WORKING PAPERS IN ECONOMICS

No. 01/20

THOMAS DE HAAN

ELICITING BELIEF DISTRIBUTIONS USING A RANDOM TWO-LEVEL PARTITIONING OF THE STATE SPACE

Department of Economics
UNIVERSITY OF BERGEN

Eliciting belief distributions using a random two-level partitioning of the state space

Thomas de Haan*

6th January 2020

Abstract

I introduce a new method to incentivise the elicitation of belief distributions, the Random Partitions method. With this method, an agent's payoff not only depends on the realised state and the elicited distribution, but also on a randomised two-level partitioning of the state-space. The method creates a binary lottery payoff structure where reports closer to an agent's true belief distribution generate a higher probability to earn a high payout. The randomisation of the state-space partitioning ensures that the agent is incentivised to report correctly across the entire belief distribution. I compare the introduced Random Partitions method with both the well known Quadratic Scoring Rule, and a method based on the Becker-DeGroot-Marschak procedure and argue that the Random Partitions method gives substantially stronger truth-telling incentives to agents in situations where there are many states/bins.

Keywords: Belief elicitation, randomized state/event space partitioning, proper scoring rules.

*University of Bergen, Norway. I would like to thank Peter Wakker, Glenn Harrison, Daniel Friedman, Fred Schroyen, Theo Offerman, Randolph Sloof, Stefan Trautmann and participants of the 2019 Bayesian Crowd conference in Rotterdam for their helpful comments and suggestions.

1 Introduction

Eliciting the beliefs of an expert to better map uncertainty regarding a future event can be a valuable aid to making economic decisions. One could think of the distribution of the oil price at a future date, predictions of the future size of an industry or next year's inflation level in a country. To ensure that an agent reports his/her true beliefs, economists have developed several incentivised mechanisms to elicit beliefs and expectations (see e.g. Schlag et al. 2015 for a literature review). This includes a number of methods to elicit means, variances or most likely intervals (Schlag & van der Weele 2015) of a distribution.

When it comes to incentivised methods to elicit an agent's complete belief distribution, most methods are based on Proper Scoring Rules (Savage 1971) with as primary application the Quadratic Scoring Rule (Brier 1950 and e.g. Selten 1998) (QSR). Proper scoring rules can be used to elicit entire continuous belief distributions, but also allow for a (more practical) application where the state space is separated in a number of intervals, or "bins" (Matheson and Winkler 1976 and e.g. Harrison et al. 2017). One known alternative elicitation method is described by Karni (2009) which uses a randomisation mechanism similar to the Becker-DeGroot-Marschak (Becker et al. 1964) procedure to elicit a probability belief for one possible event, which could be used to separately elicit the likelihood of every possible event of an entire distribution.

Although proper scoring rules to elicit complete subjective belief distributions exist in the economics literature, their application to incentivise forcasters for their predictions appears to be rare in practice.[1] Among a number of reasons (see e.g. Armantier and Treich 2013), one critical disadvantage, I will identify and prove, plaguing the most used discrete scoring rule, the QSR, is that it's incentives strength decreases as the number of states over which a distribution is elicted gets larger. Eliciting a forcaster's belief distribution for a future inflation rate or commodity price will likely entail a state space divided into many 'bins', and a mechanism such as the QSR performs particularly poor in such an enviroment when it comes to providing sharp incentives (see also Harrison 1989 for an illustration on the importance of incentive sharpness in an experimental economics setting). The availability of an elictation method that can provide incentives at a level

[1]Harrison and Phillips (2014) report one example using the discrete version of the quadratic scoring rule with chief risk officers

which is robust to changing the fineness of the state-space "grid" could be a step forward to making proper elicitation methods a viable method to incentivize forcasters.

In this paper I will introduce a new method, the "Random Partitions" method, to incentivise the elicitation of a belief distribution. In the mechanism, the agent will either receive a payoff of zero or receive a fixed prize Y and her elicited response will (in part) determine the probability that the agent receives the payoff Y. This implies that the method falls in the class of "binary lottery procedures" which have the property that an agent's attitude towards risk will not influence the optimality of reporting her true beliefs. The probability that an agent will receive a payoff will depend on the elicited distribution and the realised state, but also on a random two-level partitioning of the state space. The principal will, before asking the agent for her beliefs, specify a procedure to draw a random two-level partitioning of the state space, which should be drawn independently of the elicited distribution. After the agent reports her beliefs, the principal (or an independent third party) draws and makes public a 'two-level partioning' of the state space. When the actual realisation of the state is publicly revealed, it will fall in one of the subsets of the partition of the state space. Conditional on the state belonging to one of the subsets of the first level of the partition, the elicited belief distribution makes a prediction which subset of the second level of the partition the true state most likely belongs to. If the realised state indeed falls in the subset the elicited distribution gives the most weight here, then the agent receives the reward Y and otherwise not.

The idea behind the Random Partitions method is that it poses a very large (possibly infinite) number of potential questions to the agent about which events she thinks have a higher likelihood to occur than others. Given that via the random partitioning and the uncertainty of the true state, the type of question the elicited belief distribution will have to answer is unknown, the Random Partitions method simultaneously poses many potential tests to the elicited belief distribution, ensuring that the agent is incentivised to report truthfully across the entire belief distribution.[2]

This paper will show for the Random partitions method that reporting one's true belief distribution is a best response for the agent and will specify under what conditions reporting a deviating

[2]In a recent paper, Chambers and Lambert (2019) appeal to a similar intuition to construct an incentive compatible mechanism for eliciting expected changes in beliefs in a dynamic setting. They attribute the origin of this mechanism idea, used first in the context of eliciting preferences over many possible choices, to Allais (1953).

distribution is *not* a best response. Next the paper will identify the partition-drawing procedures for which this procedure implements an agent's true beliefs distribution as a *unique* best response for continuous belief distributions on an interval. Finally for the case of eliciting the beliefs over an interval with a finite (but possibly large) number of bins, the paper reports simulations that demonstrate the sensitivity of the expected payoff for this elicitation method the agent misreporting her belief distribution. These simulations allow to compare the introduced Random Partitions method to a normalized version of the QSR and a normalized version of the BDM/Karni elicitation procedure. The results show that the strength of the 'truthfully reporting' incentives of the Random Partitions method remains robust, while for the Quadratic Scoring Rule and the Becker-DeGroot-Marschak-based procedure, the payoff consequences for misreporting decrease as the number of relevant states or bins increases.

Section 2 describes the Random Partitions procedure and gives a simple 3-state numerical example, section 3 states and proves the main theoretical results, section 4 reports the results of the simulation exercise and section 5 concludes.

2 Defining the Random Partitions elicitation procedure

Assume an agent with any preferences such that she prefers a lottery $(Y, p; 0, 1-p)$ with $Y > 0$ to a lottery $(Y, q; 0, 1-q)$ if $p > q$, and that satisfy expected utility. Furthermore assume that this agent holds a belief distribution F over a state space S. Let Ω_S be the set of all possible *finite two-level partitions* of state space S where a finite two-level partition is a partition $\{e_1, e_{2,}, .., e_n\}$ of S into a finite number n of subsets (or 'events') e_i, with for each event e_i another partitioning $\{e_{i1}, e_{i2,}, .., e_{im_i}\}$ of e_i into a finite number m_i of events e_{ij}. Let Γ be a probability distribution over Ω_S.

The random partitions elicitation proceeds as follows:

1. The principal announces the payoff randomisation procedure, which will imply a distribution Γ over Ω_S.

2. The agent reports a belief distribution G over the state space S.

3. The principal randomly draws, using the announced distribution Γ, a finite two-level partitioning of S.

4. The value of the true state $s^* \in S$ is realised.

5. The true state s^* will lie in one of the events, say e_i, of the drawn partition of S. Furthermore s^* will lie in one of the subsets e_{ij} of the partition of e_i. The elicited distribution G assigns a highest probability to one of the subsets of e_i, call this subset e_{iG}. If $s^* \in e_{iG}$ then the agent receives an amount $Y > 0$ and otherwise the agent earns 0.

2.1 A three-state example

To give an intuition how this elicitation mechanism works I will now present a simple numerical example of eliciting the belief from an agent over a set of three states. The example is to illuminate that submitting one's true belief distribution is a best response, that truthful reporting is not neccesarily a unique best reponse, especially in a three state setting, and why the two-level partition feature of the mechanism can help generate stronger truth reporting incentives, as compared to rewards based on a one-level partition variant.

Assume a state space S containing three states $S = \{a, b, c\}$. Also assume an agent who holds the following beliefs F over S: $F(a) = 0.45, F(b) = 0.3, F(c) = 0.25$.

To illustrate how a two-level random partitioning of S can be constructed, Let $P_1(\{e_1, e_2..., e_n\}\})$ with $e_1 \cup e_2, .., \cup e_n = S$ be the probability under Γ that the first level partition becomes$\{e_1, e_2..., e_n\}$. Also let $P_2(\{e_{i1}, e_{i2}..., e_{im}\}\})$ with $e_{i1} \cup e_{i2}, .., \cup e_{im} = e_i$ be the probability under Γ that, given a first partition draw containing e_i as a subset, the second level partition becomes $\{e_{i1}, e_{i2}..., e_{im}\}$(for ease I only specify the probabilities and partitions that have a strictly positive probability under Γ).

One possible set of partition drawing probabilities is described in table 1 and will be used as an example:

Table 1: Example of two-level partitioning procedure of $S = \{a, b, c\}$

$P_1(\{\{a, b, c\}\}) = 0.4$	$P_2(\{\{a, b\}, \{c\}\}) = 0.2$	$P_2(\{\{a\}, \{b\}\}) = 1$
$P_1(\{\{a, b\}, \{c\}\}) = 0.2$	$P_2(\{\{a, c\}, \{b\}\}) = 0.2$	$P_2(\{\{a\}, \{c\}\}) = 1$
$P_1(\{\{a, c\}, \{b\}\}) = 0.2$	$P_2(\{\{b, c\}, \{a\}\}) = 0.2$	$P_2(\{\{b\}, \{c\}\}) = 1$
$P_1(\{\{b, c\}, \{a\}\}) = 0.2$	$P_2(\{\{a\}, \{b\}, \{c\}\}) = 0.4$	

Given this true belief distribution F and this specific distribution Γ of drawing the double partition we can calculate the difference in expected payoffs for the agent if she reports a distribution G with $G(a) = 0.56, G(b) = 0.2, G(c) = 0.24$ as opposed to reporting her true beliefs F. We can immediately see how reporting such a different distribution could lead to a different payoff outcome. Say for example the first level partition that is drawn is$\{\{a, b, c\}\}$, and the second level partition draw is$\{\{b, c\}, \{a\}\}$. Under distribution F the event set$\{b, c\}$ is more likely than event a, but reporting distribution G would lead to event a being considered more likely than the union of events b and c, and so reporting G would mean betting on event a in the case of this double partition draw.

Now we can calculate, given the double partition distribution Γ as specified in table 1 and true beliefs F, that the probability for the agent to receive prize $Y$when reporting F is 67.2 and when reporting G is 61% (calculations are in Appendix B). Note that, using this procedure, any reported belief with $G(a) > G(b) > G(c)$ and $G(\{b, c\}) > G(a)$ would yield the same payoff as reporting F, therefore clearly given such a coarse state space the random partition method does not provide incentives that make reporting truthfully a *unique* best response. We will see in proposition 3 however that for elicitations over a continuous interval the method does provide strict truthtelling incentives and the simulations in section 4 will show that for spaces with a higher number of states the random partitions method provides strong and near-strict truthtelling incentives.

A potential question one could have when seeing the definition of the random partitions method is "why use a two-level partition and not just a single level partitioning of S?". The advantage of a two level partitioning of the state space is that it allows one to ask the question "which of these subsets (or intervals) does the agent's submitted distribution regard as more likely" at a local level in the distribution. In the example state c is wrongly considered more likely than state b by distribution G, however as $G(a) > G(\{b, c\})$ there will not be any single partition of S where the relative likelihood of state b compared to state c can be tested. Using the same notation as before, but setting $P_1(\{\{a,b,c\}\}) = 1$, we can calculate the different expected payoffs of reporting either F or G for a single level partition example. The partition probabilities for this example are shown in table 2.

Table 2: Implementing a single level partition of $S = \{a,b,c\}$

$P_1(\{\{a,b,c\}\}) = 1$	$P_2(\{\{a,b\},\{c\}\}) = 0.2$
	$P_2(\{\{a,c\},\{b\}\}) = 0.2$
	$P_2(\{\{b,c\},\{a\}\}) = 0.2$
	$P_2(\{\{a\},\{b\},\{c\}\}) = 0.4$

We now get that the expected probability of earning a prize Y is 58%when reporting F and 56% when reporting G (again, see for calculations Appendix B). Looking at the relative intensity of the incentives of the two examples, reporting F instead of G would constitute a 10.2% increase in the expected payoff for the double partition setup versus a 3.6% increase in the single level partition setup. This example shows that the ability of the double partition procedure to put the likelihoods implied by a submitted distribution locally to the test can increase the overall strength of the incentives to encourage truthful reporting.

3 Theoretical results

3.1 General results

The following two propositions describe that reporting her true belief distribution is a best response for the agent and specify the conditions under which reporting a different distribution than the true belief is *not* a best response.

Proposition 1*: It is a best response for the agent to report her true belief distribution F.*

Proof: For each subset e_i of any drawn partition of S, the agent would maximize the probability to win the prize Y, conditional on e_i to contain the realisation s^*, by selecting as paying-out sub-subset e_{iG} the one she beliefs to have the highest probability of being realized given that $s^* \in e_i$. As the partitions and sub-partitions are drawn independently of the reported distribution G the sub-subset e_{ij} with the highest probability conditional on $s^* \in e_i$ is simply the sub-subset e_{ij} of e_i

with the highest probability according to the agent's beliefs F. Therefore reporting $G = F$ would select the winning-probability maximizing subset e_{ij} for each possible partition and sub-partition draw and is therefore a best response of the agent for this elicitation procedure.■

This procedure does not in general guarantee that for the agent, reporting her true beliefs distribution is the *unique* best response. However we can show to what extent the procedure is incentive compatible by identifying what types of belief responses G would not be a best response.

Let ζ be a subset of Ω_S and define $e_{iG} = argmax_{e_{ij} \in e_i} G(e_{ij})$ *and* $e_{iF} = argmax_{e_{ij} \in e_i} F(e_{ij})$.

Proposition 2*: A reported belief distribution G is not a best response if there exists a $\zeta \in \Omega_S$ with $\Gamma(\zeta) > 0$ where for each element in ζ there exists a subset e_i and sub-partition $\{e_{i1}, e_{i,2}, .., e_{i,m_i}\}$ with $F(e_i) > 0$ and $F(e_{iF}) > F(e_{iG})$.*

Proof: The agent can improve her expected payoff, according to her own beliefs F, by reporting F instead of G. We know for each possible pair of $e_i \in S$ and partition $\{e_{i1}, e_{i,2}, .., e_{i,m_i}\}$ of e_i, reporting F maximizes the probability of receiving reward Y given that $s^* \in e_i$. So reporting F instead of G does not lower the expected chance of winning Y conditional for any double partition draw. However there is a positive probability that a double partition is drawn for which reporting F leads to a different subset selection e_{iF} compared to reporting G, leading to a nonzero increase, conditional on this double partition being drawn, of the probability under beliefs F that the agent will earn Y after reporting F instead of G as $F(e_{iF}) > F(e_{iG})$.■

Proposition 2 shows how this procedure limits the type of distributions an agent could optimally enter in step 2 of the elicitation procedure. Essentially if there are double partition draws for which entering a deviating distribution would mean 'betting' on a less likely event e_{ij} compared to entering one's true distribution, then the agent would regret such a distribution choice after step 3 in the procedure and such a distribution would not be a best response to the elicitation. Because of this, any best response belief distribution must be very similar to the agent's true beliefs as soon as there are a substantial number of states to randomize over that are have positive probabilities of occuring according to the agent's true belief distribution.

The fact that the double partition draw is random means that to be a best response, an elicited distribution should be optimal for every possible 'dilemma' the entered beliefs distribution might have to answer once the partitions are drawn. The following subsection demonstrates a specific setting where the procedure elicits the agent's true beliefs as a *unique* best response.

3.2 Result for eliciting a continuous belief distribution over a finite interval

Now I will show how an implementation of the randomization procedure Γ for generating double partitions of the state space S will elicit an agent's true beliefs as a *unique* best response if the agent holds and can report a continuous belief distribution over an interval.

Assume the state space S is a closed interval $[x_{min}, x_{max}]$ on the real line and the agent holds a continuous belief distribution F over this interval. Furthermore, define Γ to be a *subinterval-robust* distribution over Ω_S if there exists an $\epsilon > 0$ such that for every possible set of $\{z_i, z_o, z_p, z_q, z_r\}$ with $x_{min} \leq z_i < z_o < z_p < z_q < z_r \leq x_{max}$ the distribution Γ assigns a positive probability to the set of finite double partitions containing an interval $[x_i, x_j]$ with $z_i - \epsilon \leq x_i \leq z_i + \epsilon$ and $z_j - \epsilon \leq x_j \leq z_j + \epsilon$ as a subset e_i of S and $[x_o, x_p]$ and $[x_q, x_r]$ with $z_o - \epsilon \leq x_o \leq z_o$, $z_p - \epsilon \leq x_p \leq z_p$, $z_q - \epsilon \leq x_q \leq z_q$ and $z_r - \epsilon \leq x_r \leq z_r$ as subsets e_p and e_q of e_i.[3]

Proposition 3*: It is a unique best response for the agent to report her true continuous belief distribution F over the interval $[x_{min}, x_{max}]$ if the elicitation procedure uses a distribution Γ that is subinterval-robust.*

Proof: Given that F and G are continuous belief distributions over an interval $[x_{min}, x_{max}]$, if there is a value $x^*$with $f(x^*) \neq g(x^*)$ then there must exist an set of $\epsilon, z_i, z_o, z_p, z_q, z_r$ with $\epsilon > 0$ so that, for all sets e_i equal to $[x_i, x_j]$ with $z_i - \epsilon \leq x_i \leq z_i + \epsilon$ and $z_j - \epsilon \leq x_j \leq z_j + \epsilon$ containing subsets $e_{io} = [x_o, x_p]$ and $e_{io} = [x_q, x_r]$ with $z_o - \epsilon \leq x_o \leq z_o$,$z_p - \epsilon \leq x_p \leq z_p$, $z_q - \epsilon \leq x_q \leq z_q$, we have $argmax_{e_{ij} \in e_i} G(e_{ij}) = e_{io}$ *and* $argmax_{e_{ij} \in e_i} F(e_{ij}) = e_{ir}$ with $F(e_{ir}) > F(e_{io})$. If Γ assigns a positive probability to the set of finite double partitions containing an interval $[x_i, x_j]$ and sub-

[3]Note that a random partition draw here still contains only a finite number of intervals, even if the state space is continuous. One drawing procedure that would implement a distribution Γ with these properties is one where the first level partition creates 3 intervals, with each cut generated by a uniform random draw, and 5 intervals created at the second level partition, again with each cutoff uniformly drawn within the specific level one partition subinterval. Many alternative procedures would also satisfy the requirements of proposition 3.

intervals$[x_o, x_p]$ and $[x_q, x_r]$ from the described set, then from proposition 2 it follows that any belief distribution G with $f(x^*) \neq g(x^*)$ for some value $x^*$is not a best response and so reporting the true belief distribution F is the unique best response.■

4 Comparison of the Random Partitions elicitation method with other belief elicitation methods

In this section compares the incentive strength of the proposed Random Partitions mechanism with two other methods, the Quadratic Scoring Rule, and the Becker DeGroot Marchak-based procedure as described by Karni (2009). First, I will state a proposition for the Quadratic Scoring Rule showing that the strength of incentives tends to zero as the grid gets finer. Second, I will show the results of a simulation excercise showing the payoff consequences of misreporting ones true beliefs for the three elicitation methods. What the simulations will show is that even for discrete distributions with a finite number of states (where the Random Partitions method does not formally implement the true belief distribution as a *unique* best response) the Random Partitions method still provides clear incentives to report a distribution very close to one's true beliefs. Even more, we will see that if our state space contains a large number of bins, then even in this discrete setting the incentives provided by the Random Partitions elicitation mechanism will be stronger than those of a normalized version of the Quadratic scoring rule and the Karni method.[4]

Both the quadratic scoring rule and the Karni method allow for normalization of the potential payoffs between 0 and 1.[5] This has the advantage, as explained in Schlag et al. (2015) (see also Roth and Malouf 1979, McKelvey 1990, Sandroni, A. & Shmaya 2013, Hossain and Okui 2013, and Harrison et al. (2013 and 2014)), that it allows the normalized score to be translated into a lottery probability to win a fixed price Y. This would make the elicitation method independent of

[4]Stronger incentives here means that given a certain true belief distribution F, how much do the expected payoffs for the agent decrease if she reports a different distribution G. Another interpretation of stronger incentives could be to what extent will the expected payout for an agent increase if she is able to reduce the entropy level of her true beliefs distribution.

[5]Assuming here again that the agent's preferences over lotteries satisfy the independence axiom. See for example Schlag & Van der Weele (2015) or Sandroni, A. & Shmaya (2013).

an agent's risk attitudes (an issue addressed by Winkler and Murphy, 1970).[6] Here it also allows for an easy direct comparison between the different elicitation mechanisms.

4.1 Decreasing payoff sensitivity for the Quadratic Scoring Rule

The following proposition establishes that the strength of the incentives for the discrete Quadratic Scoring Rule decrease asymptotically as the grid used for the elicitation gets finer.

Proposition 4: *Assume an agent holds a continuous belief distribution F over the interval $[x_{min}, x_{max}]$ with lenght l, which is divided into an equally spaced grid of n intervals. As n increases, the difference in expected Quadratic Scoring Rule payoff of reporting an agent's truthfull belief probability vector and reporting an arbitrary different probability vector will asymptotically converge to 0.*

The proof of this proposition is described in Appendix C.

4.2 Description of the simulation procedure

For the simulation exercise we assume that the state space is the interval $[0, 200]$ and that the agent holds a belief distribution equal to the truncated normal distribution with mean 100 and standard deviation 15. The simulation runs 100000 iterations of the elicitation procedure for each method. An iteration consists of drawing a realization from the actual belief distribution that the agent holds, and an application of each of the considered elicitation methods (which possibly entails further random draws to construct a partitioning of the interval). Throughout, the state space will be divided up in a number of equally sized bins. The simulations will be performed for different bin-width sizes, namely bins with width 10, 1 and 0.1.

4.2.1 Procedure for the Random Partitions method

The application of the proposed procedure here generates a double partition in the following way. The two level partition for the simulation was generated by drawing n uniformly random numbers,

[6] But note that Harrison et al. 2017 show the correction needed when eliciting beliefs using a QSR on risk-averse agents is limited when eliciting entire distributions. Next to this there have been other ways proposed to correct proper scoring rules for risk attitudes and probability weighting, see e.g. Offerman et al 2009.

sorted to form $\{x_1, x_2..., x_{n+2}\}$,with $x_1 = 0$ and $x_{n+2} = 200$ for the first level partition, and m uniformly random numbers, $\{x_{i1}, x_{i2}..., x_{im}\}$, between x_i and x_{i+1} for the second level partition, with n drawn uniformly from $1, 2, 3, 4$ and m uniformly from $4, 5, .., 30$.[7]

4.2.2 Procedure for the normalized Quadratic Scoring Rule

Given a vector $\mathbf{p}$ of elicited probabilities for each of the n bins that the state space is divided in, the quadratic scoring rule, after the true state is revealed to be part of bin i is equal to $\alpha + \beta \cdot p_i - \sum_{j=1}^{n} 0.5 \cdot \beta \cdot p_j^2$. To normalize this score, α is set equal to 0.5 and β equal to 1.

4.2.3 Procedure for the normalized BDM/Karni method

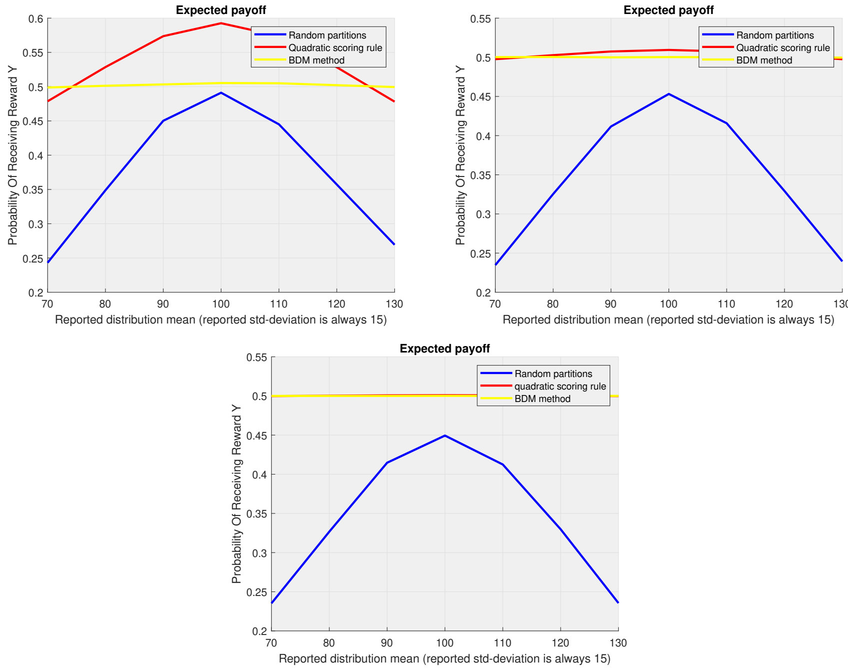
For the BDM/Karni method, following Karni (2009) , after the agent has submitted her elicitation, for each bin a separate random number r_i is drawn from the uniform distribution between 0 and 1. For each bin, if $r_i > p_i$ with p_i the probability assigned to bin i by the elicited distribution, then for this bin a lottery is played out where the agent has a probability of r_i to receive a prize Z. If instead, $r_i \leq p_i$ then the agent will receive a prize Z if the true state s^* falls in bin i. In order to normalize the procedure such that the agent receives at most a payout of 1 and at minimum a payout of 0, I choose $Z = \frac{1}{n}$ with n the number of bins.

4.3 Simulation results

Figures 1 and 2 below illustrate the simulation results for the different elicitation procedures. For Figure 1, simulated scores from the three methods are generated for truncated normal distributions all based on a standard-deviation of 15, but with means ranging from 70 to 130. For Figure 2, simulated scores from the three methods are generated for truncated normal distributions all based on a mean of 100, but with the standard-deviation ranging from 1 to 100.

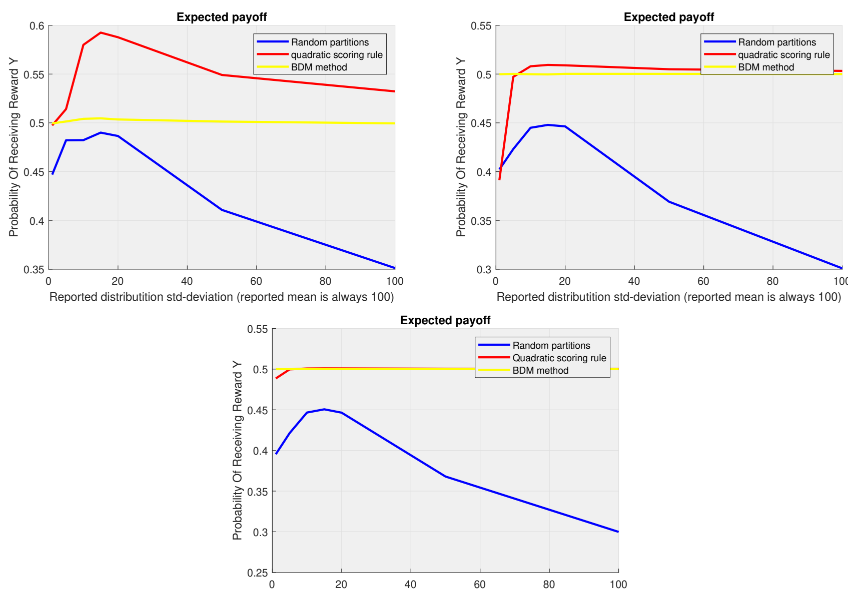
[7] The principal is free to select a partition procedure on each level that is more likely to either partition an interval in a few large subintervals or a larger number of smaller intervals. For example, having more and smaller intervals in the second level of the partition would lower the probability that the agent receives prize Y.

Figure 1: Expected payoff sensitivity to wrong reporting of the mean



Note: Simulation comparison based on reporting distributions with correct std-deviation but varying means. The bin widths used for the grapsh are respectively 10,1 and 0.1.

Figure 2: Expected payoff sensitivity to wrong reporting of the variance



Note: Simulation comparison based on reporting distributions with correct mean but varying std-deviations. The bin widths used for the grapsh are respectively 10,1 and 0.1.

As we can see from the figures, both the quadratic scoring rule and the Karni elicitation provide very 'flat' payoffs for bin-widths smaller than 10, but the Random Partitions method keeps, it's payoff sensitivity to the agent misreporting her true beliefs. This could be a desirable property in situations where one wants to elicit the beliefs of an expert on the future value of a variable where both small shifts in the variable could matter, but it's also clear that any expert belief distribution would have a substantial degree of variance.[8] Appendix A shows the results of more simulations with cases of multi-peaked and beta belief distributions.

5 Conclusion

This paper has introduced a new mechanism to incentivise belief elicitation, the Random Partitions method. Three theoretical results have shown for this method (i) it always is a best response for an agent to report her true beliefs, (ii) the conditions under which reporting a distribution different from one's true beliefs is *not* a best response, and (iii), the procedure can be used to elicit beliefs for a continuous distribution over a finite interval such that reporting her true beliefs is the agent's *unique* best response. Furthermore I have reported simulation results which show that also for a distribution over finite-state space, this method provides strong truth-telling incentives when compared to the standardly used Quadratic Scoring Rule (Brier 1950 and e.g. Selten 1998). Moreover, whereas both the Quadratic scoring rule and the BDM-based method (Becker et al. 1964 and Karni 2009) tend to provide more flat incentives as the state space grid becomes finer, the incentive precision and strenght of the proposed Random Partitions method remains robust. The results suggest that when eliciting a continuous distribution or a distribution over a state space with many bins, the proposed Random Partitions method provides sharper incentives than currently used elicitation mechanisms. One future extention to mention here is to look at the situation where an agent, *before* reporting her beliefs, has the opportunity to aquire costly information that will reduce the variance (or information entropy) of her belief distribution. A question to investigate (theoretically and potentially experimentally) is to what degree agents' willingness to pay to aquire information is influenced by the incentive strength provided a belief elicitation mechanism. This could provide another foundation for principals to prefer using elicitation mechanisms, such as the Random Partitions method, that provide substantial truth-telling incentives.

[8] One example of an economic indicator likely to have these properties would be the future oil price.

References

[1] Allais, M. (1953), Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école americaine, *Econometrica*, 21 (4), 503–546.

[2] Armantier, O, and N. Treich (2013), Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging", *European Economic Review*, 62, 17-40.

[3] Becker, G., M. DeGroot, and J. Marschak (1964), "Measuring Utility by a single-response sequential method", *Behavioral Science*, 9, 226-236.

[4] Brier, G.W. (1950), "Verification of Forecasts Expressed in Terms of Probability", *Monthly Weather Review*, 78 (1), 1-3.

[5] Chambers, C.P., and N.S. Lambert (2019), "Dynamic Belief Elicitation", Working paper.

[6] Gneiting, T., & Raftery, A. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. Journal of the American Statistical Association, 102(477), 359-378.

[7] Harrison, G. (1989), "Theory and Misbehavior of First-Price Auctions", *American Economic Review*, 79,4,749-762.

[8] Harrison, G. W., Martínez-Correa, J., & Swarthout, J. T. (2013), "Inducing risk neutral preferences with binary lotteries: A reconsideration", *Journal of Economic Behavior & Organization*, 94, 145-159.

[9] Harrison, G., J. Martínez-Correa, J.T. Swarthout, and E.R. Ulm (2017), "Scoring rules for subjective probability distributions", *Journal of Economic Behavior & Organization*, 134, 430-448.

[10] Harrison, G. W., & Phillips, R. D. (2014). Subjective beliefs and statistical forecasts of financial risks: The chief risk officer project. In Contemporary Challenges in Risk Management (pp. 163-202). Palgrave Macmillan, London.

[11] Hossain, Tanjim, and Ryo Okui. "The binarized scoring rule." *Review of Economic Studies*, 80.3 (2013): 984-1001.

[12] Karni, E. (2009), "A mechanism for eliciting probabilities", *Econometrica*, 77 (2), 603-606.

[13] Manski, C. (2004). Measuring Expectations. Econometrica, 72(5), 1329-1376.

[14] Matheson, J.E. , and R.L. Winkler (1976) "Scoring rules for continuous probability distributions". *Management Science*, 22,1087-1096.

[15] McKelvey, R.D., and T. Page (1990) Public and Private Information: An Experimental Study of Information Pooling. *Econometrica*, 58, 1321-1339.

[16] Murphy, A. H., & Winkler, R. L. (1970), "Scoring rules in probability assessment and evaluation", *Acta psychologica*, 34, 273-286.

[17] Offerman, T., Sonnemans, J., Van de Kuilen, G., & Wakker, P. P. (2009), "A truth serum for non-bayesians: Correcting proper scoring rules for risk attitudes", *The Review of Economic Studies*, 76(4), 1461-1489.

[18] Roth, A.E., and M.W.K. Malouf, (1979), "Game-theoretic models and the role of information in bargaining", *Psychological Review*, 86, 574-594

[19] Sandroni, A. & Shmaya, E., 2013. Eliciting Beliefs by Paying in Chance. *Economic Theory Bulletin*, 1(1), pp. 33–37.

[20] Savage, L.J. (1971), "Elicitation of personal probabilities and expectations", *Journal of the American Statistical Association*, 66, 783-801.

[21] Schlag, K.H., J. Tremewan, and J.J. van der Weele (2015), "A penny for your thoughts: a survey of methods for eliciting beliefs", *Experimental Economics*, 18, 457-490.

[22] Schlag, K. H., & van der Weele, J. J. v. (2015). A method to elicit beliefs as most likely intervals. *Judgment and Decision Making*, 10(5), 456-468.

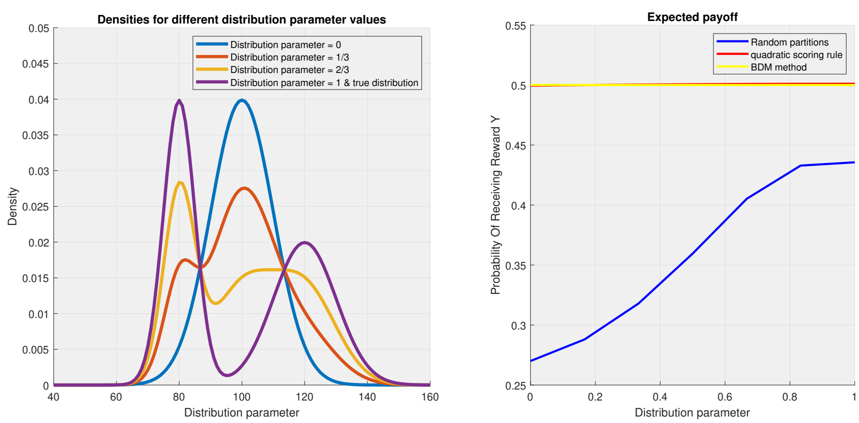
[23] Selten, R. (1998), "Axiomatic characterisation of the quadratic scoring rule", *Experimental Economics,* 1, 43-62.

[24] Winkler, R. L., & Murphy, A. H. (1970). Nonlinear utility and the probability score. *Journal of Applied Meteorology*, 9(1), 143-148.

Appendix A, further simulation examples

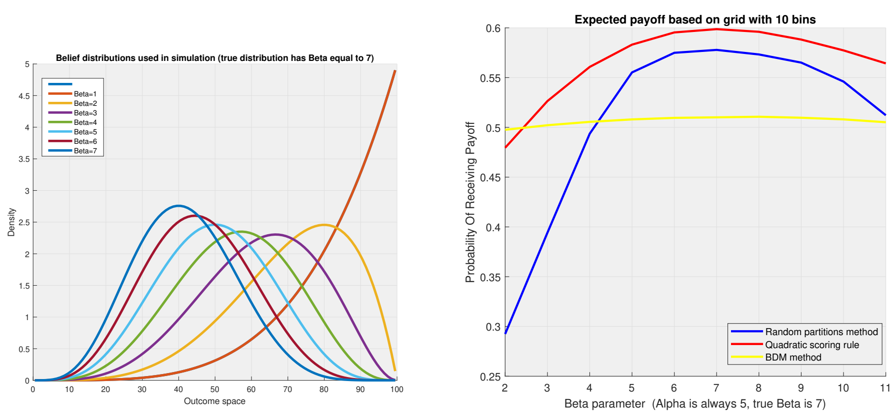
Simulations are again based on counting the the share of elicitations ending in the agent receiving a reward based on 100000 iterations.

Figure 3: Simulation results for assymmetric-multimodal distributions



The simulations are based on bins with with 0.1

Figure 4: Simulation results using Beta distribution beliefs



Appendix B, calculations for the three-state example

Calculations of the three-state example in section 2:

If the agent reports her true belief probabilities $(F(a) = 0.45, F(b) = 0.3, F(c) = 0.25)$, then under these beliefs the expected probability of receiving reward Y will be

$$F(a) \cdot (P(\{\{a\},\{b\},\{c\}\}) + P(\{\{a,b\},\{c\}\}) + P(\{\{a,c\},\{b\}\}))$$

$$+F(b) \cdot (P(\{\{a,b\},\{c\}\}) + P(\{\{b,c\},\{a\}\})) + F(c) \cdot (P(\{\{a,c\},\{b\}\}) + P(\{\{b,c\},\{a\}\}))$$

$$= 0.45 \cdot 0.8 + 0.3 \cdot 0.4 + 0.25 \cdot 0.4 = 0.58$$

Now instead if the agent submits the beliefs $(G(a) = 0.56, G(b) = 0.2, G(c) = 0.24)$ then the expected payout will be

$$F(a) \cdot (P(\{\{a\},\{b\},\{c\}\}) + P(\{\{a,b\},\{c\}\}) + P(\{\{a,c\},\{b\}\}) + P(\{\{b,c\},\{a\}\}))$$

$$+F(b) \cdot (P(\{\{a,b\},\{c\}\})) + F(c) \cdot (P(\{\{a,c\},\{b\}\}))$$

$$= 0.45 \cdot 1 + 0.3 \cdot 0.2 + 0.25 \cdot 0.2 = 0.56$$

Second in case of the double partition with the above probabilities, again starting from the true beliefs $(F(a) = 0.45, F(b) = 0.3, F(c) = 0.25)$, given honest/accurate reporting the expected payoff probabilitie will be:

$$F(a) \cdot (P_1(\{\{a,b\},\{c\}\}) + P_1(\{\{a,c\},\{b\}\}) + P_1(\{\{b,c\},\{a\}\}) + P_1(\{\{a,b,c\}\}) \cdot (1 - P_2(\{\{b,c\},\{a\}\})))$$

$$+F(b) \cdot (P_1(\{\{a,c\},\{b\}\}) + P_1(\{\{b,c\},\{a\}\}) + P_1(\{\{a,b,c\}\}) \cdot (P_2(\{\{b,c\},\{a\}\}) + P_2(\{\{a,b\},\{c\}\})))$$

$$+F(c) \cdot (P_1(\{\{a,b\},\{c\}\}) + P_1(\{\{a,b,c\}\}) \cdot (P_2(\{\{b,c\},\{a\}\}) + P_2(\{\{a,c\},\{b\}\})))$$

$$= 0.45 \cdot ((0.2 + 0.2 + 0.2) + 0.4 \cdot (1 - 0.2))$$

$$+0.3 \cdot (0.2 + 0.2 + 0.4 \cdot (0.2 + 0.2))$$

$$+0.25 \cdot (0.2 + 0.4 \cdot (0.2 + 0.2))$$

$$= 0.672$$

Now instead if the agent submits the beliefs $(G(a) = 0.56, G(b) = 0.2, G(c) = 0.24)$ then the expected payout will be

$$F(a) + F(b) \cdot P_1(\{\{a,c\},\{b\}\}) + F(c) \cdot (P_1(\{\{a,b\},\{c\}\}) + P_1(\{\{b,c\},\{a\}\}))$$

$$0.45 + 0.3 \cdot 0.2 + 0.25 \cdot 0.4 = 0.61$$

So we get stronger incentives in the two-level partition case, also relatively (3.45% expected payoff loss versus 9.23% in the double partition case here).

Appendix C, proof of bin-size sensitivity of the quadratic scoring rule

Let the agent hold a continuous belief distribution Fover the interval $[x_{min}, x_{max}]$ with finite length l. We consider the scoring rule $\alpha + \beta \cdot p_i - \sum_{j=1}^{n} 0.5 \cdot \beta \cdot p_j^2$. Assume an agent holds a belief vector $\mathbf{p}$, but reports a vector $\boldsymbol{r}$the expected payoff in the payoff for the quadratic scoring rule, assuming $\alpha = 0.5$ and $\beta = 1$ will be:

$$\sum_{k=1}^{n} p_i \cdot \left(0.5 + r_i - \sum_{j=1}^{n} 0.5 \cdot r_j^2\right)$$

$$= 0.5 + \left(\sum_{k=1}^{n} p_i \cdot r_i\right) - 0.5 \cdot \left(\sum_{j=1}^{n} r_j^2\right)$$

So the agent's loss compared to reporting $\mathbf{p}$will be

$$0.5 + \left(\sum_{k=1}^{n} p_i \cdot r_i\right) - 0.5 \cdot \left(\sum_{j=1}^{n} r_j^2\right) - \left(0.5 + \left(\sum_{k=1}^{n} p_i^2\right) - 0.5 \cdot \left(\sum_{j=1}^{n} p_j^2\right)\right)$$

$$= \left(\sum_{k=1}^{n} p_i \cdot r_i\right) - 0.5 \cdot \left(\sum_{j=1}^{n} r_j^2\right) - 0.5 \cdot \left(\sum_{j=1}^{n} p_i^2\right) = \sum_{j=1}^{n} -0.5 \cdot (p_j - r_j)^2$$

Which gives us the "loss" function (see also the derivation of Selten,1998):

$$\sum_{j=1}^{n} -0.5 \cdot (p_j - r_j)^2$$

Define $\bar{x}_i$ as the upper cutoff of bin i in the grid partition of the interval and $\underline{x}_i$ as the lower cutoff of bin i. Now assume probability vectors $\mathbf{p}$ and $\boldsymbol{r}$ are derived from continious distributions F and G with pdf's f and g, created by having p_i be equal to $\int_{\underline{\mathrm{x}}_\mathrm{i}}^{\bar{x}_i} f(x)\,dx$ and r_i be equal to $\int_{\underline{\mathrm{x}}_\mathrm{i}}^{\bar{x}_i} g(x)\,dx$. As the bin width goes to 0 and n approaches infinity, p_j converges to $\frac{l}{n} \cdot f\left(\frac{\bar{x}_i + \underline{x}_i}{2}\right)$ and r_j converges to $\frac{l}{n} \cdot g\left(\frac{\bar{x}_i + \underline{x}_i}{2}\right)$

This gives us

$$lim_{n\to\infty} \sum_{j=1}^{n} -0.5 \cdot (p_j - r_j)^2 =$$

$$lim_{n\to\infty} - 0.5 \cdot \sum_{j=1}^{n} \left(\frac{l}{n} \cdot \left(f\left(\frac{\bar{x}_i + \underline{x}_i}{2}\right) - g\left(\frac{\bar{x}_i + \underline{x}_i}{2}\right)\right)\right)^2$$

$$= lim_{n\to\infty} - 0.5 \cdot \frac{l^2}{n^2} \sum_{j=1}^{n} \left(f\left(\frac{\bar{x}_i + \underline{x}_i}{2}\right) - g\left(\frac{\bar{x}_i + \underline{x}_i}{2}\right)\right)^2$$

$$= lim_{n\to\infty} - 0.5 \cdot \frac{l}{n} \cdot \left(\frac{l}{n} \cdot \sum_{j=1}^{n} \left(f\left(\frac{\bar{x}_i + \underline{x}_i}{2}\right) - g\left(\frac{\bar{x}_i + \underline{x}_i}{2}\right)\right)^2\right)$$

$$= -0.5 \cdot \frac{l}{\infty} \cdot \int_{x_{min}}^{x_{max}} (f(x) - g(x))^2\,dx = 0$$

As long as f and g are distributions with $\int_{x_{min}}^{x_{max}} \left(f\left(x\right) - g\left(x\right)\right)^2 dx$ attaining a finite value, we have for given underlying F and G

$$lim_{n\to\infty} \sum_{j=1}^{n} -0.5 \cdot \left(p_j - r_j\right)^2 = 0$$

Thus the incentive strength of the quadratic scoring rule tends to 0 as the number of bins increase (and consequently the bin width decreases).■

Department of Economics
University of Bergen
PO BOX 7802
5020 Bergen
Visitor address: Fosswinckels gate 14
Phone: +47 5558 9200
www.uib.no/econ/