

Sins of Omission: Legislator (Mis)Representation in Constituent Communications

Lindsey Cormack*
Department of Politics
New York University
19 West 4th Street, 2nd Floor
New York, NY 10012-1119

October 21, 2013

Abstract

Do legislators accurately communicate their decisions in Washington to constituents? In this paper, I propose a theory of legislator to constituent communication that anticipates a relationship between the bundle of votes a legislator chooses to reveal and the partisan composition of her constituency. To test this theory, I use an original dataset of over 40,000 official legislator to constituent communications containing 30,000 vote revelations from the 111th Congress. After presenting a series of empirical descriptions of this new data, I find evidence substantiating my theory that the extent to which a legislator endeavors to make herself look more ideologically extreme in communications varies systematically with the ratio of base to swing voters in her district. This result is contrasted with an analysis of voting extremism that finds that the ideological preferences of donors better explains roll call voting patterns. Additionally, I provide evidence that nearly a fifth of all U.S. adults have subscribed to these types of communications at some point in their lives, providing an additional reason to study the strategies employed in modern legislator to constituent communications.

*I would like to thank Sanford Gordon, Howard Rosenthal, Jonathan Nagler, Pat Egan, Justin Grimmer, Pablo Fernandez-Vazquez, Adam Bonica, Nick Beauchamp and the participants at the 2013 American Political Science Association annual meeting for their valuable input.

1 Introduction

How do members of Congress represent their votes to constituents? If all we knew about legislators was what they told us, would our understanding of their ideology differ from what we would learn relying on direct observation of their voting behavior? The answers to these questions are important for our understanding of democratic accountability because strategic communication techniques may alter voters' ability to assess the ideological positions of their representatives. To hold legislators accountable, voters need an accurate understanding of who their representatives are, and if legislators attempt to shift perceptions of their ideology by selectively revealing choice votes, voters may end up feeling more informed, even though their knowledge may be strategically skewed.

Most research on legislator ideology focuses on roll call votes, yet since the 1970s researchers have posited that legislators strategically communicate select information to constituents in an effort to enhance reelection prospects (Mayhew 1974, Fiorina 1977, Fenno 1978, Franklin 1991, Vavreck 2009). The strategy often outlined by this research focuses on touting of pork directed to the home district or efforts on behalf of an important constituent industry. Yet voters do not rely simply on the amount of money sent to the district when evaluating incumbents; they likely also care about the political ideology of their legislators.

In this paper, I present a theory of *strategic vote revelation* that yields a testable hypothesis about how differently situated legislators present their votes to constituents. Put simply, when deciding whether to reveal a vote that splits a legislator's constituents, a legislator faces an effective choice of whom to alienate. A long line of research indicates that the relevant and most likely voters for each legislator can be divided into two factions : base and swing voters (Cox & McCubbins 1986, Lindbeck & Weibull 1987, Bartels 1998, Cox 2006). Given this division, I hypothesize that legislators who anticipate a higher marginal risk of vote abstention or defection associated with alienating base voters will tend to omit votes

in which they sided with the preferences of swing voters, and vice versa. The aggregate effect of this self-censoring is for legislators to present an ideological picture in communications that differs systematically from that suggested by their voting behavior. I situate my theory of strategic communication against literature analyzing ties between financial donors and legislator voting behavior to assess the sources for the potential divergence of voting and communicating (Bartels 2008, Gilens 2009, Hacker & Pierson 2011, Bonica et al. 2013). I seek to provide answers to the following questions: 1.) Do legislators engage in systematic ideological misrepresentation in their communications? 2.) If so, which factors are most related to voting strategy and which are most related to communication strategy?

To test my theory, I analyze an original dataset of over 40,000 legislator-to-constituent communications sent during the 111th Congress, using a methodological innovation building on the ideal point estimation technique of Clinton, Jackman & Rivers (2004).¹ During this period, 95% of Representatives and 85% of Senators sent official electronic messages to constituents. Using the votes a legislator reveals in her communications, I estimate a *communicated ideal point* (CIP) and compare this to her *behavioral ideal point* (BIP), which is based on her full voting history. It is useful to study these communications because unlike media reports, broadcast interviews, or floor speeches – all of which are subject to external constraints, i.e. different agenda setters, specific interview questions, and time limits that may compel certain types of speech – official messages are optional and the scope of each message is at the discretion of the sender.² Additionally, strategic vote revelation may have

¹Specific communications analyzed here are E-newsletters and Real Simple Syndication (RSS) feeds from August 2009 on. Official communications are those sent from an Member of Congress' .gov email address or RSS feeds from the house.gov or senate.gov websites. I subscribed to each e-newsletter and RSS feed with a dummy e-mail account. For each legislator that did not have both forms of communications at the initial time of collection, I checked once a month to add to the subscription list. This collection has been approved as exempt by NYU's Internal Review Board. There are now well over 150,000 communications in my database and they are available upon request and will be made publicly available upon publication of research using the data. For the 111th Congress analyzed here, the data are from August 2009 - January 2011. Roughly 25% of e-newsletter subscriptions require an in-district zip code, for 15% of those subscriptions the messages are made publicly available on official websites, therefore 10% of the subscriptions rely on a false zip code; which is the first zip code listed for a legislator's district office. No RSS feeds have any subscriber restrictions.

²Save for restrictions on soliciting donor money, explicitly encouraging constituents to vote for a specific candidate, and reelection black out dates.

far reaching consequences on voter knowledge as somewhere between 14-19% of US adults report that they have subscribed to official congressional communications at some point.

This new data produces a few interesting results. First, I find that there is a positive correlation between CIP and BIP estimates, but this relationship is not perfect and, on average, the measures are distinct. Second, I observe an overall leftward shift, with CIP estimates being more extreme for Democrats and more moderate for Republicans than BIP estimates. Third, I find that the likelihood that a legislator presents a CIP that is more extreme than her BIP is conditional on district characteristics, most importantly the ratio of base to swing voters. Fourth, I confirm findings that relate donor ideological extremity to voting extremity. I conclude with a discussion of why different factors likely contribute to voting strategy versus communication strategy and the implications of this divergence.

Overall, the direction of ideological misrepresentation is consistent with my theoretical expectation that a legislator will attempt to appear more ideologically similar to whichever subset of voters – base or swing – that she considers more important for reelection. This is the case even when accounting for other possible factors thought to influence the communication choices of a legislator. In contrast, I find that a distinct set of explanatory variables better accounts for voting behavior, suggesting that legislators choose voting and communication strategies with different influences in mind.

2 Constituent Communications

To date, with notable exceptions, few studies have been able use more than a sample of communications to empirically test the communication tactics of legislators.³ This is no doubt

³Studies employing the universe of a given communication media include, (Grimmer 2013, Grimmer forthcoming, Williams & Gulati 2010, Grimmer & King 2010, Grimmer 2010, Adler, Gent & Overmeyer 1998). In addition, there are many well done studies using selected texts, (Wigley 2011, Sellers 2010, Neiheisel & Niebler 2010, Ludwig 2010, Gulati & Williams 2010, Druckman, Kifer & Parkin 2010, Herrnson, Stokes-Brown & Hindman 2007, Xenos & Foot 2005, Trammell & Williams 2004, Lipinski 2004, Richardson,

because collecting and coding such a vast corpus is time consuming and previously computationally unfeasible (Grimmer & King 2010). Thus, despite great interest in congressional research, we know surprisingly little about the strategic communication of legislators.

In an early study of political communication, Mayhew (1974) distinguishes three types of communication tactics: credit claiming, advertising, and position taking. Credit claiming is a relatively well studied behavior, which includes the promotion of particularized benefits to constituents ostensibly obtained by a legislator. This tactic has been expertly analyzed by Grimmer (2010, 2013, forthcoming). Advertising, as defined by Mayhew, has not been subject to much scholarly interest because it simply involves an attempt to increase name recognition of an incumbent. Position taking, which is characterized by the stating of policy or ideological preferences, has been attended to in various manners, but the specific focus on vote revelation as a constituent communication strategy has not yet been explored in the literature (Box-Steffensmeier, Arnold & Zorn 1997, Koger 2003, Xenos & Foot 2005, Highton & Rocca 2005, Bovitz & Carson 2006). Vote revelation, which I address in this paper, is a subset of position taking as it relies on communicating a specific vote associated with a position, rather than simply stating an ideological preference without policy action.

Despite the lack of a literature on strategic vote revelation, a considerable theoretical literature aims to explain different facets of strategic *information* revelation. While not explicitly considering vote revelation, this work provides guidance on what actions rational, election-minded legislators will take. A legislator will attempt to influence her constituency's understanding of her position and should send a message only when there is a true belief in the possibility of persuading her audience (Austen-Smith 1992*a*, Austen-Smith 1992*b*, Landa & Meirowitz 2009). Second, a legislator will attempt to reveal information aligned with the wishes of her electorate (Cox & McCubbins 1986, Heidhues & Lagerlof 2003, Meirowitz 2005, Gratton 2010). Thus, a legislator will craft communications with an expectation of

Daugherty & Freeman 2001, Yiannakis 1982, Fenno 1978).

influence over voter perceptions and, as such, the methods she uses to accomplish this goal are worthwhile indicators of intent and deserving of academic attention.

3 A Heuristic Model of Strategic Vote Revelation

In this section I describe the considerations that should lead differently situated legislators to reveal certain types of votes. The basic intuition is as follows: Each legislator has a constituency made of base (co-partisan) and swing (independent) voters.⁴ These voters can be characterized by the median ideological preference of each group. For a legislator, publicizing a vote can alienate both groups; please both; or alienate one while pleasing the other. The willingness to publicize a vote in the third case will vary according to the marginal benefit of pleasing one group compared to the cost of alienating the other. Over time, the aggregate result of such legislator calculations generates a kind of selection bias in which a legislator can appear more extreme or more moderate than her full voting history suggests.⁵

3.1 Assumptions

Base voters are those who will likely vote for the incumbent by virtue of party labels or previous commitments, but may abstain. Swing voters may support an incumbent, but may also support a challenger or abstain. I hold that the median base voter is more ideologically extreme than the median swing voter.

I assume a voter prefers an incumbent legislator who is ideologically closer to himself *ceteris paribus* and the legislator knows this (Downs 1957). I also assume a voter's ability and

⁴There is of course another segment of the constituency that will not support an incumbent legislator, those who support the candidate from the opposing party. These voters are not in the relevant set of constituents when evaluating likely reelection votes.

⁵If a legislator alienates both groups, she is not expected to reveal that vote and if she pleases both groups she is expected to reveal that vote assuming the perceived benefit outweighs the cost of creating and sending a message in the first place.

willingness to expend resources to accurately learn legislator ideology is small (Achen 1975, Delli-Carpini & Keeter 1996, Feldman 2009). Lastly, I assume that voters are not negatively introspective, meaning that voters do not make inferences about legislator ideology from the absence of communication (Dickson, Hafer & Landa 2008).

I assume legislators wish to be re-elected above all other goals (Mayhew 1974). I also assume that legislators vote sincerely, but communicate strategically to maximize electoral prospects. Of course, the assumption that legislators vote sincerely in every circumstance is unrealistic (Poole & Rosenthal 1997). It is meant to capture, in reduced form, the idea that while a host of influences – including donors, party pressures, as well as the legislator’s personal ideology and that of her broader electorate – may affect her vote choice, the decision of how to communicate is predominantly about appealing to voters. I estimate models of both sorts of behavior, communicating and voting, to assess the plausibility of this assumption.

Knowing voters are in costly information environment, a legislator will attempt to influence voters’ perceptions to remind her base to vote for her, convince swing voters to choose her, and generally serve to lower information costs associated with voting for all potential supporters (Aldrich 1995). Vote revelation content is verifiable and I assume a legislator will not lie about her vote.⁶ Lastly, I assume that sending each vote revealing message is associated with some small cost made up of staff time used drafting a message, potentially cluttering a voter’s inbox with too many messages resulting in annoyance, and the opportunity cost of not talking about some other issue.

⁶I never observe lying in the dataset. Constituents and bloggers occasionally write into their local papers, and papers have published accounts of inaccuracies in e-communications effectively deterring such actions (Rock 2008, Jonson 2009).

3.2 Audience Considerations

There is no consensus that legislators are writing specifically to either base or swing voters. It is also not clear if either group is more likely to receive these communications, or if either group is more likely to change voting intention based on vote revelations. Unlike campaign communications, which are sometimes quite recipient-specific in their content as a result of micro-targeting efforts (Levy 2008), official communications are nearly always the same regardless of the receiver. The only potential for constituents of the same district to see different official messages arises if a legislator asks subscribers to select different topics of interest to limit the types of messages that each subscriber receives. This does not threaten my analyses for three reasons: 1. Very few legislators ($\approx 7\%$) include this option on their subscriptions. 2. Even when legislators have such an option for subscription, many eschew writing specific category letters and instead just compose and send general letters that all subscribers receive. 3. If a legislator has such a system, I subscribed to all categories.

Interviews with press secretaries, the recorded nature of these messages, and legislator encouragement to forward the messages suggest that the audience is broader than just base voters. Yet, a study of electronic message recipients from state level political office *campaigns* indicates that recipients tend to be more extreme than the general population (Herrnson, Stokes-Brown & Hindman 2007).⁷ To find out how many and what types of people sign up for official e-newsletters and RSS feeds, I placed a question on the 2012 Cooperative Congressional Election Study (CCES).⁸ Approximately 19% of respondents reported signing up for official messages from their Representative and 14% from their Senators.⁹ Those who

⁷While the notion of the ongoing campaign has validity, there is evidence suggesting legislators communicate differently in active campaigning outlets versus official government communications if for no other reason than the legal constraints prohibiting explicit vote seeking in official communications (Glassman 2007).

⁸The CCES draws from a nationally representative survey population for the full survey. The specific question posed to a randomly selected 1,000 respondents was, “Have you ever subscribed to email updates such as an e-newsletter or Real Simple Syndication (RSS) feed from any of the following elected officials?” followed by the name of their Representative, junior Senator, and senior Senator.

⁹For any legislator, it is unlikely that 19% of their district subscribes to e-newsletters and RSS feeds at the same time. The question asks respondents if they have *ever* signed up for these types of communications.

indicated they had subscribed to such communications are, on average, older, more educated, wealthier, more politically active, slightly more politically extreme, and are more likely to vote in primary elections, to approve of their incumbent legislators, and to think higher of Congress in general than non-subscribers. Thus, the only specific survey data on this media produces evidence indicating that the audiences for these types of messages tend to be more ideologically extreme. Yet this need not mean that all legislators write their messages with just base voters in mind; if I observe legislators moderating, this provides some evidence of strategic calculation and understanding that more than base voters may read these messages.

3.3 Hypothesized Legislator Strategy

Consider a one-dimensional ideological space with the leftmost end representing liberal and the rightmost end representing conservative. To keep matters simple, I focus on the following legislator and voter ideological arrangement corresponding to an incumbent Democrat situated between her median base and median swing voters.



In this example, the space represented is roughly one half of the whole spectrum. While variations are possible for any given vote, this is the most likely arrangement for the overall organization of voters and incumbents across policies. The assumption that a legislator is between her median base and median swing voters is justifiable because if an incumbent is more extreme than her base - unless the base is a 50%+1 of all voters - it is unlikely that she would continue to be electorally successful; the same is true if she were more moderate than the median swing voter.¹⁰ Empirically, self-identified Democrats are an absolute ma-

Interviews with press secretaries put the upper bound estimation for the amount of subscribers at any given time around 10%.

¹⁰Given that political elites tend to be more extreme than average voters, this assumption may not be fully met in every instance. However, the electoral calculation posited above; in addition to recent research on mass citizen preferences by party ID, indicates that there is often a substantial gap between the preferences of the

majority as opposed to self-identified Republicans and Independents in only 14% of districts and self-identified Republicans hold an absolute majority in only 6% of all districts.¹¹

On the one-dimensional ideological continuum each legislator has her personal ideal point (PIP) which is not observable. What is observable, estimable, and most often studied, is voting behavior that can be summarized by an ideal point based on the full non-unanimous voting record of a legislator. I call this the *behavioral ideal point* or (BIP). As contrasted to the BIP, the *communicated ideal point* or (CIP) is the ideological position voters would perceive if only the votes a legislator reveals are used in generating an ideal point estimate.

For each bill that is considered for a vote in Congress there is a “cutpoint” or a place on the ideological continuum that splits people who would prefer the bill outcome from those who would instead prefer the status quo (Poole & Rosenthal 1985). For each vote that splits the median base and swing preferences, a legislator must decide if the benefit of publicizing her vote – which sides with the preferences of just one part of the constituency – outweighs the potential cost of alienating voters who disagree with her. Over the series of votes a legislator takes, if a legislator perceives her base as more important than her swing voters in her attempt to get reelected, she will present a CIP that is more extreme than her BIP. Conversely, if a legislator perceives swing voters to be more important electorally, she will present a CIP that is more moderate than her BIP.¹²

median Democrat, Republican and Independent in each state and therefore this assumption is reasonable (Tausanovitch & Warshaw 2012). The theoretical predictions remain the same if the legislator is more moderate than both base and swing medians. If the legislator is more extreme than both medians, the predicted outcomes change slightly and I leave pursuing this possibility to a future project.

¹¹Data from the 2008 National Election Pool Exit Polls. Knowing that 20% of all districts have a one party voter majority does not necessarily imply that legislators from such districts are more extreme than their median base voter. In fact, if a legislator is in such a district, her optimal ideal point strategy will be close to the median base voter and may therefore be slightly more moderate or slightly more extreme, but not radically more extreme.

¹²The empirical strategy adopted in this paper cannot disentangle the distinction between the possibility that different parts of the electorate act on the revealed information at different rates or the different parts of the electorate might simply hear the revealed information in greater or lesser quantities. That is to say, if I observe all legislators attempting to appear more extreme, that likely indicates that the base is more important in crafting a vote revealing strategy, but I cannot infer whether the reason for this legislator choice is based on the fact that base voters react to vote revealing more strongly than swing voters or if base voters are just more likely to be the anticipated audience for legislator communications.

To assess the importance of her competing constituencies, a legislator must have a sense of the size and likelihood of voting for both base and swing voters. There is no *a priori* reason for assuming the likelihood of receiving and reacting to messages is higher for either part of the constituency, and there is no completely reliable way of knowing who is more likely to vote. This means a legislator must rely on past elections to inform her decisions about the relative size of each voting constituency.

4 Data and Empirical Strategy

The data I use to measure vote revelation are all the official e-newsletters and Real Simple Syndication (RSS) feeds sent in the 111th Congress from August 2009 through the end of the term, for a total of 40,957 messages. The messages are from an original dataset I collected from every Representative and Senator who sends official e-newsletters and/or RSS feeds. In order to generate the specific measures of vote revelation, I first ran a computer sweep over every message to identify all messages containing a potential vote revelation based on a search of key terms.¹³ Second, I read each message flagged in the computer sweep to confirm a vote from the 111th Congress was actually revealed. This resulted in 15,690 messages that were read by either myself or a highly trained team of research assistants. Each specific vote was recorded. Although each coder had to train until reaching a inter-coder reliability rate above 90%, 98% of all messages were either coded by me or double coded by me and a research assistant. Finally, I created a roll call matrix consisting only of revealed votes and compared this data to the full history of roll call votes.¹⁴

To measure the likely relative importance of different constituencies for each legislator, I use ratio measures of the party identification of voters who turned out in the 2008 general election. Data on the composition of each electorate are obtained from the 2008

¹³Specific terms and processes are detailed in Appendix A.

¹⁴Roll call vote data for 111th Congress are compiled by Jeff Lewis and Keith Poole at voteview.com.

National Election Pool Exit Polls conducted by Edison Research (2008).¹⁵ For each district and state I create a ratio measure of Democrat identifiers (base) to Independents (swing), which are the relevant constituencies for Democratic legislators, and a ratio of Republican identifiers (base) to Independents (swing) for Republican legislators.¹⁶ Table 1 displays the average ratios broken down by party of the incumbent. By only using data from actual as opposed to registered voters, I can better estimate the relevant constituencies for reelection – actual voters – from the entire constituency a legislator is responsible for representing.

[Table 1 about here]

While the 2008 exit polls are conducted in each state, not every congressional district has enough respondents to release the crosstabs on partisanship. Of the 88% of districts and states that have legislators who sent vote revelations there are exit poll data for 346 districts. Theoretically, this data missingness should be orthogonal to legislator strategy, so while it is regrettable, it should not bias the estimated parameters or change the expected values of any estimated quantities of interest.¹⁷

4.1 Empirical Implementation

To test my hypothesis I need two measures for each legislator: a behavioral ideal point (BIP), based on all the roll call votes a legislator takes; and a communicated ideal point (CIP),

¹⁵This data is made from the individual datasets for each of the fifty states and the District of Columbia. Sample sizes are between 800-2000 for each state.

¹⁶Respondents who identified as “other” (4%) are treated as Independents because there is no *a priori* reason to assume they are irrelevant. Results remain consistent if such respondents are eliminated altogether.

¹⁷Correlations between a dummy measuring missingness and the dependent variable indicating whether a legislator’s CIP is more extreme than her BIP indeed demonstrates that there is no discernible relationship. The coefficients for Democrats: -0.0034, Republicans: -0.0037. Additionally, there may be concerns that the ratios of partisans are poorly estimated since I use an exit poll. The analyses reported here have also been conducted using Catalyst voter file data of party identification via vote registration and the 2008 CCES data of self-reported party ID to estimate the co-partisan to independent ratio. Results remain substantively similar, but state and district coverage is greatest with exit poll data.

based only on the votes a legislator reveals to constituents. To estimate these quantities I employ a modification on the standard two-parameter IRT voting model pioneered by Clinton, Jackman, and Rivers (2004) to estimate latent ideal points for legislators from their observed voting and communicating behavior.

Described informally, I perform standard two-step Bayesian ideal point estimation and then add an additional step to incorporate the restricted subset of revealed votes. After setting reasonable starting values, bill parameters are estimated using legislator BIP values, then these bill parameters are used in fine tuning the legislator BIP estimates and then the cycle begins again using the values estimated in the last iteration and so on for 10,000 iterations. For the additional step used to estimate CIP, I use the bill parameters that are estimated based on the full voting history to generate a set of CIP estimates. Figure 1 provides a simple diagram of the process and a detailed explanation is in appendix B.

[Figure 1 about here]

5 Results

Owing to the newness of this data, I first describe the frequency of vote revelations and the basic relationship of BIP to CIP estimates. I subsequently test my theory of strategic vote revelation and present the results.

5.1 Heterogeneity in the Frequency of Revelation

Despite recent reports arguing that Democrat campaigns are much more technologically advanced than Republican campaigns (Rutenberg 2013), I find that members of each party have set up official email and RSS feeds at rates that are not significantly different; 94% of

Republicans and 91% of Democrats send messages. However, in terms of the quantity and content of the messages there are significant differences. Figure 2 shows histograms of the numbers of messages sent and unique votes revealed in the 111th Congress by party.

[Figure 2 about here]

There is considerable heterogeneity in the number of votes revealed, ranging from 0-172. Of legislators who choose to send any messages, the mean number of messages sent from Republicans is 89 and from Democrats it is 73. Despite sending more messages, Republicans are less likely to reveal votes within those messages. Democrats reveal 19 unique votes on average while Republicans reveal 15.¹⁸ Additionally, the difference in the percentage of messages that involve vote revealing is also significantly different; 47% of all messages sent by Democrats contain vote revelations, but only 41% of Republican sent messages do.

5.2 The Relationship of Communicated Ideal Points to Behavioral Ideal Points

To show how the two ideal point estimates compare for each legislator, Figure 3 presents a scatter plot comparing BIP to CIP points in the House and Senate.

[Figure 3 about here]

A 45-degree line is overlaid on each graph in figure 3 highlighting two interesting points about the comparison between BIP and CIP. First, while highly related, the measures are not perfectly correlated. The correlation coefficient for these two measures is 0.89 in the Senate and 0.88 in the House. Second, there is not a one size fits all communication

¹⁸This difference is significant with a p-value of 0.02.

strategy; not all legislators try to appear more moderate or extreme and there is a good deal of variation the direction and level of ideological misrepresentation each legislator pursues.

Figure 4 depicts kernel density plots of CIP and BIP by party. Figure 4 confirms that BIP estimates are quite polarized, as indicated by the higher peaks and distinct valleys between the two parties. The CIP estimates, on the other hand, are spread much more evenly, with many legislators filling the previous gap. Also apparent is the greater difference between the CIP and BIP distributions in the House than in the Senate.

[Figure 4 about here]

Table 2 provides a breakdown of the average direction and magnitude of ideological misrepresentation by party. The majority of Republicans present CIPs that are more moderate than their BIPs, and the majority of Democrats present more extreme CIPs than BIPs. This amounts to an overall leftward shift. The magnitudes of these shifts are not trivial. Given the range of observed behavioral ideal points for Republicans, the average leftward shift is roughly 24% of the scale; and for Democrats the average leftward shift is 22%.

[Table 2 about here]

This finding is interesting, and while not posited by my theory, plausibly fits my expectations if Democratic legislators tend to come from districts with a higher ratio of base to swing voters and Republican legislators tend to come from districts with lower ratios of base to swing voters and this is all that matters when crafting a communication strategy. On the other hand, it could be that other factors contribute to messaging strategy decisions and a focus simply on the base to swing voter ratio omits some forces at work. Or yet another possibility is that while the base to swing ratio is important for all legislators, it may affect one party more than the other because additional influences such as party pressures to

talk about certain topics may moderate the raw influence of the the theorized relationship between a legislator’s electorate and her communication strategy. Looking back at table 1, there is some evidence that Democratic legislators tend to have higher base to swing ratios than Republicans, but this difference is quite small. Before constructing a larger model, I ran a simple regression with the outcome variable as an indicator signaling that a legislator presented a more extreme CIP than her BIP on the base to swing ratio, a party indicator, and an interaction of the two. After predicting the likelihood that a legislator would present a more extreme CIP than her BIP, I find *some* support for the notion that, while the ratio of base to swing voters matters for both Democrats and Republicans, they exhibit different sensitivities to the measure. Figure 5 plots the predicted values from this regression by party.

[Figure 5 about here]

Figure 5 shows two important things. First, there is a positive relationship between the ratio of base to swing voters for both parties, but the starting likelihood of presenting a more extreme CIP and BIP are different for Democrats and Republicans. The baseline probability that a Democrat will present a more extreme CIP than BIP at the lowest observed values of the ratio measure is over 50% indicating that this strategy is more likely used by Democrats than Republicans regardless of their constituency. Republicans are more likely to present a more extreme CIP once the base to swing ratio reaches roughly 3:1. This may explain why the shifts reported in table 2 are leftward for both Democrats and Republicans. Second, while the sensitivities to the ratios are different for each party, the the coefficient on the interaction term in the regression was not significant and, therefore, I do not conclude that influence of the ratio of base to swing voters on the decision to strategically reveal votes.¹⁹

Part of the scattering and overlapping effects in these initial figures is due to the increased error in estimating CIP points because there are more missing observations.²⁰ This

¹⁹Given this, in the more detailed model estimated in the next section, I control for party, but I still expect that the ratio of base to swing voters will effect legislators of both parties in the same direction.

²⁰The average standard error associated with BIP is 0.032, the average for CIP is 0.321.

may give rise to a concern that CIP is just a noisy BIP measure and not actually indicative of different underlying phenomena. If 50% of legislators present a more extreme CIP and 50% a more moderate, this would not necessarily be interesting because such a result could be chalked up to chance and measurement error. I address this concern in a few ways. First, by using a paired t-test across all legislators, I find the average difference between the CIP and BIP to be roughly 0.10 and I can reject the null that they are the same, on average, with a p-value of 0.00. Second, it is in the ability to predict which legislators choose which strategy that makes a stronger case for purposeful behavior by legislators and further supports the theory that the BIP and CIP estimates are not random or unduly marred by measurement error. In the next section, I do just that by presenting results that show that the composition of the electorate is a strong predictor of legislator communication strategy.

5.3 The Effect of Constituency Composition on Communicated Ideology

To test the hypothesis relating the composition of a constituency to a legislator's CIP, I estimate a probit model where the dependent variable is an indicator that is 1 when her CIP is more extreme than her BIP. An indicator rather than a distance measure is theoretically preferable because a legislator ought to know that she is attempting to look more extreme or more moderate, but precise knowledge of *how* much more moderate or extreme would be a far more difficult strategy for a legislator to implement.²¹ Additionally, since no one member has control over the agenda, specific fine tuning is much less realistic than the decision to alter one's perceived ideology in either direction. A simple model, with no controls (not presented here) indicates that the ratio of base to swing voters is significantly related to the likelihood of presenting a more extreme CIP than BIP ($p = 0.02$). As a more rigorous test of my theory that addresses potential alternate explanations for vote revelation and

²¹However, when using a standard OLS regression of the distance, I find results are substantively similar.

voting strategies, I perform two additional regression analyses. First, I estimate the probit model described above on the electorate ratio measure as well as a series of variables that are likely related to communication strategies. Second, to contextualize the results from the first regression, I model extremism in voting as measured by the mean-centered BIP squared on the same explanatory variables used in the first regression. Here I employ ordinary least squares (OLS) owing to the continuous nature of the dependent variable. In addition to providing context, this analysis allows me to explore the possibility that similar influences contribute both to how a legislator votes and communicates.

The first control is the ideology of financial donors. I use Bonica's (2013) Campaign Finance Scores to quantify the ideological pull of each legislator's donor base. The variable provides a basic test comparing the influence of donors versus that of likely voters. A second control is a measure of overall electorate ideology that comes from Tausanovitch and Warshaw (2012), who use multi-level regression with post-stratification to estimate the ideological preferences in each district and state by pooling many surveys. While my theory – and previously cited work – argues the relevant set of voters for a legislator to consider is her base and swing voters, some may argue that the composition of the overall electorate is what legislators ought to consider, so I therefore, include this measure. In both of these analyses, the constituent ideology and finance scores are included in their raw form and they are both mean-centered and squared to capture extremity for each value. I also control for the ideological extremity of each legislator's voting history by including BIP squared and well as BIP in the first model estimating the likelihood that a legislator presents a more extreme CIP than BIP as it could be that those who actually are more extreme in their voting also wish to look more extreme in their constituent communications.

I use another set of controls measuring the variety of audiences for each legislator from the 2000 Census. Specifically, I include the percentage of the voting age population that is under the poverty line, the percentage of black and Hispanic residents in each district,

and the median age of each district. This is because legislators from districts with larger minority populations may choose to emphasize different types of voting decisions than those from more homogeneously white districts, or those from poorer or older districts may face distinct pressures than legislators from wealthier or younger districts.

I use legislator-specific controls such as race, sex, seniority, and leadership. Rocca & Sanchez (2008) find that racial minorities tend to sponsor and cosponsor bills less than their white male counterparts and perhaps these members also adopt different communication strategies. The control for legislator seniority is grounded in the finding that junior members tend to put in more upfront time on constituent contact and may therefore also use different communication strategies (Fenno 1978, Cover 1980). There is an indicator signaling that a legislator is a party leader because that these individuals, by definition, have a more public role that likely influences their communication strategies. Additionally, I include an indicator signaling whether a legislator ran for reelection in 2010 to check against the possibility that retiring legislators may communicate in a manner different from colleagues seeking reelection.

There are also various institutional settings that may affect communication strategies. Different qualities of state maintained voter file data could influence how a legislator presents herself because as the quality of data increases, a legislator can have more certainty about her perceptions of the electorate (Hersh 2011). Understanding this, I include a series of dummy variables indicating different types of voter file data kept by each state.²² Another pertinent concern for legislators is the type of primary system. Closed primaries are generally more partisan and may therefore contribute to legislator communication strategies, with an expectation that closed systems are correlated with more extreme styles of communication. I include an indicator for whether a state or district has a closed primary system. Lastly, I include party and chamber indicators. Table 3 presents the results.

²²Indicators signify if a state has (1) no partisan data in the voter record (2) primary ballot selection data (3) partisan registration (4) partisan registration and primary ballot selection data.

[Table 3 about here]

Column 1 of table 3 contains the results for the test of my theory that as the ratio of base to swing voters increases, so does the likelihood that a legislator will present a more extreme ideology in her communications than reflected in her full voting history. The coefficient on the ratio measure is positive and significant, providing support for the expected relationship. There are three other measures that display significant relationships with this communication strategy, a legislator's voting extremity, the percent of Hispanic residents in the district, and party. As a legislator votes in a more extreme manner, she is less likely to communicate in a way that is even more extreme. This could perhaps be because legislators who vote extreme may feel the need to appear more moderate in communications. Alternatively, once someone votes in a very extreme manner it may be difficult to craft an image that is even more extreme. The next observed conditional relationship indicates that as a legislator's district includes more Hispanic residents, she is more likely to try to appear more extreme in communications than in voting. The reasons for this are not entirely clear, but it could be that legislators from more heavily Hispanic districts may have greater pressures to discuss votes on topics such as immigration reform which are almost always highly polarizing votes, and therefore lead to a presentation of self that is extreme. Lastly, I find that being a Democrat is related to presenting a more extreme CIP. This result follows from the simple analyses from before and could either speak to an overall party messaging and vote focusing strategy employed by legislators broadly, or may be linked to similar constituent assessment strategies of Democratic legislators.

In stark contrast, the results in column 2 of the model of extremism in voting behavior do not show a significant relationship between the base to swing ratio measure, and instead donor ideology, overall district ideology, poverty levels, whether a legislator is Hispanic, seniority, and chamber better account for voting. I set my analysis of communication against voting with an anticipation that I would find a baseline relationship between donor

ideology and voting as reported in the previously cited literature that would likely be absent in communication strategies. The other variables that exhibit significant relationships with voting however, coincide with previous research and add some nuance to our understanding of voting. I find that as a donor base is more extreme, so too is a legislator's voting. Additionally, I find that there is relationship between voting extremity and donor conservatism. A similar relationship links the overall electorates' preferences and voting extremism; the more extreme the overall electorate is, the more extreme voting pattern a legislator has. I also find that Hispanic legislators have more extreme voting records than non-Hispanic legislators and those with higher poverty rates in their district tend to be more moderate in their voting. As the 111th congress occurred right after the emergence of the "Tea Party" movement, it is perhaps unsurprising that Democrats tend to have less extreme voting than Republicans. The conditional relationships on chamber and seniority are somewhat unexpected, as much of the research on polarization indicates that the House hosts more extreme members than the Senate and it is not clear that senior members tend to be more extreme than junior members, especially in the recent Congresses where polarization is thought to increase via replacement when new members take seats from older, more reasonable members. A detailed pursuit of why these institutional variables appear to be related to voting in this manner is beyond the scope of this work and instead I now turn to simulation analyses to better interpret how predictively useful the key variables of interest are.

To better interpret the impacts of voter ratios on communications and donor ideology on voting I ran two simulations. I estimated each model and then simulated 10,000 values of the resultant parameters, sampling from the estimated distribution each time using CLARIFY (Tomz, Wittenberg & King 2003). I set all independent variables to their means and then manipulated the base to swing voter ratio in order to better assess the conditional effect of this explanatory variable on the likelihood that a legislator has a CIP that is more extreme than her BIP.²³ I then manipulated the donor ideology and overall electorate

²³I use an average hypothetical legislator for the simulations. This is a Democratic Representative who

ideology in the model of voting extremism.

Going from the minimum to the maximum observed base to swing voter ratio corresponds to a 40% increase (s.e. 0.12) in the likelihood that a legislator attempts to appear more extreme in communications. These findings suggest that legislators tailor messages to their electoral audiences in an attempt to better match relevant voters and is especially compelling because other measures found to be strongly related to voting – donor ideology and district ideology – do not exhibit a significant relationship with communication strategy, *ceteris paribus*. In the next section, I discuss how my findings are situated amongst previous work and how they can inform future research on the divergence between voting and communication strategies, representation, electorate awareness, and legislator accountability.

6 Discussion: Voting versus Communicating

The preceding analyses and results substantiate my theory that a legislator will communicate an ideology that diverges from her behavioral ideology in a manner that makes her appear better aligned with the most electorally valuable faction of her reelection constituency. To contextualize this finding I also check to see if there are different contributing factors to actual roll call voting versus communicating to constituents about votes. With this strategy, I find evidence of a possible and plausible reason for the divergence of CIP and BIP.

Previous work has found that donors influence voting behavior, and my results confirm this, but there has been virtually no work on how legislators attempt to temper the

has a district with a citizen ideology of 0.05 and a squared citizen ideology of 0.06, a campaign finance score of -0.06 and squared CF score of 1.61, behavioral ideology of -0.11 BIP squared of 0.42, with both primary ballot and partisan registration voter data, an open primary system, 11% of the voting age population below the poverty level, a median district age of 36, 11% black citizens, 7% Hispanic citizens, is white, non-Hispanic, and a male who has served for 10 years, revealed 18 votes and is running for reelection. The average Republican has a district with a citizen ideology of 0.18, 10% of the voting age population below the poverty level, a median district age of 36, 9% black citizens, 6% Hispanic citizens, is white, non-Hispanic, and a male who has served for 10 years, and is running for reelection, and is not a legislative leader.

perception of her votes to constituents via strategic communication. Because of their size, power, and intense interest in actual policy, donors are much more likely and able to hold legislators accountable for each vote they take compared to the average voter. When communicating directly to constituents however, legislators can craft an image that may be virtually free from donor pressures. It is reasonable to expect that the large monied influences on roll call voting will be pleased with the actual roll call votes and understand that strategic communicative posturing is merely used to attract voters in the district. In my last estimates presented in table 3, I find support for this explanation. How a legislator wishes to be perceived in her district is significantly related to the ratio of base to swing voters, and not donor ideology. However, when looking at the extremism of actual votes cast, donor and overall district ideology play a much stronger roll.

Something not entirely anticipated was the influence of overall electoral extremity and voting behavior. This finding may help calm normative concerns that legislator's vote for their donors then lie to their voters. While I find support that donors influence voting, I also find that overall electorate ideology is significantly related to how a legislator votes. So while the ideological shift that occurs in communications is still present, it may be interpreted as less nefarious. Perhaps legislators have some sense of duty to the ideological preferences of their districts as a whole when casting votes, but this consideration falls away when creating a communication strategy in favor of aiming at a more narrow set of likely voters. Future research into the motivations and perceptions of legislators on their constituents can use the different effects reported here as a starting point.

6.1 Political Representation

Comparing BIP and CIP estimates across the whole Congress, I find that there are significant differences between the two. What does this pattern of ideological misrepresentation mean for political representation? One implication is that the frequency and direction with which

this ideological misrepresentation occurs signals that the effort is likely intentional. The intent behind creating a different image and what this misrepresentation looks like across the Congress is worthwhile to consider. Assuming each legislator is rational and election minded, general ideological misrepresentation should not come as a shock but the overall pattern of this distortion may be somewhat surprising.

I conjecture that when asked what type of misrepresentation is expected, the majority of political observers would either answer that legislators of both parties would moderate in an attempt to appeal to the most voters, or that most legislators would try to appear more extreme if the understanding is that base voters make up the bulk of e-newsletter and RSS feed readership and legislators are increasingly facing primary challengers. Instead of legislators of all stripes moving either towards the extremes or running to the center, I find that there is generally a leftward shift across *both* parties. The reasons for this are not entirely clear, but the analyses presented here do offer some guidance.

This partisan difference in strategy is not because legislators come from widely different types of districts in terms of base to swing voter ratios, because this is not the case. It is also not explained by the argument that one party responds to the base to swing voter ratio and one party does not, as I find legislators of both parties are receptive to base to swing voter ratios when selecting what types of votes to reveal. It could be that my assumption that places a legislator between her base and swing medians is incorrect and that most Democratic legislators are actually more moderate than the voters they perceive as most important to secure reelection and that most Republican legislators are actually more extreme. If this were true, finding that legislators across the spectrum try to appear more liberal makes perfect sense, as everyone would just be chasing their supporters. By squaring the measure of behavioral ideology to gauge legislator extremity, and squaring the MRP estimates of each district from Tausanovitch & Warshaw (2012) to measure overall district extremity, there may be some support for this supposition. Republican legislators have an average extremity

score of 0.54 while Democrats have an average score of 0.34, which makes for a statistically significant difference with Republicans being more extreme. Constituents on the other hand, show a reversed statistically significant difference with Democrat constituents more extreme than Republicans constituents on average (0.10 to 0.05). The problem with this quick comparison is that constituents and legislators are not measured on the same scale, and a true test of this theory is beyond the scope of this paper. However, this is a plausible explanation given this first cut and a broader empirical investigation could be used in the future.

6.2 Electorate Information and Accountability

Knowing that voter ignorance is widespread, and that a legislator can misrepresent her ideological position to voters, makes voting harder for citizens. This type of strategic revelation raises important normative questions for political theorists. Is the attempt to alter the perception of one's ideology a form of dishonest misrepresentation or is this all fair and expected in a representative democracy? The evidence of specifically tailored messaging strategies is important to move forward that discussion. There are also questions of constituent knowledge and trust. Is it the case that the more a legislator reveals her votes, her electorate is more knowledgeable about her positions or perceives her to be more trustworthy?

There is an argument that vote revelation is better for constituent awareness because each decision is an additional bit of information for constituents to use in deciding how to vote. On the other hand, if vote revelation is intended to throw off constituents by misrepresenting a legislator's ideology or imbue voters with a false sense of knowledge, the value of these additional bits of information becomes suspect.

Lastly, are these attempts at altering perceptions effective and do they amount to a winning strategy? To answer these questions, this data in conjunction with survey research and election analysis should enhance our understanding of why legislators engage ideological

misrepresentation. Do legislators who misrepresent their ideology in a manner aligning with voters of their district do better electorally than others? While not perfect counterfactuals for each other, the actions of different legislators who are otherwise situated in similar districts can be used to leverage the greater impact of ideological misrepresentation.

7 Conclusion

Strategic communication is a skill on which legislators must increasingly rely when connecting with constituents. In a world full of political information, a legislator has a unique opportunity to craft a specialized image in the minds of voters that may not always align with the reality of her voting history by using direct communications. This paper offers a theory of how differently situated legislators will selectively reveal certain votes and omit others. I argue that a legislator will engage in ideological misrepresentation to appear to be a better fit to the most electorally relevant parts of her constituency.

Using a new dataset of legislator communications I am able to test this theory and answer empirical questions previously left to the realm of theoretical or case work analysis. I find that most legislators seize the opportunity to communicate with citizens directly using e-newsletters and RSS feeds, bypassing the filter of the media. I also find that there are partisan strategies in using these communications, with Republicans sending more messages on average, and Democrats revealing more votes on average. Overall, Democratic legislators attempt to appear more liberal and Republican legislators attempt to appear more moderate. The number of legislators that engage in ideological misrepresentation, the differences between communicated and behavioral ideal points, as well as the finding that donor ideology strongly influences roll call voting but not communicating, leads me to conclude that this misrepresentation is intentional.

With an innovative analytical technique, I measure how the vote revealing content

of these sorts of messages differs from a legislator's full voting history. I find that the votes that a legislator reveals and the ideological implication of those choices differs significantly from the full voting history of a legislator. Specifically, as the ratio of base to swing voters increases so does the probability that a legislator will present a communicated ideology that is more extreme than her behavioral ideology.

Lastly, this new dataset holds potential for many more research areas; the theory and subsequent tests presented here serve to highlight the utility of electronic message data. Political scientists have long supposed that members of Congress present themselves in a way that may not truly reflect their acts in office, but for the most part this supposition remained untested for lack of systemic data. Having a running text of what a legislator tries to communicate to her constituents offers a very powerful, and increasingly easier to analyze dataset. As this dataset accrues roughly 80 new messages per day, the potential to study topic trends, vote revelation strategy, and presentation of self, and many more issues is great. The analysis presented here touches just one part of the strategic political communication and offers a small step in fulfilling the hope that political scientists, "spend a little less of our time explaining votes and a little more time explaining explanations." -Fenno (1978).

References

- Achen, Christopher H. 1975. "Mass Political Attitudes and the Survey Response." *The American Political Science Review* 69(4):1218–1231.
- Adler, Scott E., Chariti E. Gent & Cary B. Overmeyer. 1998. "The Home Style Homepage: Legislator Use of the World Wide Web for Constituency Contact." *Legislative Studies Quarterly* 23(4):585–595.
- Aldrich, John H. 1995. *Why Parties? The Origin and Transformation of Political Parties in America*. American Politics and Political Economy Series The University of Chicago Press.
- Austen-Smith, David. 1992a. "Explaining the Vote: Constituency Constraints on Sophisticated Voting." *American Journal of Political Science* 36(1):68–95.

- Austen-Smith, David. 1992b. "Strategic Models of Talk in Political Decision Making." *International Political Science Review* 13(1):45–58.
- Bartels, Larry M. 1998. *Politicians and Party Politics*. The Johns Hopkins University Press chapter Where the Ducks Are: Voting Power in a Party System, pp. 43–79.
- Bartels, Larry M. 2008. *Unequal Democracy: The Political Economy of the New Gilded Age*. Princeton University Press.
- Bonica, Adam. 2013. "Ideology and Interests in the Political Marketplace." *American Journal of Political Science* 57(2):294–311.
- Bonica, Adam, Nolan McCarty, Keith T. Poole & Howard Rosenthal. 2013. "Why Hasn't Democracy Slowed Rising Inequality?" *Journal of Economic Perspectives* 27(3):103–124.
- Bovitz, Gregory L. & Jamie L. Carson. 2006. "Position-Taking and Electoral Accountability in the U.S. House of Representatives." *Political Research Quarterly* 59(2):297–312.
- Box-Steffensmeier, Janet M., Laura W. Arnold & Christopher J. W. Zorn. 1997. "The Strategic Timing of Position Taking in Congress: A Study of the North American Free Trade Agreement." *American Political Science Review* 91(2):324–338.
- Clinton, Joshua D., Simon Jackman & Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *The American Political Science Review* 98(2):355–370.
- Cover, Albert D. 1980. "Contacting Congressional Constituents: Some Patterns of Perquisite Use." *American Journal of Political Science* 24(1):125–135.
- Cox, G. W. & M. D. McCubbins. 1986. "Electoral Politics as a Redistributive Game." *The Journal of Politics* 59:17–45.
- Cox, Gary W. 2006. Swing Voters, Core Voters and Distributive Politics. In *Paper presented at the Conference on Representation and Popular Rule*.
- Delli-Carpini, Michael X. & Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, CT: Yale University Press.
- Dickson, Eric S., Catherine Hafer & Dimitri Landa. 2008. "Cognition and Strategy: A Deliberation Experiment." *The Journal of Politics* 70(4):974–989.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.
- Druckman, James N., Martin J. Kifer & Michael Parkin. 2010. "Timeless Strategy Meets New Medium: Going Negative on Congressional Campaign Web Sites, 2002-2006." *Political Communication* 27(1):88–103.
- Edison Research. 2008. National Election Pool 2008 Exit Polls. Technical report The Roper Center for Public Opinion Research, University of Connecticut.

- Feldman, Stanley. 2009. "Apathy in America, 1960-1984: Causes and Consequences of Citizen Political Indifference." *Public Opinion Quarterly* 52(2):270.
- Fenno, Richard. 1978. *Home Style: House Members in Their Districts*. Addison-Wesley Educational Publishers Inc.
- Fiorina, Morris P. 1977. *Congress: Keystone of the Washington Establishment*. Yale University Press.
- Franklin, Charles H. 1991. "Eschewing Obfuscation? Campaigns and the Perception of U.S. Senate Incumbents." *The American Political Science Review* 85(4):1193–1214.
- Gilens, Martin. 2009. "Preference Gaps and Inequality in Representation." *Political Science & Politics* 42(2):335–341.
- Glassman, Matthew Eric. 2007. "Franking Privilege: Historical Development and Options for Change." *CRS Report for Congress* .
- Gratton, Gabriele. 2010. "Electoral Competition and Information Aggregation." *Job Market Paper - Boston University* .
- Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18:1–35.
- Grimmer, Justin. 2013. "Appropriators not Position Takers: The Distorting Effects of Electoral Incentives on Congressional Representation." *American Journal of Political Science* .
- Grimmer, Justin. forthcoming. *Representational Style: What Legislators Say and Why It Matters*.
- Grimmer, Justin & Gary King. 2010. "Quantitative Discovery from Qualitative Information: A General-Purpose Document Clustering Methodology." *MPSA* 2010.
- Gulati, Girish J. & Christine B. Williams. 2010. "Congressional Candidates' Use of YouTube in 2008: Its Frequency and Rationale." *Journal of Information Technology and Politics* 7(2):93–109.
- Hacker, Jacob S. & Paul Pierson. 2011. *Winner-Take-All Politics: How Washington Made the Rich Richer—and Turned Its Back on the Middle Class*. Simon and Schuster.
- Heidhues, Paul & Johan Lagerlof. 2003. "Hiding Information in Electoral Competition." *Games and Economic Behavior* 42(1):48–74.
- Herrnson, Paul S., Atiya Kai Stokes-Brown & Matthew Hindman. 2007. "Campaign Politics and the Digital Divide: Constituency Characteristics, Strategic Considerations, and Candidate Internet Use in State Legislative Elections." *Political Research Quarterly* 60(1):31–42.

- Hersh, Eitan. 2011. "At the Mercy of Data: Campaigns' Reliance on Available Information in Mobilizing Supporters."
- Highton, Benjamin & Michael S. Rocca. 2005. "Beyond the Roll-Call Arena: The Determinants of Position Taking in Congress." *Political Research Quarterly* 58(2):303–316.
- Jonson, Edward. 2009. "True cost of reform: not accurately portrayed."
- King, Gary, Michael Tomz & Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2):347–61.
- Koger, Gregory. 2003. "Position-Taking and Cosponsorship in the U.S. House." *Legislative Studies Quarterly* 28:225–246.
- Landa, Dimitri & Adam Meirowitz. 2009. "Game Theory, Information, and Deliberative Democracy." *American Journal of Political Science* 53(2):427–444.
- Levy, Steven. 2008. "In Every Voter, A 'Microtarget'." *The Washington Post* p. D01.
- Lindbeck, A. & J. W. Weibull. 1987. "Balanced-budget Redistribution as the Outcome of Political Competition." *Public Choice* 54:273–297.
- Lipinski, Daniel. 2004. *Congressional Communication*. Ann Arbor: University of Michigan Press.
- Ludwig, Mark D. 2010. "Online Communication of Candidate Traits During the 2008 Presidential Election." Presented to the Annual Meeting of the Midwest Political Science Association, Chicago, April 2010.
- Mayhew, David R. 1974. *Congress: The Electoral Connection*. New Haven and London: Yale University Press.
- Meirowitz, Adam. 2005. "Keeping the other candidate guessing: Electoral competition when preferences are private information." *Public Choice* 122:299–318.
- Neiheisel, Jacob R. & Sarah Niebler. 2010. "The Strategic Use of Party Brand Labels in Political Advertising, 1998-2008." MPSA 2010.
- Poole, Keith T. & Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29:357–84.
- Poole, Keith T. & Howard Rosenthal. 1997. *Congress a Political-Economic History of Roll Call Voting*. Oxford University Press.
- Raftery, Adrian E. & Steven M. Lewis. 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall chapter 7, pp. 115–130.
- Richardson, Lilliard E., Linda M. Daugherty & Patricia K. Freeman. 2001. "E-Mail Contact between Constituents and State Legislators." *State & Local Government Review* 33(1):23–31.

- Rocca, Michael S. & Gabriel R. Sanchez. 2008. "The Effect of Race and Ethnicity on Bill Sponsorship and Cosponsorship in Congress." *American Politics Research* 36(1):130–152.
- Rock, Vivkie. 2008. "Ensign Sends Out Reprinted Editorial From The Washington Post."
- Rutenberg, Jim. 2013. "Data You Can Believe In: The Obama Campaign's Digital Masterminds Cash In." *The New York Times Magazine* .
- Sellers, Patrick. 2010. *Cycles of Spin*. Cambridge University Press.
- Tausanovitch, Chris & Christopher Warshaw. 2012. "Measuring Constituent Policy Preferences in Congress, State Legislatures and Cities." *Journal of Politics* .
- Tierney, Luke. 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall chapter Introduction to General State-Space Markov Chain Theory, pp. 59–74.
- Tomz, Michael, Jason Wittenberg & Gary King. 2003. "CLARIFY: Software for Interpreting and Presenting Statistical Results. Version 2.1." Cambridge, MA: Harvard University.
- Trammell, Kaye D. & Andrew Paul Williams. 2004. "Beyond Direct Mail: Evaluating Candidate E-Mail Messages in the 2002 Florida Gubernatorial Campaign." *Journal of E-Government* 1(1):105–122.
- Vavreck, Lynn. 2009. *The Message Matters*. Princeton University Press.
- Wigley, Shelley. 2011. "Telling your own bad news: Eliot Spitzer and a test of the stealing thunder strategy." *Public Relations Review* 37:50–56.
- Williams, Christine B. & Girish J. Gulati. 2010. "Communicating with Constituents in 140 Characters or Less: Twitter and the Diffusion of Technology Innovation in the United States Congress." Prepared for delivery at the 2010 Annual Meeting of the Midwest Political Science Association, Chicago, IL, April 22-5, 2009.
- Xenos, Michael A. & Kirsten A. Foot. 2005. "Politics As Usual, or Politics Unusual? Position Taking and Dialogue on Campaign Websites in the 2002 U.S. Elections." *Journal of Communication* 55(1):169–185.
- Yiannakis, Diana Evans. 1982. "House Members' Communication Styles: Newsletters and Press Releases." *The Journal of Politics* 44(4):1049–1071.

Figure 1

All Roll Call Votes

Revealed Roll Call Votes

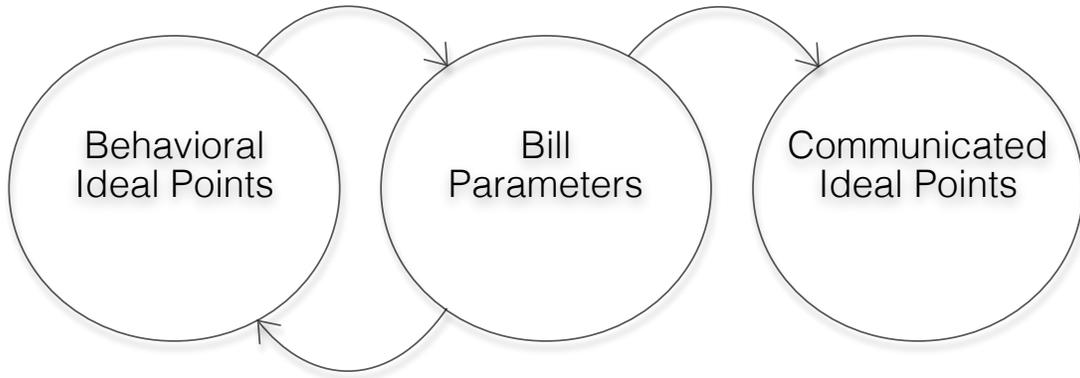


Figure 2

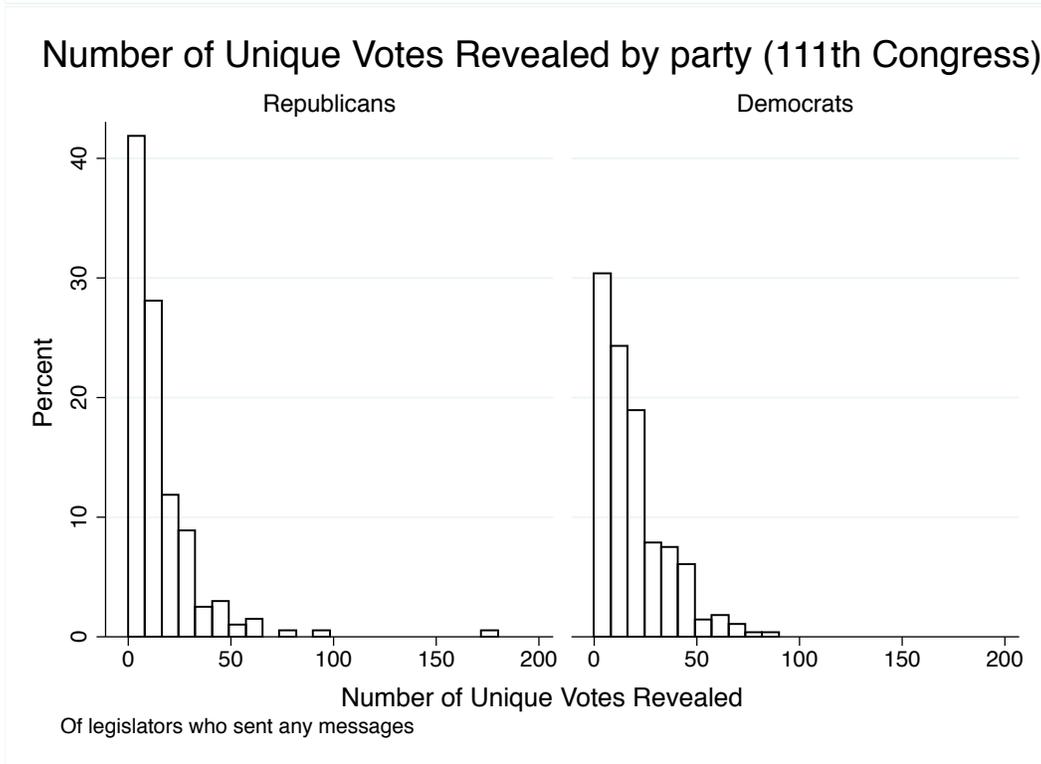
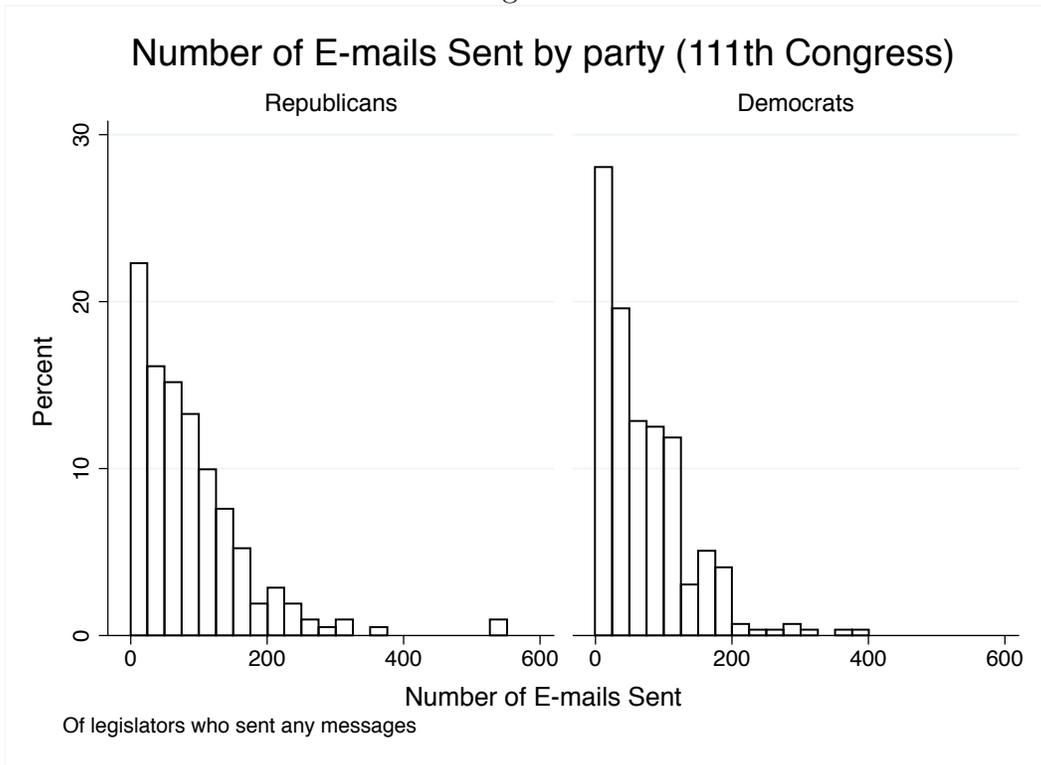


Figure 3

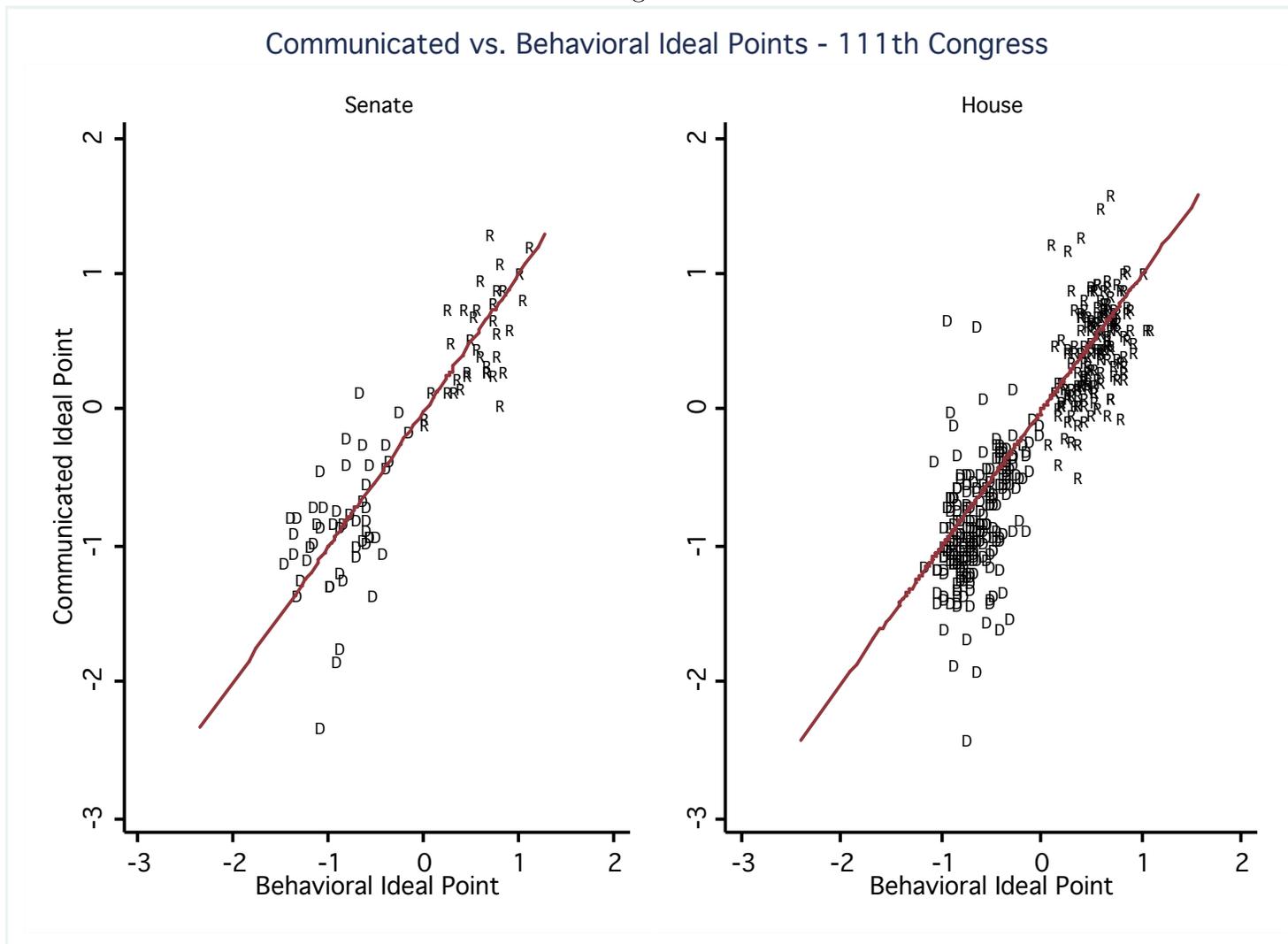


Figure 4

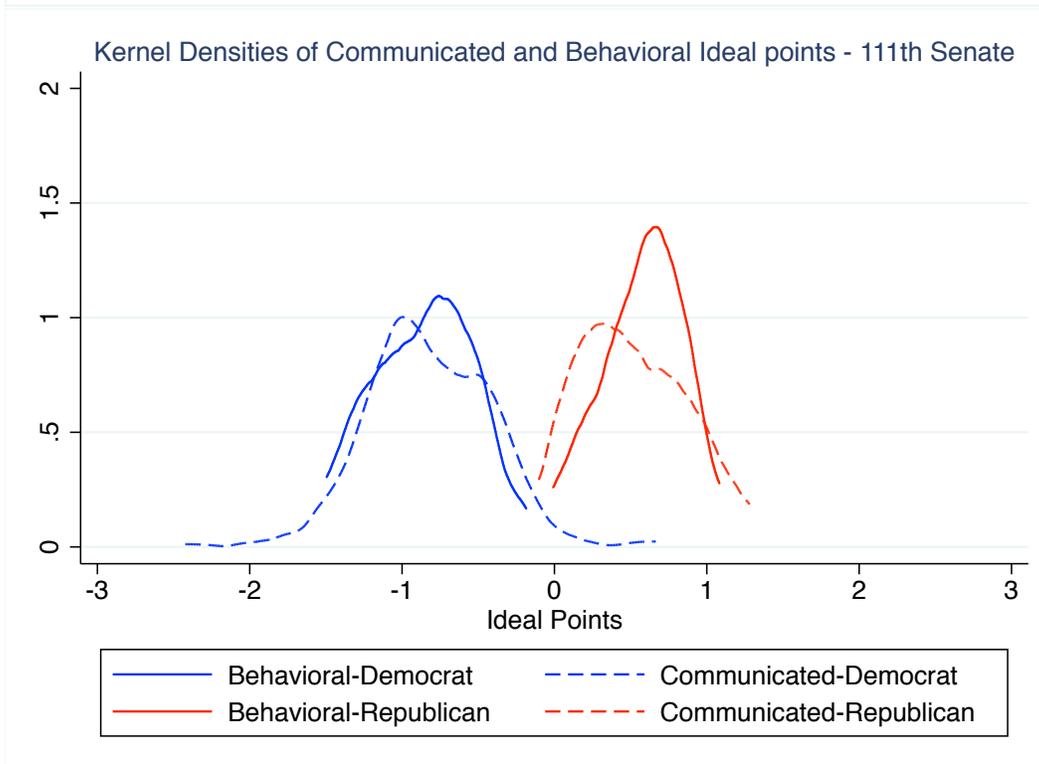
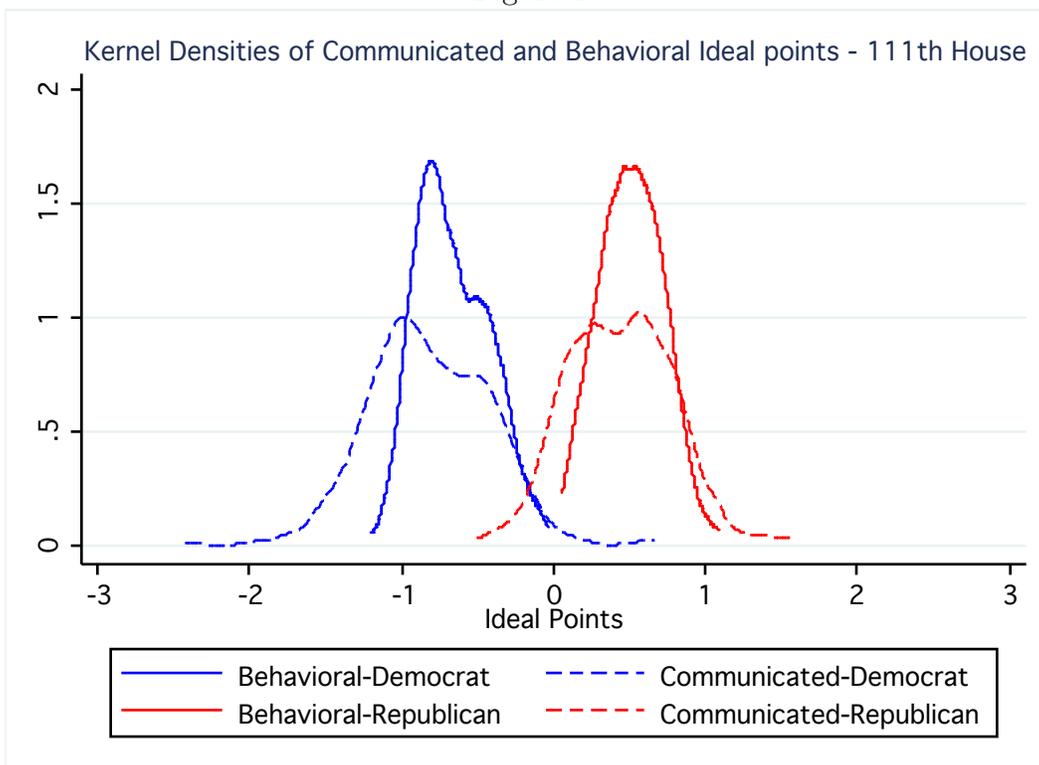


Figure 5

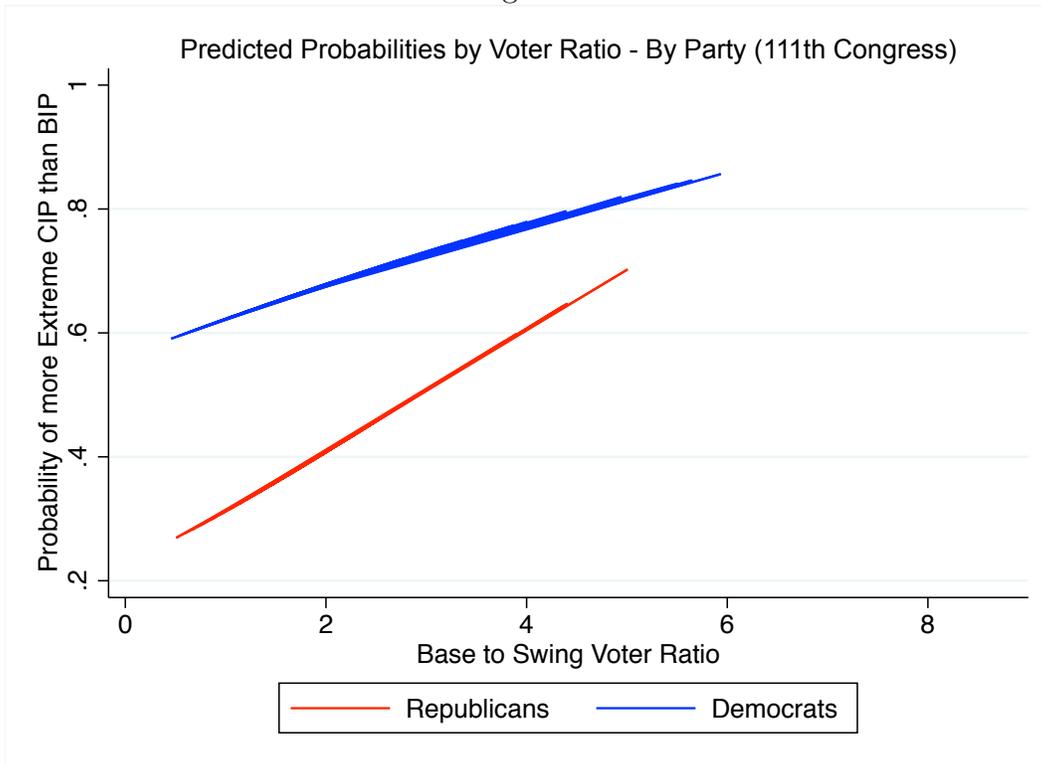


Table 1: Ratio of Partisans to Independents - 2008 National Exit Polls

Ratio Measure	Minimum	Max	Mean	(s.e.)	N
Democrats to Independents					
All Legislators	0.39	5.93	1.56	(0.90)	346
Rep. Legislators	0.39	5.20	1.41	(0.83)	158
Dem. Legislators	0.47	5.93	1.69	(0.93)	188
Republicans to Independents					
All Legislators	0.11	5.00	1.32	(0.73)	346
Rep. Legislators	0.51	5.00	1.63	(0.77)	158
Dem. Legislators	0.11	3.35	1.05	(0.57)	188

Table 2: Ideological Misrepresentation by Party - 111th Congress

Type of Ideological Misrepresentation	Republicans	Avg. BIP-CIP Diff.	Democrats	Avg. BIP-CIP Diff.
More Extreme CIP than BIP	32%	+0.24	66%	-0.32
More Moderate CIP than BIP	68%	-0.26	34%	+0.26

Table 3: More Extreme CIP vs. BIP Points Hypothesis and Voting Behavior - 111th Congress

	More Extreme Communication (Probit)	(s.e.)	More Extreme Voting (OLS)	(s.e.)
Base to Swing Voter Ratio	0.36*	(0.12)	0.03	(0.02)
Campaign Finance Score	0.34	(0.41)	0.55*	(0.08)
Campaign Finance Score ²	0.16	(0.16)	0.23*	(0.03)
Overall District Ideology	-0.49	(0.67)	-0.01	(0.11)
Overall District Ideology ²	-0.14	(1.04)	0.63*	(0.15)
BIP	0.41	(0.42)	--	--
BIP ²	-1.08*	(0.31)	--	--
Percent under Poverty Line	-0.01	(0.03)	-0.01*	(0.00)
Median District Age	-0.06	(0.04)	0.01	(0.01)
Percent Black in District	-0.01	(0.01)	0.00	(0.00)
Percent Hispanic in District	0.03*	(0.01)	0.00	(0.00)
Black Legislator	-0.29	(0.51)	0.08	(0.07)
Hispanic Legislator	-0.41	(1.05)	0.23*	(0.07)
Female Legislator	0.18	(0.21)	-0.03	(0.04)
Seniority	-0.01	(0.01)	0.01*	(0.00)
Party Leader	1.23	(0.66)	-0.03	(0.07)
Running in 2010	-0.14	0.33	0.00	(0.08)
Primary Ballot	0.32	(0.24)	-0.00	(0.04)
Partisan Registration	0.00	(0.27)	0.02	(0.05)
Partisan Registration & Primary Ballot	0.07	(0.25)	0.04	(0.05)
Democrat	0.98*	(0.49)	-0.05	(0.08)
Senate	-0.13	(0.22)	0.24*	(0.05)
Closed Primary System	-0.19	(0.23)	-0.08	(0.04)
Constant	1.56	(1.48)	0.31	(0.32)
(Pseudo) R ²	0.16		0.24	
N	316		316	

White standard errors in parentheses.

* $p < 0.05$

A Appendix - Text Coding Information

A vote is considered to be revealed if one of two things occurs:

1.) A legislator explicitly announces how she intends to vote on specific legislation, or how she voted on specific legislation.

2.) A legislator explicitly announces that she authored, introduced, sponsored, or co-sponsored a bill in which a vote was taken.²⁴

Each vote is only counted once per message.

To determine if a message contains a vote revelation each of the 40,957 messages were processed by a computer to search for the terms: vot*, cosponsor*, co-sponsor*, co-author*, coauthor*, reintroduc*, sponsor*, introduc*, author*, legislat*, bill, act, resolution.²⁵ This process resulted in 21,968 messages. Each message was then read by the author to confirm that the terms picked up actually referred to a vote in the 111th Congress. Human processing was necessary at this step because occasionally the computer would flag messages with content of the sort, “This day, in 1920 women were granted the right to **vote**.” These messages are not indicative of any specific action taken on the part of the legislator in the 111th Congress so they are discarded before moving on to the next stage of processing.

After all of the computer flagged messages were human processed there were 15,865 confirmed messages with vote revelations. This number was reduced to 15,690 when only looking at voting members of the Congress.²⁶ Each of the remaining 15,690 messages were read by either the author or a highly trained team of research assistants. Human coding was necessary at this step because of the less than straightforward communication styles of many legislators.

For each message the coder read the full text and used online sources to find the bill number and specific name of the bill on which a vote was revealed.²⁷ This step was the most time consuming part of the task because legislators are prone to speak in opaque language when referring to their votes. The following is an example from the data.

In November 2010 a Congresswoman sent a message that was included in the potential vote revealing bin because it contained these sentences:

²⁴There is only one instance in which a legislator co-sponsored a bill that he did not eventually vote for, and this legislator sent an e-newsletter explaining the discrepancy. He is coded to have revealed a no vote on this legislation.

²⁵A series of Boolean operators and message filters written by author and available upon request. Appending an asterisk to a word is the regular expression quantifier to search for the root and any type of ending. For instance vot* will return: vote, voting, voted, voter, votes, votable, voteable, voteless, and so on.

²⁶Non-voting delegates from D.C. Puerto Rico, the Virgin Islands, and the Northern Mariana Islands were coded when they indicated how they would have voted on certain bills, but they are removed from the analysis.

²⁷thomas.gov and opencongress.org

- “Congresswoman X urges Congress not to give up on extending unemployment benefits, which fell short of a two-thirds majority **vote**.”
- “I’m very disappointed that there weren’t enough **votes** today to extend unemployment benefits and I think its outrageous if Congress allows them to expire next month, just as winter is beginning and a few weeks before Christmas, X said.”
- “Although the **vote** on the bill in the House today was 258-154 in favor, because it was brought up under ‘fast-track’ rules, it didnt achieve the necessary two-thirds majority to pass.”

In the remaining text of the message, Congresswoman X revealed that she was a co-sponsor of the bill in question by saying, “X co-sponsored a bill that would extend unemployment benefits by three months.” and then decries that the super majority of the chamber did not vote for this.²⁸ Using the date and vote description, “extend unemployment benefits” the researcher can then go to online sources to figure out exactly what bill the legislator is talking about and then record the revealed vote in the data matrix.

In this example, opencongress.org helps find the specific vote because the website records votes by day. After looking at all votes on the day the legislator indicates, it is clear that the Emergency Unemployment Compensation Continuation Act (H.R. 6419) is the match for the vote legislator X is referring to for three reasons. The vote totals match exactly the vote totals in the message, the author of the message voted on this bill in the manner she described, and she is also a listed co-sponsor of the bill. With this additional confirming information, the coder can enter the revealed vote.²⁹

B Appendix - Technical Details

I obtained communicated ideal points (CIP) and behavioral ideal points (BIP) estimates based on a modification of a two-parameter item response model applied to roll call votes as detailed in Clinton, Jackman, and Rivers (2004). Consider a standard one-dimensional spatial voting model with $i = 1, \dots, n$ legislators each with a quadratic utility function over policy outcomes and $j = 1, \dots, m$ roll call votes. Each roll call can be characterized by two points corresponding to a ‘Yea’ (ζ_j) vote and ‘Nay’ (ψ_j) vote and each legislator is said to have some latent ideological point (x_i) that may be estimated by modeling each vote choice as a probabilistic decision where the utility for a Yea vote on bill j for legislator i can be expressed as:

$$U_i(y_{ij}) = -\|x_i - \zeta_j\|^2 + v_{ij}$$

²⁸While this is not a simple vote claim of the, “I vote for bill X”, the content and context of the message allows coders (and importantly, constituents) to infer Congresswoman’s vote in this situation.

²⁹Because of the necessary detective work in matching vote revelations to the actual roll call data, it was not possible to blind the gender of the legislators to the coders. To hedge against biased coding, there was no specific research aim communicated when messages were coded other than the desire to make a comprehensive list of all votes revealed by every legislator.

where v_{ij} is the stochastic element of the utility in voting Yea and $\|\bullet\|$ is the euclidean norm. The corresponding utility for a Nay vote is:

$$U_i(y_{ij}) = -\|x_i - \psi_j\|^2 + \eta_{ij}$$

where η_{ij} represents the stochastic component of voting Nay. Assume $(v_{ij}) \sim N(0, 1)$ and $(\eta_{ij}) \sim N(0, 1)$. Given that a legislator is a rational utility optimizer, she will vote yes on bill j if $U_i(\zeta_j) > U_i(\psi_j)$. Following Clinton, Jackman, and Rivers (2004), this means that the problem of identifying legislator ideal points can be more compactly expressed as:

$$P(y_{ij}^*) = \Phi(\beta_j' x_i = \alpha_j)$$

where Φ is the probit function, β_j is the direction of the j^{th} proposal relative to the Nay outcome and α_j is the discriminability of each vote to differentiate legislator ideal points.

This set up is straightforward, but unidentified given that the only observable data are the votes legislators take and the votes legislators communicate to constituents. To overcome this problem I rely on a series of Bayesian Markov Chain Monte Carlo (MCMC) simulations and a few non-controversial assumptions to identify the model.

The initial step in this approach is to adopt priors for each of the parameters to be estimated. Throughout, I used intentionally vague priors.³⁰ To obtain estimates of the parameters of interest, the MCMC algorithm successively samples the posterior density of each variable, conditional on previous values of all the other variables over t iterations, where $t = 10,000$ in my implementation. Each iteration proceeds as follows.

First, estimate a value from the density of utility differentials for each Yea versus Nay voting decisions for each legislator (y_{ij}^*) conditional on $\beta_j^{(t-1)}$, $\alpha_j^{(t-1)}$, $x_i^{(t-1)}$, and y_{ij} where $(t-1)$ indicates the value of each quantity at the previous iteration or the starting value when $t = 1$. Notice that y_{ij} does not change over iterations because this represents the actual voting history data. Because of the probabilistic nature of the voting model I must estimate (y_{ij}^*) putting the actual voting data to use by constraining these estimates such that at each iteration y_{ij}^* is sampled from one of two densities dependent on the actual vote. If $y_{ij} = 1$ then:

$$(y_{ij}^*)|(x_i^{(t-1)}, \beta_j^{(t-1)}, \alpha_j^{(t-1)}) \sim N(\mu_{ij}, 1)I(y_{ij}^* \geq 0)$$

if the $y_{ij} = 0$:

$$(y_{ij}^*)|(x_i^{(t-1)}, \beta_j^{(t-1)}, \alpha_j^{(t-1)}) \sim N(\mu_{ij}, 1)I(y_{ij}^* < 0)$$

where I is an indicator function. These are equivalent to a truncated normal distribution.

Second, estimate the bill parameters, β_j, α_j given the previous values for all other quantities for each bill. This is essentially regressing y_{ij}^* on $x_i^{(t-1)}$ where the resulting values for the constant and coefficient make up the posterior densities for β_j and α_j . Third, estimate

³⁰However, it has been shown that as the data get large – as in roll call records – and as the number of iterations increases, these prior starting values contribute less and less to end results (Tierney 1996, Raftery & Lewis 1996).

x_i given the previous $y_{ij}^*, \beta_j, \alpha_j$ values by running a regression for each legislator over the (t-1) values for the bill parameters.

If the algorithm were to stop here it would follow the approach of Clinton, Jackman, and Rivers (2004) to the letter. However, I need to estimate an additional value of the communicated ideal points (xc_i). To do so, I need to have a different set of y_{ij}^* 's based on the observed data of communicated votes. Call communicated votes yc_{ij} , and call the utility differential for legislator i on roll call j between voting yea and nay on those communicated votes yc_{ij}^* subject to the same constraints as listed in the first step above.

Instead of re-running the algorithm using only the communicated votes, I embed the estimation of yc_{ij}^* and xc_i into the the aforementioned sequence. Assuming $g()$ is a sampling function of part of each section of joint posterior density, the steps that make up the modified MCMC can be compactly listed as:

$$g((y_{ij}^*)|y_{ij}, x_i^{(t-1)}, \beta_j^{(t-1)}, \alpha_j^{(t-1)}) \quad (1)$$

$$g((yc_{ij}^*)|yc_{ij}, xc_i^{(t-1)}, \beta_j^{(t-1)}, \alpha_j^{(t-1)}) \quad (2)$$

$$g((\beta_j, \alpha_j)|y_{ij}^{*(t-1)}, x_i^{(t-1)},) \quad (3)$$

$$g((x_i)|y_{ij}^*, \beta_j^{(t-1)}, \alpha_j^{(t-1)}) \quad (4)$$

$$g((xc_i)|yc_{ij}^*, \beta_j^{(t-1)}, \alpha_j^{(t-1)}) \quad (5)$$

This is repeated 10,000 times.

This modification allows the bill parameters to be estimated using the full set of votes, and then uses those results to inform the communicated ideal point rather than generating a new set of bill parameters using only communicated votes. This is important because the features of the roll calls does not change when a legislator chooses to reveal her votes on selected bills; only her communicated ideological placement changes.

Lastly, to orient the scale, I pin two legislators and the ends of the one-dimensional ideological space. This is a non-controversial assumption necessary to identify the model. The most liberal and most conservative legislator as determined by DW-NOMINATE scores over all votes taken in the 111th Congress are the pins.³¹ Using this model, I estimate the BIPs, bill cut points, and the CIPs along with standard errors for each point of interest.

³¹The analysis and rank ordering of legislators are reported by Royce Carroll, Jeff Lewis, James Lo, Nolan McCarty, Keith Poole, and Howard Rosenthal at voteview.com