1. Expert Judgment Elicitation Protocols

Michela Catenacci, Valentina Bosetti, Giulia Fiorese and Elena Verdolini

While past data analysis can provide extremely valuable information on past dynamics and technology development, it is paramount to recognize the distorting effect of uncertainty on innovating firms’ behaviors and its determining role on current and future innovation processes. In order to account for uncertainty and to fill the lack of empirical or modeling data, the ICARUS project resorted to experts’ elicitations, which have been successfully used to collect information on future trends of technology costs.

Expert judgments are the expressions of informed opinion that experts make based on their knowledge and experience with respect to technical problems (Hogarth, 1987; Morgan and Henrion, 1990; Cooke, 1991). Eliciting experts’ judgments means collecting subjective probabilities that a specific event will take place in the future, through specific methods of verbal or written communication. Experts’ judgements are particularly useful and are often required in probabilistic decision-making and in the evaluation of risks. They can fill the lack of information or complement other available data based on models’ predictions, thus providing an additional source of information.

Expert judgement elicitation has been successfully used in the past to inform policy-makers, especially in the field of energy (Apostolakis, 1990). One prominent example in this respect is the study by the European Commission and the United States Nuclear Regulatory Commission during the 1990s focusing on nuclear power plants and the uncertainty surrounding accident consequence codes (Cooke and Goossens, 2004).

The literature on decision analysis provides interesting theories on the techniques that should be applied to elicit expert judgments under uncertainty, support risk evaluation and inform a transparent decision-making process, especially if historical data is scarce and cannot inform on future developments, as is the case for energy projections (Morgan and Keith, 2008), nuclear engineering (Cooke and Goossens, 1999), climate change impacts and policy analysis (Morgan and Keith, 1995) and environmental policy in general (Morgan and Henrion, 1990).
Applications to clean energy technologies started only recently. The methodology described in the present volume, and applied throughout the ICARUS project, was structured considering and complementing the existing applications of expert elicitation to low-carbon energy technologies, which are still relatively scarce but can provide useful insights regarding many important uncertainties in policy analysis. Baker et al. (2009a) and Curtright et al. (2008) use expert elicitation to analyse the uncertain role of RD&D investments in leading solar technologies to the commercial success. Baker and Keisler (2011) apply the same techniques to assess the effect of RD&D funding on the factors that determine the cost of cellulosic biofuels, while Baker et al. (2009b) and Chan et al. (2011) focus on carbon capture and storage technologies. Finally, Baker et al. (2010) ask experts’ probabilities on the future of battery technology for electric drive vehicles.

Our protocol complements these applications with a number of innovative elements. First, our protocol involves a larger number of experts, with different knowledge and professional background, in order to gain a clearer vision of each technology, from research to demonstration, to deployment and to diffusion. Second, this represents the first elicitation of European experts on energy technologies. Previous applications of expert elicitation techniques to the analysis of carbon-free energy technologies focused almost entirely on the USA, thus ignoring important players in the innovation process and important markets for the deployment of these technologies. Third, this elicitation broadens the elicitation scope to the potential of carbon-free technologies in Europe and worldwide. While in ICARUS the experts are asked to assess the potential for cost reductions conditional on RD&D investments in European countries, they are also invited to consider the potential for technology deployment and diffusion worldwide. In particular, looking into knowledge spillovers and technology transfer dynamics, experts indicate where commercial breakthrough is expected to occur, and how each technology would spread across different countries and regions of the world. Fourthly, the protocols consider both mature and new generation technologies by assessing the evolution of a broad set of options for each technology (e.g. when assessing the production electricity from solar technologies, we considered both mature technologies such as Crystalline-silicon photovoltaic (PV) and Concentrated Solar Power, and innovative technologies such as Third Generation PV and Organic PV). Conversely, most existing studies focus on more diffused and developed options. Finally, by investigating the non-technical conditions that could set back technological diffusion into the market, the elicitation goes beyond cost reductions conditional on RD&D investment levels, and looks into further barriers that could prevent diffusion even if the technology became cost-competitive with traditional options.
Since the seminal work by Tversky and Kahneman (1974), a growing body of literature has substantiated our knowledge on heuristics and biases which affect judgments under uncertainty. In spite of their deep and recognized knowledge of the subject, experts can be subject to the same cognitive and motivational biases as all human beings. Such ‘bias’ refers to a skewing of the experts’ estimate from some reference point and can be of different nature. In presence of uncertainty, for example, people often anchor their response to the ease with which they can imagine an event occurring or to some ‘central’ value. Biased estimates might result also from the heuristic procedure of ‘availability’, linked to the ease with which experts could imagine an event occurring, or from the ‘representativeness’ of an event, which can sometimes influence its judged probability. Experts are also often overconfident in the precision of their estimates, underestimating their uncertainty.

Protocols and techniques have been defined to minimize such biases (Clemen and Reilly, 2001; Keeney and von Winterfeldt, 1991; Meyer and Booker, 1991; Morgan and Henrion, 1990; O’Hagan et al., 2006; Phillips, 1999; Walls and Quigley, 2001). Some crucial features that should be carefully included in a robust elicitation process are: carrying out the survey in person or by ensuring a direct connection with each participant, defining the metrics and the object of the elicitation process as accurately as possible, warning the experts about main biases and heuristics and training them to the elicitation exercise, avoiding anchoring numbers and using tools to represent uncertainty that facilitate its quantification.

We carefully addressed each of those issues and designed two different protocols, considering the complexity of the analysis, the presence of multiple interests and perspectives and the need to account for the uncertainty characterizing technological development. The first protocol (protocol A), described in the next section of the volume, was defined and applied to elicit experts’ information on a group of technologies: solar photovoltaic and concentrating solar power technologies, biomass for the production of electricity, biofuels for transportation and batteries for electric drive vehicles. The second protocol (protocol B) was structured in collaboration with a group of researchers of the Harvard University to carry out a parallel analysis of nuclear energy technology in the European Union and in the United States.

Both protocols included, as the first step, a careful selection of the experts, to gather a mix of respondents with strong scientific backgrounds and sound empirical knowledge. Indeed, since the success of the elicitation crucially depends not only on the expertise and the technical background, but also on the personality of each expert as well as on her ability to provide probabilistic judgments (O’Hagan et al., 2006), our selection managed to engage a balanced group of experts who represented the major perspectives and fields.
of knowledge (engineers, economists and policy makers) but with heterogeneous backgrounds. Academia, institutions and the private sector were represented in a balanced way to ensure a thorough analysis of both basic and applied research issues as well as policy implications. The level of expertise of each selected expert was carefully assessed considering tangible evidence such as publications and direct involvement in projects related to research and development of the technology under assessment. The experts’ profiles were compiled according to the above exposed criteria, and then a first core group of experts was identified. Those experts were called ‘seeds’, and they were asked to point out other experts to involve in the elicitation exercise, according to the so-called ‘snowball sampling technique’ (Salganik and Heckathorn, 2004; Giupponi et al., 2006). The basic idea behind this technique is that respondents are not selected from a sampling frame, but from the linked network of existing members of the sample. This method was used to complete a sound and reliable list of experts for each technology, ensuring the inclusion of all the relevant perspectives.

1.1 PROTOCOL A: INTERVIEWS WITH A DOUBLE QUESTION ON COSTS

As described in Figure 1.1, protocol A was composed of three parts: an introductory phase where we selected the main components of the elicitation process, a ‘warm up’ phase were we discussed with each expert on the purpose of the exercise, on the applied methodology and on the analyzed technology, followed by the submission of the questionnaire through individual interviews.

In the introductory phase, we firstly defined the purpose of the exercise and then selected the specific issues for investigation. Within each topic, we identified concrete and detailed questions. We selected individual interviews as the most suitable elicitation approach in our case. While the literature considers group interaction among experts as the best method for gathering a large amount of qualitative data, it also suggests individual interviews as the most suitable approach to obtain detailed quantitative data as well as to analyse the expert’s problem-solving process (Meyer and Booker, 1991). Moreover, the face-to-face interviews avoid the potential biases from group dynamics which can break the link between subjective perception and underlying objective reality (Baddeley et al., 2004). Unfortunately, face-to-face elicitations are extremely expensive and time-consuming. As a consequence, they are often substituted with less effective techniques, such as the submission of questionnaires by email or by phone. These modes of communication do not allow the analyst to carefully monitor the whole
exercise and to control the occurrence of biases and errors in the estimates (O’Hagan et al., 2006).

The ICARUS project gave us the possibility to test an effective solution aimed at combining the advantages of a face-to-face interview, with the possibility to save money, time and carbon emissions. We carried out half of the interviews in person, and half through web conferences, using meeting software that could be accessed from everywhere, where the expert was invited to join an online meeting, and we could present information and share documents and applications. The results of the two methods were equally successful and satisfying both for the experts and for the research team.

To test the questionnaires and the elicitation process, we carried out a few pilot interviews with experts from the academia and the private sector. This process was crucial to refine and condense the various sections of the questionnaire. We started each interview with a warm-up phase designed to share with the experts the available data from the literature on the current trend of public RD&D investments, on the evolution of costs and on the state of development of each technology. Experts were also carefully briefed on sources of biases and difficulties in assessing probabilities.

Subsequently, experts were asked to self-assess their expertise on a scale from one to five with respect to the families of technologies included in each survey. This step was fundamental because it provided insights on the possible biases in experts’ responses, which could be dictated by their

Figure 1.1 Structure of protocol A and questionnaire’s contents
preference or better knowledge of a specific technology among the ones surveyed.

With the aim of ‘tuning in’ the experts and avoiding the possibility of overlooking or exaggerating problems, we asked them to carefully assess the current status and cost of the different technologies and all potential barriers to cost reduction and technology diffusion. We put specific emphasis on each different cost component and asked each expert to carefully consider the combination of events which would lead to cost improvements. To help them concentrate on the role of RD&D investments in the success, we let them allocate a hypothetical budget among the possible technical choices within each technology.

The core part of the questionnaire asked the experts to assess if and under what conditions the costs of each technology would eventually become competitive with traditional fossil fuel technologies. The estimates of the 2030 cost were collected conditional on three alternative scenarios of public EU RD&D funding. The first scenario assumed that the current annual EU public RD&D level of investment as a share of GDP would be maintained throughout 2030. The second scenario assumed an immediate 50 percent increase of public EU RD&D investment, then sustained until 2030. In the third scenario, annual public EU RD&D would scale up to twice the current levels and be maintained there up to 2030 (Scenario ‘+100% RD&D’). We asked the experts to provide cost estimates assuming that annual expenditure would be constant between 2011 and 2030.

To control for different sources of bias, we chose to elicit cost projections through a combination of different methods applied in the literature: we started by directly asking the experts to estimate future costs in the three RD&D funding scenarios. To control for cognitive heuristics (such as availability, anchoring and overconfidence) we asked the experts to provide the 10th, 50th and 90th percentiles estimates, with the 50th percentile representing the most likely value.

Since experts typically think in terms of technological endpoints and not in terms of electricity costs, we provided them with a formula deriving the cost of the technology from specific technical factors (see the questionnaire in the Annex chapter for the formulas applied in the solar survey). Experts who did not feel at ease with directly providing monetary estimates were free to use spreadsheet to compute electricity costs given alternative assumptions about key technical characteristics, and therefore to estimate how improvements in technical factors would result in lower monetary costs.

To verify the consistency of the experts’ responses, we also asked each expert to estimate the probability that the cost of the technology in 2030 would be lower than some representative threshold, conditional on the same alternative RD&D investment scenarios outlined above. Asking experts the
follow up question on the likelihood of reaching specific cost targets allowed
us to control for the risks linked with the direct elicitation of subjective
probabilities. One of the most important risks is the possibility to collect
highly context-dependent answers which do not represent a true, mathematical,
probability measure (Meyer and Booker, 1991). Examples of distortion in the
elicited probabilities are: treating low-probability events as impossible;
overconfidence and other anchoring effects; and overweighting salient events
(O’Hagan et al., 2006). This part of the question also helped to increase the
amount of elicited information, and deepened the discussion with the experts,
hence improving our perception of the experts’ beliefs. In cases where the
elicited cost value and the probability were inconsistent, we engaged the
experts in follow-up questions to verify what the true estimate was.

The last section of the questionnaire finally addressed the technology
transfer dynamics, knowledge spillovers, externalities and market barriers
which could affect technology diffusion, and therefore the potential for
success of each technology through multiple-choice and open questions. This
allowed us to collect crucial ancillary information on the experts’ knowledge
and perspective.

1.2 PROTOCOL B: WEB QUESTIONNAIRE AND
WORKSHOP

To gather information on the RD&D needs for the future of nuclear energy,
the future cost and performance of nuclear technologies, and on the major
barriers to large-scale deployment of nuclear energy, we conducted a survey
of nuclear experts in coordination with Harvard University. The survey asked
experts how much they would recommend that their governments spend on
nuclear energy RD&D, what progress in cost and performance might be
expected by 2030 if those recommendations were followed and what other
factors might constrain or promote future nuclear energy growth.

We carefully selected leading experts from the European Union (EU) and
the United States (US) to take part in the expert elicitation surveys during the
summer and fall of 2010. In order to reach a wide cluster of experts from all
over the EU and US, we chose to submit web questionnaires.

To reduce bias in the experts’ estimates, the introductory section of the
questionnaire was designed to train the experts on probability and bias
concepts. Specific response modes – in particular the use of percentiles to
capture uncertainty versus the probability of meeting certain cost goals; the
graphic supports for the relationship between cost, RD&D levels and
uncertainty; and the graphic supports for budget allocation – were chosen to
facilitate speedy completion and correct interpretation of the elicitation. The
feedback from experts on these visual tools was positive. After experts completed the online elicitation, they received a summary of all the answers, giving them a chance to provide feedback.

Figure 1.2 summarizes the structure of the process followed to conduct the surveys, while Figure 1.3 shows the structure of the online elicitation instrument.

Experts’ judgments can be influenced by clarity and survey design, as well as by motivational biases caused by incentives or pressures that lead people to provide answers that do not entirely reflect their beliefs. Sources of motivational biases can include, for example, social pressure, from the interviewer or from the group of experts (‘group think’). The online questionnaire allowed us to avoid the group-think bias, and was designed to minimize the occurrence of cognitive biases in the experts’ probability estimates. To address these biases, our survey instrument included, among other things: (a) a section encouraging experts to think about the lowest possible and highest possible costs and about the range of events that may affect costs before asking them to introduce their 50th, 10th and 90th percentiles; (b) a background information section to ensure that experts had a range of information available to them to help them overcome anchoring (though the background information could have also provided a common set of anchors, most experts projected costs that were higher than those presented in the background section, indicating that the experts did not anchor strongly enough).

---

### Figure 1.2 Structure of protocol B employed in the design of the online elicitation and group discussion

<table>
<thead>
<tr>
<th>Introductory phase</th>
<th>Individual questionnaires</th>
<th>Workshop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experts selection and engagement</td>
<td>Definition of the questions’ areas</td>
<td>Discussion on the main thematic areas and results</td>
</tr>
<tr>
<td>Pilot tests and corrections</td>
<td>Submission of online questionnaire</td>
<td>Experts’ revision of their answers</td>
</tr>
<tr>
<td></td>
<td>Follow up summary reports and phone calls</td>
<td>Presentation of the final results</td>
</tr>
</tbody>
</table>

---
Figure 1.3 Structure of the individual online elicitation instrument

on the numbers presented in the background information); and (c) a section explaining overconfidence with a historical example of overconfidence in estimates.

The elicitation instrument included several graphic displays to help experts visualize their answers. For example, experts were able to see their 10th, 50th and 90th percentile estimates of cost for different RD&D levels in a single graph and had 100 chips (representing percentages of their recommended RD&D investment) that they could allocate in a ‘board game’ that included different nuclear technology areas and innovation stages. The design of the survey involved several choices, including expert selection, the mode of the survey, the structure of the elicitation, the visualization of the answers and the phrasing of the questions, among others. These choices were made over a couple of months with the help of two nuclear experts, who tested the surveys and provided feedback on everything from terminology to the appropriateness and clarity of questions and the tool.

Among other sources of bias, the experts’ background and expertise is obviously a key driver of their judgment. Motivational biases could have come into play if experts thought that recommending large amounts of RD&D funding for their areas may result in more funds for research in their area. To determine whether experts consistently recommended larger RD&D investments for their areas of expertise, we included a section in the survey asking experts to rate their level of expertise on different reactor technologies and cross-cutting research areas.
In April 2011, the research teams held a workshop in Venice (Italy) with a subset of the participating EU and US experts to present and discuss the results of the elicitations in an effort to understand where there is consensus and where the most important disputes and uncertainties lie. The group workshop allowed us to verify the online survey findings, discussing areas where experts may have had different interpretations of the questions, and allowing the experts to modify their views through discussion and debate.

The two-step procedure, soliciting experts individually and then following up in a group discussion, made it possible to identify key issues that could arise when each of the two methodologies is followed as a stand-alone procedure and was a relatively cost-efficient protocol.

NOTES

1. The expert had therefore the opportunity to compile the questionnaires online, while discussing questions and commenting the survey. As she/he modified the document, we could see changes and intervene in real time to clarify doubts or to ask for more specific information. All interviews could be taped not to lose the ancillary information, which could be used to check and contextualize the experts’ answers and, if necessary, to structure follow-up phone calls. The expert could see the interviewers through the webcam, and this helped to recreate the overall effect of a face-to-face situation. The web meeting approach resulted effective and user-friendly, ensuring the defensibility of the elicitation outputs and the robustness of the consequent analyses, at the same time making a larger number of interviews and follow-ups possible.

2. Also Baker et al. (2009a) elicit experts’ probabilities. They surveyed US experts in second generation PV technologies and, by assessing the evolution of specific technical endpoints, such as efficiency, lifetime and cost of manufacturing over the next 40 years, they collect probabilities associated to the 2050 costs of electricity being below a certain threshold.

REFERENCES


Baker, E., H. Chon and J. Keisler (2010), ‘Battery technology for electric and hybrid...