ELICITATION RECORD – Part 2 – Distribution

Quartile Method

Elicitation title	As in the Part 1 form
Session	As in the Part 1 form
Date	As in the Part 1 form
Quantity	The uncertain quantity whose distribution is to be elicited
Start time	Time when this part of the elicitation started

Definition	Repeat the definition of this quantity from Part 1. Give it a symbol to facilitate the recording of judgements about it. It will be called X in these notes.
Evidence	Review the evidence specifically about X. (Refer to principal sources, but do not repeat lots of detail here.) [Although the evidence base has been set out in the Part 1 form, the facilitator should ask each expert to consider which items are of relevance to this quantity.] *As in Part 1, this step is to avoid the 'availability heuristic', in which experts rely only on a subset of evidence that comes readily to mind.*
Plausible range	Elicit from the experts the range of plausible values for X. In these notes, L is the lower bound of the range and U the upper bound. It may be useful to record absolute, logical bounds, but the objective here is to identify a range such that it is extremely unlikely (but not necessarily impossible) that X lies outside. Record how the range has been informed by the evidence base. [This range should not be unnecessarily wide, but it is important that it should not be too narrow. This should be a joint judgement of the experts, such that they all believe that X is extremely unlikely to be outside (U, L). After an initial specification of the range, the facilitator should probe by asking something like "Suppose an experiment produced a value [something below L or above U] for X; would this have to have been a flawed experiment, or might there be a way that X could have this value?"] *There is substantial evidence that experts tend to be over-confident, in the sense that they do not allow enough probability for extreme values of X. Psychologists suggest that experts develop mental models that allow them to understand and predict the processes in their area of expertise. Experts themselves may recognise instances where the model does not*

	apply, but part of their overconfidence in elicitation may be due to conditioning on inappropriate models. The suggested probing question invites them to think outside their models.
Median	The next steps should be done by each expert **separately**, without discussion. Each expert should specify their median value for X. This is a value such that they think 'X lies below the median' and 'X lies above the median' are equally likely propositions. Formally, if M is the median, then $P(X < M) = 0.5$. [The facilitator should instruct the experts to write down their own median values, but not to reveal them yet. Nothing should be written in this field until after the upper and lower quartiles have also been elicited.] *The judgement of equal probability is generally found to be simple for experts, and is not subject to systematic biases.*
Upper and lower quartiles	Each expert should now specify their lower quartile by considering the range from L to M and dividing it into two equally likely intervals. Formally, if Q1 is the lower quartile, $P(L < X < Q1) = P(Q1 < X < M) = 0.25$. Similarly, each expert should specify their upper quartile Q3 by dividing the range from M to U into equally likely intervals. Then $P(M < X < Q3) = P(Q3 < X < U) = 0.25$. Before deciding definitely on these values, experts should be asked to check that they regard each of the four ranges (L to Q1, Q1 to M, M to Q3 and Q3 to U) as equally likely. [When asking for the lower quartile, the facilitator should tell the experts to exclude for the moment the possibility that X is above M, and they should concentrate on the range from L to M. He/she should also point out that generally experts would feel that values of X close to M are more probable than values close to L, and so Q1 will typically be nearer to M than to L. In choosing Q1, the experts are determining how much more likely values near M are. Similar instructions should be given about Q3. When all the experts have written down their medians and quartiles, the facilitator should ask them to reveal their values, and should fill in this box and the one above in the record.] *Psychologists suggest that people often make judgements by adjusting a previous judgement, and that when they do this they typically do not adjust far enough. The first judgement is called the anchor, and the process is called the 'anchoring and adjustment' heuristic. Anchoring and adjustment is suggested as one reason for over-confidence in the assessment of probability intervals, when experts are anchored by a previous central estimate (such as the median), and do not adjust far enough away from this. The SHELF protocols always elicit first*

	the plausible range, and then ask experts to make such judgements relative to both the central value and the plausible range. In this way, they are anchored on both sides, which tends to cancel the effect. *However, another common cause of poor judgement is the 'range-frequency compromise,' according to which when thinking about probabilities within a range people tend to want to share probability reasonably evenly across the range. The effect of this is that the two anchors will tend to cause experts to put their quartiles in the middle of the range under consideration. This is why the facilitator should point out the essential unevenness of probability, and suggest a value closer to M. This needs to be done carefully, so as not to influence the experts too much.*
Fitting	The facilitator fits a distribution to each of the experts' assessments. (The distributions should be specified here, and if possible shown as density functions. If it is not straightforward to paste plots of the density functions into this record, they can be provided as an attachment, which should then be listed at the end.) [The facilitator should choose an appropriate family of distributions, and then fit the distribution by choosing parameters that give probabilities matching the elicited bounds, median and quartiles as closely as possible. Since L and U are not necessarily absolute bounds, there are 6 probabilities to match – four of 0.25 and two of 0.0. The distributions should be shown to the experts, but at this stage we do not invite revision (unless the expert is insistent that the plotted distribution badly distorts his/her beliefs) or provide any other feedback. If feasible, the facilitator should compute the median and quartiles of an equally-weighted average of the density functions. Ideally, these should not be revealed to the experts immediately, but may be used at the facilitator's discretion in the next stage.] *This stage of separate elicitations ensures that the initial divergence of opinion between the experts is recorded. The facilitator can refer to these if the group elicitation appears to be neglecting part of the original range of belief. There is evidence that group elicitation can itself lead to over-confidence, perhaps because the process of reaching consensus induces a false sense of decisiveness. So this step in the SHELF process allows the facilitator to see any narrowing of uncertainty, and to check that this is justified by the sharing of knowledge that has taken place. In particular, having the median and quartiles of the averaged densities gives the*

	facilitator a quick check on the degree to which subsequent discussion leads to a narrowing of uncertainty. *The process of averaging the density functions is known as the linear opinion pool (with equal weights). It is one of the formulae which proponents of eliciting separately from experts use to combine the resulting distributions. We use it in SHELF simply as a benchmark.*
Group elicitation	After discussion of the different distributions, and sharing of knowledge and reasoning about the differences, record group consensus values for the median and quartiles. [This discussion can take a substantial amount of time. The facilitator should only cut it short if he/she feels that the experts are no longer exchanging information and arguments but are just repeating opinions. The agreed median will inevitably be some sort of compromise. Before discussion, there are two components of uncertainty in the group – uncertainty that each expert has and is expressed in that expert's quartiles, as well as variability between the experts' judgements. The agreed quartiles should reflect the group's overall uncertainty that remains after the discussion. The facilitator needs to manage the discussion so that divergent views are properly considered, and to ensure that strong personalities and/or groups of people with overlapping experiences do not inappropriately dominate. In general, we would expect the group debate and sharing of opinions to result in a distribution that is narrower than the linear opinion pool, but it should not be much narrower or markedly different in other ways, unless the discussion has clearly altered individual experts' opinions. In some situations, experts may not be able to reach consensus. It may then be necessary for the facilitator to elicit two (or more) distributions, representing consensus views within the opposing factions.]
Fitting and feedback	Record here the (potentially iterative) process of fitting, feedback and revision of the group judgements. [The facilitator first fits a distribution to the group's L, Q1, M, Q3 and U values. This should be shown to the experts, and the fitted probabilities compared with the elicited probabilities. The experts are invited to consider whether the fit is close enough, or whether some values might be varied in order to fit others (that are believed to be more pivotal) better. The facilitator then feeds back to the experts some implied probabilities in the fitted distribution, such as the 10^{th} and 90^{th} percentiles. The experts are invited to consider whether these are reasonable reflections of the group's knowledge.

	If revision is needed, this may be followed by further rounds of fitting and feedback until the experts are comfortable with the fitted distribution and its implications.]
Chosen distribution	Record and show here the finally fitted distribution.
Discussion	The facilitator should record here any difficulties that arose during the elicitation of this distribution. Also the experts' reactions to the process and to the finally fitted distribution. *The elicitation record should be open about any concerns with the finished distribution. The SHELF protocols are designed to avoid many of the pitfalls of elicitation, but no process is perfect. It is important to be critical and realistic about the result. Nevertheless, it is important also to remember that, despite whatever deficiencies it might have, the elicited distribution is our best attempt. It has been developed using a robust protocol, and since expert knowledge is needed in the wider enterprise there is no alternative!*

End time	Time when elicitation of this distribution was completed.
Attachments	List any attachments, e.g. plots of distributions.