

On The Meaning of Survey Reports of Roll Call “Votes”*

Forthcoming, *American Journal of Political Science*

Seth J. Hill
University of California, San Diego[†]

Gregory A. Huber
Yale University[‡]

September 4, 2018

Short title: Survey Reports of Roll Call “Votes”

Keywords: Representation; roll call votes; survey response; Congress

*We acknowledge research support from the Center for the Study of American Politics and the Institution for Social and Policy Studies at Yale. We thank anonymous reviewers, the Editors, Joe Bafumi, David Broockman, John Bullock, Josh Clinton, Al Fang, Andrew Gooch, Eitan Hersh, Stephen Jessee, Gabe Lenz, Neil Malhotra, and Chris Tausanovitch for feedback.

[†]Department of Political Science, 9500 Gilman Drive #0521, La Jolla, CA 92093-0521; sjhill@ucsd.edu, <http://www.sethjhill.com>.

[‡]Department of Political Science and Institution for Social and Policy Studies, 77 Prospect Street, PO Box 208209, New Haven, CT 06520-8209; gregory.huber@yale.edu, <http://huber.research.yale.edu>.

Abstract: Contemporary efforts to evaluate representation often compare survey measures of how citizens say they would vote on legislation to what elected officials do in office. These comparisons generally suggest poor representation. We argue here that this common design is unlikely to effectively evaluate representation because responses to survey questions differ in important aspects from voting in legislatures. Measurement error and construct validity undermine the comparison. Three survey experiments show that providing partisan and non-partisan information readily available to legislators materially changes respondents' expressed preferences on roll call votes. With information, expressed policy positions are both less centrist and more closely matched to legislator behavior in their preferred party. Respondents also appear aware of their own lack of knowledge in evaluating roll call policy votes. The treatment effect of information decreases in confidence judging policy in that area. We show similar patterns for respondent opinions on Supreme Court decisions.

Replication Materials: The data, code, and any additional materials required to replicate all analyses in this article are available on the American Journal of Political Science Dataverse within the Harvard Dataverse Network, at: <https://doi.org/10.7910/DVN/SSEN5A>.

Word count: 7,914.

Understanding representation is a central and longstanding research topic in political science. How much do elected officials do what citizens want? Among the most common empirical assessments of the quality of representation are studies that calculate the correspondence between an elected official's behavior and the preferences of his or her constituents. Two potential ways to measure correspondence are to ask (1) would the citizen do something different if granted the legislator's vote, or (2) would she punish or reward the legislator on the basis of the legislator's vote if she were aware of it? Recently, both forms of research have turned to comparing the roll call votes cast by members of Congress with survey measures of the preferences of constituents on those same roll calls. To measure the first version of representation, the researcher compares whether survey respondents' expressed preferences comport with how their member voted. To measure the second version of representation, scholars inform respondents how the member voted and see if it changes respondents' evaluations of, or vote intentions for, the member. In this article, we argue that both of these approaches are subject to large survey measurement challenges that call into question inferences drawn about representation.

The key challenge in comparing the choices representatives make to what constituents want is creating equivalent measures between the two contexts (e.g., Achen, 1977). For example, if a survey respondent believes that abortion rights should be restricted "in some situations," how should that preference map onto voting on a particular bill concerning abortion policy? Scholars have turned to survey measures designed to directly place citizens' choices in the same policy space as their legislative counterparts. The most promising approach in this regard is survey roll call measures that ask respondents how they would vote on the exact piece of legislation that has been considered in the legislature. Such measures, used either in isolation or as a means to bridge a larger set of survey items into a common ideological space with legislator behavior, are presumed to provide measures of citizen preferences that are comparable to legislative roll call votes because the survey respondent is making the same binary choice as the legislator. With such measures, researchers have conducted several important studies assessing the quality of representation, with a canonical finding being that legislators appear substantially more extreme than both the electorate

as a whole and their partisan counterparts in the general population (Bafumi and Herron, 2010).

However, there are many reasons that a survey response to a roll call item might not accurately capture the same contextual and strategic dynamics associated with voting in the US Congress. For example, respondents may not appreciate the importance of partisan agenda control and building legislative coalitions, bargaining before an electoral audience, long- versus short-term tradeoffs, etc. Although these limitations are acknowledged in some prior work, this paper argues that such challenges are specific instances of more general sources of systematic error in survey roll call measures. By *systematic error*, we mean a combination of instrumentation-induced measurement error (question wording, etc.) and a lack of construct validity (survey respondents do not possess a preference comparable to legislators on roll call votes).

Each of these manifestations of error call into question the assumption that survey roll call votes yield estimates of citizen preferences that are appropriate to compare to legislator behavior in evaluating quality of representation. Respondents and policymakers are likely not in the same policy space when one is answering survey questions and the other is casting roll call votes. Our argument is that the central problem confronting such comparisons is that voters know very little about the items they are asked to “vote on” in a survey context. As a consequence, responses to roll call items likely have many of the same features that the survey design literature argues challenge the validity and reliability of survey measurement more generally (e.g., Groves et al., 2009). These “top of the head” responses to survey roll call votes are therefore unlikely to accurately reflect how individuals would act either in more consequential settings or with greater reflection and information. A strong take on our argument – advocated by some of our anonymous referees – would conclude that respondents can never approximate the decisions made by legislators. While we are not prepared to go this far, we do advocate that applied scholars think much more carefully about what they are measuring and what they are assuming.

To demonstrate the importance of this argument for measuring citizen preferences, we present results from three survey experiments with more than 9,000 respondents who voted on survey roll call items similar to those used in previous studies. Departing from most prior work, we

randomly varied the amount and type of information available to respondents when they answered the questions. We show that three different types of simple information about the bill materially change the level of support survey respondents express for that roll call. We first examine the effect of providing party margins on the vote when it was taken in Congress. The party margin provides rough information about the strategic, policy, and political implications surrounding the roll call vote that was surely available to a member of Congress when voting but not as readily available to most survey respondents. If the treatment systematically changes the rate of support for the roll call vote among survey respondents, it suggests the survey responses absent that information may not accurately describe how the subject would vote in the legislature.

Prior work in public opinion finds that party cues affect measured survey responses in a wide variety of contexts. While the reasons for these effects are contested, in light of this literature we believe it is also useful to test whether party cues generate partisan fealty and whether non-partisan cues also alter survey roll call responses. First, we show below that the effect of providing information about party voting on a specific bill is moderated by how confident respondents feel evaluating policy in that domain. The treatment effect on roll call support of providing the party vote margin is decreasing in the confidence respondents report in judging policy in the issue area of the roll call. This suggests that, rather than blind obedience, citizens use additional information when they know their own judgments are less informed.

Second, we investigate the effect of providing two non-partisan heuristics on the level of support survey respondents express on roll call votes. In two conditions, we informed respondents about the overall margin in the chamber (i.e., without the party breakdown), and we supplied the budgetary implications of a bill as assessed by the non-partisan Congressional Budget Office (CBO). The CBO score gives individuals a piece of information, apart from the domain-specific policy content of a bill, about how a bill is likely to affect overall government revenue and spending, information that few respondents likely have at the time they are surveyed.

Finally, to demonstrate the general implications of this work for efforts to place elite and mass opinion in the same policy space, we present a parallel party cue experiment in which respondents

express opinions on important Supreme Court cases.

To briefly summarize our findings, we show that providing either partisan or non-partisan information changes expressed preferences about important roll call votes cast by Congress and partisan information affects votes on important decisions by the Supreme Court. The votes we selected from our study vary from “easy” issues on which individuals likely have considered their own opinions (e.g., abortion) to “hard” issues where many individuals have likely never considered the issue or understand its technical details (e.g., requiring financial advisors to act as fiduciaries). They also span votes that are partisan (e.g., defunding Planned Parenthood) to issues that split the parties internally (e.g., requiring warrants for electronic surveillance). Substantively, we find that elected officials appear less extreme relative to their partisan counterparts in the public when the public’s preferences are measured in the condition with the party split, suggesting current findings about “leapfrog representation” (Bafumi and Herron, 2010) may be partially an artifact of survey error. More broadly, our results provide experimental evidence that using standard roll call survey items to create estimates of citizen preferences comparable to legislator behavior does not address the fundamental problem of placing the two sets of choices in a common space.

In addition to providing insights for ongoing work about how best to measure citizen preferences and compare them to legislative behavior, this research also has implications for work on representation more generally. Survey data provide a useful metric for understanding citizen preferences if those responses are informative of those preferences. If, as we find, however, measured policy preferences are sensitive to context, it raises important doubts as to how seriously analysts should (and citizens do) treat those responses. We provide some preliminary ideas about ways to move forward in light of this finding in our concluding discussion.

Our results and the core of our design also point toward future research that can decompose the meaning and consequences of informational cues about policies. Exploring the influence of different types of information (e.g., group or leadership endorsements, substantive policy information) across policy areas and types of legislative action (e.g., procedural versus final passage votes) in an experimental context can help us understand when and under what conditions our tools of

measurement provide estimates of citizen preferences that are more comparable to the legislators' policy space. That sort of work, in conjunction with the results presented in this paper, also speaks to research that uses statistical models such as item-response theory to scale respondent policy preferences into underlying spaces. While one advantage to these procedures is to aggregate individual measurement error, our findings suggest that measurement error is systematic. We explore item-response models briefly in the Appendix, but our design provides a framework to evaluate the influence of additional policy information on scaling procedures.

Assessing Representation Using Survey Roll Call Measures

Evaluating representation is a rich and diverse area of ongoing research in political science. While there are many different definitions of what constitutes “good” representation or how to measure public preferences, a large body of scholarship in recent decades has focused on evaluating the quality of representation empirically by assessing the correspondence between what elected officials do in office and what the public says they would like those elected official to do. Building on this work, we adopt a simple definition that allows us to make comparisons between a representative's behavior and a citizen's preferences: Does a representative behave in the manner that the voter would if she were granted the legislator's vote? Thus, we follow closely the example of Bafumi and Herron (2010) when “we characterize the extent to which there is congruence between the preferences of voters and the preferences of their corresponding members of Congress (p. 519).”¹

The quantitative study of representation was invigorated by the rise of public opinion surveys. This technology allowed researchers the possibility of directly soliciting citizen policy preferences and comparing those measures to elite behaviors. Seminal in this line of work was Miller and Stokes (1963), which used survey data to compare district-level estimates of public opinion with corresponding behavior of sitting members of Congress.

In response to criticism about existing survey measures of preferences being subject to a variety

¹ This definition sets aside issues of coordination, multidimensionality, and cycling, as well as alternative definitions of representation.

of sampling and measurement error (e.g., Achen, 1977, 1978; Clinton, 2006; Wright, Erikson, and McIver, 1985), scholars have subsequently embraced the idea of asking ordinary citizens how they would vote on specific roll call votes, what we call “survey roll call measures.” The promise of such measures is transparent: Because survey respondents are now making the same binary choice as legislators, their responses are presumably informative of their preferences over outcomes in the same way as legislators’ votes (Lax and Phillips, 2009; Broockman, 2016). To assess the nature of representation, scholars can then compare member votes on specific roll calls to survey respondent preferences elicited about those same roll calls or use a set of roll call votes in joint scaling procedures to provide more precise measures of citizen and legislator preferences in a common space (e.g., Bafumi and Herron, 2010; Clinton, 2006; Gerber and Lewis, 2004; Tausanovitch and Warshaw, 2013). Cumulatively, all of these methods aim to generate measures of citizen preferences comparable to member behavior to assess representation. In general, findings from these different approaches suggest measurable but imperfect congruence between what voters want and how representatives behave.

Approaches that use survey roll call measures to assess representation rely on a built in but often unstated assumption about measurement. If the survey response reflects how the citizen would vote in the legislature, the survey response measures the citizen’s preferences in the policy space of the legislator. If this assumption does not hold, however, then the relevance of survey roll call measures for assessing the quality of representation is in question.

There are a number of reasons that the context a member of Congress faces when voting on a particular bill is different from the context faced by a survey respondent. Setting aside normal issues associated with survey respondents being inattentive (which would itself likely generate measurement error), perhaps the most salient difference between ordinary citizens and legislators is the information they have about the choice they are making. For legislators, their job is on the line when making public roll call votes (Mayhew, 1974). As such, they devote substantial attention and resources to understanding the policy choices they face and the larger political context associated with a yea or nay vote, anticipating how the public will react to that choice (Arnold,

1990). By contrast, ordinary survey respondents may not readily know the particular details of a piece of policy legislation (or procedural vote), the nature of status quo policy, or other strategic considerations that are on the line when voting (for example, how this vote affect subsequent votes on amendments or other bills, bargaining with the president, or elections). Particularly in the contemporary US Congress where many of these concerns are tightly tied to party coalitions and party conflict (e.g., Cox and McCubbins, 1993), survey respondents are unlikely to have access to partisan-relevant information that affects how members of Congress vote on those bills. This list of contextual differences is of course not exhaustive, but generally the incentives facing a member of Congress when voting on a bill are likely very different from those facing a survey respondent.

What remains uncertain is whether altering the context of the survey response to more closely resemble that of the legislature affects patterns of survey response and subsequent estimates of citizen and legislator preferences. If survey responses are largely unaffected by additional information, researchers are warranted in treating survey roll call measures as more meaningful measures of citizen policy preferences.²

Previous research finds that survey responses to political questions change when subjects are provided with cues or information that can substitute for their own lack of policy knowledge. These results are in keeping with the broader literature on survey design that shows that responses to survey questions vary considerably in their reliability, stability, and validity (e.g., Groves et al., 2009; Zaller and Feldman, 1992). At one extreme, for example, survey respondents are willing to express opinions on non-existent bills (Bishop, Tuchfarber, and Oldendick, 1986). The survey design literature generally recommends careful investigation of any survey item to evaluate its value for measuring underlying constructs. Because such empirical work has not been applied to survey roll call measures, however, and because the arguments we discuss above imply such measures are likely subject to substantial error, many of the inferences drawn to date from designs using these measures rest on untested assumptions.³

² Of course one cannot rule out the possibility that some additional informational intervention would affect measured preferences. Theory should guide what information is likely to affect citizens opinions.

³ One indication of the challenges for this assumption is revealed in recent work by Jessee (2016) and Lewis and Tausanovitch (N.d.), who examine the performance of the joint scaling procedures that use multiple survey items and

Suppose, for example, that survey respondents with less information were less likely to “vote” with their party because they did not understand the issue at stake, reducing the correlation between partisanship and measured preferences. This measurement error would have the consequence of making voters appear more centrist (less partisan) than the better-measured preferences of members of legislatures and create a disconnect in representation, exactly as much of the existing literature finds (Bafumi and Herron, 2010; Hill and Tausanovitch, 2015).

Survey responses in general have been shown to be particularly responsive to partisan cues. While this is sometimes taken as evidence that citizens lack true preferences and are instead beholden to elite persuasion (e.g., Zaller, 1992) – a finding that nonetheless calls into question the idea that preferences measured in the absence of those cues reflect how citizens want their elected representatives to behave – citizens also appear self-aware of their own lack of detailed knowledge. For example, survey respondents distinguish among issues, being less likely to reward or punish elected officials in policy domains where they acknowledge that they do not understand the complexity of mapping policy tools to policy outcomes (Gerber et al., 2011). Furthermore, when citizens have a sound informational basis for their expressed opinions, partisan cues appear to have modest effects on expressed citizen preferences (Bullock, 2011). A study related to ours presents evidence that informational interventions do change survey responses to roll call vote items (Lauderdale, 2013). Our study presents different non-partisan cues, considers confidence in policy knowledge as a moderator, breaks down results by party and roll call vote, and considers Supreme Court cases to provide further empirical evidence in support of our argument.

Research design and data collection

In order to assess the importance of the arguments raised in the previous sections, we fielded three survey experiments in which we varied the amount of information provided to respondents when answering survey roll call measures. These questions allow us to estimate citizen preferences

roll call votes to place legislators and survey respondents in the same policy space. Both find that the underlying structure generating legislative votes and survey responses is statistically distinct, but come to different conclusions about the continued usefulness of joint scaling. This suggests that the validity of survey roll call measures cannot be determined by statistical procedures alone.

obtained under different informational conditions. None of these conditions fully replicate what a legislator faces. But, if we find that information readily available to the legislator influences survey responses, being exposed to the legislator’s environment in whole seems likely to do so as well.

The key intervention in each experiment is similar. In our control groups, respondents answered standard survey roll call measures. In our treatment groups, respondents answer the same questions but with one of three kinds of additional information. The “party split” treatment provides the party breakdown of the vote in Congress on that piece of legislation. The intuition for this particular intervention is that it provides a highly salient cue that is almost certainly available to sitting members of Congress when deciding their own actions. Party splits therefore likely reflect both immediate policy differences between the parties, strategic considerations around compromise legislation, amendments, and status quo policy, and the larger political environment in which strategic elites perform before the electoral audience. The party split conveys a range of considerations that are likely relevant for citizens when forming their own opinions about a roll call vote as if they were sitting in Congress. In Appendix Figure [A8](#), we provide evidence that party split information is likely novel to respondents by showing a very weak relationship across roll calls between actual party splits and respondent guesses about the split (in the control [no information] group).

A second treatment is a “chamber split” condition, which provides the breakdown of the vote for the full chamber (House or Senate) in which it was cast. This piece of information is less partisan than the party split treatment, instead providing information about how popular the bill was in the chamber as a whole. While participants may bring their own beliefs about the party breakdown to the intervention and attempt to infer party positions, this is a less precise partisan cue than the party split.

A third treatment provides a synopsis of the Congressional Budget Office (CBO) analysis of the bill voted upon. We simplified and shortened the CBO reports to provide paragraph-length analyses for participants to read. This piece of information should be almost entirely non-partisan, consistent with the mandate for the CBO.

The first study, fielded online with YouGov’s Cooperative Congressional Election Study (CCES) surrounding the 2014 midterm election, asked respondents about eight roll call votes. The second study, fielded online with Survey Sampling International (SSI), expanded the project by asking respondents in 2016 for their opinions across 12 recent roll call votes. The first two studies randomized respondents into the party split or control conditions only. The third study, fielded online with Lucid, Inc. in late 2017 and early 2018, asked respondents for their opinions on 11 roll call votes along with the Supreme Court cases used in Malhotra and Jessee (2014). The third study randomized participants into all four conditions for legislative voting and to two conditions for judicial decisions. We detail the design and sample of each in Appendix Sections A, B, and C, and refer to them as Study 1 (CCES), Study 2 (SSI), and Study 3 (Lucid). For each sample, we use post-stratification weights for all analysis to improve representativeness. The Appendix presents an extensive set of robustness tests related to weighting and variable coding, and provides the text of all treatments.

Roll call question wording for each study is the measure fielded on the CCES Common Content, “Congress considered many important bills over the last few years. For each of the following tell us whether you support or oppose the legislation in principle.”⁴ Each piece of legislation was described using a short bill title followed by a brief description. For example, one item was:

US-Korea Free Trade: Implements the United States-Korea Free Trade Agreement.
(Emphasis in original)

Respondents could indicate either that they “supported” or “opposed” the bill.

In the party split condition, for each item we added the observed party split in the chamber:

US-Korea Free Trade: Implements the United States-Korea Free Trade Agreement.
91% of Republicans voted in favor of the bill, and 31% of Democrats voted in favor of the bill. (Emphasis and underlining in original)

The chamber split condition displayed the phrase “Of members in the [House/Senate] voting on the bill, xx% voted in favor of the bill” instead of the breakdown by party. The CBO condition

⁴ Note that we did not design this question wording and so it does not line up exactly with our definition of representation. We use the standard wording to help our results speak to existing work using these items.

included a synopsis of the text from the CBO report on the bill without any information on the vote outcome.

For each study, we use a between-subjects design to identify the treatment effect of information (party split, chamber split, or CBO score) relative to subjects in the control condition. On the SSI study, we additionally asked respondents at the end of the survey about their confidence to evaluate policy across a set of policy areas that match to the roll call votes asked.

Results: Responses to Survey Roll Call Measures Affected by Party Split Information

In each of the three studies, respondent support for roll call votes cast in Congress is materially affected by the delivery of information about the party split. In particular, we focus on partisan respondents (including partisan “leaners”) because it is for these respondents that partisan voting patterns may provide signals about (more informed) preferences on the vote. In nearly all cases, partisans’ average responses move toward the observed party position in the legislature.

For all three studies, we present our analysis graphically with parallel statistical analysis in Appendix Tables [A6](#), [A7](#), and [A8](#). In [Figure 1](#), we present the effect of the party split intervention on support for each roll call vote from Study 1. For each bill (horizontal axis) reading left to right, we begin by plotting the proportion of Democrats (circles) and Republicans (squares) supporting the bill in the control condition. Next, connected to these points by solid lines are the proportions of Democrats and Republicans supporting the bill in the party split condition. Finally, the last pair of points is the proportion of Democrats and Republicans voting yea on the bill on the House floor, the information provided to respondents in the party split condition.

[[Figure 1](#) about here]

For example, the leftmost frame shows that Democrats moved from supporting the Bipartisan Budget Bill of 2013 at a rate of 76 percent in the control condition to a rate of 79 percent in the party split condition, with 84 percent of Democratic members of the House voting yea. Republican respondents also move towards the rate of Republican members (73 percent yea), from 69 percent

supporting in the control condition to 71 percent in the party split condition.

Two important patterns emerge. First, both Democrats and Republicans expressed support for each bill at rates closer to that observed in the House for their party when informed of the House vote for seven of eight items (88%). For Democrats, the exception is the Simpson-Bowles compromise budget, which was not supported at high rates by either party but was seen, according to some commentators, as particularly anathema to Republicans because it would have raised taxes.⁵ For Republicans, the exception is the bill to end the government shutdown and raise the debt ceiling in 2013, which was endorsed by the Republican leadership and all Democrats, but supported by only about 40% of Republican House members. After learning that the bill was supported at a much higher rate by Democrats than Republicans, Republicans become less supportive.

Second, responses in the party split condition show greater divergence in support between identifiers of the two parties. Apart from the Bipartisan Budget Bill of 2013 where actual partisan divergence in the legislature was low, roll call support diverges more when informed of the party split. Averaging across these eight bills, the average party split in the House is about 56 points. In the control condition, the average party split among survey respondents is 26 points. In the party split condition, this gap grows to 41 points, an increase of more than 60%.

Figure 2 presents parallel results for Study 2. All roll call votes in Study 2 are from the Senate, so the last pair of points shows the proportion of Senate Democrats and Senate Republicans voting yea. Here our analysis also excludes respondents who failed an attention screener.⁶

[Figure 2 about here]

The results shown in Figure 2 are consistent with those in Figure 1. Providing the vote split in the legislature changes expressed support for the bill in the public. Across the 24 party-vote observations in Figure 2, in 17 of them (71%) partisan respondents move toward the observed party vote in the treatment condition.

⁵ See, for example, <https://www.atr.org/conservative-movement-united-against-simpson-bowles-a6822> (retrieved July 12, 2016), summarizing the conservative opposition to the House vote.

⁶ See Appendix Sections B and F for details on the screener and results including all respondents. We include only respondents who pass the screener to weed out survey respondents who are not engaged with the survey instrument and whose responses are therefore likely uninformative of the theoretical arguments we are interested in testing.

Additionally, as with Study 1, support in the party split condition diverges in the party split condition (this is true on average, and for 11 of the 12 individual bills). The average party split in the Senate is about 62 points. Among respondents, the average split moves from 28 points (control) to 40 points (party split), an increase of more than 40%.

Finally, Figure 3 presents support for roll calls in Study 3 from the control and party split conditions. Study 3 roll call votes come from both the House and the Senate. The patterns for Study 3 are consistent with the results from Studies 1 and 2.

[Figure 3 about here]

We selected the roll calls in Study 3 to cover easy issues such as abortion and hard issues such as the fiduciary rule, and also with varying magnitude of party split. One might think that on easy issues participants would have more crystalized opinions less likely influenced by party cues (Tesler, 2015). However, our results suggest that ease of issue is not strongly related to the party split treatment effect. We see notable movement on easier issues like gender discrimination, death tax repeal, and firearm background checks, and little movement on the hard issue of Puerto Rican debt. Appendix Table A8 does show that the largest point estimates obtain on bonus depreciation, presumably a hard issue, but we do not observe an obvious pattern. Instability is present even in high salience and easy issue domains, showing the concerns we identify with roll call items are not restricted to cases where votes in Congress are technically difficult or on novel issues.

Overall, the evidence from the three studies shows that survey roll call measures may understate the degree to which partisans support what their legislators are doing in Congress. When told how the parties in the chamber vote on an issue, expressed support for bills before Congress polarize by party. In the next section, we show that other, non-partisan information readily available to legislators can also change what survey respondents say about important legislation before Congress.

Results: Responses to Survey Roll Call Measures Affected by Non-Partisan Information

In this section, we present treatment effects from Study 3 for the chamber split and CBO conditions. For consistency with our earlier presentation, in each case we compare partisans' responses, by party, to their responses in the control condition.

We begin in Figure 4 by plotting the difference in partisan support between the control condition and the chamber split condition. While there is often movement towards the overall split, treatment effects are heterogeneous. This is consistent with the chamber split itself having multifaceted meaning (for example, respondents may perceive bare majorities as indicating a bill that passed with Republican support over unified Democratic opposition, or as a bipartisan vote drawing support from both parties). Of the eleven bills, five have treatment effects with $p < .05$.

[Figure 4 about here]

In Figure 5, we present the treatment effects for the five bills that had a CBO score. The treatment text in each case – excluding the agricultural bill – presented the net deficit impact of the bill, in all cases increasing the deficit. The treatment decreases average support in each case, with an average effect of about eight points. Democrats become less supportive of the Bonus depreciation, Medicare/SCHIP, and Small Business Tax bills in the CBO condition, and Republicans become less supportive of Death Tax Repeal and Medicare/SCHIP. Of the five bills, two have treatment effects with $p < .05$

[Figure 5 about here]

Why does new information induce changes in expressed roll-call preferences?

Our main result is that citizen responses to survey roll call measures are unstable in the face of changes in available information. Citizens respond to information about how the parties voted in the legislature, the overall chamber split, and information from the Congressional Budget Office by changing their reported preferences on those same votes.

How should one interpret these patterns? Building on previous work (Gerber et al., 2011), we

examine whether self-assessed citizen confidence to evaluate policy moderates the treatment effect of the party split. If how people respond to the informational interventions is correlated with lack of confidence to evaluate good policy in a particular policy domain, it implies that the response to this information arises partly because citizens look to heuristics in situations where they understand their own lack of knowledge. We note that this analysis looks for heterogeneity in the treatment effect by factors not subject to randomization.

We calculate measures of average confidence for each party for each roll call item from Study 2. These measures come from a battery at the end of the study:

There are many different problems that government policies are designed to fix. For each of the following policy areas, we'd like to know how confident you are that you could distinguish good from bad policies. That is, how confident are you that you could evaluate policy in each area?

Respondents could choose Not at all confident (scored 0), A little confident (1), Somewhat confident (2), or Very confident (3).

To avoid treatment affecting measured confidence, we calculate average confidence for each bill and party using responses only among respondents assigned to the control condition. We then assess whether the average effect of the vote split treatment (the absolute difference in support for each item in the party split condition compared to the control condition), is larger for survey roll call measures in policy domains where partisans are less confident. In this aggregate analysis we estimate:

$$\text{Absolute Treatment Effect}_{i,j} = \beta_0 + \beta_1 \times \text{Average Confidence}_{i,j} + \beta_3 \times \text{Republican}_j + \epsilon_{i,j}, \quad (1)$$

using OLS regression where i indexes each bill and j whether the subgroup is Republicans (versus Democrats). We have one observation for each of 22 roll call-cross-party groups.

Parameter estimates appear in Table 1. The dependent variable absolute treatment effect ranges from .004 to .28, with a mean of .09. The average confidence measure ranges from 1.19 to 1.84, with a mean of 1.47. The point estimate in column (1) shows that as confidence increases, the

average treatment effect decreases. Subjects respond more to the vote split in those policy domains where self-assessed confidence is low. Partitioning by partisanship in columns (2) and (3) produces similar results, though with some suggestion of more moderation by confidence for Democrats. None of the estimates is individually statistically significant in this small sample.

[Table 1 about here]

Because this is observational analysis, one concern is that confidence might simply be a proxy for importance. For this reason, at the end of our survey we also asked respondents to gauge the *importance* of acting in each of the same policy areas for which we assessed confidence. Respondents could choose Not at all important (scored 0), Of little importance (1), Somewhat important (2), or One of the most important issues (3).⁷ As before, we calculate average importance in the control group. Including this measure appear in columns (4) through (6), the coefficient on the confidence measure remains negative, although in the pooled estimate the coefficient shrinks by about 24% (B=-.084, s.e.=.10).

Cumulatively these results are uncertain with large sampling variability. But point estimates suggest that survey respondents are more responsive to the treatment in those domains where confidence to evaluate policy is lower. This result also helps to rule out the possibility that respondents simply abandon their own well-formed and deeply-held opinions to align themselves with their party.

Results: Survey Supreme Court Opinions Affected by Party Split Information

The logic of using equivalent votes by elites and members of the mass public to understand representation is not limited to studies of the US Congress. Recently, work has also sought to understand the correspondence between judicial behavior, specifically voting by US Supreme Court Judges, and citizen preferences by asking citizens to cast their votes on cases previously considered by the Court (Jessee and Malhotra, 2013; Malhotra and Jessee, 2014). In Appendix Section D, we

⁷ At the aggregate bill x party level, confidence and importance are correlated at $\rho = .41$ ($p = .06$). See Appendix Figure A7 and discussion.

show that expressed opinions by survey respondents on Supreme Court cases are affected by a treatment providing information about the vote split between judges appointed by Republican and Democratic presidents.⁸ The opinion of each justice, along with their partisan background is readily available to other members on the court, but may not be available to survey respondents when they consider Supreme Court decisions. (Moreover, they may have no opinion about the relevant Constitutional issues at hand.) Supreme Court cases may be a conservative setting to find effects of information because many of these issues are social policy questions where citizens might hold strong opinions and many cases are well publicized. Nonetheless, we find that in 15 of 18 vote-party observations (83%), expressed support in the treatment condition is closer to the observed split for that party's justices. In sum, the measurement challenge we argue affects survey roll call measures seems to arise in the judicial realm as well.

Roll Call Survey Questions, Representation, and Polarization

In this section, we consider if the more-informed responses in the party split condition lead to different conclusions about the status of representation in the United States. We show that providing party split information lessens the estimated divergence between Congress and their constituents.

It is generally accepted that members of the House and Senate have polarized more than their constituents (e.g., McCarty, Poole, and Rosenthal, 2006; Hill and Tausanovitch, 2015). Almost all comparisons, however, rely on the survey responses of citizens without contextual or policy-specific information. If representatives do represent constituent interests and our argument about information holds merit, then the positions of citizens in the party split condition should more closely align with the positions of their representatives. The positions may not fully align, of course, because we have only delivered one of many pieces of information available to legislators.

We start by considering how often members and citizens vote with the party-plurality position on the 12 roll calls from Study 2.⁹ Figure 6 presents distributions of the number of votes cast with

⁸ Because judges do not use party labels in their day to day deliberations, we provide the partisanship of the president who appointed each justice.

⁹ In Appendix Section E, we scale the survey responses in Study 2 using item-response theory models separately by assigned condition (similar to the approach in Lauderdale, 2013). Although our statistical power is limited with

the Senate Democratic position, defined as yea if the rate of yea votes among Democratic senators is greater than or equal to the rate of yea votes among Republicans on the bill, nay otherwise. Each line connects the proportion of senators, control respondents, and party split respondents, by party, for each category of votes with the Democratic position.

[Figure 6 about here]

Providing citizens the party split makes rates of voting the party position look more like the rates of senators. More than 60 percent of Senate Republicans voted yea on 2 or 3 of these bills, with 20 percent voting yea on 0 or 1 (dotted density to the left). Most Republicans in the control condition support the Democratic position on 4 to 7 of these roll calls (solid density line), compared to the modal Republican in the party split condition (dash dot line) supporting the Democratic position on 2 to 3 bills. Democratic respondents in the party split condition share the same mode with Democratic Senators, supporting the Democratic position on 10 or 11 bills, while the modal Democrat in the control condition supports 8 or 9 bills. For both parties, the distribution of votes shifts away from the center in the party split condition.

Figure 6 and Appendix Section E confirm the longstanding pattern that polarization among citizens is notably lower than that observed among members of the Senate. However, providing the party split makes divergence between partisans larger and closer to that observed in the Senate. An open question is whether more information or a different context would further lessen the apparent disconnect between citizens and representatives.

Conclusion

We hypothesized that survey roll call measures may be limited in their value for measuring the quality of representation due to important problems of measurement. In particular, we argued that responses to those items may be poor proxies for citizen preferences because survey respondents lack sufficient information to express meaningful “votes.”

Our experiments show that each of three informational interventions affects measured citizen

only 12 items, we find that providing the party split cue moves the distribution of preferences among citizens towards the distribution of preferences of Senators.

preferences, suggesting that citizen survey responses in the absence of such information may not accurately measure what voters want their legislators to do. If that is the case, much of the work concluding failures of representation – including differential representation by income or other characteristics – rests on assumptions about valid measurement of citizen preferences. Moreover, such concerns do not appear limited to legislative behavior – we document similar instability in measures of citizen preferences on Supreme Court decisions.

Our design provides one path forward: if the representational benchmark is described as how the citizen would vote if she were acting as the legislator (or judge), we can develop new instruments that enrich the survey context to make it more similar to that of voting in the legislature. While it is impossible in the current Constitutional system to fully replicate the hypothetical experiment of citizens voting in a legislature, our theoretical approach can be used by scholars to justify comparisons of survey responses to representative behavior.¹⁰ We advocate that future work take more seriously issues of measurement.

As practical advice, if scholars use roll call and similar items to measure citizen preferences, they should embed in their studies tests of how sensitive their measures are to informational interventions such as those we examine here. Stability in measured preferences in the face of different information would help demonstrate the desirable construct validity, stability, and reliability emphasized in the survey design literature. Such work may also be a productive path towards understanding the systematic sources of measurement error in survey responses. One important question for future research is how the single pieces of information we delivered in these experiments compare to combinations of information.

This research also has important implications for those undertaking joint scaling. Joint scaling requires similar item parameters to map policy ideology from one arena to another. Our evidence here challenges this assumption for survey roll call measures and roll call votes in Congress. We argue that a key distinguishing feature of citizens is less information than legislators. It may be

¹⁰ Voting on ballot propositions may provide similar leverage, although it is notable that many citizens abstain, rely on party or other informational cues to substitute for lack of substantive knowledge, or simply vote no, implying that even those votes are made with self-awareness of lack of policy knowledge. Institutions such as Ireland's 2012 Convention on the Constitution, which randomly selected citizens to serve, may also provide some leverage.

that bridging across elite institutions (e.g., bureaucrats to courts or courts to executives) is more appropriate given greater information for those actors. On the other hand, even with similar levels of information, it is possible the votes in the different arenas have sufficiently different meanings or strategic contexts that joint scaling of elites suffers from systematic measurement error similar to that we identify here.

Moving forward, we argue that more empirical work with a cautious eye towards issues of measurement is needed to evaluate how well democratic representation functions. Accepting citizens' responses to survey questions as valid measures of how they prefer their members act relies on important assumptions that appear in question, making current conclusions about representation derived from those measures less certain. Future work might investigate alternative informational interventions, combinations of information, or explore qualitatively or experimentally what voters say or reveal by their behavior that they want from their representatives.

References

- Achen, Christopher H. 1977. "Measuring Representation: Perils of the Correlation Coefficient." *American Journal of Political Science* 21(4): 805–815.
- Achen, Christopher H. 1978. "Measuring Representation." *American Journal of Political Science* 22(3): 475–510.
- Arnold, R. Douglas. 1990. *The Logic of Congressional Action*. New Haven: Yale University Press.
- Bafumi, Joseph, and Michael C. Herron. 2010. "Leapfrog Representation and Extremism: A Study of American Voters and Their Members in Congress." *American Political Science Review* 104(3): 519–542.
- Bishop, George F., Alfred J. Tuchfarber, and Robert W. Oldendick. 1986. "Opinions on Fictitious Issues: The Pressure to Answer Survey Questions." *Public Opinion Quarterly* 50(2): 240–250.
- Broockman, David E. 2016. "Approaches to Studying Policy Representation." *Legislative Studies Quarterly* 41(1): 181–215.
- Bullock, John G. 2011. "Elite Influence on Public Opinion in an Informed Electorate." *American Political Science Review* 105(3): 496–515.
- Clinton, Joshua D. 2006. "Representation in Congress: Constituents and Roll Calls in the 106th House." *Journal of Politics* 68(2): 397–409.
- Cox, Gary W., and Mathew McCubbins. 1993. *Legislative Leviathan: Party Government in the House*. Berkeley and Los Angeles, CA: University of California Press.

- Gerber, Alan S., Gregory A. Huber, David Doherty, and Conor M. Dowling. 2011. "Citizens' Policy Confidence and Electoral Punishment: A Neglected Dimension of Electoral Accountability." *Journal of Politics* 73(4): 1206–1224.
- Gerber, Elisabeth R., and Jeffrey B. Lewis. 2004. "Beyond the Median: Voter Preferences, District Heterogeneity, and Political Representation." *Journal of Political Economy* 112(6): 1364–1383.
- Groves, Robert M., Floyd J. Fowler, Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology*. Hoboken, NJ: John Wiley & Sons.
- Hill, Seth J., and Chris Tausanovitch. 2015. "A Disconnect in Representation? Comparison of Trends in Congressional and Public Polarization." *Journal of Politics* 77(4): 1058–1075.
- Jessee, Stephen. 2016. "(How) Can We Estimate the Ideology of Citizens and Political Elites on the Same Scale?" *American Journal of Political Science* 60(4): 1108–1124.
- Jessee, Stephen, and Neil Malhotra. 2013. "Public (Mis)Perceptions Of Supreme Court Ideology: A Method For Directly Comparing Citizens And Justices." *Public Opinion Quarterly* 77(2): 619–634.
- Lauderdale, Benjamin E. 2013. "Does Inattention to Political Debate Explain the Polarization Gap Between the U.S. Congress and Public?" *Public Opinion Quarterly* 77: 2–23.
- Lax, Jeffrey R., and Justin H. Phillips. 2009. "How Should We Estimate Public Opinion in the States?" *American Journal of Political Science* 53(1): 107–121.
- Lewis, Jeffrey B., and Chris Tausanovitch. N.d. "When Does Joint Scaling Allow For Direct Comparisons of Preferences?" Working paper, University of California, Los Angeles, available at <http://idealpoint.tahk.us/papers/lewisTausanovitch.pdf>.
- Malhotra, Neil, and Stephen Jessee. 2014. "Ideological Proximity and Support for The Supreme Court." *Political Behavior* 36: 817–846.
- Mayhew, David R. 1974. *Congress: The Electoral Connection*. New Haven, CT: Yale University Press.
- McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. Cambridge, MA: The MIT Press.
- Miller, Warren E., and Donald E. Stokes. 1963. "Constituency Influence in Congress." *American Political Science Review* 57(1): pp. 45–56.
- Tausanovitch, Chris, and Christopher Warshaw. 2013. "Measuring Constituent Policy Preferences in Congress, State Legislatures, and Cities." *Journal of Politics* 75(2): 330–342.
- Tesler, Michael. 2015. "Priming Predispositions and Changing Policy Positions: An Account of When Mass Opinion Is Primed or Changed." *American Journal of Political Science* 59(4): 806–824.

Wright, Gerald C., Robert S. Erikson, and John P. McIver. 1985. "Measuring State Partisanship and Ideology with Survey Data." *Journal of Politics* 47(2): 469–489.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. New York: Cambridge University Press.

Zaller, John, and Stanley Feldman. 1992. "A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences." *American Journal of Political Science* 36(3): 579–616.

Table 1: Policy confidence attenuates treatment effect of information, Study 2

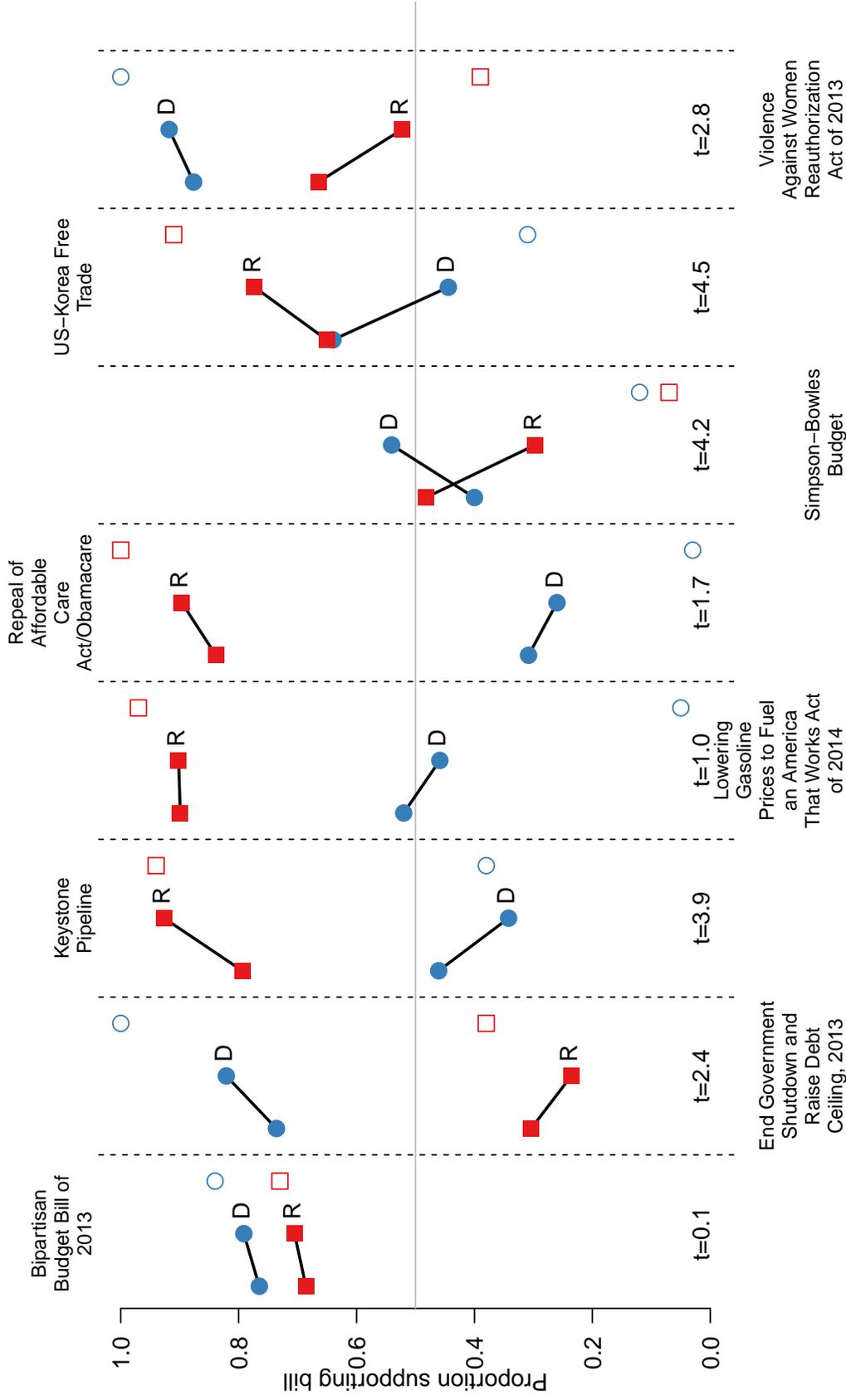
| VARIABLES | (1) All | (2) Dems | (3) Reps | (4) All | (5) Dems | (6) Reps |
|--|------------------|-----------------|------------------|------------------|-------------------|------------------|
| Average confidence in control condition by party | -0.11 (0.09) | -0.13 (0.12) | -0.030 (0.16) | -0.084 (0.10) | -0.12 (0.15) | -0.044 (0.16) |
| Average importance in control condition by party | | | | -0.038 (0.05) | -0.0061 (0.09) | -0.063 (0.06) |
| Republican respondent | -0.017 (0.03) | | | -0.021 (0.03) | | |
| Constant | 0.26 (0.14) | 0.29 (0.19) | 0.14 (0.24) | 0.31 (0.16) | 0.30 (0.22) | 0.29 (0.27) |
| Observations | 22 | 11 | 11 | 22 | 11 | 11 |
| R-squared | 0.072 | 0.106 | 0.004 | 0.101 | 0.107 | 0.135 |

Standard errors in parentheses

** p<0.01, * p<0.05

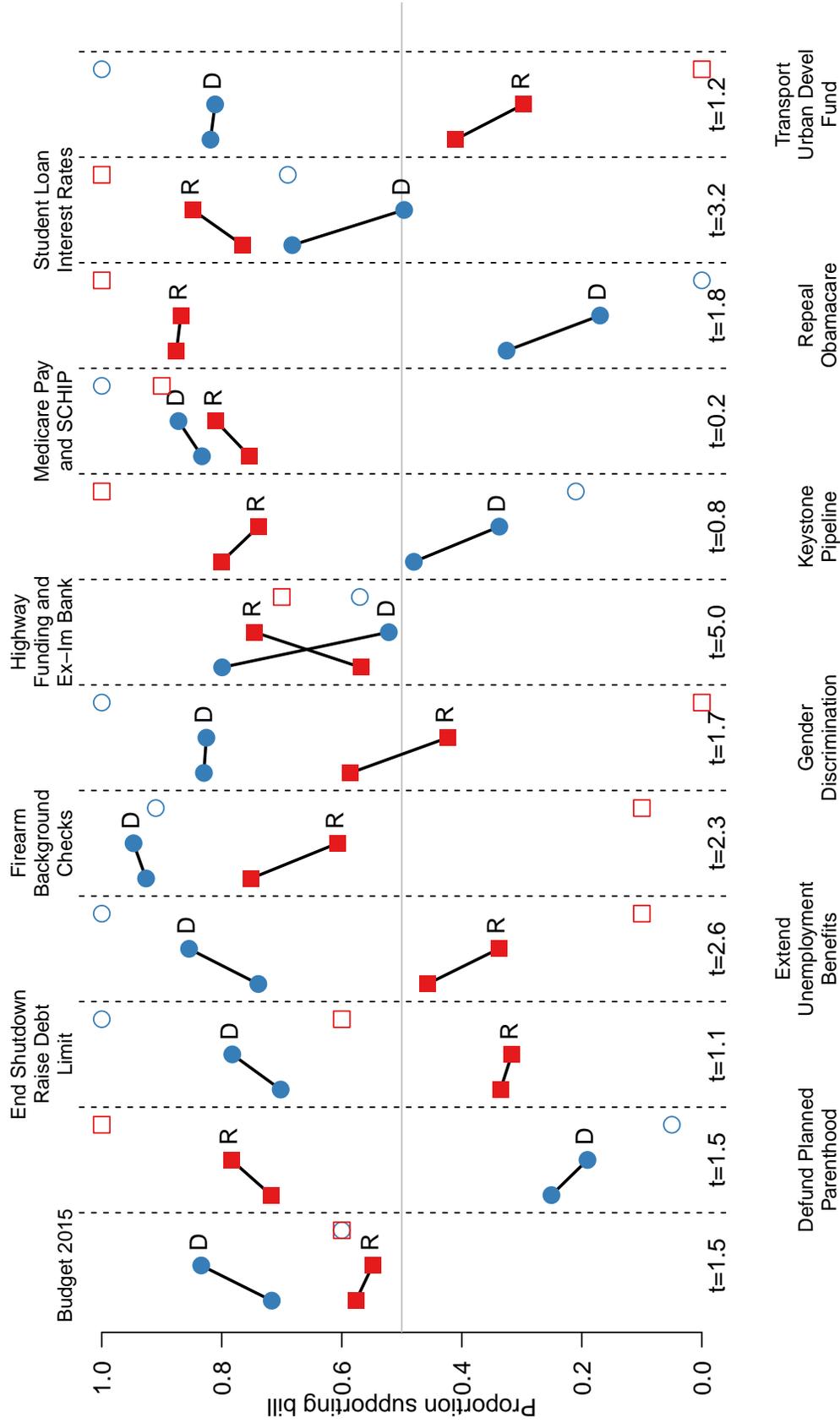
Note: OLS coefficients. Dependent variable is absolute value of treatment effect of providing party split on roll call support by party and bill.

Figure 1: Support for roll call with and without party split information, Study 1



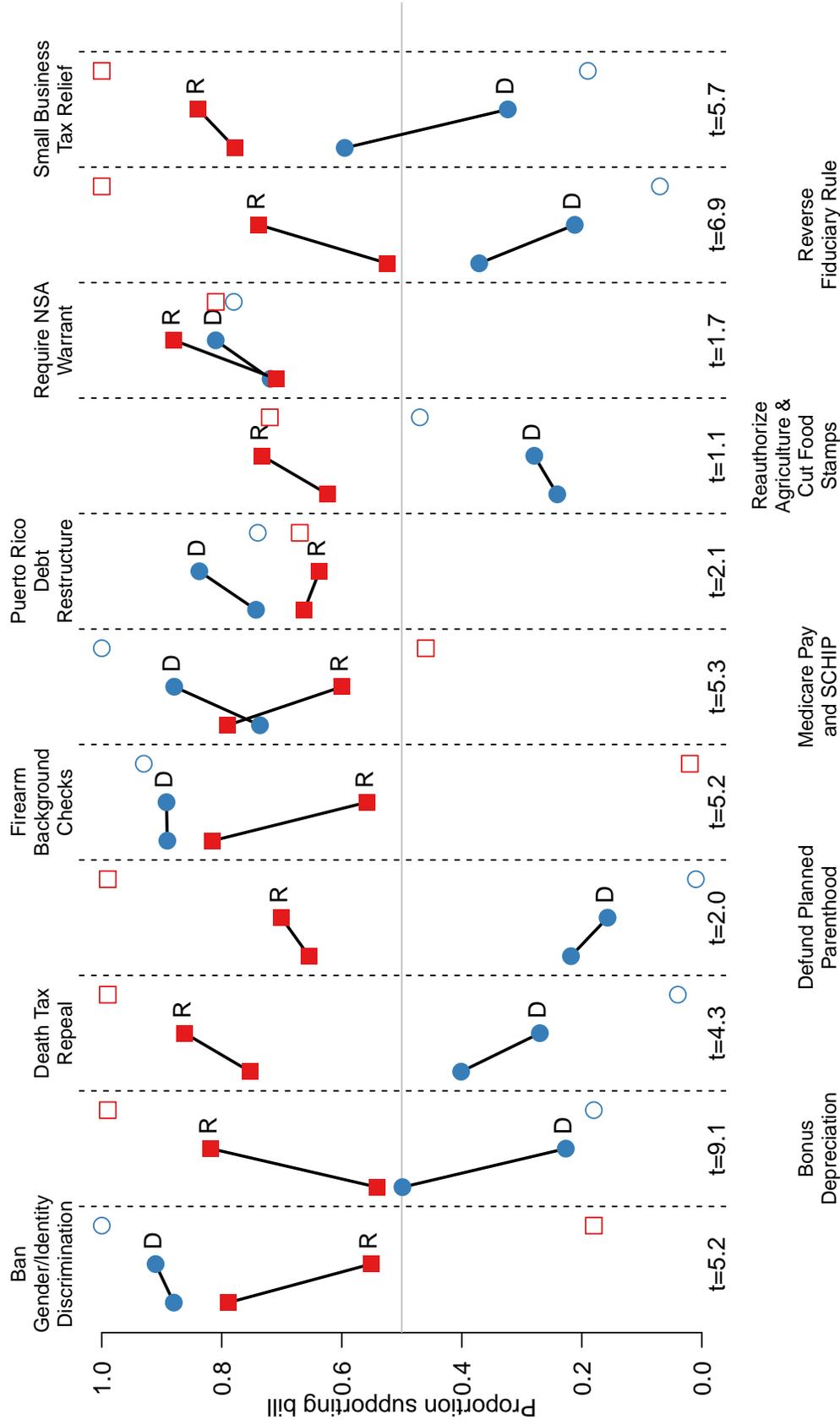
Note: Closed circles (squares) connect support among Democratic (Republican) respondents for bill from those in control condition (left) to those in the party split condition (right). Open circles (squares) are the actual rate of support among Democratic (Republican) members of the House. Absolute value of t-ratio from difference-in-difference estimate of party-times-treatment indicated at x-axis.

Figure 2: Support for roll call with and without party split information, Study 2



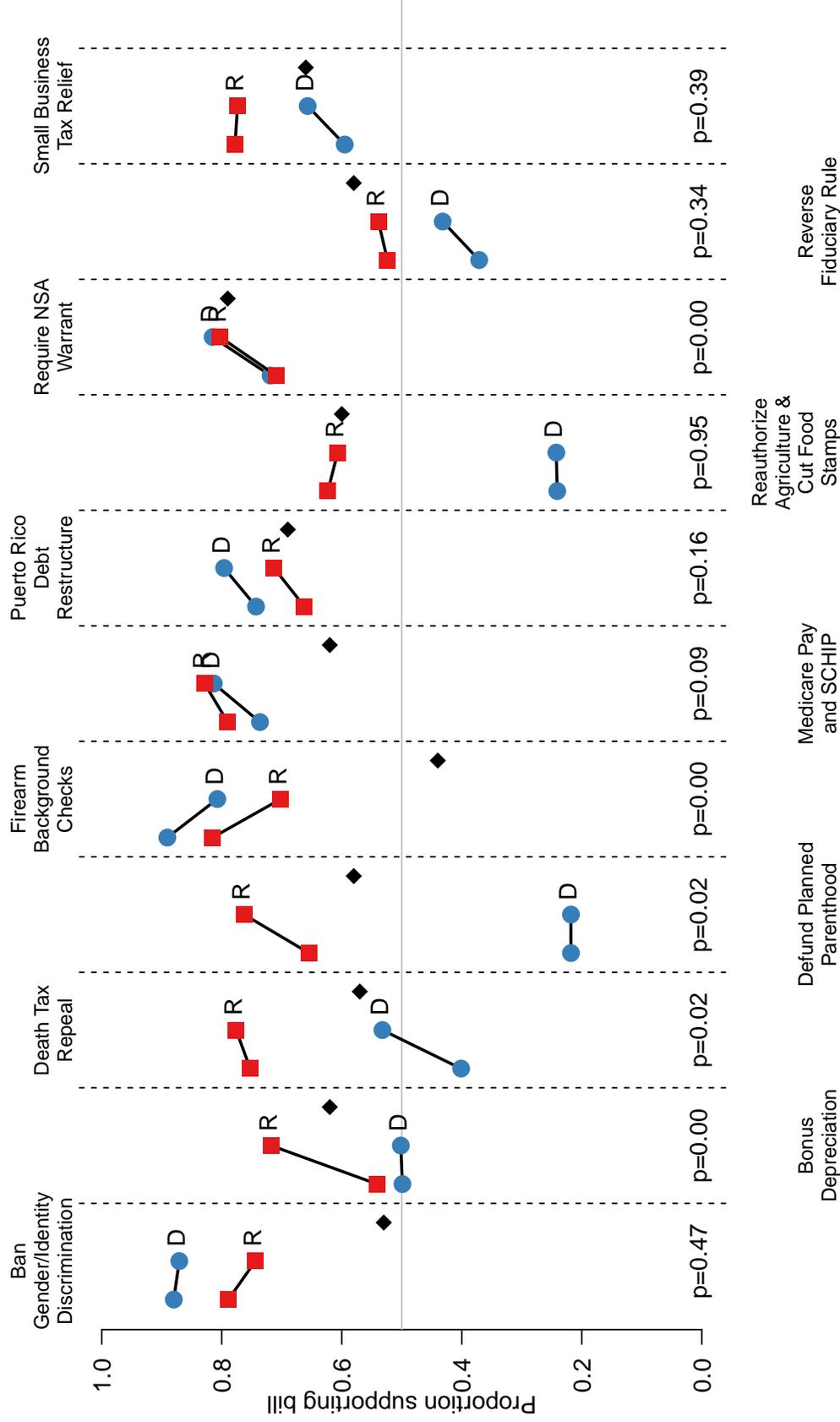
Note: Closed circles (squares) connect support among Democratic (Republican) respondents for bill from those in control condition (left) to those in the party split condition (right). Open circles (squares) are the actual rate of support among Democratic (Republican) members of the Senate. Absolute value of t-ratio from difference-in-difference estimate of party-times-treatment indicated at x-axis.

Figure 3: Support for roll call with and without party split information, Study 3



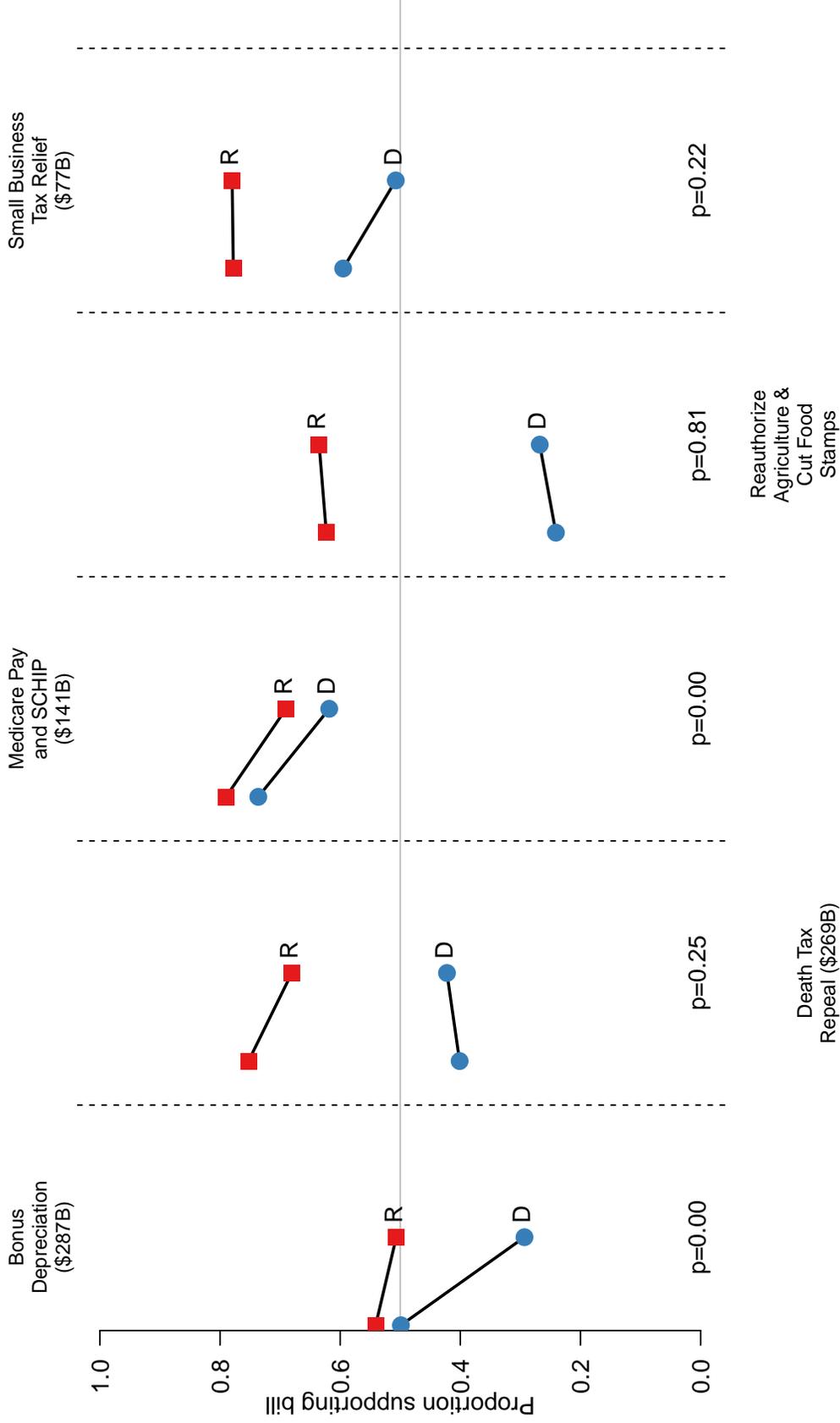
Note: Closed circles (squares) connect support among Democratic (Republican) respondents for bill from those in control condition (left) to those in the party split condition (right). Open circles (squares) are the actual rate of support among Democratic (Republican) members in the chamber. Absolute value of t-ratio from difference-in-difference estimate of party-times-treatment indicated at x-axis.

Figure 4: Support for roll call from control to Chamber Split condition, Study 3



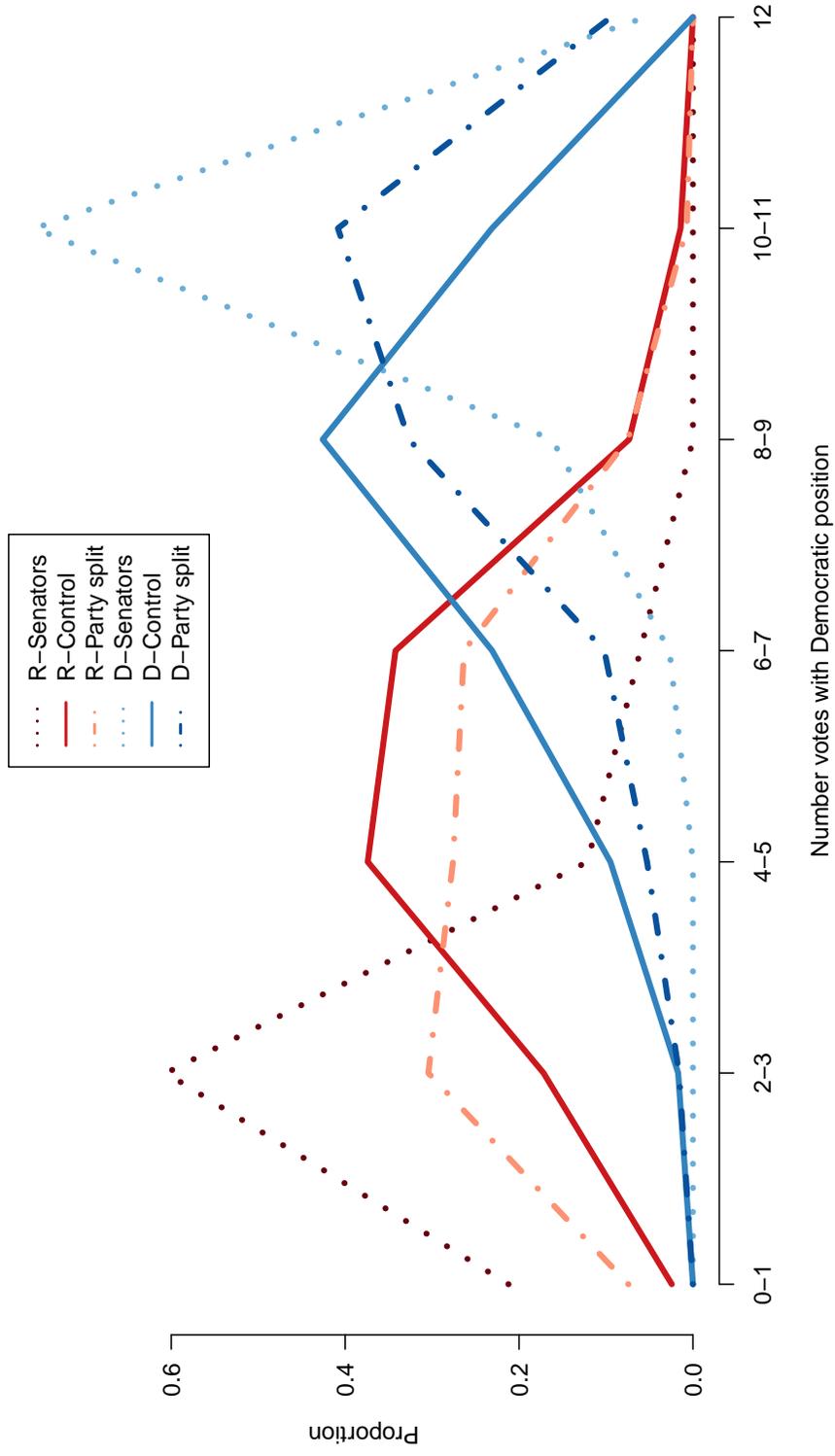
Note: Closed circles (squares) connect support among Democratic (Republican) respondents for bill from those in control condition (left) to those in the chamber split condition (right). Diamonds are the actual rate of support in the chamber. P-value for an F-test on the treatment and treatment times party regression coefficients to evaluate significance of treatment indicated at x-axis

Figure 5: Support for roll call from control to CBO condition, Study 3



Note: Closed circles (squares) connect support among Democratic (Republican) respondents for bill from those in control condition (left) to those in the CBO condition (right). CBO condition presents respondents with a synopsis of the Congressional Budget Office analysis of the legislation. Number in parenthesis is the deficit impact calculated by the CBO. There was no deficit impact presented in the analysis to the Agricultural authorization. Absolute value of t-ratio from difference-in-difference estimate of party-times-treatment indicated at x-axis.

Figure 6: Votes with Democratic position in Senate by condition and party, Study 2



Note: Each line is the distribution across number of votes with the Democrats on the 12 roll call votes in Study 2.