Energy Technology Innovation Policy

Energy Technology Expert Elicitations for Policy:

Workshops, Modeling, and Meta-analysis

By Laura Diaz Anadon, Valentina Bosetti, Gabriel Chan, Gregory Nemet, and Elena Verdolini

HARVARD Kennedy School
Belfer Center
for Science and International Affairs

Fondazione Eni
Enrico Mattei

October 2014 | #2014-08

Belfer Center for Science and International Affairs
Fondazione Eni Enrico Mattei

Discussion Paper #2014-08

John F. Kennedy School of Government
Harvard University
79 JFK Street
Cambridge, MA 02138

Laura Diaz Anadon, Valentina Bosetti, Gabriel Chan, Gregory Nemet, and Elena Verdolini

Diaz Anadon, Laura, Valentina Bosetti, Gabriel Chan, Gregory Nemet, and Elena Verdolini. "Energy Technology Expert Elicitations for Policy: Workshops, Modeling, and Meta-analysis." Discussion Paper, Belfer Center for Science and Inernational Affairs, Harvard Kennedy School. October 2014.

This paper has also been published as FEEM *Nota di Lavoro* 90.2014

Cover image: Andrew Facini, Belfer Center for Science and International Affairs

Energy Technology Innovation Policy

Energy Technology Expert Elicitations for Policy:

Workshops, Modeling, and Meta-analysis

By Laura Diaz Anadon, Valentina Bosetti, Gabriel Chan, Gregory Nemet, and Elena Verdolini

HARVARD Kennedy School
Belfer Center
for Science and International Affairs

Fondazione Eni
Enrico Mattei

October 2014 | #2014-08

Abstract

Characterizing the future performance of energy technologies can improve the development of energy policies that have net benefits under a broad set of future conditions. In particular, decisions about public investments in research, development, and demonstration (RD&D) that promote technological change can benefit from (1) an explicit consideration of the uncertainty inherent in the innovation process and (2) a systematic evaluation of the tradeoffs in investment allocations across different technologies. To shed light on these questions, over the past five years several groups in the United States and Europe have conducted expert elicitations and modeled the resulting societal benefits. In this paper, we discuss the lessons learned from the design and implementation of these initiatives in four respects. First, we discuss lessons from the development of ten energy-technology expert elicitation protocols, highlighting the challenge of matching elicitation design with a particular modeling tool. Second, we report insights from the use of expert elicitations to optimize RD&D investment portfolios. These include a discussion of the rate of decreasing marginal returns to research, the optimal level of overall investments, and the sensitivity of results to policy scenarios and selected metrics for evaluation. Third, we discuss the effect of combining online elicitation tools with in-person group discussions on the usefulness of the results. Fourth, we summarize the results of a meta-analysis of elicited data across research groups to identify the association between expert characteristics and elicitation results.

Keywords

Expert elicitations, energy technology innovation, public R&D, meta-analysis, optimization

Table of Contents

1. Introduction

Governments throughout the world justify their investments in energy technology research, development and demonstration (RD&D) (Chan, Anadon, Chan, & Lee, 2011) on the basis of three broad public policy challenges—environmental externalities, energy security, and economic competitiveness (Anadon, 2012)—in addition to the knowledge spillovers associated with scientific research in general (Arrow, 1962). Country members of the International Energy Agency[1] invested $13.7 billion PPP in public energy RD&D in 2008, which rose to $17 billion PPP (Purchasing Power Parity) in 2012 (IEA, 2013).[2] A recent review of the largest developing countries (Brazil, Russia, India, Mexico, China, and South Africa) indicates that in 2008, public energy RD&D was of a comparable scale to IEA countries, totaling $13.8 billion PPP (Gallagher, Anadon, Kempener, & Wilson, 2011).

While total energy RD&D investments are smaller than public subsidies for energy deployment,[3] the relative social benefits of RD&D investments may be larger than that of subsidies. The relatively large returns to energy RD&D are due to the long-term, high risk, and skewed benefits associated with the innovation process (Nemet, 2013). Based on this view, since 1996 expert panels in the United States (American Energy Innovation Council, 2010; NCEP, 2004; NCEP, 2007; PCAST, 1997; PCAST, 2010) and in the European Union (EERA, 2010; European Commission, 2007) have called for significant increases to public energy RD&D investments. These studies, however, offer little analytic support to justify their recommendations and often do not include substantiated estimates of benefits, costs, and associated uncertainties.

The U.S. Department of Energy (DOE), the largest single funder of energy RD&D in the United States, often conducts estimates of the expected benefits of individual RD&D programs. However, the DOE does not consistently evaluate the positive and negative interactions of its programs across its investment portfolio; for example, energy storage may complement intermittent renewables. Nor does it systematically consider uncertainty in its benefit calculations. In short, the DOE does not conduct robust consistent and transparent cost-benefit analysis to support its portfolio of RD&D investment decisions in different technology programs.

As a result of some of these shortcomings, a 2007 study of the National Research Council recommended that the DOE make probabilistic assessments of the benefits of RD&D programs when making decisions (NRC, 2007). For a short review of the literature estimating the benefits of R&D investments in energy the reader is referred to the Supplementary Information (SI). But the

[1] The IEA has 28 Member countries (Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, Republic of Korea, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom, and United States), but no data are provided for Luxembourg or the Slovak Republic. Iceland, Chile, and Mexico are OECD members, but are not IEA members.

[2] In 2012, the United States alone invested just over $4.7 billion PPP, while European countries' invested totaled $5.8 billion PPP.

[3] The United States government spent about $33.2 billion in 2010 in energy subsidies for deployment (EIA, 2011). A recent report estimated that global energy subsidies for deployment in 2007 were $483 billion (IEA, OPEC, OECD, & World Bank, 2010).

political economy conditions within an RD&D funding organization make generating credible estimates of the impact of RD&D more difficult. For example, in the case of DOE, competition between the different technology programs creates incentives for self-serving biases and erodes trust between programs. One strategy that appears feasible given DOE's existing organizational incentives is eliciting the knowledge to develop technical assumptions from external (as well as some internal) experts and integrating this knowledge into internally-acceptable assessment frameworks (Chan & Anadon, 2013).

In this vein, research groups at the Harvard Kennedy School (HKS) and at Fondazione Eni Enrico Mattei (FEEM) recently conducted expert elicitations to estimate the relationships between public RD&D investments and technology outcomes (costs and performance). The main objective was to provide insights to both DOE and EU policy makers about the allocation of RD&D funding across several technology areas: nuclear power, solar photovoltaics, concentrated solar power, biofuels, bioelectricity, vehicles, utility scale energy storage, and fossil power with and without carbon capture and storage.[4] [5] Elicitations for the US were carried out between 2009 and 2011 and were designed so that their results could be used in MARKAL (Fishbone & Abilock, 1981), a widely -used energy-economic model, to provide insights about DOE funding decisions across different programs. Elicitations for the EU were carried out by FEEM between 2009 and 2011 within the FP7 project ICARUS and designed for use in WITCH (www.witch-model.org), an integrated assessment energy model.

This paper discusses lessons emerging from these data collection and modeling efforts regarding how expert elicitations can be designed, implemented, and utilized to support decisions about the allocation of public energy RD&D investments. It also includes insights and findings from a meta-analysis of the nuclear technology elicited data, identifying how elicitation design affects results. The rest of the paper is structured as follows. Section 2 presents a literature review on the previous use of expert elicitations for energy technologies. Section 3 describes the methods used in this research, in particular, the design and implementation of expert elicitations in an energy-economic modeling context (MARKAL and WITCH) conducted by the Harvard and the FEEM groups, respectively. Section 4 discusses key insights from the analysis organized in five sub-sections. Section 5 concludes with a summary of findings and thoughts for future research.

2. Expert Elicitations of Energy Technologies to and RD&D investment decisions

Estimating the benefits of energy RD&D investments requires estimation of two relationships. First is the relationship between a given RD&D investment and individual technology outcomes, which are typically measured in terms of cost or performance. Second is the relationship between

[4] Theee results from these elicitations can be found at: Anadon et al., 2011; Anadon et al., 2012; Bosetti, et al., 2012; Bosetti et al., 2012; Catenacci et al., 2013; Chan et al., 2011; Fiorese et al., 2013.

[5] Some similar studies also utilized energy technology expert elicitations, but were not explicitly developed to provide insights about portfolios of investments at a large scale (e.g., for technology programs funded by DOE or the EU Commission) or across multiple technologies (Baker, Chon, & Keisler, 2009a; Baker, Chon, & Keisler, 2008; Baker, Chon, & Keisler, 2009b; Curtright, Morgan, & Keith, 2008).

the technology outcomes and policy goals, such as economic growth, energy prices, CO_2 emissions, or oil imports.

Expert elicitations are being increasingly used to estimate the first relationship (Anadon et al., 2011; Anadon et al., 2012; Bosetti et al., 2012; Catenacci M. et al., 2013; Chan et al., 2011; Fiorese et al., 2013); Baker, Chon, & Keisler, 2009a; Baker, Chon, & Keisler, 2008; Baker, Chon, & Keisler, 2009b; Curtright, Morgan, & Keith, 2008). These studies gather the opinions of experts on technical questions that fall within their area of expertise. Data collection is carried out using elicitation protocols carefully designed to reduce biases (Cooke, 1991; Hogarth, 1987; Morgan & Henrion, 1990; Evans, 2013).

However, few studies have designed elicitations with the objective of supporting specific energy RD&D policy decisions on a continuous basis. In addition, even though previous studies indicated the importance of protocol design and expert selection as key for elicitation results, (Keeney & Winterfeldt, 1991; Meyer & Booker, 1991; Raiffa, 1968) there are no empirical assessments in energy of the impact and size of differences in elicited results from expert selection and elicitation design (e.g., whether the survey is conducted in person, via mail, or online).

Expert elicitation estimates can also be used to inform the second relationship by introducing them as inputs to technologically-detailed models of the economy, thus linking technology outcomes to social benefits. Such an approach allows decision makers to understand how technological uncertainty propagates from the first relationship through the second relationship, providing important insights on the distribution of outcomes from RD&D. This type of two-stage analysis to support policy decisions is common in environmental policy decisions, such as those that Fann et al. (2013) inform with their analysis of approaches to estimate concentration-response functions for $PM_{2.5}$.

3. Methods

3.1 Design and implementation of expert elicitations

The Harvard studies were designed to inform the DOE on the allocation of RD&D investments across large scale technology programs, while the ICARUS project, funded by the European Research Council aimed at designing optimal allocation of the EU research budget on energy technologies, with specific attention to the role of European climate and energy policies. We highlight here some key features of both data collection efforts.

Each institution conducted six elicitations. Four of Harvard's elicitations were distributed by mail (bioenergy, utility scale storage, fossil energy and carbon capture and storage, and vehicles) and the remaining two were online (nuclear power and solar PV). Four of FEEM's elicitations were extensive in person interviews (batteries for EDV, bioenergy, biofuels, solar) and two were online (carbon capture and storage, and nuclear power).

The core objective of the elicitations in both cases was to gain insights on the relationships between public RD&D investments and technological change for specific technologies in a parameterization that could be naturally introduced into an economic model of aggregate benefits. Specifically, the Harvard elicitations included questions on experts' estimates about various technology specific cost components (e.g., overnight capital cost, operations and maintenance costs) and performance parameters (e.g., efficiency, yield, fuel efficiency) in 2030 under different DOE RD&D budgets. Two exceptions were the bioenergy survey—in which experts were given the option of providing a cost breakdown or providing an overall cost per unit of biofuel or electricity delivered—and the vehicles survey—in which experts were asked about the total purchasing cost of different types of vehicles and specific performance characteristics without a breakdown of cost components (e.g., battery cost).

FEEM's elicitations of batteries for electric vehicles, bioelectricity, biofuels, and solar power asked experts to provide an aggregated metric of the 2030 cost under different EU RD&D budgets. FEEM's carbon capture and storage (CCS) survey investigated both the cost and energy penalty of alternative CCS technologies. Finally, FEEM's nuclear survey was conducted in coordination with the Harvard study and used a two-step methodology combining an online individual elicitation with a workshop in which a subset of experts participated.[6]

The choice of the media for the elicitation is an important one which involves tradeoffs. Among other possible benefits, in-person interviews imply greater interaction between the expert and the researcher and can reduce biases and availability heuristics. However, conducting online or mail elicitations reduces costs, for both respondents and researchers, and increases flexibility, thereby expanding the pool of participants.

Independent of the media chosen, developing the elicitation protocol took around 3-5 months for both research groups, consistent with previous energy technology expert elicitations. A crucial step was testing and revising the elicitation protocol through pilot interviews and an iterative process in which a few experts on a given technology were involved.

In line with the literature, the elicitations included a background calibration section which contained a summary of the purpose of the survey, background information on either DOE's or EU current activities and investments in the technology of interest, and a statement about avoiding bias and overconfidence. All Harvard surveys and most of FEEM's surveys also asked participants to rate their own expertise in several sub-technology areas on a 6-point scale, where 6 was described as "I am one of the top experts in this technology/system" and 1 was described as "I am not familiar with this technology/system." This information was subsequently used to test for correlations between areas of expertise and recommendations for RD&D funding or particularly optimistic technology forecasts, which would have suggested experts make self-interested recommendations. The SI includes for more information on details on the elicitations.

The second half of the protocol contained the core questions of the elicitations. The Harvard studies included four sections with questions on: (1) the commercial viability, cost and performance of different technologies in 2030 under a business as usual (BAU) public RD&D funding scenario; (2) the expert's recommendation of total public investments in the technology area of interest and their recommended allocation of funds to sub-technologies, including questions about the specific technical hurdles to be addressed by their allocation; (3) how future technology costs and performance would change if their recommended RD&D investments were implemented, and how this would change under alternative RD&D investment levels; and (4) other

[6] In the survey and the workshop experts were asked questions about cost components and different performance parameters. For more detailed information the readers are referred to the papers on the Harvard and FEEM elicitations provided in Section 1.

technology-specific policies and factors affecting technology deployment. We also considered using self-rated expertise to weight experts, but we ultimately did not conduct this analysis.

The FEEM elicitations on batteries for EDV, bioenergy, biofuels, and solar, asked experts to (1) assess different technological options based on their level of maturity and possible bottleneck; (2) suggest a breakdown of public research expenditures across the different technological paths that would maximize the chance of a breakthrough; (3) provide estimates of future costs and the surrounding uncertainty conditional on different levels of public RD&D investment;[7] (4) assess the potential additional bottlenecks that additional RD&D investment could not address (i.e. concerns about competition of biofuel with food for land); and (5) assess the potential international diffusion of a given technology, if cost-competitive, to both OECD and non-OECD countries.

All elicitations included interactive visual aids. The Harvard mail surveys included a set of chips and a "board game" to help experts think through allocating their recommended budget across different technology areas and technology development "stages".[8] The Harvard and FEEM online surveys included a virtual game board and chips as well as graphical feedback for all quantitative input from the experts, allowing them to visualize probability distributions of their cost and performance estimates under the different RD&D scenarios.[9] The FEEM in person surveys allowed experts to plot their cost estimates in real time and check for the consistency of their own answer.

To evaluate the effectiveness of online elicitations, FEEM and Harvard conducted nearly identical elicitations in nuclear energy. Following the surveys, the groups convened a subset of the European and U.S. experts for a 1.5-day workshop to discuss the results of the survey and to bring forward any questions or misunderstandings that surfaced during the online elicitations. Experts discussed their answers and talked through their disagreements regarding the interpretation of the questions. Following each session of the workshop, experts were given the opportunity to privately change their answers to the survey.

Finally, both research groups worked at connecting the technical outcomes and/or the costs and uncertainty estimates to societal benefits (e.g., CO_2 emissions, energy costs, oil imports, etc.). The Harvard group selected the MARKAL model, while the FEEM worked with the WITCH model. MARKAL is a bottom up energy-economic model that is publicly-available and has institutional buy-in from many government agencies in the US and elsewhere. The use of MARKAL was coupled with an importance sampling technique which allowed changing input assumptions without requiring additional model runs, thus solving a computational constraint

[7] The FEEM elicitations asked the expert to first provide estimates of the 10th, 90th and 50th percentile of future costs. The same experts were subsequently asked to provide probabilities that under the same different RD&D scenario the cost of a given technology would be below some level chosen by the researchers. This effectively meant eliciting the same information twice, but under different format, and allowed to check the consistency of expert's responses.

[8] The Harvard board game included 100 poker chips, one for each percentage of their total recommendation, that experts allocated across sub-technology areas, which included an "other category" that allowed them to indicate additional areas. The stages of RD&D that experts could allocate across were basic research, applied research, pilots, and demonstration.

[9] The graphical feedback on the online surveys included plots of the 90th, 10th, and 50th percentile estimates for each technology and different budget scenarios, allowing experts to modify their answers as they were filling out the graphs in real-time.

faced by many decision-making entities (Pugh et al., 2011).[10] Because of this method's ability to test different input assumptions, the benefits associated with RD&D investments under more optimistic or more pessimistic experts' assumptions were estimated.[11]

3.2 Meta-analysis of expert elicitations

Anadon et al. (2013) conducted a meta-analysis of three recent nuclear expert elicitations (Abdulla, Azevedo, & Morgan, 2013; Anadon et al., 2012) given the scarcity of information regarding the impact of expert selection and elicitation design on elicitation results. Meta-analysis is a set of statistical techniques used to reconcile and aggregate the results of multiple studies testing similar hypotheses and to thus enhance the overall reliability of findings (Borenstein, Hedges, Higgins, & Rothstein, 2009; Glass, 1976). Systematic reviews and meta-analyses, which typically follow very strict rules in healthcare applications, are very systematic and time consuming (Morton, 2013). Meta-analysis complements the qualitative insights about expert selection and elicitation design and has been used in environmental economics since the 1990s (Matarazzo & Nijkamp, 1997; Nelson & Kennedy, 2009), with several recent applications in energy (Barker & Jenkins, 2007; Rose & Dormady, 2011; Zamparini & Reggiani, 2007).

Using the individual elicited values from multiple elicitations,[12] (Anadon et al., 2013) estimate how public RD&D investment affects experts' 2030 central estimates (50th percentile) and the uncertainty (defined as the difference between the 90th and the 10th percentile of expected costs, normalized by the median, (p90-p10)/p50) surrounding it, after controlling for a wide range of observed characteristics.. As a result, the study also informs on how elicitation protocol differences and expert geographical and sector characteristics affected technology outcomes. Independent variables in the central estimates and uncertainty regressions were the level of public RD&D budgets, expert background (industry, academia, and public institution), expert country (American vs. European), technology type (large-scale Gen. III/III+ designs, large-scale Gen. IV designs, and small modular reactor designs), and elicitation mode (in-person vs. online). The relationship between expected costs and RD&D investment was tested both using a log-log

---

[10] The computational challenge comes after the challenge of building internal trust and buy-in, achieving external transparency and consistency, which currently contributes to decision-making entities not estimating the benefits of RD&D investment portfolios.

[11] This approach can potentially be used to conduct other sensitivity analysis such as including experts internal to the decision making process vs. experts from stakeholder groups, experts from different countries, etc (Chan & Anadon, 2013). It also can be used to understand the sensitivity of aggregated results to decisions about whether to include or exclude the outlier expert responses (Jenni, Baker, & Nemet, 2013).

[12] The use of primary data (IPD) is considered the gold standard for systematic reviews because it avoids many of the shortcomings of aggregate meta-analysis: it enables controlling for confounding factors at the individual level and for treatment differences between studies. Moreover, using IPD the study derived results directly and independent of study reporting. This increased the aggregate power of the study, which allowed to more thoroughly scrutinize modeling assumptions (such as the presence of interactions and the linearity of associations) and explore subgroup effects (Borenstein et al., 2009; Ghersi, Berlin, & Askie, 2013; Reade et al., 2009).

It is also important to point out that expert elicitations are used to estimate the distribution of the underlying beliefs held by experts with the largest information sets over an uncertain quantity. Therefore, an expert elicitation study does not rely on asymptotic convergence of sample estimates through the collection of a large number of individual observations, but rather develops the highest quality representation of the underlying distribution among the most informed experts. In this sense the use of IPD meta-analysis that treats individual experts as single observations relies on a random sampling assumption that the original data collection did not make.

specification, usually applied in the learning-by-searching literature, and a linear specification with a squared RD&D term, in line with the literature on diminishing marginal returns to RD&D investments (Evenson & Kislev, 1976; Hall, Mairesse, & Mohnen, 2009; Popp, 2002).

4. Key Findings

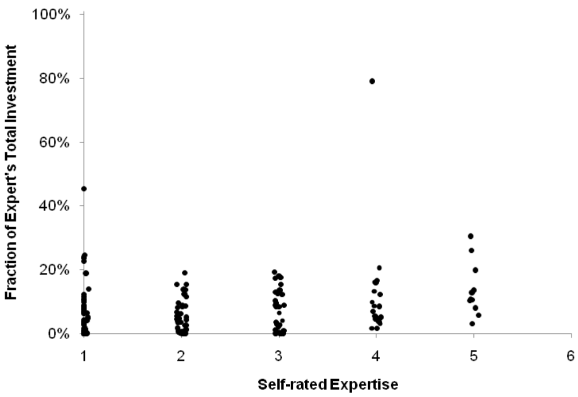
This section describes the key findings regarding the role of public RD&D on the future of energy technologies and the use of elicitations to inform the policy process.

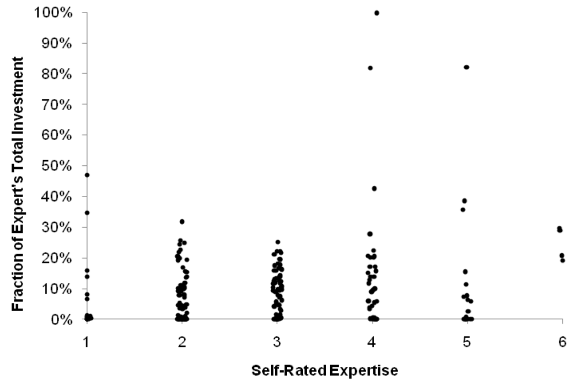
4.1 Including questions about self-rating of expertise

Including a section on self-assessed expertise in the elicitation protocol helped assess whether experts were biased towards favorable treatment for the sub-technology area which they were most knowledgeable. However, we found little evidence of experts systematically recommending greater funding levels for the technology areas with which they were most familiar (Figures 1 and 2 for US and EU experts, respectively).

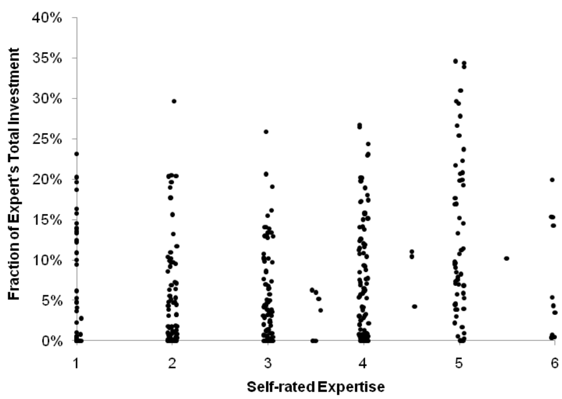
Figure 1 (following page): Analysis of expert–recommended budget allocations in areas of self-assessed expertise in Harvard elicitations. The x-axis corresponds to the self-rated expertise (1: I am not familiar with this technology; 6: I am one of the top experts in this technology). The y-axis corresponds to the fraction of the recommended budget that an expert devoted to a particular technology. The graphs represent 6 different elicitations: (a) Bioenergy; (b) Utility scale energy storage; (c) Nuclear energy; (d) Fossil energy and CCS; (e) Vehicle technologies; (f) Solar photovoltaics.

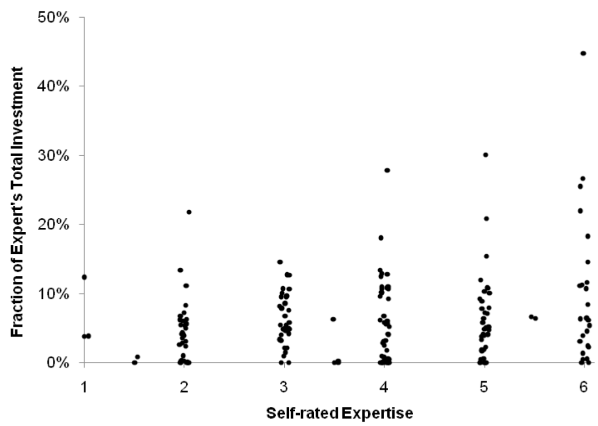
X- axis: Self-rated expertise (1: lowest; 6: highest)
Y- Axis: Fraction of expert's total investment for a particular technology area



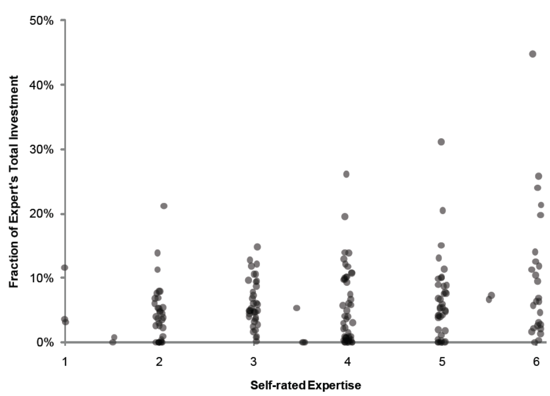


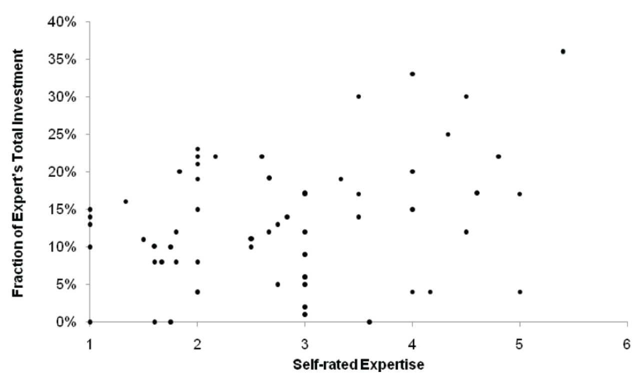
(a) Bioenergy
(b) Utility scale energy storage





(c) Nuclear energy
(d) Fossil energy and CCS



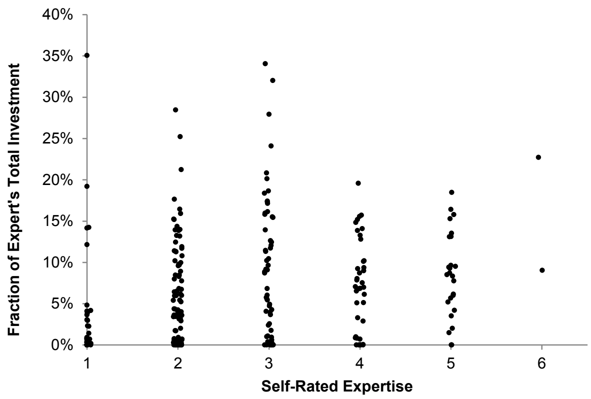


(e) Vehicle technologies

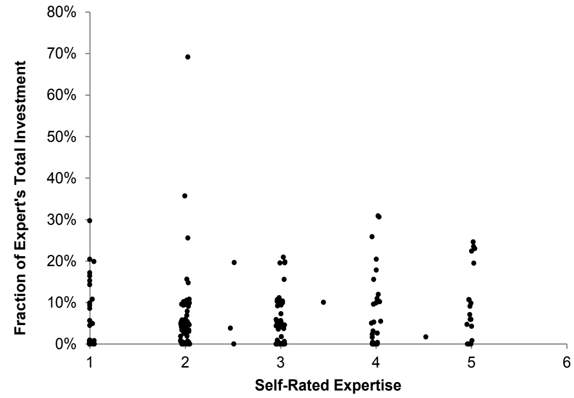
(f) Solar photovoltaics

X- axis: Self-rated expertise (1: lowest; 6: highest for nuclear, 1: lowest; 5: highest for all the other technologies)

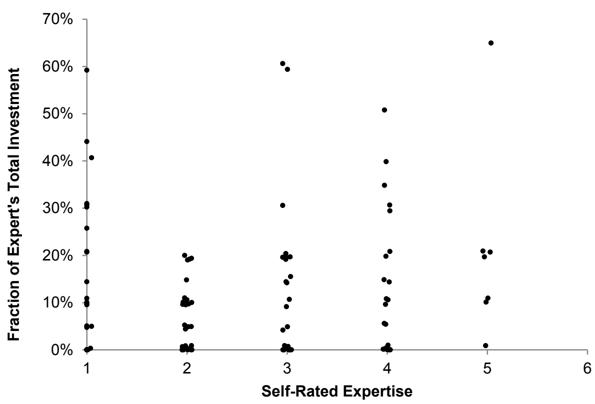
Y- Axis: Fraction of expert's total investment for a particular technology area



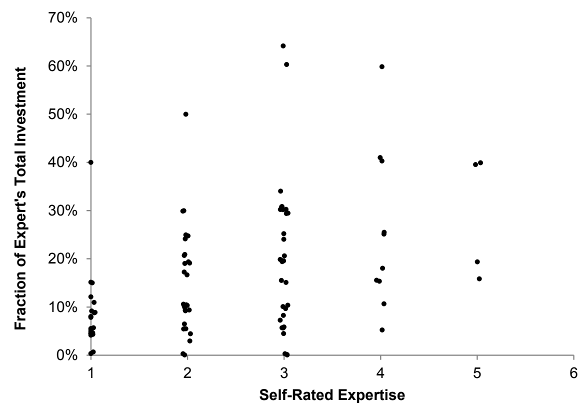
(a) Nuclear energy



(b) Biofuels



(c) Vehicle technologies



(d) Solar photovoltaics

Figure 2: Analysis of expert–recommended budget allocations in areas of self-assessed expertise in FEEM elicitations. The x-axis corresponds to the self-rated expertise (1: I am not familiar with this technology; 6: I am one of the top experts in this technology). The y-axis corresponds to the fraction of the recommended budget that an expert devoted to a particular technology. The graphs represent 4 different elicitations: (a) Nuclear energy; (b) Biofuels; (c) Vehicle technologies; (c) Solar photovoltaics.

4.2 Conducting elicitations online

Online elicitations emerge as the lowest cost option, followed by mail and in person elicitations. An extremely conservative back of the envelope calculation of the monetary benefits (i.e., excluding benefits in future years, assuming that researchers travelling to interview experts do not need accommodation, and ignoring the time and effort savings to researchers and experts) indicates that online surveys with 11 experts are at least 40% cheaper than in-person elicitations with the same number of experts. During the online surveys, some experts did contact the research team for clarification, but it is virtually impossible to rule out that the lower interaction between experts and researchers decreased the value of the information contained in the online estimates, as some experts may have found some of the questions ambiguous (even after extensive pilot testing of the elicitation instruments).

The discussion during the nuclear group workshop, which included 18 out of the 60 experts that participated in the FEEM and Harvard nuclear expert elicitations, confirmed that the online tools providing real-time feedback were useful and that expert interpretation of the questions was consistent with the researchers' intentions. In addition to the qualitative discussion, the robustness of the online elicitation tool was further validated by virtue of very few experts requesting to make changes to their original answers by the end of the workshop.[13]

Differences in the media chosen for the elicitations (online as well as by mail), RD&D scenarios, time periods and technology focus can be used to quantitatively investigate any possible systematic differences in the normalized uncertainty range ((90^{th}-10^{th})/50^{th} percentile cost estimates of the experts). Table 1 shows the results of the analysis of the normalized uncertainty range provided by the experts in the Harvard elicitations using dummy variables for online surveys and for different RD&D and technology scenarios. Table 2 shows similar regression results for the four in-person FEEM elicitations.

Model 1 in Table I shows that the normalized uncertainty range in the Harvard data is greater for online rather than paper sent by mail elicitations. However, we must note that the technology areas are perfectly collinear with the online dummy, which means that further work is needed to disentangle the effect of conducting elicitations online from the differences in normalized uncertainty across technology areas. Model 3 and Model 4 in Table I shows respectively that: (a) controlling for unobserved expert-level heterogeneity with expert fixed effects, RD&D scenarios with greater investment than the BAU RD&D scenario had significantly lower normalized uncertainty ranges; and (b) the bioenergy, storage, solar, and nuclear surveys were associated with significantly greater normalized uncertainty ranges than the fossil survey, with the smallest difference for nuclear; conversely, there was not a significant difference in the uncertainty metric between the fossil and vehicles survey.

Turning to Table II, we find in the 4 in-person FEEM higher RD&D scenarios are associated with greater normalized uncertainty ranges. This result is in contrast with the Harvard results in Table I. There are several possible explanations for this difference, none of which can be formally tested at present. One hypothesis is that U.S. experts believe that more RD&D reduces uncertainty while E.U. experts believe that it increases it. Another hypothesis is that the results depend on the framing of questions by FEEM and Harvard. Specifically, the Harvard surveys asked experts to recommend the total amount and specific allocation of RD&D investments, while the FEEM survey asked experts about fixed increases from the BAU scenario without asking them to design their ideal RD&D program. It is possible that when experts think about their ideal RD&D

[13] The workshop was divided into discussion sessions that were design to match the elicitation questions. Each session included a presentation of the results of that part of the elicitation, a moderated group discussion, and a final session in which each expert was provided with a sheet allowing him to privately make changes to his answers to that section (all nuclear experts were men).

program they have less uncertainty about the results of their recommendations. It is also possible that when experts think about the impact of RD&D on the aggregated cost of the technologies (as is the case in most of FEEM elicitations) they think about uncertainty differently when compared to components of technology cost (as is the case in most of Harvard elicitations).

Table I: Analysis of factors associated with differences in normalized uncertainty ranges in the 6 Harvard expert elicitations. The 2030 BAU RD&D scenario and the fossil technology category serve as reference points. Y = ln(uncertainty).

1. Variable	2. Model 1	3. Model 2	4. Model 3	5. Model 4
6. Online	7. 0.1430**			
	8. (0.0622)			
9. 2010 BAU		10. -0.0431	12. -0.1140**	14. -0.0657
		11. (0.0882)	13. (0.0504)	15. (0.0841)
16. 2030 recommended budget		17. - 0.0945	19. -0.1055**	21. -0.0856
		18. (0.0924)	20. (0.0479)	22. (0.0853)
23. 2030 10X recommended budget		24. -0.0259	26. -0.0948**	28. -0.0168
		25. (0.0863)	27. (0.0473)	29. (0.0800)
30. Vehicles				31. -0.1065
				32. (0.1120)
33. Bioenergy				34. 0.6310***
				35. (0.0964)
36. Storage				37. 0.7006***
				38. (0.1265)
39. Nuclear				40. 0.2574***
				41. (0.0669)
42. Solar PV				43. 0.6894***
				44. (0.0734)
45. Expert fixed effects	46. NO	47. NO	48. YES	49. NO
50. Constant	51. 0.6010***	52. 0..4987***	53. -1.4974***	54. 0.8255***
55. R-squared	56. 0.0077	57. 0.0019	58. 0.7419	59. 0.1465
60. Observations	61. 635	62. 635	63. 635	64. 635

Robust p-values in brackets

65. *** P<0.01, ** p<0.05, * p<0.1
66. Notes: The nuclear and solar PV elicitations were conducted online, and the others via mail.

The difference between the impact of RD&D on the estimates by US and EU experts highlights the complex set of factors involved when making these estimates. They therefore should be carefully considered when using the results from different elicitations on similar topics. Overall, the launch, data acquisition, and data processing for the online surveys were faster than for the paper surveys. Both groups also learned valuable lessons from the development of their first elicitations (bioenergy energy for Harvard and solar survey for FEEM) that made the development of the remaining elicitations faster.

Table II: Analysis of factors associated with differences in normalized uncertainty ranges in the 4 FEEM in person expert elicitations. The 2030 BAU RD&D scenario and the biofuels technology category serve as reference points. Y = ln(uncertainty).

	Model a1	Model a2	Model b1	Model b2	Model c1	Model c2
+50% RD&D	0.167*	0.186***	0.165*	0.186***	0.165*	0.186***
	(0.0847)	(8.50e-06)	(0.0868)	(8.50e-06)	(0.0750)	(4.86e-06)
+100% RD&D	0.305***	0.327***	0.298***	0.327***	0.298***	0.327***
	(0.00231)	(3.00e-09)	(0.00304)	(3.00e-09)	(0.00244)	(2.16e-09)
Solar			-0.110	-0.0267		
			(0.338)	(0.862)		
Vehicle			-0.151	0.380**		
			(0.152)	(0.0490)		
Cost_CSP					-0.476***	-0.0267
					(0.000400)	(0.863)
Cost_EV					-0.104	0.427**
					(0.376)	(0.0313)
Cost_PHEV					-0.197	0.333*
					(0.105)	(0.0919)
Cost_PV					0.0429	0.362**
					(0.725)	(0.0421)
Constant	-0.623***	-0.439***	-0.515***	-0.819***	-0.515***	-0.819***
	(0)	(0.00412)	(1.91e-06)	(4.16e-10)	(1.75e-06)	(5.66e-10)
Observations	161	161	161	161	161	161
R-squared	0.058	0.857	0.071	0.857	0.142	0.867
Expert FE	NO	YES	NO	YES	NO	YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Note: the “b” and “c” versions of the regression models represent different levels of aggregation in the solar and vehicle technologies elicitations.

4.3 Combining elicitations with a group workshop

The Harvard and FEEM groups carried out the same nuclear two-step elicitation in the US and in the EU. First, experts provided individual estimates through and online survey. Then, a subset of experts was involved in a workshop and group discussion (see Figure 3 for a schematic of the process). This identified issues that could arise when each of the two steps is followed as a standalone procedure.

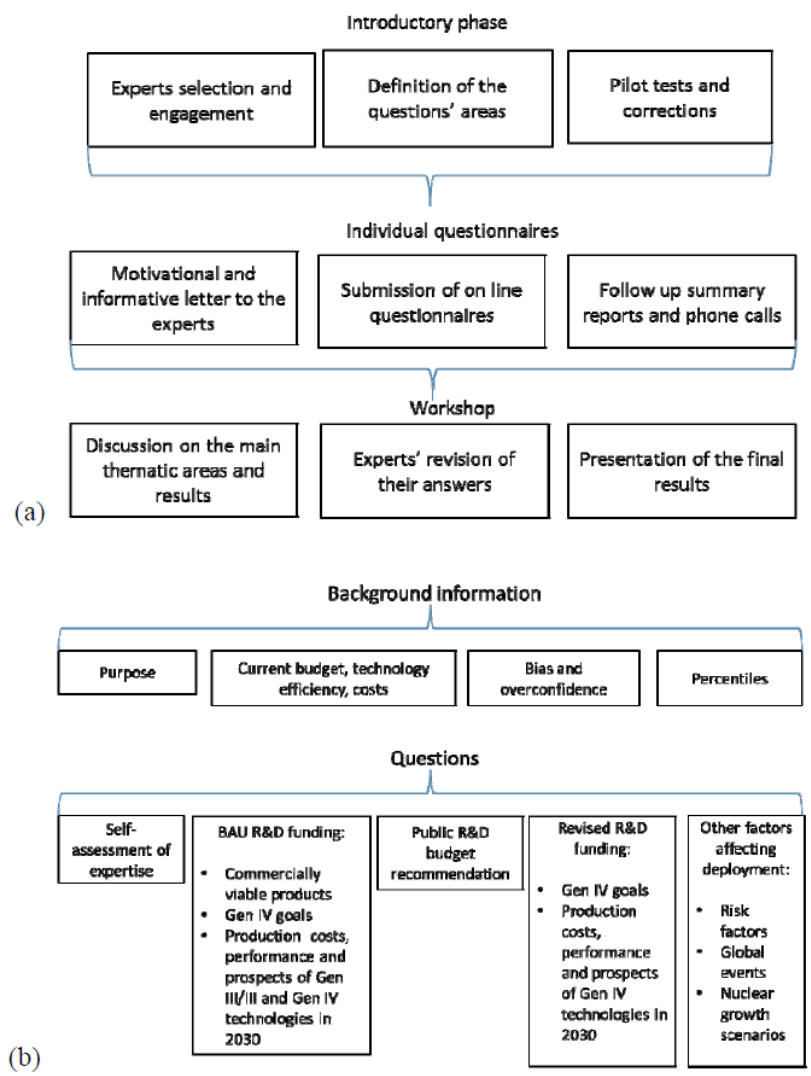


Figure 3: (a) Structure protocol employed in the design of the online elicitation and group discussion; (b) structure of the individual online elicitation instrument (Anadon et al. 2012).

As discussed in section 4.2, while cost and performance estimates did not change substantially during the workshop from the individual expert elicitations, the workshop did enrich the information obtained from the elicitations on other topics. The workshop had some impact on the stated RD&D policy objectives that recommended investments were meant to address. Experts who participated in the workshop made some changes (mainly in the form of additions), suggesting that the workshop discussion was helpful in building consensus in this area. RD&D policy objectives that gained priority after the workshop were development of SMRs, risk and safety, and proliferation resistance. EU experts also increased recommended funding for sodium-cooled fast reactors and fuels and materials.

The workshop also resulted in an improved understanding of how some experts perceived definitional and framing issues that were originally taken for granted. For example, while some experts thought of climate change mitigation as the main goal when making RD&D recommendations, others had multiple goals in mind, such as non-proliferation concerns and hydrogen production.[14] This variation in the experts' reasoning would not have been revealed had we pursued only individual elicitations. The workshop also helped clarify the reasons why U.S. experts placed more emphasis on RD&D to understand fuel cycle economics and reduce fuel cycle costs than E.U. experts and why EU experts thought that it was unlikely that there would be a market for small modular reactors (SMRs) in the future. Due to the (obviously unplanned) timing of the workshop after the Fukushima disaster, we were also able to determine that the Fukushima disaster did not alter the expert's answers regarding the future of nuclear deployment in the United States and the European Union.

Overall, the combination of the individual online elicitation and expert workshops served to validate the online tool and build consensus on parts of the survey, while allowing the research team to better understand some of the reasons behind expert answers. The combination of online tools and other tools to increase expert interaction without incurring additional costs is an area of growing interest (Siddharth, Khodyakov, Srinivasan, Straus, & Adams, 2011).

4.4 Designing expert elicitations to use as modeling inputs

Even though the elicitations were explicitly designed to provide insights about the optimal allocation and total level of RD&D investments across different technology areas, some design needs were not foreseen. Chan & Anadon (2013) identify ways to improve the elicitation to better match analysis needs.[15] First, obtaining experts' estimates of future technology cost over a very large range of RD&D investments, including RD&D ranges well-beyond current levels, can yield additional insights. Experts in the Harvard study were asked to provide estimates of 2030 technology cost and performance under a BAU RD&D funding scenario, their recommended RD&D funding level, and 10 times their recommended funding level. When designing the survey, researchers believed that this was the maximum feasible range that experts would be able to consider; 10-times the recommended levels amounted to 10 to 80 times current funding levels. Outside of this range, the elicitations could not inform the relationship between RD&D funding and technology cost and performance without heroic assumptions to extrapolate beyond the range experts were asked to consider. The funding levels selected in the Harvard work were sufficient to determine that the current RD&D investment level is too low and that, if properly allocated, $15 billion in aggregate US RD&D funding could be justified on the basis of aggregate economic benefits. However, because the calculated benefits of RD&D were so large, this range

[14] Here we include two other examples: (a) While experts displayed a clear understanding of the questions asked about cost and performance, different experts were using a different definition of "major radioactivity releases caused by an accident or sabotage." (b) Some experts thought that the Fukushima accident would fall under their personal definition of "major radioactivity release," others felt that such a description would only apply to a larger accident with more direct casualties.

[15] The work by Chan & Anadon (2013) on estimating and optimizing the benefits of energy RD&D portfolios presented here relates to three other pieces of work. Although Blanford (2009) and Davis and Owens (2003) present two frameworks to support investment decisions, they do not justify their assumptions regarding the impact of RD&D on future technology cost and performance, and they do not provide computational flexibility to allow the estimation of optimal RD&D investment levels in a range of technologies at a sufficiently small level of granularity (in the range of millions of dollars) and with the ability to optimize for different goals and risk considerations. Baker & Solak (2011) use elicitation data for three technologies not targeted to inform government investments at the program level and, unlike this work, the R&D investment optimization relies on assumptions about climate damages.

proved too small to estimate the optimal level of RD&D investment. Even though the Harvard study could calculate the rate of decreasing marginal benefits, benefits (in terms of aggregate economic surplus) were still increasing 10% faster than costs at the maximum aggregate range considered, $15 billion per year (see Figure 4). Other than aggregate economic surplus, there are many other metrics of benefits that one could use (for example, one could use avoided CO_2 emissions for benefits and incorporate opportunity costs for the RD&D costs).

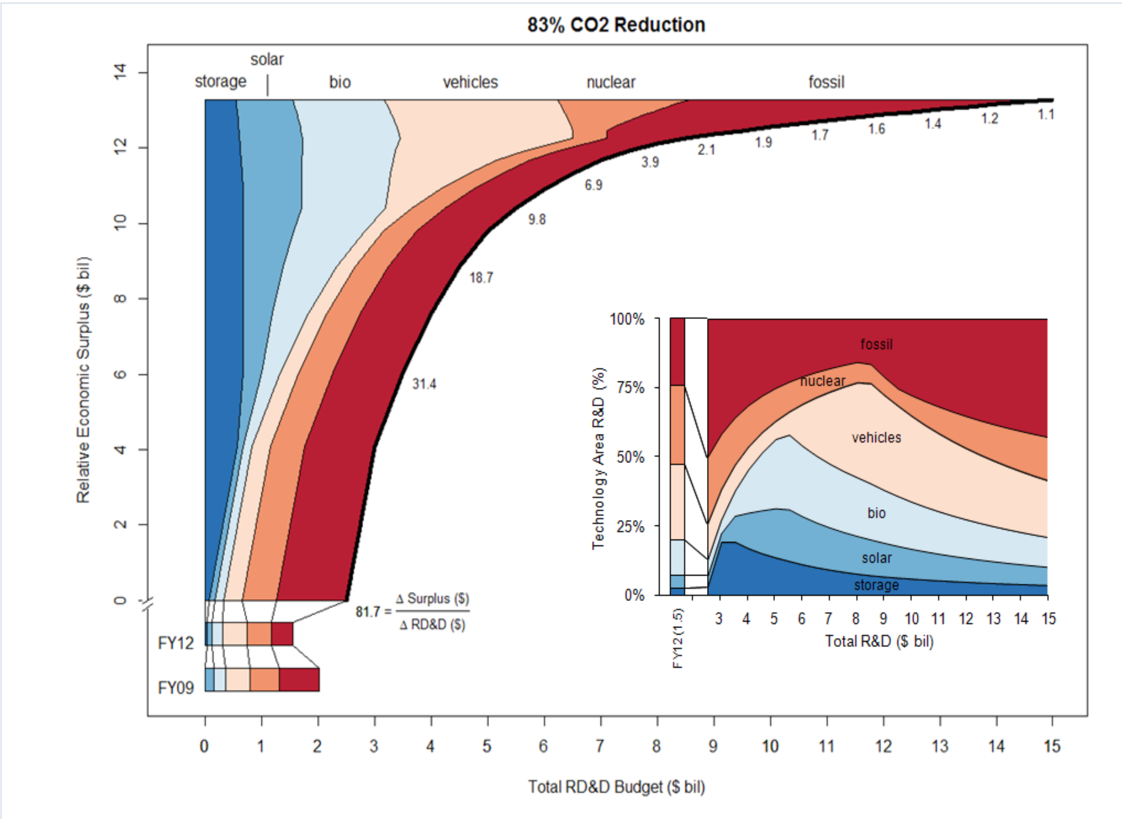


Figure 4: Optimal R&D portfolios under an 83% CO_2 reduction policy. The figure shows the allocation of RD&D funding at different RD&D budget constraints between $2.5 billion - $15 billion per year, relative to the Fiscal Year 2009 and 2012 allocations. The dark blank line in the main plots is the maximum expected increase in economic surplus (above the an arbitrary reference point, the expected surplus in the optimal $2.5bil budget) that can be attained under a given RD&D budget constraint. The small numbers along the black line are estimated marginal returns on investment, calculated by linear approximations to the derivative of the optimal expected surplus at different budgets (Chan & Anadon, 2013).

Second, future elicitations in this area should incorporate questions about the extent to which advances in a particular technology co-develop with advances in other related technologies. The Harvard researchers felt that it was reasonable to assume that future advances in some technologies would be uncorrelated with advanced in other technologies (e.g. solar photovoltaics and nuclear technologies). However, due to knowledge spillovers between technology areas, it seemed unreasonable to make this assumption for all technologies (Nemet, 2012). For example, the Harvard bioenergy technology elicitation consisted of technology processes for three bio-based fuels: gasoline substitutes, diesel substitutes, and jet fuel substitutes. Because of the similarity in the technology to produce any of the three products, assuming independence across the impact of RD&D on the future costs of these technologies did not seem reasonable. Complete independence also did not seem reasonable across other technology areas, such as utility-scale energy storage and electric or plug-in-hybrid vehicles, which could feasibly share battery technology. Thus, a correlation table was developed based on Harvard's expertise in various technology areas (see section SI4 in the SI). To inform future elicitations, the Harvard vehicles elicitation implemented a pilot approach to utilize expert knowledge to estimate cross-technology correlations. The pilot asked experts to revise their 90th, 10th, and 50th percentile estimates for a technology considering several future scenarios with different realizations of 2030 costs in a related technology. While most experts were willing and able to think through and answer these questions thoughtfully, including these questions lengthened an already long elicitation. Third, asking qualitative questions to justify experts' recommended level of investments and allocation increased our own confidence in the results and their external credibility.[16]

Fourth, the large number of experts included in the elicitations (more than 100 per research group), required substantial preprocessing before summary results could be presented. Harvard developed an importance sampling technique to reduce the computational requirements of assessing the RD&D allocations and forecasts of many different experts. However, for the parsimony of presenting results, experts' responses were eventually selected or aggregated. Anadon et al. (2011), for example, relied on three "expert scenarios", labeled, "optimistic," "middle," and "pessimistic", each of which grouped the answers of the 6 most optimistic, central, and pessimistic experts. As shown in Figure 5, even increasing RD&D investments from a BAU budget of \$2 billion to \$82 billion/year, and utilizing assumptions from the most optimistic experts, CO_2 emissions are not expected to decrease substantially from current levels. Thus, creating "expert scenarios" allowed researchers to calculate high and low bounds of benefit metrics that did not depend on the choice of expert.

[16] This is something that the Harvard group did not include in the first bioenergy elicitation, but did include in the subsequent five elicitations. For more information on what some of these qualitative questions focused on, the reader can access the links to the nuclear survey in the SI of Anadon et al. (2012).

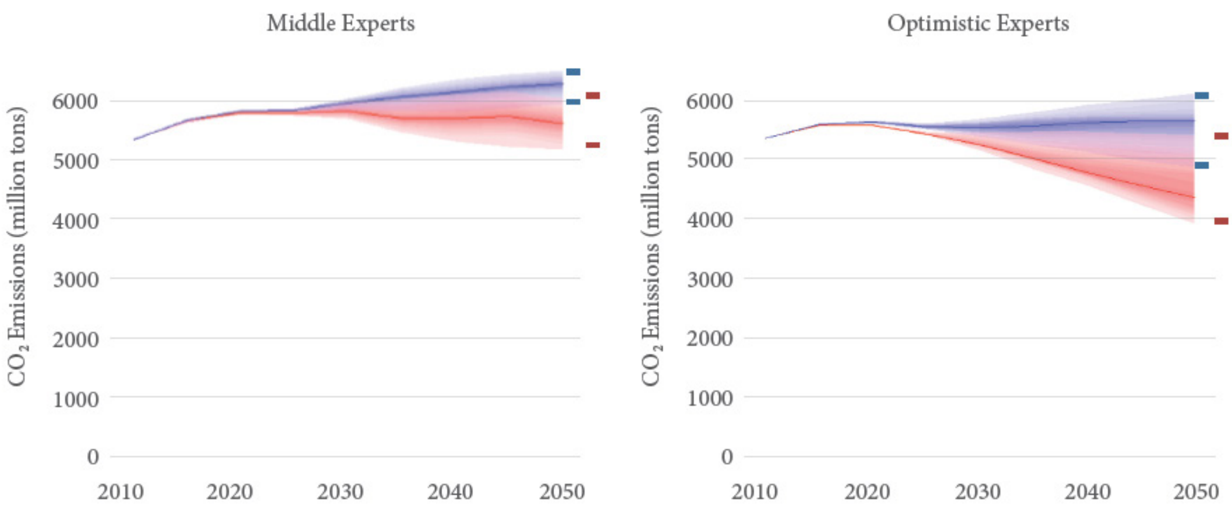


Figure 5: U.S. energy-related CO_2 emissions under (a) business-as-usual federal energy RD&D investment and no additional demand-side policies (blue) and (b) ten times the experts' average recommended federal energy RD&D investments (somewhere between $49 and $82 billion/year) (red), with no additional demand-side policies, using "middle of the road" and "optimistic" experts' technology cost projections. Note that optimistic experts were optimistic about technological progress in general, and not necessarily optimistic about the effects of RD&D (Anadon et al., 2011).

4.5 Using meta-analysis to improve elicitation usability and design

The meta-analysis of the nuclear elicitations evaluated the impact of expert selection (background and country) and elicitation design (technology granularity and online vs. in person mode) on the elicited costs of nuclear technologies. The goal of this exercise was to inform future elicitations and to better capture the experts' thinking on the impact of public nuclear RD&D on future technology costs for modeling and policy analysis.

Here, we discuss the key insights from the log-log model of the experts' central estimate of nuclear power overnight capital cost in 2030 (see section 3.2).[17] Expert composition has a qualitatively large impact on the range of estimates available for policy analysis. Controlling for expert affiliation, expert country of origin, and technology type, the coefficient of the RD&D variable increases by 25% relative to the estimated coefficient in the reduced form model (namely, RD&D on costs). On average, a doubling of the yearly public nuclear RD&D budget in the US and the EU is associated with an 8% decrease in nuclear costs in 2030, *ceteribus paribus.* Experts from public institutions have estimates of overnight capital costs that are about 14% higher on average than those of academics and that estimates from industry experts are even higher, on average around 31% higher than academics. Expected overnight capital costs are approximately 22%

[17] The results of the two non-linear models we specified, log-log and linear-quadratic, were consistent in terms of the statistical significance and sign of the estimated effects. Also, the estimated negative quadratic coefficients in the linear model are consistent with diminished returns to RD&D (and are not necessarily inconsistent with learning-by-searching).

lower for experts in the USA when compared to experts in the European Union. Technology type also is a statistically significant determinant of 2030 expected costs: overnight capital costs are expected to be higher for both Gen. IV and SMR technologies with respect to Gen. III/III+ technologies by roughly 23% and 24%, respectively.[18]

Focusing on uncertainty—defined as the 90th percentile estimate less the 10th percentile estimate, normalized by the 50th percentile estimate—higher or lower levels of RD&D investment are not systematically associated with narrower or wider uncertainty ranges. However, US experts have around 16% wider uncertainty ranges compared to EU experts. The uncertainty range for SMRs is about 14% smaller than that for large scale Gen. III/III+, suggesting that experts are more confident about their cost estimates for these systems. This was a somewhat surprising finding considering that SMRs are expected to be delivered to the site fully constructed from the manufacturing facilities, yet current experience is limited and no operating licenses have been issued in the United States or the EU.

Ongoing work is now focusing on validating these results across a wide range of technologies through a larger meta-analysis The increased variation among these studies, as well as the increase observations, will enable more precise estimation of both expert and elicitation design effects and will allow to gauge differences in experts' assumptions about the returns of RD&D in different technological areas.

5. Conclusions and future work

The findings presented in this paper provide lessons for the future design and use of expert elicitations to inform policy decisions on public RD&D investments. The findings presented in this paper stem from several pieces of work related to 10 expert elicitation exercises encompassing 6 energy technology areas and conducted between 2009 and 2011 by Harvard researchers and FEEM researchers. Below we summarize five key findings.

First, mail and online expert elicitation tools can be used to obtain expert elicitation estimates more cost-effectively than in-person interviews without introducing bias. This finding relies on insights from the expert workshop that followed FEEM and Harvard nuclear elicitations and is conditional on appropriate preparatory work by the eliciting research team. This work includes extensive background research on the topic, pilot testing the elicitation instrument, background material that discussed biases and confidence, and the utilization of numerous interactive visual aids. In particular, conducting elicitations online can contribute to an easier institutionalization of the process.

Second, asking experts to self-assess their level of expertise in specific technologies and processes, to justify their RD&D priorities, and to identify non-RD&D-related factors that would

[18] The Anadon, Nemet & Verdolini (2013) study found that the in-person variable (accounting for the observations obtained through an in-person interview as opposed to through an online tool) becomes negative and significant when expert fixed effects are included, although it is difficult to draw conclusions about this effect since it requires inclusion of unobserved expert characteristics for it to become significant. This tentative result is consistent with results in Table 1 in this paper, but the tentative nature of this analysis requires that the inn-person effects be a focus of future work assembling additional elicitation data ensuring that more than the 3% of observations are in-person.

affect the future of specific technologies, increases both the researchers' confidence in the level of intellectual engagement of the experts and in the external credibility of the results. For example, experts were not systematically recommending larger amounts of funding to their areas of expertise, providing some evidence that they were not solely motivated by self-interest.

Third, to support decisions about RD&D investments in different technology programs, it can be useful to push experts to consider a wide range of scenarios, including scenarios at the boundary of their private information set, to explore potentially-desirable scenarios far from current activities without undue extrapolation bias. In addition, elicitations should include questions to allow the deduction of correlations across technology improvements. Alternatively, researchers (or analysts) can create a separate elicitation targeting correlations.

Fourth, some important policy insights can be derived by creating scenarios without aggregating experts. Insights regarding the need to put in place additional policies beyond RD&D investments to meet CO2 emissions reductions goals, and the decreasing marginal returns to RD&D investments, were independent of whether or not modeling included experts that were optimistic, central, or pessimistic regarding forecasted 2030 technology costs.

And fifth, expert selection has a large and significant impact on elicitation results, indicating that experts from the private sector, academia, and public institutions, as well as experts from different countries, have different private information sets and beliefs. An elicitation exercise that sought to include all perspectives would need to include experts from all of these backgrounds. Further, the meta-analysis exercise allowed researchers to better understand estimates of the impact of RD&D on technology costs.

The lessons from this work are applicable not only to energy, but also to other technology areas that receive substantial government RD&D support, such as health, defense, and agriculture. Public RD&D investments in other sectors also face questions regarding the extent to which they should be guided purely by scientific merit or by mission. For example, there have been calls to increase the extent to which funding in the R&D budget of the National Institutes of Health (NIH) should consider disease burdens (see review by Sampat (2012)). This approach would require not only that greater fractions of the NIH budget be allocated to specific diseases, but also some consideration of the extent to which additional research could result in improvements. Industrial research institutions could also implement some of the insights and methods discussed in this paper, as they also deal with investing in projects with uncertain returns that will only impact their bottom line if they are diffused in the market.

Although the combination of insights from this body of work improves our confidence in the use of expert elicitations to inform RD&D decisions in the energy sector and (we would argue) beyond, there are several avenues for ongoing and future research that will further improve our understanding. Experts could be randomized into three different groups to complete the same elicitation in-person, online, or via mail, respectively to conduct a more systematic evaluation of whether there are any systematic differences in the results. Additional meta-analysis work including elicitations for energy technologies beyond nuclear energy would establish the extent to which expert background, country variables and returns to RD&D change across technologies. Ongoing work involving three major teams involved with energy economic models (GCAM at the Pacific Northwest National Laboratory, WITCH at FEEM, and MARKAL at Brookhaven

National Laboratory) is using aggregates of elicitation results from different studies. This effort will develop probability distributions of technology costs conditional on R&D levels by applying equal weights in a mixture distribution of individual expert assessments collected from major studies conducted at the University of Massachusetts Amherst, FEEM, and Harvard University. Finally, the question of whether or not to aggregate expert answers to model future technical change and the uncertainty around it was not a focus of this paper (the focus was on insights robust to different "expert scenarios"). Identifying the benefits of aggregating expert assessments may ultimately require ex-post analysis of previous elicitations against the realized technical change.

6. References

Abdulla, A., Azevedo, I. L., & Morgan, M. G. (2013). Expert elicitation of the cost of small modular reactors. *Proceedings of the National Academy of Sciences, 110*(24), 9686–9691.

American Energy Innovation Council. (2010). *A business plan for America's energy future. full report.* (June). Washington D.C., United States: American Energy Innovation Council.

Anadon, L. D., Bunn, M., Chan, G., Chan, M., Jones, C., Kempener, R., Logar, N., Narayanamurti, V. (2011). *Transforming U.S. energy innovation.* Cambridge, MA, United States: Energy Technology Innovation Policy Group, Belfer Center for Science and International Affairs, John F. Kennedy School of Government, Harvard University.

Anadon, L. D., Bosetti, V., Bunn, M., Catenacci, M., & Lee, A. (2012). Expert judgments about RD&D and the future of nuclear energy. *Environmental Science & Technology, 46*, 11497-11504.

Anadon, L. D., Nemet, G. F., & Verdolini, E. (2013). The future costs of nuclear power using multiple expert elicitations: Effects of RD&D and elicitation design. *Environmental Research Letters, 8,* 034020.

Anadon, L. D. (2012). Missions-oriented RD&D institutions in energy between 2000 and 2010: A comparative analysis of China, the United Kingdom, and the United States. *Research Policy, 41*(10), 1742–1756.

Arrow, K. J. (1962). Economic welfare and the allocation of resources for invention. In R. R. Nelson (Ed.), *The rate and direction of inventive activity* (pp. 609-626). Princeton, NJ, USA: Princeton University Press.

Baker, E., & Solak, S. (2011). Climate change and optimal energy technology R&D. *European Journal of Operations Research, 213*, 442-454.

Barker, T., & Jenkins, K. (2007). *The costs of avoiding dangerous climate change: Estimates derived from a meta-analysis of the literature.* (). New York, NY, United States: United Nations Human Development Report.

Blanford, G. J. (2009). R&D investment strategy for climate change. *Energy Economics, 31*(1), S27-S36.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, U.K.: John Wiley & Sons, Ltd.

Bosetti, V., Catenacci, M., Fiorese, G., & Verdolini, E. (2012). The future prospect of PV and CSP solar technologies: An expert elicitation survey. *Energy Policy, 49*, 308-317.

Catenacci M., Verdolini E., Bosetti B., & Fiorese G. (2013). Going electric: Expert survey on the future of battery technologies for electric vehicles. *Energy Policy, 61,* 403-413.

Chan, G., Anadon, L. D., Chan, M., & Lee, A. (2011). Expert elicitation of cost, performance, and RD&D budgets for coal power with CCS. *Energy Procedia, 4*, 2685-2692.

Chan, G., & Anadon, L. D. (2013). Utilizing expert assessments to inform allocating government energy RD&D investment portfolios. *Submitted.*

Cooke, R. L. (1991). *Experts in uncertainty: Opinion and subjective probability in science*. New York City, NY, USA: Oxford University Press.

Davis, G. A., & Owens, B. (2003). Optimizing the level of renewable electric R&D using real options analysis. *Energy Policy, 31*, 1589-1608.

EERA. (2010). *8th Framework Programme (FP8) position paper of the European Energy Research Alliance (EERA). October*. Brussels, Belgium: European Commission.

EIA. (2011). *Direct federal financial interventions and subsidies in energy in fiscal year 2010.* . Washington D.C.: U.S. Energy Information Administration, U.S. Department of Energy.

European Commission. (2007). *Communication from the commission to the council, the European parliament, the European Economic and Social Committee, and the Committee of the Regions "limiting global climate change to 2 degrees Celsius the way ahead for 2020 and beyond"* . Brussels, Belgium: European Commission, COM(2007) 2 final.

Evans, J. (2013). Introduction to Structured Expert Elicitation: A Risk Analysis Perspective. *Methods for Research Synthesis: A Cross-Disciplinary Approach.* Workshop. Harvard Center for Risk Analysis. October.

Evenson, R. E., & Kislev, Y. (1976). A stochastic model of applied research. *The Journal of Political Economy, 84*(2), 265-281.

Fann, N., Gilmore, E., Walker, K. (2013). Characterizing the long-term $PM_{2.5}$ concentration response function: Comparing the strengths and weaknesses of research synthesis approaches. *Methods for Research Synthesis: A Cross-Disciplinary Approach.* Workshop. Harvard Center for Risk Analysis. October.

Fiorese, G., Catenacci, M., Verdolini, E., & Bosetti, V. (2013). Advanced biofuels: Future perspectives from an expert elicitation survey. *Energy Policy, 56*, 293–311.

Fishbone, L. G., & Abilock, H. (1981). MARKAL, a linear-programming model for energy systems analysis: Technical description of the BNL version. *International Journal of Energy Research, 5*(4), 353-375.

Gallagher, K. S., Anadon, L. D., Kempener, R., & Wilson, C. (2011). Trends in investments in global energy research, development, and demonstration. *Wiley Interdisciplinary Reviews: Climate Change, 2*(3), 373-396.

Ghersi, D., Berlin, J. A., & Askie, L. (2013). *Cochrane Collaboration—Prospective meta-analysis methods group.*. http://pma.cochrane.org: The Cochrane Collaboration.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3-8.

Hall, B. H., Mairesse, J., & Mohnen, P. (2009). Measuring the returns to R&D. *National Bureau of Economic Research Working Paper Series, No. 15622*

Hogarth, R. M. (1987). *Judgment and choice: The psychology of decision*. Chichester, West Sussex, United Kingdom; New York City, NY, USA: Wyley.

IEA. (2013). *International Energy Agency Energy R&D Statistics.* Available at: http://www.iea.org/stats/rd.asp). Paris, France: International Energy Agency.

IEA, OPEC, OECD, & World Bank. (2010). *Analysis of the scope of energy subsidies and Suggestions for th eG-20 Initiative. Joint Report*. Toronto, Canada: International Energy Agency, Organization fo the Petroleum Exporting Countries, Organization for Economic Cooperation and Development, and The World Bank.

Jenni, K. E., Baker, E. D., & Nemet, G. F. (2013). Expert elicitations of energy penalties for carbon capture technologies. *International Journal of Greenhouse Gas Control, 12*, 136-145.

Keeney, R. L., & Winterfeldt, D. (1991). Eliciting probabilities from experts in complex technical problems. *Transactions on Engineering Management, 38*, 191-201.

Matarazzo, B., & Nijkamp, P. (1997). Meta-analysis for comparative environmental case studies: Methodological issues. *International Journal of Social Economics, 24*(7/8/9), 799-811.

Meyer, M. A., & Booker, J. M. (1991). *Eliciting and Analysing Expert Judgment: A Practical Guide*. London, U.K.: Academic Press Ltd.

Morgan, M. G., & Henrion, M. (1990). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge: Cambridge University Press.

Morton, S. (2013). Introduction to Systematic Review and Meta-Analysis: A Health Care Perspective. M*ethods for Research Synthesis: A Cross-Disciplinary Approach.* Workshop. Harvard Center for Risk Analysis. October.

NCEP. (2004). *Ending the energy stalemate. A bipartisan strategy to meet America's energy challenges.* Washington D.C., United States: National Commission on Energy Policy.

NCEP. (2007). *Energy policy recommendations to the President and the 110th Congress.* The National Commission on Energy Policy.

Nelson, J., & Kennedy, P. (2009). The use (and abuse) of meta-analysis in environmental and natural resource economics: An assessment. *Environmental & Resource Economics, 42*(3), 345-377.

Nemet, G. F. (2012). Inter-technology knowledge spillovers for energy technologies. *Energy Economics, 34*(5), 1259-1270.

Nemet, G. F. (2013). Technological change and climate-change policy. In J. Shogren (Ed.), *Encyclopedia of energy, natural resource and environmental economics* (Amsterdam, The Netherlands ed., pp. 107-116) Elsevier.

NRC. (2007). *Prospective evaluation of applied energy research and development at DOE (phase two)*. National Academies Press, Washington D.C., United States: National Research Council.

PCAST. (1997). *Federal energy research and development for the challenges of the twenty-first century.* (No. Chapter 4). Washington D.C., United States: President's Council of Advisors on Science and Technology, Executive Office of the President.

PCAST. (2010). *Report to the President on accelerating the pace of change in energy technologies through an integrated federal energy policy.* President's Council of Advisors on Science and Technology. Executive Office of the President.

Popp, D. (2002). Induced innovation and energy prices. *American Economic Review, 92*, 160-180.

Pugh, G., Clarke, L., Marlay, R., Kyle, P., Wise, M., McJeon, H., & Chan, G. (2011). Energy R&D portfolio analysis based on climate mitigation. *Energy Economics, 33*, 634-643.

Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty*. Oxford, UK: Addison-Wesley.

Reade, M. C., Delaney, A., Bailey, M. J., Harrison, D. A., Yealy, D. M., Jones, P. J., . . . Angus, D. C. (2009). Prospective meta-analysis using individual patient data in intensive care medicine. *Intensive Care Medicine, 36*(1), 1-21.

Rose, A., & Dormady, N. (2011). A meta-analysis of the economic impacts of climate change policy in the United States. *The Energy Journal, 32*(2), 143-166.

Sampat, B. N. (2012). Mission-oriented biomedical research at the NIH. *Research Policy, 41*, 1729-1741.

Siddharth, D., Khodyakov, D., Srinivasan, R., Straus, S., & Adams, J. (2011). ExpertLens: A system for eliciting opinions from a large pool of non-collocated experts with diverse knowledge. *Technological Forecasting & Social Change, 78*(8), 1426–1444.

Zamparini, L., & Reggiani, A. (2007). Meta-analysis and the value of travel time savings: A transatlantic perspective in passenger transport. *Networks & Spatial Economics, 7*(4), 377-396.

Belfer Center for Science and International Affairs
Harvard Kennedy School
79 JFK Street
Cambridge, MA 02138
Fax: (617) 495-8963
Email: belfer_center@harvard.edu
Website: http://belfercenter.org