

ELICITATION RECORD – Part 2 – Distribution

Elicitation title	As in the Part 1 form
Workshop	As in the Part 1 form
Date	As in the Part 1 form
Quantity	The uncertain quantity whose distribution is to be elicited
Anonymity	Record here the codes that will be used to identify experts in this template. For instance, “In this record, experts are identified by letters A, B, C and the facilitator by Z.”
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	<p>Review the evidence specifically about X. (Refer to principal sources, but do not repeat lots of detail here. If there is only one distribution to be elicited in this session, then simply refer to the evidence dossier.)</p> <p>[Although the evidence base has been set out in the Part 1 form, if distributions for more than one quantity are to be elicited in this session then the facilitator should ask the experts to consider which items are of relevance to this quantity X. DO NOT allow the experts to discuss the evidence.]</p> <p><i>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.</i></p>
Plausible range	<p>Each expert should write down, privately and without discussion, their lower plausible limit L and their upper plausible limit U, thereby defining their plausible range.</p> <p>[This range should not be unnecessarily wide, but it is important that it should not be too narrow. Refer to the slide set “Plausible Range” for ways to explain L and U to the experts, and to help them to challenge and refine their judgements. The experienced facilitator may prefer to present these ideas in their own way, but otherwise it is recommended that the slide set should be used directly as a presentation to the experts.]</p> <p><i>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. One possible reason might be anchoring on a central value of X that has</i></p>

	<i>already been discussed. Establishing their plausible range before asking for central values is intended to avoid this problem.</i>
Individual elicitation	<p>Method: Record the method for individual elicitation – T(ertile), Q(uartile) or R(oulette)</p> <p>[The method will have been chosen in advance by the facilitator. Refer to the document “SHELF Methods” for guidance on making this choice.]</p> <p>Judgements: Each expert should write down, privately and without discussion, their judgements as required for the chosen method.</p> <ul style="list-style-type: none"> • For the tertile method, first their median M and then their tertiles T1 and T2. • For the quartile method, first their median M and then their quartiles Q1 and Q3. • For the roulette method, their probabilities in bins. <p>[For the tertile and quartile methods, refer to the slide sets “Median”, “Tertiles” and “Quartiles” for ways to explain M, T1, T2, Q1 and Q3, to the experts, and to help them to challenge and refine their judgements. The experienced facilitator may prefer to present these ideas in their own way, but otherwise it is recommended that the slide set should be used directly as a presentation to the experts.</p> <p>For the roulette method, refer to the document “SHELF Methods” for guidance on choosing the bins, and the slide set “Roulette” for explaining to the experts how to allocate probs to bins. The experienced facilitator may prefer to present these ideas in their own way, but otherwise it is recommended that the slide set should be used directly as a presentation to the experts.]</p> <p><i>The “SHELF Methods” document covers the reasoning behind each of the different methods, and how they are based on research into the psychology of expert judgement.</i></p>
Fitting	<p>Each expert should reveal their individual judgements (including their plausible range). All of the judgements are recorded here, using anonymised codes for the experts.</p> <p>The facilitator fits a distribution to each of the experts’ assessments.</p> <p>(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 of this form.)</p>

	<p>[The facilitator should choose an appropriate family of distributions, and then fit the distribution by choosing parameters that give probabilities matching as closely as possible the elicited judgements (for instance by using the “SHELF” software).</p> <p>The distributions should be shown to the experts, but at this stage we do not invite revision (even if the expert is insistent that the plotted distribution badly distorts his/her beliefs) or provide any other feedback.</p> <p>The facilitator may compute an equally-weighted average of the density functions (for instance by using the “SHELF” software). This should NOT be revealed to the experts, but may be used at the facilitator’s discretion in the later stages.]</p> <p><i>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.</i></p> <p><i>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.</i></p>
Group discussion	<p>The experts now discuss the fitted distributions, with a view to understanding each expert’s reasoning for their judgements, and to share experience and interpretations of the evidence. A digest of the discussion should be recorded here, using anonymised codes for the experts.</p> <p>[The facilitator should prompt debate around differences between experts. Managing the group discussion is a very important skill for the facilitator. Refer to the document “Facilitator Skills” for guidance.</p> <p>The record of the discussion in this form should be sufficiently detailed to cover the main arguments advanced by the experts, without being unnecessarily long.</p> <p>Note that this is the time for experts to give their opinions about the quality and interpretation of the evidence. (They were not allowed to do so before making their individual</p>

	<p>judgements.) This is also the time to introduce any opinions that have been solicited from experts who are not present.]</p> <p><i>The group discussion is a feature of the SHELF approach. It has the benefit of allowing a synthesis of the experts' knowledge, but there are psychological influences that pose challenges for even experienced facilitators. These are covered in some depth in the document "Facilitator Skills".</i></p>
Group plausible range	<p>The experts make group "consensus" judgements of the plausible limits for X.</p> <p>[It is particularly important to begin by explaining to the experts the meaning of "consensus" judgements. Their group "consensus" judgements should be such as it would be reasonable for a Rational Impartial Observer to make having seen their individual judgements and heard their discussion. Refer to the slide set "RIO" for a full explanation of this perspective. The experienced facilitator may prefer to present the RIO perspective in their own way, but otherwise it is recommended that the slide set should be used directly as a presentation to the experts.</p> <p>The experts may take the view that RIO would consider that the plausible range should include the plausible range for each expert individually, but a narrower range may be acceptable in the light of their discussion.]</p>
Group elicitation	<p>[It may happen that the experts' individual fitted distributions were relatively similar, and that in the group discussion no great divergence of opinion emerges. In such a situation, the facilitator may propose that in place of the formal group elicitation the experts may instead agree to adopt the average of their densities (the linear opinion pool) as their "consensus" judgements. If so, simply record the Method in this box as "Linear Pool", and show the averaged density as the fitted distribution in the "Fitting and feedback" box below.]</p> <p>Method: Record the method for group elicitation – P(robability), T(ertile) or Q(uartile).</p> <p>[The method will have been chosen in advance by the facilitator. Refer to the document "SHELF Methods" for guidance on making this choice.]</p> <p>Judgements: The experts now make "consensus" judgements according to the chosen method.</p> <ul style="list-style-type: none"> For the probability method, three probabilities P1, P2 and P0 corresponding to three values of X (X1, X2 and X0) chosen by the facilitator.

	<ul style="list-style-type: none"> • For the tertile method, first their median M and then their tertiles T1 and T2. • For the quartile method, first their median M and then their quartiles Q1 and Q3. <p>[For the probability method, refer to the document “SHELF Methods” for guidance on choosing the three values X1, X2 and X0.</p> <p>For the other methods, refer to the slide sets “Median”, “Tertiles” and “Quartiles” (as appropriate) for ways to explain the judgements to the experts. If using the same method as for individual judgements, a simple reminder should be adequate. Otherwise the slide sets may be used directly as a presentation to the experts.]</p> <p><i>The “SHELF Methods” document covers the reasoning behind each of the different methods, and how they are based on research into the psychology of expert judgement.</i></p>
Fitting and feedback	<p>Record here the (potentially iterative) process of fitting, feedback and revision of the group judgements.</p> <p>[The facilitator first fits a distribution to the group’s “consensus” judgements. 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 10th and 90th 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.]</p>
Chosen distribution	Record and show here the finally fitted distribution.
Discussion	<p>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.</p> <p>[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</p>

	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.