Who Cares? Measuring Preference Intensity in a Polarized Environment

Charlotte Cavaillé, Daniel L. Chen

and Karine Van Der Straeten

Who Cares? Measuring Preference Intensity in a Polarized Environment[*]

Charlotte Cavaillé

(Ford School of Public Policy, University of Michigan)

Daniel L. Chen

(Toulouse School of Economics - IAST)

Karine Van der Straeten

(Toulouse School of Economics - IAST)

January 31, 2022

Many questions in political science require knowing not only what voters want (preference orientation) but also how much they want it (preference intensity). In this paper, we assess two methods for measuring individual differences in preference intensity. One method — issue importance items — asks respondents to self-report how important a given set of policy proposals is to them personally. Another — Quadratic Voting for Survey Research (QVSR) — gives respondents a fixed budget to 'buy' votes in favor of (against) these policy proposals, with the price for each vote increasing quadratically. We provide theoretical arguments explaining why, in a polarized environment where some respondents may feel pressured to pay lip service to the party norms, one should expect QVSR to offer a better measure of preference intensity. Using Likert items as the benchmark, we find that QVSR more consistently differentiates between intense and weak preferences, as proxied by respondents' behavior on simplified real-world tasks. Revisiting debates on the determinants of policy preferences, or the congruence between mass opinions and the policy status quo, we show that conclusions reached when using Likert items alone change once differences in preference intensity are better accounted for.

Attitudinal survey data is central to research in political science. It is used by political scientists to understand how voters form policy preferences (Zaller 1992) or to assess the congruence between policy outcomes and citizens' preferences[1] (Gilens 2012). Because policymaking involves trade-offs, knowing what people want, i.e. preference orientation, is often not enough. Also important is knowing what they want the most and what they want the least, i.e. preference intensity.

Despite its relevance, preference intensity is rarely explicitly measured in surveys. Most often, respondents are asked to evaluate a battery of policy statements using a Likert scale ranging from 'strongly agree (favor)' to 'strongly disagree (oppose).' People asked about their policy preferences using this measurement tool have only limited incentives to consider trade-offs across issues (e.g., immigration control at the expense of access to the European single market). As a result, people's answers convey little information regarding which issues they prioritize and which issues they are willing to compromise on, something we call the *abundance problem*. Another concern, particularly relevant in the American context, is the distortion introduced by partisan motives. For politicized policy issues, polarized ideological messaging and affective partisanship combine to generate bi-modal response distributions. In this case, the same response category (e.g. *strongly* agree or *strongly* disagree) might include respondents who care about the issue and respondents who do not care as intensely and are merely "paying lip service to the party norm" (Zaller 2012).[2] In the American context, partisan polarization makes it even more difficult to disentangle preference orientation and preference intensity, something we call the *bunching problem*.

The abundance problem is especially salient in the literature assessing the congruence between policymaking and mass preferences. Take, for example, policies such as a federally-mandated parental leave, an increase in the minimum wage or fiscal discipline. Preferences measured using Likert items indicate broad support for these policies (see Section 3), implying a substantive mismatch between policy outcomes and mass preferences. Yet, a mismatch on an issue a majority cares about is substantively different from a mismatch on an issue only a few care about.

---

[1] In this paper, we define 'preferences' as evaluations of statements, in this case policy-relevant statements, ranging from positive to negative. We also distinguish between 'preference orientation' and 'preference intensity.' Our definitions are in broad agreement with similar concepts in social psychology such as 'attitude orientation' and 'attitude strength.'

[2] See also Johnston, Lavine and Federico (2017) for an example of how policy issues people care about (e.g. abortion, race or immigration) predicts partisan identity, which in turn affects answers to survey questions on issues people care less about (e.g. redistributive policies).

The bunching problem is particularly concerning for students of preference formation. First, with only limited variance to leverage, end-of-scale bunching mean results are more sensitive to underlying modeling assumptions. Second, because of end-of-scale bunching, empirical strategies that rely on Likert items to measure preferences tend to favor predictors tied to partisan identity at the expense of other predictors (see Sections 1 and 3).

For research to move forward, we need better ways of measuring differences in preference intensity. The purpose of this paper is to assess, both theoretically and empirically, the advantages and limitations of two easy-to-implement methods designed to improve how we measure preference intensity.[3] One method combines a Likert item with an issue importance item that asks respondents whether or not an issue is personally important to them (*Likert+* for short). Likert+, because it also places respondents in a world of abundance and does not penalize partisan motives, should theoretically face some of the same limits as Likert items.

We also study a new way of measuring preference intensity called Quadratic Voting for Survey Research (*QVSR*), which can be easily implemented in an online survey using a free web application. This method mimics real world trade-offs by asking respondents to vote on a bundle of issues under conditions of scarcity: respondents are constrained by a fixed budget with which to 'buy' the votes. Because the price for each vote is quadratic, it becomes increasingly costly to acquire additional votes to express more intense support for (or opposition to) the same issue. The budget constraint and quadratic pricing compel respondents to arbitrate between the issues listed in the bundle and make it costly to express intense preferences by voting repeatedly for the same issue (Lalley and Weyl 2018). This set up contrasts sharply with Likert items' world of abundance in which respondents face no trade-offs and can pick end-of-scale responses (e.g. strongly favor/oppose) at no cost. With QVSR, respondents are forced to de-bunch. Under the assumption that the psychological cost of misreporting one's preference depends positively on preference intensity, we formally show that

---

[3] Note that we are interested in survey instruments designed to measure individual-level differences in preference intensity. There is a growing literature discussing how to best measure which policy issues voters care about the most using observational or experimental *voting* data i.e., a forced choice between real or hypothetical candidates or platforms (e.g. Horiuchi, Smith and Yamamoto (2018); Abramson, Koçak and Magazinnik (2019); Leeper and Robison (2018); Hanretty, Lauderdale and Vivyan (2020)). These methods recover population-level estimates under varying, and often conflicting, modeling assumptions. In contrast, we focus on individual-level differences in preference intensity, which can be aggregated into population-level descriptive statistics or used for the study of preference formation. Unfortunately, a thorough investigation of the theoretical and empirical overlap between these choice-based methods and the survey instruments discussed here is beyond the scope of this paper.

respondents are expected to prioritize issues they care about the most at the expense of issues they care about the least (Section 1).

We then examine whether the additional information (or higher entropy) obtained using Likert+ or QVSR in place of Likert items is informative of differences in preference intensity. To do so, we randomly assign individuals to take the same survey varying only the technology used to measure policy preferences. We then ask respondents to perform a number of choice tasks related to policy issues asked about in the survey (e.g. a donation to a non-profit advocating for gun control, letter writing to a senator, etc.). Respondents have the option to act or not act. If they decide to act, the tasks are designed to involve a trade-off between two policy issues. Furthermore, we measure behavior on a continuous scale such that higher values on this behavioral outcome imply more resources allocated to this issue, whether in dollars (e.g. amount donated) or effort (e.g. length of the letter). We then compare each tool's ability to discriminate between respondents whose behavior suggests more intense preferences on a given issue and those whose behavior suggests weaker preferences (Section 2).[4]

Across a total of six behavioral outcomes, the same pattern emerges: the point estimate for QVSR is larger (indicating more discriminatory power) than that for Likert+, which is in turn larger than that found for Likert. Using Likert's own discriminatory power as a benchmark, we find that for three out of six outcomes, we can reject the null that QVSR and Likert have the same discriminatory power. In contrast, on all outcomes, we cannot reject the null that Likert+'s discriminatory power is the same as Likert's. We tentatively conclude that, assuming researchers can afford the additional survey time necessary to explain respondents how QVSR works, this tool provides a significant improvement over Likert.

In the third and last section of the paper, we discuss how research on preference formation can benefit from improved measures of preference intensity. Specifically, we examine the relationship between support for a policy proposal and the likelihood of being personally affected by its implementation. We find that, while policy preferences measured using Likert items convey little information regarding the likelihood of benefiting from a given policy, policy preferences measured

---

[4] By definition, more intense preferences on one issue relative to another implies resolving resource allocation trade-offs (in terms of money, time or effort) in favor of the former not the latter.

using QVSR do. This suggests that people directly affected by a policy hold more intense preferences over this policy than individuals not affected by it. We also briefly discuss the issue of policy congruence. In particular, we show that, when using Likert, large majorities of Republicans and Democrats express support fiscal discipline. Using QVSR, we find that Republicans care more about fiscal discipline than Democrats. This asymmetry in preference intensity is often overlooked in discussions of congruence. It suggests that demand-side factors likely played an important role in the recent failed negotiations over the Build Back Better Bill.

In sum, this paper provide a first step toward better disentangling preference orientation and preference intensity. Jointly, our formal (Section 1), experimental (Section 2) and descriptive (Section 3) results show the benefits of grounding survey data in a theory of choice. Absent careful attention paid to the relationship between quantities of interest and measurement tools, many empirical findings risk being incorrectly or imprecisely interpreted.

1 Measuring Preference Intensity

To compare Likert, Likert+ and QVSR, we start from a simple formal model. In this section, we explain the model's main ingredients, intuitions, and results. A more detailed exposition can be found in Appendix A.1.

Consider K proposed policy changes. Individuals' preference on these issues are characterized by two main parameters: preference orientation and preference intensity. Preference orientation on issue $k = 1, ..., K$ captures whether the individual agrees or disagrees with the policy proposal as worded. Formally, we represent preference orientation by a real number α_{ik} in the interval $[-1, 1]$, where $\alpha_{ik} = 1$ means full agreement with the policy, and $\alpha_{ik} = -1$ full disagreement. Intermediate values correspond to intermediate opinions, something that might be due to ambivalence (e.g. support for the policy's core principle but not as stated in the survey question) or neutrality towards the issue. Preference intensity describes how much an individual cares about the proposed policy change, which we represent formally by a positive number β_{ik} in the interval $[0, 1]$, where $\beta_{ik} = 0$ means that this individual does not care about this issue, and $\beta_{ik} = 1$ means that the question is of the utmost importance to her. In the context of this paper, 'caring' about an issue captures one's

likelihood to take action on this issue, such as signing a petition in favor of a reform, liking/sharing a news story that opposes a given political message, volunteering time and effort in an NGO supporting a cause, etc.. Because of individual constraints (e.g. time, income, mental energy, etc.), or policy trade-offs (e.g. addressing a pandemic at the expense of individual liberties, fiscal constraints that limit the types of social programs that can be funded), people do not take action on all issues. Instead, people will prioritize issues they care about the most (high preference intensity) and compromise on issues they care about the least (low preference intensity).

1.1 Measuring Preference Orientation and Preference Intensity with Surveys

Next, we examine the relationship between tool-specific survey answers on the one hand, and α_{ik} and β_{ik} on the other. For expository purposes, we focus on the versions of Likert, Likert+ and QVSR used in our empirical study.

Likert Item Following best practices in survey design (Malhotra, Krosnick and Thomas 2009), we asked respondents assigned to the Likert tool about their opinion using a sequence of two branching questions: "Do you favor, oppose, or neither favor nor oppose: [Example] Giving same sex couples the legal right to adopt a child?" Respondents who pick the favor or oppose option then see the following prompt: "Do you favor [oppose] that a great deal, moderately, or a little?" Respondents who initially select "neither" are not asked a follow-up question. Likert items' wording suggests that answers collected using this survey tool capture a mix of preference orientation and intensity. Indeed, when an individual reports that she opposes giving same sex couples the legal right to adopt a child *a great deal*, she conveys not only that she opposes the proposed policy, but also that she feels strongly about it. Formally, we denote by $x_{ik} \in [-1, 1]$ the answer that individual i would give to this question under Likert if fully sincere. Later in this section, we examine what happens when respondents have the incentive to systematically deviate from their sincere views.

Likert Item Followed by Issue Importance Item (Likert+) With Likert+, there is a follow-up question: "How important is this issue to you personally?" The response options are: "extremely important; very important; somewhat important; not too important; not important at all." Assuming

6

people are fully sincere, Likert+ collects two pieces of information: x_{ik}, as with a standard Likert item, and β_{ik}, directly measured using the follow-up question. The expected improvement with Likert+ compared to Likert is straightforward: it provides a direct measure of preference intensity.

Quadratic Voting for Survey Research   QVSR draws on research on quadratic voting (Posner and Weyl 2018).[5] Each person starts with an artificial budget of credits, which she can use to 'buy' votes in favor of (against) a fixed set of proposals. Respondents can cast more than one vote per proposal. A distinctive feature of QVSR is that the cost of buying votes is quadratic: buying one vote in support of (against) one proposal costs one credit; buying two votes for the same proposal costs four credits; buying three votes costs nine credits; and so on. In our own survey, respondents assigned to the QVSR tool are given a budget of 100 credits to spend across ten different survey questions. Figure 1 shows what the survey looks like to respondents. In this example, the first three issues involve positions on the border wall, abortion and parental leave. Respondents can scroll down to vote on the other seven issues examined in the survey. The order in which the 10 proposals are presented is fully randomized. Notice that respondents can use credits in favor of or against each proposal. The cost of each vote increases according to the quadratic form and is displayed below each question. Remaining credits are displayed on the top of the screen. Respondents can go back to revise their answers until they are satisfied with how they have allocated their credits. The maximum that respondents can spend in favor of (against) any question is 10 votes (100 credits) though doing so would mean not being able to cast a vote on any of the other 9 issues. Respondents do not have to spend all their 100 credits.

With QVSR, respondents face a budget constraint that prevents her from opposing or favoring *a great deal* too many options. When the budget constraint is binding, an individual cannot report her policy preferences as she would have had under Likert or Likert+. Instead she has to deviate from her unconstrained response on each issue. In our model, we assume that the reluctance to deviate

---

[5] In this original formulation, Quadratic Voting (QV) was intended as a means of arriving at efficient social decisions when voting on policies that have a high probability of being implemented. Lalley and Weyl (2018) primarily assume that influencing policy is the main motivation of citizens. Similarly, Casella and Sanchez (Forthcoming) compare the performance of QV and Storable Votes in a realistic voting context In the theory section of their paper, they assume that voters only care about the outcome of the election (and thus use their votes to get the best possible outcomes for them). In contrast, we examine the implications of QV in a context that is quite different from voting, namely survey research, where influencing policy is probably not the main motivation of citizens.

Figure 1: Screenshot of the QVSR Version of the Survey

from one's unconstrained answer on one issue is positively correlated with preference intensity. We formally show that, compared to standard Likert items, an individual will allocate more credits to issues over which she has more intense preferences. This partially follows from our definition of preference intensity as the willingness to take action on an issue. Indeed, not reporting one's true preference on one issue liberates credits to take action —in this case publicly express one's true preference— on another issue. In other words, not reporting one's true preference on one issue is a form of compromise and people should be more (less) willing to compromise on issues they care little (a lot) about. Note that answers collected using QVSR are still expected to be a mix of α_{ik} and β_{ik}, as with Likert items. Yet QVSR votes should be more informative of preference intensity than answers measured using Likert.

With Likert+, respondents are asked about β_{ik} directly. By design, Likert+ relies on two questions instead of one, and in that sense, gathers more information than QVSR. This can be seen as a positive or a negative depending on how much one values parsimony. Another major difference is

8

that QVSR compels individuals to explicitly compare across issues. In Likert+, this comparison is only implicit: asked about the importance of a given issue, respondents could have their answers to the other nine issue importance items in mind but there is no way for the researcher to ensure that this is indeed the case. Under QVSR, the procedure is baked into the hard budget constraint.[6] However, in our view, the main advantage of QVSR over Likert+ follows from a different set of considerations pertaining to the fact that respondents are only partially sincere. We now turn to this issue.

1.2 Non-Sincere Responses: Measuring Policy Preferences in a Polarized Environment

Another potential issue with Likert items, shared by most survey instruments, is that respondents' answers might be only partially sincere. Specifically, strategic or partisan motives may induce respondents to inflate their opposition to (support for) a given reform. The same applies when respondents are asked how important an issue is to them personally. Consider the example of a Republican voter being asked whether she favors or opposes building a wall between the U.S. and Mexico and whether this issue is important to her personally. Assume further that this respondent has some doubts about the adequacy of this policy to address immigration issues, or thinks that this issue is only of secondary importance compared to more pressing economic issues. Still, since this issue is high on the agenda of the Republican party, out of loyalty to her preferred party or candidate, or profound animosity to the competing party, she may feel a psychological pressure to answer that she strongly supports this reform and that this issue is of the utmost importance (Bullock and Lenz 2019).

We model respondents as being intrinsically motivated to report their true preferences, which we call the *sincerity motive*, as well as simultaneously pursuing other objectives when answering surveys. For example, they might also want to strategically influence policy or signal a partisan identity, something we call the *signaling motive*. We assume these conflicting motives to be present whatever the survey technology.

---

[6] A previous study of how people engage with QVSR finds that respondents proceed by trial and error, moving back and forth across items to best distribute their votes and credits, and thus being forced to compare (Quarfoot et al. 2017).

Likert Item   With standard Likert items, we show that answers are typically a weighted average of the sincere answer (x_{ik}) and some signaling target. In particular, in a polarized environment where parties are quite extreme, we expect the partisan motive to push respondents to systematically inflate their views, and to report stronger support or opposition than what they actually feel. In the model, we denote by $\widehat{x}_{ik}^{L}$ the answers under standard Likert assuming mixed motives. Our expectation in a polarized environment is that $\left|\widehat{x}_{ik}^{L}\right| > |x_{ik}|$.

As a result, responses will bunch at the extremities of the response scale (strongly favor/oppose) making it difficult to distinguish between 1) respondents who strongly favor (or oppose) a given policy change and are ready to act on this preference and 2) respondents who, while expressing the same position, do not care as much and are merely "paying lip service to the party norm" (Zaller 2012), or expressing their dislike for the opposite party (Mason 2015). In other words, if political polarization is high, voters polled using a Likert item will appear to strongly disagree on key issues, when in practice, only a subset of voters really care. When the latter group is unequally distributed across parties (e.g. members of one party strongly support a policy while members of the other only weakly oppose it), Likert items can provide an inadequate picture of public opinion, one that overlooks important asymmetries in preference intensity across parties.

In a polarized two-party system where political elites provide coherent ideological bundles, such bunching at the extremities of a scale occurs on more than one key issue: a strong signaling motive implies 'strong' support for a bundle of policies associated with their party. In this case, most of the variance in the data is due to *between* party differences. The available information on differences *within* each political party is, by contrast, much smaller. Relatedly, this "bunching to the extremes" on more than one issue results in limited individual-level information regarding which policies a respondent cares about the most. Instead, she appears to care about many things, with little information on how she might address policy trade-offs central to democratic politics.

Likert+   Likert+ combines a Likert item and an issue importance item. With regards to the Likert item, the logic is the same as the one stated above ($\left|\widehat{x}_{ik}^{L+}\right| > |x_{ik}|$). Regarding the issue importance item, a similar partisan motive may push individuals away from reporting their 'true' preference intensity, β_{ik}. Denoting by $\widehat{\beta}_{ik}^{L+}$ the answers to the issue importance question under Likert+, one

might consequently expect that, assuming a strong signaling motive, $\widehat{\beta}_{ik}^{L+} > \beta_{ik}$ for issues that are high on their preferred party's agenda.

QVSR With QVSR, as explained above, if the budget constraint is binding, an individual cannot report her policy preferences as she would have had using Likert items (the $\widehat{x}_{ik}^{L}$), and she will comparatively allocate more credits to issues over which she has intense preferences. A second advantage of QVSR, which is more subtle, is that QVSR not only better captures β_{ik}, it may also provide a more accurate measure of x_{ik}, i.e. the "sincere" component of the unconstrained Likert answer. To understand why, remember that when a signaling motive is present, answers under Likert (the $\widehat{x}_{ik}^{L}$) are typically a weighted average between the sincere answer (the x_{ik} parameter) and some signaling target. The lower the strength of the sincerity motive compared to the signaling/partisan motive, the further away the Likert answers will be from the sincere answers. Now, remember that under QVSR the respondent comparatively allocates more votes to issues over which she has intense preferences. Under the assumption of a positive correlation between preference intensity and the strength of the sincerity motive, the model predicts that with QVSR, the individual will allocate more credits to issues on which the unconstrained answer ($\widehat{x}_{ik}^{L}$) is closer to the sincere answer (x_{ik}), thus better capturing the true x_{ik} parameter.

Lastly, we can qualify the predictions made earlier regarding the comparison between QVSR and Likert+. When one explicitly takes into account the fact that some respondents may be insincere and have the incentive to inflate the importance of polarizing issues, then answers to the issue importance question will be a noisy measure of the true intensity parameter. In such a context, and based on the reasoning presented above, QVSR's 'revealed preference' approach could potentially provide a better measure of preference intensity than Likert+. In the end, however, the comparison between Likert+ and QVSR depends on the weight of the sincerity motive. Given the impossibility to directly assess the latter, which tool performs best in the current U.S. context is ultimately an empirical question.

Based on the argument made in this section, we expect both QVSR and Likert+ to better measure preference intensity than a standard Likert item. However, in the current American context, where strong partisan motives compete with the sincerity motive (Bullock and Lenz 2019), we expect less

improvements from switching from Likert to Likert+ than from switching from Likert to QVSR. The comparison between QVSR and Likert+ depends on the relative weight of the sincerity motive, something we do not observe.[7]

2 Experimental Strategy and Results

Each survey technology can be thought of as a classification tool that distributes a surveyed population across a fixed number of response categories. The survey tool that best measures preference intensity is the one that best classifies (or discriminates) respondents from most to least "intense", i.e. from most to least likely to take an issue-relevant action. To compare each survey tools' classification abilities (or discriminatory power), we rely on a survey experiment.[8] Respondents (N=3666) were randomly assigned to one of the three survey tools and asked to provide their opinion on the following 10 policy issues:

Do you Favor or Oppose:
- **[sameS]** Giving same sex couples the legal right to adopt a child
- **[gunC]** Laws making it more difficult for people to buy a gun
- **[wall]** Building a wall on the U.S. Border with Mexico
- **[paidL]** Requiring employers to offer paid leave to parents of new children
- **[affA]** Preferential hiring and promotion of blacks to address past discrimination
- **[equalP]** Requiring employers to pay women and men the same amount for the same work
- **[minW]** Raising the minimum wage to $15 an hour over the next 6 years
- **[abort]** A nationwide ban on abortion with only very limited exceptions
- **[cap]** A spending cap that prevents the federal government from spending more than it takes
- **[env]** The government regulating business to protect the environment

---

[7] We tried to experimentally manipulate the weight of the sincerity motive by priming partisan identity and thus increasing the weight of the partisan signaling motive. Ultimately, the treatment effect was too weak to affect survey answers on very salient issues such as abortion, gun control or building a wall on the border with Mexico. Future experiment tests might focus instead on manipulating the sincerity motive.

[8] The main component of our survey took place from October 5 to October 9, 2018. The survey was administered to a general population of US citizens over the age of 18. The survey company, GfK-Ipsos, uses a probability-based web panel designed to be representative of the U.S. population. For an overview of the survey design, see Appendix A.2 and A.6.

Respondents assigned to the QVSR tool had to watch a video explaining how QVSR works before they answered the policy questions.[9] After expressing their opinion, respondents were asked to perform a task with "real world" consequences. Specifically, they were given the opportunity to donate lottery money to single-issue advocacy groups (e.g. immigration control, gun control). This first behavioral task elicits the willingness to incur a monetary cost to promote a political cause one agrees with. At the beginning of the task, respondents were told that, as participants to the survey, they had been automatically entered into a lottery with "a prize of $100 for 40 randomly selected respondents (among 4000 or so)." They were then prompted to imagine that they were among the winners and asked whether they wanted to donate part of their lottery money to an advocacy group. They had a choice between four advocacy groups working on two issue areas: immigration and gun control. For each issue area, we chose organizations that fall on different sides of the political divide: for and against immigration, as well as for and against gun control. Respondents could choose not to donate or to donate to one of the four advocacy groups. Whatever they did not donate, they could keep.

Four months later (between January 31 and February 18, 2019), we recontacted a random subset of these respondents and asked them to answer the same 10 survey questions using the survey tool they were assigned to in the first wave (number of responses, N= 1569).[10] We then collected information on two additional behavioral tasks.[11] First, respondents were asked to play dictator games with an individual who shares their view on one policy issue and does not share their view on a second policy issue (see details below). This second task approximates a situation in which respondents face a trade-off: reward or punish another participant —identified as an independent— who agrees with them on one issue but disagrees with them on another. To start, we asked respondents how they would behave in three dictator games involving a Republican, a Democrat and an Independent respectively (the order was randomized). Respondents had the option to donate anywhere between $0 and $100 of some lottery money (following a similar set up as the one in

---

[9] The video can be found at: `https://www.youtube.com/watch?v=GrY_RzDsqLY`.

[10] Participation in wave 2 is not predicted by treatment condition and policy preferences in wave 1, nor by partisanship. See Appendix A.2 for more on balance across treatment conditions and wave 2 participation.

[11] Wave 1 took on average 9 min to complete. The second wave enabled us to collect additional behavioral data without burdening respondents in wave 1.

wave 1). After they made their decisions, we informed them of the Independent respondent's preferences on two issues, namely opposed to the border wall and similarly opposed to gun control. We asked respondents if they wanted to change the amount they had previously decided to donate to this individual.[12] We chose this specific mix of preferences (anti-gun control and pro-immigration) because they rarely coexist in practice, and can be, as a result, expected to generate a trade-off among the majority of respondents. For example, for most Democrats, punishing the Independent for donating to a pro-gun organization meant also punishing her for donating to a pro-immigration organization.

Respondents were also given the opportunity to write to their Senators about real bills that were moving through Congress at the time of the survey, one about abortion and the other about raising the minimum wage. This third behavioral task captures the willingness to spend time and effort promoting a political cause one agrees with. Respondents were given the option to write a short text about one of two policy proposals being discussed in Congress at the time. We did not mention who the bill sponsors were, only the content of the bills. The texts provided by the respondents were then integrated into a letter, which was ultimately sent to the Senate committees in charge of reviewing the policy proposals (Adida, Lo and Platas 2018). Comments were anonymous.

Section A.6 in the Appendix provides screen shots of all three tasks.[13] Table 2 provides an overview of the outcome variables derived from these tasks and used in the remainder of the analysis. Note that throughout the paper, when we examine the relationship between survey answers and behavior, we only use answers collected in wave 1, though using data collected in wave 2 does not change the results (See Appendix A.3).[14] Table 1 summarizes the sample sizes for each wave.

---

[12] The Independent's preferences were conveyed using information on donation patterns in wave 1. In practice, respondents could only donate to one non-profit. After implementing the lottery, we picked an independent to receive the funds based on her survey answers, not her donation behavior.

[13] Note that the three behavioral tasks used in our study, while more costly than simply answering a survey question, remain relatively low-stakes. Evidence that Likert+ or QVSR outperform Likert for such low-cost outcomes suggests that they will also outperform Likert when predicting more costly behaviors.

[14] Readers should nevertheless keep in mind that, in at least two cases (dictator game and letter writing tasks), the attitudinal data is analyzed alongside behavioral data collected four months apart.

Table 1: Sample Sizes: Overview

Wave 1	Likert	Likert+	QVSR
Consented	1306	1305	1289
Valid observations	1215	1217	1116
Wave 2	Likert	Likert+	QVSR
Valid observations	523	539	470

Wave 1 dropout rate is roughly twice as large in the QVSR treatment (14 percent) than in the Likert and Likert+ conditions. Most of the dropout occurred when respondents were asked to watch the video explaining how QVSR works. We found no evidence that this attrition was correlated with key variables of interests (see Appendix A.2).

2.1 Estimation Strategy

As already mentioned, each of the three survey technologies can be thought of as an ordinal classification tool. Each tool differs from the other two in terms of 1) the total number of ordinal categories and 2) the distribution of observations across these categories. We discuss each point in turn.

Number of Response Categories Likert has 7 response categories: from -3 (strongly oppose) to 3 (strongly favor), centered around 0 (neither/nor). QVSR in theory ranges from -10 (up to 10 votes against) to 10 (up to 10 votes in favor), centered around 0 (no votes). In practice, the highest number of votes cast is 7^{15} not 10, meaning that for a given issue, QVSR sorts respondents across 15 response categories in total. In the case of Likert+, we combine the Likert and issue importance items into one single scale by multiplying values of the first Likert branching question (favor = 1, opposes = -1, neither = 0) by the values of the issue importance scale. The result is an ordinal scale with 11 response categories ranging from -5 (oppose, extremely important) to 5 (favor, extremely important) and centered around 0 (neither/nor).[16] In the remainder of the

---

[15] There are a few respondents casting 8 or 9 votes, we merge them with the '7 votes' response categories.

[16] We discuss alternative combinations of the two items in Appendix A.3.

analysis, we use normalized answers such that the lowest possible answer corresponds to zero (-3 / -5 / -7 for Likert, Likert+ and QVSR respectively) and the highest possible answer to 1 (3/5/7).

Table 2: Behavioral Outcomes and Relevant Survey Question

Variable	Description	Mean (Stand. dev.)	Survey question expected to correlate with behavior
Donation to gun-related advocacy group	Equal to the $ amount donated multiplied by 1 if donated to pro gun control and -1 if donated to anti gun control advocacy group.	9.8 (33.5)	Laws making it more difficult for people to buy a gun
Donation to immigration-related advocacy group	Equal to the $ amount donated multiplied by 1 if donated to anti immigration and -1 if donated to pro immigration advocacy group.	1.6 (28.7)	Wall on the border with Mexico
Punishment of Independent respondent (1)	Equal to the $ amount *taken off* the amount previously donated to the Independent. If respondent *gave* additional $ then amount multiplied by -1. Note that individuals who donated in the first round but did not change their answers in the second, and individuals who did not donate in the two rounds are both coded as 0.	4.9 (13.5)	Laws making it more difficult for people to buy a gun/ Wall on the U.S. border with Mexico
Punishment of Independent respondent (2)	Equal to the $ amount *taken off* the amount previously donated to the Independent as a *proportion* of the amount originally donated. If respondent *gave* additional $ then amount multiplied by -1. Note that individuals (N = 24) who did not donate in the first round but nevertheless asked to take $ away from the Independent are coded as missing (N = 24).	0.17 (0.38)	Laws making it more difficult for people to buy a gun/ Wall on the U.S. border with Mexico
Letter writing on the minimum wage bill	Equal to the length of text written (number of characters).	76 (139)	Raising the minimum wage to $15/h over the next 3 years (absolute values)
Letter writing on the abortion bill	Equal to the length of text written (number of characters).	59 (90)	A nationwide ban on abortion with only very limited exceptions (absolute values)

Distribution of Responses   As shown in the bottom panel of Figure 2, when preferences are measured using a Likert item, the distribution of answers to the gun control item is uni-modal: most answers bunch on one extreme of the scale. This pattern is somewhat less pronounced in Likert+, implying that, while most respondents strongly support gun control, not everyone believes this

issue to be personally important to them. Responses in QVSR, by design, exhibit no such bunching patterns.

More response categories and less bunching imply more information (i.e. higher entropy) for QVSR and Likert+ on the one hand than for Likert on the other.[17] One might wonder if the main advantage of Likert+ and QVSR over Likert is simply more response categories to choose from. This would suggest comparing QVSR and Likert+ to Likert items with more than 7 response categories. However, as shown by Revilla, Saris and Krosnick (2014), with Likert items, the quality of the data collected deteriorates as the number of response categories increases. In other words, the issue is not the number of response categories *per se* but the ways in which a given survey instrument distributes respondents across available response categories. In other words, of importance here is assessing the extent to which Likert+'s targeted wording and QVSR's induced scarcity generate not only more variance, but variance that conveys additional information regarding differences in preference intensity.

If Likert+ and QVSR's higher entropy is more than just noise then, comparing individuals with the highest score to individuals with intermediary scores, the behavior of the former should signal more intense preferences than the behavior of the latter. In practice, this implies a positive and monotonic relationship between response categories on the one hand, and the mean of the outcome of interest conditional on the response category on the other.

To assess the extent to which a given survey tool successfully sorts respondents into increasingly discriminating categories, we first match a given outcome Y to the relevant policy issue in the survey (e.g. donation to a gun-related organization and gun control). Table 2 provides more information on this procedure. Thanks to randomization, the distribution of outcome variables is the same across the three treatment conditions (see Appendix A.2). We regress the behavioral outcomes over the normalized survey responses (X) interacted with a categorical variable identifying the

---

[17] Shannon entropy for each of the 10 issues broken down by survey tool are provided in Appendix A.5 (see Dahl and Østerås (2010)). For all items, QVSR has higher entropy score relative to Likert. For 8 out of 10 survey items, QVSR also has higher entropy relative to Likert+.

treatment conditions (Likert is the reference category):

$$Y_i = \sigma_0 + \mu_1 D_{i,Likert+} + \mu_2 D_{i,QVSR} + \sigma_1 X_i + \sigma_2 X_i D_{i,Likert+} + \sigma_3 X_i D_{i,QVSR} + \sigma_4 J_4 + ... + \sigma_j J_j + \varepsilon_i \quad (1)$$

where $J_4, ... J_j$ are dummy variables that indicate membership in a block used for block randomization (see Appendix A.2 for more details). Regression coefficients σ_1, $\sigma_1 + \sigma_2$ and $\sigma_1 + \sigma_3$ can be interpreted as the difference between $E(Y/X = 1)$ and $E(Y/X = 0)$ for Likert, Likert+ and QVSR respectively. The better discriminating tool is the one with a larger difference between the two quantities of interest, i.e. with a larger regression coefficient.[18]

Note that, given differences in sample sizes between wave 1 and wave 2, the coefficients for the behavioral tasks recorded in wave 1 are more precisely estimated than the coefficients for the behavioral tasks recorded in wave 2.[19] Given the budget constraint, answers in QVSR are a linear combination of each other meaning that the error terms across equations are correlated. Consequently, we also estimate a seemingly unrelated regressions model (Zellner 1962), one for wave 1 outcomes and another for wave 2 outcomes. Results are reported in Appendix A.3 and align with results presented in Section 2.2.

To understand the intuition behind this estimation strategy, we turn to Figure 2, which plots average donations to the gun control charities by (normalized) survey response on the gun control question, further broken down by survey methodology. The lines captures the three regression coefficients mentioned above (see Figure 3 for the actual estimates). As shown on this figure, the regression slope is larger for QVSR than for Likert. Individuals who choose the end-of-scale response categories in Likert end up de-bunching under QVSR. This de-bunching aligns with their behavior on the donation task: people who donate less choose, on average, smaller values in QVSR than people who donate more. This is captured by the magnitude of the regression slope: individ-

---

[18] We checked for a non-parametric relationship and, for all methods, the standard errors do not rule out a monotonic relationship. In Appendix A.5, we plot, for each survey tool, the mean of the outcome of interest for all values of X and overlay a linear fit line based on the underlying data. For three outcomes, a quadratic function provides a better fit. As we discuss, this does not affect the interpretation of our results.

[19] To account for unobserved individual components that may be correlated across periods, we cluster at the participant level. Note that results remain the same whether or not we cluster at the participant level.

uals who do not donate are no longer pulling the regression slope down by 'sharing' the extreme response categories with people who care enough to donate. Using the regression coefficients as a summary statistics, we can also see that, in this case, the discrimination achieved with QVSR better aligns with preference intensity than that achieved with Likert+.

Figure 2: Estimation Strategy
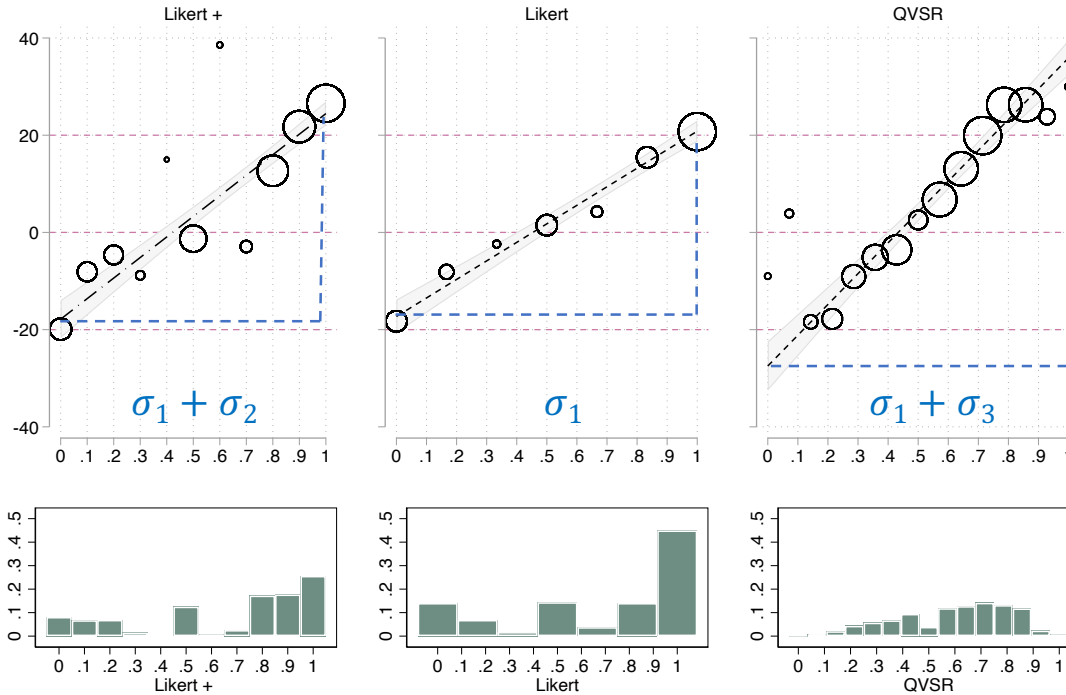
Y-axis: Donation to gun-related advocacy group. X-axis: answers (normalized) for the following item "Laws making it more difficult for people to buy a gun."
Likert, in the center of the figure provides the benchmark. A visual comparison indicates that the coefficient in Likert+ is only marginally larger than that for Likert. Notice the difference in slope between Likert and QVSR. The full estimates are available in Figure 3. Compare also the bunching in Likert and the variance gained under Likert+ and QVSR.

2.2 Results

Figure 3 plots regression coefficients obtained using equation 1. For the donation outcome, we use the normalized values of the response variables (X). When predicting the number of characters written, we use the normalized *absolute* values of the response variables (i.e. 0-3 for Likert, 0-5 for Likert+ and 0-7 for QVSR). Indeed, our outcome variable does not capture what was written about the bill (i.e. for or against), merely the overall effort spent writing about it. When predicting

punishment in the two dictator games, we use the normalized difference between responses on gun control and responses on the border wall. Higher positive values indicate that one favors gun control more intensely than one opposes the wall. Higher negative values indicated that one favors the wall more intensely than one opposes gun control. Given that the Independent recipient in the dictator game was opposed to gun control and opposed to the wall, we examine whether higher positive (negative) scores predict a higher likelihood of punishing (rewarding) the Independent.

The higher the regression coefficient in Figure 3, the better a given tool is at distinguishing between respondents with high and low preference intensity (as proxied by behavior). For wave 1 outcomes, QVSR outperforms Likert both substantively and statistically. As mentioned above, due to smaller sample sizes, results for the wave 2 tasks exhibit larger standard errors. Still, a comparison of regression coefficients suggests that QVSR has higher discrimination power than Likert: QVSR coefficients are more than twice the size of those found with Likert on 3 of the 4 outcomes.

In Section 1, we argued that asking people whether a policy issue is important to them might not be enough to distinguish strong from weak preferences. The results align with this conjecture: Likert+ also appears to carry more information on preference intensity than Likert, but its relative performance is far less consistent than that of QVSR. Indeed, Likert+'s discriminatory power is statistically indistinguishable from Likert's on all 6 outcomes. The coefficient for Likert+ is substantively larger for only 2 of the 6 outcomes, representing at best a 50% increase (against 100% or more for QVSR).

Figure 3: Regression Coefficients for Behavioral Outcomes
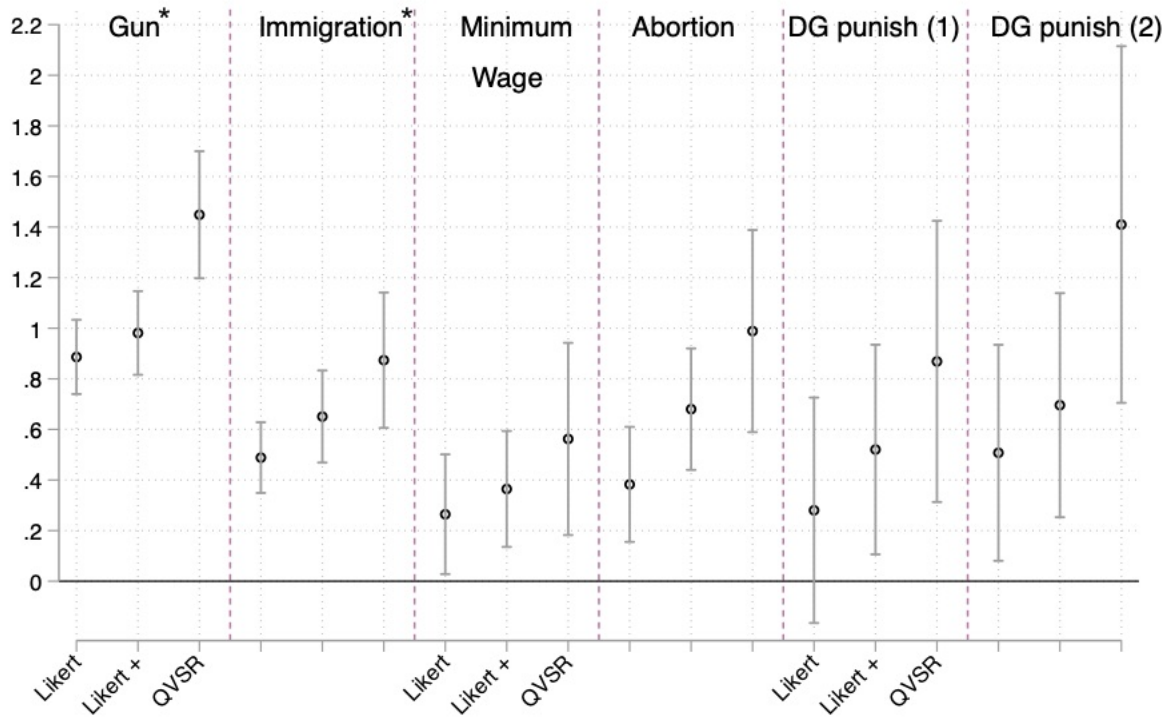
* Sample sizes for the Gun and Immigration tasks (wave 1) are double the size of the samples sizes for the other tasks (wave 2). As a result, effect sizes are more precisely estimated for these two tasks. For details on each task, see text.

Interpretation: a switch from the smallest response category (0) to the largest (1) is associated with a σ increase in Y. The increase is equal to σ times the standard deviation of Y.

For the letter writing tasks, the predictor is the normalized absolute value of the response variable. For the punishment task, the predictor is the normalized difference between the gun control and the border wall response variables. For details on each, see text.

3 Measuring Preference Intensity: So What?

Having provided evidence that votes in QVSR convey information on preference intensity that is not captured by Likert items, we now examine how important debates in political science might be affected by taking preference intensity —and how to best measure it— seriously. We revisit two lines of work, one focusing on the determinants of policy preferences and the other assessing the (lack of) congruence between mass preferences and policy outcomes. Based on Section 2 results, our preferred survey instrument for measuring individual-level differences in preference intensity is QVSR. For expository purposes, we also present results using Likert+.

3.1 The Determinants of Policy Preferences: The Role of Material Self-Interest

A common assumption in political economy is that individuals significantly affected by a policy are more likely to take action and mobilize around this policy (see for example Bouton et al. (2018)). In other words, preferences over policies that "directly affect (a voter's) rights, privileges, or lifestyle in some concrete manner" (Howe and Krosnick 2017: 328) are more intense than preferences over policies that have no direct personal implications. To the extent that QVSR better measures preference intensity, we can expect, if the latter view is correct, the correlation between personal exposure and policy preferences to be larger when measured using QVSR than when measured using Likert. By contrast, if self-interest is mostly irrelevant to attitude formation, improving the measure of preference intensity should only have minimal impact.

The assumption that material self-interest matters for attitude formation and change is routinely dismissed by public opinion scholars who, instead, emphasize non-economic modes of reasoning such as value-based or partisan-motivated reasoning (Sears and Funk 1990; Margalit 2013). This conclusion builds on horse-race studies designed to assess which, of material self-interest or partisan/value-based modes of reasoning explain the largest share of the observed variance.[20] In these studies, material self-interest's explanatory power is very limited.

While we did not design our study to settle this debate, we can use the data collected to ex-

---

[20] See Chong, Citrin and Conley (2001) and ? for exceptions.

amine how measurement strategies can affect conclusions regarding the importance of material self-interest relative to non-economic modes of reasoning. Indeed, as we argued in Section 1, the relative weight of partisan reasoning in the data generating process is itself a function of the measurement strategy used. Furthermore, as argued above, results might vary depending on how well a given measurement tool captures how much people care about a given issue.

For a visual overview of differences between measurement tools, we turn to Figure 4. It plots the relationship between gender on the one hand and subjective support for gender equality in the workplace on the other, measured using Likert, Likert+ and QVSR. Notice how, in Likert, there is very little variance in survey answers: most people appear to *strongly* support workplace gender equality. The additional variance gained by switching from Likert to Likert+ is informative of respondents' gender: women are more likely than men to be in the 'extremely important' response category. In QVSR, the de-bunching is even more consequential: there is a linear and monotonic relationship between the number of votes in QVSR and the percentage of women as a share of individuals who cast the same number of votes. In other words, Figure 4 provides preliminary evidence that conclusions about the importance of material self-interest might vary depending on whether one is using Likert or QVSR.

This de-bunching has implications when using preferences as a predictor. Indeed, with only limited variance to leverage, end-of-scale bunching in Likert (and Likert+) makes results more sensitive to underlying modeling assumptions. Compare the top panel, which assumes a linear relationship between preferences and gender, to the bottom panel, which assumes a quadratic relationship. Once the assumption of a linear relationship is relaxed, there is a significant drop in the predicted differences between the highest and lowest response categories in Likert and Likert+. In contrast, with QVSR, these differences remain mostly unchanged.

Next, we turn to other policy proposals asked in the survey and examine the relationship between exposure to the policy and survey responses. With two exceptions, measures of exposure were directly purchased from the survey firm, which has extensive information on respondents, meaning that respondents were not primed to think about their own situation. The two exceptions are information on whether one would benefit from a minimum wage increase and whether one would benefit from paid parental leave. Questions used to construct these two exposure variables were

Figure 4: Respondent's Gender and Response to Pay Equity Item
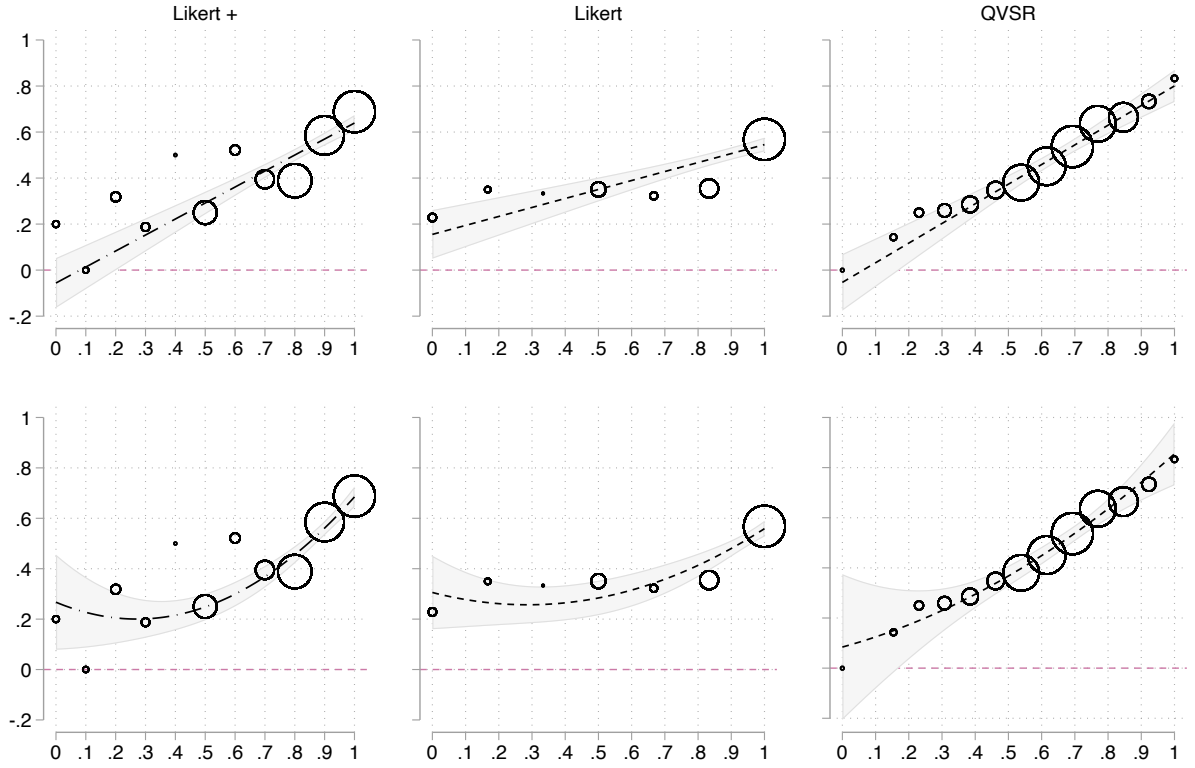
Y-axis: gender of respondent (female = 1), scatter plot represents the share of all respondents who are female conditional on having the same value on X, i.e. E(Y=1/X=x). X-axis: survey answers by survey method, normalized to vary from 0 to 1.
In the top panel, scatter plots are overlayed with a linear fit line. In the bottom panel, scatter plots are overlayed with a quadratic fit line.

asked at the beginning of the survey. Table 3 lists all exposure variables alongside the relevant policy issues.

Due to space constraints, we present results in the form of a difference in predicted probabilities. First, we model the relationship between a given exposure proxy (race, reliance on minimum wage, gun ownership, etc.) and the relevant policy preference using equation 1 modified by adding a logit link function.[21] Then we compute the predicted probability of exhibiting this exposure feature when the normalized preference variable is equal to 0 and when it is equal to 1. Figure 5 plots the difference between the two predicted values for all exposure proxies listed in Table 3.[22] Note that,

---

[21] We also estimate a seemingly unrelated regressions model to account for error correlation across equations. Results are reported in Appendix A.3 and align with results presented in Figure 5.

[22] For a detailed overview of the underlying data, as in Figure 4, see Appendix A.5

as suggested by Figure 4, the linear model used to compute these predicted values likely favors Likert and Likert+ by artificially inflating the difference between the lowest and highest preference values.

Table 3: Proxies of Personal Exposure and Relevant Policy Proposal

Policy Proposal	Exposure Proxy (Mean)
Giving same sex couples the legal right to adopt a child	Equal to 0 if heterosexual, 1 otherwise (0.07)
Laws making it more difficult for people to buy a gun	Equal to 1 if does not own gun or none in the house, 0 otherwise (0.63)
Building a wall on the US Border with Mexico	Equal to 1 if one or both parents born abroad, 0 otherwise (0.16)
Require employers to pay women and men the same amount for the same work?	Equal to 1 if female, 0 otherwise (0.51)
Preferential hiring and promotion of blacks to address past discrimination	Equal to 1 if black, $=0$ otherwise (0.13)
Require employers to offer paid leave to parents of new children	Equal to 1 if about to have a child, plans to have a child soon or already has a toddler, 0 otherwise (0.19)
Raising the minimum wage to 15\$/h over the next 3 years	Equal to 1 if would benefit from increase to 15\$/h, 0 otherwise (0.29)
A nationwide ban on abortion with only very limited exceptions	Equal to 1 if *not* a born-again, $=0$ otherwise (0.63)

Figure 5: Predictions for Proxies of Exposure
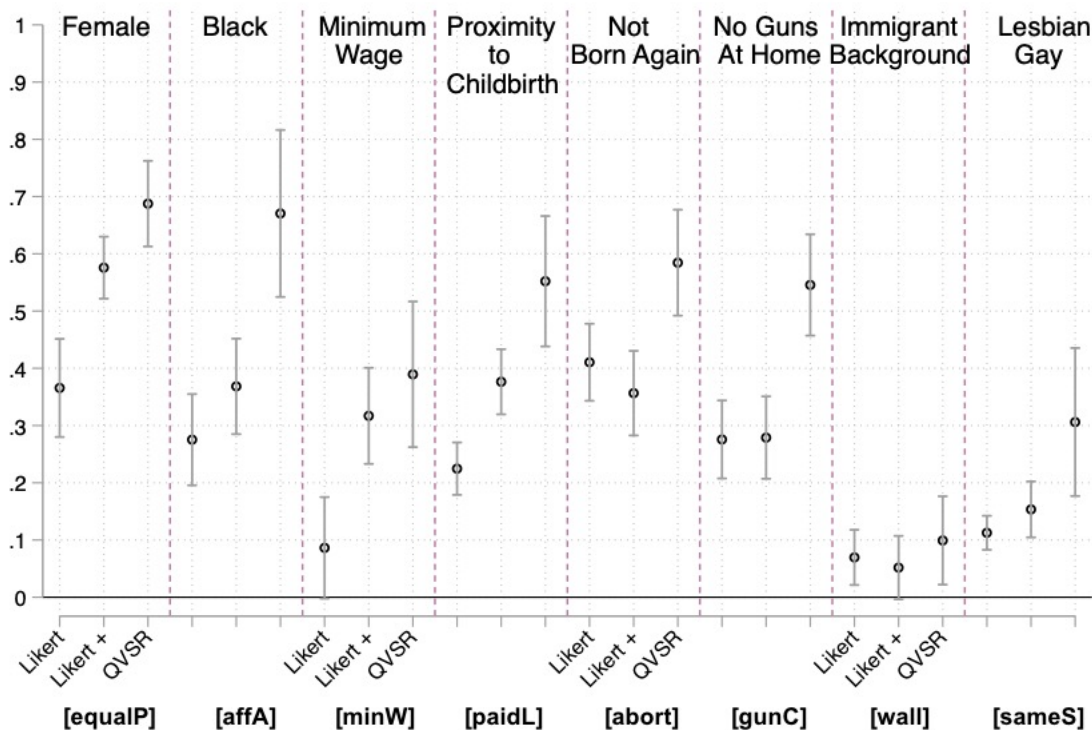
For details on each proxy, see Table 3. Note that predictors (i.e. survey responses/votes) are coded from conservative to liberal. See page 9 for wording. Outcomes are coded such that 1 (0) captures the likelihood that one will "benefit" from the liberal (conservative) policy position.

Interpretation: a switch from the smallest response category (0) to the largest (1) is associated with a σ increase (in % points) in the probability of exhibiting the mentioned feature (race, reliance on minimum wage, proximity to childbirth, etc.). Probabilities are computed using a logit regression (see text). OLS estimates are available in Appendix A.3.

As shown in Figure 5, with one exception (immigrant background), the difference in predicted values is higher for QVSR than it is for Likert. For 3 out of 8 exposure proxies, the difference in predicted values is higher for Likert+ than it is for Likert. Note that given Likert+'s wording, Likert+'s comparatively lackluster improvement is notable. Indeed, the item is worded to encourage respondents to think about how the policy might affect them *personally*. This suggest that wording is not enough to compensate for the effect of competing motives such as the partisan motive.

The differences across measurement techniques documented in Figure 5 should encourage researchers to think carefully about measurement strategies when designing studies aimed at unpacking how preferences form and change. For example, Kuziemko et al. (2015) ran a survey experiment to understand how support for redistribution is affected by information on inequality.

They find no effect of the information treatment on support for redistributive social policies, which were measured using a Likert item. A closer look at the treatment suggests that it was not merely informative but could be also interpreted as partisan: indeed it provided a very "left-wing" interpretation of inequality and its social costs. QVSR, because it penalized the partisan motive and because it better measures preference intensity (which might have been affected by the treatment) could potentially yield different results. Alternatively, a null result using QVSR would provide more solid grounds to conclude that information on inequality does not affect preference for redistribution (whether preference orientation or preference intensity).

3.2 Congruence Between Policy Outcomes and Mass Preferences

Next, we examine how reliance on different survey technologies can affect researchers' understanding of the demand side of policymaking in representative democracies. Focusing on the percentage of respondents who favor or oppose a given policy proposal, we find no evidence that QVSR returns a fundamentally different snapshot of the public's preference orientation on a given issue when compared to Likert or Likert+ (see Appendix A.4).

When assessing the relationship between mass preferences and policy outcomes, measuring the percentage of respondents who favor or oppose a policy is often not enough: researchers also need to account for asymmetry in preference intensity. For example, a common assumption with regards to gun control is that the minority of voters who oppose gun control have more intense preferences than the majority of voters who support it (e.g. Bouton et al. (2018)). Based on our data, if there is asymmetry in preference intensity, it runs in a direction opposite to the one commonly hypothesized. Indeed, with QVSR, respondents who favor gun control cast, on average, 0.6 more votes than people who oppose gun control. Overall, 43% of people in favor cast 3 votes or more, while only 13% of people opposed do the same.

Another issue is that of policy trade-offs: voters cannot "have it all" and some policies have to be prioritized at the expense of others. Put differently, a mismatch between majoritarian preferences and policy outcomes might be less concerning if the status quo roughly reflects voters' policy priorities. The example of paid parental leave is here informative. Employer-provided parental leave receives

Figure 6: Parental Leave versus Spending Cap: Republicans and Democrats
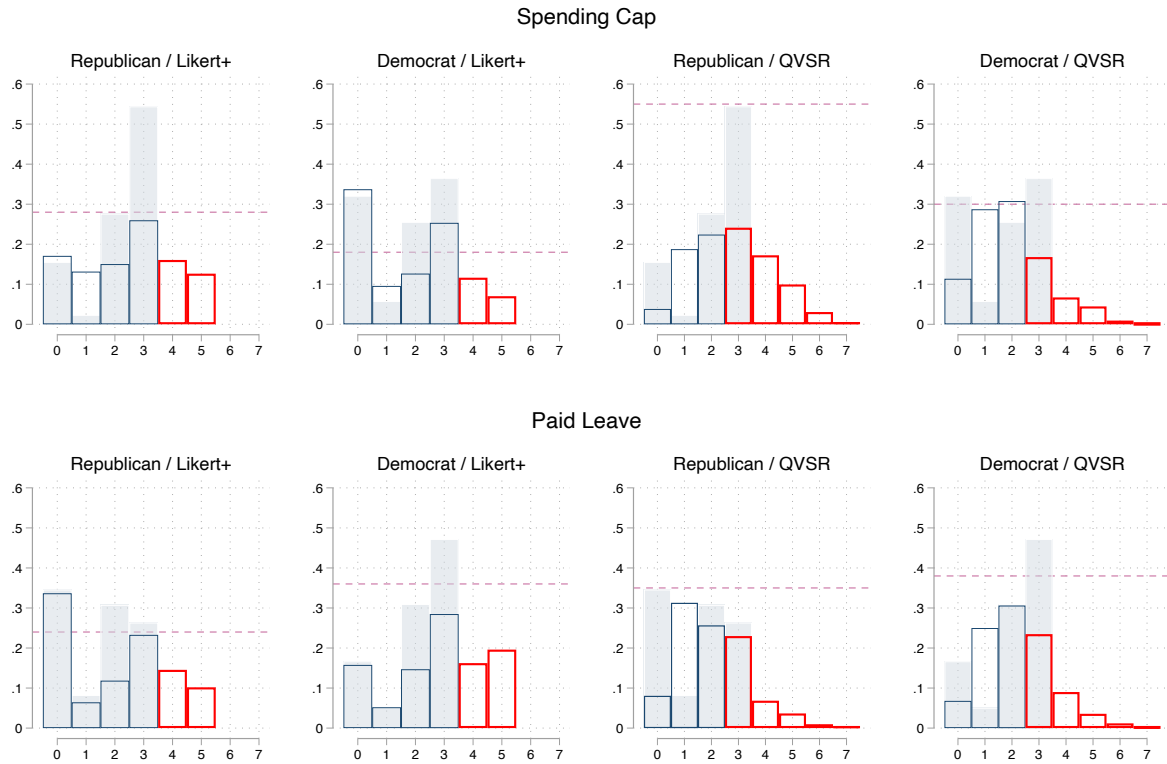
Answers in Likert+ and QVSR are overlayed over responses in Likert (the light-blue histogram). Response categories proxying for more intense preferences are plotted in red. These correspond to the 'very important' and 'extremely important' response categories in Likert+ and 3 votes or more in QVSR. The 3 votes threshold is used because it is the number of votes respondents would assign to all issues if they cared about them with similar intensity.
The horizontal dotted lines plot the share of respondents who give intense responses per the previous definition.

very broad support among voters of both parties. Given mass support, the federal government's inability to act on these issues suggests that powerful groups who represent interests affected by a change in the status quo (e.g. small businesses) are successfully lobbying legislators to prevent legislation favored by an overwhelming bipartisan majority. One way to get these groups on board is to shift part of the burden to tax payers in the form of payroll contributions and federal grants. Given voters' opposition to tax increases, the easy compromise is to use deficit spending. In other words, there is a trade-off between politically feasible expansions of parental leave and balanced budgets, something true for most social insurance programs. As discussed in the introduction and Section 1, this is where the concept of preference intensity —and tools that help measure it— come in handy: do people care more about paid leave or balance budgets?

Figure 6 plots survey answers for these two issues, broken down by survey tool and partisanship. We focus on the 8 out of 10 respondents who favor[23] both paid parental leave and a spending cap. As shown in the bottom panel, irrespective of the method used, preferences for paid parental leave appear of similar intensity in both parties. Indeed, in Likert+, roughly a third of Democrats and Republicans who do not oppose paid parental leave describe this issue as personally very or extremely important. In QVSR, a similar share casts 3 votes or more in favor of the proposal. This pattern, however, does not extend to preferences for a spending cap (top panel). In QVSR, 55% of Republican respondents cast 3 votes or more, while only 30% of Democrats do. Notice also how this asymmetry is only picked up by QVSR and not by Likert+. To summarize, based on QVSR, Republicans appear to care more about controlling federal spending than about paid parental leave. In contrast, Democrats do not appear to care more about one issue over the other. While tentative, this analysis suggests that, even for policy changes favored by voters from both parties, demand side factors can still contribute to the absence of bipartisan legislation.

Discussion and Next Steps

Despite its theoretical importance, there currently exists no routine way of measuring preference intensity, explaining why the concept is often overlooked in empirical research on public opinion. In Sections 1 and 2, we compared and contrasted two ways of measuring preference intensity: Likert+ and QVSR. We found that asking about issue importance does not convey much more information about preference intensity than simply relying on Likert items' distinction between 'favor (oppose) a great deal, moderately, or a little.' QVSR, instead of asking explicitly how important an issue is, forces respondents to trade-off across issues to measure how much people care. Our results indicate that QVSR's higher entropy is more than just noise: relative to Likert, it better discriminate between people who care more about an issue, and people who care less. Still, based on the argument presented in Section 1, it remains to be seen whether this conclusion extends to contexts in which signaling motives, especially the partisan motive, do no introduce a systematic bias.

Our arguments and results advance survey methodology in three ways. First, the model presented

---

[23] We include not only people who express support but also people who take the ambivalent position, i.e. neither/nor in Likert and Likert+ and 0 vote in QVSR.

in this paper (and in Appendix A.1) lays the foundations for evaluating how the abundance and bunching problems affect the data generation process and, ultimately, hypothesis testing. Second, to the best of our knowledge, this is the first study to use real-world behavior to evaluate the performance of different methods for measuring preference intensity. Third, and most importantly, with QVSR, we discuss one possible way to better measure preference intensity.

So what should an applied researcher take away from the results presented here? The main message is straightforward and likely unsurprising to most readers: time spent understanding how measurement strategies interfere with theory testing is usually time well spent. In Section 3, we revisited two lines of research, varying the measurement strategy to examine the latter's possible implications for the type of conclusions reached. We found that the relationship between exposure to a policy's consequences and policy preferences varies with the measurement strategy. This suggests that researchers studying the determinants of policy preferences need to think carefully about their measurement strategy: arguments emphasizing a signaling motive are likely penalized when tested using QVSR to measure preferences, while arguments emphasizing self-interest are likely penalized when tested using Likert-type items. With regards to comparing mass preferences and policy outcomes, we found noteworthy differences across measurement tools when considering the distribution of intensity. We provided an example of QVSR's benefits, focusing on the trade-off between social policy expansion and deficit spending. Note that these results are tentative: research designs better adapted to real world trade-offs are needed before any strong conclusions regarding best practices can be reached.

Additional methodological issues remain to be investigated. First, in the case of QVSR, the consequences of switching to other forms of pricing (e.g. linear instead of quadratic) have yet to be examined theoretically and empirically. Different forms of pricing have been studied in the context of voting (see Casella and Sanchez (Forthcoming)). Yet, results obtained in that literature cannot be automatically transferred to survey research: as emphasized in our model, individuals may have quite distinct motivations when voting and when answering surveys. Second, and most importantly, we have set aside the issue of interpersonal comparisons: to what extent can we assume that 3 votes for one individual (9 credits) is similar to 3 votes for another individuals? Note that this issue is a concern for most measurement tools. For example, in Likert+, not everyone imparts

the same meaning to the "extremely important" response category. Yet one might worry that, for QVSR, if the fixed budget is more constraining for some than for others, votes become difficult to compare across individuals. Based on our empirical results, whatever worsening of the ratio of noise-to-signal this issue generates, it is not enough to favor the status quo (i.e. Likert items) over QVSR. Furthermore, we found no obvious differences that issues of interpersonal comparisons affected the type of descriptive conclusions most common in studies of public opinion.

However, more should be done and we see at least two ways forward. One is to only use QVSR for research questions and hypotheses centered on what QVSR does best, i.e. measure how people navigate trade-offs on a small set of issues. For example, QVSR would be most helpful in countries in which a large part of the electorate is cross-pressured on two salient issues (i.e. voting according to one's policy preference on issue A implies voting against one's policy preference on issue B). Another way forward is to develop and test possible external numeraires, such as a small monetary lottery, with respondents buying with their voting credits the right for one or more lottery entries.

To facilitate follow-up studies, we have made available a web application enabling researchers to vary key features including pricing (e.g. linear versus quadratic) and the number of credits relative to the number of options. The resulting survey can be easily embedded into other online platforms, such as Qualtrics.[24] We hope this online tool will help generate additional evidence regarding the pros and cons of QVSR and, beyond QVSR, spur new research on theoretically-informed measurement strategies adapted to research goals.

---

[24] This web application can be found at `https://qvsr.io`

References

Abramson, Scott F, Korhan Koçak and Asya Magazinnik. 2019. What Do We Learn About Voter Preferences From Conjoint Experiments? Technical report Working paper. URL: https://scholar. princeton. edu/sites/default/files . . . .

Adida, Claire L, Adeline Lo and Melina R Platas. 2018. "Perspective taking can promote short-term inclusionary behavior toward Syrian refugees." *Proceedings of the National Academy of Sciences* 115(38):9521–9526.

Bouton, Laurent, Paola Conconi, Francisco Pino and Maurizio Zanardi. 2018. Guns, Environment, and Abortion: How Single-Minded Voters Shape Politicians' Decisions. Technical report CEPR Discussion Paper No. DP12801.

Bullock, John and Gabriel Lenz. 2019. "Partisan Bias in Surveys." *Annual Review of Political Science* 22:325–342.

Casella, Alessandra and Luis Sanchez. Forthcoming. "Democracy and Intensity of Preferences. A Test of Storable Votes and Quadratic Voting on Four California Propositions." *Journal of Politics* .

Chong, Dennis, Jack Citrin and Patricia Conley. 2001. "When Self-Interest Matters." *Political Psychology* 22(3):541–570.

Dahl, Fredrik A and Nina Østerås. 2010. "Quantifying information content in survey data by entropy." *Entropy* 12(2):161–163.

Gilens, Martin. 2012. *Affluence and influence: Economic inequality and political power in America*. Princeton University Press.

Hanretty, Chris, Benjamin Lauderdale and Nick Vivyan. 2020. "A Choice-Based Measure of Issue Importance in the Electorate." *American Journal of Political Science* 64(3):519–535.

Horiuchi, Yusaku, Daniel M Smith and Teppei Yamamoto. 2018. "Measuring voters' multidimensional policy preferences with conjoint analysis: Application to Japan's 2014 election." *Political Analysis* 26(2):190–209.

Howe, Lauren and Jon Krosnick. 2017. "Attitude Strength." *Annual Review of Psychology* 68:327–351.

Johnston, Christopher D, Howard G Lavine and Christopher M Federico. 2017. *Open versus closed:*

Personality, identity, and the politics of redistribution. Cambridge University Press.

Kuziemko, Ilyana, Michael Norton, Emmanuel Saez and Stefanie Stantcheva. 2015. "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments." *American Economic Review* 105(4):1478–1508.

Lalley, Steven and E Glen Weyl. 2018. "Quadratic Voting: How Mechanism Design Can Radicalize Democracy." *American Economic Association Papers and Proceedings* 108:33–37.

Leeper, Thomas J and Joshua Robison. 2018. "More Important, but for What Exactly? The Insignificant Role of Subjective Issue Importance in Vote Decisions." *Political Behavior* pp. 1–21.

Malhotra, Neil, Jon A Krosnick and Randall K Thomas. 2009. "Optimal Design of Branching Questions to Measure Bipolar Constructs." *Public Opinion Quarterly* 73(2):304–324.

Margalit, Yotam. 2013. "Explaining social policy preferences: Evidence from the Great Recession." *American Political Science Review* 107(01):80–103.

Mason, Lilliana. 2015. ""I Disrespectfully Agree":The Differential Effects of Partisan Sorting on Social and Issue Polarization." *American Journal of Political Science* 59(1):128–145.

Posner, Eric and E Glen Weyl. 2018. *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton: Princeton University Press.

Quarfoot, David, Douglas von Kohorn, Kevin Slavin, Rory Sutherland, David Goldstein and Ellen Konar. 2017. "Quadratic Voting in the Wild: Real People, Real Votes." *Public Choice* 172(1-2):283–303.

Revilla, Melanie, Willem Saris and Jon Krosnick. 2014. "Choosing the Number of Categories in Agree–Disagree Scales." *Sociological Methods & Research* 43(1):73–97.

Sears, D.O. and C.L. Funk. 1990. "The limited effect of economic self-interest on the political attitudes of the mass public." *Journal of Behavioral Economics* 19(3):247–271.

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

Zaller, John. 2012. "What Nature and Origins Leaves Out." *Critical Review* 24(4):569–642.

Zellner, Arnold. 1962. "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." *Journal of the American statistical Association* 57(298):348–368.

Who Cares? Measuring Preference Intensity in a Polarized Environment

Appendix

Contents

1 A model of Survey Answers

Likert scales are the standard tool to measure policy preferences. Nevertheless, there are (at least) two potential issues with standard Likert scales when one wants to measure policy preferences.

The first conceptual issue with standard Likert scales rests with what they actually try to measure: is the wording fit to capture preference intensity?

The second conceptual issue with Likert scales, and with any survey instrument more generally, is that respondents may not answer truthfully.

We propose a model that allows us to study both potential flaws, and to examine how alternative survey techniques - Likert+ and QVSR - may help address these issues.

1.1 Assumptions

We build on Cavaillé, Chen and Van Der Straeten (2019) to derive predictions about the relative performance of QVSR, Likert and Likert+ at measuring policy preferences.

Given the potential flaws of standard Likert scales that we want to address, we make two main sets of assumptions: one about policy preferences, and the second about motivations when answering survey.

Assumptions about policy preferences   Consider a number of proposed policy reforms, e.g. building a wall on the border between the U.S. and Mexico or legislating to give same sex couples the right to adopt a child.

We assume that an individual's policy preferences on each of these different reforms are characterized by two parameters: her *preference orientation*, and her *preference intensity*.

The *preference orientation* describes whether the individual agrees or not with the reform as described. It is a subjective evaluation of the "quality" of the reform, as considered from the individual's perspective. Formally, one's preference orientation on each issue $k = 1, ..., K$ can be described by a real number α_{ik} in the interval $[-1, 1]$, where 1 means perfect agreement with the reform, and -1 total disagreement with this reform. Intermediate values correspond to intermediate opinions, something for instance, that might be due to ambivalence (e.g. support for the policy principle but not its specific suggested implementation).

The *preference intensity* describes how important the issue at stake is to the individual. Formally, how much an individual cares about an issue, i.e. preference intensity, can be captured by a positive number β_{ik} in the interval $[0, 1]$, where $\beta_{ik} = 0$ means that this individual does not care about this issue, in the sense that she would not be ready to take costly action in favor or against any policy change on this issue, and $\beta_{ik} = 1$ means that the question is of the utmost importance to her.

Ideally, one would like to collect information about both parameters, since both are relevant to social scientists studying policy preferences and to policy makers.

Remark about standard Likert scales   With this framework in mind, we can ask what Likert scales actually try to measure. The wording of the Likert question in our survey is the following: "Do you favor, oppose, or neither favor nor oppose: [Example] Giving same sex couples the legal right to adopt a child?", and if favor (resp. oppose), "Do you favor (resp. oppose) that a great deal, moderately, or a little?".

Answers to Likert questions are ambiguous regarding the key parameters we are interested in measuring, that is, the preference orientation and the preference intensity as defined above. Indeed, the wording ("favor"/"oppose") seems to imply that the question should predominantly tap into the 'preference orientation' dimension. Yet, when an individual reports that she opposes giving same sex couples the legal right to adopt a child *a great deal*, it can be interpreted as meaning that she fully disagrees with the terms of this reform (preference orientation), but it may also mean that she has strong feelings about this issue (preference intensity). Formally, we will denote by $x_{ik} \in [-1, 1]$ the sincere answer that individual i would like to give

to this question, which, as discussed above, is likely to include information mostly about α_{ik} (the preference orientation) but also about β_{ik} (the preference intensity).

Assumptions about motivations when answering survey   We assume that an individual may have (at least) two (potentially) conflicting motives when answering the survey. On the one hand, she derives some intrinsic utility from answering each question sincerely. This might derive from some expressive benefits (I am happy to tell who I am, or what I stand for), or this might be induced by a psychological cost of lying. We call this motive the *"sincerity motive"*. This is the motive generally assumed in the literature using survey data.[1] On the other hand, we defend the view that she may also care about how her answers will be read and interpreted by other people, which might conflict with this sincerity motive. This additional motivation might encompass a variety of psychological mechanisms, depending on the context and the question. For example, imagine that the government is considering whether a specific reform should be adopted or not, and that a survey is conducted to measure public support for or opposition to this reform. The respondent might be willing to use her answers to the survey to influence policy making. Another motivation for the respondent might be to signal to herself, or to whoever is going to read the survey, that she has some socially desirable traits. For example, she may want to appear altruistic, non-racist, tolerant, etc. She might also want to signal a group identity. For example, if she is a Republican, and she expects Republicans to take specific positions on some issues, she may suffer a psychological cost from moving away from these typical "Republican positions". Whatever the source of this motivation, because of this *"signaling motive"*, one position is particularly attractive to the respondent, which might be different from where she really stands.[2] Note that the signaling motive is likely to be at play when a respondent is asked to report her opinion on an issue, but also when asked how important this issues is to her (under Likert+). For example, consider a Republican respondent asked how important the issue of the wall on the U.S.-Mexico border is personally to her. even if she thinks that this issue is only of secondary importance compared to more pressing economic issues, since this issue is high on the agenda of the Republican party, she may feel pressured, out of loyalty to her party, to report that the issue is of the utmost importance to her.

Under all three survey techniques, the individual is asked to answer the "Favor/oppose" question. Additionally, under Likert+, she is also asked to answer the "How important" question. Formally, we assume that, on each issue $k = 1, ..., K$, respondent i is characterized by four parameters:

- the answer she would give to the standard "Favor/oppose" question if she were to answer it sincerely, $x_{ik} \in [-1, 1]$,

- the answer to the "Favor/oppose" question that she finds the most attractive because of the signaling motive, denoted by $X_{ik} \in [-1, 1]$,

- her true preference intensity of this issue, $\beta_{ik} \in [0, 1]$,

- the answer to the "How important" question that she finds the most attractive because of the signaling motive, denoted by $B_{ik} \in [0, 1]$.

Under standard Likert and QVSR, we assume that the utility a respondent derives from answering a survey on the K issues, denoted by V_i, depends on her answers to the "Favor/oppose" questions, denoted by $\widehat{x}_i = (\widehat{x}_{i1}, ..., \widehat{x}_{iK}) \in [-1, +1]^K$, in the following way:

$$V_i(\widehat{x}_i) = -\sum_k w_{ik} \left[(1 - z_{ik}) (\widehat{x}_{ik} - x_{ik})^2 + z_{ik} (\widehat{x}_{ik} - X_{ik})^2 \right], \tag{A1}$$

with $w_{ik} > 0$ and $z_{ik} \in [0, 1]$.[3]

---

[1] One assumption commonly made by social scientists using survey data is that they provide a faithful - if noisy - measure of respondents views (See for example Achen (1975) or Ansolabehere, Rodden and Snyder (2008).

[2] Our model of signaling shares some similarities with Benabou et al. (2020), who investigate how much of a person's deep moral preferences can be retrieved from observing their choices, for instance via experiments, when (social or self) image motives are likely to inflate the extent to which agents behave pro-socially. Relatedly, Bullock et al. (2015) study systematic differences between Republican and Democrat voters in how they answer factual questions about economic facts.

[3] We use these simple quadratic forms to derive some simple closed-form solutions. See Cavaillé, Chen and Van Der Straeten (2019) for more general specifications.

Parameter z_{ik} is the relative weight of the signaling motive compared to the sincerity motive for this question. Parameters w_{ik} is meant to capture the psychological disutility that a respondent faces when she gives answers to a survey that do not exactly reflect her views. This parameter is especially important because, with QVSR, respondents are constrained in their answers by the number of credits they get. This can prevent them from expressing as extreme opinions as they might have liked to. As we will see, under a reasonable assumption regarding the relation between β_{ik} and w_{ik}, we can expect respondents to prioritize issues they care the most about.

Under Likert+, the respondent is additionally asked to report how important each issue is to her personally. We then assume that in this case, the utility a respondent derives from answering the survey, denoted by W_i, depends on both her answers to the "Favor/oppose" question, denoted as previously by $\widehat{x}_i = (\widehat{x}_{i1}, ..., \widehat{x}_{iK}) \in [-1,1]^K$, and her answers to the "How important" question, denoted by $\widehat{\beta}_i = (\widehat{\beta}_{i1}, ..., \widehat{\beta}_{iK}) \in [0,1]^K$. We make the following assumption in that case:

$$
\begin{aligned}
W_i(\widehat{x}_i, \widehat{\beta}_i) \;=\; & -\sum_k w_{ik}\left[(1 - z_{ik})\left(\widehat{x}_{ik} - x_{ik}\right)^2 + z_{ik}\left(\widehat{x}_{ik} - X_{ik}\right)^2\right] \\
& -\sum_k w'_{ik}\left[\left(1 - z'_{ik}\right)\left(\widehat{\beta}_{ik} - \beta_{ik}\right)^2 + z'_{ik}\left(\widehat{\beta}_{ik} - B_{ik}\right)^2\right],
\end{aligned}
$$

with $w'_{ik} > 0$ and $z'_{ik} \in [0,1]$. Parameter w'_{ik} is the psychological disutility that a respondent faces when she gives answers that do not exactly reflect her views when answering the "How important" questions, and parameter z'_{ik} is the relative weight of the signaling motive compared to the sincerity motive for this question.

"Survey technology"   The survey technology specifies the set of questions that are asked, and the set of answers that are admissible, that is, the set of answers the respondents can choose from. For example, under standard Likert scales, a respondent can pick any answer on a pre-determined scale (e.g. "oppose a great deal", "oppose moderately", "neither oppose nor support", "support moderately", "support a great deal"). Under QVSR, there is a maximum number of credits that the respondent can use to answer the "Favor/oppose" question, and the marginal cost of moving away from the neutral answer (here 0) increases linearly with the distance to this neutral answer.

Optimization problem   Individuals are assumed to choose answers ($\widehat{x}_i$ in all cases, and additionally $\widehat{\beta}_i$ in the case of Likert+) that maximize the utility function V_i (W_i in the case of Likert+), subject to the constraints on answers imposed by the survey technology.

Equipped with this model, we can predict how respondents answer survey depending on the survey tool. In particular, our interest will be in discussing whether the reported views are a good measure of the underlying parameters of interest: preference orientation (α_{ik}) and preference intensity (β_{ik}). In the next sections, we derive answers under Likert, Likert+ and QVSR respectively, and discuss the relative performance of each survey tool at measuring policy preferences.

1.2 Optimal responses under standard Likert scales

Under standard Likert technology, the individual can freely pick any answer she wishes to on all issues. She solves the following optimization program:

$$
\max_{\widehat{x}_i \in [-1,1]^K} V_i(\widehat{x}_i) = -\sum_k w_{ik}\left[(1 - z_{ik})\left(\widehat{x}_{ik} - x_{ik}\right)^2 + z_{ik}\left(\widehat{x}_{ik} - X_{ik}\right)^2\right].
$$

It is easy to check that the solution of the optimization program for issue k, denoted by $\widehat{x}^L_{ik}$, is:

$$
\widehat{x}^L_{ik} = (1 - z_{ik})\, x_{ik} + z_{ik} X_{ik}. \tag{A2}
$$

If $z_{ik} = 0$ (only the sincerity motive is active), the individual has no incentive to misreport her view, and $\widehat{x}_{ik}^{L} = x_{ik}$. But as soon as $z_{ik} > 0$, the individual has the incentive to move away from her true opinion in the direction of the partisan target. In particular, in a situation of intense political polarization, one might expect the $|X_{ik}|$ to be quite large, which will induce respondents to inflate their support or opposition to policy reforms (depending whether their preferred party endorse or oppose them). This might result in massive bunching for issues with a strong and clear partisan divide.

Second, note that how much the individual values her response to this question compared to other questions in the survey (parameter w_{ik}) does not influence her answers. Indeed, each question is treated in isolation.

Regarding the key parameters we are interested in measuring, that is, the preference orientation (α_{ik}) and the preference intensity (β_{ik}), answers under standard Likert scales are a mix of their sincere response x_{ik} - which itself includes mostly information about α_{ik} (the preference orientation) and also presumably some limited amount of information about β_{ik} (the preference intensity) - and their signaling target.

1.3 Optimal responses under Likert+

Under Likert+, the individual answers both the "Favor/oppose" question and the "How important" question. Question by question, the individual solves the trade-off between the sincerity motive and the signaling motive, resulting in the following responses, denoted $\widehat{x}_{ik}^{L+}$ and $\widehat{\beta}_{ik}^{L+}$:

$$\widehat{x}_{ik}^{L+} = (1 - z_{ik})\, x_{ik} + z_{ik} X_{ik} = \widehat{x}_{ik}^{L} \tag{A3}$$

$$\widehat{\beta}_{ik}^{L+} = (1 - z'_{ik})\, \beta_{ik} + z'_{ik} B_{ik} \tag{A4}$$

Answers to the "Favor/oppose" question are the same as with standard Likert scales. As with the "Favor/oppose" question, the partisan motive may induce individuals to misreport the importance of issues. In particular, in a highly polarized political landscape, some respondents may have the incentive to inflate the importance of issues that are high on the agenda of their preferred party, even if they feel only moderately concerned about them.

Comparison between Likert+ and standard Likert scales   Compared to standard Likert scales, the advantage of Likert+ is obvious: it aims at collecting two pieces of information instead of one: x_{ik} (as in standard Likert) but also β_{ik} (the preference intensity). Yet, as highlighted above, note that this direct measure of the preference intensity is likely to be only imperfect. In a highly polarized environment, respondents may have the incentive to inflate the importance of the issues that are high on their preferred party's agenda.

1.4 Optimal responses under QVSR

Under QVSR, there is a maximum number of credits (m) that the individual is allowed to spend on answers. Formally, assume that the set of feasible answers under QVSR is:

$$\left\{ \widehat{x}_i = (\widehat{x}_{i1}, ..., \widehat{x}_{iK}) \in [-1,1]^K : \sum_k \widehat{x}_{ik}^2 \leq m \right\},$$

with $0 < m < K$ (the individual cannot pick the most extreme answers of all issues). Deriving the optimal answers under QVSR is more complicated since it involves solving a constrained maximization program. The individual solves the following optimization program:

$$\max_{\widehat{x}_i \in [-1,1]^K} \mathcal{L}(\widehat{x}_i, \lambda_i) = -\sum_k w_{ik} \left[(1 - z_{ik})\, (\widehat{x}_{ik} - x_{ik})^2 + z_{ik}\, (\widehat{x}_{ik} - X_{ik})^2 \right]$$

$$+ \lambda_i \left[m - \sum_k \widehat{x}_{ik}^2 \right],$$

where λ_i is the Lagrange multiplier. One may check that first order conditions with respect to $\widehat{x}_{ik}$ yield:

$$\widehat{x}_{ik}^{QVSR} = \frac{w_{ik}(1-\beta_{ik})}{w_{ik}+\lambda_i}x_{ik} + \frac{w_{ik}(1-\beta_{ik})}{w_{ik}+\lambda_i}X_{ik} = \frac{w_{ik}}{w_{ik}+\lambda_i}\widehat{x}_{ik}^{L}.$$

If $\sum_k \left(\widehat{x}_{ik}^L\right)^2 \leq m$: responses are the same as under Likert (the budget constraint is not binding).

If $\sum_k \left(\widehat{x}_{ik}^L\right)^2 > m$, then satisfying the budget constraint implies that:

$$\sum_k \left(\frac{w_{ik}}{w_{ik}+\lambda_i}\widehat{x}_{ik}^L\right)^2 = m. \tag{A5}$$

Note that the left-hand side of the equality is strictly decreasing in λ_i, taking the value $\sum_k \left(\widehat{x}_{ik}^L\right)^2$ strictly higher than m when $\lambda_i = 0$, and converging towards 0 as λ_i goes to $+\infty$. Therefore, there exists a unique positive λ_i such that equality (A5) is satisfied.

Denoting the Lagrange multiplier at the optimum by λ_i^*, the optimal response on issue k under QVSR is therefore:

$$\begin{aligned}
\widehat{x}_{ik}^{QVSR} &= \frac{1}{1+\frac{\lambda_i^*}{w_{ik}}}\widehat{x}_{ik}^L \tag{A6}\\
&= \frac{1}{1+\frac{\lambda_i^*}{w_{ik}}}\left[(1-z_{ik})x_{ik} + z_{ik}X_{ik}\right].
\end{aligned}$$

where λ_i^* is the Lagrange multiplier at the optimum. As soon as the budget constraint is binding, compared to Likert, QVSR 'shrinks' all answers towards the neutral answer (0). Expression (A6) shows that this 'contraction' is likely to be heterogenous across issues: more credits will be given to issues with an higher w_{ik}, (see the utility function when answering surveys (A1)).

Comparison between QVSR and standard Likert scales Let us first compare the performance of QVSR and standard Likert scales. Note that both standard Likert and QVSR only ask the "Favor/oppose" question. The relative performance of QVSR vs. Likert at measuring the relevant preference parameters - to wit, the preference orientation, α_{ik}, and the preference intensity, β_{ik} - will depend on the statistical relationship between w_{ik} (how much the individual suffers when she has to deviate from reporting her unconstrained answers on issue k), β_{ik} (the intrinsic importance of the issue), and z_{ik} (the relative importance of the partisan motive compared to the sincerity motive).

Like Likert, answers under QVSR are still expected to be a mix of the respondent's preference orientation and her preference intensity. One first advantage of QVSR compared to standard Likert is that answers under QVSR also incorporates information about w_{ik} (how much the individual suffers when she has to deviate from reporting her unconstrained answers on issue k). If w_{ik} correlates with β_{ik} (the intrinsic importance of the issue) then, compared to standard Likert, answers under QVSR, are more informative of β_{ik}.

There are reasons to expect such a positive correlation between parameters w_{ik} and the importance of the issue (parameters β_{ik}). Indeed, if a respondent cares strongly about an issue, she might be expected to try her best to report an opinion that correctly reflects her views. But the two do not have to be perfectly identical. For example, a respondent may feel that a survey question is quite poorly framed, and only very imperfectly captures what is really at stake. So that even if she really cares about this issue, getting her answer right to the question might not be that important to her.

Similarly, there are reasons to expect a positive correlation between parameters w_{ik} and the relative weight of the sincerity motive (parameters $1-z_{ik}$). If on some issue, a respondent's opinion is mostly driven by the sincerity motive and intrinsic considerations (careful thinking and deliberation, personal interest) as opposed to external cues (follow the party line), one might expect that she will be more committed to adequately transmit her view when surveyed about this issue. But here again, the two parameters do not have to coincide.

If there is a positive correlation between w_{ik} (how much the individual suffers when she has to deviate from reporting her unconstrained answers on issue k) and the relative importance of the sincerity motive compared to the signaling motive $(1 - z_{ik})$, then QVSR has a second advantage over the other survey instruments. Not only does QVSR better captures the preference intensity, but it also better measures x_{ik} (the "sincere" answer to the standard Likert question). To understand why, remember that when a signaling motive is present, answers under Likert only imperfectly captures x_{ik}, since answers under Likert (the $\widehat{x}_{ik}^{L}$) are typically a weighted average between the sincere answer (the x_{ik} parameter) and some signaling target (see (A2)). The lower the strength of the sincerity motive compared to the signaling/partisan motive, the further away the Likert answers will be from the sincere answers. Now, remember that under QVSR, the individual relatively allocates more credits to issues for which w_{ik} is large (see (A6)). If one assumes that there is a positive correlation between w_{ik} and the relative strength of the sincerity motive, the model then predicts that under QVSR, the individual will allocate more credits to issues on which the unconstrained answers ($\widehat{x}_{ik}^{L}$) are closer to the sincere answer (x_{ik}), thus better capturing the true x_{ik} parameter.

Comparison between QVSR and Likert+ Compared to the Likert+ instrument, QVSR relies on only one question per issue instead of two, which might be seen as an advantage or a default depending on how much one values parsimony. Under Likert+, one gets $\widehat{x}_{ik}^{L}$ (same response to the "Favor/oppose" question as under standard Likert, see (A3)), plus an imperfect report about preference intensity $\widehat{\beta}_{ik}^{L+}$ (see (A4)). Under QVSR, one gets $\widehat{x}_{ik}^{QVSR}$, which by (A2) incorporates information about both $\widehat{x}_{ik}^{L}$ and w_{ik}.

The comparison between Likert+ and QVSR regarding which instrument best measures policy preferences is ambiguous. Indeed, one may think at first sight that with two questions instead of one, Likert+ collects more information than QVSR.

The model makes us qualify this result. When one explicitly takes into account the fact that respondents' answers to surveys are likely to be only partially sincere, the model reveals a potential advantage of QVSR compared to Likert+. Indeed, one may suspect that in a polarized environment, answers to the "How important?" question are likely to be only very noisy measures of the true intensity parameter. In such a context, an advantage of QVSR is to bring a "*revealed preference approach*" to the measurement of preference intensity. Instead of asking explicitly and openly how important an issue is, QVSR uses the fact that it forces trade-offs across issues to indirectly measure their importance. In a highly polarized environment where partisan identities are very salient, answers to the Likert+ "How important" question might not be that informative, and the revealed preference approach of QVSR -where individual actually have to *choose* across issues rather than just *reporting* whether an issue is important or not - might be superior. To conclude, the comparison between Likert+ and QVSR in terms of ability to recover meaningful information about preference intensity is ambiguous, and which performs best is an empirical question.

2 Survey Design

2.1 Overview

To recruit participants, we relied on the GFK/Ipsos KnowledgePanel. It is the oldest and largest probability-based online panel in the U.S.—with about 60,000 members. Panelists take on average two to three KnowledgePanel surveys a month, minimizing respondent fatigue and attrition. Panel participants are rewarded through the provision of free internet and a tablet to access it. Participants' consent was obtained on the first page of the survey. On this page, we provided information on the topic of the survey, the length and potential benefits from participating (entering a $100 lottery). We clearly stated that the survey was anonymous.

To allocate respondents across the three survey tools, we used a randomized-block design. We first formed 27 blocks on the basis of partisan identity (Republican, Independent, Democrat), subjective ideology (liberal, middle of the road, conservative), and vote in 2016 (Clinton-other, Trump, did not vote/too young to vote). These variables are important predictors of individuals' policy positions on politicized issues such as immigration, gay rights or budget deficits, as well as predictors of partisan identity and partisan strength. Within each block, we implemented a complete randomization. Balance tables are included below.

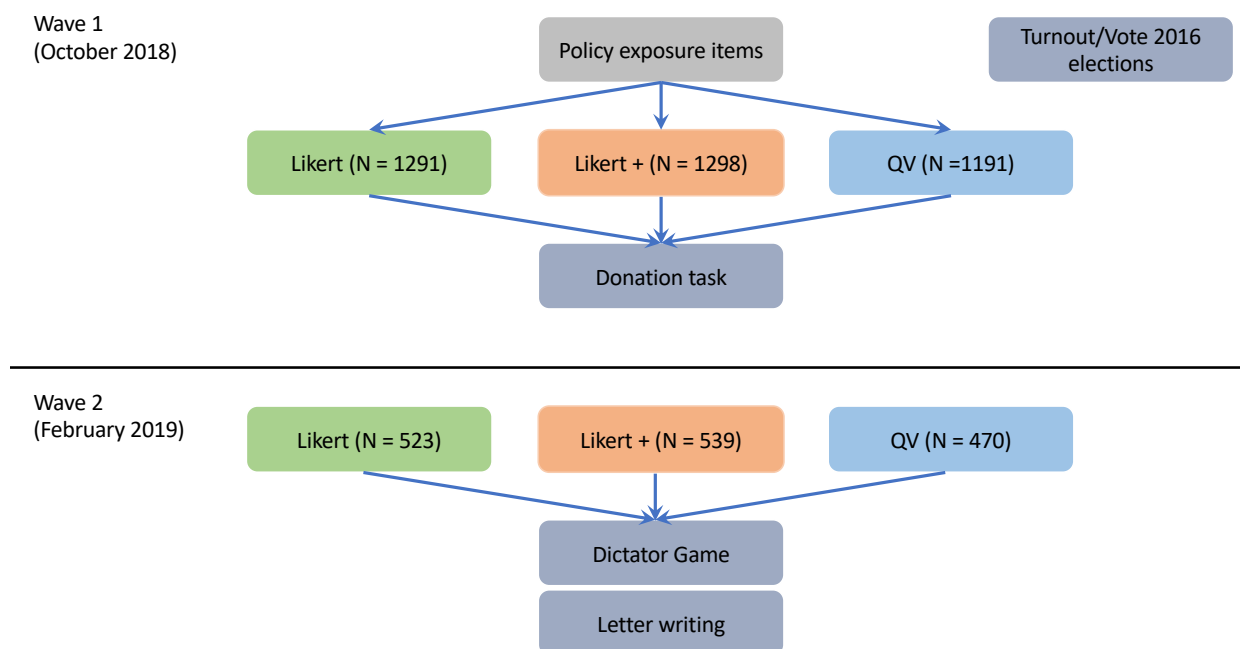**Wave 1 (October 2018)**

Policy exposure items

Turnout/Vote 2016 elections

Likert (N = 1291)　　Likert + (N = 1298)　　QV (N =1191)

Donation task

Wave 2 (February 2019)

Likert (N = 523)　　Likert + (N = 539)　　QV (N = 470)

Dictator Game

Letter writing

Table A1: Likert Treatment vs. Likert+/QVSR (pooled)

	Likert + / QVSR (pooled)	Likert	Δ
Donation (gun)	10.449	8.701	-1.748
	(33.794)	(32.869)	(1.185)
Donation (immigration)	1.216	2.014	0.798
	(29.160)	(27.847)	(1.016)
Party identity (1-7)	4.178	4.111	-0.067
	(2.219)	(2.203)	(0.079)
Ideology (1-7)	4.105	4.113	0.008
	(1.623)	(1.618)	(0.058)
HS or less	0.282	0.310	0.028*
	(0.450)	(0.463)	(0.016)
Some college	0.318	0.288	-0.030*
	(0.466)	(0.453)	(0.016)
BA or more	0.400	0.402	0.002
	(0.490)	(0.491)	(0.017)
Age	52.288	52.420	0.132
	(16.068)	(16.606)	(0.577)
Gender	1.509	1.491	-0.018
	(0.500)	(0.500)	(0.018)
White	0.756	0.758	0.002
	(0.430)	(0.429)	(0.015)
Black	0.092	0.079	-0.013
	(0.288)	(0.269)	(0.010)
Other	0.066	0.070	0.004
	(0.248)	(0.256)	(0.009)
Hispanic	0.087	0.093	0.006
	(0.282)	(0.290)	(0.010)
No exposure to minimum wage increase	0.365	0.353	-0.012
	(0.482)	(0.478)	(0.017)
Exposure to minimum wage increase	0.259	0.263	0.003
	(0.438)	(0.440)	(0.016)
Missing value on minW exposure variable	0.375	0.384	0.009
	(0.484)	(0.487)	(0.017)
Paid leave exposure score	1.579	1.567	-0.011
	(0.793)	(0.798)	(0.028)
Gun control exposure score	3.364	3.324	-0.040
	(0.870)	(0.886)	(0.031)
Immigrant parents	1.236	1.247	0.011
	(0.594)	(0.605)	(0.021)
Gay/lesbian	0.067	0.070	0.003
	(0.250)	(0.255)	(0.009)
Born again (BA)	0.635	0.609	-0.026
	(0.482)	(0.488)	(0.021)
Missing value on BA	0.297	0.317	0.020
	(0.457)	(0.465)	(0.016)
Observations	2,325	1,216	3,541

Table A2: Likert+ Treatment vs. Likert/QVSR (pooled)

	Likert / QVSR (pooled)	Likert +	Δ
Donation (gun)	9.572	10.381	0.810
	(32.640)	(35.055)	(1.186)
Donation (immigration)	1.313	1.829	0.516
	(28.184)	(29.713)	(1.017)
Party identity (1-7)	4.141	4.182	0.041
	(2.210)	(2.222)	(0.079)
Ideology (1-7)	4.118	4.088	-0.029
	(1.603)	(1.655)	(0.058)
HS or less	0.296	0.283	-0.013
	(0.457)	(0.451)	(0.016)
Some college	0.303	0.316	0.013
	(0.460)	(0.465)	(0.016)
BA or more	0.401	0.400	-0.001
	(0.490)	(0.490)	(0.017)
Age	52.474	52.066	-0.408
	(16.161)	(16.428)	(0.577)
Gender	1.496	1.515	0.019
	(0.500)	(0.500)	(0.018)
White	0.754	0.761	0.008
	(0.431)	(0.426)	(0.015)
Black	0.086	0.089	0.003
	(0.281)	(0.285)	(0.010)
Other	0.067	0.069	0.001
	(0.250)	(0.253)	(0.009)
Hispanic	0.093	0.081	-0.012
	(0.291)	(0.273)	(0.010)
No exposure to minimum wage increase	0.364	0.355	-0.010
	(0.481)	(0.479)	(0.017)
Exposure to minimum wage increase	0.258	0.266	0.008
	(0.438)	(0.442)	(0.016)
Missing value on minW exposure variable	0.378	0.380	0.002
	(0.485)	(0.485)	(0.017)
Paid leave exposure score	1.573	1.579	0.006
	(0.794)	(0.796)	(0.028)
Gun control exposure score	3.335	3.380	0.044
	(0.882)	(0.863)	(0.031)
Immigrant parents	1.239	1.241	0.002
	(0.597)	(0.599)	(0.021)
Gay/lesbian	0.071	0.061	-0.010
	(0.258)	(0.240)	(0.009)
Born again (BA)	0.617	0.644	0.027
	(0.486)	(0.479)	(0.021)
Missing value on BA	0.306	0.298	-0.008
	(0.461)	(0.458)	(0.016)
Observations	2,327	1,214	3,541

Table A3: QVSR vs. Likert/Likert+ (pooled)

	Likert / Likert + (pooled)	QVSR	Δ
Donation (gun)	9.541	10.524	0.983
	(33.982)	(32.376)	(1.213)
Donation (immigration)	1.922	0.546	-1.375
	(28.789)	(28.541)	(1.040)
Party identity (1-7)	4.147	4.174	0.027
	(2.212)	(2.218)	(0.080)
Ideology (1-7)	4.101	4.123	0.022
	(1.636)	(1.588)	(0.059)
HS or less	0.297	0.281	-0.016
	(0.457)	(0.449)	(0.017)
Some college	0.302	0.319	0.017
	(0.459)	(0.466)	(0.017)
BA or more	0.401	0.400	-0.001
	(0.490)	(0.490)	(0.018)
Age	52.243	52.532	0.289
	(16.514)	(15.669)	(0.590)
Gender	1.503	1.501	-0.002
	(0.500)	(0.500)	(0.018)
White	0.760	0.749	-0.010
	(0.427)	(0.434)	(0.016)
Black	0.084	0.094	0.010
	(0.277)	(0.292)	(0.010)
Other	0.070	0.063	-0.006
	(0.254)	(0.244)	(0.009)
Hispanic	0.087	0.093	0.006
	(0.282)	(0.291)	(0.010)
No exposure to minimum wage increase	0.354	0.376	0.022
	(0.478)	(0.485)	(0.017)
Exposure to minimum wage increase	0.264	0.253	-0.011
	(0.441)	(0.435)	(0.016)
Missing value on minW exposure variable	0.382	0.371	-0.011
	(0.486)	(0.483)	(0.018)
Paid leave exposure score	1.573	1.578	0.005
	(0.797)	(0.790)	(0.029)
Gun control exposure score	3.352	3.348	-0.005
	(0.875)	(0.877)	(0.032)
Immigrant parents	1.244	1.230	-0.014
	(0.602)	(0.588)	(0.022)
Gay/lesbian	0.065	0.073	0.008
	(0.247)	(0.261)	(0.009)
Born again	0.627	0.625	-0.001
	(0.484)	(0.484)	(0.021)
Missing value on BA	0.307	0.295	-0.012
	(0.462)	(0.456)	(0.017)
Observations	2,430	1,111	3,541

2.2 Within-Wave Attrition

Table A4 examines whether observable covariates help predict who, in the QVSR treatment group, is most likely to drop out once asked to watch the explanatory video. As highlighted using the † symbol, individuals who are not in the labor market (i.e. missing value on "minimum wage exposure score (MWES)") are less likely to drop out in wave 1. Yet, as indicated by the size of the coefficients, the implications for covariate imbalance are substantively small.

We run the same analysis for wave 2 and also examine if preferences measured in wave 1 help predict who does not finish the survey (still conditional on being assigned to QVSR). Individuals with more intense preference on the minimum wage item (see †† symbol) in wave 1 are more likely to drop out in wave 2 (conditional on having consented to participate in wave 2). Specifically, moving from 0 to 4 (90th percentile) on this value, increases the drop out probability from 6 to 16%. Note that this imbalance biases results against QVSR (since we are losing individuals with intense preferences).

Table A4: Predicting Within-Wave Dropout (QVSR only)

(OLS)	Wave 1 b/se	Wave 2 b/se
Partisanship	0.00	-0.00
	(0.00)	(0.01)
Ideology	0.00	0.02
	(0.00)	(0.01)
Some college [ref:HS only]	0.01	0.04
	(0.01)	(0.04)
BA or more	0.02	0.03
	(0.01)	(0.04)
Age	0.00	0.00*
	(0.00)	(0.00)
Gender	0.00	-0.03
	(0.01)	(0.03)
Black [ref: white]	0.03	0.03
	(0.02)	(0.06)
Other	0.02	-0.03
	(0.02)	(0.06)
Hispanic	0.01	0.03
	(0.02)	(0.06)
Minimum wage exposure score (MWES)	0.00	-0.02
	(0.01)	(0.04)
Missing value on MWES†	-0.05***	0.02
	(0.01)	(0.04)
Paid leave exposure score	0.00	0.03
	(0.01)	(0.03)
Gun control exposure score	0.01	0.03
	(0.01)	(0.02)
Immigrant parents	-0.01	-0.04
	(0.01)	(0.03)
Gay/lesbian	-0.00	0.07
	(0.02)	(0.06)
Born again (BA)	0.01	0.02
	(0.01)	(0.03)
Missing value on BA variable	-0.03*	0.07
	(0.01)	(0.04)
Difference between attitudes on gun control and border wall		0.01
		(0.00)
Intensity of attitude toward minimum wage ††		0.02*
		(0.01)
Intensity of attitudes toward abortion ban		-0.00
		(0.01)
_cons	-0.03	-0.30
	(0.05)	(0.16)
N	1213	494

† See text
†† See text

2.3 Participation in Wave 2

In total, 870 (Likert), 873 (Likert+) and 791 (QVSR) individuals were eligible to participate in wave 2.[4] In Tables A5 and A6, we examine whether observable factors predict participation in wave 2. The outcome is equal to 1 if individuals consented to participate in wave 2, 0 otherwise. Note that individuals who, while having consented, did not finish the survey are also coded as 1.[5] We also examined whether behavior and policy preferences measured in wave 1, and relevant to the analysis in wave 2, predict the likelihood of participating in wave 2. Note that, ultimately, the sample size for this second attrition analysis is smaller because of missing values on the donation task, with roughly 195 refusing to participate in the lottery.

The † symbol identifies covariates that predict wave 2 participation but with only substantively small implications for imbalances across methods. Note that the "Other" ethnic category is a residual category that is difficult to interpret. More concerning are predictors marked with the †† symbol. As shown in Table A5 that, in the QVSR treatment condition, Democrats are less likely to participate in wave 2. This might be especially concerning for the dictator game, which was set up with cross-pressured democrats in mind. Yet, looking at Table A6, there is no evidence that this pattern has implications with regards to one's attitudes (and behavior) towards gun control and the border wall, the issue at the center of the dictator game. There is also evidence that people with more intense preferences on minimum wage increase are less likely to participate in wave 2 if they have been assigned to the QVSR treatment. The converse applies to preferences toward an abortion ban (more intense, means more likely to participate). While the former biases results against QVSR, the latter biases results in its favor. Absent a systematic bias in favor or against QVSR, we find no reason to revisit our main conclusions.

---

[4] The lower number of people eligible in QVSR is the result of higher attrition in wave 1 due to the video: with fewer people who completed the survey, we have fewer people to recontact.

[5] See section 2.2 for an analysis of within-wave attrition focused on QVSR, where within-wave attrition is the highest.

Table A5: Predicting Participation in Wave 2

	Likert b/se	Likert+ b/se	QVSR b/se
Republican [ref: Strong Rep][††]	-0.02	0.15*	-0.10
	(0.06)	(0.06)	(0.07)
Lean Republican	0.01	0.10	-0.11
	(0.06)	(0.06)	(0.06)
independent	-0.06	0.02	0.04
	(0.15)	(0.13)	(0.13)
lean Democrat[††]	0.01	0.06	-0.18**
	(0.07)	(0.07)	(0.07)
Democrat[††]	0.02	0.12	-0.17*
	(0.07)	(0.07)	(0.08)
Strong Democrat	-0.05	0.03	-0.15
	(0.07)	(0.07)	(0.08)
Ideology	-0.02	-0.01	-0.02
	(0.02)	(0.01)	(0.02)
Some college [ref: HS or less]	0.04	-0.01	0.05
	(0.04)	(0.04)	(0.05)
BA or more[†]	0.04	0.04	0.09*
	(0.04)	(0.04)	(0.05)
Age[†]	0.00	0.00	0.00*
	(0.00)	(0.00)	(0.00)
Gender[††]	-0.05	-0.13***	-0.02
	(0.03)	(0.03)	(0.04)
Black [ref: White][††]	-0.20**	-0.14*	0.01
	(0.07)	(0.06)	(0.07)
Other[†]	-0.16*	-0.12	-0.04
	(0.07)	(0.07)	(0.07)
Hispanic	-0.06	0.05	-0.07
	(0.07)	(0.07)	(0.07)
Minimum wage exposure score (MWES)	-0.02	0.00	-0.02
	(0.05)	(0.04)	(0.05)
Missing value on MWES	-0.01	0.07	0.00
	(0.04)	(0.04)	(0.05)
Paid leave exposure score	-0.01	0.01	-0.00
	(0.03)	(0.03)	(0.03)
Gun control exposure score [†]	0.05**	0.01	-0.01
	(0.02)	(0.02)	(0.02)
Immigrant parents	0.00	0.03	-0.02
	(0.03)	(0.03)	(0.03)
Gay/lesbian	-0.05	-0.01	0.06
	(0.07)	(0.07)	(0.07)
Born again	0.05	-0.03	0.02
	(0.04)	(0.04)	(0.04)
Missing value on BA variable l	-0.04	-0.02	-0.05
	(0.05)	(0.05)	(0.05)
_cons	0.53**	0.59***	0.68***
	(0.18)	(0.18)	(0.19)
N	870	873	791

† See text
†† See text

Table A6: Predicting Participation in Wave 2 Using Wave 1 Donations

	Likert b/se	Likert+ b/se	QVSR b/se
Donation wave 1 (gun)	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)
Donation wave 1 (immigration)	0.00	0.00	-0.00
	(0.00)	(0.00)	(0.00)
Difference between attitudes on gun control and border wall	-0.01	0.00	-0.00
	(0.01)	(0.00)	(0.01)
Intensity of attitude toward minimum wage[††]	-0.02	-0.01	-0.03*
	(0.02)	(0.01)	(0.01)
Intensity of attitude toward abortion ban[††]	0.01	0.02	0.02*
	(0.02)	(0.01)	(0.01)
_cons	0.61***	0.60***	0.64***
	(0.05)	(0.04)	(0.05)
N	805	819	758

†† See text

3 Alternative Analyses

We present results varying the data used, coding decisions and modeling assumptions.

3.1 Using Wave 2 Preferences

Figure A1 is the same as Figure 3 in the main paper with one difference: we use wave 2 preferences instead of those measured in wave 1.

3.2 Using OLS Instead of Logit

Figure A2 is the same as Figure 5 in the main paper with one difference: we use OLS instead of a logit.

3.3 Alternative Approach to Likert+

Figure A3 is the same as Figure 3 in the main paper with one difference: we have combined the Likert item and the issue importance item by adding them such that Likert+ runs from -8 to 8. As one can see, performance of Likert+ with the alternative approach is either the same or worse.

Figure A1: Revisiting Figure 3 Using Preferences Measured in Wave 2

A14
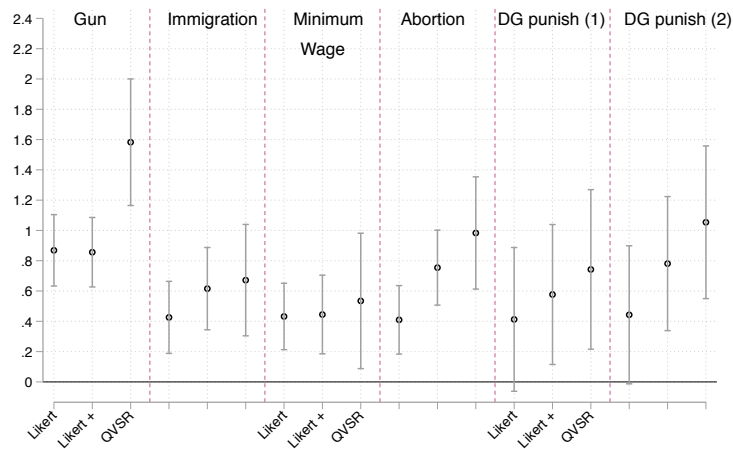
Figure A2: Revisiting Figure 5 Using OLS Instead of a Logit Link Function
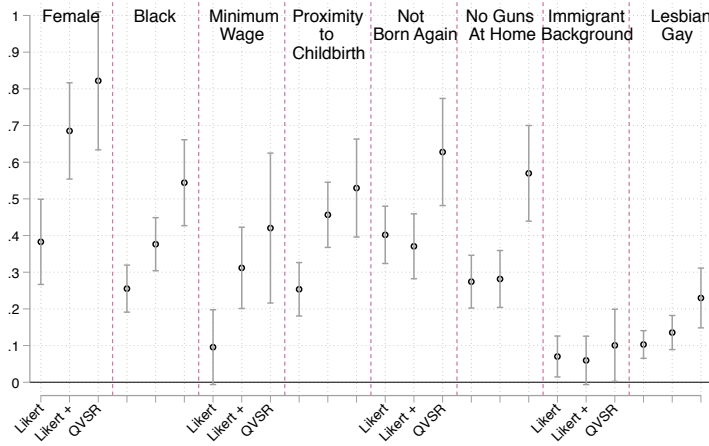
Figure A3: Revisiting Figure 3 Using a Different Coding for Likert +
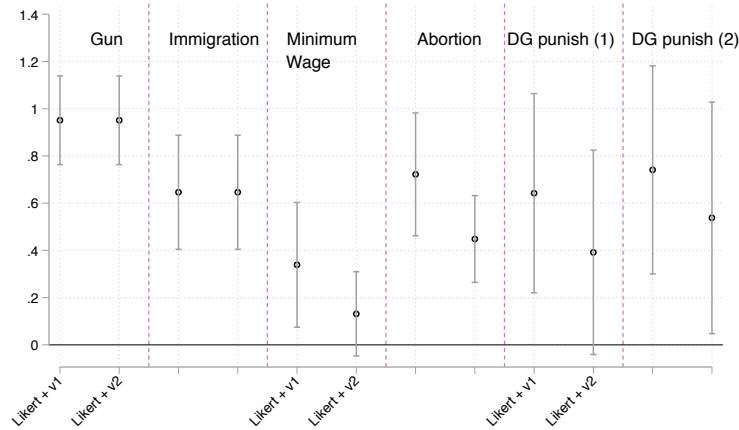
3.4 Seemingly Unrelated Regression

Because of QVSR's budget constraint, for individuals who use all their credits, votes on one issue is a linear combination of votes on other issues. As a result, the error terms across outcome-specific (or covariate-specific) equations are likely correlated. We consequently re-run the analyses underpinning Figures 3 and 5 and estimate seemingly unrelated regressions models that account for this correlation (Zellner 1962).

Table A7 summarizes results for behavioral outcomes (Figure 3), focusing first on wave 1 outcomes (top 2 rows) then on wave 2 outcomes (remaining rows). The table reports the interaction between the preference variable and a dummy variable identifying the survey methods used. Each column focuses on one pairwise comparison. The bottom row reports an F-test of the null-hypothesis that the sum of the coefficients is equal to 0. In line with results presented in the main manuscript, the coefficient for QVSR is always substantively and statistically larger than the one for Likert. This does not hold when comparing Likert and Likert+.

Table A8 reports results focusing on proxies of exposure to a policy's material consequences (Figure 5). Note that, in this case, we do not include the minimum wage dummy as this would result in dropping a third of respondents. In line with results presented in the main manuscript, the coefficient for QVSR is always substantively and statistically larger than the one for Likert. One exception is the immigrant background dummy, which —irrespective of the method— does not correlate with attitudes toward the border wall. For 3

Table A7: Coefficient Size Difference: Seemingly Unrelated Models for Behavioral Outcomes

	Wave 1 Outcomes			Wave 2 Outcomes		
	QVSR (=1) vs. Likert (=0) b/se	Likert+ vs. Likert b/se	QVSR vs. Likert+ b/se	QVSR vs. Likert b/se	Likert+ vs. Likert b/se	QVSR vs. Likert+ b/se
Gun	0.55*** (0.14)	0.10 (0.10)	0.43** (0.15)			
Immigration control	0.45*** (0.13)	0.19 (0.10)	0.25 (0.14)			
Minimum wage				0.06 (0.25)	0.06 (0.20)	0.13 (0.29)
Abortion				0.63** (0.22)	0.34 (0.18)	0.35 (0.23)
DG punish (2)				1.05** (0.38)	0.30 (0.32)	0.66 (0.37)
N	2321	2424	2327	935	1003	958
F-test	29.41	4.51	12.17	11.92	2.90	4.65

We replicate Figure 3 analysis using seemingly unrelated models. This table reports the interaction between the preference variable and a dummy variable identifying the survey methods used. For example, for the gun donation outcome, the difference between the coefficient for Likert and that for QVSR is equal to 0.55. We report results using version 2 of *DG punish* (column 6 in Figure 3). Results do not change when using version 1 instead (column 5 in Figure 3).

of the 7 exposure proxies, the Likert+ coefficient is larger than that for Likert. For 4 of the 7 exposure proxies, the QVSR coefficient is larger than that for Likert+. Irrespective of the exposure proxy, the Likert+ coefficient is never larger than that for QVSR.

Table A8: Coefficient Size Difference: Seemingly Unrelated Models for Exposure Proxies

	QVSR (=1) vs. Likert (=0) b/se	Likert+ vs. Likert b/se	QVSR vs. Likert+ b/se
Female	0.67***	0.31**	0.30*
	(0.14)	(0.11)	(0.14)
Black	0.24**	0.13*	0.10
	(0.07)	(0.05)	(0.08)
Proximity to Childbirth	0.23**	0.22**	0.03
	(0.09)	(0.07)	(0.09)
Not Born Again	0.20*	-0.05	0.27**
	(0.08)	(0.06)	(0.09)
No Guns at Home	0.31***	0.01	0.29**
	(0.09)	(0.07)	(0.09)
Immigrant Background	-0.04	-0.01	-0.02
	(0.05)	(0.03)	(0.05)
Lesbian/Gay	0.12**	0.01	0.12**
	(0.05)	(0.03)	(0.04)
N	1503	1594	1537
F-test	56.50	13.18	21.73

We replicate Figure 5 analysis using seemingly unrelated models. This table reports the interaction between the preference variable and a dummy variable identifying the survey methods used. For example, for gender, the difference between the coefficient for Likert and that for QVSR is equal to 0.67. Note that to maximize sample size and comparability with Figure 5, we do not include the minimum wage dummy (this would require dropping a third of observations).

4 Preference Orientation

Figure A4 plots the share of respondents who signal, through their response or vote, a liberal or conservative preference orientation on each of the 10 issues (based on whether they favor or oppose the reform). By design, there are not differences when orientation is measured using Likert or Likert +. QVSR returns roughly a similar picture, with one main exception: adoption by same sex couples. In Likert and Likert +, 24% of respondents chose the 'neither' response option. In QVSR, less than 4% did not vote on the issue. Instead, the majority ended up casting one vote in favor of same sex adoption (the liberal preference orientation).

Figure A4: Preference Orientation by Method
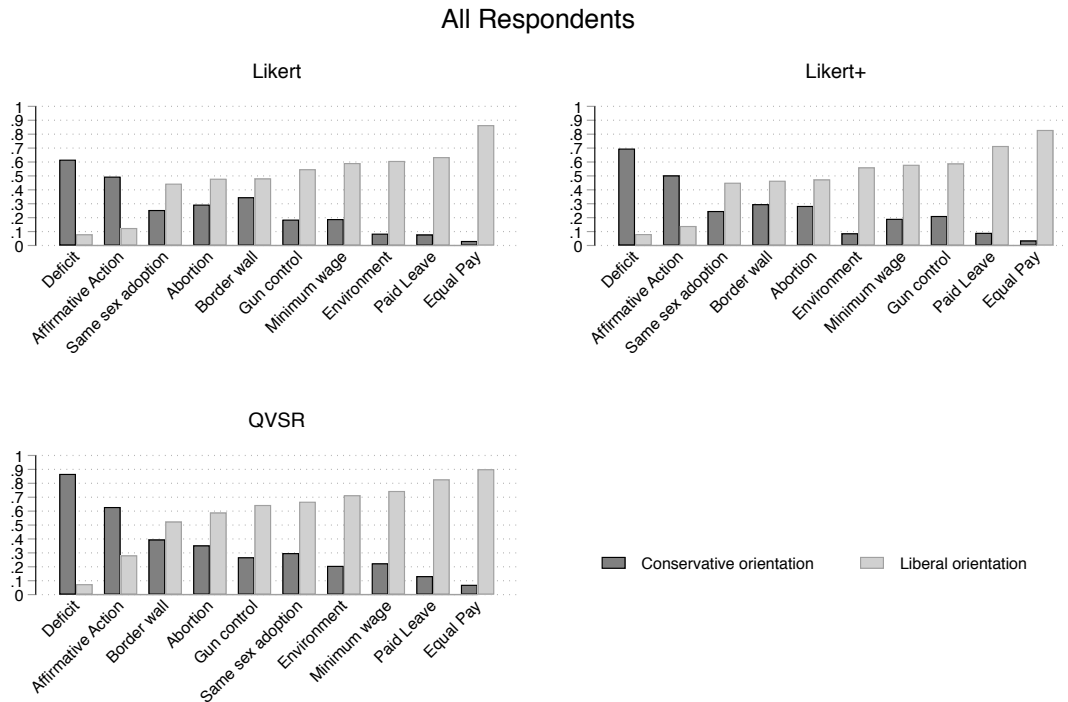
All Respondents

Note: Adding conservative and liberal shares might not add to 100% because of the 'neither" response option (with Likert and Likert+) or the "do not vote" option (i.e. spend 0 credit on this issue) with QVSR.

Figure A5 presents the same information as Figure A4, by respondents' partisanship. Based on Figure A5, we can distinguish between three types of issues. Some issues, such as equal pay and paid parental leave or the deficit cap receive very broad support among both parties. Other issues are highly partisan, such as the border wall. A third type of issue include those for which one party is quasi-unanimously in favor or against, while the other party is divided, i.e. abortion, minimum wage, same sex adoption, gun control and affirmative action.

Figure A5: Preference Orientation by Method and Partisanship
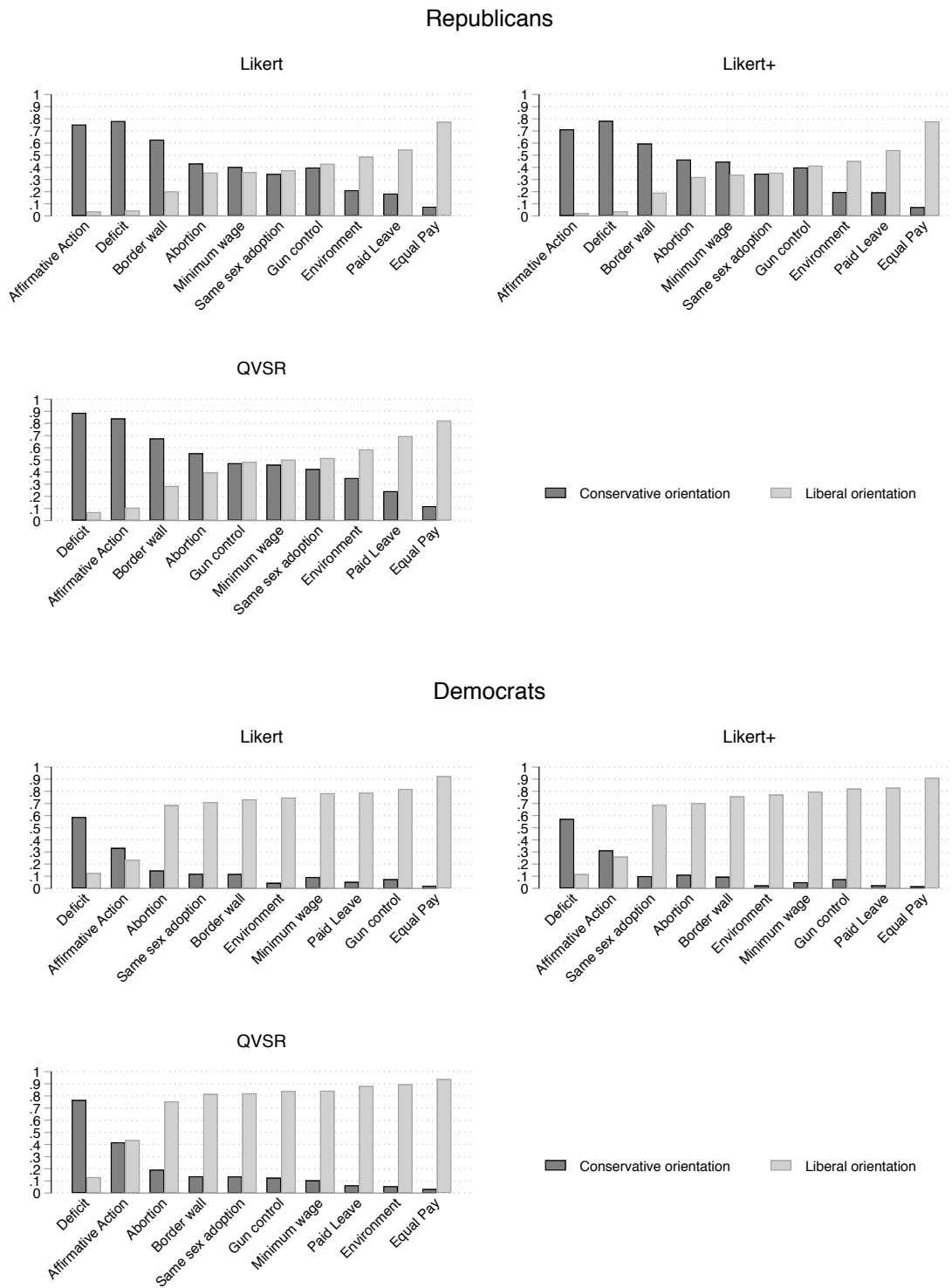
Republicans

Likert

Likert+

QVSR

■ Conservative orientation   ■ Liberal orientation

Democrats

Likert

Likert+

QVSR

■ Conservative orientation   ■ Liberal orientation

Note: Adding conservative and liberal shares might not add to 100% because of the 'neither" response option (with Likert and Likert+) or the "do not vote" option (i.e. spend 0 credit on this issue) with QVSR.

5 Descriptive Statistics and Figures

5.1 Shannon Entropy

Table A9: Entropy Scores

	Likert	Likert +	QVSR
Same sex right to adopt	1.62	2.16	2.24
Make it difficult to buy gun	1.57	2.00	2.40
Wall on the US Border	1.57	2.18	2.52
Paid leave	1.59	2.04	2.03
Preferential hiring of blacks	1.58	1.94	2.27
Pay women and men the same	1.00	1.71	2.02
Minimum wage to $15 an hour	1.67	2.14	2.24
Ban on abortion	1.55	2.13	2.48
Cap on federal spending	1.47	1.98	2.02
Regulation for environment	1.61	1.83	2.10

5.2 Scatter Plots and Histograms

In the top panel, the scatter plots represent the share of all respondents for whom Y = 1, conditional on having the same value on X. The X-axis represents survey answers by survey method, normalized to vary from 0 to 1. Each line plots the relationship between Y and X derived from a linear regression. The bottom panel shows the distribution of normalized survey answers (by survey method).

Gun-related Donations and Support for Gun Control
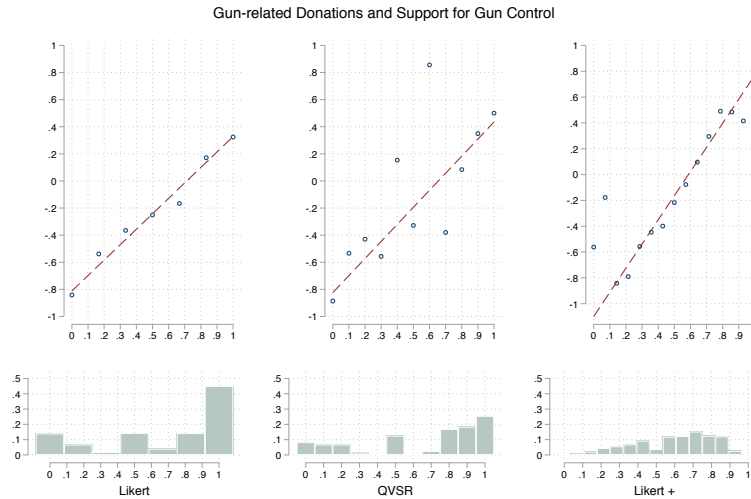
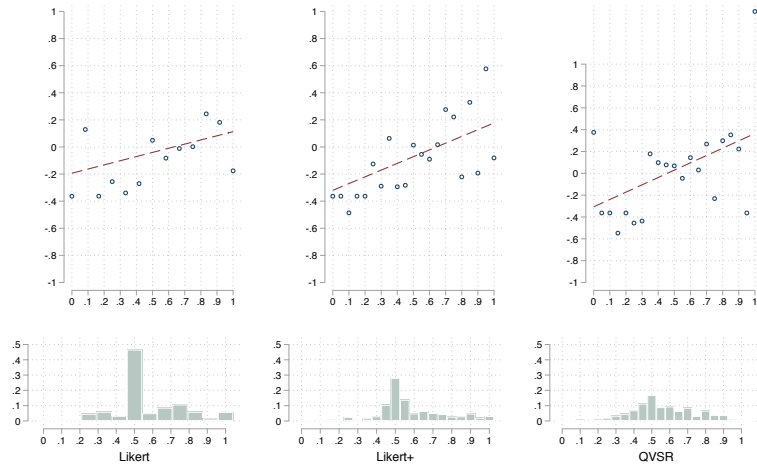Immigration-related Donations and Opposition to Borden Wall
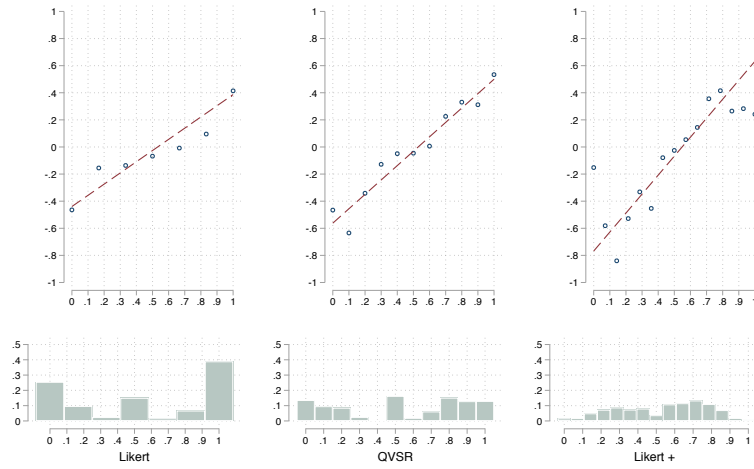
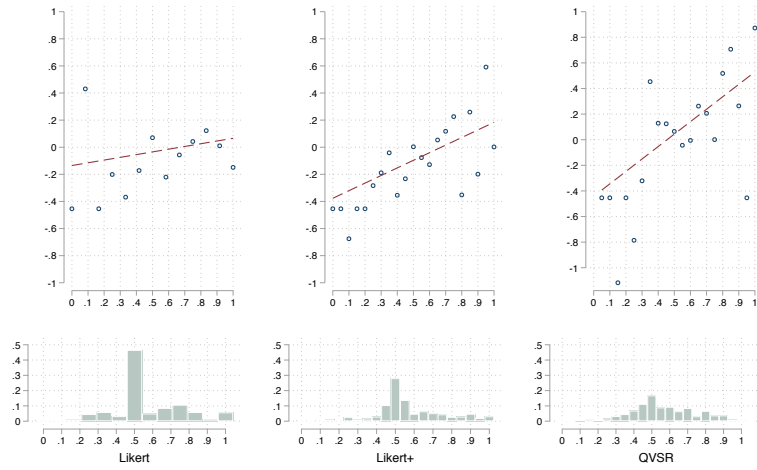Length of Writing: Minimum Wage Bill and Support for Min Wage
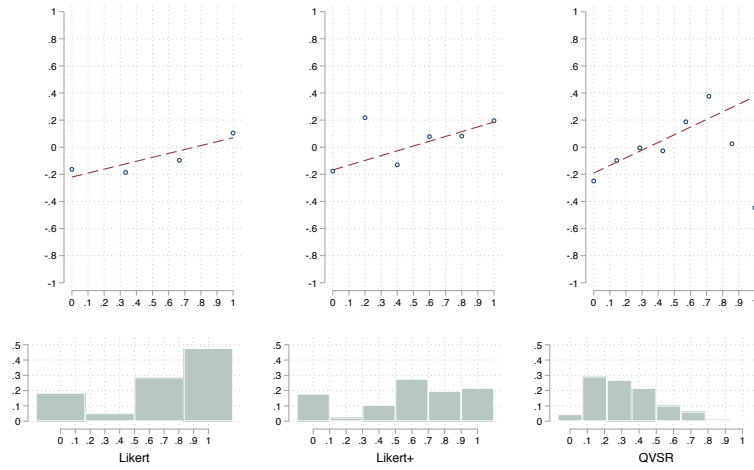
Length of Writing: Abortion Bill and Opposition to Abortion Ban
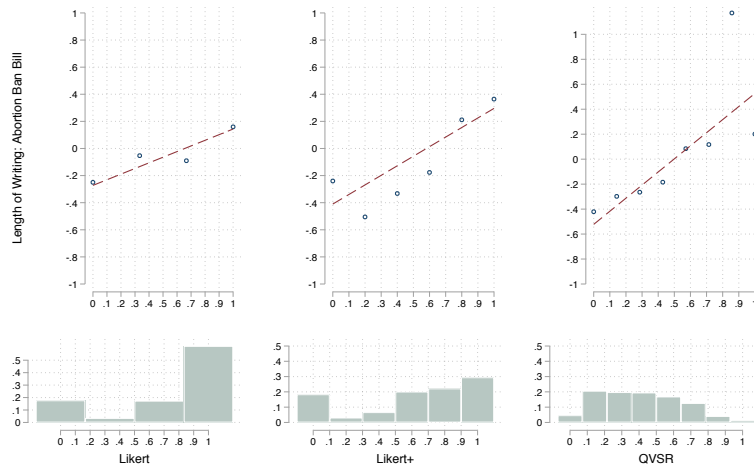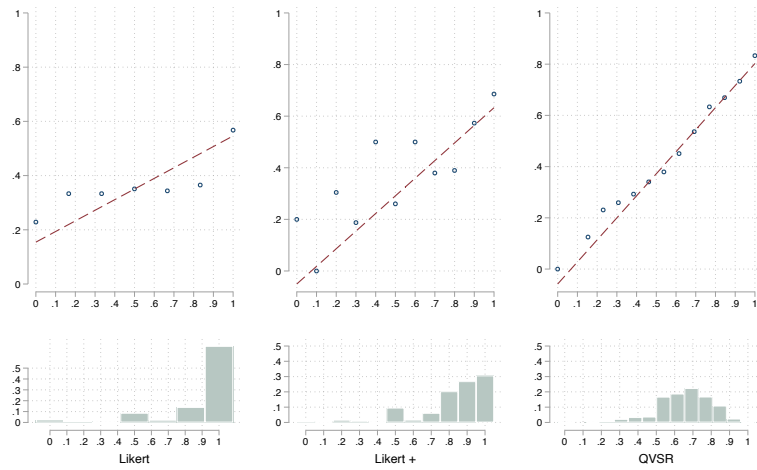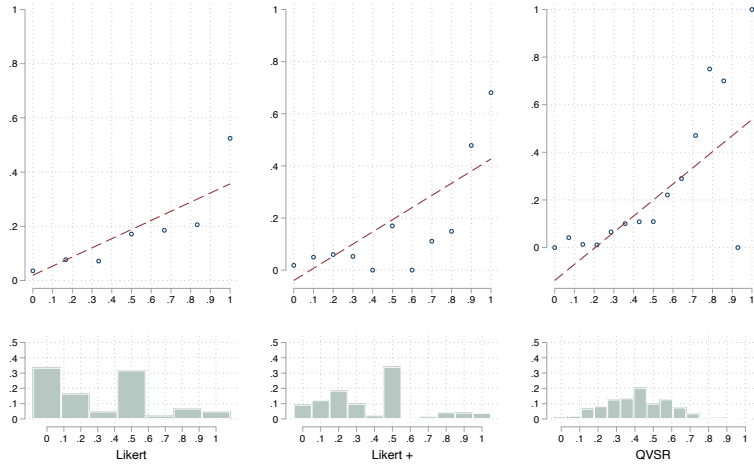
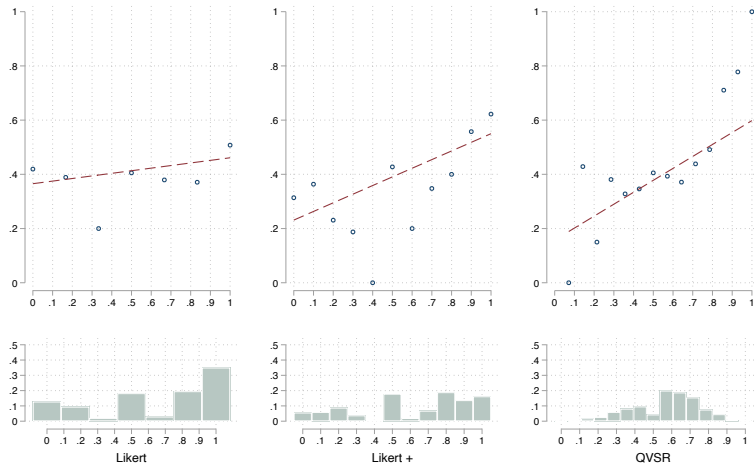Amount Taken in DG

Proportion Taken in DG
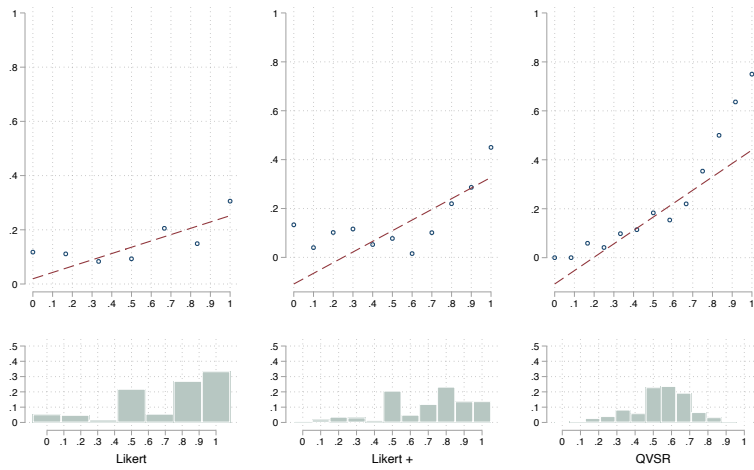
% Female and Suport for Equal Pay
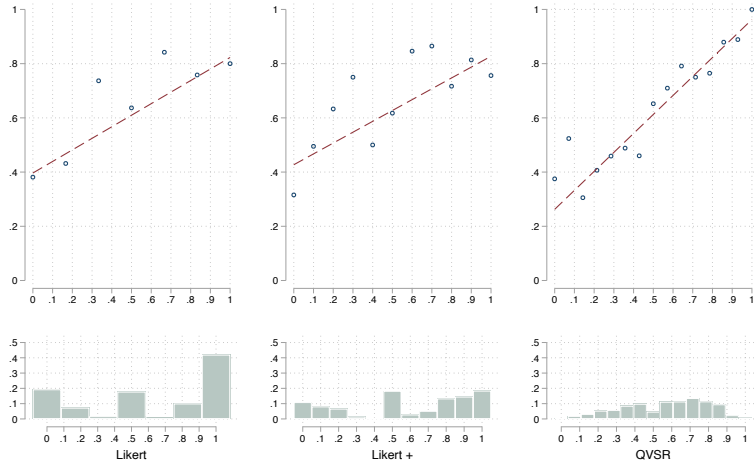
% Black and Support for Affirmative Action

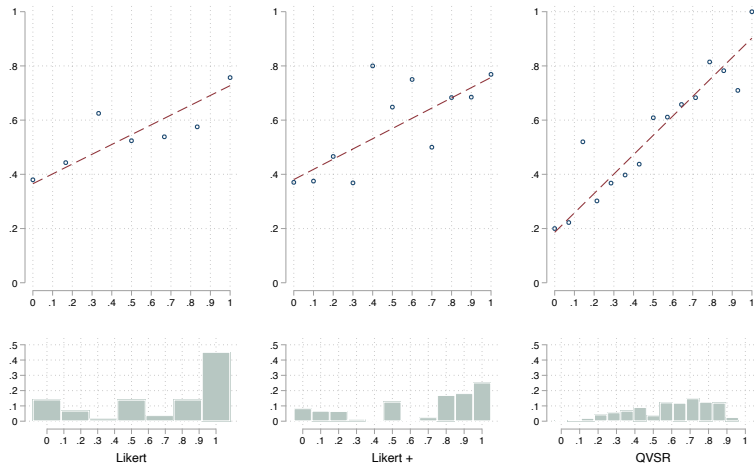% At or Close to Min Wage and Support for MinW increase

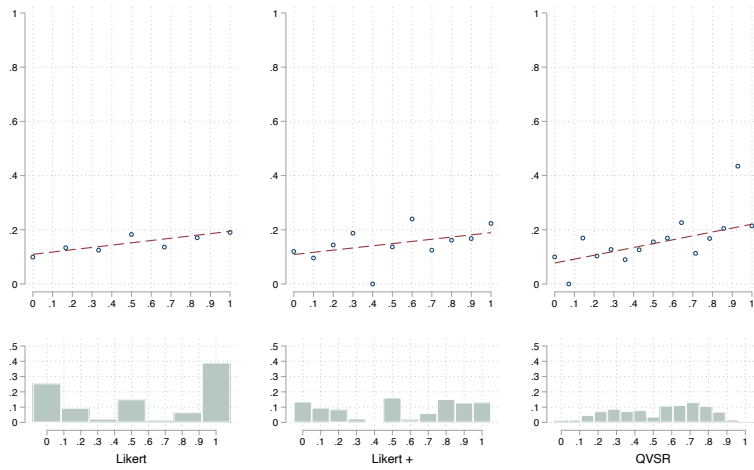% Close to Childbirth and Support for Paid leave

A23

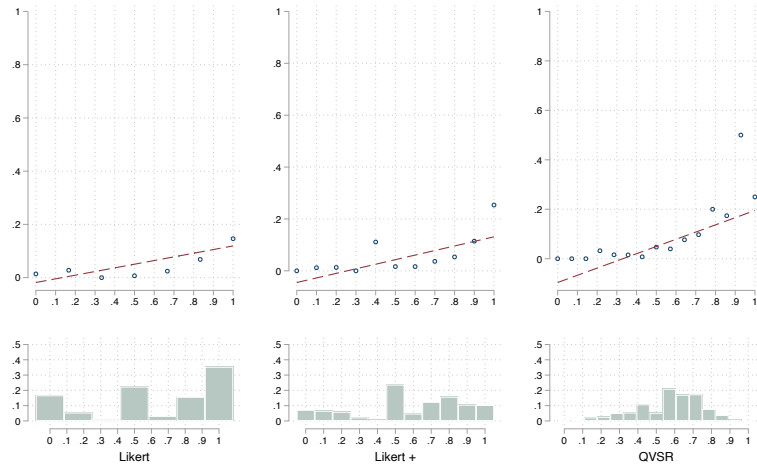% NOT Born Again and Opposition to Abortion Ban

% Does NOT Own a Gun and Support for Gun Control

% With 1 or 2 Foreign Born Parents and Support for Border Wall

% Lesbian/Gay and Support for Same Sex Adoption

Likert

Likert +

QVSR

6 Behavioral Tasks: Screenshots

6.1 Donation

Part 3/3

Please read this important information before you move to the third and last part of the survey:

At the end of the survey, the computer will randomly select **40** people among all the survey participants (40 among roughly 4000 people). Each winner will receive a bonus worth up to **$100**. Winners will be notified in the 10 days following the end of the survey.
In this section of the survey:

- We ask you to imagine that you are among the 40 lucky winners selected by the computer.
- We offer you the opportunity to donate some of the $100 bonus to one non-profit organization. What you do not donate, you can keep for yourself.
- On the next page, we provide you with more information on each organization. We then ask you whether you would like to make a donation.

If you are among the randomly chosen winners, we will pay you the bonus amount, minus your donation, in points credited to your Knowledge Panel™ account ($100 = 100,000 points).

If you would prefer to skip this part of the survey, you can do so below. Please note that respondents who do not complete this last section will not be entered into the drawing for one of the $100 bonuses.

○ Take me to next page and enter me into the drawing
○ I would like to skip this part of the survey

`>>`

Immigration policy

Numbers USA: this organization **FAVORS restrictive immigration policies** aimed at reducing both legal and illegal immigration to the United States. Members act to persuade elected officials **to FAVOR more restrictive immigration bills**. This organization also provides voters with information on candidates' positions on immigration control, encouraging them to vote for candidates that favor immigration control.

You can find more information on this organization by copying and pasting this link into a separate browser tab or windowr: `https://www.numbersusa.com`

American Immigration Council: this organization **OPPOSES restrictive immigration policies**, especially if they infringe on the rights of immigrants in the United States. Members act to educate the public on immigration issues and persuade elected officials **to OPPOSE more restrictive immigration bills.**. This organization also goes to court to defend immigrants' rights against government policies and practices.

You can find more information on this organization by copying and pasting this link into a separate browser tab or window: `https://www.americanimmigrationcouncil.org`

Non-profit organizations are dedicated to defending a wide variety of social causes. These organizations depend on the generosity of the general public to fulfill their mission.

We have picked non-profit organizations working on two policy issues previously discussed in this study. These issues are **gun policy** and **immigration policy**. For each issue, we have selected two reputable non-profit organizations, one for each side of the debate.

Please take a few minutes to familiarize yourself with these organizations.

Gun policy

Gifford Law Center to Prevent Gun Violence : this organization **FAVORS gun contro**l. Its main activities include lobbying state and federal legislatures in **FAVOR** of **bills** that **regulate the purchase, possession and use of firearms.** This organization also goes to court to defend gun control laws against legal challenges from people who oppose such laws.

You can find more information on this organization by copying and pasting this link into a separate browser tab or window: `http://lawcenter.giffords.org/`

Institute for Legislative Action: this organization **OPPOSES gun control**. Its mains activities include lobbying state and federal legislatures to **OPPOSE bills** that **regulate the purchase, possession and use of firearms.** This organization also provides voters with information on candidates' position on gun control, encouraging them to vote for candidates that oppose gun control.

You can find more information on this organization by copying and pasting this link into a separate browser tab or window: `https://www.nraila.org/`

Assume you have been selected to receive one of the $100 cash bonuses.

You have the option to donate none, all, or part of your $100 bonus to one of the non-profit organizations. Whatever you do not donate, you can keep for yourself.

Whatever you decide to donate, we will match it. For example is you donate $50, we will add another $50. We will pay the final amount to the organization of your choice, and **email you the receipt**. The donation will be **anonymous**.

Pick the non-profit you would like to make a donation to:

● Institute for Legislative Action
   (opposed to gun control)
○ Numbers USA
   (opposed to immigration)
○ Gifford Law Center to Prevent Gun Violence
   (favorable to gun control)
○ American Immigration Council
   (favorable to immigration)
○ I do not want to make a donation

How much of your $100 bonus do you wish to keep and how much do you wish to donate? Remember that we will double your donation. Reply using the response boxes below (the total must equal $100):

Donate	0
Keep for myself	0
Total	0

`>>`

6.2 Dictator Game

For this section, imagine that you have been selected by the computer to win the $100 bonus.

All winners are matched with another survey respondent, **randomly selected among all individuals who did not win the bonus.**

All you know about this person is **whether or not they identify with a political party,** and if so, **which party they identify with.**

You have the option to share some of your $100 bonus with this person. How much of your $100 bonus do you wish to keep and how much do you wish to give to this person?

Think carefully. If you win, your bonus will be shared with a real person based on your answer!

Click on the double arrows [>>] to proceed.

> >>

Imagine you are matched with an **Independent** selected among all the Independents who did not win the bonus. An Independent is someone who does not identify with any political party. You have the option to share some of your $100 bonus with this person.

How much of your $100 bonus do you wish to give to this unlucky survey respondent? Answer by filling in the blank with a number between 0 and 100.

*If I am matched with an **Independent**, I wish to give $ ____ to this person.*

> >>

Now, imagine you are matched with an **Independent**.

Based on information collected in wave 1 of this survey, you also learn that this person has donated money to two organizations:

- One organization **defends immigrants' rights** by advocating for undocumented immigrants in the United States.
- The other organization **defends gun owners' rights** by lobbying against gun control.

You previously indicated that you would give an unlucky **Independent** a portion of your bonus worth $. You would keep $100 for yourself.

If, in light of the donation information, you would like to change this amount, you can do so below. Otherwise click on the double arrows [>>] to move to the next page.

*If am matched with an **Independent** who has **donated to the two organizations described above**, I wish to give $ ____ to this person.*

You have chosen to give $0. This means that you will be receiving $100.

Is this correct?

○ No, I would like to modify my answer
○ Yes, take me to the next page

> >>

> >>

You have chosen to give $. This means that you will be receiving $100.

Is this correct?

○ No, I would like to keep my original answer ($ / $100)
○ Yes, take me to the next page

> >>

6.3 Letter Writing

The following 2 bill proposals are currently being examined by congressional committees:

<u>Official Bill Title:</u>

H.R. 7: No Taxpayer Funding for Abortion and Abortion Insurance Full Disclosure Act

<u>Summary:</u>

This bill seeks to limit the federal funds that can be used for abortion. The bill would:

- Prohibit the use of federal funds for abortion.
- Prohibit abortions at facilities owned or operated by the federal government.
- Prevent federal employees from performing abortions within the scope of their employment.

Part 3/3

You have arrived at the final section of the study.

On the next page, you will find information on 2 bill proposals that are **currently** being examined by a congressional committee. As an American citizen, you can **take action** by telling your representatives what you think about these bill proposals.

Please read through the summaries of each of the proposed bills. After you read the short summaries, we will give you the opportunity to voice your opinion on each proposed bill.

<u>Official Bill Title:</u>

S. 1242: Raise the Wage Act

<u>Summary:</u>

This bill seeks to raise the minimum wage. The bill would:

- Raise the federal minimum wage from $7.25 to $15 an hour by 2024.
- Guarantee that, going forward, the minimum wage will increase at the same rate as average wages.

This bill is expected to increase the wages of 25 percent of the U.S. workforce.

```
>>
```

In the next section, you will have the option to let Congress know what you think about one of these bills.

We will compile comments made by you and other survey respondents into a letter (one for each bill) and will send the letter to the head of the committee currently examining these bills. Your comments, if you chose to provide any, will be completely anonymous.

Would you be interested in letting Congress know your opinion on one of the bills?

○ Yes
○ No

Which bill would you like to express your opinion on?

○ Raise the Wage Act
○ No Taxpayer Funding for Abortion and Abortion Insurance Full Disclosure Act

```
>>
```

```
>>
```

You have a particular interest in the following bill proposal:

On this page, you have the opportunity to write your comments and thoughts. For instance, you can:

- Praise or criticize the bill
- Explain why this issue is important for you
- Urge members of Congress to support or oppose the bill

Tips: Your comments don't have to be complicated, just a few genuine sentences. Don't worry about typos or mistakes. They further show that you are a real person who hasn't been given a script by an activist or partisan group.

You can write your comments below:

>>

References

Achen, Christopher. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69(4):1218–1231.

Ansolabehere, Stephen, Jonathan Rodden and James Snyder. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting." *American Political Science Review* 102(2):215–232.

Benabou, Roland, Armin Falk, Luca Henkel and Jean Tirole. 2020. "Eliciting Moral Preferences: Theory and Experiments.".

Bullock, John, Alan Gerber, Seth Hill and Gregory Huber. 2015. "Partisan Bias in Factual Beliefs about Politics." *Quarterly Journal of Political Science* 10:519–578.

Cavaillé, Charlotte, Daniel L. Chen and Karine Van Der Straeten. 2019. "A Decision-Theoretic Approach to Understanding Survey Response: Likert vs. Quadratic Voting for Attitudinal Research." *University of Chicago Law Review* .

Zellner, Arnold. 1962. "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." *Journal of the American statistical Association* 57(298):348–368.