Eliciting beliefs as distributions in online surveys

Leemann, Lucas ; Stoetzer, Lukas ; Traunmueller, Richard

DOI: https://doi.org/10.1017/pan.2020.42

Originally published at:
Leemann, Lucas; Stoetzer, Lukas; Traunmueller, Richard (2017). Eliciting beliefs as distributions in online surveys. In: 7th Annual General Conference of the European Political Science Association, Milan, Italy, 22 June 2017 - 24 June 2017, EPSA.
DOI: https://doi.org/10.1017/pan.2020.42

Eliciting Beliefs as Distributions in Online Surveys*

Lucas Leemann[†] Lukas F. Stoetzer[‡] Richard Traunmüller[§]

Abstract

Citizens' beliefs about uncertain events are fundamental variables in many areas of political science. While beliefs are often conceptualized in the form of distributions, obtaining reliable measures in terms of full probability densities is a difficult task. In this letter we ask whether there is an effective way to elicit beliefs as distributions in the context of (online) surveys? Relying on experimental evidence, we evaluate the performance of five different elicitation methods designed to capture citizens' uncertain expectations. Our results suggest that an elicitation method originally proposed by Manski (2009) performs well. It reliably measures the subjective belief distribution of average citizens and is easily implemented in the context of regular (online) surveys. We expect that a wider use of this method will lead to considerable improvements in the study of citizens' expectations and beliefs.

Word count: 3,918 — according to https://leeper.shinyapps.io/pdfcount/.

* Many thanks for very helpful comments and help to Daniel Bischof, Tim Hicks, Patrick Kraft, Andreas Murr, Simon Munzert, and David Rothschild. We thank Lucien Baumgartner for impeccable research assistance. This research was partly funded by the Department of Political Science at the University of Zürich.

[†]Department of Political Science, University of Zurich, Switzerland. Email: leemann@ipz.uzh.ch, URL: http://www.lucasleemann.ch

[‡]Department of Political Science, Humboldt University of Berlin, Germany. Email: lukas.stoetzer@hu-berlin.de, URL: http://lukas-stoetzer.org/

[§]University of Mannheim, Germany. URL: www.richardtraunmueller.com, Email: traunmueller@uni-mannheim.de

1 Introduction

Citizens' beliefs about uncertain events are fundamental variables in many areas of political science, including work on attitudes (e.g. Zaller and Feldman, 1992), cognitive biases (e.g. Gerber and Green, 1999; Bartels, 2002; Bullock, 2009), ambivalence (e.g. Alvarez and Brehm, 1997), misinformation (e.g. Berinsky, 2017), or citizen forecasts (e.g. Murr, 2011; Leiter et al., 2018) to name just a few. While beliefs are often conceptualized theoretically in the form of distributions, obtaining reliable measures of these beliefs in terms of full probability densities is a difficult task (Savage, 1971; Garthwaite et al., 2005; Goldstein and Rothschild, 2014). Most survey questions are focused on the first moment of an underlying distribution and thus miss important information regarding the variance or uncertainty of beliefs.

The question we ask in this letter is whether there is an effective way to elicit the belief distributions of average citizens in the context of (online) surveys? This paper discusses five different elicitation methods designed to capture citizens' uncertain expectations. We present experimental evidence and evaluate which question format is best suited to elicit continuous beliefs as distributions from regular (i.e. non-expert) survey respondents. That is, we are interested in how well these methods capture subjective distributions when compared to a benchmark and which of these methods performs best.

Our results suggest that an elicitation method originally proposed by Manski (2009) performs well. It contains five sequential survey questions that reliably measure the subjective belief distribution of average citizens and are easily implemented in the context of regular (online) surveys. They are also easy and quick to answer and, hence, not too cost intensive in (online) surveys. We expect that a wider use of this method will lead to considerable improvements in the study of citizens' expectations and beliefs and therefore important political science theories.

The remainder of this letter proceeds as follows. The next section discusses the elicitation process. Section 3 then presents the experiment and the five elicitation approaches that are evaluated. Section 4 discusses how formats are evaluated and presents the results.

2 Eliciting Beliefs as Distributions

The elicitation of beliefs as distributions has a long tradition in statistics, psychology, and economics.[1] In political science, it is Bayesians that seek to elicit prior distributions from *experts* to inform their statistical models (Gill and Walker, 2005; Gill and Freeman, 2013). But the process of eliciting probability distributions described in these literatures is usually a time-consuming enterprise requiring careful effort even if used to learn about the beliefs of experts who may already be familiar with probabilities.

What makes the elicitation of beliefs so difficult is that average people are not used to expressing themselves in easily quantifiable ways. Not only are average citizens less likely to be familiar with the concept of probability and less used to express their expectations in terms of distributions. Lengthy elicitation protocols also do not scale well to the number of respondents required for testing political science theories about citizens' expectations and are unlikely to be part of nationally representative surveys. Thus, the central challenge is how to best translate what they think into probability distributions within the confines of standard survey methodology.

Formally, an elicitation process can involve up to four steps (Garthwaite et al., 2005). In the *setup* step the problem is defined, respondents are recruited and trained in the key concepts and procedures. The *elicitation* is the key step, where the respondent is asked to provide information about his or her subjective belief. In the *fitting* step, this elicited information is converted into a probability distribution. The final step assesses the *adequacy* of the elicited distribution and provides the opportunity for correction. The challenge we address in this letter is how to implement these steps in the context of regular (online) surveys, where both time and scale concerns as well as limited researcher-respondent interaction render the use of full elicitation protocols impractical.

Traditional elicitation methods come in three basic forms (Spetzler and Stael von Holstein, 1975). In each of these three forms subjects are asked questions and the answers represent points on a cumulative distribution function. In so-called P-methods, subjects

[1] Our paper focuses on methods that are implementable in online surveys and as such leaves out incentivized elicitation methods that are often used in economics laboratory experiments (for an overview see e.g. Schlag et al., 2015)

are provided with fixed values referring to the quantity of interest and asked to give probabilities attached to these values (e.g. what is the probability that X is below Y?). In V-methods, subjects are instead provided with pre-defined probabilities and asked to give respective values (e.g. at what value are half the observations above and half below the value?). PV-methods are more difficult and simultaneously integrate both approaches. For instance, respondents may be asked to draw a graph of a probability distribution. In this letter we evaluate several implementations of these methods in terms of (online) survey questions.

Given humans' difficulties with probabilities, eliciting beliefs as distributions is as much a psychological problem as it is a statistical one. Many cognitive human biases are well known, e.g. representativeness, availability, anchoring biases, as well as the law of small numbers and hindsight biases (Tversky and Kahneman, 1971, 1973, 1974; Kynn, 2008). But it is important to distinguish those biases in beliefs from biases introduced by elicitation methods. Psychological research suggests that while people are generally capable of estimating proportions, modes, and medians, they are less proficient at assessing the means of highly skewed distributions (Peterson and Miller, 1964) and often have serious misconceptions about variances (Garthwaite et al., 2005). People are reasonable at quantifying their opinions as credible intervals but have the tendency to imply a greater degree of confidence than is justifiable (Wallsten and Budescu, 1983; Cosmides and Tooby, 1996).

3 Experimental Set-Up

In the following we evaluate a broad set of elicitation question formats, some purely verbal and others fully interactive, taking advantage of the fact that many surveys are carried out online and provide us with the ability to use visualizations. We run a number of experiments and assess which format yields beliefs that are most consistent with an objective benchmark distribution. In the presentation of our experiments we follow the four steps of the elicitation process as described in the previous section: setup, elicitation, fitting, and adequacy check. Instead of working with arbitrary numbers, we rely on an example of citizens' beliefs over hypothetical election results. We hope that this illustrates the potential usage of the method to a political science audience.

3.1 The Setup Step

We ran experiments with a total of about 3,600 participants. We relied on Amazon Mechanical Turk (MTurk) which is widely used for scientific purposes (Berinsky et al., 2012; Mason and Suri, 2012; Thomas and Clifford, 2017).[2] Amazon Turk allowed us to carry out the experiment in a short time period and at low costs. We recruited workers advertising a study on *surveys*, *opinion polls*, and *charts*.

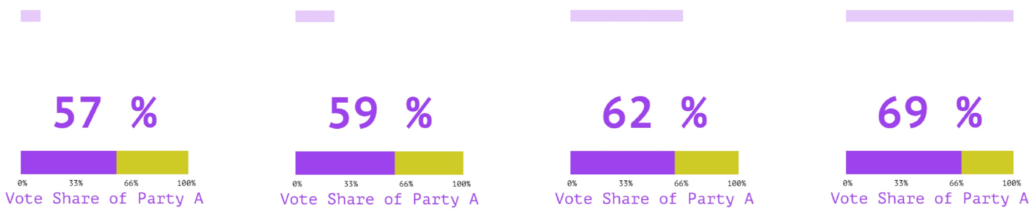

Figure 1: Four still frames from the GIF. Each still frame is shown for about half a second. The purple bar at the top is a progress bar.

We presented respondents with 100 results from hypothetical local elections that we randomly drew from pe-specified distribution. By exposing respondents to these draws, we manipulated the objective belief distribution[3] along two factors: symmetric vs. asymmetric and small vs. large variance. We rely on a Beta distribution in all four conditions but vary the shape parameters of the distribution. The symmetric small variance distribution is $\mathcal{B}(60,60)$ and the respective asymmetric distribution is $\mathcal{B}(60,30)$. For the large variance condition we rely on $\mathcal{B}(30,30)$ for the symmetric and $\mathcal{B}(30,15)$ for the asymmetric distribution. The realizations of 100 election results are presented as a short GIF where each frame shows one election outcome and is shown for about half a second. Four ran-

[2] Recent research shows that treatment effects and treatment heterogeneity can be well estimated in MTurk samples. Coppock (2018) replicates fifteen survey experiments and compares the estimates based on random samples with estimates based on a MTurk sample. In general, the two sets of estimates overlap.

[3] This works under the assumption that for our hypothetical example respondents hold no prior over the result.

dom draws are illustrated in Figure 1. This approach follows Goldstein and Rothschild (2014) who also relied on this form of visualization to present the distribution. The goal is to treat these distributions as the objective truth and to identify which question format elicits beliefs that are closest to the true distribution.

Each respondent is then also randomly assigned to elicitation question format in a simple between subjects design. There is balance across question types with respect to a number of socio-economic variables (see subsection 6.3 in the appendix). We also employ two questions that serve as attention check and each question is correctly answered by about 75% of the respondents. We show results here for all respondents that answered both questions correctly which is about 60% of the original sample. The same tables, but based on all respondents, are shown in the appendix (see subsection 6.5) and there is no substantive difference.

3.2 The Elicitation Step: Comparing Five Question Formats

We compare five question formats that differ in terms of structure and cognitive demand. Here, we only discuss each format briefly but the precise question wording can be found in the appendix (subsection 6.1).

Interval Question (1a & 1b). The first and second question formats are similar and come in a wide (# 1a) and a narrow (# 1b) version. We ask people to indicate the most likely value. In addition, respondents are asked to provide the probability that a vote outcome will be less than 40% (45% in the narrow format) and the probability of more than 60% (55% in the narrow) format.

Quantile Question (#2). The second question format asks people to identify the median. It then asks to imagine that the value is below and what the median for values below the median are, i.e. it asks for the first quartile. It then does the same but asks respondents to imagine that a value is above the median. It ends by showing people their three responses (P_{25}, P_{50}, P_{75}) and asks them whether they think that a random draw is equally likely to fall into any of these intervals: $0-P_{25}$, $P_{25}-P_{50}$, $P_{50}-P_{75}$, $P_{75}-1$. Respondent can then adapt their responses if they want to do so. Here, the forth step is possible and respondents can assess the adequacy of their responses.

Manski Question (#3). The third format relies on work by Manski (2009). After asking for the most likely value, respondents are asked to supply the lowest possible value and the highest possible value they expect. After that they are asked to estimate the probability that a random draw would be below (above) the lower (higher) bound. This has already been used in political science to elicit priors (Gill and Freeman, 2013).

Bins and Balls Question Format (#4). The last question is the latest addition to elicitation methods and exploits that a large amount of surveys are being carried out online and hence allow for new question formats. We implement a proposal by Goldstein and Rothschild (2014) where respondents have to place 100 balls into bins indicating a specific range (see Figure 4). By clicking on the + and − symbols balls are placed in a bin. Since respondents constantly see the implied distribution this is akin to an implicit adequacy check.

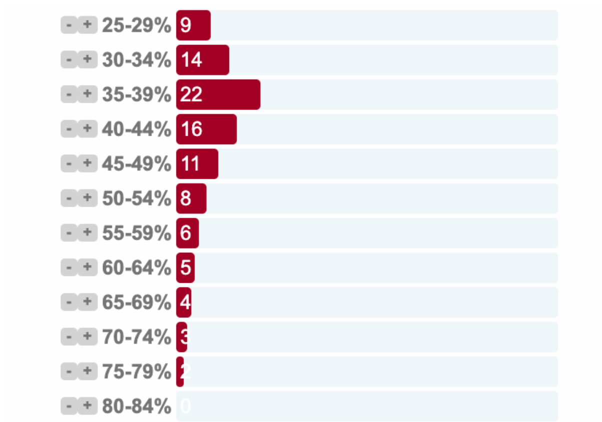

Figure 2: Screenshot of Balls and Bins. Illustration after hypothetical respondent has allocated all 100 balls.

All five question formats differ in their complexity for respondents but also in how easily they can be implemented. Some of these questions lend themselves to adding an adequacy check at the end, others do not. Table 1 allows to compare the different formats.

The number of questions is somewhat deceiving for the Bins & Balls format as it is one question but requires respondents to click 100 times to distribute the virtual balls.

Table 1: Question Formats

Question format	Most difficult concept R's need to know?	How many questions?	Adequacy check?
Interval (1a & 1b)	Quantiles	3	×
Quantile (2)	Median	4	✓
Manski (3)	Percentages	5	×
Bins & Balls (4)	Percentages	1	✓

The Interval Question as well as the Quantile Question are particularly demanding as they require an understanding of quantiles. The Manski is similar but can be expected to be less demanding since it translates the task well into easier terms (means, maximums, and minimus). Finally, the Bins and Balls Question requires the least of the respondents but is also the most demanding for researchers when it comes to implementation. The question formats further differ on whether they allow for an adequacy check. In the Quantile question, for example, respondents can incorrectly place the upper quartile below the median. This can give indication for the wrong understanding of the question. In the next section we investigate the accuracy of the elicited beliefs and discuss the experimental results.

3.3 The Fitting Step

To estimate a respondent's belief we assume a flexible parametric distribution for respondents' beliefs and estimate the parameters of the distribution such that it closely mimics the observed indicators from the different question formats. Because the sampling space of our experiment is bound between 0 and 1, we employ a Beta distribution as our parametric assumption. The Beta distribution has two shape parameters α, β. As an example we provide the derivation of the interval question format here. We present the derived likelihood functions for the other formats in the appendix (see subsection 6.2).

For the interval question we observe three values. Respondents report the mean value of their beliefs and the probabilities to observe a value below and above a certain threshold. We denote the mean with y_i and the two ($k \in (1,2)$) probabilities with p_{i1}, p_{i2}. The interval values depend on the question format and are denoted with $c = [c_1, c_2]$, where in the wide version $c = [40\%, 60\%]$ and in the narrow version $c = [45\%, 55\%]$. We assume that the values are measured with normal measurement error.[4]

$$y_i \sim \mathcal{N}(\mu_y, \sigma_y^2) \tag{1}$$

$$p_{i1} \sim \mathcal{N}(\mu_{p_1}, \sigma_p^2) \tag{2}$$

$$p_{i2} \sim \mathcal{N}(\mu_{p_2}, \sigma_p^2) \tag{3}$$

The expectations μ_y are calculated from the assumed parametric belief distribution. Here we use the same distribution as in the data generating process - a beta distribution. The beta distribution is relatively flexible and well-suited for our example with vote shares being constrained on the unit interval. The expectation for the mean from the beta are given by the two shape parameter α and β:

$$\mu_y = \frac{\alpha}{\alpha + \beta} \tag{4}$$

The expected probabilities are given by the CDF of the beta distribution, which we denote with $Q(\cdot, \alpha, \beta)$.

$$\mu_{p_1} = Q(c_1, \alpha, \beta) \tag{5}$$

$$\mu_{p_2} = 1 - Q(c_2, \alpha, \beta) \tag{6}$$

With this model we can define the Likelihood for the observed data of N respondents $Y = [[y_1, p_{i1}, p_{i2}]', \ldots, [y_N, p_{N1}, p_{N2}]']'$. We assume that all responses are identically and

[4]Including a measurement model extends existing approaches that only minimize the squared error between observed and theoretical expected value (e.g. Morris et al., 2014). This can open up the modeling framework for a set of extensions, e.g. correlated and heteroscedastic errors, hierarchical structures, and Bayesian estimation.

independently distributed, which gives the following Likelihood:

$$L(\alpha, \beta, \sigma_y^2, \sigma_p^2 \mid \boldsymbol{Y}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_y^2}} exp\left[\frac{-\left(y_i - \mu_y\right)^2}{2\sigma_y^2}\right] \prod_{k=1}^{2} \frac{1}{\sqrt{2\pi\sigma_p^2}} exp\left[\frac{-\left(p_k - \mu_{p_k}\right)^2}{2\sigma_p^2}\right]. \qquad (7)$$

To obtain MLE estimates of the parameters, the function is also maximized using R's `optim` function. The obtained estimates yield an estimate of the average beliefs in a specific condition. The goal is to identify the question format that will yield estimates closest to the true values.

3.4 The Adequacy Check Step

Assessing the adequacy of the elicited distribution by giving respondents the chance to review and correct their belief distribution is difficult, because the fitting is done 'outside' of the survey software and only after the answers have been collected. But for some formats it is still possible to provide the opportunity for correction using question filters based on respondents' answers or visual question formats. The Quantile Question for instance presents respondents with the quartiles they provided and asks if election results are equally likely to fall in each.[5] The Bins and Balls asks respondents to 'draw' their distribution and thus gives immediate feedback.

4 The Results

To evaluate the different question formats one can now compare the elicited beliefs with the true objective distribution. In Figure 6 you can see for all four combinations the true distribution and the elicited beliefs based on the various question formats. The four columns are for the different conditions (small/large variance and symmetric/asymmetric). The five rows are for the five different elicitation methods. While we have symmetric and asymmetric true distributions, we argue that the asymmetric cases are more relevant. The reason for this is that any Beta distribution will be asymmetric with the sole

[5] In Appendix 6.4 we compare the results for the Quantile question with and without the adequacy check and find that the adequacy check can improve the result considerably.

exception of those distributions were both shape parameters are equal to each other.

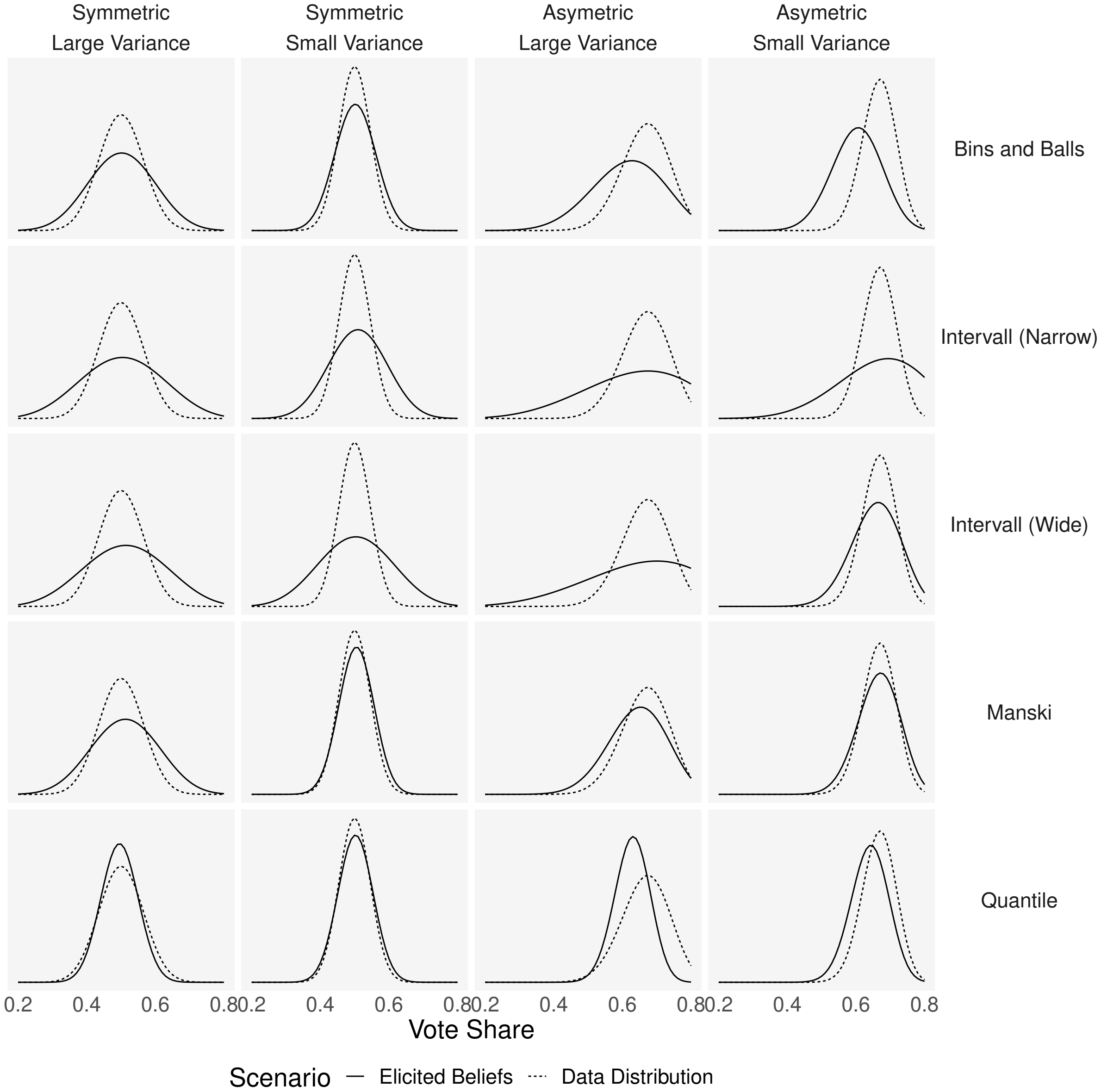

Figure 3: Comparison of Question Formats. The dotted line indicates the true distribution and the black solid line shows the average of the elicited distributions.

When looking at Figure 6 we see that most question formats are unbiased when the true distribution is symmetric. In the appendix we also present the same figure but with the average elicited belief drawn (see Figure 5). When turning our attention to the second moment we see that the interval questions tend to provide too wide beliefs. Looking at the asymmetric distributions we see that there is some bias towards $\frac{1}{2}$ but the extent varies across questions formats. As with the symmetric distributions, the interval questions tend to provide too wide beliefs.

From quickly eyeballing the visualization it seems that the Manski question as well as the quantile question are closest to the true distributions. To be able to clearly evaluate the question formats we now turn to the results in Table 2 where we show one table for each combination of the experimental factors. While we show here the average results we also present the individual results in the appendix (see subsection 6.6). Table 2 shows the implied parameters of the average elicited priors are shown as well as the Kullback-Leibler divergence, the number of observations, and the p-value of the likelihood-ratio test on whether the estimated parameters differ from the true parameter values. The smaller the KL divergence the closer is the elicited prior to the true distribution.

method	alpha	beta	KL	lr	N
Quantile	42.35	43.10	0.04	0.44	65
Bins and Balls	13.59	13.51	0.13	0.00	60
Manski	13.01	12.41	0.15	0.00	62
Interval (Narrow)	8.54	8.41	0.28	0.00	61
Interval (Wide)	8.68	8.27	0.29	0.00	69

(a) Symmetric, Large Variance

method	alpha	beta	KL	lr	N
Quantile	48.41	47.90	0.01	0.69	119
Manski	48.75	47.72	0.02	0.39	112
Bins and Balls	35.67	35.38	0.06	0.00	107
Interval (Narrow)	18.02	17.35	0.27	0.00	126
Interval (Wide)	11.03	10.88	0.45	0.00	115

(c) Symmetric, Small Variance

method	alpha	beta	KL	lr	N
Manski	19.97	11.07	0.06	0.00	65
Bins and Balls	12.75	7.96	0.23	0.00	48
Interval (Narrow)	5.95	3.39	0.41	0.00	68
Interval (Wide)	5.39	2.88	0.46	0.00	65
Quantile	55.23	32.69	0.47	0.00	65

(b) Asymmetric, Large Variance

method	alpha	beta	KL	lr	N
Manski	38.63	19.36	0.04	0.18	112
Interval (Wide)	28.39	14.79	0.11	0.00	120
Quantile	49.08	27.64	0.13	0.00	118
Bins and Balls	27.21	17.98	0.50	0.00	121
Interval (Narrow)	9.38	4.69	0.52	0.00	133

(d) Asymmetric, large Variance

Table 2: Estimates for the Elicited Beliefs

As seen in the visual inspection, the quantile question format and the Manski question seem to outperform the alternatives. If we average over all four experimental settings the Manski question has the smallest value for the Kullback-Leibler divergence followed by the quantile question. As mentioned above, the asymmetric cases are more frequent and the Manski question format beats all other alternatives for this format. In addition, the Manski question gives comparable estimates for political interested and not very interested participates. In Appendix 6.7 we describe this and show that, especially, for the Quantile question the estimates diverge for different sub-groups.

Based on these experiments and the two criteria, KL divergence and LR test, we conclude that the Manski question format outperforms the other alternatives. The Manski format provides a fairly simple approach that only requires five simple questions to elicit the prior beliefs of respondents. Unlike other formats tested here is it also straight-forward to implement as one can rely on standard questions. But even this best format is not free of shortcomings and fails to always perfectly reproduce the true underlying distribution. For large variances there was a small downward bias in the asymmetric case.

An open question is whether these results are sensitive to the precise sample composition. Do results only hold when we exclude individuals that fail attention checks (Berinsky et al., 2014)? We use two screening questions to detect inattentive respondents; one asks respondents to recall the founding organization mentioned in the introduction text and the second question ask them to recall the colors of the plots in the GIF (see Figure 1). Especially the latter is relevant as it is directly tied to the communication of the true distribution. The short answer is yes – the results hold up on the raw sample. Results are presented in subsection 6.5.

On a final note, we find only limited evidence for systematic biases in respondents' beliefs. In particular, we only observe clear over-confidence, i.e. the tendency that respondents are more certain than the objective data would warrant and therefore give variances that are too narrow, for the Quantile question. For all other formats, respondents actually express beliefs that are *less* certain than the objective benchmark would demand. This suggests that this bias may actually be driven by the elicitation method rather than being a systematic inability of respondents. However, respondents do exhibit problems with skewed distributions, although to a lesser degree when using interval questions.

5 Conclusion

This research note empirically validated five different question formats for prior elicitation. For each format we derive the estimators that allow to recover the shape parameters describing the beliefs. We run a various experiments to be able to compare the performance of these elicitation methods and find that a set of questions proposed by Manski (2009) performs very well.

This is good news for applied researchers. While all five are fairly easy to implement, the Manski questions are especially straight-forward as they only consist of asking people for numbers (e.g. mean or percentage of values below a specific threshold). There is no need for programing as in other question formats (the Bins and Balls question format requires programing in java script). In addition, the Manski format seemed to perform similarly across different subgroups which can be a relevant consideration. Finally, there is one caveat that needs to be mentioned: most of these methods assume a unimodal distribution. While this is a reasonable assumption in many circumstances it is possible that one would want to elicit beliefs that may be multimodal. In such a situation the Bins and Balls format would allow researchers to do so.

References

Alvarez, R. Michael, and John Brehm. 1997. "Are Americans Ambivalent Towards Racial Policies?" *American Journal of Political Science* 41: 345–374.

Bartels, Larry M. 2002. "Beyond the running tally: Partisan bias in political perceptions." *Political behavior* 24(2): 117–150.

Berinsky, Adam J. 2017. "Rumors and health care reform: Experiments in political misinformation." *British Journal of Political Science* 47(2): 241–262.

Berinsky, Adam J, Gregory A Huber, and Gabriel S Lenz. 2012. "Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk." *Political Analysis* 20(3): 351–368.

Berinsky, Adam J, Michele F Margolis, and Michael W Sances. 2014. "Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys." *American Journal of Political Science* 58(3): 739–753.

Bullock, John G. 2009. "Partisan bias and the Bayesian ideal in the study of public opinion." *The Journal of Politics* 71(3): 1109–1124.

Coppock, Alexander. 2018. "Generalizing from survey experiments conducted on mechanical Turk: A replication approach." *Political Science Research and Methods* pp. 1–16.

Cosmides, Leda, and John Tooby. 1996. "Are Humans Good Intuitive Statisticians after all? Rethinking some Conclusions from the Literature on Judgment under Uncertainty." *cognition* 58(1): 1–73.

Garthwaite, Paul H, Joseph B Kadane, and Anthony O'Hagan. 2005. "Statistical methods for eliciting probability distributions." *Journal of the American Statistical Association* 100(470): 680–701.

Gerber, Alan, and Donald Green. 1999. "Misperceptions About Perceptual Bias." *Annual Review of Political Science* 2(1): 189–210.

Gill, Jeff, and John R Freeman. 2013. "Dynamic elicited priors for updating covert networks." *Network Science* 1(1): 68–94.

Gill, Jeff, and Lee D Walker. 2005. "Elicited priors for Bayesian model specifications in political science research." *The Journal of Politics* 67(3): 841–872.

Goldstein, Daniel G, and David Rothschild. 2014. "Lay understanding of probability distributions." *Judgment & Decision Making* 9(1).

Kynn, Mary. 2008. "The 'heuristics and biases' bias in expert elicitation." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171(1): 239–264.

Leiter, Debra, Andreas Murr, Ericka Rascón Ramirez, and Mary Stegmaier. 2018. "Social Networks and Citizen Election Forecasting: The More Friends the Better." *International Journal of Forecasting* 34(2): 235–248.

Manski, Charles F. 2009. *Identification for prediction and decision.* Harvard University Press.

Mason, Winter, and Siddharth Suri. 2012. "Conducting behavioral research on Amazon's Mechanical Turk." *Behavior research methods* 44(1): 1–23.

Morris, David E, Jeremy E Oakley, and John A Crowe. 2014. "A web-based tool for eliciting probability distributions from experts." *Environmental Modelling & Software* 52: 1–4.

Murr, Andreas. 2011. "'Wisdom of crowds'? A decentralised election forecasting model that uses citizens' local expectations." *Electoral Studies* 30(4): 771–783.

Mutz, Diana C. 2011. *Population-Based Survey Experiments.* Princeton University Press.

Peterson, Cameron, and Alan Miller. 1964. "Mode, median, and mean as optimal strategies." *Journal of Experimental Psychology* 68(4): 363.

Savage, Leonard J. 1971. "Elicitation of personal probabilities and expectations." *Journal of the American Statistical Association* 66(336): 783–801.

Schlag, Karl H, James Tremewan, and Joël J Van der Weele. 2015. "A penny for your thoughts: a survey of methods for eliciting beliefs." *Experimental Economics* 18(3): 457–490.

Spetzler, Carl S, and Carl-Axel S Stael von Holstein. 1975. "Exceptional paper—Probability encoding in decision analysis." *Management science* 22(3): 340–358.

Thomas, Kyle A, and Scott Clifford. 2017. "Validity and mechanical turk: An assessment of exclusion methods and interactive experiments." *Computers in Human Behavior* 77: 184–197.

Tversky, Amos, and Daniel Kahneman. 1971. "Belief in the Law of Small Numbers." *Psychological Bulletin* 76(2): 105.

Tversky, Amos, and Daniel Kahneman. 1973. "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology* 5(2): 207–232.

Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under uncertainty: Heuristics and biases." *science* 185(4157): 1124–1131.

Wallsten, Thomas S, and David V Budescu. 1983. "State of the Art – Encoding Subjective Probabilities: A Psychological and Psychometric Review." *Management Science* 29(2): 151–173.

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

6 ONLINE-APPENDIX

6.1 Question Formats

- **Quantile Question**

  1. We just showed you election results for similar districts. Now, we want to know what your expectations are.

     Can you determine the median? This is the value, where vote share of party A is equally likely to be less than or larger than this value.
  2. Imagine you were told that the actual result was below your median value.

     Can you determine a new value, such that the vote share of party A is equally likely to be *between 0 percent and the new value* or *between the new value and the median value*?
  3. Imagine you were told that the actual result was above your median value.

     Can you determine yet another value, such that the vote share of party A is equally likely to be *between the median and this new value* or *between this new value and 100 percent*?
  4. At the end respondents are shown the four ranges (0-25th, 25th-50th, 50th-75th, 75th-100) and asked whether a random draw is equally likely to occur in each of them. If not, respondents can go and adjust their responses.

     (a) Consider the following four intervals: $[0,P_{25}]$, $[P_{25}, P_{50}]$, $[P_{50},P_{75}]$, $[P_{75}, 100]$.
     Is it equally likely that party A's vote share will fall in any of these intervals?
     (P_{XY} indicates the respondent's XY percentile.)

- **Wide & narrow Question** This question comes in two versions – a wide and a narrow version.

  1. We just showed you election results for similar districts. Now, we want to know what your expectations for party A's vote share are.

     What is the most likely vote share of party A? Please give your response in percentage points.
  2. What is the probability that party A will receive a vote share of less than 40 percent? (*45% in narrow format*)
  3. What is the probability that party A will receive a vote share of more than 60 percent? (*55% in narrow format*)

- **Manski Question**

  1. We just showed you election results for similar districts. Now, we want to know what your expectations for party A's vote share are.

     What is the most likely vote share of party A? Please give your response in percentage points.
  2. What do you think is a likely range of the vote share that party A will receive? Please indicate the lower bound in percentage points.

3. Now, please indicate the upper bound in percentage points.
4. What is the probability that party A will get a vote share of less than (*lower value indicated by R*) percent?
5. What is the probability that party A will get a vote share of more than (*upper value indicated by R*) percent?

- **Goldstein and Rothschild Question**

  We implemented this question by inserting Java script into the survey software - a step that should be easily replicated by anybody. We also provide the JS code here (link) and have annotated it such that it can be easily adapted.

  Figure 4 shows the full question as it is presented to respondents. By clicking on + and − buttons the various bins can be filled or emptied. Each respondent allocates 100 balls into these bins.

Figure 4: Screenshot of Balls and Bins

6.2 Estimation

To estimate the beliefs form the different question formats we develop statistical models that permit us to estimate the parameters of respondent's belief distributions. In the following we describe the Likelihoods that model the observed outcomes given parametric belief distributions for the different question formats.

6.2.1 Likelihood for the Quantile Question

For the quantile question we observe three outcomes for each respondent: The lower quartile, the median value, and the upper quartile of the respondent's belief. We denote them with $y_i = [y_{i1}, y_{i2}, y_{i3}]$, where $i \in (1, \ldots, N)$ are respondents and k refers to the different quartile questions $k \in (1,2,3)$. We assume that these values are observed with measurement error, such that:[6]

$$y_{ik} \sim \mathcal{N}(\mu_{ik}, \sigma^2). \tag{8}$$

To estimate the average belief of respondent's we assume a parametric distribution and estimate the parameters of the distribution to closely mimics the expected observed indicators. To map the beliefs the the measured values, we require the quantile function of the belief distribution. Because the sampling space of our experiment is bound between 0 and 1, we employ a Beta distribution as our parametric distribution. We dneote $Q^{-1}(q_k, \alpha, \beta)$ as the quantile function of the beta distribution. The two shape parameters α, β define the expectation and the variance of the belief. The distribution is than linked to the expectation of the observed values. If we denote the three quartiles with $q_k = [0.25, 0.5, 0.75]$, we can write:

$$\mu_k = Q^{-1}(q_k, \alpha, \beta) \tag{9}$$

With this model we define the likelihood for the observed answers for a respondent as:

$$L(\alpha, \beta, \sigma^2 \mid y_i) = \prod_{k=1}^{3} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[\frac{-\left(y_{ik} - \mu_k\right)^2}{2\sigma^2}\right] \tag{10}$$

Maximizing the Likelihood for each individual would involve to minimize the squared distance between the observe quartile measurements and the shape parameters of the beta-distribution that generate the expected quartiles. If we assume that individual responses are identical and independently distributed, we can further write the Likelihood for the full sample as:

$$L(\alpha, \beta, \sigma^2 \mid \boldsymbol{Y}) = \prod_{i=1}^{N} \prod_{k=1}^{3} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[\frac{-\left(y_{ik} - Q^{-1}(q_k, \alpha, \beta)\right)^2}{2\sigma^2}\right], \tag{11}$$

where $\boldsymbol{Y}$ is $(N \times K)$ matrix with all respondents' responses $[y_1, \ldots, y_N]'$. The function is maximized with respect to the parameters α, β, σ using R's optim function.

6.2.2 Likelihood for the Interval Question

For the interval question we observe three values. Respondents report the mean value of their beliefs and the probabilities to observe a value below and above a certain threshold. We denote the mean with y_i and the two ($k \in (1,2)$) probabilities with p_{i1}, p_{i2}. The interval values depend on the question format and are denoted with $c = [c_1, c_2]$, where in the wide version $c = [40\%, 60\%]$ and in the narrow version $c = [45\%, 55\%]$. We assume that the values are measured with normal measurement error.

[6]We assume that the measurement errors are normal distributed with the same error variance and no covariance between the errors.

$$y_i \sim \mathcal{N}(\mu_y, \sigma_y^2) \tag{12}$$
$$p_{i1} \sim \mathcal{N}(\mu_{p_1}, \sigma_p^2) \tag{13}$$
$$p_{i2} \sim \mathcal{N}(\mu_{p_2}, \sigma_p^2) \tag{14}$$

The expectations μ_y are calculated from the assumed parametric belief distribution. Here we use the same distribution as in the data generating process - a beta distribution. The beta distribution is relatively flexible and well-suited for our example with vote shares being constrained on the unit interval. It is general possible to use other parametric distributions, like a normal distribution, instead. In practical applications it would be sensible to try different distributions and compare their relative fit. The expectation for the mean from the beta are given by the two shape parameter α and β:

$$\mu_y = \frac{\alpha}{\alpha + \beta} \tag{15}$$

The expected probabilities are given by the CDF of the beta distribution, which we denote with $Q(\cdot, \alpha, \beta)$.

$$\mu_{p_1} = Q(c_1, \alpha, \beta) \tag{16}$$
$$\mu_{p_2} = 1 - Q(c_2, \alpha, \beta) \tag{17}$$

With this model we can define the Likelihood for the observed data of N respondents $Y = [[y_1, p_{i1}, p_{i2}]', \ldots, [y_N, p_{N1}, p_{N2}]']'$. We assume that all responses are identically and independently distributed, which gives the following Likelihood:

$$L(\alpha, \beta, \sigma_y^2, \sigma_p^2 \mid \boldsymbol{Y}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_y^2}} exp\left[\frac{-(y_i - \mu_y)^2}{2\sigma_y^2}\right] \prod_{k=1}^{2} \frac{1}{\sqrt{2\pi\sigma_p^2}} exp\left[\frac{-(p_k - \mu_{p_k})^2}{2\sigma_p^2}\right]. \tag{18}$$

To obtain MLE estimates of the parameters, the function is also maximized using R's `optim` function. The obtained estimates yield an estimate of the average beliefs in a specific condition. The goal is to identify the question format that will yield estimates closest to the true values.

6.2.3 Likelihood for Manski Question

For the Manski Question we observe five measures of respondents' belief. We measure three $k \in 1, 2, 3$ values: the mean value (which we denote with y_{i1}), and the lower and upper bound values (which we denote with y_{i2} and y_{i3}). In addition we measure two probabilities of observing values below and above the bounds (which we denote with p_{i1} and p_{i2}). We assume that the values are measured with error and assume that the errors are identical and independently normal distributed.

$$y_{ik} \sim \mathcal{N}(\mu_k, \sigma^2) \tag{19}$$

the expectations μ_k are calculated from the assumed parametric distribution of respondents' belief. In our analysis we work with the beta distribution, which yields a simple expression for the mean value. The observed lower and upper bounds, given the assigned probabilities, can be calculated from the quantile function of the Beta distribution, which we denote as $Q^{-1}(\cdot, \alpha, \beta)$. The expectations of the measurement model are given by:

$$\mu_{i1} = \frac{\alpha}{\alpha + \beta} \tag{20}$$

$$\mu_{i2} = Q^{-1}(p_{i1}, \alpha, \beta) \tag{21}$$

$$\mu_{i3} = 1 - Q^{-1}(p_{i2}, \alpha, \beta) \tag{22}$$

The Likelihood is given by the normal measurement error and the respective expectation generating functions. We collapse the measured values and the probabilities in a matrix ($\boldsymbol{Y} = [[y_{11}, y_{12}, y_{13}, p_{11}, p_{12}]', \ldots, [y_{N1}, y_{N2}, y_{N3}, p_{N1}, p_{N2}]']'$. Assuming that the observed values are independent permits us to write the Likelihood as:

$$L(\alpha, \beta, \sigma \mid \boldsymbol{Y}) = \prod_{i=1}^{N} \prod_{k=1}^{3} \frac{1}{\sqrt{2\pi\sigma^2}} exp \left[\frac{-\left(y_{ik} - \mu_{ik}\right)^2}{2\sigma^2} \right], \tag{23}$$

which is numerically maximized with respect to the parameters using R's optim function.

6.2.4 Likelihood for Bins and Ball Question

The bins and balls question has a slightly different structure compared to the quantile questions. For this question we observe the number of balls a respondent decides to place into K bins that each covers an exclusive interval. The intervals are given by ordered cut points $c_1, \ldots, c_C$. There is one cut-point more than categories $C = K + 1$ as the question format can have lower and upper bounds.[7] The number of balls out of $B = 100$ that a respondent places in a bin is denoted with y_{ik}. We assume that the measured placements are binomial distributed, with a certain probability π_k.

$$y_{ik} \sim \mathcal{B}(\pi_k, B) \tag{24}$$

The probabilities are calculated from the CDF of the assumed parametric belief distribution. The CDF of the Beta distribution is given by $Q(\cdot, \alpha, \beta)$. With this we calculate the probability that a respondent places balls in each bin, as:

$$\pi_k = Q(c_{k+1}, \alpha, \beta) - Q(c_k, \alpha, \beta). \tag{25}$$

Assuming that the observed values are conditionally independent, and combining all observed placements $\boldsymbol{Y} = [[y_{11}, \ldots, y_{1K}]', \ldots, [y_{N1}, \ldots, y_{NK}]']'$ gives the following Likelihood:

[7] C is the number of cut-points, which in our question format is $C = 13$ and the corresponding cut points are $0.25, 0.3, 0.35, 0.4, 0.45, \ldots, 0.85$.

$$L(\alpha, \beta \mid \boldsymbol{Y}) = \prod_{i=1}^{N} \prod_{k=1}^{K} \binom{B}{y_{ik}} \pi_k^{y_{ik}} (1 - \pi_k)^{B - y_{ik}} \tag{26}$$

We numerically maximize the likelihood to obtain MLE estimates of the shape parameters.

6.3 Balance

The following four tables present balance checks in terms of covariate averages and standard deviations for each experimental condition. These checks are based on the full data before reducing the data set to only observations which passed both attention checks. Please note that we refrain from the ill-advised practice of statistically *testing* for mean differences (Mutz, 2011). Overall, we find treatment conditions to be well balanced.

Format	n	Female	Age	University	Political Interest
Quantile	196	0.45	41.40	0.61	3.03
		(0.50)	(11.73)	(0.49)	(0.76)
Interval (Wide)	205	0.44	40.33	0.60	3.00
		(0.50)	(13.76)	(0.49)	(0.82)
Interval (Narrow)	205	0.48	41.39	0.57	3.09
		(0.50)	(12.26)	(0.50)	(0.78)
Manski	201	0.49	41.69	0.60	3.04
		(0.50)	(11.64)	(0.49)	(0.81)
Bins and Balls	189	0.49	40.81	0.61	2.99
		(0.50)	(11.63)	(0.49)	(0.81)

Table 3: Balance Check for Symmetric Distribution. Means and Standard Deviations in Parentheses.

Format	n	Female	Age	University	Political Interest
Quantile	100	0.43	38.04	0.67	2.93
		(0.50)	(11.74)	(0.47)	(0.84)
Interval (Wide)	102	0.40	39.29	0.53	2.92
		(0.49)	(12.29)	(0.50)	(0.86)
Interval (Narrow)	97	0.40	40.66	0.60	2.95
		(0.49)	(12.66)	(0.49)	(0.85)
Manski	106	0.42	41.20	0.56	2.93
		(0.50)	(12.39)	(0.50)	(0.80)
Bins and Balls	102	0.41	41.35	0.65	3.08
		(0.49)	(13.58)	(0.48)	(0.83)

Table 4: Balance Check for Symmetric Distribution with Large Variance. Means and Standard Deviations in Parentheses

Format	n	Female	Age	University	Political Interest
Quantile	201	0.48	41.29	0.63	3.00
		(0.50)	(12.24)	(0.48)	(0.73)
Interval (Wide)	197	0.40	40.43	0.59	3.04
		(0.49)	(11.61)	(0.49)	(0.77)
Interval (Narrow)	206	0.47	39.47	0.59	3.02
		(0.50)	(12.13)	(0.49)	(0.80)
Manski	203	0.46	39.93	0.59	3.04
		(0.50)	(12.30)	(0.49)	(0.78)
Bins and Balls	196	0.43	41.44	0.64	3.10
		(0.50)	(12.07)	(0.48)	(0.80)

Table 5: Balance Check for Asymmetric Distribution. Means and Standard Deviations in Parentheses.

Format	n	Female	Age	University	Political Interest
Quantile	102.00	0.54	41.49	0.72	3.09
		(0.50)	(12.40)	(0.45)	(0.76)
Interval (Wide)	104.00	0.44	38.80	0.72	2.88
		(0.50)	(10.72)	(0.45)	(0.75)
Interval (Narrow)	101.00	0.45	39.73	0.65	2.99
		(0.50)	(12.18)	(0.48)	(0.83)
Manski	98.00	0.48	41.00	0.68	2.91
		(0.50)	(13.26)	(0.47)	(0.90)
Bins and Balls	95.00	0.53	42.15	0.49	3.00
		(0.50)	(12.26)	(0.50)	(0.77)

Table 6: Balance Check for Asymmetric Distribution with Large Variance. Means and Standard Deviations in Parentheses.

6.4 Evaluating Adequacy Check for Quantile Question

Variance	Distribution	AdequacyCheck	alpha	beta	KL	N
Large Variance	Asymmetric	Yes	49.08	27.64	0.13	118
Large Variance	Asymmetric	No	158.50	91.07	0.98	130
Large Variance	Symmetric	Yes	48.41	47.90	0.01	119
Large Variance	Symmetric	No	34.21	34.59	0.07	115

Table 7: Estimates for the Elicited Beliefs. Comparing Quantile with and without Adequacy Check for the Large Variance Scenarios

6.5 No Screening

These two tables present the same information as is shown in Table 2. The data here is based on all results, i.e the raw data before reducing the data set to only observations which pass both attention checks.

method	alpha	beta	KL	lr	N
Quantile	45.54	47.29	0.07	0.14	100
Bins and Balls	15.27	15.04	0.10	0.00	102
Manski	13.57	13.34	0.13	0.00	106
Interval (Wide)	8.80	8.53	0.27	0.00	97
Interval (Narrow)	7.13	6.78	0.36	0.00	102

(a) Symmetric, Large Variance

method	alpha	beta	KL	lr	N
Quantile	59.69	60.15	0.00	0.85	196
Manski	50.73	49.12	0.02	0.04	201
Bins and Balls	26.76	26.82	0.13	0.00	189
Interval (Wide)	15.98	15.40	0.31	0.00	205
Interval (Narrow)	11.56	11.36	0.43	0.00	205

(c) Symmetric, Small Variance

method	alpha	beta	KL	lr	N
Manski	22.79	13.46	0.12	0.00	98
Bins and Balls	11.49	7.16	0.24	0.00	95
Interval (Wide)	5.82	3.34	0.42	0.00	101
Quantile	60.58	35.75	0.54	0.00	102
Interval (Narrow)	3.40	1.82	0.67	0.00	104

(b) Asymmetric, Large Variance

method	alpha	beta	KL	lr	N
Manski	38.79	20.29	0.05	0.00	203
Quantile	56.95	31.47	0.10	0.00	201
Bins and Balls	18.27	12.45	0.54	0.00	196
Interval (Wide)	7.85	3.98	0.60	0.00	206
Interval (Narrow)	6.56	3.12	0.70	0.00	197

(d) Asymmetric, large Variance

Table 8: Estimates for the Elicited Beliefs. No Screen

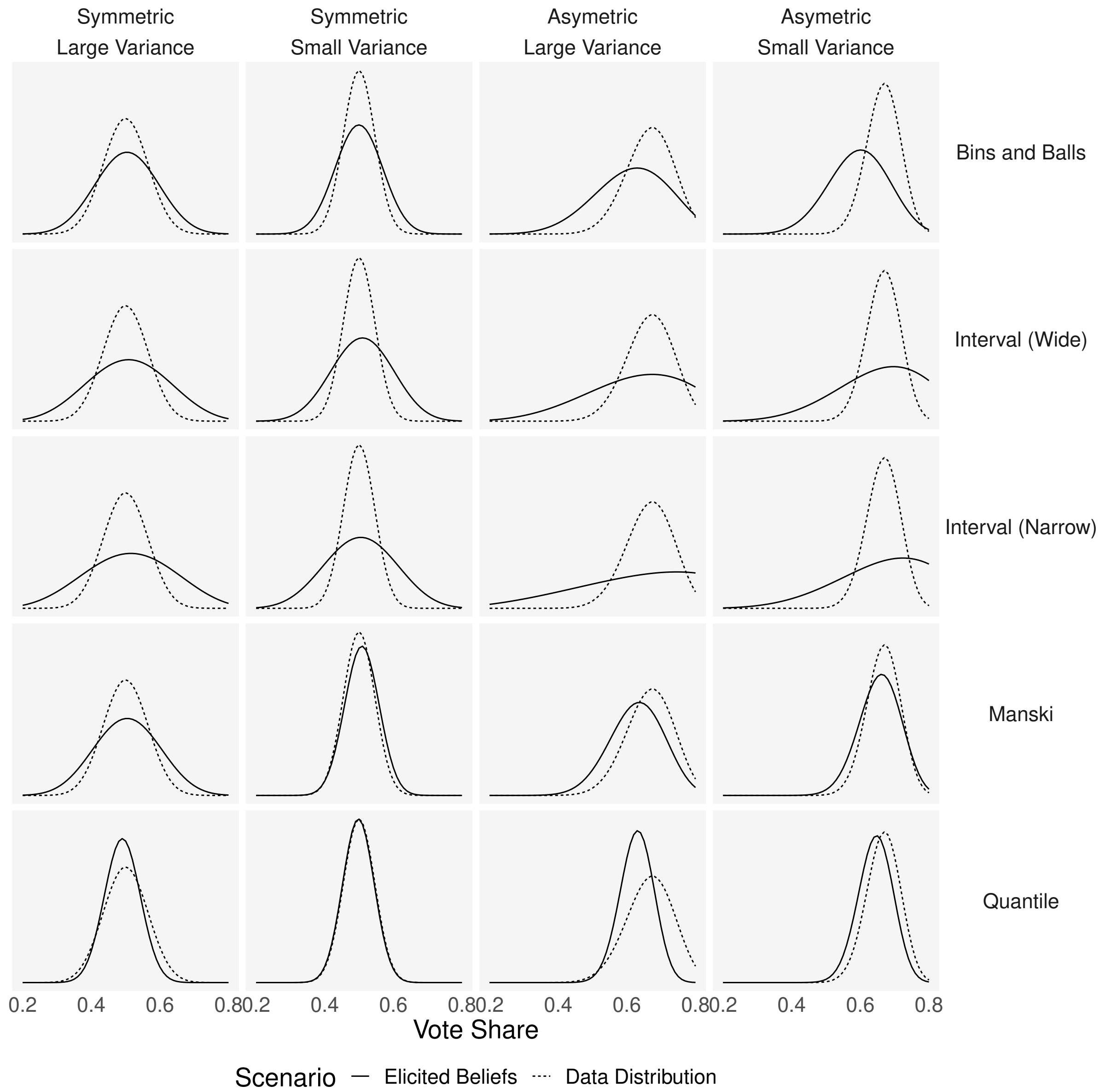

Figure 5: Comparison of Question Formats. No screen. The dotted line indicates the true distribution and the black solid line shows the average of the elicited distributions.

6.6 Individual Beliefs

The question formats and estimation method can also be used to obtain individual beliefs. We use th same Maximum Likelihood approach as described in section 6.2, but allow for individual shape parameters: $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_N]$ and $\boldsymbol{\beta} = [\beta_i, \ldots, \beta_N]$. To illustrate, consider the the Individual Likelihood for the Quantile question:

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma \mid \boldsymbol{Y}) = \prod_{i=1}^{N} \prod_{k=1}^{3} \frac{1}{\sqrt{2\pi\sigma^2}} exp \left[\frac{-\left(y_{ik} - Q^{-1}(q_k, \alpha_i, \beta_i)\right)^2}{2\sigma^2} \right], \tag{27}$$

where now we introduce a subscript for the shape parameters of the Beta quantile function $Q^{-1}(q_k, \alpha_i, \beta_i)$.

We obtain estimates for respondent specific shape parameters by numerically maximizing the Likelihood function. We first estimate the shape parameters for each respondent, and afterwards estimate the error variance terms for the Likelihood function. We repeat until convergence in the error variances.[8]

6.6.1 Results

type	method	median	qlow	qhigh
Symmetric	Bins and Balls	0.13	0.06	0.26
Symmetric	Interval (Narrow)	0.27	0.05	0.81
Symmetric	Manski	0.36	0.23	0.66
Symmetric	Interval (Wide)	0.45	0.14	1.33
Symmetric	Quantile	0.50	0.15	1.92

(a) Symmetric, Large Variance

type	method	median	qlow	qhigh
Symmetric	Bins and Balls	0.10	0.03	0.23
Symmetric	Manski	0.25	0.11	0.45
Symmetric	Interval (Wide)	0.36	0.12	1.27
Symmetric	Interval (Narrow)	0.37	0.11	1.11
Symmetric	Quantile	0.48	0.17	1.17

(c) Symmetric, Small Variance

type	method	median	qlow	qhigh
Asymetric	Bins and Balls	0.27	0.12	0.62
Asymetric	Manski	0.27	0.09	0.53
Asymetric	Quantile	0.51	0.16	1.48
Asymetric	Interval (Wide)	0.64	0.63	0.73
Asymetric	Interval (Narrow)	0.84	0.37	2.26

(b) Asymmetric, Large Variance

type	method	median	qlow	qhigh
Asymetric	Manski	0.15	0.05	0.38
Asymetric	Quantile	0.44	0.14	0.91
Asymetric	Interval (Wide)	0.96	0.95	0.96
Asymetric	Interval (Narrow)	0.96	0.95	0.96
Asymetric	Bins and Balls	0.99	0.55	1.54

(d) Asymmetric, large Variance

Table 9: KL Divergence for Individual Beliefs for different Scenarios

[8]Some responses patterns do not yield estimates of sensible shape parameters. For example, if a respondent reports a lower quartile of 0.50 and a Median of 0.45 the maximization of the function will be impossible. We exclude respondents with such inconsistent answering behaviors.

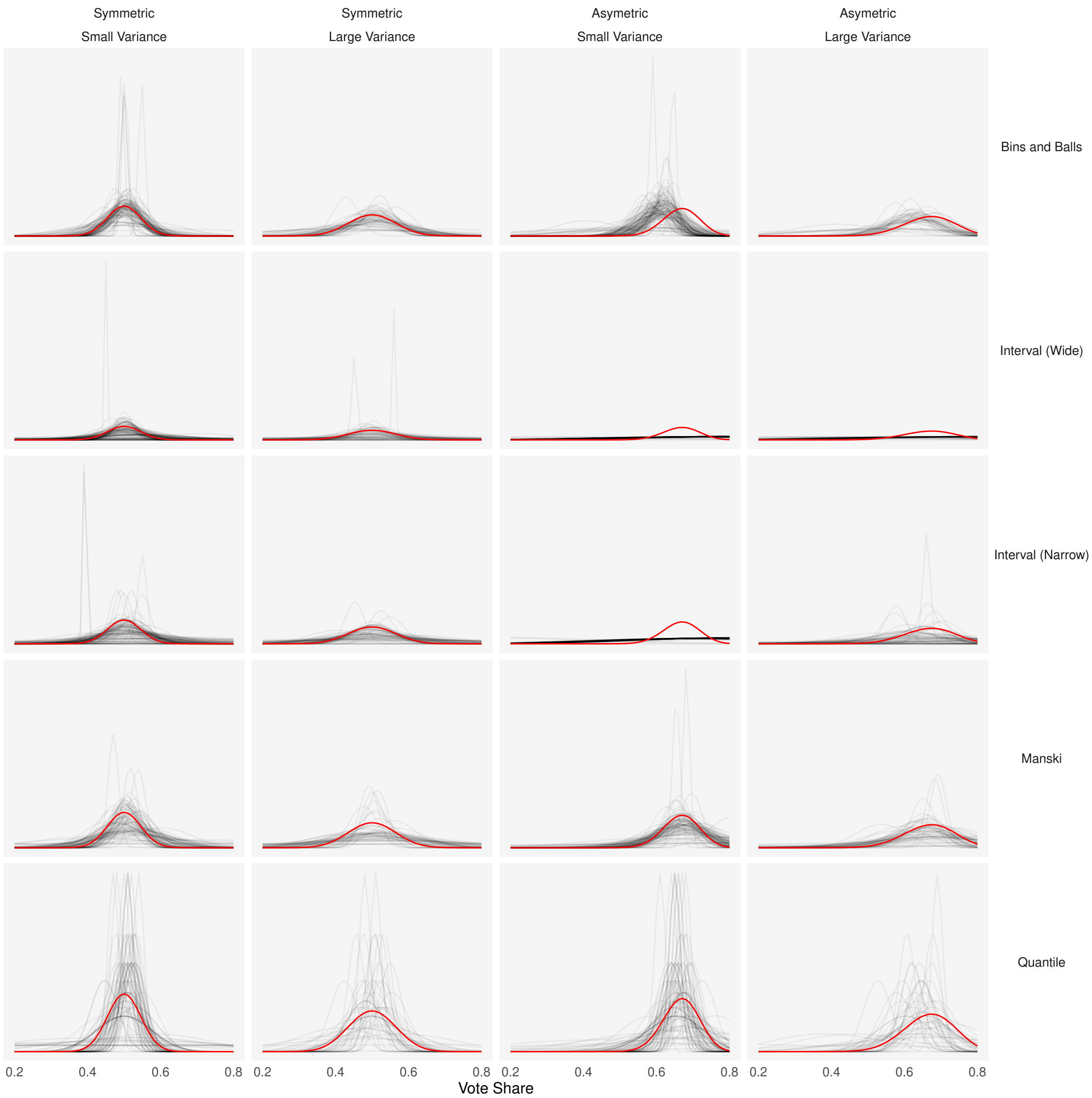

Figure 6: Individual Beliefs. The grey lines indicate individual elicited beliefs. The red line indicates the true distribution.

method	median	qlow	qhigh
Manski	0.25	0.09	0.49
Bins and Balls	0.30	0.09	0.90
Quantile	0.48	0.14	1.20
Interval (Wide)	0.92	0.43	0.96
Interval (Narrow)	0.93	0.25	0.98

Table 10: KL Divergence for Individual Beliefs over different Scenarios

6.7 Sub-Samples Political Interest

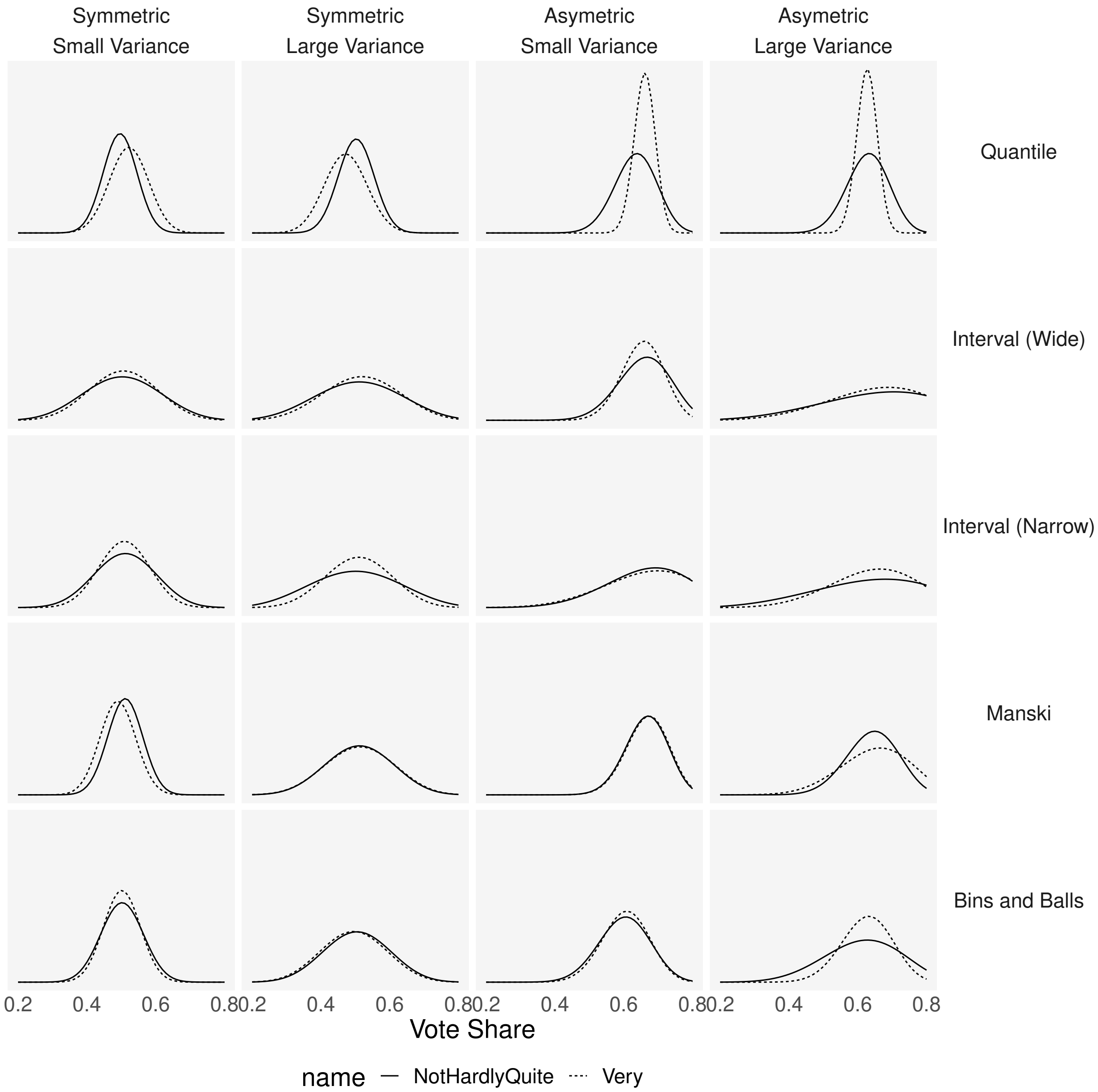

Figure 7: Estimated Beliefs for Sub-Groups of Political Interest

method	NotHardlyQuite_alpha	NotHardlyQuite_beta	NotHardlyQuite_KL	Very_alpha	Very_beta	Very_KL
Quantile	39.09	23.07	0.24	166.57	99.32	2.70
Interval (Wide)	5.01	2.68	0.49	6.75	3.60	0.36
Interval (Narrow)	5.04	2.90	0.48	9.27	5.17	0.25
Manski	25.16	14.05	0.06	13.69	7.38	0.12
Bins and Balls	10.98	6.91	0.26	26.79	16.02	0.16

(a) Symmetric, Large Variance

method	NotHardlyQuite_alpha	NotHardlyQuite_beta	NotHardlyQuite_KL	Very_alpha	Very_beta	Very_KL
Quantile	39.09	23.07	0.24	166.57	99.32	2.70
Interval (Wide)	5.01	2.68	0.49	6.75	3.60	0.36
Interval (Narrow)	5.04	2.90	0.48	9.27	5.17	0.25
Manski	25.16	14.05	0.06	13.69	7.38	0.12
Bins and Balls	10.98	6.91	0.26	26.79	16.02	0.16

(b) Asymmetric, Large Variance

method	NotHardlyQuite_alpha	NotHardlyQuite_beta	NotHardlyQuite_KL	Very_alpha	Very_beta	Very_KL
Quantile	51.63	52.38	0.01	40.24	36.87	0.12
Interval (Wide)	10.18	10.08	0.48	13.15	12.84	0.38
Interval (Narrow)	15.88	15.20	0.32	23.63	22.84	0.18
Manski	49.91	47.72	0.03	45.31	47.19	0.03
Bins and Balls	33.50	33.21	0.07	44.41	44.11	0.02

(c) Symmetric, Small Variance

method	NotHardlyQuite_alpha	NotHardlyQuite_beta	NotHardlyQuite_KL	Very_alpha	Very_beta	Very_KL
Quantile	39.07	22.54	0.17	159.98	82.26	0.38
Interval (Wide)	24.78	12.82	0.15	38.88	20.56	0.06
Interval (Narrow)	9.79	4.90	0.50	8.38	4.17	0.57
Manski	38.64	19.43	0.04	38.58	19.17	0.04
Bins and Balls	25.90	17.20	0.50	30.70	20.06	0.48

(d) Asymmetric, large Variance

Table 11: Estimates for Sub-groups of Political Interest