

Who Cares?
Measuring Differences in Preference Intensity
Online Appendix

Charlotte Cavaille*

Ford School of Public Policy, University of Michigan

Daniel L. Chen

Toulouse School of Economics - IAST - CNRS

Karine Van der Straeten

Toulouse School of Economics - IAST - CNRS

* cavaille@umich.edu

Contents

A. Preferences and Attitudes: Disciplinary Differences	A1
B. Modeling Survey Answers	A2
B.1. Assumptions about policy preferences	A2
B.2. Preferences and Attitudes	A3
B.3. Measuring policy preferences	A4
B.4. Assumptions about motivations when answering survey	A5
B.5. Optimization problem under different survey technologies	A5
B.6. Optimal responses under standard Likert items	A7
B.7. Optimal responses under Likert+	A7
B.8. Optimal responses under QVSR	A8
C. Survey Design	A11
C.1. Overview	A11
C.2. Completion Rate	A15
C.3. Participation in Wave 2	A15
C.4. Comparing Survey Experience Across Instruments	A20
D. Using Wave 2 Preferences	A21
E. Alternative Approaches to Likert+	A21
F. Descriptive Statistics and Figures for the Behavioral Tasks	A24
F.1. Shannon Entropy	A24
F.2. Scatter Plots and Histograms	A25
G. Material Self-Interest	A28
H. Behavioral Tasks: Screenshots	A31
H.1. Donation	A31
H.2. Dictator Game	A32
H.3. Letter Writing	A33

A. Preferences and Attitudes: Disciplinary Differences

Both social psychology (which provides micro foundations for public opinion studies) and economics (which provides micro foundations for political economy) start from the assumption that people act purposefully based on a subjective mental state that, if known, can help researchers explain people's decision to behave in one way versus another. Much work has been done trying to best characterize and measure this mental state in ways that are amenable to theory development and testing (i.e., predicting who will choose what in what context). Still, the social psychology and economics approaches to the study of opinion-driven behavior diverge philosophically in at least two ways.

First, in the economics approach, there are no such things as single-object "true" attitudes (though see Hausman (2023), chapter 1). Instead, all researchers can recover are preference rankings inferred from *observable choices between costly alternatives*. This difference follows from economics' emphasis on decision making under constraints.¹ A forced-choice measurement strategy naturally follows from this approach.

Second, the two fields differ in how they handle subjective survey data. For many scholars in the economics tradition, this type of data raises many red flags. First, changes in survey answers following simple manipulations indicate that cognitive engagement is very low (see Hanretty, Lauderdale and Vivyan (2020) –p. 3– for a discussion). Second, competing motives beyond sincerity (e.g., changing one's answer to not appear bigoted) introduce systematic measurement bias that interfere with the empirical analysis (Bertrand and Mullainathan 2003). For many economists, in light of subjective data's limits, the best one can do is indirectly infer preferences using costly real world decisions.

Social psychologists, in contrast, tend to work under the assumption that, when answering survey questions, people's main goal is to provide researchers with their most sincere, if imperfect, answer. While well aware of the cognitive processes lamented by economists, social psychologists approach them less as a source of measurement error and bias and more as objects of study important for what they reveal about human cognition (e.g., contrast Bertrand and Mullainathan (2003) and Zaller (1992)).

These differences partly originate in each field's own understanding of what theory generation and testing entail. To simplify, economists rely on models, that is, a set of assumptions and predicates that are neither true nor false and are used to develop testable claims about human decision making, itself conceived as a choice between competing alternatives. To the extent that human cognition is upstream of behavior, and unobservable, it is best tackled using a model whose truth is less important than its ability to predict (i.e., explain) observable individual decisions. Social psychologists, in contrast, are more committed to closing the gap between the fictional entities entailed by models and the "true" cognitive processes underpinning human behavior.

Our goal is not to arbitrate, settle or bridge this philosophical gap. In light of our interest in explaining how people choose between competing courses of action, the economics approach "à la QVSR" provides, at least conceptually, a natural fit. Still, the tools developed by social psychologists have withstood the test of time and there is enough research on attitude extremity and importance showing that existing attitudinal measures capture meaningful individual-level differences.

In Appendix B, we propose a model (per economists' definition) to investigate the claim that motives beyond sincerity shape people's survey answers. We posit the existence of two motives —report-

¹ The set of alternatives over which preferences are defined reflect these constraints as well as people's beliefs about these constraints.

ing one’s true opinion and reporting the opinion one associates with being a member of a given identity group—and discuss how survey methodology affects the weight respondents give to one motive over the other. We conclude that if people are “sincere,” then shifting to QVSR might not be worth the effort. However, in contexts in which competing motives (e.g., the partisan motive) affect people’s answers, then QVSR should, theoretically, provide a better measure. Note that this is only one angle through which to compare different survey methodologies. Another model might have focused on differences in cognitive engagement instead.

The model we propose in Appendix B is different from Zaller’s in Zaller and Feldman (1992) and Zaller (1992). Zaller seeks to provide a model of the survey answer that can explain a long list of well documented empirical patterns (including patterns of stability and instability, framing effects and the impact of situation cues more broadly). In that sense, Zaller is committed to finding the model that best fits the data (the type of “gap closing” mentioned earlier). Note that in Zaller’s case, the data to explain tends to be generated by measurement tools of the *simply-ask* family. In contrast, we start from a simple model of how people express an opinion (e.g., two competing motives) and examine how answers provided vary across survey technologies, *assuming our simple model is true*.

B. Modeling Survey Answers

B.1. 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. Our goal is to measure information on policy preferences, that is, the *extent* to which an individual prefers a proposed policy reform over the status quo. Formally, we capture one’s preference for each policy proposal $k = 1, \dots, K$ using a real number u_{ik} in the interval $[-1, 1]$, where 1 means the respondent strongly prefers the reform over the status quo and -1 means they strongly prefer the status quo over the proposed reform. We define u_{ik} as the difference between respondent i ’s utility if the reform is implemented and their utility under the status quo.² We assume that the likelihood of taking costly action in favor of the reform k and against the status quo increases as u_{ik} gets closer to 1. Conversely, the likelihood of taking costly action against the reform and in favor of the status quo increases as u_{ik} gets closer to -1 .

For expository purposes one can decompose u_{ik} into two terms: 1) an indicator variable which captures whether u_{ik} is positive or negative (preference orientation), and 2) the absolute value of u_{ik} (preference intensity). Throughout, we mostly focus on u_{ik} as defined above. When discussing preference intensity, we are referring to $|u_{ik}|$.

Preference orientation changes with the specific content of the proposed policy and how it compares to the status quo. Preference intensity varies not only with the specific content of the proposed policy but also with the relative importance of the policy domain. Take, for example, two individuals with moderate preferences in favor of a given reform. One individual might feel strongly about the policy domain yet be quite satisfied with the status quo, resulting in moderately intense preferences. Another might not care much about a policy domain yet be dissatisfied enough with the status quo to be moderately supportive of the reform. People with the most intense preferences will include those who care strongly about a given policy domain and are highly dissatisfied with the status

² Note that an individual’s utility from a proposed reform is shaped by both material and selfish considerations (e.g. how much an individual will personally benefit from a rise in the minimum wage) as well as non-material and other-regarding considerations (e.g., the reform aligns with their moral values and thus constitutes an improvement compared to the status quo).

quo.

To illustrate this point, assume that any policy on each issue k can be described by a number t_k in the interval $[0,1]$. For example, if issue k is regulation of abortion, $t_k = 1$ could mean full legalization of abortion, whereas $t_k = 0$ means a complete ban of abortion whatever the circumstances. Intermediate numbers could mean a ban with some exceptions. On issue k , denote the status quo by t_k^{SQ} and the reform by t_k^R . We assume that an individual's policy preference on each of these different policy dimension is characterized by two parameters: her 'ideal policy', denoted by t_{ik} , and how important this policy domain is to her. For example, independently of the specific reforms being discussed, an individual can think that gender-related issues are much more important than immigration-related issues. Formally, this 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 the type of issue related to dimension k , and $\beta_{ik} = 1$ means that it is of the utmost importance to her. Assuming quadratic preferences, the utility of respondent i if reform k is implemented is $-\beta_{ik} (t_{ik} - t_k^R)^2 + cst$, while under the status quo her utility is $-\beta_{ik} (t_{ik} - t_k^{SQ})^2 + cst$. Therefore the net value of reform to her is:

$$\begin{aligned}
 u_{ik} &= -\beta_{ik} (t_{ik} - t_k^R)^2 + \beta_{ik} (t_{ik} - t_k^{SQ})^2 \\
 &= \underbrace{\beta_{ik}}_{\text{domain importance}} * \underbrace{(t_k^R - t_k^{SQ}) * \left(t_{ik} - \frac{t_k^R + t_k^{SQ}}{2}\right)}_{\text{spatial alignment}}
 \end{aligned}$$

The second line in this equation makes transparent the determinants of preference orientation: preference orientation is positive if and only if the reform is closer to the individual's ideal point than the status quo. Preference intensity ($|u_{ik}|$) is the product of three terms: (i) how much one cares about this general field of policy (β_{ik}), (ii) whether the reform is far from the status quo ($|t_k^R - t_k^{SQ}|$) – otherwise, we are talking about mild insignificant changes anyway, (iii) whether the individual's ideal point is much closer to one alternative than to the other ($|t_{ik} - \frac{t_k^R + t_k^{SQ}}{2}|$), indeed, if the individual's ideal point is right between the status quo and the reform, she is almost indifferent. The last two components together form the "spatial alignment" component.

B.2. Preferences and Attitudes

Building on the above exposition, we can now relate preference intensity on the one hand, to other concepts in social psychology that have been tied to opinion-congruent behavior on the other. These concepts are attitude extremity, importance and strength.

An attitude is a psychological predisposition to favor or disfavor a particular object (e.g., a policy). While attitude extremity captures "the degree to which the person likes or dislikes the object," attitude importance captures "an individual's subjective judgment of the significance he or she attaches to his or her attitude" (Howe and Krosnick 2017: 329). Attitude extremity and importance are two important features of what Krosnick and Abelson (1992) call 'attitude strength,' which they define as the extent to which a given attitude "affects one's cognition or behavior." Krosnick and Abelson identify five sub-attributes of attitudes, all contributing to attitude strength. Given our interest in explaining attitude-congruent action, we leave aside features of attitude strength (e.g., attitude accessibility) most relevant for cognition (Howe and Krosnick 2017).

When measuring attitude extremity and importance, researchers rely on the *simply-ask* approach

described in the main paper. Respondents are asked to specify “how strongly” they favor (or oppose) a policy (the Likert item) and “how important” this policy is to them personally or, alternatively, “how much they (personally) care” about it (the personal importance item).

There is no 1-to-1 overlap between u_{ik} , domain importance and spatial alignment on the one hand, and these social psychology concepts on the other. As we detail in Appendix A, each reflect a very different conceptualizations of people’s opinion broadly defined. One might tentatively claim that attitudes are closer to what we have called t_{ik} , a respondent’s ideal policy. Still, conceptually t_{ik} assumes a high degree of opinion “specificity,” an assumption that attitudes are conceptually designed to relax.

What we can more easily do is relate the measurement tools commonly used to capture attitude extremity and importance (that is Likert and personal importance items) to u_{ik} , domain importance and spatial alignment. Due to differences in wording and scales, only Likert is informative of preference orientation *and* preference intensity, while the personal importance item is only informative of preference intensity. Both domain importance and spatial alignment shape answers to Likert and issue importance items. Put differently, the extent to which someone strongly supports a policy proposal and finds it personally important will be a function of their dissatisfaction with the status quo and how much they want to prioritize this policy domain. But wording matters. For example, a personal importance item that asks about a policy domain in general instead of a specific policy would be much more informative of domain importance, especially compared to Likert.

As we state in the manuscript, our definition of preference intensity does not negate the existence of attitudes, attitude extremity and attitude importance, it simply supplants it: given that both extremity and importance affect preference intensity, our goal is to measure the total sum of this effect, which we call preference intensity.

Note that a concept such as attitude strength does not provide an adequate substitute for preference intensity, this for two reasons. First, attitude strength is a broad umbrella term that covers attitude features that are relevant to behavior as well as features relevant to cognition. Because we are only interested in explaining opinion-congruent behavior, not cognition, we needed a different concept. Second, social psychologists relate strong attitudes to stable attitudes that are hard to change. In contrast, based on our definition of $|u_{ik}|$, preference intensity can vary over time depending, for example, on changes in the status quo.

B.3. Measuring policy preferences

To measure u_{ik} , researchers commonly rely on variants of the Likert item, which directly asks people how intensely they favor/oppose a given policy proposal. With Likert+, respondents are also asked to convey “how important” a given proposal is to them personally. This item wording helps identify people who care about a given policy domain, independently of the proposal being discussed. Given the wording of the issue importance item, we can also reasonably expect that people dissatisfied with the status quo and extremely favorable to a proposed reform will be more likely to convey in a survey that a proposed reform is personally important to them. This additional information might add to what is already collected by Likert. It might also be redundant, an issue we explore empirically in the paper.

There are two potential concerns with standard Likert and issue importance items. One is that respondents have only limited incentives to consider real-world trade-offs, something we call the *abundance problem*. As a result, their answers carry only limited information about how they might behave when faced with conflicting alternatives.

The second concern is that respondents may not answer truthfully. Among the reasons why respondents may not answer truthfully is the distortion introduced by partisan motives. In the current American context, one can reasonably expect polarizing ideological messaging and affective partisanship to combine to generate bi-modal response distributions on highly salient issues, something we call the *bunching problem*. In this case, the same response categories (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).

Next we discuss Likert and Likert+'s strength and weaknesses in light of these concerns and discuss the conditions under which QVSR might provide a comparatively better measure.

B.4. Assumptions about motivations when answering survey

We assume that an individual may have (at least) two (potentially) conflicting motives when asked about their policy preferences in a survey. On the one hand, they derive some intrinsic utility from answering each question sincerely. They might, for example, derive from some expressive benefits. Alternatively, this might be induced by a psychological cost of not behaving sincerely. We call this motive the *sincerity motive*. On the other hand, an individual may also care about how their answers will be read and interpreted by other people, which might conflict with this sincerity motive.

This second motivation encompasses a variety of psychological mechanisms, which can vary 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 themselves, or to whoever is going to read the survey, that they have some socially desirable traits. For example, they may want to appear altruistic or non-racist. They may also want to signal a group identity. For example, if they are a Republican, and they expect Republicans to take specific positions on some issues, they 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 they really stand.³

B.5. Optimization problem under different survey technologies

A researcher's goal is to probe out of respondents their most sincere answers to whether they favor or oppose a set of policy K . Under both Likert and QVSR, they are asked to answer a question about said support or opposition. What varies is the technology used to record the respondents' answers. In the model presented next, we assume that the sincere answer to the Likert item provides the best available proxy for u_{ik} , at least when using state preferences. Our goal is to examine whether adding a "How important" question or using QVSR's technology to answer the favor/oppose question does a better job than Likert at recovering this sincere answer.

Formally, in addition to previously defined $u_{ik} \in [-1, 1]$, we assume that, on each issue $k = 1, \dots, K$, respondent i is characterized by the following parameters:

³ 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.

- the sincere answer they would give to the benchmark Likert item on issue k , $u_{ik}^L \in [-1, 1]$,
- the answer to the standard Likert item that she finds the most attractive because of the signaling motive, denoted by $U_{ik}^L \in [-1, 1]$,
- the answer she would give to the personal importance item if she were to answer it sincerely, $Imp_{ik} \in [0, 1]$,
- the answer to the personal importance item that she finds the most attractive because of the signaling motive, denoted by $IMP_{ik} \in [0, 1]$.

As discussed above, both $|u_{ik}^L|$ and Imp_{ik} can be thought of as proxies for the intensity of preference ($|u_{ik}|$), which is our main quantity of interest. Both contain information about the spatial alignment and the policy domain importance, with Imp_{ik} arguably containing relatively more information about domain importance than $|u_{ik}^L|$.

Now, let's consider the K issues jointly. Under standard Likert, 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 "Extent of support/opposition" questions, denoted by $\hat{u}_i = (\hat{u}_{i1}, \dots, \hat{u}_{iK}) \in [-1, +1]^K$, in the following way:

$$V_i(\hat{u}_i) = -\sum_k w_{ik} \left[(1 - z_{ik}) \left(\hat{u}_{ik} - u_{ik}^L \right)^2 + z_{ik} \left(\hat{u}_{ik} - U_{ik}^L \right)^2 \right], \quad (1)$$

with $w_{ik} > 0$ and $z_{ik} \in [0, 1]$.⁴

QVSR seeks to recover the same quantities of interest as a battery of Likert items. This means that respondents, whether exposed to a battery of Likert items or to QVSR are trying to maximize the same utility function as above. Parameter z_{ik} is the relative weight of the signaling motive compared to the sincerity motive for this question. Parameters w_{ik} are meant to capture the psychological disutility that a respondent faces when she gives answers to a survey that do not exactly reflect her views. These parameters are 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 z_{ik} and w_{ik} , we can expect respondents to prioritize issues for which the sincerity motive is more important, which will explain why QVSR might be able to better measure u_{ik} than Likert.

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 $\hat{u}_i = (\hat{u}_{i1}, \dots, \hat{u}_{iK}) \in [-1, 1]^K$, and her answers to the personal importance item, denoted by $\hat{y}_i = (\hat{y}_{i1}, \dots, \hat{y}_{iK}) \in [0, 1]^K$. We make the following assumption in that case:

$$\begin{aligned} W_i(\hat{u}_i, \widehat{Imp}_i) &= -\sum_k w_{ik} \left[(1 - z_{ik}) \left(\hat{u}_{ik} - u_{ik}^L \right)^2 + z_{ik} \left(\hat{u}_{ik} - U_{ik}^L \right)^2 \right] \\ &\quad -\sum_k w'_{ik} \left[(1 - z'_{ik}) \left(\widehat{Imp}_{ik} - Imp_{ik} \right)^2 + z'_{ik} \left(\widehat{Imp}_{ik} - IMP_{ik} \right)^2 \right], \end{aligned}$$

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 personal impor-

⁴ We use these simple quadratic forms to derive some simple closed-form solutions. See Cavallé, Chen and Van Der Straeten (2019) for a more general specifications.

tance questions, and parameter z'_{ik} is the relative weight of the signaling motive compared to the sincerity motive for this question. Because we are interested in constructing a one-dimensional measure of u_{ik} , we will explain below how we use the answers to both questions to get a measure of u_{ik}^{L+} with Likert+.

Optimization problem Individuals are assumed to choose answers 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 questions depending on the survey tool. In particular, our interest will be in discussing whether the reported views are a good measure of $|u_{ik}|$. 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.

B.6. Optimal responses under standard Likert items

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

$$\max_{\hat{u}_i \in [-1,1]^K} V_i(\hat{u}_i) = - \sum_k w_{ik} \left[(1 - z_{ik}) (\hat{u}_{ik} - u_{ik}^L)^2 + z_{ik} (\hat{u}_{ik} - U_{ik}^L)^2 \right].$$

It is easy to check that the solution of the optimization program for issue k , denoted by \hat{u}_{ik}^L , is:

$$\hat{u}_{ik}^L = (1 - z_{ik}) u_{ik}^L + z_{ik} U_{ik}^L. \quad (2)$$

If $z_{ik} = 0$ (only the sincerity motive is active), the individual has no incentive to misreport her view, and $\hat{x}_{ik}^L = u_{ik}^L$. 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 $|U_{ik}^L|$ 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.

B.7. 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.

Answers to the "Favor/oppose" question are the same as with standard Likert items. Answers to the personal importance item, denoted \widehat{Imp}_{ik} :

$$\widehat{Imp}_{ik} = (1 - z'_{ik}) Imp_{ik} + z'_{ik} IMP_{ik}$$

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.

Because we are interested in constructing a one-dimensional measure of u_{ik} , in the empirical part of the paper, we will study different ways of combining \widehat{u}_{ik}^L and \widehat{Imp}_{ik} to get a measure of u_{ik}^L with Likert+:

$$\widehat{u}_{ik}^{L+} = F\left(\widehat{u}_{ik}^L, \widehat{Imp}_{ik}\right)$$

Comparison between Likert+ and standard Likert items Compared to standard Likert items, the advantage of Likert+ is obvious: it aims at collecting two pieces of information instead of one: one about u_{ik}^L (as in standard Likert) but also one about Imp_{ik} (the issue importance). If one knows about to best combine these two items, one can only improve upon Likert. Yet, as highlighted above, note that this direct measure of the issue importance 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.

B.8. 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:

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

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:

$$\begin{aligned} \max_{\widehat{u}_i \in [-1, 1]^K} \mathcal{L}(\widehat{u}_i, \lambda_i) &= - \sum_k w_{ik} \left[(1 - z_{ik}) \left(\widehat{u}_{ik} - u_{ik}^L \right)^2 + z_{ik} \left(\widehat{u}_{ik} - U_{ik}^L \right)^2 \right] \\ &+ \lambda_i \left[m - \sum_k \widehat{u}_{ik}^2 \right], \end{aligned}$$

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

$$\widehat{u}_{ik}^{QVSR} = \frac{w_{ik}(1 - z_{ik})}{w_{ik} + \lambda_i} u_{ik}^L + \frac{w_{ik}z_{ik}}{w_{ik} + \lambda_i} U_{ik}^L = \frac{w_{ik}}{w_{ik} + \lambda_i} \widehat{u}_{ik}^L.$$

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

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

$$\sum_k \left(\frac{w_{ik}}{w_{ik} + \lambda_i} \hat{u}_{ik}^L \right)^2 = m. \quad (3)$$

Note that the left-hand side of the equality is strictly decreasing in λ_i , taking the value $\sum_k (\hat{u}_{ik}^L)^2$ 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 (3) is satisfied.

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

$$\begin{aligned} \hat{u}_{ik}^{QVSR} &= \frac{1}{1 + \frac{\lambda_i^*}{w_{ik}}} \hat{u}_{ik}^L \\ &= \frac{1}{1 + \frac{\lambda_i^*}{w_{ik}}} \left[(1 - z_{ik}) u_{ik}^L + z_{ik} U_{ik}^L \right]. \end{aligned} \quad (4)$$

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 (4) 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 (1)).

Comparison between QVSR and standard Likert items Note that both standard Likert and QVSR only ask the "Favor/oppose" question. The main difference of QVSR compared to standard Likert is that answers under QVSR also incorporates information about w_{ik} , that is, how much the individual suffers when she has to deviate from reporting her unconstrained answers on issue k .

Remember that our goal is to identify the hypothetical sincere answer to the "Favor/oppose" question as simply worded with Likert. Whether QVSR provides a better measure of u_{ik}^L than the observed Likert answer depends on the correlation between w_{ik} (as defined above) and the relative importance of the sincerity motive compared to the signaling motive ($1 - z_{ik}$). If there is a positive correlation between w_{ik} and $1 - z_{ik}$, then QVSR can be expected to better measure u_{ik}^L . To understand why, remember that when a signaling motive is present, answers under Likert only imperfectly captures u_{ik}^L , since answers under Likert (the \hat{x}_{ik}^L) are typically a weighted average between the sincere answer (the u_{ik}^L parameter) and some signaling target (see (2)). 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 (4)). 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 (\hat{u}_{ik}^L) are closer to the sincere answer (u_{ik}^L), thus better capturing the true u_{ik}^L parameters.

There are reasons to expect such 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 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 has thought a lot about the issue and her sincerity motive is strong, getting her answer right to the

question might not be that important to her.

As discussed above, under QVSR, the individual relatively allocates more credits to issues for which w_{ik} is large. Another reason why QVSR might perform better than Likert at measuring the preference of intensity is that w_{ik} might independently capture relevant information about how much the individual cares about the issue. There are reasons to expect such a positive correlation between parameters w_{ik} and the (true) importance of the issue (parameters Imp_{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. If w_{ik} correlates with Imp_{ik} (the intrinsic importance of the issue) then, compared to standard Likert, answers under QVSR, are also incorporate information of Imp_{ik} . QVSR can be seen as a way to also collect information about issue importance, but not by asking the question directly (as with Likert+) but by forcing respondents to choose across issues.

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 \hat{u}_{ik}^L (same response to the "Favor/oppose" question as under standard Likert), plus an imperfect report about the personal issue importance \widehat{Imp}_{ik} . Under QVSR, one gets \hat{u}_{ik}^{QVSR} , which by (2) incorporates information about both \hat{u}_{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" importance parameter. In such a context, an advantage of QVSR is to bring a "forced choice" / revealed preference approach to the measurement of issue importance. 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+ personal importance item 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.

C. Survey Design

C.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. Sample size and balance tables are included below and discussed in the next section.

Table C1: Sample Sizes: Overview

Wave 1	Likert	Likert+	QVSR
Number of times survey was started (dup incl)	1392	1391	1397
Number of individuals who started the survey (dup excl)	1357	1349	1325
Duplicate rate	3	3	5
Donation task (valid obs.**)	1257	1259	1163
Gender (valid obs.**)	1333	1333	1251
Drop out rate (as a share of ind.*)	2	1	6
Loss rate (as a share of all obs.)	4	4	10
Wave 2	Likert	Likert+	QVSR
DG (valid obs.**)	513	532	493
Letter Writing (valid obs.**)	518	540	508

* GfK-Ipsos’ carefully-maintained representative panel ensures high completion rates.

** We drop observations that could not be matched with GfK-Ipsos’ background information data or that have a missing value on key outcome variables and predictors. We also drop duplicates, keeping the first observation recorded (time-wise).

Table C2: Likert Treatment vs. Likert+/QVSR (pooled)

Variable	(1) Likert+/QVSR (pooled)	(2) Likert	(3) Difference
Donation_gun	10.316 (33.885)	9.035 (32.824)	-1.281 (1.165)
Donation_immi	1.290 (29.013)	1.646 (28.543)	0.356 (1.003)
Party_ID	4.178 (2.213)	4.115 (2.208)	-0.064 (0.077)
Ideology	4.106 (1.622)	4.128 (1.610)	0.023 (0.056)
HS_or_less	0.280 (0.449)	0.310 (0.462)	0.029* (0.016)
Some_college	0.318 (0.466)	0.289 (0.453)	-0.029* (0.016)
BA	0.402 (0.490)	0.402 (0.490)	-0.000 (0.017)
Age	52.281 (16.091)	52.581 (16.616)	0.299 (0.568)
Gender	1.506 (0.500)	1.492 (0.500)	-0.015 (0.017)
White	0.754 (0.431)	0.760 (0.427)	0.006 (0.015)
Black	0.092 (0.289)	0.079 (0.270)	-0.013 (0.010)
Other	0.067 (0.250)	0.071 (0.256)	0.003 (0.009)
Hispanic	0.087 (0.281)	0.090 (0.286)	0.003 (0.010)
Exp_minW	0.259 (0.438)	0.262 (0.440)	0.003 (0.015)
MissV_minW	0.374 (0.484)	0.388 (0.488)	0.015 (0.017)
Paid_leave_exp	1.581 (0.793)	1.566 (0.797)	-0.014 (0.028)
Gun_exp	0.626 (0.484)	0.605 (0.489)	-0.021 (0.017)
Immi_exp	1.242 (0.601)	1.243 (0.600)	0.001 (0.021)
Sexual_orientation	0.066 (0.249)	0.068 (0.252)	0.002 (0.009)
Born_again	0.631 (0.483)	0.614 (0.487)	-0.016 (0.020)
MissV_BA	0.299 (0.458)	0.320 (0.467)	0.020 (0.016)
Observations	2,422	1,257	3,679

* $p < .05$, ** $p < .01$ *** $p < .001$.

Exp_minW: = 1 if would benefit from min wage increase. *Miss_minW*: = 1 if did not answer min wage items. *Paid_leave_exp*: proximity to childbirth score (see Fig.5 notes in main manuscript). *Gun_exp*: see Table G10. *Immi_exp*: score capturing how many relatives are immigrants. *MissV_BA*: = 1 if missing value on religion questions used to identify evangelical Christians ("born again").

Table C3: Likert+ Treatment vs. Likert/QVSR (pooled)

Variable	(1) Likert/QVSR (pooled)	(2) Likert+	(3) Difference
Donation_gun	9.543 (32.688)	10.523 (35.088)	0.980 (1.165)
Donation_immi	1.306 (28.354)	1.614 (29.790)	0.308 (1.003)
Party_ID	4.144 (2.211)	4.182 (2.212)	0.038 (0.077)
Ideology	4.124 (1.598)	4.093 (1.657)	-0.031 (0.056)
HS_or_less	0.295 (0.456)	0.282 (0.450)	-0.012 (0.016)
Some_college	0.303 (0.460)	0.317 (0.466)	0.014 (0.016)
BA	0.403 (0.491)	0.401 (0.490)	-0.002 (0.017)
Age	52.563 (16.178)	52.038 (16.448)	-0.525 (0.567)
Gender	1.495 (0.500)	1.514 (0.500)	0.019 (0.017)
White	0.754 (0.431)	0.761 (0.427)	0.007 (0.015)
Black	0.086 (0.281)	0.091 (0.287)	0.005 (0.010)
Other	0.069 (0.253)	0.068 (0.251)	-0.001 (0.009)
Hispanic	0.092 (0.288)	0.080 (0.272)	-0.011 (0.010)
Exp_minW	0.260 (0.438)	0.261 (0.439)	0.001 (0.015)
MissV_minW	0.376 (0.484)	0.384 (0.487)	0.009 (0.017)
Paid_leave_exp	1.571 (0.793)	1.584 (0.798)	0.013 (0.028)
Gun_exp	0.609 (0.488)	0.636 (0.481)	0.026 (0.017)
Immi_exp	1.241 (0.599)	1.244 (0.604)	0.003 (0.021)
Sexual_orientation	0.070 (0.256)	0.061 (0.239)	-0.010 (0.009)
Born_again	0.618 (0.486)	0.638 (0.481)	0.020 (0.020)
MissV_BA	0.310 (0.462)	0.300 (0.459)	-0.009 (0.016)
Observations	2,420	1,259	3,679

* $p < .05$, ** $p < .01$ *** $p < .001$.

See notes Table C2.

Table C4: QVSR Treatment vs. Likert/Likert+ (pooled)

Variable	(1) Likert/Likert+ (pooled)	(2) QVSR	(3) Difference
Donation_gun	9.779 (33.977)	10.092 (32.546)	0.313 (1.189)
Donation_immi	1.630 (29.168)	0.939 (28.157)	-0.691 (1.023)
Party_ID	4.148 (2.210)	4.175 (2.214)	0.026 (0.079)
Ideology	4.111 (1.634)	4.119 (1.585)	0.009 (0.058)
HS_or_less	0.296 (0.456)	0.278 (0.448)	-0.017 (0.016)
Some_college	0.303 (0.460)	0.318 (0.466)	0.015 (0.016)
BA	0.401 (0.490)	0.404 (0.491)	0.002 (0.017)
Age	52.309 (16.531)	52.545 (15.699)	0.236 (0.579)
Gender	1.503 (0.500)	1.498 (0.500)	-0.005 (0.018)
White	0.761 (0.427)	0.747 (0.435)	-0.014 (0.015)
Black	0.085 (0.279)	0.093 (0.291)	0.008 (0.010)
Other	0.069 (0.254)	0.067 (0.249)	-0.003 (0.009)
Hispanic	0.085 (0.279)	0.093 (0.291)	0.008 (0.010)
Exp_minW	0.261 (0.439)	0.257 (0.437)	-0.004 (0.016)
MissV_minW	0.386 (0.487)	0.362 (0.481)	-0.024 (0.017)
Paid_leave_exp	1.575 (0.797)	1.577 (0.789)	0.002 (0.028)
Gun_exp	0.620 (0.485)	0.615 (0.487)	-0.006 (0.017)
Immi_exp	1.244 (0.602)	1.239 (0.598)	-0.004 (0.021)
Sexual_orientation	0.064 (0.245)	0.073 (0.260)	0.008 (0.009)
Born_again	0.626 (0.484)	0.623 (0.485)	-0.004 (0.021)
MissV_BA	0.310 (0.463)	0.298 (0.458)	-0.012 (0.016)
Observations	2,516	1,163	3,679

* $p < .05$, ** $p < .01$ *** $p < .001$.

See notes Table C2.

C.2. Completion Rate

For each 100 survey taking attempts, 4 do not provide a usable observation when respondents are assigned to Likert or Likert+. That number increases to 10 when respondents are assigned to QVSR. In other words, the loss rate is 6 percentage points higher in QVSR than with more traditional survey methodologies (see Table C1, *loss rate* row). In QVSR, the share of respondents who start the survey more than once is 2 percentage points larger than in Likert and Likert+ (see Table C1, *Duplicate rate* row). The smaller sample size for QVSR has consequently two origins: more duplicates (i.e., the same individual finishing the survey more than once) and more attrition-related losses (starting but not finishing the survey).

The larger number of duplicates is plausibly tied to respondents wanting to go back and take the QVSR portion of the survey again. This could be out of interest for the tool or because respondents wanted to “do it differently” once they had a better understanding of the QVSR. For the analysis, we only keep the first observation (using the time-stamp provided by Qualtric).

The main source of attrition is the segment of the survey in which respondents are asked their preferences using QVSR. We examine whether observable covariates predict who, conditional on starting the survey, ends up in the final dataset and whether this varies across treatments. In this analysis, we use all survey-taking attempts as our reference population (including duplicates). The analysis consequently captures who, conditional on starting the survey, is more likely to either not finish the survey or produce a duplicate observation. As Table C5 shows, none of the available covariates predict attrition in QVSR. For Likert and Likert+, race, income and race predict drop out but effect sizes are substantively too small to generate imbalances between treatment conditions.

In a previous study, Quarfoot et al. (2017) find no evidence of differential attrition. What explains then the 6 percentage points difference between QVSR and other methods in our study? One explanation might be our own survey hardware, which required respondents to switch between two platforms. This could have introduced glitches (e.g., a longer wait time when transitioning from Qualtrics to the QVSR platform, which discouraged some survey takers). Given no systematic differences between who finishes the survey and who drops out, one interpretation is that these glitches were more or less random and affected individuals in ways uncorrelated with their individual features. This interpretation aligns with evidence that the number of people re-starting the test is higher in QVSR. Alternatively, the higher drop out rate is a defining feature of QVSR and needs to be factored in when using QVSR. If this is the case, then follow-up studies will have to confirm whether the absence of attrition predictors replicates in samples beyond the one used in this study.

C.3. Participation in Wave 2

We contacted respondents who had no missing values in the preferences and donation variables. In total, 1,257 (Likert), 1,259 (Likert+) and 1,163 (QVSR) individuals were eligible to participate in wave 2. Among people contacted, an average of 42% took the wave 2 survey, with a low of 40% for people assigned to Likert in wave 1 and a high of 45% for people assigned to QVSR in wave 1. This difference could be due to the treatment: people exposed to QVSR in wave 1 could be more likely to willingly participate in a follow-up survey. We caution against this interpretation: because of a slightly higher attrition rate, it is necessary to reach out to more QVSR-treated respondents in order to meet response targets.

In Table C6, we examine whether observable covariates predict participation in wave 2 (conditional on being eligible to participate). In Table C7, we examine whether donation behavior and policy preferences measured in wave 1, and relevant to the analysis in wave 2, predict the likelihood of

Table C5: Predicting Within-Wave Dropout

	(1)	(2)	(3)
	Likert	Likert+	QVSR
	b/se	b/se	b/se
Party_ID	0.004 (0.003)	0.002 (0.003)	0.010 (0.005)
Ideology	-0.000 (0.004)	-0.005 (0.005)	0.014 (0.007)
Ref: HS or less			
Some_College	-0.004 (0.012)	0.009 (0.014)	-0.029 (0.021)
BA	-0.001 (0.013)	0.007 (0.015)	-0.018 (0.022)
Age	0.000 (0.000)	0.001* (0.000)	0.001 (0.001)
Gender	0.003 (0.010)	0.012 (0.011)	0.018 (0.016)
Ref: White			
Black	0.042* (0.019)	-0.009 (0.020)	0.004 (0.031)
Other	-0.001 (0.020)	0.022 (0.022)	-0.021 (0.035)
Hispanic	-0.005 (0.018)	0.019 (0.021)	-0.052 (0.032)
Exp_minW	-0.000 (0.013)	0.003 (0.015)	-0.010 (0.023)
MissV_minW	0.005 (0.013)	-0.021 (0.014)	0.010 (0.022)
Paid_leave_exp	-0.001 (0.002)	-0.000 (0.002)	0.001 (0.003)
Gun_exp	0.010 (0.010)	-0.016 (0.011)	0.001 (0.017)
Immi_exp	0.006 (0.009)	-0.003 (0.010)	0.019 (0.015)
Sexual_orientation	0.003 (0.019)	-0.021 (0.023)	0.012 (0.032)
Born_again	-0.003 (0.012)	0.002 (0.013)	0.020 (0.021)
MissV_BA	0.013 (0.014)	-0.007 (0.016)	0.024 (0.024)
Income	-0.001 (0.001)	-0.003* (0.001)	-0.001 (0.002)
cons	-0.013 (0.043)	0.040 (0.049)	-0.090 (0.075)
N	1377	1378	1382

* $p < .05$, ** $p < .01$ *** $p < .001$.

participating in wave 2. The outcome is equal to 1 if individuals consented to participate in wave 2, 0 otherwise.

In the Likert+ treatment, age and gender predict wave 2 participation but with only substantively small implications for imbalances across methods. In the Likert treatment, race and gun ownership predict wave 2 participation. We consequently compare the coefficients on race and gun ownership across treatments. We find evidence that gun owners and white respondents were statistically more likely to participate in wave 2 (relative to non gun owners and minority respondents) if they were assigned to Likert in wave 1 (not shown). These imbalances are of limited concern given that the analyses focusing on preferences for gun control and affirmative action only use wave 1 data, not wave 2 data.

Turning to Table C7, there is no evidence that attitudes in wave 1 predict participation in wave 2. On the one hand, people who cast more votes on the wall issue are more likely to participate in wave 2 when assigned to QVSR. On the other hand, the coefficient is substantively small. Furthermore, there is no evidence that border wall attitudes differentially predict wave 2 participation across the three treatment conditions (not shown).

Table C6: Predicting Participation in Wave 2

	(1)	(2)	(3)
	Likert	Likert_plus	QVSR
	b/se	b/se	b/se
Partisanship	0.003 (0.009)	0.001 (0.009)	-0.017 (0.010)
Ideology	-0.010 (0.013)	-0.005 (0.013)	-0.023 (0.013)
Some_College	0.040 (0.037)	0.007 (0.037)	0.056 (0.039)
BA	-0.021 (0.038)	0.050 (0.040)	0.047 (0.041)
Age	0.001 (0.001)	0.003* (0.001)	0.003 (0.002)
Gender	-0.025 (0.028)	-0.058* (0.029)	-0.026 (0.030)
Black	-0.114* (0.056)	-0.082 (0.052)	0.038 (0.056)
Other	-0.137* (0.058)	-0.072 (0.060)	0.044 (0.064)
Hispanic	-0.064 (0.054)	0.040 (0.057)	-0.084 (0.057)
Exp_minW	-0.001 (0.039)	0.011 (0.039)	-0.004 (0.042)
MissV_minW	0.029 (0.037)	0.049 (0.038)	0.050 (0.041)
Paid_leave_exp	-0.002 (0.026)	0.022 (0.026)	0.006 (0.027)
Gun_exp	0.097** (0.030)	-0.020 (0.031)	0.037 (0.032)
Immi_exp	0.006 (0.026)	0.032 (0.026)	-0.006 (0.028)
Sexual_orientation	-0.035 (0.057)	0.047 (0.061)	0.045 (0.060)
Born_again	-0.025 (0.036)	-0.005 (0.036)	-0.028 (0.038)
MissV_BA	-0.001 (0.042)	-0.049 (0.043)	-0.074 (0.043)
Income	0.007 (0.004)	0.002 (0.004)	0.004 (0.004)
cons	0.288 (0.152)	0.265 (0.156)	0.421* (0.165)
N	1245	1249	1155

Table C7: Predicting Participation in Wave 2 Using Wave 1 Donations and Opinions

	(1)	(2)	(3)
	Likert	Likert_plus	QVSR
	b/se	b/se	b/se
Donation_gun	0.000 (0.000)	-0.000 (0.000)	0.000 (0.000)
Donation_immi	0.000 (0.000)	0.000 (0.000)	-0.000 (0.001)
Opinion_gun_abs	-0.007 (0.013)	-0.008 (0.009)	0.004 (0.009)
Opinion_wall_abs	0.025 (0.013)	0.004 (0.009)	0.022* (0.009)
Opinion_minW_abs	-0.004 (0.013)	-0.002 (0.009)	-0.018 (0.011)
Opinion_abortion_abs	0.014 (0.012)	0.009 (0.008)	0.005 (0.009)
cons	0.335*** (0.049)	0.413*** (0.044)	0.396*** (0.053)
N	1257	1259	1163

C.4. Comparing Survey Experience Across Instruments

A survey instrument is valuable for what it measures. But the quality of the measure is not the only criterion researchers use when choosing a survey tool. Also important is the additional costs induced by relying on one instrument instead of another. In this section, we discuss three costs: implementation, user experience that affects completion rate and user experience that affects survey time. High implementation costs, lower completion rates and longer survey time affect the costs of a project. We also discuss user experience that affects survey engagement, which could affect how carefully respondents think before providing an answer.

Implementation costs: When we ran our study, we had to hire someone to code the QVSR segment of the survey. To avoid other researchers having to pay these costs, we have made a QVSR web application available for free (qvrs.io, see also charlottecavaille.com/qvrs/). This web application easily integrates with Qualtrics or other survey platforms. Implementation is consequently easy and cheap for all three methods.

Completion rate: We found that QVSR affects completion rate. The higher attrition rate for QVSR could be due to technical issues (e.g., redirect delays between survey platforms), the explanatory video (e.g., respondents are unwilling to sit through it), or the survey tool itself (e.g., respondents find it too complicated and drop out). Given that Quarfoot et al. (2017) did not find differences in completion rates, we speculate that user experience tied to delays switching back and forth between Qualtrics and the QVSR platform is at least partly to blame. We have tried to minimize this concern when designing the new web app. Follow-up studies will need to confirm what the average attrition rate is and how technical improvements can help minimize it.

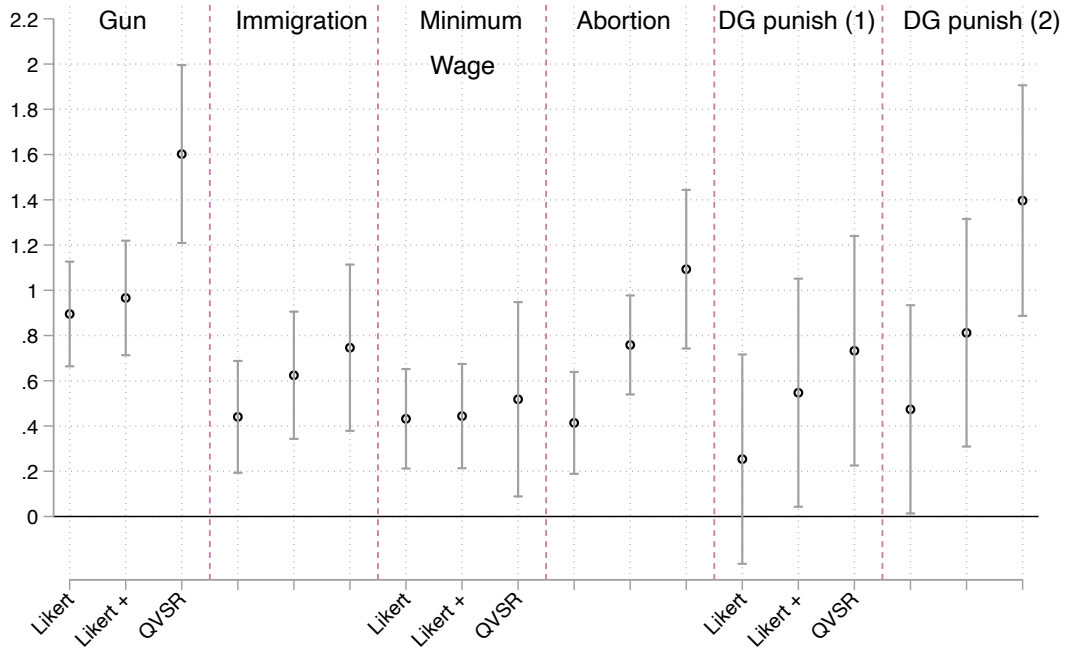
Survey time: Answering 10 item using QVSR takes about the same time as answering 10 Likert items following by 10 issue importance items (Likert+). On average, both QVSR and Likert+ take 30 seconds more to answer, when compared to 10 Likert items. The cost in terms of survey time is higher for QVSR because of the 90 second explanatory video. Related to Likert, QVSR thus costs on average 2 more minutes of survey time.

Survey engagement: While QVSR increases survey costs, these costs might have benefits in terms of survey engagement. Put differently, more time could also mean more careful answers as, by design, people cannot click through the survey at a high speed. This issue is discussed in detail by Cavallé, Chen and Van Der Straeten (N.d.) who provide evidence that answer quality improves under QVSR for respondents with a high school degree only.

D. Using Wave 2 Preferences

Figure D1 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.

Figure D1: Revisiting Figure 3 Using Preferences Measured in Wave 2



E. Alternative Approaches to Likert+

In the manuscript, to combine information collected using Likert and personal importance items, we rely on a multiplicative approach (**multi**). The result is a multiplicative scale ranging from -15 to 15 . In this section, we replicate the main results using alternative ways of combining information from these two items. We consider four versions of the Likert+ scale.

Importance only (ImpO) Personal importance items improve on Likert by explicitly asking respondents how much they care about a given issue. Using information from the first step of the Likert item (favor/neither/oppose), we construct a scale ranging from -5 (oppose, extremely important) to $+5$ (favor, extremely important) and centered around 0 (neither-nor). This version of Likert+ discards information from the second step in the 2-step Likert item. By comparing the **ImpO** and **multi** versions of the scale, we can assess the extent to which the information contained in the second step of the Likert item is redundant with the one provided by the personal importance item.

Additive (Add) We combine the Likert item and the personal importance item by adding them such that this version of Likert+ runs from -8 to 8 .

Y-hat linear (YHL) In Table E8, we replicate an analysis commonly found in studies of attitude strength (Carsey and Layman 2006; Miller, Krosnick and Fabrigar 2017). We regress the outcomes used in Figures 3 through 5 on the Likert and personal importance variables, allowing for a linear interaction between the two. In all cases but one (letter writing on minimum wage), the interaction term is positive and significant at the 10% level or better. The **multi** version does not use information from the first order terms and assumes an interaction coefficient that is equal across policy proposals. In the **YHL** version of the index, we combine the two items using the same linear combination and weights as in Table E8 (columns (1) through (4)). In practice, this amounts to computing predicted values, hence the name Y-hat.

Table E8: OLS with interaction term

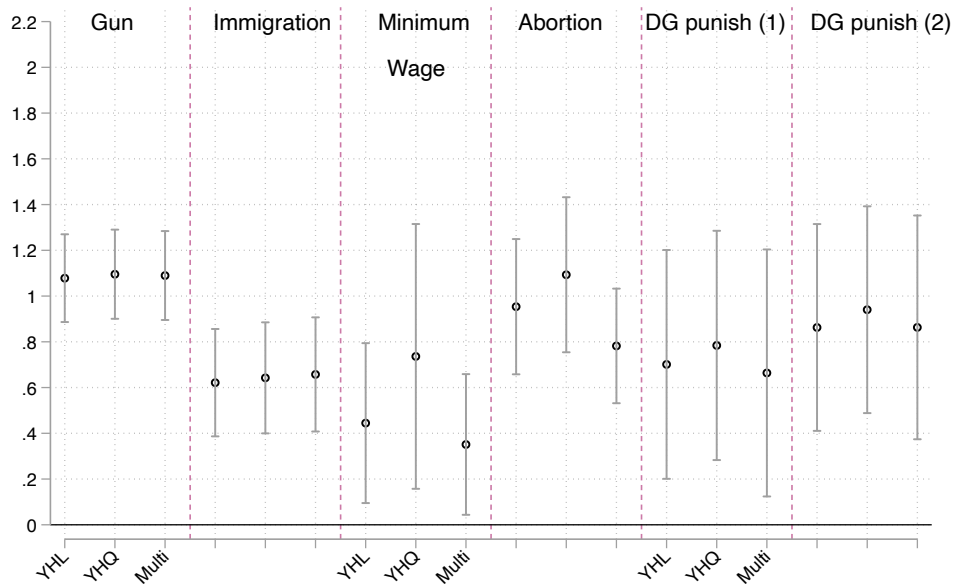
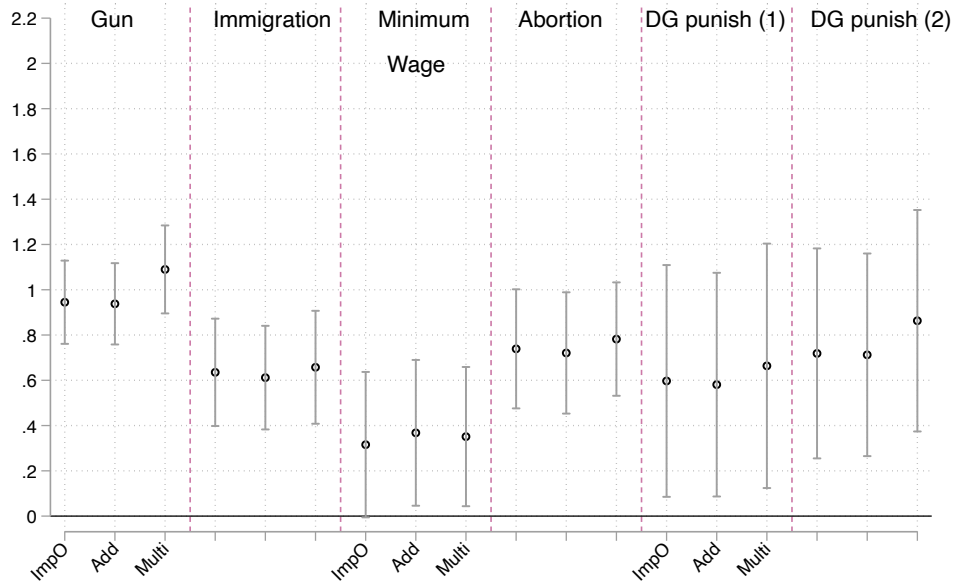
	(1) Donation (gun) b/se	(2) Donation (immi) b/se	(3) Letter (abortion) b/se	(4) Letter (minW) b/se	(5) Gender b/se	(6) Prox. to childbirth b/se
Likert	-0.00 (0.05)					
Issue_importance	0.02 (0.03)					
Interaction	0.05*** (0.01)					
Likert		0.07 (0.04)				
Issue_importance		0.00 (0.02)				
Interaction		0.02+ (0.01)				
Likert_abs			-0.23+ (0.13)			
Issue_importance			-0.08 (0.12)			
Interaction			0.11* (0.04)			
Likert_abs				0.12 (0.15)		
Issue_importance				0.04 (0.13)		
Interaction				-0.01 (0.05)		
Likert					0.01 (0.03)	
Issue_importance					0.05* (0.02)	
Interaction					0.02+ (0.01)	
Likert						-0.06 (0.03)
Issue_importance						0.10*** (0.02)
Interaction						0.04*** (0.01)
cons	-0.29** (0.11)	-0.05 (0.09)	-0.09 (0.31)	-0.30 (0.34)	0.17* (0.07)	1.16*** (0.07)
<i>N</i>	1264	1265	541	542	1338	1343
adj. <i>R</i> ²	0.187	0.133	0.072	0.006	0.094	0.114

+ $p < .10$ * $p < .05$, ** $p < .01$ *** $p < .001$.

Y-hat quadratic (YHQ) The **YHQ** version of Likert+ is similar to **YHL** except that we allow for a non-linear relationship between the personal importance variable and the outcome. Given QVSR's quadratic set up, we model this non-linear relationship using a quadratic term.

Figure E2 reproduces the analysis in Figure 3 using the multiplicative approach as a benchmark and

Figure E2: Comparing Alternative Versions of Likert+ (1)



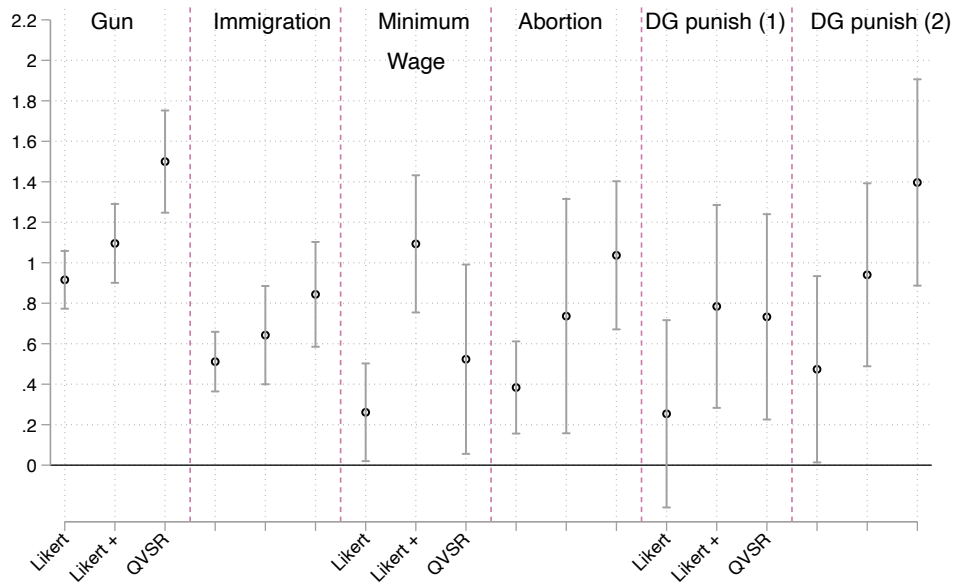
comparing it to the four versions discussed in this section. In all cases but one, the estimate for the alternative approaches are either of similar size or smaller, indicating that, for our horse race, we use a version of Likert+ that is not biased against this measurement strategy.

The exception is the **YHQ** version of Likert+. To further assess whether our results would change using this alternative index, we re-run the main analysis, comparing the YHQ index to Likert and QVSR. The results, presented in Figure E3, show that this version of the index is better than QVSR at discriminating between respondents when the letter writing task (minW bill) is used as a proxy for intense preferences. This result suggests that the personal importance item is, for this outcome, penalized by using a simple functional form to combine information from Likert and personal

importance items.

Still, the implications for measuring preference intensity are limited. First, the advantages of **YHQ** are limited to one outcome only. Second, while QVSR returns a simple interpretable index, reliance on **YHQ** to measure preference intensity outputs an index that cannot be interpreted in a straightforward way. Finally, to build the **YHQ** scale, we had to derive weights from a regression, using our behavioral variables as outcomes. This has two implications. First, the weights are themselves derived by leveraging information from our validation outcome, something not true of Likert and QVSR. The fact that, even under these ideal circumstances, Likert+ does not perform better than QVSR in its **YHL** and **YHQ** versions is evidence in favor of QVSR. Second, and most importantly, a researcher deciding which measure to choose from will not have access to a validation outcome, precluding reliance on the **YHQ** version of Likert. If they did, then it would be better to simply use this outcome as a measure of preference intensity.

Figure E3: Revisiting Figure 3 Using the **YHQ** version of Likert+



F. Descriptive Statistics and Figures for the Behavioral Tasks

F.1. Shannon Entropy

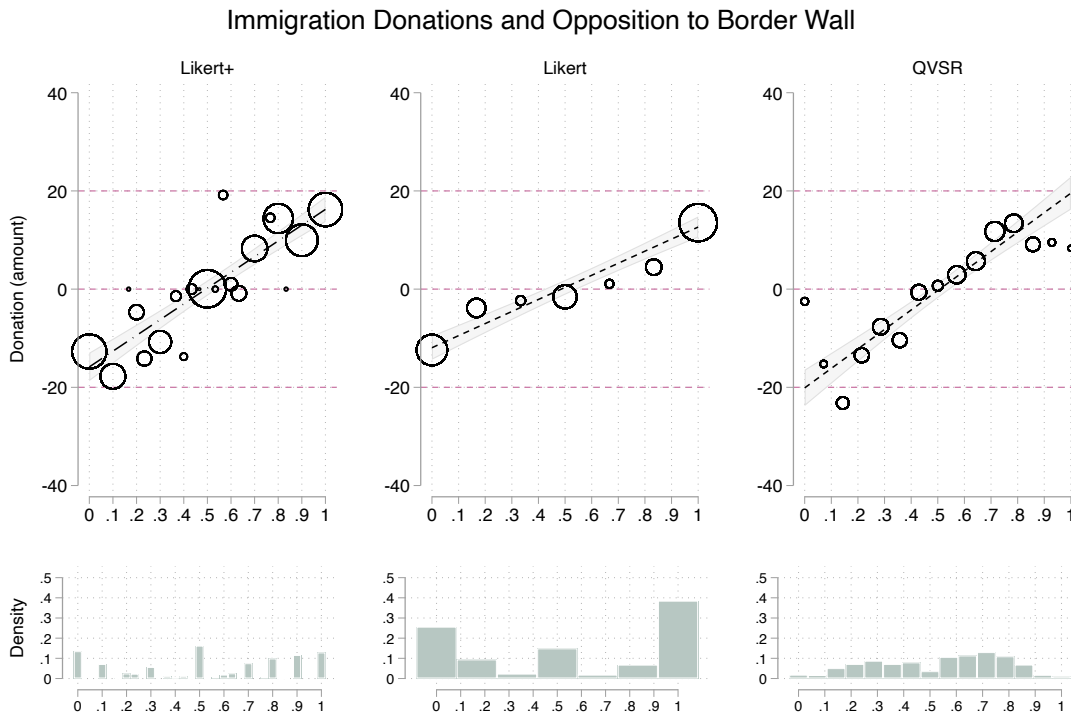
Table F9: Entropy Scores

	Likert	Likert +	QVSR
Same sex right to adopt	1.62	2.16	2.24
Make it difficult to buy gun	1.57	1.99	2.41
Wall on the US Border	1.58	2.18	2.52
Paid leave	1.59	2.03	2.03
Preferential hiring of blacks	1.58	1.94	2.27
Pay women and men the same	1.00	1.71	2.03
Minimum wage to 15 an hour	1.67	2.15	2.25
Ban on abortion	1.56	2.13	2.49
Cap on federal spending	1.47	1.99	2.04
Regulation for environment	1.61	1.84	2.12

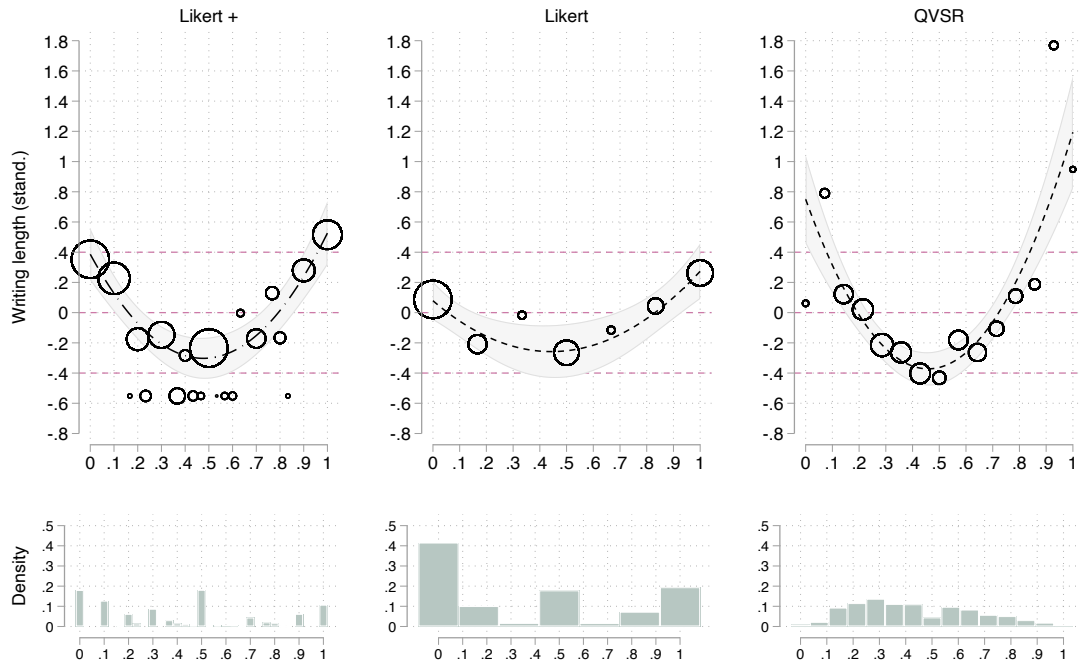
F.2. Scatter Plots and Histograms

The figures below plot average outcome within a given survey response bin. The outcomes are described in Table 2 in the main manuscript, alongside the relevant survey item used. Scatter plots are overlaid with a linear fit line. In each figure, the bottom panel shows the distribution of normalized survey answers.

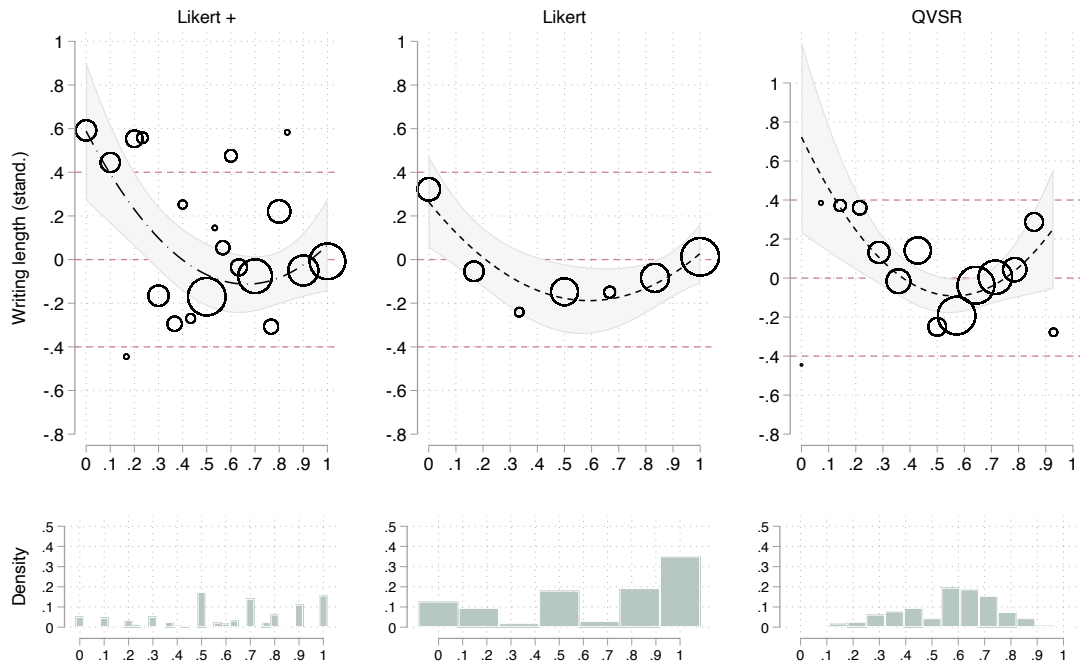
Note that, for information purposes, and in contrast to Figure 3 in the main manuscript, the figures for the letter writing outcomes plot the full scale of the response variable instead of the absolute value. This way of plotting shows that letter writing length increases at both ends of the scales (strong in favor or strong against abortion restriction). In this case, scatter plots are overlaid with a quadratic fit line



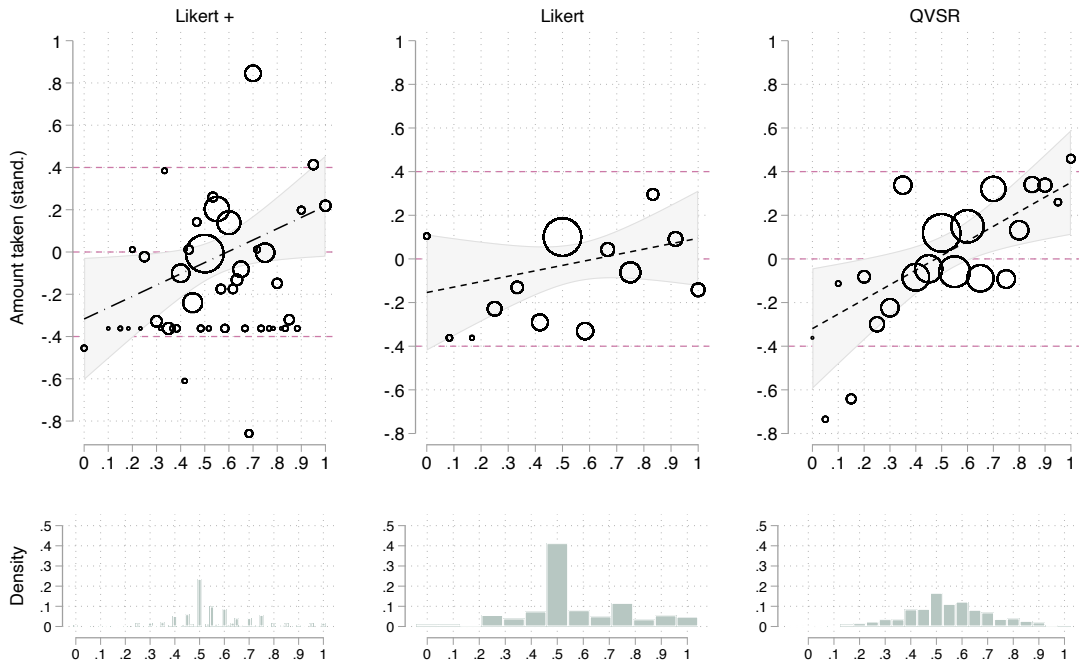
Length of Writing: Abortion Bill and Support for Abortion



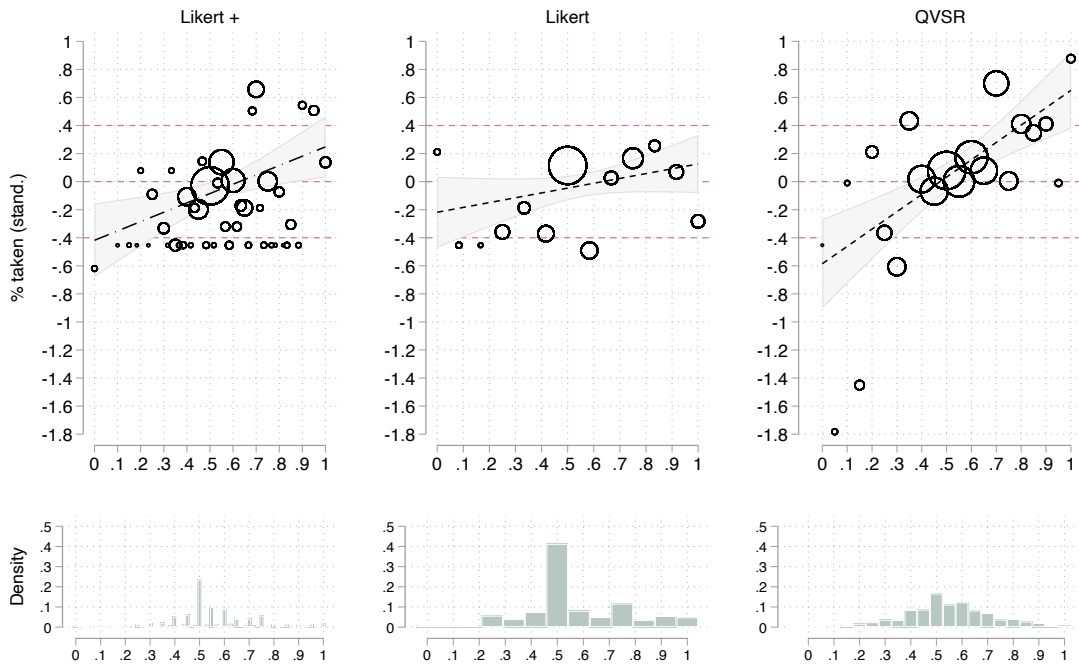
Length of Writing: minW Bill and Support for minW



Amount Taken in DG and Difference Gun - Immi

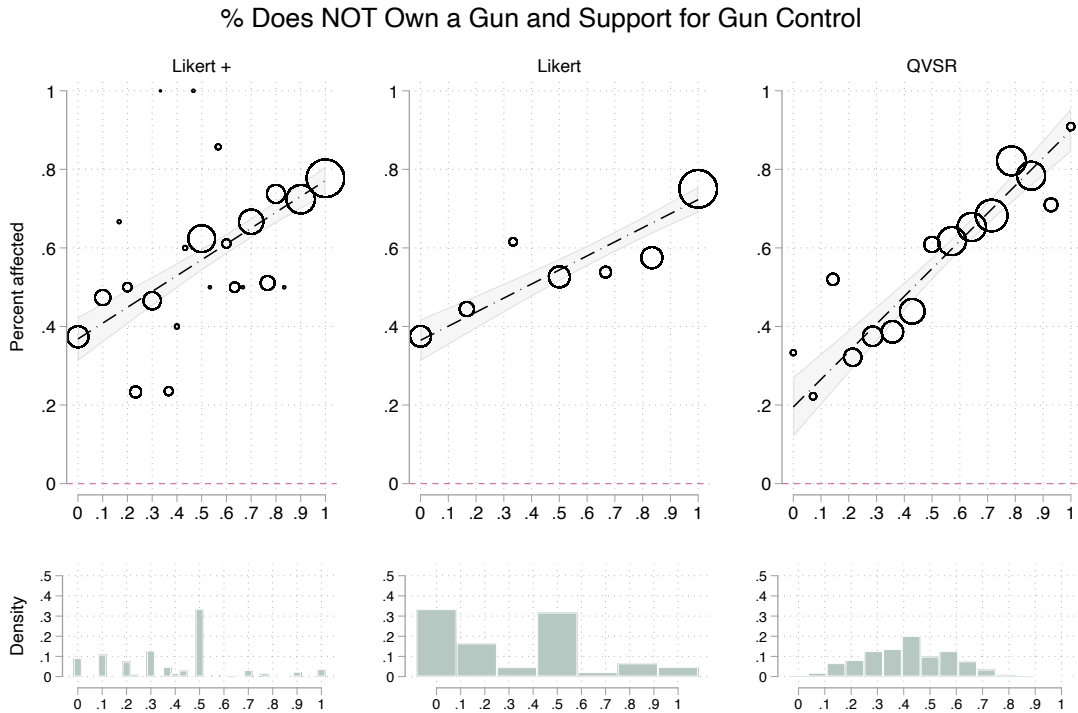


Perc. Taken in DG and Difference Gun - Immi

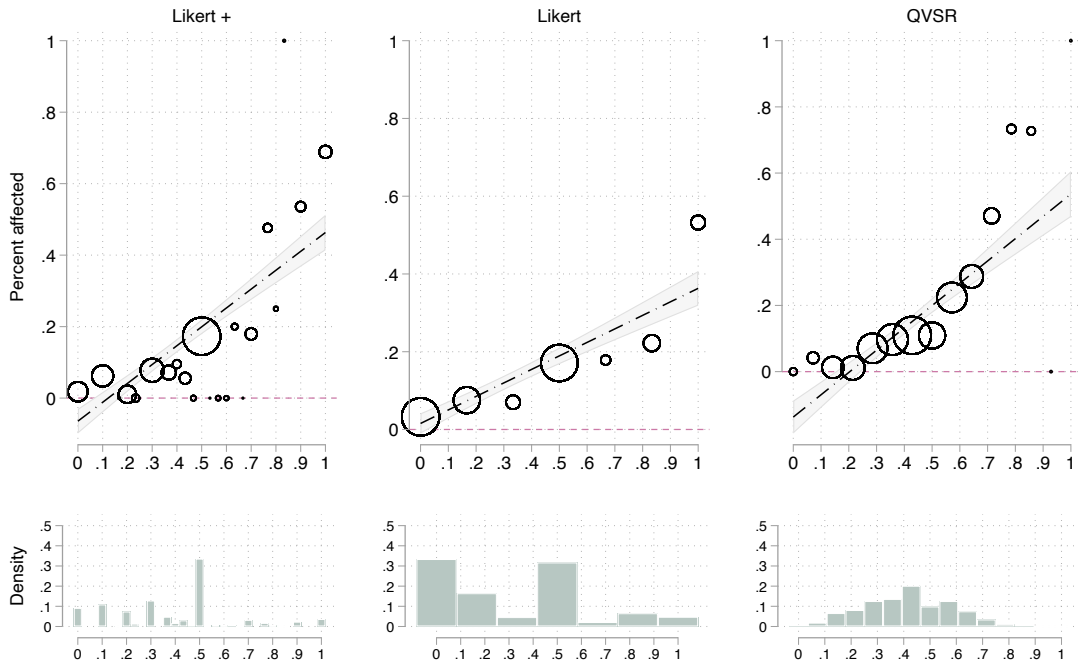


G. Material Self-Interest

In this section, we examine the relationship between preferences for a given policy propose and personal exposure to this policy. The outcome used to proxy for exposure are described in Table G10, alongside the relevant survey item. The figures below plot average outcome within a given survey response bin. Scatter plots are overlaid with a linear fit line. In each figure, the bottom panel shows the distribution of normalized survey answers. Difference across methods are also evaluated in Table G11 using a SUR model. Two exposure proxies are discussed in the main manuscript, namely gender and proximity to childbirth. The remaining three are gun ownership, race and whether one's wage is at or below minimum wage. In this latter case, sample sizes are smaller as we drop a third of respondents who did not feel comfortable answering our question on wage levels. 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 is true of 4 coefficients out of 5 when comparing Likert+ to Likert. For all of the 5 exposure proxies, the QVSR coefficient is larger than that for Likert+, but this difference is significant for only one coefficient. Irrespective of the exposure proxy, the Likert+ coefficient is never larger than that for QVSR.



% Black and Support for Affirmative Action



% At or Close to Min Wage and Support for Min Wage Increase

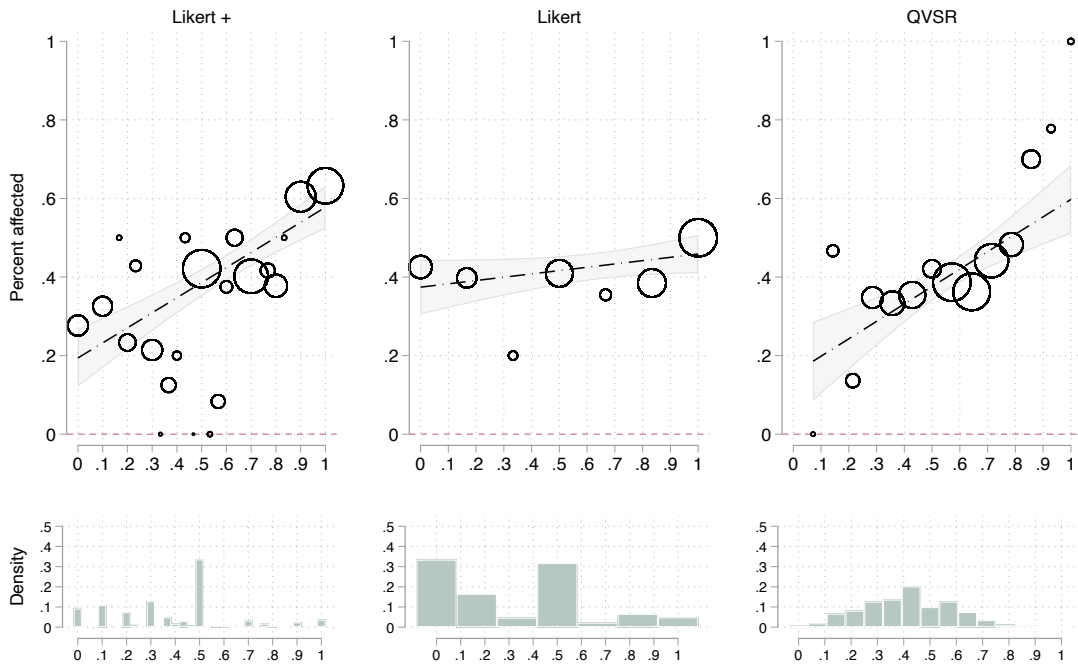


Table G10: Proxies of Personal Exposure and Relevant Policy Proposal

Policy Proposal	Exposure Proxy (Mean)
Laws making it more difficult for people to buy a gun	Equal to 1 if does not own gun or none in the house does, 0 otherwise (0.63)
Preferential hiring and promotion of blacks to address past discrimination	Equal to 1 if black, = 0 otherwise (0.13)
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)

Table G11: Differences in Coefficient Size (SUR): Exposure Outcomes table

	QVSR vs. Likert b/se	Likert + vs. Likert b/se	QVSR vs. Likert + b/se
Gender	0.58*** (0.15)	0.36** (0.12)	0.17 (0.15)
Prox_childB	0.88*** (0.22)	0.65*** (0.18)	0.29 (0.23)
No_Gun	0.26** (0.10)	-0.04 (0.07)	0.29** (0.10)
Black	0.26** (0.08)	0.15* (0.06)	0.15 (0.09)
Min_wage	0.34** (0.12)	0.27*** (0.08)	0.08 (0.12)
F test	52	9	30
Prob > F	0.000	0.004	0.000
Prob not rej. the null	< 0.000	< 0.000	< 0.008

* $p < .05$, ** $p < .01$ *** $p < .001$. We replicate Table 2 in the main paper using exposure proxies (instead of behavior outcomes). 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 0.58. Bottom row: F-test of the null-hypothesis that the sum of the coefficients is equal to 0.

H. Behavioral Tasks: Screenshots

H.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 window: <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 control**. 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 main 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 if 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	<input type="text" value="0"/>
Keep for myself	<input type="text" value="0"/>
Total	<input type="text" value="0"/>

>>

H.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

>>

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

H.3. Letter Writing

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.

>>

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

>>

To contact us: cc1933@georgetown.edu

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

Official Bill Title:

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

Summary:

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.

Official Bill Title:

S. 1242: Raise the Wage Act

Summary:

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.

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

- Benabou, Roland, Armin Falk, Luca Henkel and Jean Tirole. 2020. "Eliciting Moral Preferences: Theory and Experiments."
- Bertrand, Marianne and Sendhil Mullainathan. 2003. "Enjoying the quiet life? Corporate governance and managerial preferences." *Journal of political Economy* 111(5):1043–1075.
- 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.
- Carsey, Thomas M and Geoffrey C Layman. 2006. "Changing sides or changing minds? Party identification and policy preferences in the American electorate." *American Journal of Political Science* 50(2):464–477.
- 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* .
- Cavaillé, Charlotte, Daniel L. Chen and Karine Van Der Straeten. N.d. "Willingness to Say: Optimal Survey Design for Prediction." Toulouse School of Economics, working paper.
- 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.
- Hausman, Daniel M. 2023. *The inexact and separate science of economics*. Cambridge University Press.
- Howe, Lauren and Jon Krosnick. 2017. "Attitude Strength." *Annual Review of Psychology* 68:327–351.
- Krosnick, Jon and Robert Abelson. 1992. The Case for Measuring Attitude Strength in Surveys. In *Questions about Questions: Inquiries into the Cognitive Bases of Surveys*, ed. J.M. Tamur. Russell Sage Foundation pp. 177–203.
- Miller, Joanne M, Jon Krosnick and Leandre A Fabrigar. 2017. The origins of Policy Issue Salience. In *Political psychology: New explorations*, ed. Jon A Krosnick, I-Chant A Chiang and Tobias H Stark. Routledge pp. 125–171.
- 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.
- Zaller, John. 1992. *The Nature and Origins of Mass Opinion*. Cambridge University Press.
- Zaller, John. 2012. "What Nature and Origins Leaves Out." *Critical Review* 24(4):569–642.
- Zaller, John and Stanley Feldman. 1992. "A simple theory of the survey response: Answering questions versus revealing preferences." *American journal of political science* pp. 579–616.