On The Meaning of Survey Reports of Roll Call Votes Not Cast in a Legislature

Seth J. Hill
University of California, San Diego*

Gregory A. Huber
Yale University†

February 22, 2017

Abstract: Contemporary efforts to evaluate the quality of representation often use survey measures of how citizens say they would vote on legislation to compare citizen preferences to what elected officials do in office. These comparisons often suggest poor representation. We argue here that this common design is unlikely to effectively evaluate representation because responses to survey questions, even on roll call votes, differ in important aspects from voting in legislatures. This leads to systematic measurement error that undermines the key assumption that survey responses measure preferences in the same policy space as legislative behavior. Results from two survey experiments show that providing information readily available to legislators though not readily available to survey respondents materially changes respondents’ expressed preferences on roll call votes. Expressed policy positions in this “informed” survey condition indicate survey respondents are both more extreme than commonly estimated and more closely matched to legislator behavior in their preferred party. Information increases party splits among respondents by 40 and 60 percent. We also show that respondents appear aware of their own lack of knowledge in evaluating roll call policy votes and that the size of the treatment effect of information decreases in the confidence judging policy in that area.

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

*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.
Understanding representation is a central and longstanding research topic in political science. Among the most common ways to assess the quality of representation empirically are studies that calculate the concordance between an elected official’s behavior and constituent preferences. Roughly, do elected officials do what citizens want, or, stated slightly differently, would the citizen do something different if granted the legislator’s vote? While there are myriad ways in which to attempt to answer this question, one common analytic approach is to compare the roll call votes of elected members of Congress with the preferences of constituents as measured using survey data. The key challenge in this exercise (e.g., Achen, 1977) is creating comparable measures between the two contexts. For example, if a survey respondent believes that abortion rights should be restricted in some situations, how should that preference map onto voting for or against a particular bill concerning abortion policy?

In response to this sort of criticism, over the last decade scholars have turned to survey measures that are designed to explicitly place citizens’ choices in the same “policy space” as their legislative counterparts. The most promising in this regard are “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 ideology that are directly comparable to legislative roll call votes because the survey respondent is making the same binary choice as the legislator. With those measures in hand, 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.

While there may be 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 (e.g., the importance of partisan agenda control and building legislative coalitions, bargaining before an

---

1 This use of survey data was itself a response to criticism about using aggregate district-level characteristics as proxies for presumed voter preferences.

2 There is a robust subsequent body of research building on, and calling into question, this finding. We discuss the relevant work below.
electoral audience, long- versus short-term strategic tradeoffs, etc.) and these limitations are acknowledged in some prior work, this paper identifies a more basic source of systematic measurement error in survey assessments of voters’ preferences about roll call votes. This measurement error calls into question the assumption that responses to roll call survey items yield estimates of citizen preferences that are appropriate to compare to legislator behavior in evaluating the quality of representation. The particular problem is that voters know very little about the items they are asked to “vote on” in a survey context. As a consequence, their responses may not reflect how they would vote if they were fully informed about the partisan and strategic context of the legislature. More bluntly, a survey item asking citizens how they would vote on a roll call item may be a substantively poor measure of how the citizen would like her legislator to behave or would herself vote if sitting in Congress. In a low-information survey context, many respondents are likely expressing non-attitudes with little relevance for the quality of representation, undercutting the assumption that because these roll-call items ask survey respondents to make the same binary choices as a legislator that they therefore measure preferences in the same policy space.

To demonstrate the importance of this argument for measuring citizen preferences, we present results from two survey experiments in which we randomly varied the amount of information available to survey respondents when providing their responses to survey roll call items. The specific information that we provide—party margins on the vote as taken in Congress—is a proxy for a range of party cues and contextual information likely available to a member of Congress when voting but not readily available to most survey respondents. The party margin is one of many potential pieces of information that could be delivered, but it is also one of the most simple and potentially most powerful. The party margin provides (rough) information to the subject about the policy and political implications surrounding the roll call vote. If the treatment systematically changes the rate of support for the roll call vote among survey respondents, it suggests the survey response absent that information may not accurately describe how the subject would vote were they acting in the legislator’s position with access the legislator’s information.

To briefly summarize our findings, we show that providing information about party votes in
Congress affects patterns of citizen preferences measured using roll call survey items and increases party divisions in these responses by around 50 percent. Providing this information increases the predictive power of partisanship in explaining survey responses and increases the concordance between partisan patterns of survey responses and observed party splits in Congress. Substantively, we find that elected officials no longer appear substantially more extreme than their partisan counterparts in the mass public when the public’s preferences are measured in the informed condition. Overall, these data provide experimental evidence that using standard roll call survey items to create comparable estimates of citizen preferences and legislator behavior does not address the fundamental problem of placing the two sets of choices in a common space within which one can assess the quality of representation.

We discuss one important objection to this interpretation. This objection is that showing survey responses change when respondents are told party vote margins merely indicates that citizens display blind partisan obedience rather than substantive policy preferences. There is a longstanding literature suggesting many policy opinions as measured in surveys are not deeply held (e.g., Zaller, 1992) and are shaped by details of survey design. In light of this instability in survey responses, we believe it is unlikely that policy opinions measured when less information is provided are more informative of true citizen preferences than when more information is available. Nonetheless, one might argue that our design of providing vote margins is a party cue that causes people to change their survey response despite having held strong and coherent policy preferences on the issue absent the cue. We note, however, recent research has shown that survey responses are much less affected by partisan cues that previously presumed (Bullock, 2011) and that survey respondents seem to vary the rewards for elected officials by self-assessments of their own confidence on the policy in question (Gerber et al., 2011). 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. That is, policy views change more in response to the party cue when respondents report less confidence in evaluating the issue. This suggests that, rather than blind obedience, our treatment effects are consistent with citizens using additional information to help evaluate policy.
when they know they do not have the full information to evaluate that policy.

This research provides novel insights for three areas of ongoing work in political science. The first is work on how best to measure citizen preferences and place those preferences in a ideological space that is comparable to legislative behavior. The second is about assessing the quality of representation more generally. An important normative question is how stable and meaningful individual citizen preferences over specific policies are. If, as we find, policy preferences are more difficult to measure than commonly assumed, it calls for more research to understand how to evaluate the quality of representation using survey measures of preferences. Finally, our results speak to the meaning of survey responses to questions about political policy preferences.

Our results and the core of our design point toward future research that can unpack the meaning and consequences of informational cues about political 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 therefore help explore when and under what conditions our tools of measurement provided estimates of citizen preferences that are more aligned with legislators’ policy space. Those results, as well as the results presented in this paper, also speak to work 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 away individual measurement error, our findings suggest that measurement error may be systematic in addition to random. We explore these issues briefly in an Appendix, but our design provides a framework for future work to evaluate the influence of additional policy information on scaling procedures.

**Assessing Representation Using Survey Responses to Roll Call Items**

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 offi-
cials 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).” This definition sets aside issues of coordination, multidimensionality, and cycling, as well as alternative definitions of representation, as does most of the empirical literature on the question.

While there were some early efforts to evaluate representation using district-level characteristics (aggregate data) to impute individual-level preferences (e.g., Turner, 1951; MacRae, 1958), 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 them to elite behaviors. Seminal in this line of work was Miller and Stokes (1963), which used American National Election Survey data to compare district-level estimates of public opinion with corresponding data about the behaviors of sitting members of Congress.

A number of criticisms were directed against this empirical approach (for example, that it was vulnerable to measurement error because it relied on a small number of survey respondents to estimate district preferences), but the most fundamental critique of this method was raised in two seminal papers (Achen, 1977, 1978) that argued that the correlation between these estimates of citizen preferences and legislator behavior was not informative of the quality of representation. Most generally, Achen argues that it is impossible to meaningfully compare the two measures unless they are constructed in the same policy space. It is not clear what a “high” versus a “low” response on one survey item – for example preferring expansive abortion rights over more restrictive policy

---

3 By contrast, one criticism of the earlier work was that it assumes a preference that isn’t directly measured and it was therefore difficult to characterize the magnitude of any departure from perfect representation. This limitation, along with the technological innovation of the opinion survey, has mostly left this mode of inquiry behind (though see Fowler and Hall, 2016, for a recent application).

4 See Clinton (2006) or Wright, Erikson, and McIver (1985) on the challenges of sampling error in making these comparisons.
– would mean in the binary legislative policy of supporting or opposing a particular bill brought before Congress.\footnote{For a similar criticism of research evaluating the effect of citizen initiative procedures on the relationship between citizen preferences and policy, see Matsusaka (2001).}

In light of this criticism about standard survey measures and roll call votes not being comparable, scholars have subsequently embraced the idea of asking ordinary citizens how they would vote on roll calls, i.e. “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 (e.g., Ansolabehere and Jones, 2010). Additionally, survey responses and votes on a common set of roll call votes can serve as bridging items in joint scaling procedures that use a broader range of roll call votes (for legislators) and survey items (for citizens) to provide more precise measures of citizen and legislator preferences (e.g., Ansolabehere, Snyder, Jr., and Stewart III, 2001; Bafumi and Herron, 2010; Clinton, 2006; Gerber and Lewis, 2004; Tausanovitch and Warshaw, 2013). Cumulatively, all of these methods aim to measure citizen preferences in the same space as member behavior so that the correspondence between the two can be used to assess the quality of representation. In general, findings from these different approaches suggest measurable but imperfect congruence between what voters want and what representatives do.

Approaches that use “roll call” survey items 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 and the two can be compared. If this assumption does not hold, however, then the relevance of roll call items for assessing the quality of representation is in question. Without the assumption, roll call survey items do not address the core criticisms of Achen that measures could not be accurately compared between the legislative and survey domains. One indication of the challenges for this assumption is revealed in work by Jessee (2016) and Lewis and Tausanovitch
who examine the performance of the joint scaling procedures that use multiple survey items and roll call votes to place legislators and survey respondents in the same policy space. Jessee (2016) finds that the relationship between ideology and votes is different for legislators (roll call votes) and citizens (roll call survey items), but argues that joint scaling can be useful for creating estimates in a common space if methods are applied to evaluate substantive and theoretical concerns. Lewis and Tausanovitch (N.d.) show that the dimensionality underlying citizen preferences is statistically distinct from that underlying elite preferences, leading them to conclude that joint scaling using survey roll calls fails to produce a common issue space.

In part, the disagreement between Jessee (2016) and Lewis and Tausanovitch (N.d.) reflects the ambiguity of any approach that relies on the scaling process itself to assess the validity of using roll call survey responses to place voters in the legislator’s policy space. Roughly, the differences between legislators and citizens in the mapping of ideology to outcomes in these scaling procedures might reflect either imperfect representation or that the roll call survey items are not serving the purpose they were meant to.

In light of this ambiguity, we take a different approach, which is to test whether survey responses to roll call survey items as normally fielded produce stable estimates of citizen preferences. There are a number of reasons that the context a member of Congress faces when voting on a particular bill are different from the survey context faced by an ordinary respondent. If these contextual features have a meaningful effect on measured citizen preferences then it implies that roll call survey items as normally used do not by themselves yield comparable ideology estimates across the two domains. 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 “yes” or “no” vote, anticipating how the
public will likely 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 will 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. While this list is not exhaustive, the general point is that incentives and choices facing a member of Congress when voting on a bill are likely very different from those a survey respondent faces. 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 ideology.

Supporting the notion that survey response to standard roll-call items may not accurately reflect citizen policy preferences is the large literature showing that survey responses are often affected by even minor alterations in design and question wording. At one extreme, for example, survey respondents are willing to express opinions on non-existent bills (Bishop, Tuchfarber, and Oldendick, 1986). In general, survey responses change when subjects are provided with cues or information that can substitute for their own lack of policy knowledge. While this is sometimes taken as evidence that citizens lack true preferences and are instead beholden to elite persuasion (e.g., Zaller, 1992), citizens also appear sophisticated and self-aware of their own lack of detailed knowledge. For example, survey respondents distinguish among issue positions they take, being less likely to reward or punish elected officials in policy domains where they understand 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 very modest effects on expressed citizen preferences, implying that concerns about partisanship dominating sincere policy views are largely misplaced (Bullock, 2011).

Overall, what is unclear from prior research using survey roll call measures is whether those
items accomplish their goal of providing an informative measure of citizen preferences in the same policy space as legislators. While there are many reasons survey respondents may not otherwise fully appreciate and simulate the legislative context when offering their survey responses, existing evidence suggests survey responses are sometimes but not always affected by additional information. In the next section, we therefore present the design for an experiment that tests the sensitivity of survey responses to roll call items to informational cues that allow voters to more closely mirror the legislative context. Doing so allows us to understand whether standard roll call survey items provide a response that is informative to the core question of how the citizen would vote if she were to take the legislator’s place.

**Research design and data collection**

In order to assess the importance of the questions and arguments raised in the previous sections about the challenge of measuring citizen preferences, we fielded two survey experiments in which we randomly varied the amount of information provided to respondents when answering questions about support for legislation considered by Congress. These questions allow us to estimate citizen preferences using survey responses obtained under these different informational conditions and compare them to legislator roll call voting in Congress. If the amount of information presented to survey respondents affects expressed preferences over roll calls (and comparisons of citizen to legislator polarization), then it suggests that standard (low-information) survey responses may yield misleading pictures of the relationship between citizen preferences and member behavior.

The key intervention in both experiments is the same. In our control groups, survey respondents answered survey roll call items that mirror those commonly used in studies that jointly scale legislators and citizens. In our treatment groups, by contrast, we provide one additional salient piece of information: 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 (and their strategic
considerations around compromise legislation, amendments, and status quo policy) and the larger political environment in which strategic elites perform before the electoral audience. For this reason, this information 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. Of course, it is by no means an exhaustive treatment because it does not convey the full complexity of either the policy or strategic political environment, but it is nonetheless likely a useful proxy for many facets of these considerations.

The first study asked respondents for their support on a few roll call votes surrounding the 2014 midterm election, and the second study expanded the project by asking respondents in 2016 for their opinions across 12 recent roll call votes. We detail each study below.

**Study 1: 2014 Cooperative Congressional Election Study**

Our survey experiment for Study 1 was embedded in 3 team modules fielded on the 2014 Cooperative Congressional Election Study (CCES). Respondents were asked whether they supported or opposed a randomly selected subset of 8 different roll call votes, listed in Table 1. Congress had voted on each in a recent session. We selected these items to follow the set of roll calls survey items from the CCES common content and to vary in both their subject matter and the degree to which voting in the House divided the parties.

Respondents were first asked whether or not they supported two of these items, selected at random, using the standard CCES common-content question wording, “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.” 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. Respondents were

---

6 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.
<table>
<thead>
<tr>
<th>Bill Title</th>
<th>Long Description</th>
<th>Roll Call #</th>
<th>Democrat margin (Y-N%)</th>
<th>Republican margin (Y-N%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeal of Affordable Care Act / Obamacare</td>
<td>Repeals the Patient Protection and Affordable Care Act and health care-related provisions in the Health Care and Education Reconciliation Act of 2010.</td>
<td>460 (July 11, 2012)</td>
<td>3-97</td>
<td>100-0</td>
</tr>
<tr>
<td>US-Korea Free Trade</td>
<td>Implements the United States-Korea Free Trade Agreement.</td>
<td>783 (October 12, 2011)</td>
<td>31-69</td>
<td>91-9</td>
</tr>
<tr>
<td>Simpson-Bowles Budget</td>
<td>Adopt budget proposal endorsed by the Simpson-Bowles Commission.</td>
<td>145 (March 28, 2012)</td>
<td>12-88</td>
<td>7-93</td>
</tr>
<tr>
<td>Keystone Pipeline</td>
<td>Extends Federal aid for highways and requires the Federal Energy Regulatory Commission to approve the Keystone Pipeline within 30 days.</td>
<td>170 (April 18, 2012)</td>
<td>38-62</td>
<td>94-6</td>
</tr>
<tr>
<td>Lowering Gasoline Prices to Fuel an America That Works Act of 2014</td>
<td>Implements a new program to lease space off of America’s coasts to drill for new oil and gas resources.</td>
<td>368 (June 26, 2014)</td>
<td>5-95</td>
<td>97-3</td>
</tr>
<tr>
<td>Bipartisan Budget Bill of 2013</td>
<td>Implements a budget compromise to reduce some mandatory spending cuts and funds the federal government for fiscal years 2013 and 2014.</td>
<td>640 (December 12, 2013)</td>
<td>84-16</td>
<td>73-27</td>
</tr>
<tr>
<td>Violence Against Women Reauthorization Act of 2013</td>
<td>Provides $1.6 billion toward investigation and prosecution of violent crimes against women and establishes the Office on Violence Against Women within the Department of Justice.</td>
<td>47 (February 28, 2013)</td>
<td>100-0</td>
<td>39-61</td>
</tr>
<tr>
<td>End Government Shutdown and Raise Debt Ceiling, 2013</td>
<td>Ends the government shutdown that began October 2013 and increases the federal debt limit.</td>
<td>550 (October 16, 2013)</td>
<td>100-0</td>
<td>38-62</td>
</tr>
</tbody>
</table>
“soft-forced” to choose one of these options. For simplicity, we refer to these roll call questions as “standard” survey roll call items.

One-third of respondents, selected at random, were then asked how they would have voted on four additional items, selected at random from those items not chosen for the “standard” items. The question prompt and response options were the same. However, for each item we added a brief summary of the observed pattern of partisan voting in the House. Thus, the US-Korea Free Trade item shown above would have instead appeared as follows:

**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)

For simplicity, we refer to these roll call questions as “informed” survey roll call items.

Our total sample includes 3,456 respondents. 2,300 individuals were assigned to receive only the two standard survey items, and the additional 1,156 individuals received both the two standard items and four of the informed items.7

**Study 2: 2016 Survey Sampling International**

Study 2 replicate and extends the Study 1 design. All roll call votes selected for Study 2 were cast in the 113th or 114th Senates and are detailed in Table 2. We selected these roll call votes from all final passage votes to vary on topic as well as party splits in the Senate.8 Additionally, we included a longer (post-treatment) battery on perceptions of policy importance and expertise to evaluate policy by issue area, whose use we detail below.

Study 2 was a simple between-subject design. Subjects were assigned at random either to the informed or standard condition, and all 12 of their roll call questions were of that type.9 We

---

7 All CCES analysis uses the provided post-stratification weights and is restricted to those respondents who answered all of their assigned roll call items. Of those assigned to two items, 96.1% answered both and 2.4% answered one. Of those assigned to six items, 93% answered all 6, 4.2% answered 5 items, and the remaining 2.9% answered 4 or fewer items. Patterns of non-response do not differ consistently by policy area across the two conditions. Average rates of non-response are 2.6% for the standard survey items and 2.2% for the informed survey items.

8 To assist with summarizing the bill for participants, we sent our bill summaries to six colleagues who are experts in congressional politics. We are grateful to these colleagues for helping us clarify the language of the items.

9 One advantage of this design over Study 1 is that treatment condition is therefore uncorrelated with response order.
Table 2: Twelve roll call votes used in Study 2

<table>
<thead>
<tr>
<th>Bill Title</th>
<th>Long Description</th>
<th>Roll Call</th>
<th>Democrat margin (Y-N%)</th>
<th>Republican margin (Y-N%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expand Existing Background Checks for Firearm Sales</td>
<td>Require federal background checks for gun sales that take place at gun shows or via the internet, the same requirement that exists for sales from regular brick and mortar gun stores.</td>
<td>97 (April 17, 2013)</td>
<td>91-9</td>
<td>9-91</td>
</tr>
<tr>
<td>Set Federal Student Loan Interest Rates</td>
<td>Set federal student loan interest rates, raising rates relative to recent rates but decreasing them compared to the rates that were in force because an old law had expired.</td>
<td>185 (July 24, 2013)</td>
<td>69-31</td>
<td>98-2</td>
</tr>
<tr>
<td>Allow a Vote on Funding Transportation and Urban Development</td>
<td>Support a motion to end debate and allow a final vote on a bill that would fund at a level of $54 billion for one year Transportation and Housing and Urban Development.</td>
<td>199 (August 01, 2013)</td>
<td>100-0</td>
<td>2-98</td>
</tr>
<tr>
<td>End Government Shutdown and Raise Government Debt Limit</td>
<td>End the government shutdown of October 2013 by funding the government for three months and also allowing it to borrow money.</td>
<td>219 (October 16, 2013)</td>
<td>100-0</td>
<td>60-40</td>
</tr>
<tr>
<td>Extend Federal Unemployment Benefits</td>
<td>Extend existing federal unemployment benefits for a minimum of an additional 5 months.</td>
<td>392 (April 07, 2014)</td>
<td>100-0</td>
<td>14-86</td>
</tr>
<tr>
<td>Allow a Vote on Changing the Standard for Determining Gender Discrimination in the Workplace</td>
<td>Support a motion to end debate and allow a final vote on a bill that would require employers to show that any wage gaps between men and women with similar jobs and qualifications have a business justification.</td>
<td>553 (September 15, 2014)</td>
<td>100-0</td>
<td>0-100</td>
</tr>
<tr>
<td>Approve 2015 Budget and Fund Government for 2015</td>
<td>Agree to a measure that would fund almost all federal government agencies for fiscal year 2015.</td>
<td>645 (December 13, 2014)</td>
<td>60-40</td>
<td>57-43</td>
</tr>
<tr>
<td>Approve Keystone XL Pipeline</td>
<td>Allow TransCanada to construct the 1,179-mile Keystone XL pipeline that would carry oil from Canada’s tar sands to refineries in Texas.</td>
<td>49 (January 29, 2015)</td>
<td>21-79</td>
<td>100-0</td>
</tr>
<tr>
<td>Revise Medicare Physician Payment Rates and Reauthorize Child Health Insurance Program</td>
<td>Change the rules used to calculate physician payments so that doctors who see Medicare patients did not experience large drops in the amount the government paid them for providing care and fund for two years the program that provides free or low-cost insurance for low-income children and families.</td>
<td>144 (April 14, 2015)</td>
<td>100-0</td>
<td>85-15</td>
</tr>
<tr>
<td>Pass the FAST Act and Extend the Export-Import Bank</td>
<td>Authorize 6 years of federal spending on highways and other transit programs and extend programs to use federal funds to finance and insure foreign purchases of American goods.</td>
<td>260 (July 30, 2015)</td>
<td>57-43</td>
<td>72-28</td>
</tr>
<tr>
<td>Allow a Vote on Banning Federal Funding for Planned Parenthood</td>
<td>Support a motion to end debate and allow a final vote on a bill that would prevent any federal money from going to Planned Parenthood.</td>
<td>262 (August 03, 2015)</td>
<td>5-95</td>
<td>96-4</td>
</tr>
<tr>
<td>Repeal ObamaCare</td>
<td>Repeal the Affordable Care Act health care program by removing the federal health insurance requirement, eliminating associated taxes, and eliminating federal subsidies for low-income individuals to purchase insurance. Also bans federal funding of Planned Parenthood for one year.</td>
<td>329 (December 03, 2015)</td>
<td>0-100</td>
<td>96-4</td>
</tr>
</tbody>
</table>
fielded the survey through Survey Sampling International (SSI), a firm that maintains an online panel whose demographics approximate a nationally representative sample. Our sample includes 1,464 respondents who participated in May and June of 2016. We exclude from our analysis 347 respondents who failed an attention screener in the middle of the survey because these subjects appear less engaged and are therefore unlikely to provide meaningful responses, yielding a final sample of 1,117 SSI participants.

**Results: Responses to Survey Roll Call Votes Items Affected by Information about Member Votes**

In this section, we present the main results of the two survey experiments. We show that providing information about the outcome of a roll call vote in Congress changes the pattern of support among survey respondents. In particular, we show that the partisan breakdown among respondents moves towards the observed party split in the legislature in nearly all cases. This suggests that standard roll call survey measures do not accurately reflect how citizens want their representatives in the legislature to vote.

**Study 1**

In Figure 1, we present the effect of our informational intervention on support for each roll call vote from Study 1. Specifically, for each bill (horizontal axis) we plot the proportion of survey respondents and members of the House supporting the bill, by party. (Partisans are defined to include “leaners.” For analysis excluding leaners, as well as for analysis of “pure” independents, see the appendix.) Reading left to right, we begin by plotting the proportion of Democrats (blue

---

10 Although the sample is approximately representative of the American population, we found some demographic and political variables did not match population targets very well, and so constructed post-stratification weights that make the SSI sample approximate the Pew Research Center 2015 Governance Survey, a random-digit dial telephone sample of about 6,000 fielded in September 2015. All analysis uses these weights. We detail the weighting procedure in Appendix Section A.

11 The 12 roll calls were asked in random order across three screens each with four items. After the first screen of four items but before the second, we screened for attention by asking the respondents which of four roll calls they had just given their opinion about. Only one of the four had actually been asked on the first screen, and we use only those respondents who identified this roll call correctly. This sort of screen likely includes some respondent who simply guessed which of the four was the right answer. Results for all respondents appear in the Appendix.
circles) and Republicans (red squares) supporting the bill in the “standard” roll call condition. Next, connected to these points by solid lines are the proportions of Democrats and Republicans supporting the bill in the “informed” condition. (These are between, rather than within subject comparisons, because different individuals were assigned to each condition.) Finally, the last pair of points, indicated by hollow markers, is the proportion of Democrats and Republicans voting yea on the bill on the House floor. This is the information provided to respondents in the “informed” survey item.

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 standard condition to a rate of 79 percent in the informed 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 standard condition to 71 percent in the informed condition.

Comparing across the 8 bills shown in Figure 1, several 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 they were informed of the House vote for 7 of 8 items. 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.12 Democrats become substantively more supportive of this bill (and Republicans less) after learning it had low support from both parties in the House. 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 which was supported by only about 40% of Republican House members (close to the minimum necessary for passage). After learning that the bill was supported at a much higher rate by Democrats than Republicans they become less supportive.13

12 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.
13 At the time of the House vote, Republicans expressed fears that voting for the compromise measure would invite primary challengers from more conservative Republicans. See, e.g., http://articles.latimes.com/2011/jul/21/nation/la-na-debt-talks-20110721 (Retrieved July 12, 2016). Nonetheless, Republican elites viewed its passage as necessary to avoid being blamed for (again) shutting down the government.
Figure 1: Support for roll call with and without information, Study 1

Note: Closed circles (squares) connect support among Democratic (Republican) respondents for bill from those in standard condition (left) to those in the informed condition (right). Open circles (squares) are the actual rate of support among Democratic (Republican) members of the House. These rates were presented to respondents in the informed condition. In all but two party-roll call votes (Simpson-Bowles for Democratic respondents, End Shutdown for Republican respondents), support moves towards the legislative margin with information.
Second, responses in the informed condition are more polarized than in the uninformed condition. Apart from the Bipartisan Budget Bill of 2013 (where actual partisan divergence in the legislature was low and treatment effects are modest), expressed partisan roll call preferences diverge more when informed of how the House voted than in the standard survey item. How large are these effects? Averaging across these 8 bills, the average party split in the House is about 56 points. With the uninformed survey item, the average party split among survey respondents is 26 points. When informed of how the House voted, this gap grows to 41 points, an increase of more than 60%. For half of the bills, informed survey respondents are at least as polarized\textsuperscript{14} as the US House. For the remaining 4 bills, the average party split in the informed condition is 17 points larger than in the standard survey item.

Overall, this evidence shows that standard roll-call survey items may understate the degree to which partisans support what their (polarized) legislators are doing in the US House. When simply told how the parties in the House vote on an issue, citizens and legislators are both more closely aligned and more polarized than when similar assessments are made using uniformed survey items.\textsuperscript{15}

**Study 2**

Figure 2 presents results for Study 2 paralleling the Figure 1 format. All the roll call votes in Study 2 are from the Senate, so the last pair of points shows the proportion of Democrats and Republicans voting yea on the bill on the Senate floor. Levels of support are the proportion of respondents who indicated they support the roll call for each of the 12 roll calls (analysis is weighted and restricted to those who passed the attention screener).\textsuperscript{16}

The results in Figure 2 are consistent with those in Figure 1. Providing simple information about the vote split in the legislature changes the expressed support for the bill in the public.

\textsuperscript{14}Either larger or within 5 points of the observed House split.

\textsuperscript{15}In Appendix Figure A2, we replicate our analysis separately for partisans whose ideology is aligned with their party orientation (i.e., Democrats who are liberal or moderate and Republicans who are conservative or moderate) and those whose ideology is at odds with their party orientation.

\textsuperscript{16}Appendix Figure A3 includes all respondents regardless of how they answered the screener question. Once again, analysis excluding leaners and of pure independents appears in the Appendix. Appendix Figure A4 presents a bar plot of these differences with confidence intervals.
Figure 2: Support for roll call with and without information, Study 2

Note: Closed circles (squares) connect support among Democratic (Republican) respondents for bill from those in standard condition (left) to those in the informed condition (right). Open circles (squares) are the actual rate of support among Democratic (Republican) members of the Senate. These rates were presented to respondents in the informed condition. Figure limited to respondents who passed attention screener.
In addition, in most cases we observe polarization of partisan preferences in the informational relative to the standard condition. Treatment effects are more dramatic for some votes than others, notably for the bills on extending unemployment benefits, firearm background checks, gender discrimination, and changing the formula for student loan interest calculations. On the bill pairing highway funding and reauthorization of the Export-Import Bank, the size of the split between Democrat and Republican is about the same in both conditions, but flips to become aligned with Senate voting in the informed condition.

As with Study 1, responses in the informed condition are more polarized than in the uninformed condition. Averaging across these 12 bills, the average party split in the Senate is about 62 points. With the standard survey item, the average party split among survey respondents is 28 points. When informed of how the Senate voted, this gap grows to 40 points, an increase of more than 40%.

Figure 2 nonetheless shows that the Senate is more polarized than voters in the informed survey condition; in most cases, the Senate split is outside of the population split. We note, however, that our treatment aims to mediate only one of the many differences between the two contexts. It may be that more information would push the split more towards that observed among the legislators. Ultimately, though, the evidence of changing public support in response to simple information illuminates our argument that using the standard responses as a benchmark against which to compare Senate voting is likely to substantially underestimate party divergence in the mass public.

**Effect of Roll Call Survey Question Design on Assessments of Representation**

The results from the previous section establish that responses to standard survey questions on roll call votes depend upon the information available to respondents when answering them. This is prima facie evidence that these responses may not accurately reflect how citizens actually want their representatives to vote on these bills. In this section, we consider if the more-informed responses lead to different conclusions about the status of representation in the United States. Many claim based on evidence from the standard question wording that members of Congress are less
responsive to voters than they should be. In particular, members are more polarized than voters. We show that providing one piece of additional information to survey respondents lessens the divergence of members of Congress relative to their constituents. An implication is that more information might lessen this divergence even further, weakening the evidentiary basis for the empirical claim that members do not represent voters.

It is generally accepted that members of the House and Senate vote less moderately and with more polarization than their constituents prefer. This finding seems to hold for other elites and over many years (e.g., McCarty, Poole, and Rosenthal, 2006; Hill and Tausanovitch, 2015).\footnote{We note that this “delegate” definition of representation is not the only potential form of representation of constituent interests in a democracy.} Almost all of this evidence, however, relies on comparing survey responses of citizens without much contextual or policy-specific information to the votes cast by members of legislatures, who do have a great deal of information. If our argument about information is true and if representatives do represent constituent interests, then the positions of citizens in the informed condition of our experiment should more closely align with the positions of their representatives.

We start by considering how often members and citizens vote on the 12 roll calls from Study 2 with the party-preferred condition. Figure 3 presents kernel density estimates of the number of votes cast with the Senate Democratic position. We define the Democratic position 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. We superimpose densities of senators and citizens on the same graphic with separate densities for respondents in the standard versus informed condition among respondents who passed the screener.\footnote{Appendix Figure A5 presents this figure with all respondents regardless of screener.} The top frame plots densities for Democratic senators and respondents, the bottom Republicans.

Providing citizens with information about the vote split in the Senate makes their rate of voting the party position look more like the rates of senators. The modal Democratic respondent in the standard condition votes with the Democratic position on less than 9 of the 12 roll call votes, while the modal informed Democrat votes with the Democratic position on near 10 votes, close to the
Figure 3: Votes with Democratic position in Senate by condition and party, Study 2

Democratic respondents and Senators

Republican respondents and Senators

Note: Each line is a density of the distribution across number of votes with the Democrats on the 12 roll call votes in study 2. Limited to respondents and senators who voted on all 12 roll call votes, and to respondents who passed the screener.
modal Democratic senator. The modal Republican citizen in the standard condition votes with the Democratic position on almost 6 of these roll calls, compared to modal informed Republican voting with the Democrats on about 2 roll calls. The informed Republican citizen mode is actually more extreme than the Republican senator mode. For both parties, the distribution of votes shifts away from the center in the informed condition.

In total, Figure 3 confirms the longstanding pattern that partisan polarization exhibited with standard survey questions among citizens is notably less than that observed among actual votes cast by members of the Senate. However, providing a simple piece of additional information makes divergence among partisans appear more polarized and closer to that observed in the Senate.

In Appendix Section B, we scale the survey responses in Study 2 using item-response theory models separately by assigned condition. Although our statistical power is limited with only 12 items (we scale Senators using only 12 items to maintain comparability), we find that providing the informational cue moves the distribution of preferences among citizens towards the distribution of preferences of Senators. In particular, the tails of the citizen distribution are more similar to the tails of the Senator distribution in the informed condition – in both conditions, the median is well represented and the filibuster pivots are modestly represented.

**Why do party-relevant cues induce changes in expressed roll-call preferences?**

Our core result is that expressed citizen preferences on roll call vote survey items are unstable in the face of changes in the information available. Partisan citizens respond to information about how members voted in the legislature by changing their reported preferences on those same votes. More specifically, this information yields survey responses more aligned with the behavior of their party’s elected officials. This means that conclusions drawn from survey responses about the quality and nature of the relationship between mass and elite preferences are affected by the survey context.

How should one interpret these patterns? Such an assessment necessarily takes us away from the clear identification provided by the experimental results, although we believe that it is impor-
tant to simply reiterate the core finding and what it means. Many define representation as the correspondence between what representatives do and what their constituents prefer. Our finding is that assessments of the quality and nature of representation are affected by relatively minor alterations in how we measure the preferences of the public. Thus, our work reveals a consequential violation of an important assumption in prior work that uses survey roll call questions to measure mass preferences in the same policy space as legislative behavior.

It is important to take seriously potential objections to our findings. The most likely critique is that evidence of changing preferences in response to party cues indicates blind partisan obedience. To restate this argument in its strongest form, finding that people abandon their own beliefs in the face of party relevant cues may mean that partisan ties cause people to ignore their own preferences in favor of responding in a manner consistent with their party. Or, even more strongly, individuals abandon personal policy preferences for party.

On the one hand, we agree that our results provide clear evidence that people respond to party-relevant cues. But the second part of the criticism, that individuals readily abandon their own well-developed and deeply held positions, has not been demonstrated. In some ways, this criticism boils down to an assumption that people’s survey responses to roll call items *in the absence of any additional information* are more informative of how individuals would vote if they were fully informed than when they are told how the legislature actually voted. We think this is unlikely given past work suggesting that survey respondents often know little about the details of what is going on in Congress and will express opinions even on fictitious bills. Additionally, work by Bullock (2011) shows that the effects of party-relevant cues are largest when respondents know little, but are much smaller when individuals have a strong basis for their own policy views. Nonetheless, none of the data shown so far proves that party does not override sincere policy positions (although if voters really do care more about party than policy, that itself has implications for assessing the quality of representation).

In light of this potential objection, we collected data that helps to understand the degree to which the patterns we observe are reflective of sophisticated reasoning or blind partisan obed-
ence. Specifically, we included in Study 2 survey measures to identify the conditions under which expressed opinions change in response to partisan-relevant information. Following Gerber et al. (2011), we asked respondents about their confidence to evaluate policy in the policy domains corresponding to the roll call vote survey items. These measures are designed to capture the degree to which people understand their (in)ability to map desired policy outcomes to particular policy proposals in different policy domains. If how much people respond to the roll call margin information is correlated with lack of confidence to evaluate good policy in a particular policy domain, it suggests that the response to these cues is more than simple blind partisan obedience.

At the end of Study 2, we asked participants how confident they were in their ability to identify good public policy:

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). Following these items, we also asked respondents to gauge the importance of acting in those same policy areas:

Now, for this same list of policy areas, we’d like to know how important it is to you what government does in that area. Compared to all other policy areas (not just the ones listed below), how important is government policy in this area to you?

Respondents could choose Not at all important (scored 0), Of little importance (1), Somewhat important (2), or One of the most important issues (3).

In Appendix Figure A7 we present the raw individual-level confidence and importance responses for each policy domain. These data show that while individuals think most of these policy areas are important, they are less confident in their ability to identify good policy in each domain. Furthermore, the relationship between confidence and importance is modest, showing that just because someone thinks a policy area is important does not mean that they think they can identify good policy proposals in that domain.
Figure 4: Relationship of confidence evaluating policy to treatment effect of information, Study 2

Note: Each point represents the average confidence participants in the control condition expressed in evaluating policy in that area (x-axis) compared to the average treatment effect of learning the vote split on the roll call in that same area (y-axis). Dashed line is a loess smooth, suggesting that participants change their beliefs about correct policy in response to information on the vote split more when they have less confidence in evaluating policy in that realm.
Next, we examine whether differences in confidence explain the treatment effect of partisan-relevant information. Because our confidence measures are asked post-treatment (e.g., after individuals express responses to the roll call items in either the standard or informed condition), we cannot estimate whether individuals who are more or less confident respond more to the informational treatment. Instead we use the average response to the confidence item in the control condition for each policy domain and party to calculate an average (across-respondent) confidence score in that domain. We then plot in Figure 4 for each party and policy domain the relationship between average confidence (the horizontal axis) and the absolute size of the estimated treatment effect of providing the Senate roll-call split on survey roll call responses (the vertical axis). (Note that while we have 12 roll calls, two are in the same policy domain and so in one policy domain the data point is the average absolute effect of providing information on two roll call items.) For both Democrats (the left panel) and Republicans (the right panel) the line is downward sloping, indicating that as respondents on average become more confident they respond less to the partisan relevant information, just as one would expect given sophisticated responses to partisan relevant information.

Cumulatively, these data suggest survey respondents do not simply abandon their own well-formed and deeply-held opinions to align themselves with their party (see also Bullock, 2011). While this is an observational comparison, we find that citizens change their reported roll call preferences more in policy domains where they are on average less confident in their own ability to evaluate policy. This is consistent with the argument that citizens are not blindly obedient to party cues but instead rely to a greater degree on heuristics when they feel they have less knowledge to evaluate that policy. Thus, even if one evaluates representation using a delegate model in which the appropriate benchmark is the correspondence between the legislator’s behavior and citizens’

19 That said, pooling each confidence evaluation across individuals and regressing on an indicator for treatment assignment yields a coefficient estimate of 0.035 with a clustered standard error of 0.041, indicating very minor and uncertain influence of treatment on confidence.

20 The y-axis is simply the difference in mean support for the roll call between the treatment and standard conditions.

21 In a pooled linear regression (N=22) predicting the absolute size of the average treatment effect for each bill (11) and party (Dem/Rep) as a function of average confidence and an indicator for partisanship, the coefficient on average confidence is -.19 (S.E.=.11, p = .08). Adding average importance to the model reduces the estimated effect of confidence somewhat, to -.14 (S.E.=.11, p = .20).
preferences, the combination of the instability of survey responses with self-expressed lack of confidence in knowledge of the right policy implies that comparisons of legislative vote outcomes to expressed levels of survey support using standard roll call items are likely misleading.

Conclusion

Our argument and experimental evidence here suggest that the standard (without information) survey response to a policy question may not be a good measure of 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 – may require some revision. Our design provides a path forward: with a definition of representation being how the citizen would vote were they to act as the legislator, we can develop new instruments that enrich the survey context to make it more similar to that of voting in the legislature. While it may be near impossible to fully implement the experiment of citizens voting in a legislature (though, ballot propositions may provide some opportunities), our theoretical approach can be used by scholars to clarify and justify the choices made in relating survey responses to representative behavior. Our results show that even small increases in similarity of context can lead to different conclusions about representation, and so we advocate that future work take more seriously issues of measurement.

Our results suggest that for those fielding these kinds of studies, it would be valuable to test the robustness of conclusions to varying the information available to respondents as they make their choices within the survey. Scholars should embed randomizations such as ours in their studies to see how conclusions differ between more and less informed conditions. This may be a productive path towards understanding the systematic sources of measurement error in survey responses and affect our evaluations of the quality of representation and the nature of citizen policy preferences.

One important question for future research is how the single piece of information we delivered in these experiments compares to other potential pieces of information. On one hand, the party split in the legislature may be a very relevant cue for many voters. On the other hand, the party split is a function of many different considerations, and it may be that many voters, particularly those less
attached to either of the parties, would exhibit greater movement with different heuristics. Future work exploring the influence of other informational interventions can help us understand if the effects here are nearer to a lower or an upper bound on movement.

This research also has important implications for those undertaking joint scaling. The assumption of joint scaling is that the item parameters for the actors in different arenas are on average the same, which allows mapping of ideology from one arena to another. Our evidence here suggests this at least does not hold for citizens and members of the House and Senate. We argue that a key distinguishing feature of citizens is less information than legislators. It may be 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 that joint scaling suffers from similar challenges as those we identify here.

Our argument also speaks to theories on representation and accountability.\textsuperscript{22} Many models of accountability revolve around making inferences about different politicians’ types, oftentimes their policy type. These models assume that voters have sufficient knowledge about their own type to know how to compare the types of competing choices for office. Our results imply that voters are uncertain about their own preferences on specific policy questions. Accountability models might explore the conditions under which more general inferences about characteristics can work if voters are not confident about their own preferences on policy.

Moving forward, we argue that more empirical work with a cautious eye towards issues of measurement is needed to consider the functioning of democratic representation. Accepting citizens’ responses to survey question as valid measures of how they prefer their members act relies on important assumptions that make current conclusions about representation uncertain. Future work might investigate alternative informational interventions, or explore more qualitatively or experimentally what voters say or reveal behaviorally that they want from their representatives.

\textsuperscript{22} For a review of the formal literature on accountability, see Ashworth (2012).
References


Appendix

A Weighting to Pew Governance Survey

To construct weights to make the SSI sample look like the sample to the Pew Governance Survey, we asked six questions of the SSI sample equivalent to those asked of the Pew sample. We use these six variables (age, gender, state of residence, level of education, 7-point party identification, and 5-point self-reported ideology) with the rake function from the R library survey (R Development Core Team, 2015; Lumley, 2011) to construct post-stratification weights. The Pew survey itself has post-stratification weights to Census targets, which we use to construct the target distribution for our weighting. We trim the resulting weights to range from 1/8 to 8 to limit variance. The case with the largest pre-trimmed weight was a 55-64 year old male from New Jersey with a high school degree who reported being a conservative Republican. The case with the smallest pre-trimmed weight was a 55-64 year old female from Vermont with a postgraduate degree and a very liberal Democrat.

B IRT estimates of representation

One standard approach to evaluating the quality of representation is to use IRT models to summarize the preferences of individuals and representatives across issues and then to compare those summaries (e.g. Bafumi and Herron, 2010; Hill and Tausanovitch, 2015; Tausanovitch and Warshaw, 2013). The IRT models help mitigate measurement error in each individual item and have been found to be a fair single summary of member votes across thousands of bills (Poole and Rosenthal, 1997). In this section, we follow this standard practice while examining how evaluations of representation from IRT models vary when respondents are voting on the bills (from Study 2) with and without additional information.

We implement an IRT voting model using the R package pscl (Jackman, 2012). We scale the 12 roll call votes cast by each respondent into the same space as the set of senators who voted on those bills. We summarize the implementation below. To place the respondents in the same space as the Senators, we first scaled the Senators by themselves on the 12 roll calls. We then fixed the item parameters estimated from the Senate-only model and applied them to the joint models of Senators and respondents, yielding respondent ideal points in the Senate-space.

Because we use Markov chain Monte Carlo methods for the IRT model, we are able to summarize our posterior beliefs about multiple statistics of polarization. In particular, the United States Senate is a super-majoritarian legislature. We consider how well the distribution of ideal points in the Senate represents the distribution of ideal points in the public, particularly at percentiles of the Senate distribution that correspond to important veto points in the legislature (i.e., the median and the filibuster pivots, Krehbiel, 1998). We consider whether the estimated distribution of citizen preferences (in percentiles) appears more aligned with that in the Senate when citizens are in the informational condition than the standard condition.

In Figure A1, we plot the location of quantiles for citizen and Senator distributions of ideal points separately for citizen distributions in the standard (top) and informed (bottom) condition.\footnote{Appendix Figure A6 includes all respondents regardless of screener.} We characterize features of the posterior distribution of these quantiles for each population. Each point is the posterior median ideal point at that quantile, with lines extending to the posterior 95 percent credible interval. We summarize ideal points for the institutionally-relevant 41st, 50th,
and 60th percentiles of each distribution, along with the more extreme 10th, 25th, 75th, and 90th percentiles.

The top frame with respondents in the standard condition exhibits the conventional pattern of Senators more polarized than members of the public (e.g., Hill and Tausanovitch, 2015). The ideal point at the 0.1 quantile for citizens has a posterior median of -0.4, while for Senators the 0.1 quantile is one standard deviation more extreme at -1.35. Likewise, the 0.9 Senate quantile is 1.28 compared to 0.29 for citizens. The 0.25 and 0.75 quantiles show similar polarization of legislators relative to their constituents. The slope of ideal point to quantile is notably attenuated for the citizens relative to the Senate, suggesting some breakdown in the representation of preferences.

With respect to the institutional rules of the Senate, invoking cloture requires the votes of 3/5ths of the chamber to proceed to considering most bills. Figure A1 shows that the filibuster generates more status quo bias among the observed set of Senator ideal points than among the set of citizen ideal points in the standard condition. Among citizens, the filibuster would have little influence on the set of status quos available to be modified by the legislature. The 0.41, 0.5, and 0.6 quantiles posterior medians are -0.04, 0.01, and 0.06. The Senate filibuster interval, in contrast, ranges from -0.29 to 0.33. Inside this region reside a set of status quo policies that could change in a legislature with the citizens’ ideal points but that could not overcome the filibuster with the Senators’ ideal points.

The quantiles of the citizen distribution in the informed condition are less divergent with those in the Senate. The posterior median 0.1 quantile for citizens is -0.92, half a standard deviation closer to the Senate quantile than in the standard condition. The posterior median 0.9 quantile in the informed condition is 0.55, a quarter standard deviation closer to the Senate. The 0.25 quantile moves from -0.14 in the standard condition to -0.25 informed, and the 0.75 from 0.14 to 0.15.

The filibuster interval for citizens in the informed condition, however, is as narrow as in the standard condition. The posterior medians for the 0.41, 0.5, and 0.6 quantiles are -0.1, -0.04, and 0.02. The informed condition thus appears to change the location of the more extreme quantiles of the citizen distribution, but does not have as large an influence on the location of the center of the distribution. The slope in the informed condition is closer to that in the Senate than the slope in the standard condition.

In sum, the IRT models suggest that providing a single piece of information leads to a population distribution of ideal points that moves towards the Senate distribution, in particular with fatter tails more consistent with the bimodal distribution in the Senate, suggesting that more information and/or contexts making the survey environment more similar to that facing legislators would lead to policy positions closer to the votes we observe in the national legislature.

B.1 Details of IRT model

We jointly scaled the respondents with the 117 members of the 113th and 114th Senates who voted on some of these 12 roll call votes using the Bayesian Markov chain Monte Carlo (MCMC) ideal() in the pscl library in R (Jackman, 2012). To place the respondents in the same space as the Senators, we first scaled the Senators by themselves on the 12 roll calls. We then fix the item parameters estimated from the Senate-only model and apply them to the joint models of Senators and respondents. This creates distributions of ideal points in the space implied by the item

---

24 The item parameters were fixed by setting the prior mean to the posterior mean from the Senate-only model, the prior variance to 100e-3, and no normalization to the distribution of ideal points. All models were burned in for 150,000 iterations, and then 200,000 samples were taken, thinned by 20 yielding 10,000 posterior values summarizing
Figure A1: Change in representative divergence at selected quantiles with information, Study 2

Note: Points represent the estimated ideal point at quantiles of respondent and Senate posterior distributions (posterior median with 95 percent posterior credible intervals). Figure limited to respondents who passed attention screener. Posterior quantiles of the respondent distribution are closer to quantiles of the Senate distribution when respondents are provided information.
parameters from a Senate-only model on the assumption that the item parameters are the same for Senators and respondents. Note that the Senators will have mean zero and unit variance in these joint scalings, but not necessarily the respondents. Although our argument in this essay is that the item parameters for Senators and respondents are *not* equivalent, this procedure allows us to assess how information changes conclusions drawn from the standard joint scaling approach.

### C Additional tables and figures

One concern with the graphical presentation in Figure 1 is that it sorts individuals only on the basis of their partisanship. In fact, some partisans may be “cross-pressured” because their ideological views are inconsistent with their partisan orientation. For this reason, in Appendix Figure A2 we replicate our analysis separately for partisans whose ideology is aligned with their party orientation (i.e., Democrats who are liberal or moderate and Republicans who are conservative or moderate) and those whose ideology is at odds with their party orientation. For the aligned partisans, their behavior closely follows those shown in the pooled Figure 1 analysis. For cross-pressured partisans, the picture is more complicated. There are too few cross-pressured Republicans in our sample for reliable analysis, but for cross-pressured Democrats, they are both generally more conservative and move toward the Republican position on two issues when informed of the House vote. These two bills, on the Keystone Pipeline and the bill described as lower gas taxes, are two notable cases in which the Republican leadership pushed bills that presented policy options targeting unpopular Democratic policies.

Figure A7 plots, for each policy area, the distribution of self-assessed policy importance (on the horizontal axis) and confidence (the vertical axis). We add to this plot the average importance score (the vertical grey line), the average confidence score (the horizontal grey line), and the relationship between the two (the dashed black line, which is from a local polynomial smoother). Each dot is a single survey respondent’s answers, with the points jittered for clarity. Several important patterns emerge.

Figure A8 plots, for each policy area, the relationship between self-assessed confidence (the vertical axis) and policy importance (horizontal axis). Each black line is a loess smooth of the individual relationship for that policy area. We indicate the average importance score (vertical grey lines) and average confidence score (horizontal grey lines) for each policy area. We also present the tabulation of each response at each value on the two axes, for example only 6% of responses to the question about policy importance indicated the policy was “not at all important.” Several important patterns emerge.

First, on average, respondents think most policy areas are important. The average importance score across all policy areas is 2.14, which is slightly more than somewhat important. Only 21% of evaluations scored the policies as little or not at all important. Second, while respondents think policy in these areas is important, they are on average less confident in their ability to pick policies that give them what they want. The average confidence score is 1.56, which is roughly halfway between a little and somewhat competent. While 41% of evaluations indicated the policy area “one of the most important,” only 22% of evaluations indicated the individual felt “very confident” that they could distinguish good from bad policies. Further, the loess smooths show that there is only

---

25 Appendix Figure A7 plots the full distributions.
Figure A2: Support for roll call with and without information by party-ideology cross-pressure

Note: Closed circles (squares) connect support among Democratic (Republican) respondents for bill from those in standard condition (left) to those in the informed condition (right). Open circles (squares) are the actual rate of support among Democratic (Republican) members of the House. The top frame presents support for non-cross-pressured partisans, liberal and moderate Democrats and moderate and conservative Republicans. The bottom frame presents cross-pressured conservative Democrats; there are too few liberal Republicans to plot.
Figure A3: Support for roll call with and without information, Study 2, All respondents

Note: Closed circles (squares) connect support among Democratic (Republican) respondents for bill from those in standard condition (left) to those in the informed condition (right). Open circles (squares) are the actual rate of support among Democratic (Republican) members of the Senate. These rates were presented to respondents in the informed condition.
Figure A4: Change in party divergence on roll call with and without information, Study 2

Note: Each bar presents the change in the difference in support for the yea position on the roll call between Democratic and Republican identifiers in the informational condition relative to the control condition. Estimated with an OLS difference in difference regression with robust standard errors and stratification weights. Confidence bars extend to 95 percent. The negative value for the Highway Funding bill is due to the partisans switching majority positions from the standard to the informational condition (see Figure 2).
Figure A5: Votes with Democratic side in Senate by condition and party, Study 2, All respondents

Democratic respondents and Senators

Republican respondents and Senators

Note: Each line is a density of the distribution across number of votes with the Democrats on the 12 roll call votes in study 2. Limited to respondents and Senators who voted on all 12 roll call votes.
Figure A6: Change in relative polarization with information, Study 2, All respondents

Note: Points represent the estimated ideal point at quantiles of respondent and Senate posterior distributions (posterior median with 95 percent posterior credible intervals). Posterior quantiles of the respondent distribution are closer to quantiles of the Senate distribution when respondents are provided information.
Note: Compares each respondent’s assessment of the importance of a policy area to their own confidence that they are able to evaluate policy in that area. The policy areas correspond roughly to the 12 roll call votes of Study 2.
Figure A8: Relationship of confidence to importance by policy area, Study 2

Note: Each line is a loess smooth of respondents’ assessment of their own confidence that they are able to evaluate policy in an area to the importance they ascribe to policy in that area. The policy areas correspond roughly to the 12 roll call votes of Study 2. Vertical and horizontal lines indicate the mean confidence and importance for each policy. Percentage of responses in each category across roll calls indicated on each axis.
a weak positive relationship between believing a policy area is important and believing one can identify good public policy. Thus, it is not the case that simply thinking something is important means individuals have great confidence that they can pick which policies are best in that area.