion blame* for producing those conditions. Psychologists define blame as the impulse to punish an actor for a negative outcome over which the actor has some direct control (Alicke 2000; Brewin and Shapiro 1984; Weiner 1995).[5] The distinction between responsibility without blame and responsibility with blame is important because it implies potentially different voter behavior. For example, voters may hold the president responsible for employment levels, but do not *blame* him for conditions inherited from the last administration. Voters may, however, blame a political party that blocked job-creating legislation, then punish them at election time. Treating blame as a distinct concept provides a richer view of voter behavior.

Blame and reputation are interrelated, but not synonymous; distinct models of each are necessary. Repeated games often model reputational concerns as beliefs about an opponent's promised future actions, given the opponent's history of behavior, or unobservable payoff-relevant types (Abreu and Gul 2000; Celetani et al. 1996; Embrey, Fréchette, and Lehrer 2014; Kreps, Milgrom, et al. 1982; Kreps and Wilson 1982; Roth 1985). These models do not address punishment as a distinct mechanism, and reputational concerns arise only between bargainers, since no audience is present. Another form of reputation arises in games of incomplete information, such as in Groseclose and McCarty (2001), who incorporate both blame and an audience in a model of veto bargaining. Voters,

[5] This is a counterfactual-based definition of causality. A caused B (in whole or part) if B would not have happened without A's presence. The *desire* to punish this causal connection is sufficient for psychology, though we choose to make it tangible.

as the audience, apply blame by lowering approval ratings for the president if he vetoes a certain bill. In this model, the voters assign both full responsibility, and thus full blame, to only one side of what is a strategic interaction. In contrast, our experiment allows the audience to blame neither, one, or both parties according to their assessment of both actors' choices. International relations studies of bargaining often include audiences that can impose a cost on political leaders as a function of the bargaining outcome (Fearon 1994; Tomz 2007a,b; Weeks 2008). Moreover, we follow Fearon (1994) in employing a war of attrition model. Domestic politics, too, are often described as wars of attrition (Wawro and Schickler 2006), and have been explicitly modeled this way (Kousser and Phillips 2009; Kovenock and Roberson 2009). We complement this by making the cost mechanism – a subject of some debate (Schultz 2001; Trachtenberg 2012) – explicit.

Our operationalization of blame is similar to the use of punishment in public goods experiments, making them appropriate bases on which to build our examination of standoffs and players' outcome preferences.[6] We choose to make blame costless to the observer, to render the application of punishment independent of the game's structure. That said, experiments report significant use of punishment even when the punisher incurs a cost to do so (**fehr2000**; Fehr and Gächter 2002; Masclet et al. 2003). Punishment has proven useful in studying a wide array of social situations, despite debates about what the findings mean (Guala 2012). Blame may prove a similarly useful experimental tool for understanding political phenomena.

3.2 EXPERIMENTAL DESIGN AND PROCEDURES

To investigate how the presence of an audience and the potential for punishment affects behavior during a standoff, and how an audience might actually punish the standoff contestants, we designed a simple experiment that contrasts standoffs with and without an audience. For our framework, we chose to use a war of attrition game in which waiting

[6]The general public goods literature is expansive. For a useful summary of punishment and cooperation in experimental settings, see Chaudhuri (2011).

is costly. When we add an observer, the observer's payoffs depend on the outcome of the attrition game, and we operationalize "blame" by allowing the observer to punish the players by deducting points. We also vary the alignment of interest between the observer and the players by modeling outcomes as spatial policies and varying the ideal point of the observer.

We chose the war of attrition framework rather than a sequential bargaining game for two reasons. Substantively, we are interested in the duration of standoffs, as we believe that recent episodes of presidential-congressional and inter-cameral negotiating over the federal budget, debt limit, and sequestration seemed to resemble a process in which each side stakes out a position and waits for the other to concede rather than a process of negotiation that leads to compromise outcomes. From a design standpoint, the war of attrition framework allows us to keep proposal power symmetric between the standoff contestants. This is important because we are interested in understanding how observers endogenously assign blame as a function of the choice to prolong or end the stalemate, free of considerations arising from exogenous structural features of a standard bargaining situation (such as proposal power or a first mover advantage).

The experiment utilized a within-subject design and consisted of two parts. In Part 1, subjects played a war of attrition game without an audience, while in Part 2, they played the same game with an observer who was permitted to administer ex post punishment. In each part, they played the game multiple times. We read the instructions aloud and had subjects complete a comprehension quiz prior to each part of the experiment. Thus, in Part 1, subjects did not know the rules of the game they would be playing in Part 2, thereby precluding any incentives to form reputations in anticipation of Part 2. We also implemented standard design features to minimize the interaction between rounds, ensuring as much as possible the one-shot nature of the incentives of each play of the game. First, we used anonymous, random rematching of groups between rounds so that subjects never knew which of the other subjects they were paired with. This feature reduces any reputational incentives subjects may have. Second, we randomly selected one round from the entire session to count as payment. This feature eliminates wealth effects and ensures that a subject's payoffs in later rounds are independent of their payoffs in earlier rounds.

In Part 1, subjects played 12 rounds of a two-player war of attrition game without an audience. In our experimental instructions, we referred to the war of attrition as a "Staring Contest Game" to emphasize its dynamic nature and to convey the intuition that the player who waits the longest in the game obtains a better outcome. We can think of the player who "stares" the longest as "winning" the game and the player who "blinks" first as "losing." Although we referred to the game as a Staring Contest, we did not otherwise use the terms "staring," "blinking," "winning," or "losing" to describe the game.

The Staring Contest Game is played by two players, which we designate Player A and Player B. Each subject played 6 rounds as Player A and 6 rounds as Player B. The game lasts for 30 seconds, and each player's only decision is when to "concede" by ending the game. In our graphic interface for the game — programmed in z-tree (Fischbacher 2007) — each player sees a timer bar that decreases in size at 1-second intervals, and chooses to concede by moving the cursor over the timer bar. We designed the interface in this way so to avoid audible clicks of the mouse that would signal to other subjects when players in other groups conceded. If both players concede at the same time, we use a random tie-breaking rule.

When the Staring Contest is played dynamically, as in our design, we only observe the actual concession time of the game for the losing player. We do not observe the time at which the winning player would have conceded, so in a way, our data are censored. The standard solution to this kind of censoring problem in experiments with sequential games is to use the "strategy method," in which subjects do not play the game dynamically but instead indicate their complete strategies before other players' choices are revealed. But the strategy method removes the dynamic element of the game that interests us. Thus, we opted to use a mixture of dynamic play and the strategy method. At the beginning of each round of the Staring Contest, we asked each player to state an "Intended Stopping Time" (between 0 and 30 seconds), and there was a 1 in 10 chance that we would implement the Intended Stopping Time as the player's actual stopping time. The realization of the

Intended Stopping Time was independent across players and periods. We intended the small chance that the Intended Stopping Time would be implemented as a way to make the choice meaningful (rather than completely hypothetical) while ensuring that subjects would play the dynamic version of the game most of the time.

We described the outcomes and payoffs of the game in terms of a one-dimensional spatial model of policy.[7] If Player A wins, the "outcome number" of the game is $a = 10$, while if Player B wins, the outcome is $b = 90$. If neither player concedes, we consider the outcome of the game to have broken down in so far as neither side gets the outcome they prefer despite both holding out.[8] To construct the players' payoffs in the Staring Contest, each player is described as having a "target number," which we can think of as the player's ideal point. We assign A the target number 0, and B the target number 100. If we denote the outcome of the game by $x \in \{a, b, \phi\}$ (where ϕ denotes disagreement) and the game ends at time t (in seconds), then player i's payoff (denominated in points) is given by the function

$$u_i(x,t) = \begin{cases} 350 - |x - \theta_i| - t & \text{if } x \neq \phi \\ 190 & \text{if } x = \phi \end{cases}$$

where θ_i denotes the player's target number. Note that "winning" the Staring Contest results in payoffs between 310 and 340 points, while "losing" results in payoffs between 230 and 260 points. Disagreement is the worst possible outcome, and it is equally "bad" for both players in order to preserve the symmetry of the players' incentives in the game. Substantively, neither side of a political standoff needs to know the actual value of getting one policy or the other. Instead, these actors only need an ordinal ranking: the party getting exactly the policy they want is the best outcome; any policy between their preference and the other party's is worse; a policy exactly at the other party's preference is worse still; finally, the outcome corresponding to a government shutdown in which *no* policy is enacted, not even the status quo, is the worst possible outcome for either party.

[7] When we introduce an observer in Part 2, this allows us to manipulate the alignment of preferences between the observer and the two contestants.

[8] We do not use the terms "breakdown", "bargaining failure", "standoff", or any similar term in describing this outcome to subjects. We only refer to a default payoff if neither player concedes.

In each round of Part 2, randomly matched groups of three participants played the Staring Contest Game with an audience. We designated two players in each group to be the "contestants" (Players A and B) in the Staring Contest and the third player to be an "observer." Every subject played 6 rounds in Part 2 as Player A, 6 rounds as Player B, and 6 rounds as the observer. Play in each round consisted of two stages. The contestants first played the Staring Contest Game exactly as they did in Part 1. The observers then learned the outcome and their payoffs from the Staring Contest and chose how to allocate "blame" or "punishment" by deducting points from one or both contestants' payoffs. The observer has the ability to allocate punishment to one, both, or neither of the contestants and could deduct any amount of points as long as the sum of deductions was between 0 and 100 points.[9] We chose to make punishment costless rather than costly so we could observe the maximum amount of punishment subjects might be willing to give.

The contestants' payoffs in this version of the game are their Staring Contest payoffs minus the observer's deduction, $u_i(x,t) - d_i$, where d_i is an integer between 0 and 100 and the sum $d_A + d_B$ is between 0 and 100. The observer's payoff is given by the same function $u(x,t)$ that describes the contestant's payoffs in Part 1, except that we varied the ideal point of the observer, $\theta_O \in \{0, 25, 50\}$, in order to vary the alignment of interests between the observer and contestants. When $\theta_O = 50$, the observer's interests are aligned equally with Player A and Player B. When $\theta_O = 25$, the observer's interests are more aligned with Player A than Player B. And when $\theta_O = 0$, then the observer is completely aligned with Player A.[10] This alignment models partisan bias in the audience, doing so only for one side to keep the experimental conditions tractable. Information about the observer's ideal point was announced to all players at the beginning of every round and therefore common knowledge. Each subject played two rounds in each role with each of the possible observer ideal points.

[9] Van De Ven and Villeval (2014) develop a deception game with an observer that can reveal a player's lie. Both this model and ours make the player outcomes dependent on the choice of a second-mover audience, though the negotiating parties in our model cannot take unobserved actions.

[10] Note that the observer can make conceding the worst possible outcome by deducting 80 points or more from the conceding player.

A total of 72 subjects participated in five sessions of the experiment, which took place at the Pittsburgh Experimental Economics Laboratory (PEEL).[11] Subjects were recruited from PEEL's general subject pool, gave informed consent according to standard procedures, and were privately paid in cash at the conclusion of the experiment (including a $5 show-up fee).

3.3 THEORETICAL EXPECTATIONS

In our theoretical analysis, we first consider how payoff-maximizing subjects would play the baseline Staring Contest Game and then consider how the observer's actions might affect concessions and timing. Even though subjects play the game in real time, we can analyze the game as a simultaneous move game in which a player's strategy is the amount of time (in integers) he or she waits until conceding, denoted by t_i for $i \in \{A, B\}$.

In the game without the observer, each player always has an incentive to wait longer than the other. For any strategy for Player B, $t_B < 30$, Player A's best response is to wait a little longer and to concede at any time t_A such that $t_A > t_B$. To see why, note that A's payoff for waiting longer than B is $340 - t_B$ while conceding at $t_A < t_B$ yields a payoff of $260 - t_A$. Since $t_B - t_A$ is at most 30 points and the difference between winning and losing is 80 points, it is better to wait and win the contest than it is to end the game quickly on the losing side (because $340 - t_B > 260 - t_A$). If both players wait until the last second, $t_i = 30$, then the random tie-breaking rule implies an equal likelihood of obtaining a payoff of 310 (from winning) and a payoff of 230 (from losing). The expected payoff when both players wait until the very last second to concede is therefore 270, which is still preferable to conceding immediately and obtaining 260 points. Thus, the Nash equilibrium is for both players to wait until the last second, $t_i = 30$, to concede. Our first prediction is a simple application of this result, provided subjects are risk neutral. However, if subjects

[11] One subject voluntarily left the experiment during the last rounds of Part 1, which also forced us to involuntarily dismiss additional subjects. We use data from this session in the following results. The number of subjects is reduced to 66 in cases where the excluded subjects did not participate. The basic results do not change if these data are excluded.

are sufficiently risk averse, then they will prefer conceding immediately and receiving a payoff of 260 points with certainty to waiting until the end of the game and receiving a risky outcome with an expected value of 270 points. In this case, we would expect immediate concessions. It then follows that if there is heterogeneity in subjects' risk preferences, we would expect to see a mixture of concession times at the extremes of 0 and 30.[12]

Hypothesis 3.1. *In the Staring Contest Game without an observer, risk neutral players will wait until the last moment before conceding while sufficiently risk averse players will concede immediately. A mixture of risk neutral and risk averse players will produce waiting times at the extremes of 0 and 30 seconds.*

Now consider the game when an observer is present. In the spirit of backward induction, we formulate expectations about what a payoff-maximizing observer would do and then discuss how contestants would change their behavior in anticipation of the observer's response. Because the observer moves last and takes an action that has no effect on her own payoff, she will be indifferent between allocations of punishment along the path of play. In other words, any allocation of punishment can be supported in equilibrium. However, the observer is not indifferent between all possible outcomes and can design a punishment strategy that, if rationally anticipated by the contestants, induces the best possible Staring Contest outcome for herself.

The level of blame used by an observer is endogenous and accounts for expressions of partisan bias. The extent to which voters strictly side with one party or another, however, is unclear since political ideology and party identification are not synonymous for the modern U.S. electorate (Treier and Hillygus 2009). The gap between ideology and identification is naturally embodied in our experiment: Observers can allocate blame

[12]Modeling the Staring Contest as a game of incomplete information would not substantively change these hypotheses. Consider two players with unknown types, defined as their level of risk acceptance. If player A believes her opponent to be more risk-seeking, and thus that Player B would be willing to out-wait her, then it is strictly better for A to concede as soon as possible since waiting is costly and waiting until $t = 30$ results in a toss-up. Conversely, if Player A believes Player B to be less risk-seeking, then at every time t Player A's incentive is to continue waiting until $t = t + 1$ because she expects Player B will concede at t, continuing until $t = 30$. (The argument for Player B's choices are symmetric.) Thus, as with the above model, we expect concessions to occur at the beginning or the end of the Staring Contest.

based on ideological distance between the outcome and their preference, party loyalty to Player A regardless of the strength of affiliation, or some combination of both.

When the observer's ideal point is $\theta_O = 50$, she is indifferent between the outcomes $a = 10$ and $b = 90$ and thus cares only about minimizing the duration of the Staring Contest. That is, she does not care which contestant wins as long as $t_A = 0$ or $t_B = 0$. The observer can force the contestants to end the game immediately by deducting at least 80 points from the winning player. To see how this encourages the desired behavior, note that deducting 80 points from the winner implies that the winner's payoff will be $260 - t$. Note also that the loser's payoff is $260 - t$, so the observer's strategy effectively makes the contestants indifferent between winning and losing.[13] If contestants rationally anticipate the observer's strategy, their best response is to end the game at $t = 0$, yielding the maximum payoff of 260 points.

In rounds where the observer's ideal point is either $\theta_O = 25$ or $\theta_O = 0$, her ideal outcome is for Player B to concede at $t = 0$ and for Player A to win the contest, which yields the outcome $a = 10$. The observer can therefore design a strategy that encourages Player B to concede at $t_B = 0$ by deducting at least 80 points from B if and only if B wins the Staring Contest. To illustrate how this strategy works, consider any outcome where A is the first to concede at $t_A > 0$. If the observer deducts the maximum 100 points from Player B, then B's payoff would be $240 - t_A$. By deviating to some other time $t_B < t_A$, Player B increases his payoff to $260 - t_B$ and can maximize his payoff by stopping immediately at $t_B = 0$. To see that the stopping times $t_B = 0$ and $t_A > 0$ constitute an equilibrium given the observer's punishment strategy, note that this yields payoffs of 340 for Player A and 260 for Player B. Player A will not deviate to $t_A = 0$ and risk obtaining the lower payoff. Similarly, Player B will not deviate to any $t_B > t_A$ because doing so would incur the observer's punishment and yield a lower payoff of $240 - t_B$.

The following hypotheses summarize the effects of introducing an audience on contestants' behavior and the ways in which we expect payoff-maximizing observers to play. In terms of standoff outcomes, the effect of the audience should be to decrease the observed

[13]For this strategy to work, the observer must also punish the winning player even if $t = 0$ because otherwise each contestant has an incentive to deviate to $t_i > 0$ given that the other contestant $j \neq i$ concedes immediately at $t_j = 0$.

and intended stopping times (Hypothesis 3.2). When the observer's preferences align with Player A's, we also expect to see standoff behavior and outcomes in which Player A wins the Staring Contest more often than Player B (Hypothesis 3.3). In terms of blame and punishment, observers will generally punish the winning player in order to create incentives that minimize delay in standoffs (Hypothesis 3.4). Finally, observers will direct punishment towards Player B when she prefers Player A win the Staring Contest (Hypothesis 3.5).

Hypothesis 3.2. *Waiting times will be shorter and the Staring Contest is more likely to end immediately with an observer than without an observer.*

Hypothesis 3.3. *As the distance between the observer's ideal point and Player A's ideal point decreases, Player A's waiting time will increase, Player B's waiting time will decrease, and Player A is more likely to win the Staring Contest.*

Hypothesis 3.4. *When* $\theta_O = 50$*, the observer is more likely to punish the winning player than the losing player.*

Hypothesis 3.5. *When* $\theta_O = 25$ *or* $\theta_O = 0$*, the observer is more likely to punish Player B for winning than Player A.*

3.4 FINDINGS

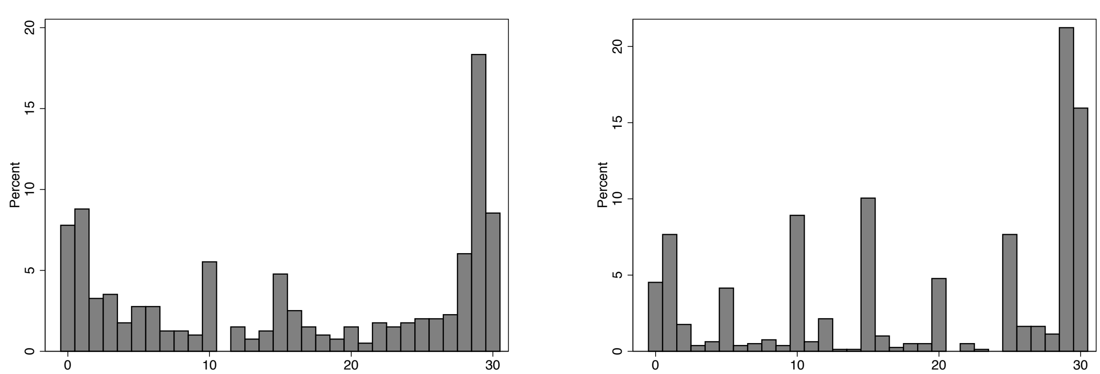
The results of the experiment support several of our hypotheses. We find support for the two major aspects of the theoretical predictions regarding the outcomes of the Staring Contest and the amount of punishment levied by the observer. In our analysis, we also classify the types of strategies observers use and assess how behavior correlates with individual characteristics.

3.4.1 Outcomes

In the baseline Staring Contest without an observer, 94% of plays of the game result in a concession. Overall, the outcomes are close to evenly split between Player A winning

(50%) and Player B winning (44%), but with a slight advantage for Player A in terms of outcomes. Figure 1 shows the distribution of actual and intended concession times in the baseline Staring Contest without an observer. Actual stopping times follow a bimodal distribution, with the majority of contests ending around $t = 0$ and $t = 30$. The game ends immediately (at $t = 0$) in 8% of rounds played and at the last possible second (at $t = 30$) in 9% of rounds played. If we allow for some error in waiting times, we find that 25% of rounds played end within the first 5 seconds while 35% end within the last 5 seconds. Intended waiting times are roughly similar. Six percent of subjects indicated they intended to concede immediately (at $t = 0$), while 16% indicated they intended to wait until the very end (at $t = 30$). Further, 61% of all intended stopping times fall within the first five seconds (19%) or the last five seconds (42%). Thus, without an observer, the majority of games ended—or subjects stated their willingness to end the game—at the very beginning or end of the waiting period, generally consistent with the prediction stated in Hypothesis 3.2.[14]

Figure 1: **Actual Stopping Times Bimodal, Intended Times Cluster**



(a) **Actual Stopping Times** (b) **Intended Stopping Times**

Turning now to the effects of introducing the observer on Staring Contest outcomes, we find mixed support for our hypotheses. Table 2 presents the average actual and

[14]Subjects do exhibit some learning effects as they play more rounds. Specifically, they are more likely to concede near $t = 30$ the more rounds they play. However, we believe this has no substantive effect on our findings. Details can be found in Appendix B.1.

Table 2: **Average Actual and Intended Stopping Times by Treatment**

		Actual	**Intended Player A**	**Intended Player B**
No Observer	Mean	16.4	20	18
	N	398	426	426
Observer 50	Mean	15.7	19.7	19
	N	128	132	132
Observer 25	Mean	14.2	19.6	17.5
	N	122	132	132
Observer 0	Mean	14.8	18.7	18.2
	N	125	132	132
Total	Mean	15.7	19.7	18.1
	N	773	822	822

intended stopping times by observer condition. Although there is a very slight decrease in actual waiting times, the average waiting times are nearly identical across treatments and between the contestants. These results suggest that introducing an observer has little effect on the duration of the standoff. In every condition, we also note that the intended stopping times are always longer than the actual waiting times, and Player A's intended stopping times are longer than Player B's. For a more rigorous test of whether the observer affects standoff durations, we estimate regressions of stopping times on a set of treatment variables with subject fixed effects. The results, presented in Table 3, indeed suggest that the observer has no statistically significant effect on average waiting times. The exception is a statistically significant decrease in Player A's intended waiting time when the observer's ideal point is 0, but this effect is the opposite what the theoretical analysis predicts for Player A. Our analysis of average waiting times lends no support to Hypothesis 3.2 or Hypothesis 3.3.

Table 3: **Effects of Treatment on Contestant Choice of Stopping Times**

	(1) **Actual**	(2) **Intended Player A**	(3) **Intended Player B**
Observer 0	-1.62	-1.63*	-0.15
	(1.22)	(0.76)	(0.82)
Observer 25	-2.14	-0.73	-0.86
	(1.24)	(0.76)	(0.82)
Observer 50	-0.73	-0.65	0.60
	(1.21)	(0.76)	(0.82)
Constant	16.39***	20.15***	18.15***
	(0.60)	(0.37)	(0.40)
R^2	0.01	0.01	0.00
Observations	773	822	822

Standard errors in parentheses.

No Observer treatment omitted as reference category.

Models (2) and (3) use subject fixed effects.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

When we look at the distribution of waiting times instead of the average in the observer conditions, shown in Figure 2, there do appear to be differences between the conditions in terms of the proportion of early and late concessions. The number of players conceding immediately (or very early) is higher under all observer treatments than under the no observer treatment. In rounds when there is an observer, the proportion of rounds with immediate concessions ($t = 0$) more than doubles, increasing from 8% to $18 - 22\%$. There is also an increase in terms of the proportion of rounds that end early (within the first 5 seconds) from 25% in the baseline to $32 - 37\%$ with an observer.[15] Table 4 shows that these effects are statistically significant. The first column presents the estimates for a probit model in which the dependent variable is an indicator for whether the game ends immediately at $t = 0$; the probability the game ends immediately is significantly higher for each of the observer conditions. The dependent variable for the second column is an indicator for whether the game ends early (within the first 5 seconds); while the coefficients are all positive, the effect of the observer when $\theta_O = 25$ is significant at the 0.05 level, the effect is significant at the 0.10 level for $\theta = 0$. The dependent variable for the third column is an indicator for whether the game ended late (the last five seconds); there are no significant difference from the No Observer treatment. The dependent variable for the final model is an indicator for whether the game stopped at the last possible moment ($t = 30$). Note that the coefficients in column four are negative, and the effect of the observer when $\theta_O = 50$ is significant at the 0.05 level, indicating that, under this treatment, the game is less likely to end at the last second. In contrast to the analysis of average waiting times, our analysis of immediate concessions suggests some support for Hypothesis 3.2.

Figure 3 shows the distribution of outcomes from the Staring Contest by observer condition. The most notable result is that the presence of an observer appears to have little, if any, influence on the outcomes of the contest. There appear to be slight differences in the proportion of rounds that end in breakdown, but since there is no clear pattern, these may simply be noise. In terms of whether Player A or Player B wins the contest

[15] Conducting Kolmogorov-Smirnov equality of distribution tests shows that the distribution under the No Observer treatment is statistically significantly different from each of the observer treatments ($p < 0.01$ for each treatment).

Figure 2: **Distribution of Actual Stopping Times**

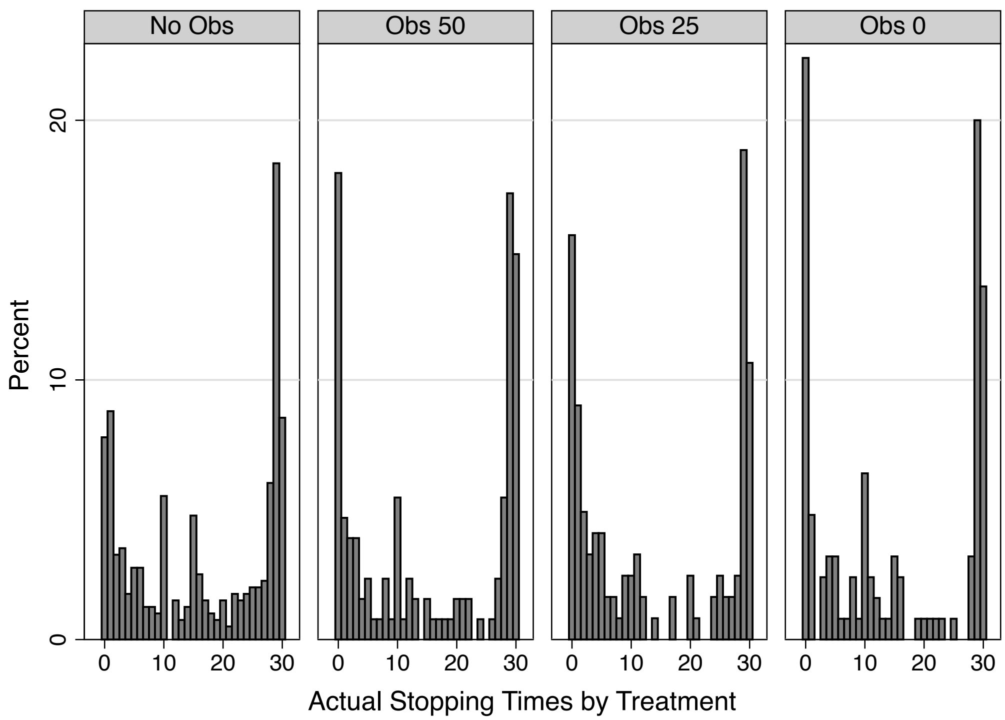


Table 4: **Effects of Treatment on Likelihood of an Extreme Stopping Time**

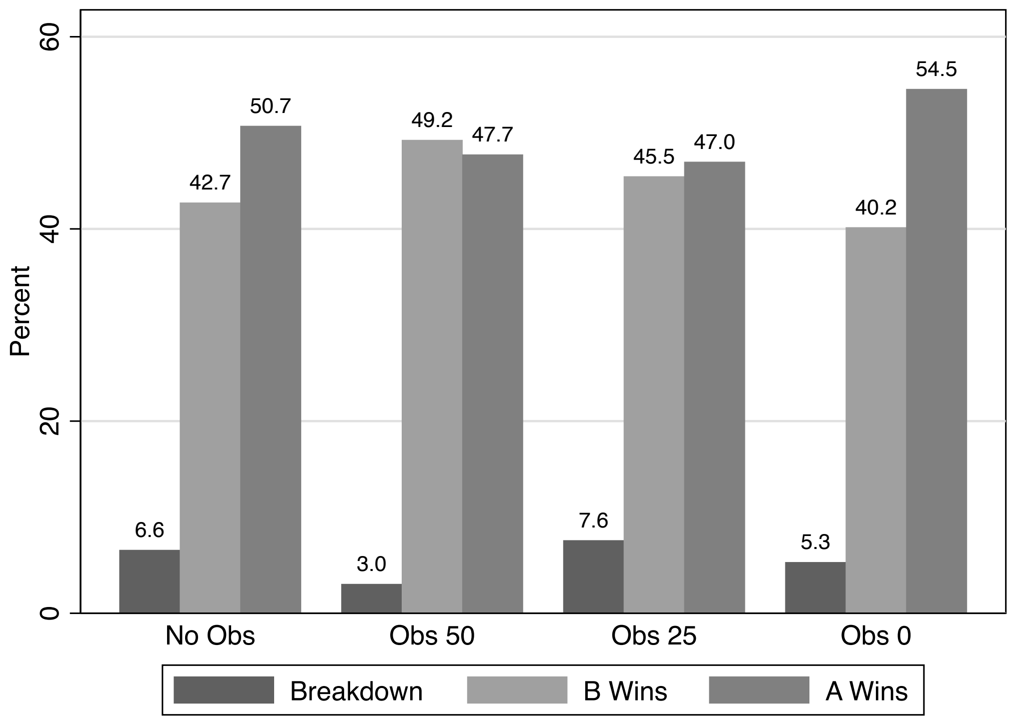
	(1) **Immediately**	(2) **First 5 Seconds**	(3) **Last 5 Seconds**	(4) **Last Second**
Observer 50	0.50^{**}	0.20	0.09	-0.33^{*}
	(0.16)	(0.13)	(0.13)	(0.16)
Observer 25	0.41^{*}	0.34^{*}	-0.05	-0.12
	(0.17)	(0.13)	(0.13)	(0.18)
Observer 0	0.66^{***}	0.23^{+}	-0.01	-0.27
	(0.16)	(0.13)	(0.13)	(0.17)
Constant	-1.42^{***}	-0.67^{***}	-0.33^{***}	1.37^{***}
	(0.09)	(0.07)	(0.06)	(0.09)
Observations	773	773	773	773
Pseudo R^2	0.04	0.01	0.00	0.01

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

with an observer, it does appear that Player A is less likely to concede as the alignment between the observer and Player A increases. Furthermore, Player A wins the contest more frequently when the observer's ideal point is $\theta_O = 0$ than in any other conditions. However, these differences are not statistically significant, so there is little support for Hypothesis 3.3.

Figure 3: **Distribution of Outcomes by Treatment**



To summarize, we find modest effects of the presence of an observer on the duration of the standoff. The basic outcomes of the staring contest (who wins or loses) appear to be insensitive to the addition of an observer. However, the distribution of stopping times changes: when an observer is present, subjects are more likely to end the game immediately.

3.4.2 Punishment

Broadly speaking, we find that observers allocate punishment in ways consistent with our theoretical expectations. They appear to punish a contestant for winning the contest more than they punish a contestant for losing, and they also appear to punish Player B for winning more often than they punish Player A. We also find evidence that observers punish both contestants heavily for breakdown.[16]

In Table 5 we present the average number of points the observer deducts from each contestant by outcome and treatment condition. The pattern of behavior appears to lend support to both Hypothesis 3.4 and Hypothesis 3.5. First, we see that observers punish the winning player more than they punish the losing player. When Player A wins, the observer deducts an average of $4.4 - 7.3$ points from A and $15.4 - 27$ points from B. Similarly, when Player B wins the Staring Contest, the observer deducts an average of $6.4 - 22.5$ points from A and $23 - 40$ points from B. These patterns are consistent with Hypothesis 3.4. Second, the deductions from Player B appear to be much more sensitive to the outcome than deductions from Player A. For example, when the observer's ideal point is $\theta_O = 0$, the observer deducts nearly 10 times more points from Player B when B wins as when A wins. In contrast, the observer deducts about the same number of points from A for both outcomes. Table 5 also suggests that Player B's punishment for winning increases as the alignment between the observer and Player A increases, while Player A's punishment for winning decreases. These results appear to be consistent with Hypothesis 3.5.

Another pattern we observe in the data is that observers tend to punish both contestants heavily in the (relatively rare) case of breakdown. When neither player concedes in the Staring Contest, the observer deducts between 36.4 and 47.5 points from Player A and between 37.5 and 43.6 points from Player B. These deductions are roughly equal for both players and do not appear to depend on the alignment of interest between the observer and Player A. We did not anticipate this result in our theoretical analysis because the

[16] Subjects do not exhibit any learning effects due to punishment. The allocation of punishment by subjects is effectively equal before and after the first time subjects receives punishment themselves. Details can be found in the Appendix.

observer does not gain from punishing the contestants when there is breakdown; it is already the worst possible outcome and contestants should seek to avoid it.

One possible explanation for these behaviors is inequity aversion (Hatfield et al. 1978): Subjects dislike the resulting point distribution, and attempt to smooth it out by deducting points from the winning player. While our model does not account for it, caring about equality can be rational, and economic models of behavior often incorporate these preferences (Engelmann and Strobel 2004; Fehr and Schmidt 1999; Goeree and Holt 2000). This does not rule out instrumental uses of punishment such as deducting points from both sides in the event of breakdown. Rather, we suggest that deviations from our precise definition of rational play are themselves behaviorally coherent and interesting.

Table 5: **Average Deductions by Contest, Outcomes, and Treatment**

		Observer Ideal Point		
Contest Outcome		Obs 0	Obs 25	Obs 50
A Wins	Punishment to A	15.4	21.7	27
	Punishment to B	4.4	4.4	7.3
B Wins	Punishment to A	17.5	22.5	6.4
	Punishment to B	40	32.8	23
Breakdown	Punishment to A	36.4	47.5	37.5
	Punishment to B	43.6	43.5	37.5

Table 6 presents regression analyses of the observer's punishment behavior. In these models, we regress a measure of punishment on the actual stopping time, a set of dummy variables for the standoff outcomes, and a set of dummy variables for the observer's ideal point. The excluded category is the case where A wins and the observer's ideal point is $\theta_O = 50$. We used subject-level fixed effects to account for potential heterogeneity in observers' average deductions. The dependent variable for the first column is the total amount of punishment used by the observer. We find that total punishment is increasing in the duration of the Staring Contest, which is consistent with the notion that observers seek to encourage contestants to end the game early. More substantial differences in punishment behavior, consistent with the averages reported in Table 5, arise from differences in contest outcomes. Observers use the least amount of punishment when A wins the Staring Contest, use more total punishment when B wins, and use the most punishment when there is breakdown. We also find that observers use more punishment when $\theta_O = 25$ than when $\theta_O = 50$ but that there is no difference in total punishment when $\theta_O = 0$.

Table 6: **Effects of Outcomes and Treatments on the Use of Punishment**

	(1) **Total Blame**	(2) **Blame to A**	(3) **Blame to B**	(4) **% of All Blame to A**
Actual Stopping Time	0.36**	0.04	0.32*	-0.00
	(0.13)	(0.12)	(0.12)	(0.00)
B Wins	17.07***	-7.37**	24.45***	-0.48***
	(3.04)	(2.79)	(2.80)	(0.05)
Breakdown	64.08***	25.82***	38.26***	-0.22*
	(7.21)	(6.62)	(6.62)	(0.11)
Obs 0	5.23	-1.07	6.30*	0.03
	(3.40)	(3.12)	(3.12)	(0.06)
Obs 25	9.12**	5.42	3.70	0.11*
	(3.40)	(3.12)	(3.13)	(0.06)
Constant	17.07***	19.39***	-2.32	0.79***
	(3.48)	(3.20)	(3.20)	(0.06)
R^2	0.13	0.05	0.20	0.38
R^2_w	0.29	0.10	0.26	0.40
R^2_b	0.00	0.01	0.06	0.31
N	396	396	396	223
N_g	66	66	66	57

Standard errors in parentheses.
Observer Ideal 50 and A Wins omitted as reference categories.
All models use subject-level fixed effects.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In the second and third columns of Table 6, we examine the effects of standoff outcomes and observer treatments on the points deducted from each player separately. Interestingly, we see that the observer deducts significantly fewer points from A when A loses than when

A wins. This appears to be consistent with Hypothesis 3.4. However, we do not see any significant treatment effects of varying the observer's ideal point in column 2. Inequity aversion may play a role in this result: the observer may value a roughly equal point distribution for Players A and B more highly than she values any particular outcome of the contest. Different weightings on equity versus strictly rational play (as we define it) does not disprove our hypotheses — we cannot parse the competing theories given our data. In column 3, we see that B is punished significantly more for winning than losing, which is also consistent with Hypothesis 3.4. And in this case, we also find that the amount of punishment is slightly greater when $\theta_O = 0$, which provides some support for Hypothesis 3.5.

The dependent variable in the fourth column is the percentage of total punishment used in the round applied to Player A. This captures the relative amount of blame assigned to Player A whenever punishment is used. These results are consistent with the results in columns 2 and 3. The intercept is 0.79, which means that Player A receives most of the punishment when A wins in the symmetric condition $\theta_O = 50$. But when Player B wins, the negative and statistically significant coefficient implies that punishment shifts from A to B—that is, to the winning player. Further, the coefficient for breakdown implies that the punishment is approximately equal. There is some evidence of an effect of varying the observer's ideal point, as the proportion of punishment applied to Player A increases when $\theta_O = 25$ relative to the symmetric case where $\theta_O = 50$, but it is in the wrong direction predicted by the theoretical analysis. Overall, the regression analysis implies support for Hypothesis 3.4. These choices in punishment allocation are, however, also consistent with a subject attempting to produce a more equitable final distribution of points.

3.4.3 Classification of Punishment Strategies

In this section, we provide additional insight into punishment behavior by classifying whether the observed allocations of blame are consistent with theoretical predictions or with other empirical regularities. We also classify subjects using these same categories by identifying subjects' modal punishment strategies. We learn that, while adherence to strict rationality (in terms of the magnitude of punishment) is low, most punishment follows the theoretical argument to punish the Staring Contest winner. This analysis provides further qualified support for our hypotheses concerning blame.

Following our theoretical analysis, we classify punishment as generally rational if the observer deducts points from the winner. More specifically, we say that punishment is *strongly rational* if it creates the strongest incentives for the game to end at $t = 0$: when the observer deducts 80 points or more from B when B wins (in any observer condition) or from A when A wins (but only if $\theta_O = 50$) and deducts 0 points from the loser. Weakening this requirement, we say that punishment is *weakly rational* if it targets the contestant according to the alignment of the observer's interest but less than the full magnitude required to induce a difference in contestants' behavior (i.e., less than 80 points). An even weaker version of this is for observers to *punish the winner* more than the loser under any condition (regardless of θ_0). These categories of rationality are also coded according to observer ideal point; all results can be read as pertaining to rational play that accounts for each particular value of θ_O.

We use two additional categories to describe allocations that also appear frequently in our data. Punishment is *equal* if the observer deducts an equal number of points from both players. We also classify rounds where *no punishment* is used at all.

We coded each instance of punishment with an indicator for the category that identified the allocation in that round. Next, we categorized each subject by finding their modal category. (Note that codings are not mutually exclusive at the level of observation since, for example, rational punishment is a subset of weakly rational punishment, nor are they mutually exclusive at the subject level in the case of ties.) This gives us a distribution of categories of play across rounds, and a distribution of general types of play across subjects. Table 7 presents these results.

We first note that subjects do not shy away from punishing the contestants. Non-zero punishment is used in over half of all periods. When punishment is non-zero, punishing the winner appears to be the prevailing strategy.[17] Although a small percentage of

[17] One indicative comment from subjects in the post-experiment questionnaire explicitly noted "I would deduct points from the player that did not concede. I was hoping this would make players more likely to concede in future rounds when I would be playing against them." Although this comment might

Table 7: **Distribution of Play and Subjects by Category**

Behavior	Percent of Rounds ($N = 375$)	Percent of Subjects ($N = 66$)
Rational Punishment	11%	2%
Weakly Rational Punishment	26%	6%
Punish the Winner	37%	32%
Equal Punishment	9%	9%
No Punishment	46%	52%

observations fit the requirements of strongly rational play (11%), we do find that a quarter of observations can be classified as weakly rational (26%) and that even more rounds are consistent with punishing the winner (37%). When we condition on the use of non-zero punishment, 21% of such observations can be classified as strongly rational, nearly half (48%) count as weakly rational, and more than two-thirds (69%) are consistent with punishing the winner. We observe very few rounds with equal punishment (9%), while the observer uses no punishment whatsoever in slightly less than half of the rounds (46%). The subject-level classifications are generally consistent with the observation-level findings except that few subjects consistently use strongly rational punishment. This does not, indeed cannot, include explicit attempts at leveling the final point distribution in our definition of rationality. Overall, our classification analysis finds a high level of punishment targeted at the contest winner, which constitutes further qualitative support for our hypotheses.

suggest repeated play considerations, it is also consistent with subjects' understanding of learning and experience—that behavior adjusts to incentives over time. In contrast, the comments of subjects who did not use any punishment suggested they did so expecting reciprocity, for example: "I never deducted anyone's points, didn't want it to happen to me[.]"

3.4.4 Individual Differences in Behavior

At the end of the experiment, subjects provided their gender, ideology, party affiliation, and a description of their approach (if any) to playing the game.[18] We can use these data to assess whether behavior in the Staring Contest varies with individual-level characteristics, and indeed, we find that differences are associated with gender and ideology. In general, we find evidence that women used more intermediate values for intended stopping times, that women use less punishment relative to men, and that being more conservative is associated with greater allocation of blame in general.

Recall that subjects entered an intended waiting time that had a small chance (1 in 10) of being implemented as their actual stopping time. These times are a form of commitment to conceding by a certain time in a non-trivial number of rounds, and thus are a significant strategic consideration for subjects. Plotting the distribution of these times by gender demonstrates a distinct difference in the use of stopping times. Figure 4 shows women enter extreme stopping times less frequently than men, providing times most often at intervals of five seconds.[19] The most striking difference is that men enter $t = 30$ for the intended stopping time over twice as often, and enter $t = 29$ almost as frequently as they enter $t = 30$. Note also that times within 5 seconds of each extreme (that is, $t \in [0, 5)$ and $t \in (25, 30]$) make up a larger proportion of the distribution for women than for men.[20]

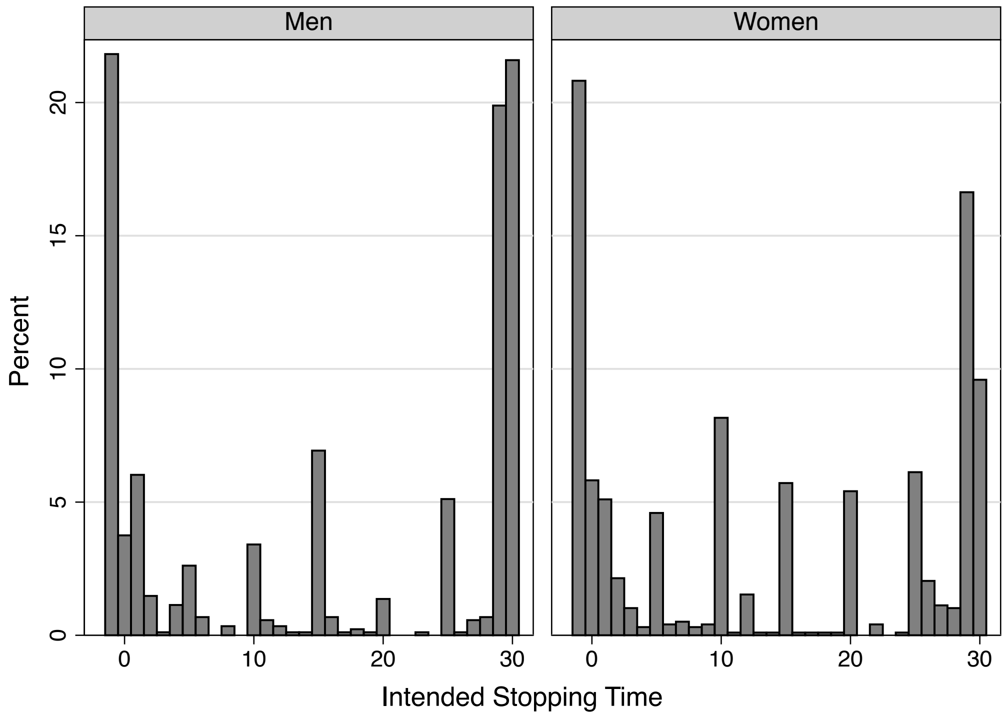
Women exhibit different strategies for play than men, allocating less punishment overall and deducting points in a more equitable manner. First, we show the gender-based differences in the use of punishment. Table 8 summarizes the deductions made by men and women, given the outcome of the contest and the observer ideal points. In line with Hypothesis 3.5, both men and women tend to punish B more for winning than A when the observer's ideal is 0 or 25. However, except for A winning when the observer's ideal is

[18]Subjects completed questionnaires via computer. They comprised a mixture of item selection and free-entry text fields. Note that six subjects did not complete the experiment and therefore did not fill out the post-experiment questionnaire. Thus, there are 66 subjects for this section of the analysis.

[19]A Kolmogorov-Smirnov test rejects (at the $p < 0.001$ level) the null hypothesis that these distributions are equal.

[20]Regression analysis supports the presence of a gender-based difference. However, the negative and significant coefficient on the gender variable only shows that women enter lower intended stopping times on average, a fact already obvious in the graph.

Figure 4: **Distribution of Intended Stopping Times, by Gender**



25, women deduct considerably fewer points than men in the same conditions. Further, when the observer's ideal point is 0 or 50, men punish breakdown severely, while women punish breakdown either moderately (for $\theta = 0$) or not at all (for $\theta = 50$).

Women appear more even-handed in their allocation of blame than men. Table 9 reprises Table 7 by categorizing play at the subject-level, disaggregated by gender. None of the women in our sample use strongly or weakly rational punishment, while fewer women punish the winner. Conversely, none of the men use equal punishment and women are more likely than men to refrain from using any punishment at all.[21] These findings support recent findings that men tend to play more (economically) optimal, or competitive,

[21] Some subjects might consider multiple rounds of the contest as a repeated game. If so, it is possible that either evenly distributing punishment or deducting no points at all could be a rational choice. Alternatively, the unwillingness to use punishment may stem from altruism, reciprocity, or other-regarding concerns.

Table 8: **Average Deductions by Outcomes, Treatment, and Gender**

		Observer Ideal Point		
Contest Outcome		0	25	50
A Wins	Women	18.5	30.9	23.7
	Men	21.3	20.7	44.7
B Wins	Women	48.3	37.5	30.1
	Men	67	70.9	28.5
Breakdown	Women	53.3	94.3	0
	Men	100	83.3	100

strategies in games of shared value and resource allocation than women, who are more likely to adopt egalitarian, or cooperative, strategies (Croson and Buchan 1999; Kennelly and Fantino 2007; Van den Assem, Van Dolder, and Thaler 2012). Furthermore, our results extend the reach of such findings to explicitly political scenarios.

Table 9: **Distribution of Rational Play by Gender**

Behavior	Men ($N = 32$)	Women ($N = 34$)
Rational Punishment	3%	0%
Weakly Rational Punishment	13%	0%
Punish the Winner	41%	24%
Equal Punishment	0%	18%
No Punishment	41%	62%

Subject ideology also appears to correlate with differences in the use of punishment. Table 10 shows that more conservative subjects punish more heavily. The distribution of ideological types is less even than for gender, with 23 liberal subjects, 29 moderate, and only 14 conservative, making the estimates for conservatives noisier. Differences in punishment are still notable. Biased observers increase their use of punishment as they become more ideologically conservative. When the observer's ideal point is either 0 or 25, punishment always increases under the same outcome: moderates punish more than

Table 10: **Average Deductions By Outcomes, Treatment, and Ideology**

		Observer Ideal Point		
Contest Outcome		0	25	50
A Wins	Liberal	13.7	20.2	34
	Moderate	23.6	25.2	37.9
	Conservative	24.4	37.2	28.3
B Wins	Liberal	40.3	39.3	30
	Moderate	62.8	56	32.2
	Conservative	68.9	81.2	21.5
Breakdown	Liberal	30	100	66.7
	Moderate	100	87.5	100
	Conservative	100	80	—

liberals, conservatives punish more than moderates. However, for conservative, neutral observers, the trend reverses: they use less punishment than even liberals under the same outcome. In line with rational play, Player B is punished more severely for winning than Player A. The only deviation is the conservative, neutral observer who punishes Player B less for winning than in any other treatment or outcome. Finally, breakdown is punished most severely by all observer ideologies, with the exception of liberal subjects when $\theta = 0$.

Examining rational play by ideology further supports our finding that subjects rarely engage in *strictly* rational play, but do frequently apportion most blame to the Staring Contest winner. Table 11 presents these results. Only liberal subjects can be classified as using strictly rational punishment, and then only rarely (4%). Liberals are also most likely to employ a weakly rational strategy (9%). However, conservative subjects are the most likely to punish the winner, almost doubling the frequency for liberals (43% vs 22%, respectively). Also, conservative observers are the most likely of the ideological types to dole out equal punishment to all players (14%), at half-again the rate of liberals (9%), and double that of moderates (7%).

Individual differences in play and punishment exist between men and women, and among ideological types. Women and men conceded at different rates, and responded

Table 11: **Distribution of Rational Play by Ideology**

Behavior	Liberals ($N = 23$)	Moderates ($N = 29$)	Conservatives ($N = 14$)
Rational Punishment	4%	0%	0%
Weakly Rational Punishment	9%	3%	7%
Punish the Winner	22%	34%	43%
Equal Punishment	9%	7%	14%
No Punishment	65%	41%	50%

differently when playing the same role; women concede roughly evenly as Player A or B, whereas men concede less as Player A. The presence of this effect is independent of subjects' strategic considerations, in contrast to work on gender in bargaining that generally makes gender known to all parties (Cadsby, Servátka, and Song 2010; Holm 2005; Putnam and Jones 1982; Sutter et al. 2009). Further research on this effect would contribute to understanding differences that arise for intrinsic reasons rather than from strategic responses to gender. Ideology also matters in this study, with growing conservatism correlating with both greater use of punishment in general, and an unwillingness to concede as Player A. Similar to gender, ideology in bargaining experiments is usually modeled as common knowledge (Banas and Parks 2002; Wade-Benzoni et al. 2002) but is important to understand as a personal trait affecting evaluations of politics.[22]

3.5 DISCUSSION

In this paper, we examined how blame could be used as a mechanism for voters to control their representatives, focusing on how observers apportion blame and how the anticipation of blame affects political outcomes. We found that, in our experimental setting, observers use blame as punishment for delaying the resolution of a political standoff. Overall levels

[22] One interesting exception is Knight and Ensminger (1998), which considers personal ideology in the development of social norms, rather than through mutual evaluation of a bargaining partner's ideology.

of punishment increase when observers' interests are more closely aligned with one side of the fight, though no one is entirely spared and both sides are heavily punished for breakdowns. Thus, our findings suggest that interests alone do not determine the target of blame, cutting against the conventional wisdom that observers rarely blame copartisans for bad outcomes, as this behavior is broadly consistent with a rational strategy observers can use to induce better outcomes for themselves.

We also found that while blame does not significantly alter Staring Contest *outcomes*, it does induce representatives to reduce the *duration* of standoffs. The rational anticipation of blame drives players to terminate contests earlier, thereby reducing waiting costs for all players. Shorter delays can be seen as a "bipartisan" improvement in our experiment as well as in political reality. But breakdown still occurs, of course. Our results suggest breakdown also arises when political elites fail to fully anticipate the consequences of protracted disagreements.

While the experimental evidence generally supports our hypotheses, inequity aversion may help explain deviations. Punishment is widely used, but is only mildly affected by observer ideal point treatments; subjects may care more about the final point allocation than their induced affiliation with Player A. Furthermore, "Equal Punishment" is a rarely used strategy for players of any gender or ideology. Deducting the same number of points from both sides does not satisfy a desire for equality of outcomes. A different structure for the Staring Contest would be necessary to fully tease out the prevalence of these incentives (and to separate them from risk aversion (Carlsson, Daruvala, and Johansson-Stenman 2005). The gap between our predictions of rational play and actual subject choices cannot be read simply as the inability to predict outcomes based on our model. Rather, we take these findings to suggest that subjects demonstrate behavioral regularities that could be, but in the current form are not, reflected in our model of political standoffs.

Extensions to our work would move beyond our focus on how voter blame impacts negotiating behavior during standoffs that are already underway. A single-shot game highlighted how an audience to the standoff decides to allot punishment at a single point in time. Of course, politics is frequently a repeated game, and expanding our experiment may add to the substantive findings reported here. Clearly, some subjects already considered the potential impact of punishment in one round on their outcomes in future rounds. In addition, pre-standoff communication might introduce interesting incentives, asking the observer to consider the possible avoidance of the standoff when allocating blame. This paper serves as a foundation for examining the important, and recurring, interaction between elected representatives and the expressed desire of voters during high-stakes political brinkmanship In 2011, Standard & Poor's downgraded the U.S. Treasury partly because Congress' partisan showdowns repeatedly threatened default on loan service payments.[23] Prompt resolution of subsequent debt limit battles convinced another ratings firm to return the U.S. credit rating to full AAA status.[24] Similarly, in 1995, Speaker Newt Gingrich stood firm against President Bill Clinton over the federal budget, bolstered by the belief that the public would blame the president for a shutdown (Drew 1997). Gingrich held out, and the government shut down for 27 days. While media labeled Gingrich the winner, the public clearly blamed Republicans more than Democrats.[25] Speaker John Boehner overestimated public support for his stance, refused to blink, and allowed the government to shut down again. Here too, despite Boehner "winning" by (temporarily) preventing an increase in the debt ceiling, the public laid the majority of blame at his feet.[26] Following these episodes, Gingrich and Boehner seemed to learn their lesson, as later battles ended in favor of their opponents.[27]

[23]Damian Paletta and Matt Phillips, "S&P Strips U.S. of Top Credit Rating", *Wall Street Journal*, 6 August, 2011.

[24]Jim Puzzanghera, "Fitch takes U.S. credit rating off downgrade watch after debt deal", *Los Angeles Times*, 21 March, 2014.

[25]CNN/USA/Gallup Poll, "Americans blame GOP for budget mess", *CNN*, 15 November 1995.

[26]Paul Steinhauser, "CNN Poll: GOP would bear the brunt of shutdown blame", *CNN Politics*, 30 September 2013.

[27]Paul Kane, Robert Costa, Ed O'Keefe, "House passes 'clean' debt-ceiling bill, ending two-week showdown", *Washington Post*, 11 February, 2014.

4.0 INFORMATION TRANSMISSION IN CONGRESSIONAL HEARINGS

Hearings are the most public expression of Congress' oversight function. They provide ample opportunity for members of Congress to engage in all three famed Mayhew (1974) activities: conducting a single hearing places the member at the center of a news-worthy event, argues for the member's ownership of the issue, and lets the member stake out a position on the topic at hand. Those called to testify may burnish their reputation for expertise, rehabilitate a tarnished public image, or simply demonstrate their importance to the policy making process (Esterling 2004; Leyden 1995). Hearings also allow committees to demonstrate a commitment to policy specialization to the full chamber (Diermeier and Feddersen 2000). The question remains, however, as to whether the hearings are valuable as vehicles for oversight *in and of themselves*.

Hearings shape political attention to important policy issues (Edwards and Wood 1999; Robinson and Appel 1979), and consume a considerable about of legislative time. Whether or not relevant and useful policy information is conveyed during these hearings is an important facet of understanding congressional control over executive branch activities. While delegation of policy information to the bureaucracy is predicated on asymmetries of information and a desire for efficient resource allocation (Bendor, Glazer, and Hammond 2001; Epstein and O'Halloran 1999; Holmstrom 1984; Krehbiel 1991; Patty 2005), Congress cannot adequately audit policy implementation if it cannot extract accurate reporting from its agents (Gailmard 2009; Gailmard and Patty 2013). The primary result of this study shows that, while congressional hearings can enable this extraction, meaningful information exchange is far more likely to occur when both the legislative and executive representatives agree on policy. Put another way, when the interests of committee members and witnesses diverge, hearings serve primarily as political theater.

The informativeness of hearings has received less attention than its importance warrants, perhaps in part because of the difficulty of conducting a meaningful content analysis on a significant sample. Most research on the topic of congressional hearing content focuses on the "metadata" related to the hearing itself, including the policy topic, the convening committee or subcommittee, or the broad affiliation of external attendees (Degregorio 1992; Leyden 1995). One notable example of exploring informativeness finds that hearings do not have instrumental value in getting bills passed (Brasher 2006). Of course, that study only examines whether the hearing was necessary for the subsequent bill passage, and cannot address whether anything within the hearing was itself informative. After all, the dynamics of passing legislation are largely independent of how revelatory a hearing may have been. Del Sesto (1980) demonstrates the considerable effort required to perform hand-coded content analysis of hearings, but supports the value of doing so by clearly identifying how the committee acquired information through the hearing process (even if it is sometimes ignored (Bradley 1980)). Sabatier and Whiteman (1985) shows that extensive information flows between members of congress, their staffs, and committee staffers, as well as the executive branch and private organizations, but assumes that whatever is exchanged is actually useful information (*vice* being simply ceremonial interactions intended as a signal to any observers). I tackle the informativeness problem directly using automated text analysis to quantify the policy content of hearings across both chambers and every committee over multiple sessions (limited only by the public availability of transcripts). Such a broad sample of data starts filling in the gap in knowledge about the effectiveness of hearings for congressional oversight.

The rest of this paper proceeds as follows. First, I build my theoretical structure and resulting hypotheses on the foundation of Crawford and Sobel (1982) by showing that a single hearing is a useful approximation of a signaling game between a Committee Chair and an executive branch witness. The next section explains how a speaker's choice of specificity in language operationalizes the concept of message partitioning from the signaling game. I then describe my data, and the method for calculating language specificity from hearing texts. Results from a model of information transmission are presented, followed by a discussion of the importance of these results. Ultimately, this paper demonstrates that the level of information conveyed in Congressional hearings is a direct result of the ideological distance between a Committee Chair and a witness. We learn that, precisely when the legislature may wish to exert greater influence – when the two disagree about policy and Congress hopes to reorient its efforts – is when representatives of the executive branch shade their private information in an attempt to retain autonomy.

4.1 HEARINGS AS SIGNALING

The actual conduct of a hearing closely resembles the strategic situation Crawford and Sobel (1982) (henceforth, CS) analyze in their theory of information transmission.[1] I lean on their theoretical structure to identify the driving mechanisms and incentives for participants in congressional hearings. Actors, incentives, choices, and the play of the game map naturally from the theoretical to empirical setting. Of course, the comparison is not perfect, and I note the most significant discrepancies along with the reasons they do not significantly diminish the usefulness of this approach.

At least two roles appear in any signaling game – the receiver, R, and the sender, S – which are filled in this study by the committee chair of a hearing as R, and any actor from Executive Branch agencies as S. There are, of course, any number of other people who may appear in hearings: private sector business leaders, athletes, academics, or any other individual subject to Congress' subpoena power.[2] In the analysis, however, I restrict the set of senders to Executive Branch representatives for two reasons. First, the functioning of oversight pertains almost entirely to legislative control of bureaucracies. Second, data on policy preferences (ideal points) for actors outside these two roles is insufficient to use in analysis. Considering R, congressional committees comprise members of both political

[1] Analogies to non-strategic signaling games (e.g. Spence (1973)) are obvious, but entail the specification of exogenous costs that are not applicable in this setting.

[2] For clarity and consistency, I use the pronouns "she/her" when referring to chairs/receivers, and "he/him" when referring to witnesses/senders.

parties and a range of ideological concerns even within one party. Considering the chair of the committee as the central figure, however, does not remove these interests from the game. Chairs coordinate the request for, structure of, and witness attendance at the hearing in conjunction with the interested parties in the committee (Leyden 1992). During a hearing the chair guides the pace of interaction, manages time allotted to committee members and witnesses, and holds ultimate say over the hearing's duration. Because of this control over the forum, the chair is the focal consideration for any witness called to testify.

Each side of the signaling game has distinct interests, with chairs representing the committee, the chamber, and their own constituencies, and witnesses representing multi-faceted organizations, or speaking for a community that may hold diverse views. While this appears to unduly condense a large cast of characters, a theoretical model of hearings that concentrates on the two actors with single utility functions (even when left largely abstract here) is not as limiting as it may first appear. Testimony requires distilling diverse interests and messages into a unified statement delivered by a single person. Multiple witnesses in a hearing may represent similar positions, but in each case the speaker has chosen to reveal a particular amount of information during the hearing. Witnesses have every incentive to capture the position of the interest they represent, and indeed spend considerable time and effort crafting statements, since they are opportunities to lobby for a position (Leyden 1995), demonstrate competence to the audience, and guide policy choices (Esterling 2004). Each side of the game, then, can be usefully said to have an ideal point that guides their behavior during the hearing.

The two sides in the CS game are not equally informed about an issue, and neither are the actors in a hearing. A better-informed bureaucracy is a fundamental, and largely non-controversial, aspect of the delegation literature (Bendor, Glazer, and Hammond 2001; Bendor and Meirowitz 2004; Gailmard 2002; Huber and Shipan 2002; Kiewiet and M D McCubbins 1991; Volden 2002). The committee chair is, ostensibly, seeking to remedy this information asymmetry by requiring the better-informed actor to report what they know about some topic. The potential effects of a policy choice is only one possible topic, as committees routinely demand information about executive agency management choices, private industry practices, or even the behavior of a key individual. I retain this central feature by assuming that any witness delivering testimony is the relatively more knowledgeable sender, and the committee chair is the receiver who must update her beliefs about the information delivered before taking action. Crucially, the witness' statements are thus conditioned by his expectation about the chair's knowledge.

Each player in the hearing has the opportunity to make a strategic choice. The sender is the CS game moves first, and systematically introduces noise into his message by indicating only a range in which the true state of the world lay. In equilibrium, the amount of noise — "partitioning" in CS — is a direct function of the distance between the sender's and the receiver's preferences. Again, the connection to hearings is straightforward: both the committee chair and the witness have individual interests that may diverge. Bureaucratic agencies have both their own interests and the need to consider the needs of their ultimate sponsor, the President. I leverage this connection through measurements of committee chair and agency ideal points (Chen and Johnson 2015; K. Poole and Rosenthal 2007).

Divergence between ideal outcomes for a witness and a committee chairs stems from numerous sources. First, and perhaps most prominent, is ideological differences between the committee chair and the president, who is generally assumed to exert strong influence over the policy goals of an agency (Huber and Shipan 2002). This control is not perfect since agencies develop a measure of autonomy in defining and pursuing goals, though this does not necessarily narrow the gap between the legislature and the agency (Carpenter 2001, 2010). The passage of time is itself a wedge, as the interests of new generations of congress drift from an agency's original mandate (Macey 1992). Non-bureaucratic actors may not have such direct disagreements with the committee, but are certainly concerned with any legislative results of the process. That said, both the CS game and my empirical measures do not *require* positive divergence. For CS, perfect alignment coincides with perfect information transmission. My measure allows for identical ideological positions and perfect precision (though they are not, in actuality, ever seen).

After hearing the message, the receiver updates her beliefs, then makes a choice that affects utility for both actors in the game. For hearings, this choice is the adoption, or not, of new legislation that governs the life of the witness in some way.[3] Both the committee chair and the witness have preferences over the outcome. The chair may have some desire for good policy, or may simply be looking for instrumental utility obtained by position-taking (Mayhew 1974) or communicating competence to the full chamber (Diermeier and Feddersen 2000). Witnesses may have hopes of successfully arguing for a new policy (a lobbyist, perhaps), or may hope to avoid budgetary restrictions (agency leadership). Strictly within the signaling game played in the hearing, testimony is "cheap talk," but this does not mean the witness relates nonsense or lies. He does, however, attempt to send the least precise information possible, in order to retain some of the information advantage. The witness choice of precision, and conditions that drive it, is the central concern of this paper so the definition of precision is explained in detail below.

Indeed, the process of holding a hearing looks like the sequence of play in the CS game. First, S observes his private information, then chooses a message. R updates her beliefs, then takes an action. The sender moving first is both theoretically important and empirically valid: witnesses collect relevant information to craft their testimony, then deliver it, to the congressional attendees.[4] After hearing all relevant testimony, the committee decides whether and how to pursue legislative change.

Rather than focus on the technical equilibria, the relevant insight is found in the comparative static: congruence of actor interests. Only when the two sides have perfectly aligned interests should anyone expect to see fully truthful communication, otherwise the sender will choose a noisy message. To be specific, introducing some amount of noise in the message is almost always a best response for S, considering the potential action R may take. As the chair and witness preferences diverge, the amount of noise should grow, or, equivalently, the amount of precision should decrease. Further, and just as important for studying hearings, is the fact that amount of information transmitted is zero only in

[3]To use this choice as the receiver's action only requires the *possibility* of new legislation. The sender expects, correctly, that unwanted changes are possible, and conditions his signal accordingly. Indeed, the sender may begrudgingly change his behavior to better suit the receiver in order to prevent legislation the sender dislikes even more; observing altered legislation to force this change is not necessarily on the equilibrium path.

[4]There is no chance in the CS game for either side to opt out and avoid the exchange. Similarly, and because of the compelling legal force of Congress' subpoena power, I make the assumption that appearing at a hearing is a *fait comply* for the witness.

the extreme case of precisely opposite preferences. Thus the CS actors are most likely to coordinate on an equilibrium that always includes some positive amount of information. This paper operationalizes this insight by measuring the level of precision in testimony, then explores the impact of preference divergence, and other institutional factors, on that precision.[5]

The empirical implications of this result for hearings are straightforward, so I state them here as the first hypothesis I will test:

Hypothesis 4.1. *As witness and chair ideal points converge, the witness will deliver more information during a hearing.*

I develop the full operationalization of ideal points and information in the next section.

Despite other signaling game models presenting technically and substantively opposing views of how ideological difference affects information transmission, I rely here on the CS framework, and choose to avoid any sense of a "horserace". That said, this work does contribute to understanding why CS is the best starting point for investigating hearing informativeness. Specifically, work such as Gilligan and Krehbiel (1989), Krishna (2001), Battaglini (2002), or Minozzi (2011) present direct challenges to the CS finding: ideological distance is not found to matter for the amount of information transmitted. However, those models posit interactions that are rarely, if ever, present in actual congressional hearings that include only agency representatives. Agency witnesses address a range of issues on a single topic, and are not repeatedly placed in pairs (or groups) of diametrically opposed experts at opposition with each other on the same exact issue. The possibility of transmitting information in just this latter contest is a central concern of those papers. Second, the physical process of a hearing is a succession of witnesses testifying either one-by-one (or, rarely, in highly-related small groups such as the highest ranking officers of a corporation) in set order, and not, as is the case in the models noted, simultaneously. That is, equilibria suggesting that ideological distance does not decrease (or, indeed, increase) information revelation are predicated on each sender choosing a message without

[5] Diermeier and Feddersen (2000) use a model where the informativeness of a hearing is exogenously given. The CS model makes the level of information endogenous, and my measure of information corresponds to the CS model's basic features.

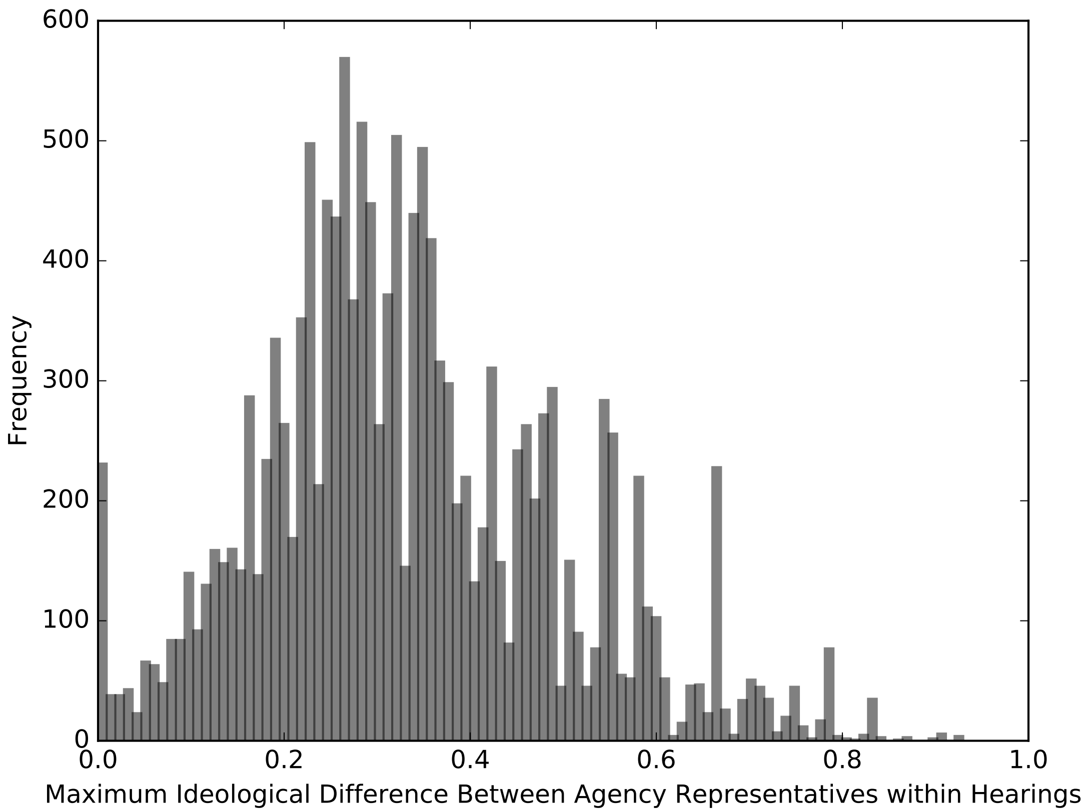
knowledge of the other's. This effective simultaneity of messaging substantially changes the sender's incentives, and with it the receiver's beliefs and ultimate choice of action, while also not reflecting the process of testifying-through-interrogation each witness undergoes in a hearing.[6]

Descriptive verisimilitude is not a useful discriminator of formal models. Instead, I use the data from actual hearings to demonstrate that within-hearing transmission goes from relatively similar agency witnesses to the committee chairs, thus supporting the use of CS versus other models. Figure 5 shows the distribution of the difference between the largest and smallest ideal points of the representatives testifying withing a *single* hearing. This can be thought of as the largest ideological range between witnesses that occured during the hearing. Across all hearings in the dataset, the difference between agency representative ideal points is quite small, predominantly below 0.4. While that distance is itself enough to separate viewpoints in a technical sense (after all, a Democrat and a Republican may be different parties while still having ideal points closer than that), this indicates that, substantively, the speakers hold kindred views (because the conservative Democrat and liberal Republican have similar policy views by definition). Of course, this does not rule out the possibility that, across multiple hearings, chairs may array opposing viewpoints to gather diverse opinions. That situation is different than the one I present in this paper, and thus does not argue against relying on the CS model.

While the similarity between the CS game and a hearing are clear, there are several limitations. First, I cannot know with certainty the true level of knowledge either side has about the policy under question. Oleszek (2004) notes that committee chairs may conduct extensive review of a policy before conducting a hearing. In this case, the hearing functions more like an audit with the chair measuring the witness' truthfulness. Even so, the witness will be aware of this effort since committee staff are engaging in background research. Thus, even if the informational asymmetry is small or non-existent, testimony is still based in part on the witness' expectation of the chair's knowledge.

[6] To anticipate a further objection: while a committee's preparation for the hearing, and thus the witness' preparation for testifying, is simultaneous, testimony is observable by all players, and thus adaptable at the time of delivery. Witnesses can, and often do, deviate from their prepared remarks.

Figure 5: **Distribution of Agency Witness Ideal Point Ranges**



Second, the final choice of the receiver is opaque. Because I concentrate here entirely on the process within the hearing, I do not include any eventual change in legislation. While Brasher (2006) look at the impact hearings have on the likelihood of new legislation being offered, my design does not require validation with this data. Like the issue with the information held by the chair, the important mechanism is the senders' expectations about the receiver's actions; the action itself is identified in equilibrium.

A second hypothesis specifically about the amount of information delivered in hearing testimony follows from informational theories of legislative organization (Krehbiel 1991). The committee structure in Congress encourages policy specialization, with members requesting committee assignments that comport with their constituent and personal interests and further developing expertise through the creation of legislation within a policy arena. The longer a member serves on the committee, the greater her topic knowledge, as compared to both the average member in the full chamber and the newer members of her committee. The important implication here is this information makes her a better auditor of any messaging about that topic delivered during a hearing. Simply put, the longer a committee member has served, the better she is at knowing if a witness is being forthcoming. Moreover, this fact is known to the witness, who develops a statement that is only as specific as it needs to be, without risking being caught in a falsehood. Stating this as a testable hypothesis:

Hypothesis 4.2. *Witnesses will deliver more informative testimony in hearings with committee chairs who have served more terms.*

In sum, congressional hearings are usefully modeled by the canonical CS signaling game, allowing me to employ the logic of the game's play and solution to investigate the informative value of testimony. Preference divergence, as a comparative static, informs my examination of testimony. Note that I do not mean to test the underlying theory of the CS game; rather I mean that my theoretical expectations are founded on the insights the CS model provides. The next step is defining how the actual content of the messages is measured.

4.2 LANGUAGE AS INFORMATION

On top of the clear correlation between the signaling model and the hearing, I develop a measure of the political heart of the exchange: the policy information being transmitted. Specifically, I quantify the *language specificity* (Li and Nenkova 2015) of all speakers in a hearing, including both congress members conducting the hearing and witnesses giving statements. First I explain how this serves as an empirical instantiation of the CS game's messaging, then I provide the technical details for arriving at the measure.

Spoken testimony and signaling can both be measured in terms of specificity. Messages in the CS game transmit information about some variable, the true value of which is known to the sender. A key result for CS is that a message, in equilibrium, will actually transmit a *range* of values, partitioning the variable's domain into multiple sets. S says, in effect, the value of the variable may be below some value, above some value, or lie in an interval. The size of that partition containing the true value (or, at least the range in which S claims the value sits) is a measure of the message's specificity.

Natural language can be understood using the same concept of specificity. Consider, as an example, the following statements from then-Representative of the State of Ohio, James Garfield made during Congress' 1873 investigation of the *Crédit Mobilier* scandal. First, Garfield makes specific claims about statements made to him, the amount of a loan, and individual contacts:

> "Nothing was ever said to me by Mr. Train or Mr. Ames to indicate or imply that the Credit Mobilier was or could be in any way connected with the legislation of Congress for the Pacific Railroad or for any other purpose. Mr. Ames never gave, nor offered to give, me any stock or other valuable thing as a gift. I once asked and obtained from him, and afterward repaid to him, a loan of $300; that amount is the only valuable thing I ever received from or delivered to him." –Garfield (1873)

Garfield's statement could be called a fine partitioning of the underlying information: he states a clear legal view, recounts the kind and number of items received from Mr. Train and Mr. Ames, and a specific amount of money. The second statement again deals with Garfield's securing of loans and his personal contacts, but is far less specific:

> "I should think it was in the session of 1868. I had been to Europe the fall before, and was in debt, and borrowed several sums of money at different times and from different persons. This loan from Mr. Ames was not at his instance. I made the request myself. I think I had asked one or two persons before him for the loan." –Garfield (1873)

This time Garfield chooses a coarser partitioning for his message: his recollection of a date is vague, notes receiving "several" loans instead of a precise number, and then notes an indeterminate number of people he had previously ask. I argue that the difference is a strategic choices made by any S when delivering testimony.

The strategic nature of vagueness is made explicit in more recent work on signaling. Blume and Board (2014) analyze a signaling game where vague messages admit multiple interpretations of the same signal. While this message-based uncertainty differs from the CS type-based uncertainty, the important point for my purpose is the same: uncertainty in messaging is reflected in language choice. Further, this uncertainty can be quantified and utilized to study strategic information transmission. Indeed, prior work uses this very method to study the level of information in arguments, a clear companion to the kind of exchange CS capture with their model (Swanson, Ecker, and Walker 2015). The next section provides details on how I measure information in hearings, as well as other relevant measures that impact the game played during hearings.

4.3 DATA AND METHODS

My primary unit of observation is a single witness in a given hearing. For each hearing, there is an associated committee that convened the hearing, and the chair of that committee. These, then, are the two players in the CS signaling game: the witness is the sender, and the chair is the receiver. The key comparative static in the CS game is the difference between the sender and the receiver's ideological positions, and the amount of information delivered by the sender (technically, how finely partitioned the message space is). Working with this theory thus requires data on the ideological positions of the two sides of the game, as well as a measure of information.

My primary source of data is the text of Congressional hearings for all committees in a session of Congress. For reasons of data availability, I use the publicly-available transcripts of Congressional hearings from the $105^{th} - 112^{th}$ Congresses (covering $1993 - 2012$).[7] Hearing transcripts include not only all recorded verbal exchanges, but also all the relevant details to identify the committee (or subcommittee, or both) convening the hearing, committee (and subcommittee) members, hearing title and date, as well as information about all present witnesses, including any organization they may represent.[8]

Data for the congressional side of the game are taken from standard sources. Ideological scores for members of Congress are the standard Common Space DW Nominate scores from Carroll et al. (2008). Committees and their Chairs are obtained from Stewart III and Woon (n.d.). I use scores from the 105^{th} to the 112^{th} Congresses, reflecting the overlap between the congresses for which full sets of hearing texts, updated ideological scores, and committee data are available.

Ideological positions for hearing witnesses are obtained from Chen and Johnson (2015).[9] I restrict the set of witnesses to only those people who are, at the time of testifying, representing an executive branch agency for two reasons. First, the theory presented here applies most directly to actors whose utility Congress can directly affect. Certainly, Congress might affect the operations and financial standing of private organizations through actions like changes in tax codes, fiscal reporting to the Securities and Exchange Commission, or other legal actions. That said, the empirical difficulty of obtaining or developing ideological preferences for private sector actors is the second reason I restrict the set of witnesses. Using ideological estimates of an organization as a substitute for personal assumes the person testifying represents the general interests, and thus ideological position, of the full agency. Multiple people usually engage in crafting the testimony, and these people are most frequently at the higher ranks of the agency, making it plausible that the message delivered is a distillation of agency views, which the ideological scores

[7]These transcripts are made available from the U.S. Government Printing Office's Federal Data System.

[8]These transcripts were obtained via automated download. The metadata concerning each hearing were obtained in two ways: 1) extracted via text processing, and 2) parsing XML-delineated data provided by the GPO. Both methods were used because of severe limitations in the XML.

[9]Clinton and Lewis (2007) offer alternative measures of agency ideology. I opt to not use these since they are not directly comparable to the DW Nominate scores I use for the committee chairs.

reflect. My key explanatory variable is thus *Ideological Distance*, which is measured as the absolute value of the difference between ideal points for chairs and witnesses. The resulting data set has a total of 130670 observations, with 81191 over 8342 hearings for the House of Representatives, and 49479 over 6757 hearings for the Senate.

The next step is to define a measure of information in the context of congressional hearings. To achieve this, I use the recorded text of statements and exchanges the occur within a hearing. For each speaker in the hearing, I extract all the words they utter, and combine it into a single set. Each speaker, then, has an associated corpus of words that reflect their strategic choice of messaging. I capture this strategic behavior by calculating a level of language specificity used in each speaker's testimony.[10] Witnesses choose how much detail their testimony contains, opting to either provide vague abstractions, or to relate precise dates, program names, dollar amounts, locations, and other facets of agency behavior. Legislators armed with precise information can exert greater pressure on witnesses that might attempt to gloss over areas in which their work does not satisfy the committee. And, of course, both sides are aware the other is making a strategic choice. Moreover, while there are usually multiple witnesses in a hearing, I consider the informational content of each speaker to be independent of the content provided by other speakers.[11]

I employ standard techniques in natural language processing and text analysis to develop my main dependent variable: *Speaker Specificity*. Language specificity is calculated for each speaker in each hearing for both chambers and all relevant committees. This calculation is the output of a supervised learning process that categorizes a single sentence as either specific or general. More precisely, the algorithm from Li and Nenkova (2015) uses a logistic regression to produce the posterior probability that the sentence would

[10]Committee chair and member speech is also collected and analyzed. However, their language specificity has no theoretical importance. Congressional speakers may appear to give vague statements either because they are uninformed about a topic, or because the witness has provided sufficient information and no detailed questioning is required. This observational equivalence makes including language specificity for Congressional speakers unnecessary.

[11]This is not the same as saying all witness testimony is entirely independent. Witnesses are coordinated by committee staff, so the subject matter of one expert is likely to depend on another, perhaps in order to avoid redundancy. My measure of information is independent of topic, however. A violation of my assumption, then, requires that one witness' choice about *specificity* is conditioned on the testimony of other witnesses.

be classified as specific or general by a human reader. The *Penn Discourse Tree Bank 2.0* serves as the reference corpus, the full details for which can be found in Prasad et al. (2008). The basic process entails trained readers coding each sentence in the reference text as either "general" or "specific," along with tagging them with lexical information. A machine learning algorithm is then run, or "trained," to find the features of the sentences that most highly correlate with the specific/general classification. A new body of text, such as a corpus congressional hearing sentences, is then surveyed for these features. Classification is carried out by running regressions run where the features are predictor variables and the dependent variable is the likelihood of being tagged specific or general. Also important to this project is that the sentences in this training corpus are gathered from a news source that aims its writing to an educated audience. I expect the level of discourse in a congressional hearing to be roughly equivalent, making this set of texts a valid basis for comparison. The posterior probabilities generated by this process are for single *sentences*, so I use the process noted in Equation 1 of Li and Nenkova (2015) to develop a specificity score for each *speaker*. This score should be read as the *word-average specificity* for that speaker, within a single hearing. In general, this scores the amount of policy-specific information relayed during the speaker's testimony (for witnesses) or questioning (for committee members).

An example from an in-sample hearing is useful at this point. The following sentences are from a House of Representatives hearing titled "Long-Term Care and Medicaid: Spiraling Costs and the Need for Reform," convened by the Subcommittee on Health of the Committee on Energy and Commerce, during the 109^{th} Congress. Both sentences are made by Dr. Mark McClellan, then Administrator of the Centers for Medicare and Medicaid Services. The first sentence is given a value of 0.996, meaning it has a greater than 99% chance of being classified as specific if given to a human reader:

> "No, when Secretary Thompson was involved in all of those discussions, and people looked at what the independent CMS actuaries, and what the CBO analysts had to say, we went for the approach that was going to get the best costs for up to date access to medications, and that is what we are implementing right now."

The second sentence, by contrast, has a value of only 0.047:

> "And there are few Medicare beneficiaries who have more to gain from the new benefit than our dual eligibles, who are often getting very fragmented care."

In contrast to the first, the second sentence is far less precise about the topic under discussion, makes no reference to particular people, does not mention any organizations, and has only a general subject. In the language of the signaling game, the second sentence is has a larger partition than the first. Of course, while these sentences are part of the full testimony Dr. McClellan gave, the reported scores for each sentence are *independent* of each other and the rest of his testimony. To rate a speaker for an entire hearing, I use a modification of the basic Li and Nenkova algorithm. The full details are given in Equation 1 from Li and Nenkova (2015), but can be summarized as follows: the measure condenses all statements made by a single speaker into a $1 \times N$ array of sentences, weights the score for sentence $i \in N$ by the length of that sentence, sums the N weighted scores, then divides this value by the total number of sentences in a speech (effectively normalizing the scores according to the amount the speaker talked).

Information in hearings is not an absolute, independent entity, so I choose to translate the specificity values into a z-score for each congressional session, making the values a within-congress comparison. The z-score transformation normalizes the level of specificity for each speaker against a general level of discourse precision in a given congress. I expect that the general makeup of any one session (including partisan control, ideological distribution, committee assignments, and any aspect of internal structure that can be chosen by Congress at the outset of each session) will, in some part, drive a general level of specificity. The z-score for the speaker thus indicates how informative the speaker was relative to all other speakers. This mitigates any effects of external conditions that might affect all hearings in some way, but that is not present in the text of the hearings themselves. For example, if members of Congress respond to low levels of public approval by becoming more aggressive in hearings, independent of topic, this would affect all hearings during the period of low approval. A by-congress normalization accounts for the greater specificity in all hearings, and focuses attention on the relative difference. I also transform the specificity variable into a z-score across the entire data set for purposes of checking robustness.

To test my hypotheses about the amount of information relayed in a hearing, I employ a simple linear regression model. Because I a directly measuring the theoretical entities – ideology and information – posited by a causal relationship, I can interpret the outcome of my econometric model as causal itself. As noted, my dependent variable, *Speaker Specificity* is a quantification of information, and is a continuous random variable taking on values between 0 and 1. The z-score transformation means the variable can take values on the entire real line, though in practice the values fall in $[-3.2, 4.06]$. The key independent variable, *Ideological Distance*, is a continuous random variable taking values between 0 (indicating perfect ideological alignment) and 2 (the greatest possible ideological difference). The second hypothesis concerns the amount of time a member of congress has served. I use the seniority measure in Stewart III and Woon (2005) to identify the member's period of service.

The data also include multiple characteristics of the hearings themselves. This includes date, hearing title, reference to any bills in either chamber, and, importantly, the involvement of any subcommittee. Numerous hearing connect the content of the proceedings to a specific subcommittee within the larger committee, and note the subcommittee membership. While this has considerable importance for the conduct of congressional business, and likely impacts how the larger chamber interprets the work of the committee (in the sense of signaling used by Diermeier and Feddersen (2000)), any committee chair is automatically a member of any subcommittee; I still use the committee chair's ideal point even when the hearings indicate the presence of a subcommittee.

4.4 FINDINGS

In the following sections, I demonstrate the support my hypotheses receive under several specifications of the relationship between ideological congruence and information transmission. First I examine only the bivariate relationship posited in the theoretical model. Second, I expand the relationship to include several additional variables that could arguably moderate the theoretical relationship.

4.4.1 Ideology and Informativeness

According to my results, the amount of information related in a hearing, measured by specificity of speaker language, increases as the ideological distance between the Committee Chair and the witness decreases. Crawford and Sobel (1982) argument applies well to the setting of congressional hearings, as the messages sent by the witness are more finely partitioned with a narrowing of the gap between sender and receiver. In this section I detail the strength of this relationship, and show that it is robust to several plausible specifications.

I use a simple linear regression model to test the relationship between ideological distance and speaker specificity. One advantage of building on the theoretical CS model is the analytically clean econometric model that it entails. Table 12 shows the results of running an ordinary least squares model for the key dependent variable and three versions of the dependent variable. The clear result is support for Hypothesis 4.1 across all versions of the model. The first version uses just the raw *Speaker Specificity* as the dependent variable. The next two models use transformed versions of *Speaker Specificity*, first transformed relative to each session of Congress (Within-Congress), and second transformed relative to the full range of sessions (Between-Congress). Recall that the basic specificity variable takes a value in $[0, 1]$, while the transformed variables are centered on 0, with values expressing the number of standard deviations from that mean. The direct interpretation of the coefficients are thus different among these models. That said, the general results are the same across transformations: increasing ideological distance results in a decrease in the amount of information delivered.[12] Comparing one transformed variable to the other, the largest effect is obtained with the Between-Congress (BC) Z-Score, where a unit increase in ideological distance results in a decrease of specificity of 0.20, or one-fifth of a standard deviation. The substantive effect is thus large, given that the entire range of the BC variable's support spans roughly four standard deviations in either direction from the mean. I only use the normalized versions of the dependent model in the rest of the analysis.

[12] This language is expressly causal. The model is explicitly meant to use empirical data to test a proposed causal relationship among the theoretical entities (Morton and Williams 2010).

Table 12: **Increasing Ideological Difference Decreases Information**

Dependent Variable:	Specificity (Raw)	Within-Congress Z-Score	Between-Congress Z-Score
Ideological Distance	-0.03^{***}	-0.14^{***}	-0.20^{***}
	(0.00)	(0.01)	(0.01)
Intercept	0.45^{***}	0.05^{***}	0.08^{***}
	(0.00)	(0.01)	(0.01)
N	128529	128529	128529

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Hypothesis 4.2 posits a positive relationship between the amount of time a Committee Chair has served and the amount of information delivered. Again using a simple linear model, Table 13 shows that there is a statistically significant relationship. Chair seniority is denoted in integers corresponding to the number of Congressional sessions the individual has been in place. A unit increase in Chair Seniority, then, has a positive, but substantively small effect on the overall amount of information transmitted. This result says little about the theory of committee specialization writ large. Since hearings are convened at the prerogative of chairs, the chair may choose not to hear about topics for which they are relatively confident the full chamber recognizes their expertise. Within a chair's tenure, a single or limited number of hearings might suffice to signal policy specialization, reducing the value of subsequent sessions. The observed set of hearings then, are chosen to convey ongoing rather than prior specialization, so that seniority has no effect for a given hearing on any particular topic. Thus, communicating specialization may motivate convening the hearing, but does not have an effect on the testimony delivered.

While the House is larger in size, and thus has more capacity to, and does, hold more hearings, the effect of ideological distance on information is present in both chambers. Table 14 shows that ideological differences result in less information being delivered in both chambers. The effect in the House is a little under double that of the effect in the

Table 13: **Increasing Chair Seniority Modestly Increases Information**

Dependent Variable:	Within-Congress Z-Score	Between-Congress Z-Score
Chair Seniority	0.002***	0.003***
	(0.00)	(0.00)
Intercept	−0.03***	−0.04***
	(0.01)	(0.01)
R^2	0.001	0.001
N	130670	130670

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Senate (−0.15 vs. −0.09, respectively). The number of observations is also approaching only twice that of the Senate, despite the number of members of the House being over four times that of the Senate.

Congress ostensibly uses hearings to elicit information from witnesses in order to minimize the information asymmetry, at least enough to enable evaluation of some policy choice. The amount of information, as these results show, are highly dependent on the ideological congruence between a Committee Chair and the witness (in all these cases, a representative of an executive branch agency). This dynamic is reminiscent of the ally principle common to many models of principal-agent relationships (Bendor and Meirowitz 2004). Protecting their information advantage helps bureaucracies retain, and even enhance, autonomy over policy choices and implementation (Aberbach 1990; Carpenter 2001; Shipan 2005). In this view the Committee Chair is the principal, and the witness the agent. As the preference of the principal and the agent converge, there is less need for the former to exert post-delegation pressure on the latter (Huber and Shipan 2002; Mathew D. McCubbins, Noll, and Weingast 1987). Under these conditions, information may be more freely given by the agent, since the expectation is a policy outcome that is more favorable than under a more ideologically disparate Committee Chair.

Table 14: **Ideological Distance Matters in Both Chambers**

Dependent Variable:	Speaker Specificity (WC Transformation)	
	House	Senate
Ideological Distance	-0.15^{***}	-0.09^{***}
	(0.02)	(0.02)
Intercept	0.06^{***}	0.05^{***}
	(0.01)	(0.01)
R^2	0.001	0.001
N	79592	48937

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.4.2 Additional Influences on Specificity

Several facets of Congressional hearings may serve to enhance or dampen the relationship between ideological congruence and information transmission. I note several of these for which data can be obtained through the hearing itself, or via external sources.

First, a contest for control in the US system of separated and shared powers colors how the government operates in most, if not all, areas of influence. Divided government – when one chamber's majority party is different from that of the president – may spur Congress to greater pugnacity in calling, and conducting hearings. I thus include *Divided Government* as a binary variable equal to 1 when the majority party of the chamber holding the hearing is different than that of the president. With a bureaucracy under opposition control, in the case of divided government, I expect that the agency representative will be less forthcoming, and thus have lower specificity. Stated directly, I expect *Divided Government* to exert a negative influence on the amount of information revealed.

Furthermore, the tenacity with which Congress pursues its agenda may be a function of the size of the majority party's advantage. A greater majority may convince the party members that policy initiatives are more likely to become legislation, and should be bolstered with more hearings, and with more ideologically-compatible witnesses. I thus include *Majority Strength* as a count variable indicating the number of members in the majority party minus the number in the minority party. The agency representative would also expect a greater chance of policy change as the majority increases, and should provide better information in order to prevent the committee from forming policy based on ignorance or misinformation. As *Majority Strength* increases, then, I expect the amount of information to increase.

Finally, the actual conduct of a hearing has one considerable deviation from the CS game: the actors actually more often engage in a dialog, rather than a single witness delivering her testimony. (Uncontested statements do happen with regularity, however. Appropriations hearings frequently entail agency representatives reading without interruption lengthy texts about the fiscal performance of the agency over the past year.) Committee chairs have control over who speaks, when, and for how long. They can, and do, regularly interrupt testimony to pose questions or contest the witness' statement. Moreover, they can do this themselves, or allow other members of the committee to do so. The amount this kind of back-and-forth occurs in a single hearing can be captured by calculating the number of times a speaker in a hearing made a discrete remark (defined as a single uninterrupted string of words), and dividing it by the number of remarks made in the entire hearing. I call this variable *Interactivity*, as it captures how many times the speaker had to stop their testimony to clarify a point, answer a question (posed by any person), or restart due to some kind of interruption. While the name of the variable is chosen to be free of bias, I argue that these interactions arise because the chair is dissatisfied with how forthcoming the witness is, and thus the frequency of these interruptions should positively correlate with information transmission. More prosaically, if the chair likes what she hears, the witness will go on without interruption.

The results of these expanded models are presented in Table 15. I include each new variable with the previous two predictors in a step-wise fashion. This is to show not only the coefficient for the newly included variable, but how each addition affects the two variables of primary interest, *Ideological Distance* and *Seniority*. Note that I use the

Within-Congress transformation of the dependent variable *Speaker Specificity* in all the extended models.

The basic results hold through all iterations of the model: increasing ideological distance decreases the amount of information transmitted, while seniority has the opposite (though still small) effect. Further, for *Majority Strength* and *Interactivity*, the relationship is in the posited direction. The most prominent outcome here is the substantive size of the coefficient on the *Interactivity* variable: 0.704. The implication is that aggressive exchanges in hearings produce better information. Particularly newsworthy hearings, despite their propensity to invite grandstanding, may result in quality policy information as the witnesses are subject to intense scrutiny. As noted, though, we cannot expect better *policy* as a result of contentious hearings. Witnesses may provide a great deal of new and useful insights, but there is no guarantee the committee members employ them in legislation. The coefficient on *Divided Government* is different than I proposed – different parties in power in the executive and congress correlates positively with the amount of information delivered. Perhaps committee members are more aggressive in their interrogation of witnesses, so that *Interactivity* is increasing in *Divided Government*, though I do note that the move from unified to divided party control is substantively small.

4.5 DISCUSSION

In this paper, I explore whether or not hearings provide members of Congress with information necessary for successful policy making. I find strong evidence that Congressional hearings can be more than a propaganda channel Truman (1951), but only when the two sides of the microphone are in agreement about the subject at hand. Of course, if both sides are in agreement, any policy change will satisfy the very target of the policy, and thus is indistinguishable from capitulation. As interests diverge, Congress' oversight wanes in effectiveness in as much as the information they can extract from the bureaucracy is

Table 15: **Expanded Models**

Model:	1	2	3	4	5
Ideological Distance	-0.132^{***}	-0.113^{***}	-0.115^{***}	-0.115^{***}	-0.109^{***}
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)
Divided Government			0.004	0.005	0.008
			(0.004)	(0.005)	(0.005)
Interactivity					0.697^{***}
					(0.055)
Majority Strength				0.000	-0.000
				(0.000)	(0.000)
Seniority		0.001^{***}	0.001^{***}	0.001^{***}	0.001^{***}
		(0.000)	(0.000)	(0.000)	(0.000)
Intercept	0.012	-0.012	-0.014	-0.020	-0.082^{***}
	(0.009)	(0.012)	(0.012)	(0.016)	(0.017)
Agency Fixed Effects	✓	✓	✓	✓	✓
R^2	0.001	0.001	0.001	0.001	0.003
N	128529	128529	128529	128529	128529

Standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

curtailed. Given substantial enough ideological differences, the hearing is little more than a media ploy. Policy change may, in such a case, be built on purely political desires.

These conclusions stem from developing an empirical analog for a theoretical model first proposed by Crawford and Sobel (1982). To build a structural model that captures the key entity of the CS game – message partitioning – I subjected the text of thousands of Congressional hearings to text analysis methods that quantified the specificity of the language used in delivered testimony. This technique has promise in areas across political science, quantifying the content of political speeches, party manifestos, campaign advertising, legislation, ballot initiatives, and could even complement previous work with text as data (such as press releases used by members of Congress to communicate with their constituents (Justin Grimmer 2010)). Moreover, the approach I take does not use the standard bag of words assumption about text (J Grimmer and Stewart 2013). Instead, I account for the context of the recorded statements by showing that policy information is directly related to language specificity, which itself is a function of the strategic dynamic between a witness and a committee chair. This allows me to compare the behavior of speakers across topics, congressional sessions, and executive branch agencies.

The main substantive contribution of my paper is partially resolving a long-standing uncertainty about the value of hearings for their stated purpose: to assess, and if needed restrain, the activity of the bureaucracy. My demonstration of the applicability of the CS model to hearings, and the success of its predictions, is simultaneously academically reassuring and politically troubling. We have a satisfying abstract model of the process, though it forces us to confront the implication that those hearings may serve little purpose beyond political theater. Rhetorical aggressiveness may rescue the policy value of a hearing by dragging information out of a reluctant witness; the need to engage in such questioning, however, is only likely when the amount of information being delivered started at a low point. The positive relationship between information and ideological similarity is technically and substantively akin to the ally principle in the study of delegation, agreeing with both theoretical and experimental evidence (Dickhaut, McCabe, and Mukherji 1995).

Extension of this work would build on the language model itself. While there is ample reason to believe that major newspapers provide a useful corpus to build the classification model, developing an original corpus of policy-related material would undoubtedly improve the algorithm's success. Hearing text could serve as the training data itself, even incorporating information about the policy area being discussed. Understanding more about committees as preference outliers, for instance, could be addressed by augmenting my approach with information about a committee chair's specific interests – members with rural constituencies pursuing farm-related legislation, for example (Krehbiel 1990; Overby and Kazee 2000). Regardless of the specific direction, text analysis built on a rigorous theory of language usage provides a new method to address open questions about how information shapes political decision making.

5.0 CONCLUSION

The foregoing three essays could fairly be said to be as much about *misinformation* as about information. At the heart of each essay is a question whose answer hinges not on an actor having or using just *any* information, but the *right* information. Candidates try to share their true selves with just a select group, hoping other voters will judge them based on their incomplete or incorrect views. People locked in negotiation may fail to fully understand the implications of delay, and develop incorrect beliefs about how much an audience is willing to punish their behavior. And political agents prevaricate to maintain their relative informational advantage. That said, the work here builds on a foundation that provides structure and guidance. The rational choice tradition often leaves information as a highly abstract, almost formless concept. And rightly so, for many cases. This dissertation, however, is predicated on the view that the specific forms of that information are often critical for understanding how its acquisition, use, and manipulation affect political outcomes. By rooting the information in a specific context – campaigns, negotiations, and congressional hearings in this case – I have shown that explicit consideration of content increases our understanding of several important political phenomena.

The first essay, “Dog Whistling,” extends the standard concept of message delivery by making the content of the message a function, where the resulting information is probabilistic, and conditional on audience traits. The function is a simple, but powerful extension. Audience traits, and the fact that speakers tailor their messages according to them, is a standard concept in cognitive psychology (where it is termed “audience design”) (Clark and Murphy 1982). Messaging could thus be a function of any number of politically salient traits, such as race, education, or political sophistication. This conditional aspect of messaging also makes empirical exploration of the concept more tangible: the group distinction provides a clear divide on which to test the reception of messages. We could learn how and when speech is divisive in more than just religious examples (Albertson 2014). The second essay, "Time and Punishment," focuses on blame as a strategic tool such as voters might have over their elected representatives. Prior work on blame either limits the target (Groseclose and McCarty 2001), or employs survey methods that do not allow for endogenous allocation of blame based on controlled levels of information (Brown 2010; Malhotra and Kuo 2007). Making blame a concrete, and measurable, is a distinct contribution that, again, links the formal modeling approach with work on the psychology of voters. Moreover, given the frequency of standoff-like political fights in the current U.S. Congress, we address a highly topical concern with an experiment that lets us observe how people assess culpability for costly delay or even total breakdown. Extending this work would import a dynamic setting, allowing for repeated negotiations and multiple opportunities for punishment. "Information Transmission in Congressional Hearings" presents both the most direct test of political information and the most technical contribution of this dissertation. Viewing messages as abstract spaces to be partitioned by a speaker is, at first glance, an almost abstruse concept that is amenable only to the most theoretical of work. Linking these messages to actual language, however, is possible and useful. I exploited this link to examine whether or not oversight hearings in congress can truly be used by political principals to extract information from their agents. Given the pervasiveness of the principal-agent relationship in politics, and thus political science, the method I applied has application in a vast array of topics. From the American perspective: the speeches delivered by legislators on the chamber floor could, once measured by my process, be identified as truly attempting to impart policy specifics, or toothless political bromides; the informativeness of campaign advertising, the subject of considerable scrutiny, could be compared at a far more precise level, and across a far greater sample, than has been done previously; the informational substance of comments on proposed regulation given during notice-and-comment periods could be scrutinized for their impact on final regulatory language. Beyond the technical merits, this essay contributes to work on cross-branch relations and representativeness of the bureaucracy:

the ability to extract useful information from the executive branch diminishes as the two sides diverge on their policy preferences, making it hard for the elected body to exert adequate control over the unelected one.

The centrality of information to political activity is manifest, its impact on all arenas clear and profound. Consider a simple historical example that fuses this dissertation's distinct lines of inquiry into an illustrative whole: the Watergate Hearings. Starting in June 1972, the nation was gripped by news of a break-in at the Democratic National Committee's (DNC) headquarters, and the subsequent revelations that President Richard Nixon not only knew of this particular crime, but may have actively participated in planning, and then concealing, the planting of listening devices in the offices. As the press reported on the crimes, and the complicated relationships among the burglars caught in the DNC offices, the Republican party leadership, and the president's own staff, Senator Howard Baker posed this question to Nixon legal counsel John Dean during a Congressional hearing:

> "What did the president know, and when did he know it?"

Baker's concise inquiry still echoes through committee hearings and on cable news because it coalesced the entire scandal into a single, fundamental problem statement. Nixon's fate rested on his specific knowledge about criminal activities conducted during his campaign. The central concern with the investigation was uncovering and evaluating that knowledge. Baker's question was posed in frustration, since the information Dean and his colleagues provided during the hearing was, according to Baker, woefully inadequate to answer the question.

This inadequacy should come as no surprise to a reader of my dissertation. These hearings, comprising two nearly diametrically opposed parties, were bound to produce little of informational value. Even aggressive questioning is of little use when the starting point for transmission is so close to zero. After Baker's famous line in October, Nixon went on to coin his own notorious phrase:

> "I am not a crook."

This was meant to defend against allegations that Nixon had personally profited from various deals, tax evasion, and even possible blackmail. Notably, the word crook is unlikely to trigger the thought of approving illegal spying, but does conjure up thoughts of stolen money. The difference is subtle, but important: the term addresses the financial gains portion of the Watergate scandal, but *not* the issues around abuses of power. For a political actor so calculating in his approach to every aspect of the presidency, and so driven by concerns of public perception, these carefully chosen words were meant to direct attention away from one topic and towards another. And, of course, the final outcome of the investigation process was Nixon's resignation, an act that avoided the onrushing political cliff of impeachment. Perhaps Nixon and his aides fought so long, holding fast to their respective postions and ultimately delaying the inevitable, because they based their choices on an incorrect belief about where voters will lay the most blame, just as we suggest negotiators might. Information – its concealment, manipulation, and ultimate revelation – brought down a president, and shaped politics for generations (Kutler 1992; T. H. White 1975).

APPENDIX A

DOG WHISTLING PROOFS

A.1 PROOF OF PROPOSITION 1: NO "FULLY TRUTHFUL" SEPARATING EQUILIBRIA

No equilibria exist where both Candidate 1 types send messages that match their actual type.

A.1.1 Strategy

In a separating equilibrium Candidate 1's strategy is to send a message that matches her underlying type, which I refer to as being "truthful":

$$\sigma(C_{1_\gamma}) = \begin{cases} m = o & \text{if} \quad \gamma = O \\ m = i & \text{if} \quad \gamma = I \end{cases} \tag{A.1}$$

A.1.2 Voter Beliefs and Best Responses

There are two types of voters to consider, out-group and in-group. Denote beliefs by $\mu_\beta, \beta \in \{O, I\}$. In the standard Bayesian way, voters update their beliefs upon hearing a message from C1. First, the out-group voter beliefs about C_1's type, given the message they hear and C1's strategy, are:

$$
\begin{aligned}
\mu_O(C_{1_O} \mid h_o) &= \frac{Pr(h_o \mid C_{1_O})Pr(C_{1_O})}{Pr(h_o \mid C_{1_O})Pr(C_{1_O}) + Pr(h_o \mid C_{1_I})Pr(C_{1_I})} \\
&= \frac{1 \cdot g}{(1 \cdot g) + (0 \cdot (1-g))} \\
&= 1 \\
\mu_O(C_{1_O} \mid h_i) &= \frac{Pr(h_i \mid C_{1_O})Pr(C_{1_O})}{Pr(h_i \mid C_{1_O})Pr(C_{1_O}) + Pr(h_i \mid C_{1_I})Pr(C_{1_I})} \\
&= \frac{0 \cdot g}{(0 \cdot g) + (1 \cdot g)} \\
&= 0 \\
\mu_O(C_{1_I} \mid h_i) &= \frac{Pr(h_i \mid C_{1_I})Pr(C_{1_I})}{Pr(h_i \mid C_{1_I})Pr(C_{1_I}) + Pr(h_i \mid C_{1_O})Pr(C_{1_O})} \\
&= \frac{1 \cdot (1-g)}{(1 \cdot (1-g)) + (0 \cdot g)} \\
&= 1 \\
\mu_O(C_{1_O} \mid h_i) &= \frac{Pr(h_i \mid C_{1_O})Pr(C_{1_O})}{Pr(h_i \mid C_{1_O})Pr(C_{1_O}) + Pr(h_i \mid C_{1_I})Pr(C_{1_I})} \\
&= \frac{0 \cdot g}{(0 \cdot g) + (1 \cdot g)} \\
&= 0
\end{aligned}
\tag{A.2}
$$

The same process holds for in-group voters hearing each message, giving: $\mu_I(C_{1_I} \mid h_i) = \mu_I(C_{1_O} \mid h_o) = 1$, and $\mu_I(C_{1_I} \mid h_o) = \mu_I(C_{1_O} \mid h_i) = 0$. Put succinctly, voters believe what they hear, placing no weight on the probability that the candidate has a group membership other than what they hear C_1 claim.

Voters always cast votes for the candidate that provides them the highest utility. Here, this is also the candidate with the highest probability of matching their group type. This can be seen in the comparison of expected utility for voting for C_1 or C_2. I again start by considering out-group voters, both when $C_{2_\gamma} = O$ and when $C_{2_\gamma} = I$.

$$
\begin{aligned}
EU_{v_O}(C_1 \mid h_o) &= \mu_O(C_{1_O} \mid h_o)U_{v_O}(C_{1_O}) + \mu_O(C_{1_I} \mid h_o)U_{v_O}(C_{1_I}) \quad = \quad 0 \\
U_{v_O}(C2 \mid C_{2_O}) &= 0 \\
U_{v_O}(C2 \mid C_{2_I}) &= -1
\end{aligned}
\tag{A.3}
$$

When out-group voters hear C_1 claim out-group status (s_o), they believe the message with certainty. This means that their expected utility for voting for C_1 is the same as their utility for voting for a C_{2_O}. In this case voters choose randomly: $Pr(v_o = C1) = 1/2$. Voters clearly prefer a C_1 that sent s_o against a C_{2_I}, and will vote accordingly. The same process applies to in-group voters considering a C_1 that sent s_i, versus the two types of C_{2_γ}. Thus, voters will cast their votes for the candidate whom they are certain match their group-type, and will decide via toss-up if both candidates match. Note that this tie-breaking rule also applies if the voter believes candidates have the same type, and both are of a different type than the voter. Together, this means that the voters' best responses are:

$$BR_{v_O}(C_{2_\gamma}, \mu_O) = \begin{cases} C_1 & \text{if} \quad h_o, C_{2_I} \\ \frac{1}{2} & \text{if} \quad h \in \{o, i\}, h_y = C_{2_\gamma} \\ C2 & \text{if} \quad h_i, C_{2_O} \end{cases} \tag{A.4}$$

$$BR_{v_I}(C_{2_\gamma}, \mu_I) = \begin{cases} C_1 & \text{if} \quad h_i, C_{2_O} \\ \frac{1}{2} & \text{if} \quad h \in \{o, i\}, h_y = C_{2_\gamma} \\ C2 & \text{if} \quad h_o, C_{2_I} \end{cases} \tag{A.5}$$

A.1.3 C_1 Utilities

If Candidate 1 is truthful, she sends a message that matches her underlying type, which in turn affect the voter's choice. Candidate 1's utility is thus also a function of Candidate 2's type. Table 16 lists all the permutations of messages and C_2 types.

Table 16 shows utility for each permutation. When C_1's message differs from C_2's

Table 16: **Utilities for Combinations of Candidate Types and C_1 Messages**

Candidate 2 Type	Candidate 1 Message	Candidate 1 Utility
C_{2_O}	s_o	$\frac{1}{2}$
C_{2_O}	s_i	v_I
C_{2_I}	s_o	v_O
C_{2_I}	s_i	$\frac{1}{2}$

type, C_1 captures the votes of all those voter types that match the message. When C_1's message matches C_2's type, the voters decide via toss-up, resulting in an expected vote share of half of all voters.

Except when voter groups are perfectly balanced ($v_o = v_I$), the "minority" type of Candidate 1 (the type whose group matches the smaller voter population) will have a profitable deviation from the separating strategy. Since no distributional constraints are placed on voter types, either group-type may dominate the other. To maximize utility, the minority type of C_1 should match her message to C_2's type. For instance, when $v_I = 0.4$, C_{1_I} sending s_i against a C_{2_O} earns a utility of 0.4, which is clearly less than the 0.5 C_1 receives for making the voters believe the candidates are similar. Indeed, when the voters are not evenly balanced, the minority type of Candidate 1 always improves her utility by matching her message to the majority voter type: either she becomes the lone candidate (voters believe is) in the majority, and thus gains the larger share of votes, or she moves from obtaining only the minority voters to splitting all votes evenly.

The knife-edge case of a balanced voter population supports the proposed separating equilibrium. However, this support disappears in the event of dog whistling.

A.1.4 Using a Dog Whistle

Candidate 1 also has the option of using a dog whistle instead of clearly claiming one group status or the other. For this candidate equilibrium, I only need to show that the use of a dog whistle is a better strategy for C_1 than claiming her true group membership when the voting population is strictly balanced.[1] First, I work through how the dog whistle impacts the play of the game.

Dog whistling introduces a probabilistic aspect to message reception. However, given the separating strategy, this has no effect on voter beliefs about C_1's type. Voters still believe that C_1 is the type claimed in the message they *hear* (as defined in Equation (2.1)). Voters cast their votes according to Equations (A.4) and (A.5), though this yield different utility for C_1 since the votes are a function of what message the voter hears.

Consider C_{1_O} deviating to using a dog whistle instead of the clear message s_o. Table 17 gives the new utilities for C_{1_O}. Using a dog whistle results in C_1 getting votes from: all those voters that hear a message that matches their type when $C_{1_\gamma} \neq C_{2_\gamma}$; half the

[1] The presence of the deviation to claiming the majority group membership (when C_{1_γ} is in the minority) is enough to prove the absence of separating equilibria in all other cases.

Table 17: **Utilities for C_{1_O} Whistling**

Candidate 2 Type	Candidate 1 Message	Candidate 1 Utility
C_{2_O}	s_w	$pv_I + \frac{1}{2}(1-q)v_O + \frac{1}{2}(1-p)v_I$
C_{2_I}	s_w	$(1-q)v_O + \frac{1}{2}pv_I + \frac{1}{2}qv_O$

votes of those voters that hear a message indicating $C_{1_\gamma} = C_{2_\gamma}$. C_1, importantly, loses some voters that she would otherwise gain, specifically those voters that match her type but who hear $m \neq v_\beta$. The case for C_{1_I} is identical.

Fixing $v_I = v_O = 0.5$ for the case of interest, C_1's utilities in Table 17 clearly present profitable deviations for either type of C_1 against either type of C_2. Thus, in the case of a strictly balanced voting population, both types of C_1 would opt to use a dog whistle rather than sending a clear message about her type (either accurate or inaccurate). A dog whistle may increase utility for one or both C_1 types when the audience is unbalanced. However, this only provides another example of deviation, and would not be more informative for the purpose of this analysis. At minimum, when the audience groups are unbalanced, the minority type of C_1 will choose to deviate.

In summary, there are no fully separating equilibria. At least one type of C_1 will have an incentive to deviate. If one voter group-type is larger than the other, the type of C_1 corresponding to this smaller audience will, at least, deviate to sending a (clear) message that differs from her true group type.

A.2 PROOF OF PROPOSITION 2: TWO SEPARATING EQUILIBRIA

A.2.1 Strategies

For this equilibrium, one type of C_1 chooses to whistle, while the other always sends a truthful message. For the moment, assume that the out-group type of C_1 chooses to whistle, while the in-group type chooses to truthfully reveal her type. For this proposed equilibrium:

$$\sigma(C_{1_\gamma}) = \begin{cases} s_w & \text{if} \quad \gamma = O \\ s_i & \text{if} \quad \gamma = I \end{cases} \tag{A.6}$$

A.2.2 Voter Beliefs and Best Responses

There are two types of voters to consider, out-group and in-group. Denote beliefs by $\mu_\beta, \beta \in \{O, I\}$. Voters update with the standard Bayesian process (where possible), though one difference arises because of the whistling portion of the strategy. The likelihood term in the Bayes equation is neither certainty (as it would be for a strategy positing a speaker always issuing one message) nor zero (a it would be in a strategy positing a speaker never issuing a particular message). In the case of whistling as part of C_1's strategy, the likelihood equals the voter's group-type's probability of hearing a particular message (as defined in Equation (2.1)). Thus, out-group voters beliefs, given C_1's strategy, are:

$$\begin{aligned}
\mu_O(C_{1_O} \mid h_o) &= \frac{Pr(s_o \mid C_{1_O})Pr(C_{1_O})}{Pr(h_o \mid C_{1_O})Pr(C_{1_O}) + Pr(h_o \mid C_{1_I})Pr(C_{1_I})} \\
&= \frac{(1-q)\cdot g}{((1-q)\cdot g) + (0\cdot(1-g))} = 1 \Rightarrow \\
\mu_O(C_{1_I} \mid h_o) &= 0 \\
\mu_O(C_{1_O} \mid h_i) &= \frac{Pr(h_i \mid C_{1_O})Pr(C_{1_O})}{Pr(h_i \mid C_{1_O})Pr(C_{1_O}) + Pr(h_i \mid C_{1_I})Pr(C_{1_I})} \\
&= \frac{qg}{gq - g + 1} \\
\mu_O(C_{1_I} \mid h_i) &= \frac{Pr(h_i \mid C_{1_I})Pr(C_{1_I})}{Pr(h_i \mid C_{1_I})Pr(C_{1_I}) + Pr(h_i \mid C_{1_O})Pr(C_{1_O})} \\
&= \frac{1-g}{gq + 1 - g}
\end{aligned} \tag{A.7}$$

These beliefs show how whistling produces voter uncertainty. Only C_{1_O} ever sends s_o, so voters are certain about the candidate type if they hear h_o. However, there are two ways for voters to hear h_i, so voters cannot perfectly infer C_1's type. Thus, voters who hear h_i have posterior beliefs that differ from their prior beliefs, and that are strictly between 0 and 1. The same process holds for in-group voters. The group-based difference in the probability of hearing particular message changes the likelihood term when h_i, yielding $\mu_I(C_{1_O} \mid h_i) = \frac{pg}{pg-g+1}$, and $\mu_I(C_{1_I} \mid h_i) = \frac{1-g}{pg-g+1}$.

Voters always cast votes for the candidate they believe will maximize their utility. This is also the candidate that with the highest subjective probability of matching their group-type. This can be seen in the comparison of expected utility of voting for C_1 or the utility of voting for C_2. I again start by considering out-group voters, both when $C_{2_\gamma} = O$ and when $C_{2_\gamma} = I$.

$$
\begin{array}{rcl cl}
EU_{v_O}(C_1 \mid h_o) & = & \mu_O(C_{1_\gamma} = O \mid s_o)U_{v_O}(C_{1_O}) + \mu_O(C_{1_I} \mid s_o)U_{v_O}(C_{1_I}) & = & 0 \\
EU_{v_O}(C_1 \mid h_i) & = & \mu_O(C_{1_\gamma} = O \mid s_i)U_{v_O}(C_{1_O}) + \mu_O(C_{1_I} \mid s_i)U_{v_O}(C_{1_I}) & = & -\dfrac{gq}{gq+1-g} \\
U_{v_O}(C2 \mid C_{2_O}, h_o) & = & 0 & & \\
U_{v_O}(C2 \mid C_{2_I}, h_o) & = & -1 & &
\end{array}
\tag{A.8}
$$

In the even that both C_{1_O} and C_{2_O}, v_O that hear h_o choose by tossing a fair coin. However, v_O that hear h_i have $\mu_O(C1 = O \mid h_i) < 1$, so they will cast votes for C_{2_O}. Note that v_o will never vote for C_{2_I}, no matter what message they hear. The comparison for in-group voters is symmetric, so v_I that hear h_i will choose via coin-toss between C_1 and C_{2_I}, while v_I that hear h_o will always vote for C_{2_I}.

Comparing what they hear and C_2 types, voter best responses are given by:

$$
BR_{v_O}(C_{2_\gamma}, \mu_O) = \begin{cases} \dfrac{1}{2} & \text{if} \quad h_o \quad \& \quad C_{2_O} \\ C2 & \text{if} \quad h_i \quad \& \quad C_{2_O} \\ C_1 & \text{if} \quad h \in \{o, i\} \quad \& \quad C_{2_I} \end{cases} \tag{A.9}
$$

$$
BR_{v_I}(C_{2_\gamma}, \mu_I) = \begin{cases} \dfrac{1}{2} & \text{if} \quad h_o \quad \& \quad C_{2_O} \\ C2 & \text{if} \quad h \in \{o, i\} \quad \& \quad C_{2_I} \\ C_1 & \text{if} \quad h_i \quad \& \quad C_{2_O} \end{cases} \tag{A.10}
$$

Table 18: **Utilities for Combinations of Candidate Types and C_1 Messages**

Candidate 1 Type	Candidate 1 Message	Candidate 2 Type	Candidate 1 Utility
C_{1_O}	s_w	C_{2_O}	$\frac{1}{2}(1-q)v_O + \frac{1}{2}(1-p)v_I + pv_I$
	s_i		v_I
	s_o		$\frac{1}{2}$
C_{1_O}	s_w	C_{2_I}	v_O
	s_i		v_O
	s_o		v_O
C_{1_I}	s_w	C_{2_O}	$\frac{1}{2}(1-q)v_O + \frac{1}{2}(1-p)v_I + pv_I$
	s_i		v_I
	s_o		$\frac{1}{2}$
C_{1_I}	s_w	C_{2_I}	v_O
	s_i		v_O
	s_o		v_O

A.2.3 C_1 Utilities

Using the voter best responses, Table 18 shows C_1's utility under permutations of message and C_2 type.

First, note that C_1's only changes as a function of her message when C_{2_O}. This asymmetry (the lack of different utilities when C_{2_I}) is a result of C_1's strategy in this equilibrium. The opposite — C_1's utility is the same for any message sent against a C_{2_I} — holds when the strategy has only C_{1_I} sending s_w. When C_{2_O}, dog whistling is the optimal message when $U_{C_{1_O}}(s_w) \geq U_{C_{1_O}}(s_i) \geq U_{C_{1_O}}(s_o)$:

$$\begin{aligned} \frac{1}{2}(1-q)v_O + \frac{1}{2}(1-p)v_I + pv_I \quad &\geq \quad v_I \quad \geq \quad \frac{1}{2} \\ &\Rightarrow \\ \frac{-1+q}{-2+p+q} \quad &\geq \quad v_I \quad \geq \quad \frac{q}{p+q} \end{aligned} \tag{A.11}$$

This expression gives bounds on the size of v_I for s_w to be supported as the message sent by C_{1_O}. However, when this holds true for C_{1_O}, it can also hold true for C_{1_I}. When it does, C_1's strategy is not optimal since C_{1_I} would have an incentive to switch from always sending s_i to sending s_w. The only time it does *not* hold is in the knife-edge situation where $U_{C_{1_O}}(s_w) = U_{C_{1_O}}(s_i) > U_{C_{1_O}}(s_o)$ (assuming indifference is decided in favor of the stated strategy). This is only true when $p = \frac{q(1-v_I)+2v_I-1}{v_I}$.

Taken together, this indicates that an equilibrium exists where only C_{1_O} uses a dog whistle, C_{1_I} reveals her true type, voter beliefs are as given in Equation (A.7), voter choices are as given in Equations (A.9) and (A.10), and $p = \frac{q(1-v_I)+2v_I-1}{v_I}$. For any other value of p, C_{1_I} will deviate to sending s_w.

The argument, and thus the result, is similar for the case where C_1's strategy is:

$$\sigma(C_{1_\gamma}) = \begin{cases} s_o & \text{if} \quad \gamma = O \\ s_w & \text{if} \quad \gamma = I \end{cases} . \tag{A.12}$$

The result pins down the specific value, $q = \frac{p(1-v_O)+2v_O-1}{v_O}$, that must hold to support the equilibrium. Otherwise, C_{1_O} will deviate to using a dog whistle.

A.3 PROOF OF PROPOSITION 3: POOLING EQUILIBRIA

There exist three equilibria where both types of C_1 pool on sending the same message, either s_o, s_i, or s_w.

A.3.1 Strategy

To pool, Candidate 1's strategy is to send the same message to voters, regardless of her type. There are two possible equilibria, one where both types send s_o, and one where both types send s_i. I begin with the former case:

$$\sigma(C_{1_\gamma}) = \ s_o \quad \forall \quad C_\gamma \in \{O, I\} \tag{A.13}$$

A.3.2 Voter Beliefs and Best Responses

There are two types of voters to consider, out-group and in-group. Denote beliefs by $\mu_\beta, \beta \in \{O, I\}$. In the standard Bayesian way, voters update their beliefs upon hearing a message from C1. Updating, in this proposed equilibrium, produces posterior beliefs for voters that equal their prior beliefs. Given the message they hear and C1's strategy, these beliefs are:

$$\begin{aligned}
\mu_O(C_{1_O} \mid h_o) &= \frac{Pr(h_o \mid C_{1_O})Pr(C_{1_O})}{Pr(h_o \mid C_{1_O})Pr(C_{1_O}) + Pr(h_o \mid C_{1_I})Pr(C_{1_I})} \\
&= \frac{1 \cdot g}{(1 \cdot g) + (1 \cdot (1-g))} = g \Rightarrow \\
\mu_O(C_{1_I} \mid h_o) &= 1 - g
\end{aligned} \tag{A.14}$$

Under the proposed pooling strategy, voters learn nothing from hearing the message h_o, and hearing h_i is off the path of play, so Bayes' Rule does not apply. The same process holds for in-group voters, giving: $\mu_I(C_{1_O} \mid h_o) = g$, and $\mu_I(C_{1_I} \mid h_o) = 1 - g$ (and undefined beliefs in the event h_i).

Voters always cast votes for the candidate they believe will maximize their utility. This is also the candidate with the highest posterior probability of matching their group type. This can be seen in the comparison of expected utility of voting for C_1 or the utility of voting for C_2. I again start by considering out-group voters, both when $C_{2_\gamma} = O$ and when $C_{2_\gamma} = I$.

$$\begin{array}{lclcl} EU_{v_O}(C_1 \mid h_o) & = & \mu_O(C_{1_O} \mid h_o)U_{v_O}(C_{1_O}) + \mu_O(C_{1_I} \mid h_o)U_{v_O}(C_{1_I}) & = & -g \\ U_{v_O}(C2 \mid C_{2_O}) & = & 0 & & \\ U_{v_O}(C2 \mid C_{2_I}) & = & -1 & & \end{array} \tag{A.15}$$

Voting for C_1 after hearing h_o has an expected value equal to the basic probability of C_{1_O}. By assumption, $g < 1$, making it always better for v_O types to vote for C_2, whom they know to be $\gamma = O$ with certainty. Out-group voters will vote for C_1 when $C_{2_\gamma} = I$. The comparison for in-group voters is symmetric.

In the event that C_1 sends s_i, I pick one set of any beliefs off the equilibrium path. For out-group voters and C_{2_O}, any $\mu_O(C_{1_O} \mid s_i)$ means that $EU_{v_O}(C_{1_O} \mid s_i) \in (0,1) < U_{v_o}(C_{2_O} \mid s_i) = 1$. That is, there are no beliefs that induce the out-group voters to switch to voting for the candidate they are certain is out-group to the candidate that is only possibly out-group. Conversely, when $C_{2_\gamma} = I$, any positive $\mu_{v_o}(C_{1_O} \mid s_i)$ means $EU_{v_O}(C1 \mid s_o) \in (0,1) > U_{v_O}(C_{2_I}) = 0$. Any positive chance that C_1 is a Different type than C_2 is enough to get that group of voters to support C_1.

This means that the voters' best responses are:

$$BR_{v_O}(C_{2_\gamma}, \mu_O) = \begin{cases} C_1 & \text{if} \quad h_o \quad \& \quad C_{2_I} \\ C2 & \text{if} \quad h_o \quad \& \quad C_{2_O} \\ C_1 & \text{if} \quad h_i \quad \& \quad C_{2_I} \, \forall \, \mu_{v_O}(C_{1_O} \mid h_i) \in (0,1) \\ C2 & \text{if} \quad h_i \quad \& \quad C_{2_O} \, \forall \, \mu_{v_O}(C_{1_O} \mid h_i) \in (0,1) \end{cases} \tag{A.16}$$

$$BR_{v_I}(C_{2_\gamma}, \mu_I) = \begin{cases} C2 & \text{if} \quad h_o \quad \& \quad C_{2_I} \\ C_1 & \text{if} \quad h_o \quad \& \quad C_{2_O} \\ C2 & \text{if} \quad h_i \quad \& \quad C_{2_I} \, \forall \, \mu_{v_I}(C_{1_O} \mid h_i) \in (0,1) \\ C_1 & \text{if} \quad h_i \quad \& \quad C_{2_O} \, \forall \, \mu_{v_I}(C_{1_O} \mid h_i) \in (0,1) \end{cases} \tag{A.17}$$

A.3.3 C_1 Utilities

Using the voter's best responses, Table 19 shows C_1's utility under permutations of message and C_2 type. The format of the table highlights the fact that, given a pair $\{C_{1_\gamma}, C_{2_\gamma}\}$, both messages yield C_1 the same utility. Further, that holds true for both types of C_1. Neither type of C_1 can improve her utility by switching messages against a given type of C_2.

Table 19: **Utilities for Combinations of Candidate Types and C_1 Messages**

Candidate 1 Type	Candidate 1 Message	Candidate 2 Type	Candidate 1 Utility
C_{1_O}	s_o	C_{2_O}	v_I
	s_i		v_I
C_{1_I}	s_o	C_{2_O}	v_I
	s_i		v_I
C_{1_O}	s_o	C_{2_I}	v_O
	s_i		v_O
C_{1_I}	s_o	C_{2_I}	v_O
	s_i		v_O

In words, this says that C_1 always obtains the share of voters opposite C_2's type, regardless of C_1's message. Thus, there is no incentive for either type of C_1 to deviate to a different (clear) message.

Candidate 1 also has the option of using a dog whistle (s_w) instead of clearly claiming one group status or another. One of either C_1 type may switch to whistling, and I start with the case where C_{1_O} whistles.

Dog whistling introduces a probabilistic aspect to message reception. However, given the pooling strategy, whistling does not change voter beliefs about C_1's type. When the candidates are not pooling on either s_i or s_o, then regardless of what the voters hear *hear*, their posterior beliefs about C_{1_γ} equal their prior beliefs (though, candidate pooling may require off the equilibrium path voter beliefs as noted above). Recall that voters hear messages according to Equation (2.1), and cast votes according to Equations (A.16) and (A.17).

Table 20: **Utilities for Combinations of Candidate Types and C_1 Messages**

Candidate 1 Type	Candidate 1 Message	Candidate 2 Type	Candidate 1 Utility
C_{1_O}	s_o	C_{2_O}	v_I
	s_i		v_I
	s_w		v_I
C_{1_I}	s_o	C_{2_O}	v_I
	s_i		v_I
	s_w		v_I
C_{1_O}	s_o	C_{2_I}	v_O
	s_i		v_O
	s_w		v_O
C_{1_I}	s_o	C_{2_I}	v_O
	s_i		v_O
	s_w		v_O

Fix C_{2_O}, and let C_{1_O} send s_w. Then, C_1 obtains v_I votes, just as she did by sending either s_o or s_i. The pv_I portion of in-group voters that hear h_i prefer voting for C_1 since they have a positive belief $\mu_I(C_{1_I} \mid h_i)$, despite the message being off the equilibrium path. The $(1-p)v_I$ portion of in-group voters that hear s_o have belief $\mu_I(C_{1_I} \mid h_o) = 1 - g > 0$ (because $g \in (0,1)$). Similarly, v_O voters prefer voting for C_{2_O} regardless of the message they hear. Table 20 shows that the logic holds for all combinations of C_{1_γ} and C_{2_γ}, so that C_1, under the pooling strategy, always gets the same utility, regardless of the message sent (clear or whistle).

The above argument began by assuming C_1 types pool on s_o. The analysis shows there are equilibria where: $\sigma(C_{1_\gamma}) : s_o \, \forall \, \gamma \in \{O, I\}$, voter beliefs follow Equation (A.14), and voters strategies follow Equations (A.16) and (A.17). The argument for pooling on s_i is symmetric, proving the existence of two types of pooling equilibria.

A.3.4 Pooling on Whistling

The use of a whistle requires restating the analysis to make the result clear.

A.3.5 Strategies

This equilibrium has both types of Candidate 1 always sending a dog whistle:

$$\sigma(C_{1_\gamma}) : \quad (s_w) \,\forall\, \gamma \in \{O, I\}. \tag{A.18}$$

A.3.6 Voter Beliefs and Best Responses

There are two types of voters to consider, out-group and in-group. Denote beliefs by $\mu_\beta, \beta \in \{O, I\}$. Voters update with the standard Bayesian process (where possible), though one difference arises because of the whistling portion of the strategy. The likelihood term in the Bayes equation is not certainty (as it would be for a strategy positing a speaker always issuing one message) or zero (a it would be in a strategy positing a speaker never issuing a particular message). In the case of whistling as part of C_1's strategy, the likelihood equals a voter's group-type probability of hearing a particular message (as defined in Equation (2.1)). Thus, out-group voters beliefs, given C_1's strategy, are:

$$\begin{aligned}
\mu_O(C_{1_O} \mid h_o) &= \frac{Pr(h_o \mid C_{1_O})Pr(C_{1_O})}{Pr(h_o \mid C_{1_O})Pr(C_{1_O}) + Pr(h_o \mid C_{1_I})Pr(C_{1_I})} \\
&= \frac{g(1-q)}{g(1-q) + (1-g)(1-q)} = g \\
\mu_O(C_{1_I} \mid h_o) &= \frac{Pr(h_o \mid C_{1_I})Pr(C_{1_O})}{Pr(h_o \mid C_{1_O})Pr(C_{1_O}) + Pr(h_o \mid C_{1_I})Pr(C_{1_I})} \\
&= \frac{(1-g)(1-q)}{(1-g)(1-q) + g(1-q)} = 1-g \\
\mu_O(C_{1_O} \mid h_i) &= \frac{Pr(h_i \mid C_{1_O})Pr(C_{1_O})}{Pr(h_i \mid C_{1_O})Pr(C_{1_O}) + Pr(h_i \mid C_{1_I})Pr(C_{1_I})} \\
&= \frac{gq}{gq + (1-g)q} = g \\
\mu_O(C_{1_I} \mid h_i) &= \frac{Pr(h_i \mid C_{1_I})Pr(C_{1_I})}{Pr(h_i \mid C_{1_I})Pr(C_{1_I}) + Pr(h_i \mid C_{1_O})Pr(C_{1_O})} \\
&= \frac{(1-g)q}{(1-g)q + gq} = 1-g
\end{aligned} \tag{A.19}$$

When both types of C_{1_γ} send s_w, there are two ways for voters to hear h_i and h_o. This is because of the probabilistic element to hearing each message. Voter beliefs are functions of their group-type propensity to hear either h_o or h_o. The same process holds for in-group voters, and result in the following beliefs:

$$\begin{array}{rcl}\mu_I(C_{1_O} \mid h_o) & = & \dfrac{g(1-p)}{g(1-p)+(1-g)(1-p)} = g \\ \mu_I(C_{1_I} \mid h_o) & = & \dfrac{(1-g)(1-p)}{(1-g)(1-p)+g(1-p)} = 1-g \\ \mu_I(C_{1_O} \mid h_i) & = & \dfrac{gp}{gp+(1-g)p} = g \\ \mu_I(C_{1_I} \mid h_i) & = & \dfrac{(1-g)p}{(1-g)p+gp} = 1-g\end{array} \tag{A.20}$$

Voters always cast votes for the candidate they believe will maximize their utility. This is also the candidate that with the highest subjective probability of matching their group-type. This can be seen in the comparison of expected utility of voting for C_1 or the utility of voting for C_2.

$$\begin{array}{rcccl}EU_{v_O}(C_1 \mid h_o) & = & \mu_O(C_{1_O} \mid h_o)U_{v_O}(C_{1_O}) + \mu_O(C_{1_I} \mid h_o)U_{v_O}(C_{1_I}) & = & -g \\ EU_{v_O}(C_1 \mid h_i) & = & \mu_O(C_{1_O} \mid h_i)U_{v_O}(C_{1_O}) + \mu_O(C_{1_I} \mid h_i)U_{v_O}(C_{1_I}) & = & -g \\ U_{v_O}(C2 \mid C_{2_O}) & = & 0 & & \\ U_{v_O}(C2 \mid C_{2_I}) & = & -1 & & \end{array} \tag{A.21}$$

$$\begin{array}{rcccl}EU_{v_I}(C_1 \mid h_o) & = & \mu_I(C_{1_O} \mid h_o)U_{v_I}(C_{1_O}) + \mu_I(C_{1_I} \mid h_o)U_{v_I}(C_{1_I}) & = & -(1-g) \\ EU_{v_I}(C_1 \mid h_i) & = & \mu_I(C_{1_O} \mid h_i)U_{v_I}(C_{1_O}) + \mu_I(C_{1_I} \mid h_i)U_{v_I}(C_{1_I}) & = & -(1-g) \\ U_{v_I}(C2 \mid C_{2_O}) & = & -1 & & \\ U_{v_I}(C2 \mid C_{2_I}) & = & 0 & & \end{array} \tag{A.22}$$

Pooling on dog whistling makes voters uncertain about C_{1_γ} given whichever message they hear. By assumption, $g \in (0,1)$, so there are no prior beliefs that, after updating, make voter posterior beliefs equal 1 or 0. For voters, then, any chance $C_{1_\gamma} = v_\beta$ makes the expected utility of voting for C_1 higher than the utility of voting for C_2 when $C_{2_\gamma} \neq v_\beta$. This yields the following best responses for out- and in-group voters:

$$BR_{v_O}(C_{2_\gamma}, \mu_O) = \begin{cases} C_1 & \text{if} \quad h \in \{o,i\} \quad \& \quad C_{2_I} \\ C2 & \text{if} \quad h \in \{o,i\} \quad \& \quad C_{2_O} \end{cases} \tag{A.23}$$

$$BR_{v_I}(C_{2_\gamma}, \mu_I) = \begin{cases} C_1 & \text{if} \quad h \in \{o,i\} \quad \& \quad C_{2_O} \\ C2 & \text{if} \quad h \in \{o,i\} \quad \& \quad C_{2_I} \end{cases} \tag{A.24}$$

Table 21: **Utilities for Combinations of Candidate Types and C_1 Messages**

Candidate 1 Type	Candidate 1 Message	Candidate 2 Type	Candidate 1 Utility
C_{1_O}	s_w	C_{2_O}	v_I
	s_i		v_I
	s_o		v_I
C_{1_O}	s_w	C_{2_I}	v_O
	s_i		v_O
	s_o		v_O
C_{1_I}	s_w	C_{2_O}	v_I
	s_i		v_I
	s_o		v_I
C_{1_I}	s_w	C_{2_I}	v_O
	s_i		v_O
	s_o		v_O

A.3.7 C_1 Utilities

Using the voter best responses, Table 21 shows C_1's utility under permutations of message and C_2 type.

For any pair of message and C_{2_γ}, C_1 obtains the same utility, regardless of the message sent. This occurs because voters know C_{2_γ} with certainty and voter for her when $v_\beta = C_{2_\gamma}$, and will take any positive gamble on C_{1_γ} when $v_\beta \neq C_{2_\gamma}$.

The above argument shows that there is an equilibrium where C_{1_γ} sends s_w for $\gamma \in \{O, I\}$, supported by voter actions given in Equations (A.16) and (A.17) and beliefs as given in Equation (A.14).

Together, there are thus three equilibria where C_{1_γ} pools on messages, one each for sending $s_\{i, o, w\}$.

A.4 PROOFS OF TWO-DIMENSION MODEL RESULTS

This version of the model has two dimensions that characterize the candidates and the voters: one ideological dimension, one group dimension. In addition, there is a weighting on the voters' ideological dimension, denoted α, that allows for voters to be more or less ideological, relative to their group. That is, when there is exactly one dimension on which the voter and the candidate are not equal, an $\alpha > 1$ indicates the voter prefers voting for a candidate whose ideology matches his but does belong to the voter's group. An $\alpha < 1$ indicates the opposite: with exactly one mismatched dimension between voter and candidate, the voter prefers the candidate that matches him on the group dimension. An $\alpha = 1$ indicates indifference.

A.5 PROOF OF PROPOSITION 4: ONE TYPE OF C_1 WHISTLING

There exist equilibria where the type of C_1 that perfectly matches the type of C_2 ($C_{1_{\tau,\gamma}} = C_{2_{\tau,\gamma}}$), and only that type of C_1, uses a dog whistle, while all other types of C_1 ($C_{1_{\tau,\gamma}} \neq C_{2_{\tau,\gamma}}$) reveal their true ideological and group types.

A.5.1 Strategy

C_1's strategy in this proposed equilibrium is to use a dog whistle if, and only if, her type is the exact same as C_2's type:

$$\sigma(C_{1_{\tau,\gamma}}) = \begin{cases} s_{x,y} = \{\tau, \gamma\} & \text{if} \quad C_{1_{\tau,\gamma}} \neq C_{2_{\tau,\gamma}} \\ s_{x,y} = \{\tau, w\} & \text{if} \quad C_{1_{\tau,\gamma}} = C_{2_{\tau,\gamma}} \end{cases} \tag{A.25}$$

A.5.2 Voter Beliefs and Best Responses

Updating on beliefs occurs in the standard manner, following Bayes' Theorem. When only one candidate type uses a dog whistle, voter beliefs shift from the common priors over group-type to specific beliefs that are functions of the voter groups' differing probabilities of hearing a candidate claim in-group status ($h_{x,y} = \{\tau, i\}$). Consider the case where $C_{1_{M,I}}$ chooses the strategy $\sigma(C_{1_{M,I}}) : s_{x,y} = \{m, w\}$, while all other types choose $\sigma(C_{1_{\neg MI}}) : s_{x,y} = \{\tau, \gamma\}$. Then the voter beliefs are given as:

$$
\begin{aligned}
\mu_I(C_{1_{M,I}} \mid h_{m,i}) &= \tfrac{(1-g)pt}{(1-g)pt+tg} \\
\mu_I(C_{1_{M,I}} \mid h_{m,o}) &= \tfrac{(1-g)(1-p)t}{(1-g)(1-p)t+gt} \\
\mu_O(C_{1_{M,O}} \mid h_{m,i}) &= \tfrac{(1-g)qt}{(1-g)qt+gt} \\
\mu_O(C_{1_{M,O}} \mid h_{m,o}) &= \tfrac{(1-g)(1-q)t}{(1-g)(1-q)t+gt}
\end{aligned}
\tag{A.26}
$$

The calculations for all permutations of voter types, candidate types, and messages are similar.

Voters always cast votes for the candidate they believe will maximize their utility. In this version of the model, however Considering the beliefs noted in Eq. (A.26), voter expected utility is straightforward. Since the proof proceeds by enumeration, first consider the case where C_1 is competing against a $C_{2_{E,O}}$. Again, C_2's type is fixed and known, and C_2 sends a message that equals her true type by assumption. Considering the equilibrium strategy, this means that $C_{1_{E,O}}$ sends $s_{x,y} = \{e, w\}$, all other types of C_1 send $s_{x,y} = C_{1_{\tau,\gamma}}$.

Table 22 shows voters' expected utilities for voting for C_1, given the message they hear. The maximum utility for each voter is found by comparing values across rows. The full share of voters $C_{1_{E,O}}$ wins given what the voters hear is found by comparing the column value under the candidate type/message pair and comparing it to what the voter would get by voting for $C_{2_{E,O}}$. Note that, for voters, the utility ranking is clear for most cases – ($0 > -1 > -1-\alpha$) – but sometimes depends on comparing α and -1, since α is allowed to be greater or lesser than -1.

Completing Table 22 and comparing utilities for each possible combination, including considering the potential value for α, results in the (rather ungainly) best response

Table 22: **Utilities for Voter Candidate Choice, Given Dog Whistling Strategy for** $C_{1_{E,O}}$

Voter Type	$EU_v(C_1 \mid h_{x,y} = \{m,o\})$	$EU_v(C_1 \mid h_{x,y} = \{m,i\})$	$EU_v(C_1 \mid h_{x,y} = \{e,o\})$	$EU_v(C_1 \mid h_{x,y} = \{e,i\})$	$EU_v(C_{2_{E,O}})$
$v_{M,O}$	0	$-\alpha$	$\mu_O^o(-1) + \mu_I^o(-1-\alpha) = -1$	$\mu_O^i(-1) + \mu_I^i(-1-\alpha)$	-1
$v_{M,I}$	$-\alpha$	0	$\mu_O^o(-1-\alpha) + \mu_I^o(-\alpha) = -1-\alpha$	$\mu_O^i(-1-\alpha) + \mu_I^i(-1)$	$-1-\alpha$
$v_{E,O}$	-1	$-1-\alpha$	$\mu_O^o(0) + \mu_I^o(-\alpha) = 0$	$\mu_O^i(0) + \mu_I^i(-\alpha)$	0
$v_{E,I}$	$-1-\alpha$	-1	$\mu_O^o(-\alpha) + \mu_I^o(0) = -\alpha$	$\mu_O^i(-\alpha) + \mu_I^i(0)$	$-\alpha$

function for the voter shown in Equation (A.27):

$$BR_{v_{\theta,\beta}}(\sigma(C_{1_{\tau,\gamma}}), C_{2_{\tau,\gamma}}, \mu) = \begin{cases} C_1 & \text{if} & h_{x,y} = v_{\theta,\beta} & \& & C_{2_\tau} \neq v_\theta & & \\ C_1 & \text{if} & h_{x,y} = v_{\theta,\beta} & \& & C_{2_{\tau,\gamma}} \neq v_{\theta,\beta} & & \\ C_1 & \text{if} & h_x = v_\theta \text{ or } h_y = v_\beta & \& & C_{2_\tau} \neq v_\theta & & \\ C_1 & \text{if} & h_x \neq v_\theta = C_{2_\tau} & \& & h_y = v_\beta \neq C_{2_\gamma} & \& & \alpha \geq 1 \\ C_1 & \text{if} & h_x = v_\theta & \& & C_{2_\gamma} \neq v_\beta & \& & \alpha \geq 1 \\ C_1 & \text{if} & h_y = v_\beta & \& & C_{2_\tau} \neq v_\theta & \& & \alpha < 1 \\ 1/2 & \text{if} & h_{x,y} = v_{\theta,\beta} & \& & C_{2_{\tau,\gamma}} = v_{\theta,\beta} & & \\ 1/2 & \text{if} & h_{x,y} \neq v_{\theta,\beta} & \& & C_{1_{\tau,\gamma}} = C_{2_{\tau,\gamma}} & & \\ C2 & \text{if} & h_x \neq v_\theta \text{ or } h_y \neq v_\beta & \& & C_{2_{\tau,\gamma}} = v_{\theta,\beta} & & \\ C2 & \text{if} & h_{x,y} \neq v_{\theta,\beta} & \& & C_{2_\tau} = v_\theta, \text{ or } C_{2_\gamma} = v_\beta, \text{ or both} & & \\ C2 & \text{if} & h_x \neq v_\theta = C_{2_\tau} & \& & h_y = v_\beta \neq C_{2_\gamma} & \& & \alpha < 1 \\ C2 & \text{if} & h_x = v_\theta & \& & C_{2_\gamma} \neq v_\beta & \& & \alpha < 1 \\ C2 & \text{if} & h_y = v_\beta & \& & C_{2_\tau} \neq v_\theta & \& & \alpha \geq 1 \end{cases} \tag{A.27}$$

This function shows which candidate the voter selects, given the combination of what the voter hears in each of the dimensions of the message, and how those align with C_2's types in both dimensions. The first line, then, says that voters will choose C_1 when they hear a message that says C_1 matches the voter's type on both dimensions, and C_2 is a mismatch on the ideological dimension ($C_{2_\tau} \neq v_\beta$). The second line indicates that voters will choose C_1 when they hear that same message in line 1, and C_2 differs from the voter on *both* dimensions. Third, voters will choose C_1 if the message they hear matches one, but not necessarily both, dimensions, while C_2 differs from the voter *at least* on the ideological dimension. This is, in effect, a weakening of the statement in the first line. It is spelled out here for clarity, and to make the derivation of the candidate utilities transparent. All the choices in this function thus indicate where matches (and mismatches) between the message heard and the incumbent C_2 produce votes for either C_1 or C_2.

A.5.3 C_1 Utilities

The utility to C_1 for sending a clear message (no whistling) is simply the sum of those voters that choose the challenger after hearing a particular message. Obviously, then, if C_1 uses a whistle, the candidate's share of voters depends on the proportion of each group that hear one part of the whistle or the other. Again there are numerous permutations of candidate types and possible messages, so I continue the previous example, then show the

Table 23: **Utilities for Each Type, Given Message and C_1's Strategy**

	$\alpha < 1$	$\alpha \geq 1$
$U_{C_{1_{M,O}}}(s_{m,o}) =$	$v_{M,O} + v_{M,I} = v_M$	v_M
$U_{C_{1_{M,I}}}(s_{m,i}) =$	$v_{M,O} + v_{M,I} = v_M$	$v_{M,I} + v_{E,I} = v_I$
$U_{C_{1_{E,O}}}(s_{e,w}) =$	Eq.(A.28)	Eq. (A.28)
$U_{C_{1_{E,O}}}(s_{e,o}) =$	$\frac{1}{2}$	$\frac{1}{2}$
$U_{C_{1_{E,I}}}(s_{e,i}) =$	$v_{M,I} + v_{E,I} = v_I$	v_I

general result. Again, set $C_{1_{E,O}} = C_{2_{E,O}}$, and let $\sigma(C_1) : \{s_{x,y} = \{e, w\}$ if $C_{1_{E,O}}$ and $s_{x,y} = \{\tau, \gamma\}$ otherwise$\}$. In that case $C_{1_{E,O}}$ wins those voters that prefer the message they hear over the clear message sent by $C_{2_{E,O}}$, and splits those voters that are indifferent between the two. Thus, by sending $s_{e,w}$, $C_{1_{E,O}}$'s expected vote share, minus the cost of the whistle, is:[2]

$$
\begin{aligned}
\tfrac{1}{2}(1-q)v_{M,O} + \tfrac{1}{2}(1-q)v_{E,O} + \tfrac{1}{2}(1-p)v_{M,I} + \tfrac{1}{2}(1-p)v_{E,I} + pv_{M,I} + pv_{E,I} - c \quad &\Rightarrow \\
\tfrac{1}{2}(1-q)(v_{M,O} + v_{E,O}) + \tfrac{1}{2}(1-p)(v_{M,I} + v_{E,I}) + p(v_{M,I} + v_{E,I}) - c \quad &\Rightarrow \\
\tfrac{1}{2}(1-q)v_O + \tfrac{1}{2}(1-p)v_I + pv_I \Rightarrow& \\
\tfrac{1}{2}(1-q)v_O + \tfrac{1}{2}(1+p)v_I - c&
\end{aligned}
\tag{A.28}
$$

Evaluating Table 22 for all voter choices produces Table 23, which shows the utilities for each type of C_1 under this strategy.

Comparing the utilities in Table 23 yields the constraints for supporting the proposed equilibrium. For example, $C_{1_{E,O}}$ prefers sending $s_{e,w}$ to sending $s_{e,o}$ when $\frac{1}{2}(1-q)v_O + \frac{1}{2}(1+p)v_I - c \geq \frac{1}{2} \Rightarrow c < \frac{pv_I - qv_O}{2}$. To guarantee $C_{1_{E,O}}$ does not switch to message $s_{e,i}$ (lying on the group dimension, yielding the utility noted for this message in Table 23, minus λ), $\frac{pv_I - qv_O}{2} > v_I - \lambda \Rightarrow \lambda > \frac{(2-p)v_I - qv_O}{2} > c$.

[2]Recall that the space for voter types is partitioned into mutually exclusive groups, and so can be summed across ideology or group identity: $v_{M,O} + v_{E,O} = v_O$, $v_{M,I} + v_{M,O} = v_M$, and so on

Table 24: **Constraints for c and λ to Support Proposed Equilibrium**

Candidate Message	Conditions			
$C_{1_{M,O}}(s_{m,o})$	$\lambda > \frac{1}{2} - v_M$,	$\lambda > v_{E,I} - v_{M,O}$,	$\lambda > \frac{v_{E,I} - v_{M,O}}{2}$,	$c > pv_{E,I} - qv_{M,O}$
$C_{1_{M,I}}(s_{m,i})$	$\lambda > \frac{1}{2} - v_M$,	$\lambda > v_{E,O} - v_{M,I}$,	$\lambda > \frac{v_{E,O} - v_{M,I}}{2}$,	$c < \frac{pv_I - qv_O}{2}$
$C_{1_{E,O}}(s_{e,w})$	$c < pv_{M,I} + pv_{E,I}$			
$C_{1_{E,I}}(s_{e,i})$	$\lambda > \frac{1}{2} - v_E$,	$\lambda > v_{M,O} - v_{E,I}$,	$\lambda > \frac{v_{M,O} - v_{E,I}}{2}$,	$c > qv_{M,O} - pv_{E,I}$

In addition to $C_{1_{E,O}}$ sending a whistle, the other types must prefer to not use a whistle. Consider $C_{1_{M,O}}$ sending $s_{m,w}$ in the proposed strategy. As with lying, voter beliefs are such that they do not question the message they hear. $C_{1_{M,O}}$'s utility is then:

$$\begin{aligned} U_{C_{1_{M,O}}}(s_{m,w}) &= v_{M,O} + v_{M,I} - c \text{ if } \alpha \geq 1 \\ U_{C_{1_{M,O}}}(s_{m,w}) &= (1-q)v_{M,O} + v_{M,I} + pv_{E,I} - c \text{ if } \alpha < 1 \end{aligned} \tag{A.29}$$

Clearly, if $\alpha \geq 1$ then $C_{1_{M,O}}$ will never use the dog whistle since it yields a strictly lower utility. When $\alpha < 1$, and voters prioritize the ideological dimension over the group dimension, solving (A.29) for c yields the cost constraint $c > pv_{E,I} - qv_{M,O}$, which prevents deviation by the other types to whistling.

Repeating this process for the other types of C_1 when $\sigma(C_1) : \{s_{e,w}$ if $C_{1_{E,O}}$ and $s_{x,y} = \{\tau, \gamma\}$ otherwise$\}$, produces a set of constraints for each type that all must be true (in addition to the model primitives) to support the strategy in equilibrium. These conditions are presented in Table 24. While there are 13 separate conditions, they do not preclude finding a level of c and λ that satisfy them all. For example, let $v_{\theta\beta}$ be distributed uniformly, so that $v_{M,O} = v_{M,I} + v_{E,O} = v_{E,I} = 0.25$, and let $p = 0.75$ and $q = 0.5$. Then, by substituting values in to the constraints, any $0.0625 < c < 0.375$ and $0 < c < \lambda$ suffices to support the equilibrium.

The proposed equilibrium requires a strategy chosen when $C_{1_{\tau,\gamma}} = C_{2_{\tau,\gamma}}$, which has just been shown to hold for $C_{1_{E,O}} = C_{2_{E,O}}$. The process for the remaining three "matching" cases is the same as above. The symmetry in the results for the constraints can be seen in Table 24. Rather than present three more tables of constraints, these are generalized

Table 25: **Constraints for c and λ to Support Proposed Equilibrium**

Candidate Types	Message	Conditions
$C_{1_{\tau,\gamma}} = C_{2_{\tau,\gamma}}$	$C_{1_{\tau,\gamma}}(s_{\tau,w})$	$c < Pr(h_y = i \mid \beta) \times v_{C_{1_{\neg\gamma}}}$
$C_{1_{\tau,\gamma}} \neq C_{2_{\tau,\gamma}}$	$C_{1_{\tau,\gamma}}(s_{\tau,\gamma})$	$\lambda > \frac{1}{2} - v_{C_{1_{\tau}}}$, $\lambda > v_{C_{1_{\neg\tau,\neg\gamma}}} - v_{C_{1_{\tau,\gamma}}}$, $c > Pr(h_y = i \mid \beta) \times (v_{C_{1_{\neg\tau,\neg\gamma}}} + v_{C_{1_{\tau,\gamma}}})$

and presented in Table 25, delineating when the voter types match or mismatch with C_1's types. The entries in the table give the ways in which candidates may match each other (or not), then the message C_1 chooses, then a general statement on the conditions required. For instance, the first line indicates that when $C_{1_{\tau,\gamma}} = C_{2_{\tau,\gamma}}$, and C_1 sends a dog whistle, then the cost of the whistle must be less than the probability that each group type of voter hears the message h_y indicates the candidate is claiming "in-group" status, times the size of the voter group that is the **opposite** of C_1's group type. In words, the cost of the whistle must be less than the size of the vote share C_1 gains by inducing uncertainty in voters of a different group type, drawing those voters to her side. Note also that because $\lambda > c$ by assumption, this condition on c is sufficient.

Therefore, there exist equilibria when C_2's type is fixed and known, C_1 adopts the strategy

$$\sigma_{C_1}(C_{2_{\tau,\gamma}}) = \begin{cases} m = \{\tau, w\} & \text{if} \quad C_{1_{\tau,\gamma}} = C_{2_{\tau,\gamma}} \\ m = \{\tau, \gamma\} & \text{otherwise} \end{cases}, \tag{A.30}$$

voter beliefs are as noted in Equation (A.26), voter actions are given by the best responses in (A.27), α takes any value, and the constraints listed in Table 25 are satisfied.

A.6 PROOF OF PROPOSITION 5: NO PARTIAL OR FULL POOLING DOG WHISTLING

There are three permutations of partial and full pooling. Types of $C_{1_{\tau,\gamma}}$ may pool by ideological type, by group type, or fully pool on both dimensions. None of these permutations are an equilibrium. The result for each of the three permutations is analytically similar, so I group them in the paper. I work through the three cases in turn below.

A.6.0.1 No Pooling by Ideological Type No equilibrium exists where each ideological type sends a whistle for their group membership, regardless of their true type.

A.6.1 C_1 Strategy

Specifically, the candidate strategy for C_1 is:

$$\sigma_{C_1}(C_{1_{\tau,\gamma}}): \quad s_{\tau,w} \quad \forall \tau \in \{M, E\} \tag{A.31}$$

The strategy indicates that both Extremist and Moderate types of C_1 will send a message revealing their true ideological type, and will always whistle their group type.

A.6.2 Voter Beliefs and Best Responses

Voters update beliefs in the standard Bayesian manner, where possible. Given C_1's strategy, voters believe the ideological portion of the message, and learn nothing from the group portion of the message: their posterior beliefs about group type are equal to their prior beliefs. As an example, the posterior beliefs for any $v_{M,O}(C_{1_{\tau,\gamma}} = \{M, O\} \mid h_{m,o}) = \frac{tg(1-q)}{(tg(1-q))+(t(1-g)(1-q))} = g$.

Voters always cast a vote for the candidate that maximizes their expected utility. This is also the candidate that has the highest posterior probability of matching their type. Voter utilities are calculated for each permutation of possible message heard and types of C_2. Table 26 shows these values for the case with $C_{2_{E,O}}$.

Table 26: Expected Utilities for Voter Choice, Given Group-Based Dog Whistling

Candidates	Voter Types			
	$v_{M,O}$	$v_{M,I}$	$v_{E,O}$	$v_{E,I}$
$EU(C1 \mid h_{m,o}) =$	$-\mu^I\alpha$	$-\mu^O\alpha$	$\mu^O(-1)+\mu^I(-1-\alpha)$	$\mu^O(-1-\alpha)+\mu^I(-1)$
$EU(C1 \mid h_{m,i}) =$	$-\mu^I\alpha$	$-\mu^O\alpha$	$\mu^O(-1)+\mu^I(-1-\alpha)$	$\mu^O(-1-\alpha)+\mu^I(-1)$
$EU(C1 \mid h_{e,o}) =$	$\mu^O(-1)+\mu^I(-1-\alpha)$	$\mu^O(-1-\alpha)+\mu^I(-1)$	$-\mu^I\alpha$	$-\mu^O\alpha$
$EU(C1 \mid h_{e,i}) =$	$\mu^O(-1)+\mu^I(-1-\alpha)$	$\mu^O(-1-\alpha)+\mu^I(-1)$	$-\mu^I\alpha$	$-\mu^O\alpha$
$EU(C_{2_{M,O}}) =$	0	$-\alpha$	-1	$-1-\alpha$
$EU(C_{2_{M,I}}) =$	$-\alpha$	0	$-1-\alpha$	-1
$EU(C_{2_{E,O}}) =$	-1	$-1-\alpha$	0	$-\alpha$
$EU(C_{2_{E,I}}) =$	$-1-\alpha$	-1	$-\alpha$	0

The remaining cases are calculated in the same fashion. Together, they result in the following best response function for the voter:

$$BR_{V_{\theta,\beta}}(C_{2_{\tau,\gamma}}, \mu, \mu) = \begin{cases} C_1 & \text{if} \quad h_x = C_{1_\tau} = C_{2_\tau},\ h_y \neq C_{2_\gamma} \\ C_1 & \text{if} \quad h_x = C_{1_\tau} \neq C_{2_\tau},\ h_y = C_{2_\gamma} \quad \& \quad \mu^I/\mu^O < \frac{1}{\alpha} \\ C_1 & \text{if} \quad h_x \neq v_\theta = C_{2_\tau} \quad \& \quad v_\beta \neq C_{2_\gamma} \\ C_1 & \text{if} \quad h_x = C_{1_\tau} \neq C_{2_\tau} \quad \& \quad \mu^I/\mu^O < \frac{1}{\alpha} \\ C_1 & \text{if} \quad v_\theta \neq C_{2_\tau} \quad \& \quad v_\beta \neq C_{2_\gamma} \\ C2 & \text{otherwise} \end{cases} \tag{A.32}$$

A.6.3 C_1 Utilities

From Table 26 I construct Table 27, showing which voters would opt for C_1, given the message they hear. This table shows immediately that there is no equilibrium where C_1 uses the strategy stated in A.31: sending a dog whistle yields the same result as a clear message, but suffers a cost of λ. In addition, given the assumption $c < \lambda$, C_1 would not opt to lie on the ideological dimension (again, given the strategy proposed).

Thus, there is no equilibrium where C_1 pools by ideological type on sending a dog whistle.

Table 27: C_1 **Utilities For Clear Messages and Dog Whistling**

Message	Candidate 2 Type				Condition
	{M,O}	{M,I}	{E,O}	{E,I}	
$U_{C1_{M,O}}(s_{m,o}) =$	v_I	$v_O - \lambda$	$v_M - \lambda$	$v_M - 2\lambda$	$\alpha < \frac{1}{1-g}$
$U_{C1_{M,I}}(s_{m,i}) =$	$v_I - \lambda$	v_O	$v_M - 2\lambda$	$v_M - \lambda$	$\alpha < \frac{1}{g}$
$U_{C1_{M,\gamma}}(s_{m,w}) =$	$v_I - c$	$v_O - c$	$v_M - c$	$v_M - c$	
$U_{C1_{E,O}}(s_{e,o}) =$	$v_E - \lambda$	$v_E - 2\lambda$	v_I	$v_O - \lambda$	$\alpha < \frac{1}{1-g}$
$U_{C1_{E,I}}(s_{e,i}) =$	$v_E - 2\lambda$	$v_E - \lambda$	$v_I - \lambda$	v_O	$\alpha < \frac{1}{g}$
$U_{C1_{E,\gamma}}(s_{e,w}) =$	$v_E - c$	$v_E - c$	$v_I - c$	$v_O - c$	

A.6.4 No Pooling By Group Type

The analysis this Proposition is identical to the analysis for ideological types, substituting the group-based indexing for the ideology-based indexing. There is no equilibrium where C_1 pools by group type on sending a dog whistle.

A.6.5 No Full Pooling

All types of C_1 whistling on both dimensions means voters are uncertain about both group types, learn nothing from the messages, and thus use their prior beliefs to select which candidate to vote for. No value is gained from the whistle, and again there is a loss of λ. Thus, there is no equilibrium where all types of C_1 send $s_{\tau,w}$.

APPENDIX B

BLAME AND CONCESSION

B.1 TIME-BASED AND PUNISHMENT-BASED LEARNING EFFECTS

Subjects may learn to play the Staring Game differently during the course of the experiment. If this occurs, and it affects the player's choices in ways that undermine our hypotheses, we must correct for them in the design of the experiment. In the case of time-based effects, we might be concerned that players may alter their strategy as they improve. Earlier rounds would then represent a different level of information than later rounds, since players have deduced the "correct" way to play the staring game. The second issue, punishment-based learning effects, suggests subjects play in one fashion until they experience punishment themselves, then play differently afterwards. The subjects learn, in other words, to be more punitive only when they have had points deducted. In this appendix we address two sources of concern – time-based effects and punishment-based effects – and show that they are not a problem.

B.1.1 Time-Based Learning Effects

Subjects that choose to concede do so more frequently during the last five seconds of the Staring Game as the number of rounds increases. While they appear to understand the basic game at the outset, after the first rounds they prefer waiting to concede near the end to conceding early.

Table 28 shows several probit models of this behavior. For each model, a player's choice to concede was categorized by either the exact or approximate time of concession. In the first column, subjects were classified as stopping at "Any Extreme" if their choice to concede was at $t < 5$ or $t > 25$. As noted in the main text, looking at time beyond the very first and last second accounts for differing subject response times. The significant coefficient shows that there is an increased probability to concede at one extreme or another as the number of periods increases. The next models examine whether the concessions are happening at one or both of these extremes. Columns 2 and 3 captures those rounds where players concede either within the first five seconds ($t = 0...4$) or just the first second, respectively. Clearly, concession at these times is not happening more frequently in later rounds. Columns 4 and 5 perform a similar analysis for stopping in the last five seconds ($t = 26...30$) and the last second, respectively. As the number of periods grows, players are more likely to concede in the last five seconds, though not at the very last second.

The presence of a time-based learning effect does not alter our substantive findings. Our treatment – the introduction of observers – occurs after a number of rounds are played, so that any effects present are constant for a player and stable during treatment. In addition, subjects learn to play according to the predictions of the Staring Game model.

B.1.2 Punishment-Based Learning Effects

Concerns about punishment-based effects are not warranted for this experiment. To show this, we categorized each round for each subject as being before or after the first round in which they received a deduction, then compared the use of punishment within subjects. Table 29 shows the average deductions taken by observers, first in those rounds before they received any punishment themselves, then after. T-tests show Punishment use is

Table 28: **Time-Based Learning Effects**

	(1) Any Extreme	(2) First 5 Seconds	(3) First Second	(4) Last 5 Seconds	(5) Last Second
Period	0.11***	0.03	0.05	0.09**	-0.06
	(0.02)	(0.03)	(0.05)	(0.03)	(0.03)
Constant	-0.31	-1.01***	-2.41***	-0.94***	1.99***
	(0.18)	(0.23)	(0.44)	(0.21)	(0.31)
$\ln(\sigma_u^2)$	-1.49**	-0.53	0.11	-0.92*	-1.07
	(0.49)	(0.37)	(0.52)	(0.38)	(0.68)
N (rounds)	398	398	398	398	398
N (subjects)	72	72	72	72	72

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

statistically equal in these two conditions, and across the targets of punishment ($p < 0.2$ for Player A, and $p < 0.9$ for Player B).

Table 29: **Punishment Use Before and After Receiving First Deductions**

	Mean Punishment to A	**Mean Punishment to B**
Pre-Punishment	16.7	17.5
	(28.1)	(28.8)
Post-Punishment	21.2	19.9
	(32.8)	(32.8)

Standard deviations in parentheses.

APPENDIX C

INFORMATION TRANSMISSION

C.1 DESCRIPTIVE STATISTICS

C.1.1 Hearings Data and Dependent Variable

Table 30 displays the number of hearings used in the final data set for this paper. This is the total number of hearings that have agency representatives engaged directly with committee chairs that were available via the Government Printing Office website; this is *not* the total number of hearings held in each congress. As the table shows, the number of hearings in the data set grows over time. This is likely due to both an increase in the number of hearings being held as well as an increase in the number of hearings with transcripts being made publicly available. (Numerous downloaded files have no content beyond the list of members of the convening committee, and so are excluded from this analysis.)

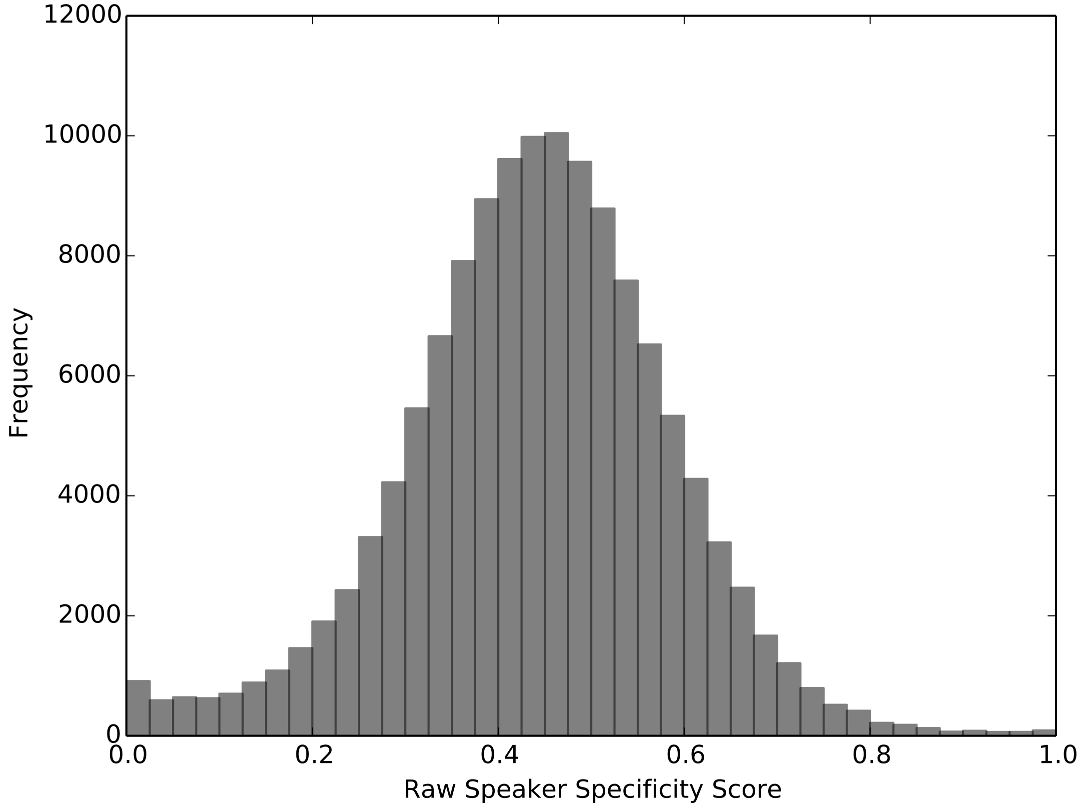
Figure 6 shows the distribution of the raw scores for the *Speaker Specificity* measure developed in the paper. Recall that this score is bounded above by 0 and below by 1. The distribution is normal, with a small spike just above 0. This occurs with speakers who are not the primary witness, but interjected to give greetings or very simple responses to questions.

Figure 7 shows the distribution of the main independent variable in the analysis: the Within-Congress Speaker Specificity value transformed into a z-score. The data show a

Table 30: **Count of Hearings in Data Set by Congress Number and Chamber**

Congress	Chamber	Count of Hearings
105	House	5028
	Senate	1785
106	House	7978
	Senate	3179
107	House	7814
	Senate	7537
108	House	6052
	Senate	6786
109	House	9522
	Senate	8007
110	House	14650
	Senate	8223
111	House	14695
	Senate	7642
112	House	15452
	Senate	6320

Figure 6: **Distribution of Raw *Speaker Specificity* Scores**



pronounced normality in distribution, probably owing to the large number of cases. The specificity measure is, itself, noisy, but this noise does not appear to be biased in one direction or another. The range of the variable is potentially infinite in both directions from 0, but empirically falls between -4 and 5.

C.1.2 Independent Variables

Table 31 shows descriptive statistics for the key independent variables in the paper. The first variable, *Chair Interactivity*, is derived from the textual hearing data. The other variables either directly stem from, or are derived from, the data sources noted in the paper.

C.2 MEASURING SPEAKER SPECIFICITY

This appendix provides a detailed description of the process of calculating speaker specificity. The basic algorithm is taken from Li and Nenkova (2015), which should be considered the primary source text, and should be referred to for further details. I reproduce some basic information here for completeness and to assist the reader in following the analysis. Moreover, some code used to develop the specificity scores were made available by the authors on their website as the Speciteller tool.[1]

The function of the Speciteller tool, as an implementation of the result published in Li and Nenkova (2015), is to take a single sentence and produce a measure of how general or specific the sentence is. To do this, the authors rely on a corpus of sentences from several sources (the Gigaword Corpus, data from Prasad et al. (2008), and the Penn Treebank) to serve as a reference for co-training a categorization algorithm. Features of the sentences are used as possible predictors: the number of words, the average number of characters in a word, named entities (proper nouns and related words), word polarity, the inverse document frequency of a word (rare words being more specific), and several others. These

[1] For those readers who cannot follow links from this document, the website is here: http://www.cis.upenn.edu/ nlp/software/speciteller.html.

Figure 7: **Distribution of Transformed *Speaker Specifity* Scores**

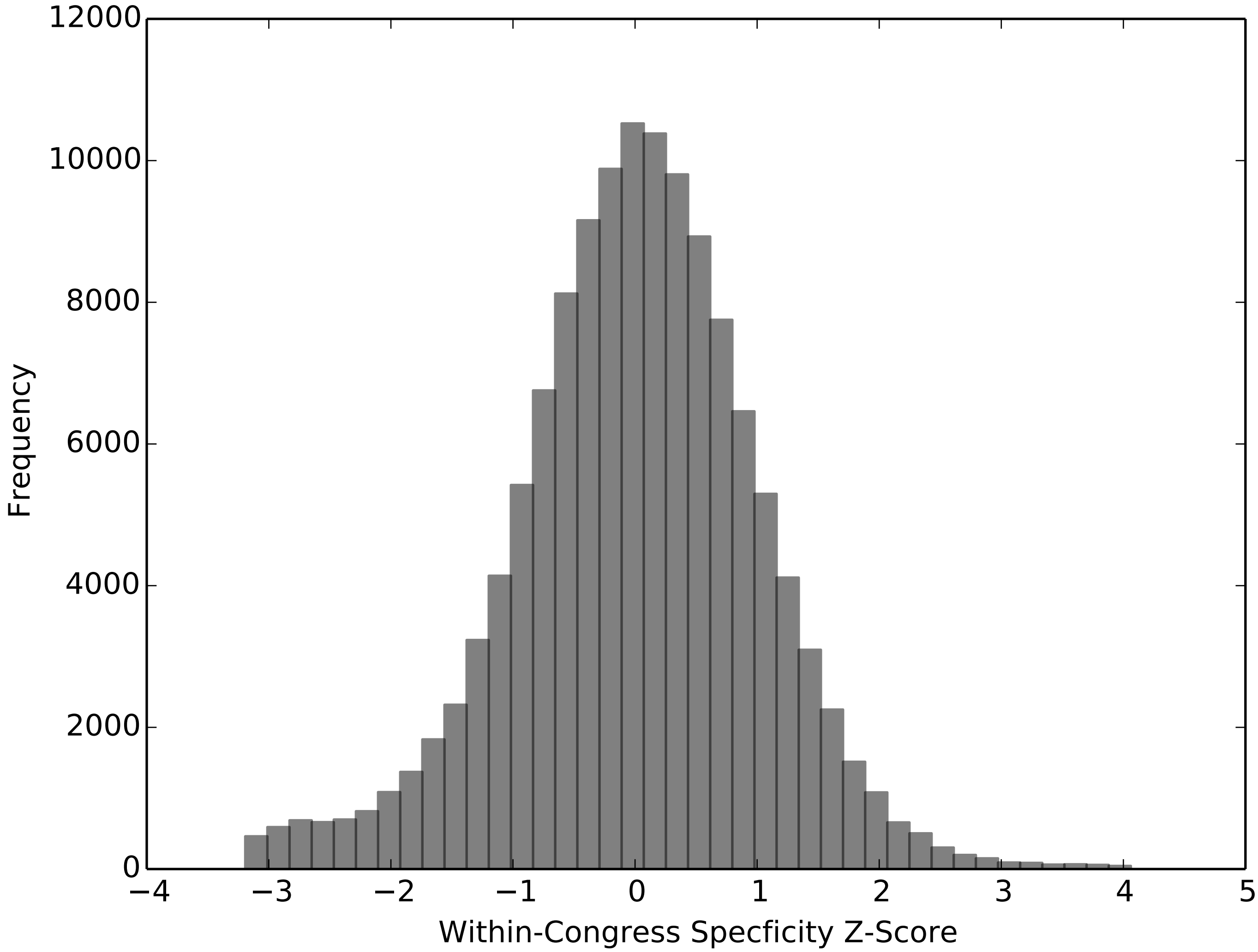


Table 31: **Summary Statistics for Key Independent Variables**

	Count	Mean	Std	Min	25%	50%	75%	Max
Chair Interactivity	130670	0.085	0.05	0.00	0.05	0.08	0.11	1.00
Ideological Distance	128529	0.39	0.22	0.00	0.22	0.39	0.55	1.19
Chair Seniority	130670	15.77	10.06	1.00	8.00	14.00	23.00	49.00
Majority Strength	130670	26.06	23.86	0	9	19	35	78

predictors are used in a logistic regression to classify them as either general or specific. The results of the regressions are compared against the actual human codings (that is, data is co-trained in semi-supervised process) to arrive at the best predictive model. In this way, a single sentence can be assigned a score that is the posterior probability for the specificity of a given sentence, and can be read as the probability of this sentence being classified as specific is....

I then apply what Li and Nenkova (2015) note as Equation 1 to extend this classification process to a whole set of sentences. The full process for arriving at a specificity score for a single speaker is as follows. First, I parse apart a hearing according to the words spoken by a single speaker. Given the interactive nature of hearings, a single person will have associated text in multiple places. Fortunately, hearing text follows a set format for identifying a speaker: an indented honorific, then a last name (sometimes hyphenated), followed by a period and a blank space. For example, here is an exchange from legislative hearing before the Subcommittee On National Parks, Recreation, And Public Lands Of The Committee On Resources, U.S. House Of Representatives:

> Mr. Roemer. Thank you, Mr. Chairman. I ask unanimous consent to have my entire statement entered into the record along with facsimiles and explanations of historic documents that we have from the Library of Congress.
> Mr. Hefley. Without objection.
> Mr. Roemer. Mr. Chairman, let me appropriately quote John Adams to start my testimony[...]

From this exchange, using regular expressions in the Python programming language, I can collect all the words spoken by, in this example, Mr. Roemer. I further split these statements by ending punctuation (periods or ellipses) to get a set of all the sentences – call this set p, and each sentence s – spoken by one witness. From this set, I can use a modification of Equation 1 in Li and Nenkova (2015)[2]:

$$Specificity(p) = \frac{1}{\sum_{s \in (p)} |s|} \sum_{s \in (p)} |s| \times Pr(\text{specific}|s) \tag{C.1}$$

The $Specificity(p)$ is the score for each speaker I use as my dependent variable. With each sentence denoted s, the length in words of each sentence is $|s|$. Finally, $Pr(\text{specific}|s)$ is the posterior probability calculated by the Speciteller process. Each speaker thus receives a score that indicates the posterior probability that their whole testimony would be considered specific or general, taking into account each sentence spoken, and normalized by the amount the speaker spoke. This makes it possible to directly compare speakers who give extensive testimony with those who give very little.

[2]The modification eliminates the process of comparing similar sentences in order to assess performance of their process. The difference in the equation is dropping a subscript i for set p that indicates a pair, the elements of which are to be used in the comparison.

APPENDIX D

BIBLIOGRAPHY

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

Abreu, Dilip and Faruk Gul (2000). "Bargaining and Reputation." In: *Econometrica* 68.1, pp. 85–117.

Albertson, Bethany L (2014). "Dog-Whistle Politics: Multivocal Communication and Religious Appeals." In: *Political Behavior*, pp. 1–24.

Alcañiz, Isabella and Timothy Hellwig (2011). "Who's to Blame? The Distribution of Responsibility in Developing Democracies." In: *British Journal of Political Science* 41.2, pp. 389–411.

Alesina, Alberto and Alex Cukierman (1990). "The Politics of Ambiguity." In: *Quarterly Journal of Economics* 105.4, pp. 829–850.

Alicke, Mark D (2000). "Culpable Control and the Psychology of Blame." In: *Psychological Bulletin* 126.4, pp. 556–574.

Banas, Joseph T and Judi M Parks (2002). "Lambs Among Lions? The Impact of Ethical Ideology on Negotiation Behaviors and Outcomes." In: *International Negotiation* 7.2, pp. 235–260.

Bartels, Larry (1996). "Uninformed Voters: Information Effects in Presidential Elections." In: *American Journal of Political Science* 40, pp. 194–230.

Bartels, Larry M (1986). "Issue Voting Under Uncertainty: An Empirical Test." In: *American Journal of Political Science* 30.4, pp. 709–728.

Battaglini, Marco (2002). "Multiple Referrals and Multidimensional Cheap Talk." In: *Econometrica* 70.4, pp. 1379–1401.

Bendor, Jonathan, Amihai Glazer, and Thomas H. Hammond (2001). "Theories of Delegation." In: *Annual Review of Political Science* 4, pp. 235–269.

Bendor, Jonathan and Adam Meirowitz (2004). "Spatial Models of Delegation." In: *American Political Science Review* 98.2, pp. 1–19.

Bengtsson, Åsa (2004). "Economic Voting: The Effect of Political Context, Volatility and Turnout on Voters' Assignment of Responsibility." In: *European Journal of Political Research* 43.5, pp. 749–767.

Blume, Andreas and Oliver Board (2009). "Intentional vagueness." In: *Erkenntnis*, pp. 1–45.

— (2014). "Intentional vagueness." In: *Erkenntnis* 79.4, pp. 855–899.

Bradley, Robert B (1980). "Motivations in Legislative Information Use." In: *Legislative studies quarterly*, pp. 393–406.

Brasher, Holly (2006). "Listening to Hearings Legislative Hearings and Legislative Outcomes." In: *American Politics Research* 34.5, pp. 583–604.

Brehm, John and Scott Gates (1997). *Working, Shirking, and Sabotage: Bureaucratic Response to a Democratic Public*. Ann Arbor, MI: University of Michigan Press.

Brewin, Chris R and David A Shapiro (1984). "Beyond Locus of Control: Attribution of Responsibility for Positive and Negative Outcomes." In: *British Journal of Psychology* 75.1, pp. 43–49.

Brown, Adam R (2010). "Are Governors Responsible for the State Economy? Partisanship, Blame, and Divided Federalism." In: *The Journal of Politics* 72.03, pp. 605–615. ISSN: 0022-3816. DOI: 10.1017/S0022381610000046.

Bush, George W (2003). *Text of President Bush's 2003 State of the Union Address*. URL: http://www.washingtonpost.com/wp-srv/onpolitics/transcripts/bushtext%5C_012803.html.

Cadsby, C Bram, Maroš Servátka, and Fei Song (2010). "Gender and Generosity: Does Degree of Anonymity or Group Gender Composition Matter?" In: *Experimental Economics* 13.3, pp. 299–308.

Callander, Steven (2007). "Bandwagons and Momentum in Sequential Voting." In: *The Review of Economic Studies* 74.3, pp. 653–684. ISSN: 00346527.

Calvert, Randall (2013). *Models of Imperfect Information in Politics*. Vol. 6. Taylor & Francis.

Campbell, Angus (1964). "Issues and Voters: Past and Present." In: *Journal of Politics* 26, pp. 745–757.

Campbell, Angus et al. (1960). *The American Voter*. New York: University Of Chicago Press, pp. -–-. ISBN: 0226092542.

Canes-Wrone, Brandice (2010). *Who Leads Whom?: Presidents, Policy, and the Public*. University of Chicago Press.

Canes-Wrone, Brandice, Michael C Herron, and Kenneth W Shotts (2001). "Leadership and Pandering: A Theory of Executive Policymaking." In: *American Journal of Political Science* 45.3, pp. 532–550.

Cannon, Lou (2005). *Governor Reagan: His Rise to Power*. Vol. 1. Public Affairs.

Carlsson, Fredrik, Dinky Daruvala, and Olof Johansson-Stenman (2005). "Are People Inequality-Averse, or Just Risk-Averse?" In: *Economica* 72.287, pp. 375–396.

Carmines, Edward G et al. (1980). "The Two Faces of Issue Voting." In: *The American Political Science Review* 74.1, pp. 78–91. ISSN: 00030554. URL: http://www.jstor.org/stable/1955648.

Carpenter, Daniel P (2001). *The Forging of Bureacuratic Autonomy: Reputations, Networks, and Policy Innovation in Executive Agencies, 1862-1928*. Princeton, NJ: Princeton University Press, p. 472. ISBN: 0691070105.

— (2010). *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the FDA*. Princeton, NJ: Princeton University Press.

Carpenter, Daniel P, Kevin M Esterling, and David M J Lazer (2004). "Friends, Brokers, and Transitivity: Who Informs Whom in Washington Politics?" In: *The Journal of Politics* 66.1, pp. 224–246. ISSN: 00223816.

Carroll, Royce et al. (2008). "Common Space DW-NOMINATE Scores with Bootstrapped Standard Errors (Joint House and Senate Scaling)." In: *Accessed via http://voteview.ucsd. edu/dwnomin_joint_house_and_senate. htm September* 28, p. 2008.

Celetani, Marco et al. (1996). "Maintaining a Reputation Against a Long-lived Opponent." In: *Econometrica: Journal of the Econometric Society* 64.3, pp. 691–704.

Chappell, Henry W (1994). "Campaign Advertising and Political Ambiguity." In: *Public Choice* 79.3/4, pp. 281–303.

Chaudhuri, Ananish (2011). "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature." In: *Experimental Economics* 14.1, pp. 47–83.

Chen, Jowei and Tim Johnson (2015). "Federal Employee Unionization and Presidential Control of the Bureaucracy: Estimating and Explaining Ideological Change in Executive Agencies." In: *Journal of Theoretical Politics* 27.1, pp. 151–174.

Chong, Dennis and James N. Druckman (2007). "Framing Theory." In: *Annual Review of Political Science* 10.1, pp. 103–126. ISSN: 1094-2939.

Chong, Dennis and James N Druckman (2013). "Counterframing Effects." In: *The Journal of Politics* 75.01, pp. 1–16.

Clark, Herbert H and Gregory L Murphy (1982). "Audience Design in Meaning and Reference." In: *Advances in Psychology* 9, pp. 287–299.

Clinton, J. D. and D. E. Lewis (2007). "Expert Opinion, Agency Characteristics, and Agency Preferences." In: *Political Analysis* 16.1, pp. 3–20. ISSN: 1047-1987. DOI: 10.1093/pan/mpm009. URL: http://pan.oxfordjournals.org/cgi/doi/10.1093/pan/mpm009.

Cobb, Michael D and James H Kuklinski (1997). "Changing minds: Political Arguments and Political Persuasion." In: *American Journal of Political Science*, pp. 88–121.

Cohen, Jeffrey E (1995). "Presidential rhetoric and the public agenda." In: *American Journal of Political Science* 39.1, pp. 87–107.

Crawford, B Y Vincent P and Joel Sobel (1982). "Strategic Information Transmission." In: *Econometrica* 50.6, pp. 1431–1451.

Croson, Rachel and Nancy Buchan (1999). "Gender and Culture: International Experimental Evidence from Trust Games." In: *American Economic Review*, pp. 386–391.

De Vries, Catherine E and Sara B Hobolt (2012). "Do Voters Blame Governments for Social Spending Cuts? Evidence from a Natural Experiment." In:

Degregorio, Christine (1992). “Leadership Approaches in Congressional Committee Hearings.” In: *Western Political Quarterly* 45.4, pp. 971–983.

Del Sesto, Steven L (1980). “Nuclear Reactor Safety and the Role of the Congressman: A Content Analysis of Congressional Hearings.” In: *Journal of Politics* 42.1, pp. 227–241.

Desmarais, Andrew (2012). *The Republican Party and the art of the dog whistle*. Denver, CO.

Dickhaut, John W, Kevin A McCabe, and Arijit Mukherji (1995). “An Experimental Study of Strategic Information Transmission.” In: *Economic Theory* 6.3, pp. 389–403.

Diermeier, Daniel and Timothy J Feddersen (2000). “Information and Congressional Hearings.” In: *American Journal of Political Science* 44.1, pp. 51–65.

Drew, Elizabeth (1997). *Showdown: The Struggle Between the Gingrich Congress and the Clinton White House*. Touchstone, p. 400. ISBN: 0684825511.

Druckman, James N (2001a). “On the Limits of Framing Effects: Who Can Frame?” In: *Journal of Politics* 23.3, pp. 225–256.

— (2001b). “The Implications of Framing Effects for Citizen Competence.” In: *Political Behavior* 23.3, pp. 225–256.

Edelman, Murray (1985). “Political Language and Political Reality.” In: *PS* 18.1, pp. 10–19.

Edsall, Thomas Byrne, Mary D Edsall, and Mary D. (1991). *Chain Reaction: The Impact of Race, Rights, and Taxes on American Politics*. New York: Norton.

Edwards, George C and B Dan Wood (1999). “Who influences whom? The president, Congress, and the media.” In: *American Political Science Review* 93.02, pp. 327–344.

Embrey, Matthew, Guillaume R Fréchette, and Steven F Lehrer (2014). “Bargaining and Reputation: An Experiment on Bargaining in the Presence of Behavioural Types.” In: *The Review of Economic Studies*, pp. –29.

Engelmann, Dirk and Martin Strobel (2004). “Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments.” In: *American Economic Review*, pp. 857–869.

Epstein, David and Sharyn O'Halloran (1999). *Delegating Powers: {A} Transaction Cost Politics Approach to Policy Making Under Separate Powers*. Cambridge University Press.

Erikson, Robert S and Gerald C Wright (1993). "Voters, Candidates and Issues in Congressional Elections." In: *Congress Reconsidered, 6th Ed.* Ed. by Lawrence C Dodd and Bruce I Oppenheimer. Washington, DC: Congressional Quarterly Press, pp. 91–116.

Esterling, Kevin M (2004). *The Political Economy of Expertise: Information and Efficiency in American National Politics*. Ann Arbor, MI: University of Michigan Press.

Farrell, Joseph, Robert Gibbons, et al. (1989). "Cheap Talk with Two Audiences." In: *The American Economic Review* 79.5, pp. 1214–1223.

Farrell, Joseph and Matthew Rabin (1996). "Cheap Talk." In: *Journal of Economic Perspectives* 10.3, pp. 103–118.

Fear, Josh (2007). *Under the Radar: Dog-whistle Politics in Australia*. Australia Institute.

Fearon, J D (1994). "Domestic Political Audiences and the Escalation of International Disputes." In: *American Political Science Review*, pp. 577–592.

Fehr, Ernst and Simon Gächter (2002). "Altruistic punishment in humans." In: *Nature* 415.6868, pp. 137–140.

Fehr, Ernst and Klaus M Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation." In: *The Quarterly Journal of Economics* 114.3, pp. 817–868.

Fischbacher, Urs (2007). "z-Tree: Zurich toolbox for ready-made economic experiments." In: *Experimental Economics* 10.2, pp. 171–178. ISSN: 1386-4157. DOI: 10.1007/s10683-006-9159-4. URL: http://dx.doi.org/10.1007/s10683-006-9159-4.

Fox, Justin and Kenneth W Shotts (2009). "Delegates or Trustees? A Theory of Political Accountability." In: *The Journal of Politics* 71.4, pp. 1225–1237. ISSN: 00223816.

Gailmard, Sean (2002). "Expertise, Subversion, and Bureaucratic Discretion." In: *Journal of Law, Economics, and Organization* 18.2, pp. 536–555. ISSN: 87566222.

— (2009). "Multiple Principals and Oversight of Bureaucratic Policy-Making." In: *Journal of Theoretical Politics* 21.2, pp. 161–186. ISSN: 0951-6298. DOI: 10.1177/0951629808100762.

Gailmard, Sean and John W. Patty (2013). *Learning While Governing: Information, Accountability, and Executive Branch Institutions*. Chicago, IL: University of Chicago Press, pp. 1–367.

Garfield, James A. (1873). *Gen. J A Garfield's Review Of The Transactions of the Credit Mobilièr Company*. Washington, D.C.: Pennsylvania State University Press.

Gates, Scott and Brian D Humes (1997). *Games, Information, and Politics: Applying Game Theoretic Models to Political Science*. University of Michigan Press.

Geertz, Clifford (1973). *The Interpretation of Cultures: Selected Essays*. Vol. 5019. Basic Books.

Gilens, Martin, Lynn Vavreck, and Martin Cohen (2007). "The Mass Media and the Public's Assessments of Presidential Candidates, 1952–2000." In: *The Journal of Politics* 69.4, pp. 1160–1175. ISSN: 0022-3816.

Gilligan, Thomas W and Keith Krehbiel (1989). "Asymmetric Information and Legislative Rules with a Heterogeneous Committee." In: *American Journal of Political Science* 33, pp. 459–490.

Glazer, Amihai (1990). "The Strategy of Candidate Ambiguity." In: *American Political Science Review* 84.1, pp. 237–241.

Gneezy, Uri (2005). "Deception: The Role of Consequences." In: *American Economic Review* 95, pp. 384–394.

Goeree, Jacob K and Charles A Holt (2000). "Asymmetric Inequality Aversion and Noisy Behavior in Alternating-offer Bargaining Games." In: *European Economic Review* 44.4, pp. 1079–1089.

Goodin, Robert E. and Michael Saward (2005). "Dog Whistles and Democratic Mandates." In: *The Political Quarterly* 76.4, pp. 471–476.

Grimmer, J and B M Stewart (2013). "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." In: *Political Analysis*, pp. 1–31. ISSN: 1047-1987.

Grimmer, Justin (2010). "Representational Style: The Central Role of Communication in Representation." PhD thesis. Harvard Univeristy.

Groseclose, Tim and Nolan McCarty (2001). “The Politics of Blame: Bargaining Before an Audience.” In: *American Journal of Political Science* 45.1, pp. 100–119.

Grossman, Michael Orlov and Ronald Eric Jr. Matthews, eds. (2009). *Perspectives on the Legacy of George W. Bush*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Guala, F (2012). “Reciprocity: Weak or strong? What Punishment Experiments Do (and Do Not) Demonstrate.” In: *Behavioral and Brain Sciences* 35.1, p. 1.

Haney-López, Ian (2014). *Dog Whistle Politics: How Coded Racial Appeals Have Wrecked the Middle Class*. Oxford University Press.

Hatfield, Elaine et al. (1978). *Equity: Theory and research.* Allyn & Bacon.

Hechter, Michael and Karl-Dieter Opp (2001). *Social Norms*. Russell Sage Foundation.

Hellwig, Timothy (2010). “Elections and the Economy.” In: *Comparing Democracies* 3, pp. 185–201.

Holm Håand Engseld, Peter kan (2005). “Choosing Bargaining Partners—An Experimental Study on the Impact of Information about Income, Status and Gender.” In: *Experimental Economics* 8.3, pp. 183–216.

Holmstrom, Bengt (1984). “On the Theory of Delegation.” In: *Bayesian Models in Economic Theory*. Ed. by Marcel Boyer and Richard Khilstrom. New York: North Holland.

Hood, Christopher (2010). *The Blame Game: Spin, Bureaucracy, and Self-Preservation in Government*. Princeton University Press.

Howard, Alton H (1994). *Songs of Faith & Praise, Gray Conventional*. Howard Books.

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

Hurkens, Sjaak and Navin Kartik (2008). “Would I Lie To You? On Social Preferences and Lying Aversion.” In: *Experimental Economics* 12.May 2006, pp. 180–192.

Hurwitz, Jon and Mark Peffley (2005). “Playing the Race Card in the Post-Willie Horton Era: The Impact of Racialized Code Words on Support for Punitive Crime Policy.” In: *Public Opinion* 69.1, pp. 99–112.

Iyengar, Shanto (1989). “How Citizens Think about National Issues: A Matter of Responsibility.” In: *American Journal of Political Science* 33.4, pp. 878–900.

Jay, Martin (2010). *The Virtues of Mendacity: on Lying in Politics*. University of Virginia Press.

Kahneman, D and A Tversky (2000). *Choices, Values and Frames*. Cambridge University Press.

Kartik, Navin (2009). "Strategic Communication with Lying Costs." In: *Review of Economic Studies* 76.4, pp. 1359–1395.

Kennelly, Arthur and Edmund Fantino (2007). "The Sharing Game: Fairness in Resource Allocation as a Function of Incentive, Gender, and Recipient Types." In: *Judgment and Decision Making* 2.3, pp. 204–216.

Kiewiet, D R and M D McCubbins (1991). *The logic of delegation: Congressional parties and the appropriations process*. University of Chicago Press.

Knight, Jack and Jean Ensminger (1998). *Conflict Over Changing Social Norms: Bargaining, Ideology, and Enforcement*. Stanford, CA: Stanford University Press, pp. 105–126.

Kousser, Thad and Justin H Phillips (2009). "Who blinks first? legislative patience and bargaining with governors." In: *Legislative Studies Quarterly* 34.1, pp. 55–86.

Kovenock, Dan and Brian Roberson (2009). "Is the 50-state strategy optimal?" In: *Journal of Theoretical Politics* 21.2, pp. 213–236.

Krehbiel, Keith (1990). "Are Congressional Committees Composed of Preference Outliers?" In: *American Political Science Review* 84.1, pp. 149–163.

— (1991). *Information and Legislative Organization*. Ann Arbor: University of Michigan Press. ISBN: 0472094602.

Kreps, David M, Paul Milgrom, et al. (1982). "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma." In: *Journal of Economic Theory* 27.2, pp. 245–252.

Kreps, David M and Robert Wilson (1982). "Reputation and Imperfect Information." In: *Journal of Economic Theory* 27.2, pp. 253–279.

Krishna, Vijay (2001). "Asymmetric Information and Legislative Rules: Some Amendments." In: *American Political Science Association*. Vol. 95. 02. Cambridge Univ Press, pp. 435–452.

Kuklinski, James H and Paul J Quirk (1998). "Reconsidering The Rational Public: Cognition, Heuristics, and Mass Opinion." In: *Elements of Political Reason*. Ed. by Arthur Lupia, Mathew McCubbins, and Samuel L Popkin. Cambridge University Press.

Kutler, Stanley I. (1992). *The Wars of Watergate: The Last Crisis of Richard Nixon*. WW Norton & Company.

Lagnado, David A and Shelley Channon (2008). "Judgments of Cause and Blame: the Effects of Intentionality and Foreseeability." In: *Cognition* 108.3, pp. 754–770. ISSN: 0010-0277. DOI: 10.1016/j.cognition.2008.06.009.

Lau, Richard R (1997). "Voting Correctly." In: *American Politcal Science Review* 91.3, pp. 585–598.

Lewis-Beck, Michael S, William G Jacoby, et al. (2008). *The American Voter Revisited*. University of Michigan Press, p. 493. ISBN: 0472050400.

Lewis-Beck, Michael S and Martin Paldam (2000). "Economic Voting: An Introduction." In: *Electoral Studies* 19.2, pp. 113–121.

Lewis-Beck, Michael S and Mary Stegmaier (2000). "Economic Determinants of Electoral Outcomes." In: *Annual Review of Political Science* 3.1, pp. 183–219.

Leyden, Kevin M (1992). "The Purpose and Politics of Congressional Committee Hearings: Who Participates and Why." In: *Ph.D. Dissertation*.

— (1995). "Interest Group Resources and Testimony at Congressional Hearings." In: *Legislative Studies Quarterly* 20.3, pp. 431–439.

Li, Junyi Jessy and Ani Nenkova (2015). "Fast and Accurate Prediction of Sentence Specificity." In: *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Lightle, John (2013). "Harmful Lie Aversion and Lie Discovery in Noisy Expert Advice Games." In: *Journal of Economic Behavior & Organization*.

Lipsky, Michael (1980). *Street-Level Bureaucracy*. Russell Sage Foundation.

López-Pérez, Raúl and Eli Spiegelman (2012). "Why do people tell the truth? Experimental evidence for pure lie aversion." In: *Experimental Economics*, pp. 1–15. ISSN: 1386-4157.

Lupia, Arthur, Mathew D McCubbins, and Samuel Popkin, eds. (2000). *Elements of Reason: Cognition, Choice, and the Bounds of Rationality*. New York, NY: Cambridge University Press.

Lupia, Arthur and Mathew Daniel McCubbins (1998). *The Democratic Dilemma: Can Citizens Learn What they Need to Know?* New York, NY: Cambridge University Press, p. 282. ISBN: 0521585937. URL: http://books.google.com/books?hl=en%5C&lr=%5C&id=2Vv6BhLC6HUC%5C&pgis=1.

Macey, Jonathan R (1992). "Organizational Design and Political Control of Administrative Agencies." In: *Journal of Law, Economics, & Organization* 8, pp. 93–110.

Malhotra, Neil and Alexander G Kuo (2007). "Attributing Blame: The Public's Response to Hurricane Katrina." In: *The Journal of Politics* 70.01, pp. 120–135. ISSN: 0022-3816. DOI: 10.1017/S0022381607080097.

Masclet, David et al. (2003). "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism." In: *American Economic Review* 93.1, pp. 366–380.

Mayhew, David R (1974). *Congress: The Electoral Connection*. New Haven: Yale University Press.

McCubbins, Mathew D., Roger G. Noll, and Barry R. Weingast (1987). "Administrative Procedures as Instruments of Political Control." In: *Journal of Law, Economics, & Organization*, pp. 243–277.

Mehta, Judith, Chris Starmer, and Robert Sugden (1994). "Focal Points in Pure Coordination Games: An Experimental Investigation." In: *Theory and Decision* 36.2, pp. 163–185.

Mendelberg, Tali (2001). *The Race Card: Campaign Strategy, Implicit Messages, and the Norm of Equality*. Princeton University Press, p. 307. ISBN: 0691070717.

Minozzi, William (2011). "A Jamming Theory of Politics." In: *The Journal of Politics* 73.02, pp. 301–315. ISSN: 1468-2508.

Minozzi, William and Jonathan Woon (2013). "Lying Aversion, Lobbying, and Context in a Strategic Communication Experiment." In: *Journal of Theoretical Politics* 25.3, pp. 309–337.

Monroe, B. L., M. P. Colaresi, and K. M. Quinn (2009). "Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict." In: *Political Analysis* 16.4, pp. 372–403.

Morton, Rebecca B. and Kenneth C. Williams (2010). *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge University Press.

Nelson, Thomas E (1999). "Group Affect and Attribution in Social Policy Opinion." In: *Journal of Politics* 61, pp. 331–362.

Nelson, Thomas E and Donald R Kinder (1996). "Issue Frames and Group-Centrism in American Public Opinion." In: *The Journal of Politics* 58.04, pp. 1055–1078. ISSN: 1468-2508.

O'Keefe, Garrett J and L Erwin Atwood (1981). "Communication and Election Campaigns." In: *Handbook of Political Communication*, pp. 329–357.

Oleszek, Walter J (2004). *Congressional Procedures and the Policy Process*. 5th. Washington, DC: CQ Press.

Olson, Mary K (1999). "Agency rulemaking, political influences, regulation, and industry compliance." In: *Journal of Law, Economics, and Organization* 15.3, pp. 573–601.

Overby, L Marvin and Thomas A Kazee (2000). "Outlying Committees in the Statehouse: An Examination of the Prevalence of Committee Outliers in State Legislatures." In: 62(3), pp. 701–728.

Page, Benjamin I (1976). "The Theory of Political Ambiguity." In: *American Political Science Review* 70.3, pp. 742–752.

Paldam, Martin (2008). "Vote and Popularity Functions." In: *Readings in Public Choice and Constitutional Political Economy*. Ed. by Charles Kershaw Rowley and Friedrich Schneider. New York: Springer, pp. 533–550.

Patty, John W. (2005). *The Strategic Spatial Calculus of Delegation and Discretion*. Typescript, Harvard University, Department of Government.

Peffley, Mark (1984). "The Voter as Juror: Attributing Responsibility forEconomic Conditions." In: *Political Behavior* 6.3, pp. 275–294. ISSN: 0190-9320. DOI: 10.1007/BF00989621.

Poole, Keith T and Howard Rosenthal (1994). "Congress and Railroad Regulation: 1874 to 1887." In: *The Regulated Economy: A Historical Approach to Political Economy*. Ed. by Claudia Dale Goldin and Gary D Libecap. Chicago, IL: University Of Chicago Press, pp. 81–120.

Poole, Keith and Howard Rosenthal (2007). *Ideology and {C}ongress*. Transaction Publishers.

Popkin, Samuel L (1991). *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.

Prasad, Rashmi et al. (2008). "The Penn Discourse TreeBank 2.0." In: *LREC*. Citeseer.

Putnam, Linda L and Tricia S Jones (1982). "Reciprocity in Negotiations: An Analysis of Bargaining Interaction." In: *Communication Monographs* 49.3, pp. 171–191.

Riker, William H (1986). *The Art of Political Manipulation*. New Haven, CT: Yale University Press.

Robinson, Michael J. and Kevin R. Appel (1979). "Network news coverage of Congress." In: *Political Science Quarterly*, pp. 407–418.

Roth, Alvin E, ed. (1985). *Game-Theoretic Models of Bargaining*. Cambridge: Cambridge Univ. Press. ISBN: 0521267579.

Sabatier, Paul and David Whiteman (1985). "Legislative Decision Making and Substantive Policy Information: Models of Information Flow." In: *Legislative Studies Quarterly*, pp. 395–421.

Sánchez-Pagés, Santiago and Marc Vorsatz (2007). "An Experimental Study of Truth-Telling in a Sender-Receiver Game." In: *Games and Economic Behavior* 61.1, pp. 86–112.

Schelling, Thomas C (1960). *The Strategy of Conflict*. Cambridge, MA.: Harvard University Press.

Schultz, K A (2001). "Looking for audience costs." In: *Journal of Conflict Resolution* 45.1, pp. 32–60.

Shaver, Kelly G (1985). *The Attribution of Blame: Causality, Responsibility, and Blameworthiness*. Springer-Verlag New York.

Shepsle, Kenneth A (1972). “The Strategy of Ambiguity: Uncertainty and Electoral Competition Author.” In: *American Political Science Review* 66.2, pp. 555–568.

Shipan, Charles R. (2005). “The Legislative Branch.” In: ed. by Paul J. Quirk and Sarah A. Binder. Oxford University Press. Chap. Congress and the Bureaucracy.

Sides, John and Lynn Vavreck (2014). *The gamble: Choice and chance in the 2012 presidential election*. Princeton University Press.

Smith, Rogers M (2013). “Identity Politics and the End of the Reagan Era.” In: *Politics, Groups, and Identities* 1.1, pp. 120–127.

Sniderman, Paul M and Sean M Theriault (2004). “The Structure of Political Argument and the Logic of Issue Framing.” In: *Studies in Public Opinion: Attitudes, Nonattitudes, Measurement Error, and Change*, pp. 133–165.

Sobel, Joel (2007). “Signaling Games.” In: *Encyclopedia of Complexity and System Science*. Ed. by Robert A. Meyers. Springer, pp. 1–28.

Spence, Michael (1973). “Job Market Signaling.” In: 87, pp. 355–374.

Splawn, Walter M W (1939). “Railroad Regulation by the Interstate Commerce Commission.” In: *The Annals of the American Academy of Political and Social Science* 201, pp. 152–164. ISSN: 0002-7162.

Stewart III, Charles and Jonathan Woon. *Congressional Committee Assignments, 103rd to 112th Congresses 1993–2011, Senate and House*. URL: http://web.mit.edu/17.251/www/data%5C_page.html.

— (2005). *Congressional Membership Data, House and Senate, 103rd-109th Congresses*. Massachusetts Institute of Technology.

Sugden, R (1995). “A Theory of Focal Points.” In: *The Economic Journal*, pp. 533–550.

Sutter, Matthias et al. (2009). “Gender Pairing and Bargaining—Beware the Same Sex!” In: *Experimental Economics* 12.3, pp. 318–331.

Swanson, Reid, Brian Ecker, and Marilyn Walker (2015). “Argument Mining: Extracting Arguments from Online Dialogue.” In: *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, p. 217.

Tomz, Michael (2007a). “Domestic Audience Costs in International Relations: An Experimental Approach.” In: *International Organization* 61.4, p. 821.

Tomz, Michael (2007b). *Reputation and International Cooperation: Soverign Debt across Three Centuries*. Princeton, NJ: Princeton University Press.

Tomz, Michael and Robert P Van Houweling (2009). "The Electoral Implications of Candidate Ambiguity." In: *American Political Science Review* 103.1, pp. 83–98.

Trachtenberg, Marc (2012). "Audience Costs: An Historical Analysis." In: *Security Studies* 21.1, pp. 3–42.

Treier, Shawn and D Sunshine Hillygus (2009). "The nature of political ideology in the contemporary electorate." In: *Public Opinion Quarterly* 73.4, pp. 679–703.

Truman, David B (1951). *The Governmental Process*. New York: Knopf.

Valentino, Nicholas A, Vincent L Hutchings, and Ismail K White (2002). "Cues that Matter: How Political Ads Prime Racial Attitudes During Campaigns." In: *American Political Science Review* 96.01, pp. 75–90. ISSN: 1537-5943.

Van De Ven, Jeroen and Marie Claire Villeval (2014). "Dishonesty under scrutiny." In: *Journal of the Economic Science Association*, pp. 1–14.

Van den Assem, Martijn J., Dennie Van Dolder, and Richard H. Thaler (2012). "Split or Steal? Cooperative Behavior when the Stakes are Large." In: *Management Science* 58.1, pp. 2–20.

Vavreck, Lynn (2009). *The Message Matters: The Economy and Presidential Campaigns*. Princeton University Press.

Vickers, John (1985). "Delegation and the Theory of the Firm." In: *Economic Journal* 95.380a, pp. 138–147.

Volden, Craig (2002). "A Formal Model of the Politics of Delegation in a Separation of Powers System." In: 46(1), pp. 111–133.

Wade-Benzoni, Kimberly A et al. (2002). "Barriers to Resolution in Ideologically Based Negotiations: The Role of Values and Institutions." In: *Academy of Management Review* 27.1, pp. 41–57.

Wawro, Gregory J and Eric Schickler (2006). *Filibuster: Obstruction and lawmaking in the US Senate*. Cambridge Univ Press.

Weale, Albert (2002). *Risk, Democratic Citizenship and Public Policy*. Vol. 2. Oxford: Oxford University Press.

Weeks, Jessica L (2008). "Autocratic Audience Costs: Regime Type and Signaling Resolve." In: *International Organization* 62.1, pp. 35–64.

Weiner, Bernard (1995). *Judgments of Responsibility: A Foundation For A Theory of Social Conduct*. New York: Guilford Press.

White, Theodore Harold (1975). *Breach of Faith: The Fall of Richard Nixon*. New York, NY: Atheneum.

Winter, Nicholas J G (2006). "Beyond Welfare: Framing and the Racialization of White Opinion on Social Security." In: *American Journal of Political Science* 50.2, pp. 400–420. ISSN: 0092-5853.

Zaller, John R (1992). *The Nature and Origins of Mass Opinion*. Cambridge University Press.