FONDAZIONE ENI ENRICO MATTEI

NOTA DI LAVORO

61.2016

Quantifying the Effects of Expert Selection and Elicitation Design on Experts' Confidence in their Judgments about Future Energy Technologies

Gregory F. Nemet, University of Wisconsin-Madison

Laura Diaz Anadon, John F. Kennedy School of Government, Harvard University

Elena Verdolini, Fondazione Eni Enrico Mattei and Centro Euro-Mediterraneo per i Cambiamenti Climatici

Mitigation, Innovation and Transformation Pathways
Series Editor: Massimo Tavoni

Quantifying the Effects of Expert Selection and Elicitation Design on Experts' Confidence in their Judgments about Future Energy Technologies

By Gregory F. Nemet, University of Wisconsin-Madison
Laura Diaz Anadon, John F. Kennedy School of Government, Harvard University
Elena Verdolini, Fondazione Eni Enrico Mattei and Centro Euro-Mediterraneo per i Cambiamenti Climatici

Summary

Expert elicitations are frequently used to characterize future technology outcomes. However their usefulness is limited, in part because: estimates across studies are not easily comparable; choices in survey design and expert selection may bias results; and over-confidence is a persistent problem. We provide quantitative evidence of how these choices affect experts' estimates of the costs of future energy technologies. We harmonize data from 19 elicitations, involving 215 experts, on the 2030 costs of 5 energy technologies: nuclear, biofuels, bioelectricity, solar, and carbon capture. We control for expert characteristics, survey design, and public R&D investment levels on which the elicited values are conditional. We find that, on average, when experts respond to elicitations in person, they ascribe lower confidence (larger uncertainty) to their estimates than when responding via mail or online. In-person interviews also produce more optimistic assessments of best-case (10th percentile) outcomes. The impacts of expert affiliation—government, private sector, or academic—and geography—US or EU—are also significant; academics and US experts have lower confidence than other types of experts. Higher R&D investment levels have no effect on the confidence of experts' judgments. R&D reduces both the median and breakthrough (10th percentile) cost estimates, although the size of the effect varies across technologies. These results indicate the source, direction, and size of bias in energy technology elicitations. They also point to the technology specificity of some of the effects. These biases should be seriously considered, both in interpreting the results of existing elicitations and in designing new ones.

Keywords: Expert Elicitations, Uncertainty, Energy Technologies, Heuristic Biases, Survey Design

JEL Classification: O13, O14, Q4

This is the pre-peer reviewed version of the following article: Nemet, G., Anadon L. D. and Verdolini E. (2016) "Quantifying the effects of expert selection and elicitation design on experts' confidence in their judgments about future energy technologies", Risk Analysis, which has been published in final form at doi:10.1111/risa.12604. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Address for correspondence:
Gregory F. Nemet
University of Wisconsin–Madison
1225 Observatory Drive
Madison, WI 53706
USA
E-mail: nemet@wisc.edu

The opinions expressed in this paper do not necessarily reflect the position of
Fondazione Eni Enrico Mattei
Corso Magenta, 63, 20123 Milano (I), web site: www.feem.it, e-mail: working.papers@feem.it

This is the pre-peer reviewed version of the following article: **Nemet, G., Anadon L. D. and Verdolini E. (2016) "Quantifying the effects of expert selection and elicitation design on experts' confidence in their judgments about future energy technologies", Risk Analysis,** *which has been published in final form at* **doi:10.1111/risa.12604.** *This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.*

Quantifying the effects of expert selection and elicitation design on experts' confidence in their judgments about future energy technologies

Gregory F. Nemet[1,2,*], *Laura Diaz Anadon*[3], *Elena Verdolini*[4]

[1] La Follette School of Public Affairs, University of Wisconsin-Madison, WI 53706, USA

[2] Nelson Institute Center for Sustainability and the Global Environment (SAGE), University of Wisconsin-Madison, WI 53706, USA

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

[4] Fondazione Eni Enrico Mattei and Centro Euro-Mediterraneo sui Cambiamenti Climatici, Corso Magenta 63,20123 Milano, Italy

* Corresponding author: nemet@wisc.edu.

ABSTRACT:

Expert elicitations are frequently used to characterize future technology outcomes. However their usefulness is limited, in part because: estimates across studies are not easily comparable; choices in survey design and expert selection may bias results; and over-confidence is a persistent problem. We provide quantitative evidence of how these choices affect experts' estimates of the costs of future energy technologies. We harmonize data from 19 elicitations, involving 215 experts, on the 2030 costs of 5 energy technologies: nuclear, biofuels, bioelectricity, solar, and carbon capture. We control for expert characteristics, survey design, and public R&D investment levels on which the elicited values are conditional. We find that, on average, when experts respond to elicitations in person, they ascribe lower confidence (larger uncertainty) to their estimates than when responding via mail or online. In-person interviews also produce more optimistic assessments of best-case (10th percentile) outcomes. The impacts of expert affiliation—government, private sector, or academic—and geography—US or EU—are also significant; academics and US experts have lower confidence than other types of experts. Higher R&D investment levels have no effect on the confidence of experts' judgments. R&D reduces both the median and breakthrough (10th percentile) cost estimates, although the size of the effect varies across technologies. These results indicate the source, direction, and size of bias in energy technology elicitations. They also point to the technology specificity of some of the effects. These biases should be seriously considered, both in interpreting the results of existing elicitations and in designing new ones.

Key Words: expert elicitations, uncertainty, energy technologies, heuristic biases, survey design

1 INTRODUCTION

1.1 Technology elicitations in science and energy policy

Policy makers addressing science and innovation issues frequently confront the challenge of making decisions that affect the development of technologies when both technology outcomes and consequent social impacts are difficult to predict. Expert elicitations allow analysts to gather information from experienced professionals about the future of specific technologies that may not be available from other data sources. Protocols for data collection are designed to reduce biases and encourage considered judgments[(1-3)]; they generate a collection of experts' best estimates of future costs of a particular technology, which can be conditional on different levels of public R&D investment. Importantly, expert elicitations also provide measures of uncertainty associated with the central estimates.

Analyzing these judgments provides a rich resource with which to inform policy decisions. Indeed, expert elicitations are increasingly used in policy making, starting in the 1970s with the U.S. Environmental Protection Agency (EPA), and at least five other federal agencies and international organizations[(4)]. Policy analysis of energy system decisions typically requires explicit[(5)] or implicit[(6)] characterizations of the anticipated cost and performance of specific technologies. However, energy technologies change over time, often in ways that diverge from historical trends[(7)]. While there is significant uncertainty about future technological change in energy, there is now a large set of results from studies eliciting judgments about future technology performance and cost from technical experts.

Over the past eight years several research groups have conducted expert elicitations on the impact of public R&D on the future of important energy technologies. The perceived central role of governments in funding new energy technologies combined with the potential of expert elicitations to characterize uncertainty have driven this burst of more than twenty studies. In the U.S, for instance, the National Research Council published a strong recommendation that the U.S. Department of Energy (DOE) use expert elicitations to inform for their R&D allocation decisions to probabilistically characterize the expected outcomes of R&D investments[8].

1.2 Challenges in using elicitations for policy design

Despite the increasing use of expert elicitations in science policy contexts, the analyst or policy maker has few tools with which to interpret the various studies. This issue of how to utilize and learn from the existing expert elicitations for future elicitations extends beyond the debate that recently took place in this journal regarding how, or whether, to derive consensus from them [9-13]. Difficulty in interpretation is exacerbated by differences in protocol design (metrics, assumptions, timeframes, methods for administering the surveys) and in the backgrounds (institutional affiliation and nationality) of experts selected. Currently, we know little about whether such differences affect the elicitation results themselves, both in terms of the distribution and central estimates.

Morgan provided a comprehensive review of how to think about selecting experts, in part to reduce bias[14], but the literature has not yet provided a framework with which to compare elicitations across different technologies, nor has it produced a comprehensive set of empirical estimates of the impact and size and direction of these differences[15-21]. Two recent articles provide first steps in this direction in assessing the roles of expert and survey characteristics in elicitations of nuclear and solar costs, with a focus on central estimates [22, 23]. In this paper, we expand the analysis to five key energy technologies--nuclear fission, biofuels, bioelectricity, solar, and carbon capture and storage in coal power plans—and focus on the experts' uncertainty range and breakthrough estimates. The results we present in this paper thus provide a much clearer sense of how much variation exists in anticipated outcomes both within and across energy technologies. They also serve as a basis for improving policy making in government agencies—such as the U.S. Dept. of Energy, multiple Congressional committees and the European Commission —in facilitating their interpretation of existing energy technology elicitations. They can certainly help to improve the design of future elicitations. Decisions related to science and innovation policy, especially the consideration of technology portfolios, need characterization of the reliability of the future cost estimates and of the uncertainty surrounding them.

2 APPROACH

We collect and harmonize data from 19 expert elicitation studies that include judgments on future technology outcomes in 5 energy technology areas (nuclear fission, biofuels, and bioelectricity, solar, and carbon capture and storage (CCS) in coal power plans). These five energy technologies are widely believed to be essential for achieving cost-effective climate change mitigation and central to many countries' energy planning[(24, 25)]. The diversity of data from these elicitations provides a unique opportunity to study whether selection and survey design affect (i.e. "bias") experts' estimates and confidence. We do so by using a meta-analytic approach to extract robust conclusions about how differences in the technologies considered, as well as in expert and survey characteristics, impact expected energy costs ($/energy) and the uncertainty surrounding them. This approach is similar to previous work using regression techniques and a randomized control trial to estimate the influence of elicitation question format on expert overconfidence in the domains of infectious disease and marine ecology[(21)]. In this paper, we focus on five key energy technologies and we enlarge the investigation to other study characteristics than question format as well as to experts' background.

Table S1 in the online Supporting Information (SI) provides details on the studies included in the analysis (4 elicitations on nuclear, 5 on solar, 3 on bioelectricity, 3 on biofuels, and 4 on CCS) and on the data handling process[1]. Data on 2030 costs were obtained from the original authors, underwent a rigorous cleaning process, and were converted into units of cost per energy (*2010$/kWh*): levelized costs of

---

[1] Three studies were ultimately dropped from the analyses due to lack of specification of policy conditions.

electricity (solar), non-fuel levelized electricity costs (in the case of bioelectricity), non-fuel levelized energy costs (in the case of biofuels), levelized capital costs for nuclear, and levelized additional capital costs for CCS. Data on technology sub-type, study characteristics, and R&D scenarios on which the estimates are conditional, were also made consistent as documented in the SI.

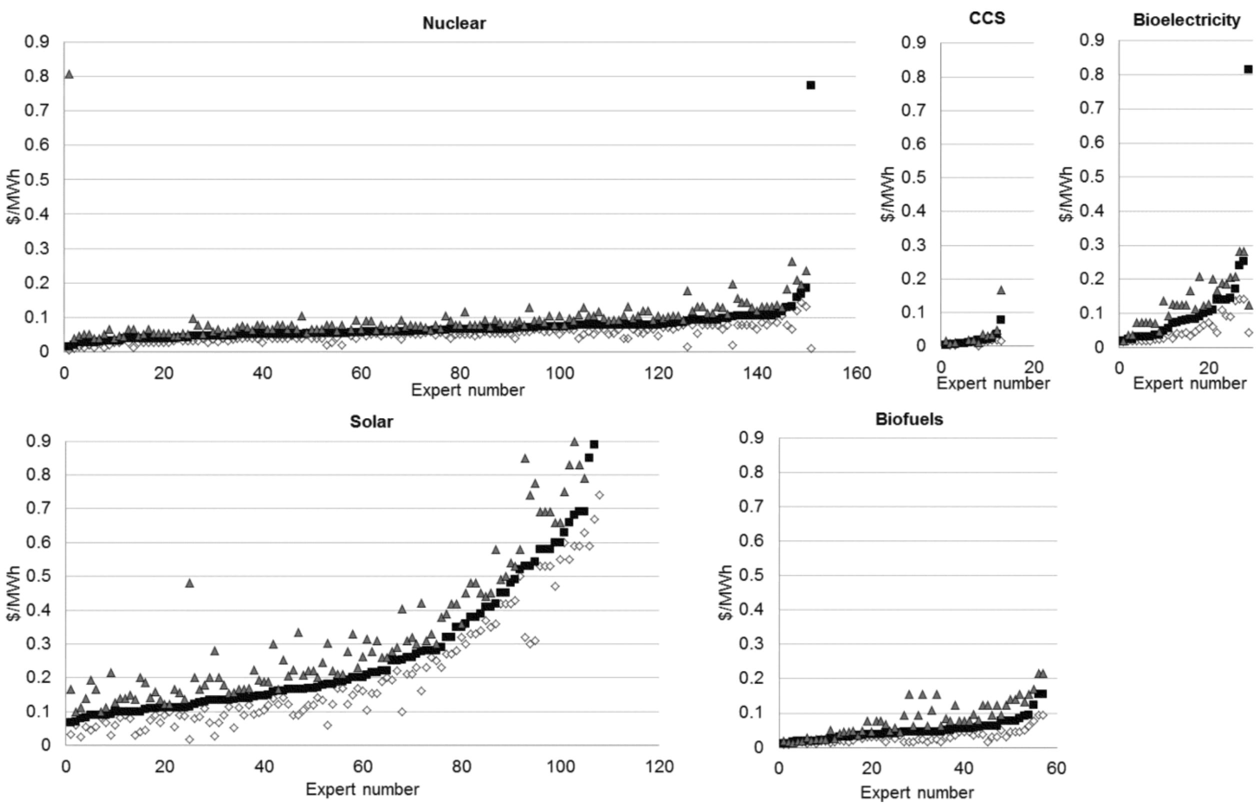

Figure 1. 2030 technology cost estimates from experts considered in the study under the "low" public R&D budget scenario. The white diamonds, black squares, and grey triangles represent, respectively, the 10th, 50th, and 90th percentile estimates. Expert estimates are ordered by increasing 50th percentile estimates (All in $/MWh).

These expert elicitations collect information on point estimates of future costs, (in this case, of 10th, 50th and 90th percentiles, with the exception of the Baker et al. studies, which asked experts about probabilities of particular goals being met and then converted the answers to percentiles), where the 50th percentile is the most likely future technology costs based on experts' judgments, and the 10th percentile gives some sense on future costs estimates under a breakthrough technology development. Figure 1 shows the elicited point estimates (50th, 10th and 90th percentiles) for the five energy technologies under the "low" R&D scenario. These data account for about one third of the observations, with the "medium" and "high" R&D scenarios accounting for the rest (see Table S4 in the SI for definitions of R&D scenarios). Table I shows the descriptive statistics of the individual participant data of the 16 expert elicitations, indicating the fraction of the observations representing various technologies, types of experts, R&D levels, etc.

Given the growing interest in consideration of uncertainty in science policy decisions(8), and the vast literature on the cognitive biases in the subjective assessment of probabilities(26), we focus here primarily on experts' confidence around central estimates. The uncertainty range ("Urange", henceforth) is defined as the difference between the 90th and 10th percentile divided by the 50th ([90th-percentile – 10th percentile]/50th percentile). It measures the percentage variation from each expert's median estimate within each of the R&D scenarios. Note that since Urange is a normalized metric, it can be pooled for all technologies, even if the standardized costs measure different parts of the technology. Hence, for this metric we present both pooled results and technology specific to show the robustness of our results to different assumptions. Conversely, elicited percentile metrics can be meaningfully compared only within technology due to the differences in what is included in the standardized costs. Figure 2 shows probability density functions of the Uranges for all data in our sample, both overall and for each technology. We also present results on the relationship between R&D investments, elicitation design, and expert selection variables on the breakthrough estimates (10th percentiles), given the interest in understanding the lower tail of future costs in various technologies separately.

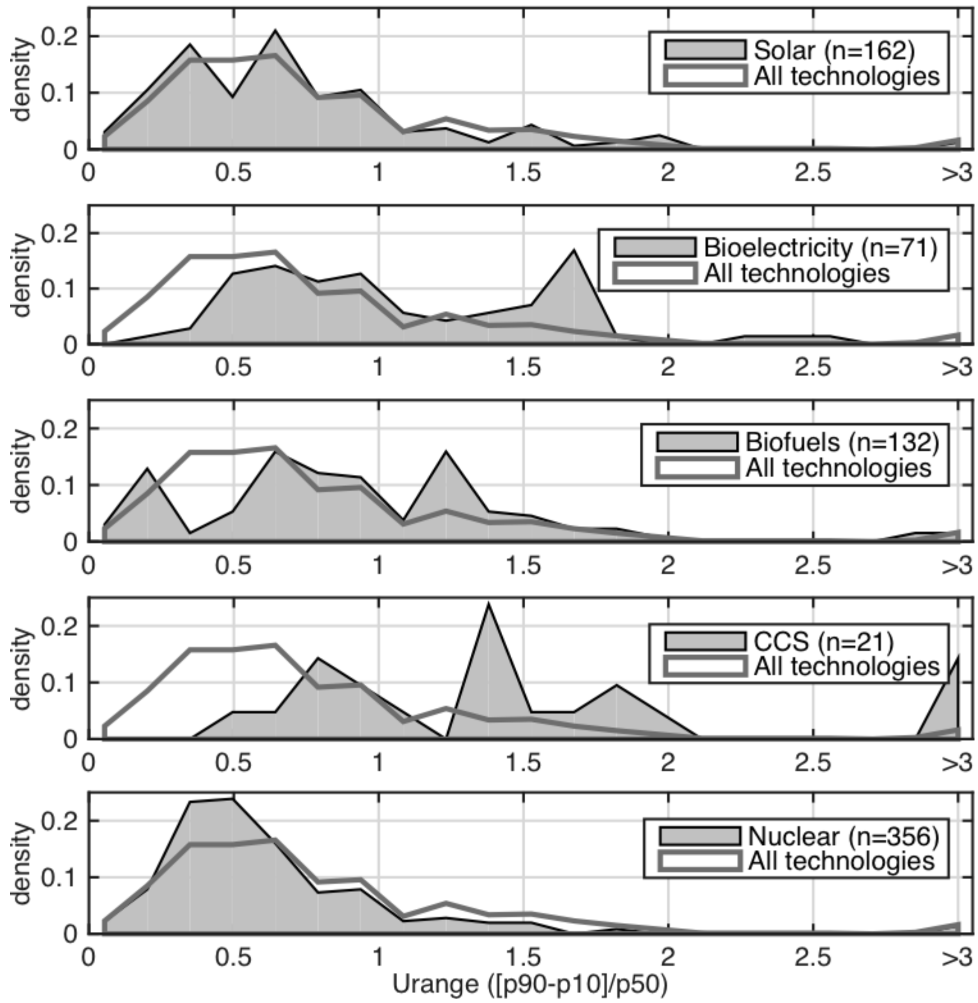

Figure 2. Distributions of uncertainty range (Urange) for all elicitations and R&D levels: pooling all five technologies and for each technology individually.

We are primarily interested in the extent to which the following four aspects may systematically affect the uncertainty range: (a) expert elicitation survey design, (b) technology characteristics, (c) expert characteristics, and (d) R&D investment levels on which the elicited values are conditional. We also evaluate the relationship between those variables and the 10th percentile for each of the five technology areas separately.

The literature on elicitation has looked at the differences in the design of elicitation protocols, highlighting in particular the importance of the expert selection phase and of the method by which the survey is administered (in person, via mail, or internet)[15-20, 27]. For instance, it devotes significant attention to issues such as the optimal number experts and the careful sampling for expert selection. It also points to the advantages of in-person elicitations[2, 17]. During in person interviews the researcher can devote more time to "debiasing" and can provide opportunities for interviewers to ask follow-up questions that prompt experts to consider a wider range of possible outcomes. However, in-person elicitations are far more costly and time-consuming, for both subjects and investigators. Researchers address this trade-off by carefully designing mail or online elicitation protocols. To date, there is no empirical quantitative evidence indicating whether this is a relevant consideration and what bias these different elicitation designs may have on confidence.

Uncertainty ranges across studies also vary due to the diversity in the technologies considered. Figure 1, for instance, highlights the wide diversity of expert opinion on the 50th percentile 2030 estimates of solar technologies under the low funding scenario. Possible explanations for such differences include the maturity of a given technology, the extent to which learning-by-doing has improved costs in the past, the number of technological paths which have already been explored, and the specific efficiency of each technological path. Some technologies appear to have rather firm lower bounds, e.g. due to thermodynamics, while others may not.

Similarly, expert background (e.g., institutional affiliation and country of residence) is likely to affect cost estimates[27, 28], but no study has systematically evaluated this claim. Moreover, elicited data is likely to be subject to availability and anchoring heuristics associated with experts' environment and experiences[26].

Finally, the suggested public R&D investment levels included in the elicitation can have an impact on uncertainty. The direction of this effect is largely an empirical question. In fact, it is possible that experts may make estimates with wider uncertainty ranges (lower confidence) under higher R&D investment assumptions if they have difficulties imagining outcomes that are far away from the actual state of the world. However, experts could also have narrower uncertainty ranges, if they expect higher R&D to solve technical issues that are unresolved under scenarios with lower R&D investments.

Understanding whether experts believe R&D will reduce or increase uncertainty will contribute to designing more effective public R&D portfolios. For example, the usefulness of public R&D includes not only the benefits to the performance of specific technologies but also the ability to make better R&D allocation decisions in future periods. The relationship between survey design, expert selection, and R&D investments and 10th percentile estimates is also policy-relevant, given the importance of governments providing technology options for the future that may not be funded by private actors alone[(29)], perhaps due to risk aversion or because they differ radically from current designs.

To estimate these effects given disparate studies we conduct a meta-analysis: a set of statistical techniques used to aggregate the results of multiple studies testing similar hypotheses and to thus enhance the overall reliability of findings[(30, 31)]. This approach accounts for differences across studies and provides results that are dependent on a consistent set of conditions across observations. This technique has been used in environmental economics since the 1990s[(32, 33)], with several recent applications in energy[(34-37)]. The use of individual primary data, which we rely on for this study, 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[(21, 31, 38-40)].

While uncertainty is the main focus of our study (Urange=Y1), we also investigate the impact of the aforementioned variables on the 10th percentile estimate (p10=Y2) as a proxy for the experts' views on what may be the outcome of technology breakthroughs for each of the five technology areas separately. In the SI we include results on the relationship between the variables of interest and the 50th percentile estimate. Our basic specification reads as follows:

$$\ln(Y_{itpr}) = \alpha + \beta \ln(S) + \gamma \ln(T) + \delta \ln(E) + \theta \ln(R) + \vartheta_{itp} + \varepsilon_{iptr}$$

where *i* indicates expert, *t* technology, *p* a given subtechnology, *r* the specific R&D scenario on which the elicited metric is conditional. We regress Urange (or alternatively p10) on the independent variables described above and summarized in Table I. Specifically, S is a dummy variable equal to one if the survey was conducted in person; T are dummy variables indicating the technology focus of the specific elicitation, with solar as the reference category; E are dummy variables indicating the expert was from academia or the public sector, with private sector being the reference category; and R are variables indicating the R&D scenario with which each estimate is associated. Due to the wide diversity in R&D funding levels associated with the expert estimates, we propose two specifications for R. In the first, dummy variables indicate medium and high funding (with business-as-usual funding being the reference). In a second specification, we use the continuous R&D variable (in dollars) and explore the possibility of diminishing marginal returns by including the squared term in the regression. We use random effects models in which each observations is a combination of expert and sub-technology, observed over different R&D scenarios, to control for expert effects, in addition to the other control variables related to survey design, expert selection, and R&D investment level[(23)]. Standard errors are clustered at the level of expert.

3 RESULTS

3.1 Estimating Uncertainty

The results of regressions for uncertainty range are presented in Table II. We first estimate the model by pooling data for all technologies (Models 1 and 2), and then for each of the five technologies separately. The first set of results has broad implications for the use of scientific expert opinion to integrate in science policy decisions, as well as for the design of elicitation protocols. The technology-specific results, conversely, are more applicable to decisions or elicitations covering the different five technologies considered. We drop results for three studies due to missing data (see SI for study information).

3.2 Relationship between Urange and survey design characteristics

Models 1 and 2 in Table II, which pool all technologies and include technology and random effects, indicate that elicitations conducted in person have uncertainty ranges that are 33% greater than those that were conducted online or over the mail, on average and *ceteris paribus*. This positive and statistically significant result (at a 1% level) is robust to conducting technology specific analysis (as shown in Models 3-7), the only exception being the positive but not statistically significant coefficient for bioelectricity (Model 5).

3.3 Relationship between Urange and technology categories

Pooled Models 1 and 2 in Table II show that, on average, Urange in solar is statistically different from those in the other four technology categories. Urange is on average roughly 17%, 19%, 18% and 63% higher in the case of nuclear, bioelectricity, biofuels, and CCS experts, respectively. That different technologies are associated with different perceptions of uncertainty is not surprising, but to the best of our knowledge this is the first empirical assessment of the extent to which experts' confidence is greater in some technologies versus others. In the specific case of the technologies considered here, the small number of new constructions in both nuclear and CCS is a likely source of their higher uncertainty.

3.4 Relationship between Urange and expert characteristics

In Model 1 (Table I), the coefficient associated with experts in academia suggests that, on average and across technologies, their Uranges are roughly 12% greater than those in the private sector. Public experts are associated with higher Uranges that those in the private sector, but the estimate is statistically significant at acceptable levels only when using a continuous R&D variable (Model 2). Finally, EU experts are, on average, more confident, with an uncertainty range that is 12% lower than that of US experts. These coefficients are also significant when the continuous variable for R&D investment is used (Model 2). Looking at the technology specific regressions, however, it is evident that the significance of the result is mostly attributable to experts in biofuel technologies (Model 5). European experts in all other technologies are associated with lower Uranges, but this result is not statistically significant.

3.5 Relationship between Urange and R&D variables

As shown in Table II, the different R&D scenarios upon which the cost estimates are conditional do not have a significant impact on experts' confidence. However, the higher R&D scenarios are associated with more uncertain estimates (lower confidence) around future costs for solar, and less uncertain estimates (higher confidence) for biofuels. These effects may be due to increasing R&D investments pushing researchers to expand the range of technological possibilities for solar, whereas in biofuels experts are more certain about the possibilities due to a focus on particular technical bottlenecks to overcome. In this respect, note that solar is the technology for which R&D has the largest effect on median (p50) future costs (see Table S8 in the SI). Conversely, for biofuels the medium and high R&D scenarios are associated with the lowest uncertainty ranges.

We acknowledge that the range of observed characteristics we control for in our regression is unlikely to account for all variation beyond the core technical judgments we are attempting to elicit. For example, in the SI we discuss our attempt to evaluate the impact of two additional elicitation variables of interest to the meta-analysis literature: the year in which the study was conducted and whether or not the results were published in the peer-reviewed literature. We were however unable to determine their impact in a robust manner. First, all elicitations were carried out only a few years apart, providing very little variation. Second, in a few cases the year of elicitation was different between different studies, resulting in collinearity issues with other variables. The same is true for the "published" variable vis-a-vis the E.U. and in-person variables for some technologies.

3.6 Estimating breakthrough outcomes

We regress the 10^{th} percentile cost estimate—which approximates the best outcome that the experts could imagine—on the independent variables separately by technology, and include results in Table III. As expected, the R&D variables are negative (higher R&D scenarios are associated with lower p10 values) and statistically significant, with the exception of CCS, which has only 18 observations. It is also notable that the size of the coefficient is quite different for all technologies in Models 1-5. The difference between the size of the coefficients can be due to the fact that the R&D bins relate to different R&D investments for different technologies as well as to different beliefs regarding the impact that R&D investments will have on future outcomes. In particular, the coefficient of the impact of the mid and high R&D scenarios on p10 is largest in the case of solar power.

Similarly to Urange, the effect of the in-person variable on p10 is also relatively robust across the different technologies: it is negative and significant in the solar, nuclear and biofuel regressions (those with the largest number of observations), positive but not significant for bioelectricity, and negative but not significant for CCS. Overall, these results suggest that in-person elicitations are likely to result in more optimistic p10 estimates compared to mail or internet elicitations. Lower p10 values is one of the reasons why in-person is associated with greater uncertainty ranges (lower confidence), as discussed in Section 3.5.

Unlike the results for Urange, the results for expert selection variables differ by technology area. Academic experts provided more optimistic p10 estimates for nuclear and CCS, and more pessimistic estimates for biofuels than their industry counterparts (see Models 2, 3, and 4, respectively, in Table III). And public sector experts provided more optimistic p10 estimates for nuclear and bioelectricity, and more pessimistic estimates for biofuels than their industry counterparts. EU experts are more pessimistic about p10 than their US counterparts in bioenergy and biofuels. This suggests that previous experiences are more conducive to differences in the sign of perceptions of breakthroughs than on the uncertainty range in the technology areas evaluated in this study.

4 DISCUSSION

These results show that decisions made in elicitation design and in the selection of experts can bias study results. If what we are really interested in is the effects of public R&D investment, as the NRC study suggests[(8)], we need to be wary of choices that have the potential to bias findings. Consequently, this analysis of multiple elicitation studies provide a basis with which to increase the reliability of elicited technology performance values by controlling for design choices, expanding the number of observations to include, and considering experts from varied geographical areas and backgrounds. Most importantly, we have quantified the relationship between expert and survey characteristics and experts' confidence in future energy technology outcomes.

We find that, on average, when experts respond to elicitations on the future costs of energy technologies in person, they ascribe lower confidence (larger uncertainty) to their estimates than when responding via mail or internet. This result is robust to a variety of specifications. Our results also indicate that in-person elicitations generate more optimistic expectations about how inexpensive the best case, or breakthrough, outcomes will be. These results, taken together, are somewhat surprising in that, in particular, the online elicitations included here took many steps to train experts about people's tendency toward overconfidence and to provide interactive tools to help them visualize the distributions generated from their individual responses. That in-person elicitations reduce confidence suggests that on-line techniques are still not yet a substitute for a well-trained interviewer who can convey the importance of thinking about extremes, ask relevant follow up questions, and prod an expert to move their thinking beyond glib responses and immediate gut feeling.. Another hypothesis is that experts who are amenable to investing their time in an in-person interview are, for some reason, more likely to consider multiple technology pathways and thus have, before the interview, different views on uncertainty. A longstanding challenge in expert elicitation has been to find ways to overcome experts' biases to think too narrowly about possible outcomes—even if many decision makers consider results with high confidence to be more useful than those with low confidence(13). These empirical results indicate that in-person interviews are more promising in addressing over-confidence, at least until on-line elicitations can incorporate some aspects of in-person interviews that at present they do not. We also find that in-person elicitations are associated with more optimistic breakthrough estimates p10. This effect is, to a large extent, what explains the results on the uncertainty range, given that an analysis of the relationship between p90 and in-person shows a very inconsistent impact of in-person (see Table S-9 in the SI). Thus, in-person interviews seem to increase the uncertainty range mainly by expanding the lower bound, rather than the upper bound.

We also find that US experts were more uncertain about future costs than EU experts and that academics are generally more uncertain than their industry counterparts. Provided that all are experts, interpreting this result for designing future elicitations is less straightforward than the in-person. If we want a broad swatch of expertise, it seems unwise to prefer Americans and academics to other types of experts in order to minimize within-expert over-confidence. Rather, this result demonstrates the need for a broad set of experts since their environment may affect their access to disparate information and thus their confidence about future technology costs. The technology-specific regressions indicate that, often, expert background is a significant predictor of confidence in specific technologies. This could be related again to availability biases—experts not engaged in taking technologies into the market may have more uncertainty regarding what may take to achieve commercialization, including uncertainties related to technology performance at scale. European experts may have lower uncertainty in general because they have had more recent experience with biofuels (a technology for which EU is statistically significant), and other technologies (for which the coefficient is also negative but not significant).

We also found that high R&D scenarios do not affect experts' confidence in the outcomes. The absence of an R&D effect contrasts the results for R&D on p10, where higher R&D investments are consistently associated with lower cost estimates. From a social perspective, public R&D investments in energy are thought provide multiple public goods: they are expected to improve technology outcomes, and also generate information that can, for example, help inform future decisions. These results, for these five technologies, provide a much stronger case that more R&D will improve technologies by 2030 than it does that more R&D will clarify expectations about which technologies will be most promising between now and then. The results of this study could be utilized in modeling exercises that aim to inform policy design related to energy technologies. These include the Energy Modeling Forum (EMF) and the Intergovernmental Panel on Climate Change (IPCC). The study authors are already interacting with the EMF group on a formal basis and expect that this interaction will help transfer the insights from this study into a range of integrated assessment models.

The insights here are derived from a relatively large set of existing energy technology expert elicitations, which were, like all elicitations, costly and time-consuming. That in-person interviews seem to address overconfidence better than on-line elicitations suggests that there may be some value of information provided by the more expensive in-person format. Scaling up efforts to perform more elicitations might ultimately be helped by comparing the benefits of lower confidence to the costs on in-person interviews, as well as to the alternative of improving online elicitations. An important opportunity of future work would be to test the results presented here in an experimental setting. Such work could build on recent work in which EU experts and US experts responded to the same online elicitation tool[(41)]. This level of control would allow researchers to identify statistically significant differences in the answers of experts in both regions and could also be used to assess the effect of question format[(21)].

We have presented evidence that analyzing multiple expert elicitations on the future of energy technologies can provide insights for survey design and expert selection. Despite the dearth of alternative means by which to estimate future technology outcomes, expert elicitations remain vulnerable to criticisms of being unrepresentative, merely subjective, and based on opinions rather than facts. If elicitations are to be considered sufficient evidence on which to stake decisions involving potentially billions of dollars of public funds they need to be credible. The results here suggest that an empirically based understanding of what drives the range of experts' responses can increase the credibility and effectiveness of expert elicitation in supporting policy decisions involving science and innovation.

REFERENCES

1. Hogarth RM. Judgement and choice the psychology of decision Chichester, UK; New York: Wiley; 1987.

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

3. Cooke RL. Experts in Uncertainty: Opinion and Subjective Probability in Science. New York: Oxford University Press; 1991.

4. EPA. Expert Elicitation Task Force White Paper. Report. Washington DC: {U.S.} Environmental Protection Agency; Prepared for the Science and Technology Policy Council, 2011.

5. Nakicenovic N, Riahi K. An Assessment of Technological Change Across Selected Energy Scenarios. Research Report. IIASA, 2002 May. Report No.: RR-02-005.

6. Nordhaus W. A Question of Balance: Weighing the Options on Global Warming Policies: Yale University Press; 2008.

7. Gallagher KS, Grubler A, Kuhl L, Nemet G, Wilson C. The Energy Technology Innovation System. Annual Review of Environment and Resources. 2012;37(1):137--62.

8. {NRC}. Prospective Evaluation of Applied Energy Research and Development at DOE (Phase Two). Washington: The National Academies Press; 2007.

9. Bolger F, Rowe G. The Aggregation of Expert Judgment: Do Good Things Come to Those Who Weight? Risk Analysis. 2014;35(1):5--11.

10. Bolger F, Rowe G. There is Data, and then there is Data: Only Experimental Evidence will Determine the Utility of Differential Weighting of Expert Judgment. Risk Analysis. 2015;35(1):21-6.

11. Cooke RM. The Aggregation of Expert Judgment: Do Good Things Come to Those Who Weight? Risk Analysis. 2015;35(1):12-5.

12. Winkler RL. Equal Versus Differential Weighting in Combining Forecasts. Risk Analysis. 2015;35(1):16--8. doi: 10.1111/risa.12302.

13. Morgan MG. Our Knowledge of the World is Often Not Simple: Policymakers Should Not Duck that Fact, But Should Deal with It. Risk Analysis. 2015;35(1):19--20.

14. Morgan MG. Use (and abuse) of expert elicitation in support of decision making for public policy. Proceedings of the National Academy of Sciences. 2014:201319946.

15. Raiffa H. Decision analysis: introductory lectures on choices under uncertainty. Oxford: Addison-Wesley; 1968.

16. Keeney RL, Winterfeldt DV. Eliciting probabilities from experts in complex technical problems. Transactions on Engineering Management. 1991;Vol. 38:191-201.

17. Meyer MA, Booker JM. Eliciting and Analysing Expert Judgment: A Practical Guide. London: Academic Press Ltd.; 1991.

18. Phillips LD. Group elicitation of probability distributions: Are many heads better than one. 1999.

19. Clemen RT, Reilly T. Making hard decisions with Decision Tools. Pacific Grove, CA: Duxbury; 2001.

20. Walls L, Quigley J. Building prior distributions to support Bayesian reliability growth modelling using expert judgement. Reliability Engineering and System Safety. 2001;74(2):117-28.

21. Speirs-Bridge A, Fidler F, McBride M, Flander L, Cumming G, Burgman M. Reducing overconfidence in the interval judgments of experts. Risk Analysis. 2010;30(3):512-23.

22. Anadon LD, Nemet G, Verdolini E. The future costs of nuclear power using multiple expert elicitations: effects of RD&D and elicitation design. Environmental Research Letters. 2013;8(3):034020.

23. Verdolini E, Anadon LD, Lu J, Nemet GF. The effects of expert selection, elicitation design, and R&D assumptions on experts' estimates of the future costs of photovoltaics. Energy Policy. 2015(in press).

24. Victor D, Zhou D. Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press; 2014.

25. Kriegler E, Weyant J, Blanford G, Krey V, Clarke L, Edmonds J, et al. The role of technology for achieving climate policy objectives: overview of the EMF 27 study on global technology and climate policy strategies. Climatic Change. 2014;123(3-4):353-67.

26. Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. Science . 1974;185(4157):1124-31.

27. O'Hagan A, C.E. Buck, A. Daneshkhan, J.R. Eiser, P.H. Garthwaite, D.J. Jenkinson, J.E. Oakey, and T. Rakow Uncertain Judgments: Eliciting Experts' Probabilities: John Wiley & Sons, Ltd.; 2006.

28. Bosetti V, Catenacci M, Fiorese G, Verdolini E. The future prospect of PV and CSP solar technologies: An expert elicitation survey. Energy Policy. 2012;49:308-17.

29. IAC. Lighting the Way: Toward a Sustainable Energy Future. The Role of Government and the Contribution of Science and Technology Amsterdam: InterAcademy Council (IAC); 2007.

30. Glass GV. Primary, Secondary, and Meta-Analysis of Research. Educational Researcher. 1976;5(10):3-8.

31. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. How a Meta-Analysis Works. Introduction to Meta-Analysis: John Wiley & Sons, Ltd; 2009. p. 1-7.

32. Nelson J, Kennedy P. The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment. Environmental & Resource Economics. 2009;42(3):345-77.

33. Matarazzo B, and Peter Nijkamp Meta-analysis for comparative environmental case studies: methodological issues. International Journal of Social Economics. 1997;Vol. 24(No. 7/8/9):799-811.

34. Rose A, Dormady N. A Meta-Analysis of the Economic Impacts of Climate Change Policy in the United States The Energy Journal. 2011;32(2):143--66.

35. Barker T, Jenkins K. The Costs of Avoiding Dangerous Climate Change: Estimates derived from a meta-analysis of the literature. Human Development Report Office Occasional Paper. United Nations Human Development Report 2007/2008, 2007.

36. Zamparini L, Reggiani A. Meta-analysis and the value of travel time savings: A transatlantic perspective in passenger transport. Netw Spat Econ. 2007;7(4):377-96.

37. Havranek T, Irsova Z, Janda K. Demand for Gasoline Is More Price-Inelastic than Commonly Thought. Berkeley, CA: 2011.

38. Reade MC, Anthony Delaney, Michael J. Bailey, David A. Harrison, Donald M. Yealy, Peter G. Jones, Kathryn M. Rowan, Rinaldo Bellomo and Derek C. Angus. Prospective meta-analysis using individual patient data in intensive care medicine. Intensive Care Medicine. 2009;Vol. 36(No. 1):1-21.

39. Broeze KA, Brent C. Opmeer, Fulco van der Veen, Patrick M. Bossuyt, Siladitya Bhattacharya, and Ben W.J. Mol Individual patient data meta-analysis: a promising approach for evidence synthesis in reproductive medicine. Human Reproduction Update. 2010;Vol.16(No.6):pp. 561–7.

40. Riley RD. Commentary: Like it and lump it? Meta-analysis using individual participant data. International Journal of Epidemiology. 2010;Vol.39:1359–61.

41. Anadon LD, Bosetti V, Bunn M, Catenacci M, Lee A. Expert Judgments about RD&D and the Future of Nuclear Energy. Environmental Science \& Technology. 2012;46(21):11497-504.

TABLES

Table I. Descriptive statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
Urange	742	0.869	2.186	0.054	57.539
P10_LEC	742	0.058	0.057	-	0.480
P50_LEC	742	0.081	0.071	0.001	0.772
RD	694	3,563	8,629	13	80,000
RD_high	742	0.278	0.448	0	1
RD_mid	742	0.255	0.436	0	1
RD_low	742	0.381	0.486	0	1
Bioelec.	742	0.096	0.294	0	1
Biofuel	742	0.178	0.383	0	1
Nuclear	742	0.480	0.500	0	1
Solar	742	0.218	0.413	0	1
CCS	742	0.028	0.166	0	1
Academia	742	0.330	0.471	0	1
Private	742	0.395	0.489	0	1
Public	742	0.275	0.447	0	1
EU	742	0.380	0.486	0	1
Inperson	742	0.287	0.453	0	1

Table II. Factors affecting the uncertainty range (Y=ln[Urange]).

Y = ln(Urange)	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	pooled	pooled	Solar	Nuclear	Bioelec.	Biofuel	CCS
Inperson	0.286***	0.280***	0.317***	0.455***	0.0637	0.333*	0.627***
	[0.000112]	[0.00101]	[1.99e-06]	[0.00326]	[0.682]	[0.0650]	[0.00382]
academia	0.111**	0.109**	0.0115	0.189**	0.0149	0.0357	0.00387
	[0.0285]	[0.0390]	[0.890]	[0.0131]	[0.912]	[0.816]	[0.992]
public	0.0630	0.0767*	-0.0673	0.156***	0.108	0.00809	-0.173
	[0.128]	[0.0853]	[0.286]	[0.00578]	[0.384]	[0.952]	[0.646]
EU	-0.130***	-0.130***	-0.0878	-0.0374	-0.100	-0.404**	
	[0.00162]	[0.00463]	[0.254]	[0.416]	[0.386]	[0.0280]	
RD_high	0.00111		0.0705***	-0.00858	0.0354	-0.138**	0.140
	[0.947]		[0.00133]	[0.641]	[0.172]	[0.0244]	[0.488]
RD_mid	0.00622		0.0339*	0.000876	0.0320	-0.0931*	0.217
	[0.698]		[0.0597]	[0.957]	[0.346]	[0.0946]	[0.441]
ln_RD		-2.818					
		[0.324]					
ln_RDsq		1.404					
		[0.325]					
Nuclear	0.155**	0.175***					
	[0.0342]	[0.00994]					
Bioelec	0.176***	0.197***					
	[0.00711]	[0.00459]					
Biofuel	0.164**	0.168**					
	[0.0217]	[0.0216]					
CCS	0.492***	0.532***					
	[0.000236]	[0.000117]					
Observations	678	694	162	322	66	110	18
Number of experts, by subtech	301	276	71	159	23	40	8
Nr Clusters	160	146	39	66	23	24	8
R2 overall	0.229		0.319	0.174	0.0382	0.128	0.378
R2 within	0.000481	0.0287	0.166	0.00362	0.0274	0.154	0.186

Clustered values in brackets

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

Table III. Factors affecting the "best" outcome (Y=ln[p10]) for individual technologies.

Y=ln(p10) VARIABLES	(1) Solar	(2) Nuclear	(3) Bioelectricity	(4) Biofuel	(5) CCS
lnperson	-0.0845***	-0.0202***	0.00709	-0.0163***	-0.00359
	[6.09e-07]	[4.00e-06]	[0.518]	[0.000342]	[0.617]
academia	0.000778	-0.0216***	-0.00691	0.00995**	-0.0102*
	[0.959]	[2.83e-07]	[0.607]	[0.0149]	[0.0714]
public	0.0194	-0.0118**	-0.0290**	0.0137**	-0.00131
	[0.213]	[0.0101]	[0.0409]	[0.0397]	[0.871]
EU	-0.0121	0.00349	0.0437***	0.0128***	
	[0.209]	[0.385]	[0.00103]	[0.000357]	
RD_high	-0.0316***	-0.00961***	-0.0169***	-0.00465*	-0.00265
	[1.63e-08]	[0]	[5.45e-06]	[0.0944]	[0.151]
RD_mid	-0.0140***	-0.00478***	-0.00808***	-0.00169	-0.00221
	[0.000188]	[0.000221]	[0.000361]	[0.508]	[0.107]
Observations	162	322	66	110	18
Number of experts, by subtech	71	159	23	40	8
Nr Clusters	39	66	23	24	8
R2 overall	0.419	0.284	0.477	0.240	0.356
R2 within	0.371	0.372	0.476	0.119	0.365

Clustered pvalues in brackets

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

Supporting Information for:

Quantifying the effects of expert selection and elicitation design on experts' confidence in their judgments about future energy technologies

Gregory F. Nemet[1,2,], Laura Diaz Anadon[3], Elena Verdolini[4]*

[1] La Follette School of Public Affairs, University of Wisconsin-Madison, WI 53706, USA
[2] Nelson Institute Center for Sustainability and the Global Environment (SAGE), University of Wisconsin-Madison, WI 53706, USA
[3] John F. Kennedy School of Government, Harvard University, 79 JFK Street, Cambridge, MA 02138

[4] Fondazione Eni Enrico Mattei and Centro Euro-Mediterraneo sui Cambiamenti Climatici, Corso Magenta 63,20123 Milano, Italy

* Corresponding author: nemet@wisc.edu.

Contents

List of Tables

1. Summary of expert elicitations included in this study

Table S-1. Characteristics of each energy technology expert elicitations study.

Technology	Study Code	Group	experts	source/publication	year of elicitation
1.Nuclear	11	Baker	4	Baker, E., H. Chon and J. M. Keisler (2008). "Advanced Nuclear Power: Combining economic analysis with expert elicitations to inform climate policy." Available at SSRN 1407048.	2007
	12	Harvard	25	*Anadon, L. D., V. Bosetti, M. Bunn, M. Catenacci and A. Lee (2012). "Expert Judgments about RD&D and the Future of Nuclear Energy." Environmental Science & Technology 46(21): 11497-11504.	2010
	13	FEEM	30	*Anadon, L. D., V. Bosetti, M. Bunn, M. Catenacci and A. Lee (2012). "Expert Judgments about RD&D and the Future of Nuclear Energy." Environmental Science & Technology 46(21): 11497-11504.	2011
	14	CMU	12	*Abdulla, A., I. L. Azevedo and M. G. Morgan (2013). "Expert assessments of the cost of light water small modular reactors." Proceedings of the National Academy of Sciences.	2011
2.Solar	21	Baker	3	*Baker, E., H. Chon and J. Keisler (2009). "Advanced solar R&D: Combining economic analysis with expert elicitations to inform climate policy." Energy Economics 31(Supplement 1): S37-S49.	2007
	22	Harvard	9	Anadon, L. D., M. Bunn, G. Chan, M. Chan, C. Jones, R. Kempener, A. Lee, N. Logar and V. Narayanamurti (2011). Transforming U.S. Energy Innovation. Cambridge, MA, Belfer Center for Science and International Affairs, Harvard Kennedy School.	2010
	23	FEEM	13	*Bosetti, V., M. Catenacci, G. Fiorese and E. Verdolini (2012). "The future prospect of PV and CSP solar technologies: An expert elicitation survey." Energy Policy 49: 308-317.	2011
	25	NearZero	22	Mason Inman, Inman, M. (2012). How Low Will Photovoltaic Prices Go? An Expert Discussion. . Near Zero (2012).	2011
	26	Curtwright	18	*Curtwright, A.E., Expert Assessments of Future Photovoltaic Technologies. Environmental Science & Technology(**2008),** 42 (24).	2008
3. Bioelectricity	31	Baker	4	Unpublished	2007
	32	Harvard	7	Anadon, L. D., M. Bunn, G. Chan, M. Chan, C. Jones, R. Kempener, A. Lee, N. Logar and V. Narayanamurti (2011). Transforming U.S. Energy Innovation. Cambridge, MA, Belfer Center for Science and International Affairs, Harvard Kennedy School.	2010
	33	FEEM	16	*Fiorese, G., M. Catenacci, V. Bosetti and E. Verdolini "The power of biomass: Experts disclose the potential for success of bioenergy technologies." Energy Policy(0).	2011
4. Biofuel	41	Baker	3	*Baker, E. and J. M. Keisler (2011). "Cellulosic biofuels: Expert views on prospects for advancement." Energy 36(1): 595-605.	2008
	42	Harvard	8	Anadon, L. D., M. Bunn, G. Chan, M. Chan, C. Jones, R. Kempener, A. Lee, N. Logar and V. Narayanamurti (2011). Transforming U.S. Energy Innovation. Cambridge, MA, Belfer Center for Science and International Affairs, Harvard Kennedy School.	2010
	43	FEEM	15	*Fiorese, G., M. Catenacci, E. Verdolini and V. Bosetti (2013). "Advanced biofuels: Future perspectives from an expert elicitation survey." Energy Policy 56(0): 293-311.	2011
5. CCS	51	Baker	3	*Baker, E., H. Chon and J. Keisler (2009). "Carbon capture and storage: combining economic analysis with expert elicitations to inform climate policy." Climatic Change 96(3): 379-408.	2007
	52	Harvard	8	Chan, G., L. D. Anadon, M. Chan and A. Lee (2011). "Expert elicitation of cost, performance, and {RD\&D} budgets for coal power with {CCS}." Energy Procedia 4: 2685-2692.	2010
	53	FEEM	10	Bosetti, V. and E. C. Ricci (2013). Future prospects of carbon capture technologies: an expert elicitation, FEEM.	2012
	57	Chung	5	*Chung, T.S., et al., Expert assessments of retrofitting coal-fired power plants with carbon dioxide capture technologies. Energy Policy (2011), doi:10.1016/j.enpol.2011.04.038	2011

Due to missing data studies 25, 53, and 57 were excluded from the analyses included in the main text.

Table S-2. Observations by technology and source.

Technology	Source: Harvard	UMass	FEEM	CMU	Total
PV	69	6	39	48	162
Bioelectricity	21	12	38	0	71
Biofuel	90	6	36	0	132
CCS	15	6	0	0	21
Nuclear Energy	162	12	172	10	356
Total	357	42	285	58	742

Table S-3. Definitions of technologies and sub-technologies included in the elicitations.

Major Technology	Sub technology	Technology Detail
1. Nuclear	1. GenIII/III+	1. LWR/GenIII/III+
	2. SMR	(both Gen III/III+ and Gen IV configurations)
	3. GenIV	
2. Solar	1. novel PV	1. novel HE
		2. inorganic, organic, 3rd gen
		3. excitonic
	2. thin-film	1. thin-film (2d-2comp)
		2. thin-film(2a-2b)
		3. thin-film(2c-2d)
	3. x-Si	1. Crystalline Silicon PV
		2. other x-Si
	4. CPV	
	5. all PV	1. PV mix
		2. PV and CSP
		3. Solar PV
	6. application	1. Residential scale
		2. Commercial scale
		3. Utility scale
3. Bioelectricity	1. all Bioelectricity	
4. Biofuel	1. Diesel substitute	
	2. Gasoline substitute	
	3. Jet Fuel substitute	
	4. all Biofuel	
5. CCS	1. pre-Combustion	
	2. post-Combustion	1. Amines
		2. Absorption
		3. Adsportion
		4. Chilled ammonia
		5. Membranes
		6. Other PC
	3. Oxyfuel	
	4. all CCS	

Note that not all elicitations reported expert estimates at the “Technology Detail” level.

Table S-4. Assignment of study-specific R&D levels to standardized R&D bins.

Technology	Study	Group	Assignment of R&D scenarios to bins			
			Not included	Low	Mid	High
1. Nuclear	11	Baker		BAU	Rec	10x
	12	Harvard	0.5X	BAU	Rec	10x
	13	FEEM	0.5X	BAU	Rec	10x
	14	CMU		BAU	--	--
2. Solar	21	Baker		Low	Mid	--
	22	Harvard	0.5X	BAU	Rec	10x
	23	FEEM		BAU	1.5x	2x
	25	NearZero	all	--	--	--
	26	Curtwright	10x RD & Deployment	status quo	--	10x
3. Bioelectricity	31	Baker		Low	Mid	High
	32	Harvard	0.5X	BAU	Rec	10x rec
	33	FEEM		Low	Mid	High
4. Biofuel	41	Baker			Mid	High
	42	Harvard	0.5X	Low	Rec	10x
	43	FEEM		Low	Mid	High
5. CCS	51	Baker		Low	Mid	High
	52	Harvard	0.5X	BAU	Rec	10x
	53	FEEM	all	--	250	--
	57	Chung	10x RD & Deployment	BAU	--	10x RD&D

2. Data cleaning

Here we detail the data cleaning issues that we have addressed in putting together the individual participant data in the dataset used in the meta-analysis:

Missing values. The elicitation datasets contained empty values, zeroes, and values marked as 'not available.' Through conversations with the authors we were able to determine that the zeroes were empty values (in some cases experts had decided that they did not want to provide an estimate), and that in some cases empty values were caused by a mistake copying data.

Outliers. In some cases elicited values seemed too low or too high. It was necessary to determine whether these values were a result of differences in units, of mistakes copying results, or of the true beliefs of the expert.

Variables not included in published elicitations. Some of the variables we investigate in the meta-analysis—e.g., expert affiliation (private sector, public institution, and academia) and nationality—were not reported in the individual participant data we obtained from elicitation authors. Obtaining this information requires dealing with confidentiality issues and more dialogue with the authors.

Inconsistencies. We corrected for inflation and ensured that costs under lower RD&D scenarios stochastically dominate costs under high RD&D scenarios.

Differences in technological specificity across and within studies. For example, some of the studies collected information on all Gen. IV nuclear systems, while others collected information on more specific reactor configurations (fast reactors and high temperature reactors). We conducted our analysis using the least common denominator for technological specificity, which we refer to as the 5 "technologies" in

Table S-3.

Using data from elicitations collecting different values. All nuclear fission elicitations collected the same key metric (overnight capital cost), but in other energy technology elicitations groups collected different metrics. For example, FEEM collected levelized cost of electricity, while Harvard collected data on different cost components, including cell efficiency and cost and inverter efficiency and cost. In this case it was necessary to construct a model to make the data comparable using common assumptions (e.g., insolation and discount rates).Details in this respect are explained below (Section 3)

Extensive interactions with elicitation authors were essential in the data cleaning process. This dialogue was facilitated by the fact that the authors have collaborations with most of the elicitation authors listed in Table 1, making it possible to engage in a back and forth to ensure that the data are accurately represented.

3. Variable definitions and descriptive statistics

Table S-5 describes the independent variables investigated in this study and the rational for their inclusion.

Table S-5. Definitions of variables used.

Technology Outcome	
p10, p50, p90	Levelized energy cost elicited conditional on probabilities of 0.10, 0.50, and 0.90 in future, e.g. 2030.
p50_2010	Median levelized energy cost, current (defined as 2010).
Norm_p50	Normalized levelized energy cost in 2030 compared to 2010. p50/p50_2010
Urange	Uncertainty range: (p90-p10)/p50
Input units	Metric being elicited. Some studies elicit levelized energy cost ($/kwh) directly while others separately elicited components of it. Section 3 provides information on assumptions used for conversion.
R&D investment characteristics	
RD level	Funding amount per year, converted into millions of constant 2010 dollars.
RD bins	Level of funding coded as high (RD1), medium (RD2), and low (RD3). See Table S3.
Survey design characteristics	
Study	Code of the 19 studies (11-57) used in meta-analysis. See Table S1.
Published	Whether study has been published in peer-reviewed journal.
Technology	Technology category, including nuclear, solar, bioelectricity, biofuel, and CCS. See Table S2
Subtech	Disaggregated technology categories, e.g. multiple types of PV under solar energy. See Table S2.
Inperson	Elicitation conducted as an in-person interview. Otherwise conducted as internet or mail survey.
Year estimate applies	Year for which p50, p10, and p90 values were elicited.
Year estimate made	Year elicitation was conducted; typically precedes publication year.
Expert Characteristics	
Expert	4-digit expert code: first 2 digits are study code, 3rd and 4th digits are for experts 01-99 within that study. Used for expert fixed effects.
Affiliation	Whether expert is affiliated with a university (academia), industry (private), or government laboratory (public).
Country	Residence of expert, converted to EU and non-EU.

4. Results of additional regressions

Table S-6. Estimates for models of Y = ln(p10), with continuous R&D.

	(1)	(3)	(5)	(7)	(9)
VARIABLES	t_pv	t_nu	t_bioe	t_biof	t_ccs
lnperson	-0.0819***	-0.0230***	-0.00673	-0.0216***	-0.00610
	[0.000238]	[2.48e-06]	[0.597]	[8.00e-05]	[0.374]
academia	0.000255	-0.0223***	-0.00440	0.0127**	-0.0117**
	[0.990]	[1.27e-07]	[0.748]	[0.0126]	[0.0231]
public	0.00961	-0.0115**	-0.0275*	0.0160**	-0.00165
	[0.678]	[0.0130]	[0.0502]	[0.0260]	[0.836]
EU	-0.0256	0.00517	0.0517***	0.0132***	0
	[0.180]	[0.190]	[0.000171]	[0.000620]	
		-	-	-	-
ln_RD	-0.0135***	0.00287***	0.00508***	0.00221***	0.000450
	[1.55e-05]	[0]	[0.00408]	[0.00771]	[0.489]
ln_RDsq					
Observations	114	356	71	132	21
Number of identif	46	159	23	40	8
Nr Clusters	25	66	23	24	8
R2 overall	0.385	0.292	0.419	0.235	0.309
R2 within	0.328	0.350	0.169	0.102	0.0676

Robust pval in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table S-7. Estimates of models for Y = ln(p50).

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	t_pv	t_nu	t_nu (no UMASS)	t_bioe	t_biof	t_ccs
lnperson	-0.0935***	0.0191	-0.00722	0.0194	-0.0150**	0.0173
	[0.00155]	[0.368]	[0.278]	[0.231]	[0.0448]	[0.436]
academia	-0.00433	-0.0124	-0.0221***	-0.0181	0.0123	-0.00442
	[0.799]	[0.204]	[0.000108]	[0.316]	[0.145]	[0.659]
public	0.0193	-0.00653	-0.00851	-0.0410*	0.0223*	-0.00624
	[0.312]	[0.252]	[0.142]	[0.0720]	[0.0502]	[0.553]
EU	0.0191	0.00682	0.00560	0.0671***	0.00561	
	[0.572]	[0.190]	[0.250]	[0.000377]	[0.185]	
RD_hi	-0.0394***	-0.0130***	-0.0131***	-0.0199***	-0.0118***	-0.0125
	[5.96e-10]	[0]	[0]	[1.31e-07]	[0.00101]	[0.236]
RD_mid	-0.0196***	-0.00679***	-0.00691***	-0.0124***	-0.00802***	-0.00810
	[7.59e-08]	[7.54e-08]	[1.27e-07]	[0.000225]	[0.00367]	[0.206]
Yearestimatemade	-0.0151					
	[0.379]					
Technology effects		yes	yes			
Observations	162	322	310	66	110	18
Number of identif	71	159	147	23	40	8
Nr Clusters	39	66	62	23	24	8
R2 overall	0.339	0.0973	0.266	0.545	0.216	0.234
R2 within	0.451	0.477	0.477	0.519	0.399	0.303

Robust pval in brackets

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

Notes: Specification 1 includes Yearestimatemade to mirror the specification present in Verdolini et al. (2015). Specifications 2 and 3 include technology dummies to mirror the specification presented in Anadon et al. (2013).

Table S-8. Estimates of models for Y = ln(p90).

	(1)	(2)	(3)	(4)	(5)
VARIABLES	t_pv	t_nu	t_bioe	t_biof	t_ccs
lnperson	-0.0376	0.0687**	0.0302**	-0.0117	0.0428
	[0.107]	[0.0307]	[0.0438]	[0.547]	[0.320]
academia	0.00596	0.00233	-0.00338	0.0234	-0.00443
	[0.808]	[0.874]	[0.826]	[0.206]	[0.699]
public	0.0240	0.00432	-0.0331	0.0320*	-0.00757
	[0.249]	[0.630]	[0.111]	[0.0951]	[0.489]
EU	-0.0530***	0.00634	0.0793***	-0.00996	
	[0.00178]	[0.401]	[4.09e-07]	[0.505]	
RD_hi	-0.0445***	-0.0178***	-0.0272***	-0.0264***	-0.0248
	[4.28e-09]	[0]	[4.01e-07]	[0.000611]	[0.240]
RD_mid	-0.0208***	-0.00902***	-0.0142***	-0.0169**	-0.0150
	[7.34e-08]	[5.64e-07]	[0.00208]	[0.0237]	[0.257]
Observations	162	322	66	110	18
Number of experts, by subtech	71	159	23	40	8
Nr Clusters	39	66	23	24	8
R2 overall	0.239	0.0878	0.659	0.204	0.321
R2 within	0.445	0.443	0.456	0.394	0.288

Clustered pval in brackets

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

5. Covariates

Table S-9. Correlation matrix of covariates.

	RD_midhi	academia	public	EU	Inperson	Publis~d	Yeares~e	t_bioe	t_biof	t_nu	t_pv	t_ccs
RD_midhi	1.0000											
academia	-0.0082	1.0000										
public	0.0266	-0.4323	1.0000									
EU	-0.0229	-0.1247	0.1957	1.0000								
Inperson	0.0932	0.0232	0.1431	0.1783	1.0000							
Published	-0.0206	-0.1325	0.2169	0.4821	0.2306	1.0000						
Yearestima~e	-0.0590	-0.1306	0.0351	0.6628	-0.3265	0.2680	1.0000					
t_bioe	0.0478	0.0152	0.0049	0.1040	0.3000	-0.1383	-0.0241	1.0000				
t_biof	0.0334	0.2579	-0.1444	-0.1029	0.0320	-0.4239	0.0332	-0.1513	1.0000			
t_nu	-0.0947	-0.1007	0.0430	0.2040	-0.4782	0.5189	0.2763	-0.3124	-0.4467	1.0000		
t_pv	0.0573	-0.1075	0.0399	-0.1718	0.3353	-0.1787	-0.2851	-0.1719	-0.2458	-0.5075	1.0000	
t_ccs	-0.0192	-0.0507	0.0951	-0.1336	-0.0005	0.1051	-0.1558	-0.0555	-0.0794	-0.1639	-0.0902	1.0000

NOTE DI LAVORO DELLA FONDAZIONE ENI ENRICO MATTEI

Fondazione Eni Enrico Mattei Working Paper Series

Our Note di Lavoro are available on the Internet at the following addresses:
http://www.feem.it/getpage.aspx?id=73&sez=Publications&padre=20&tab=1
http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=266659
http://ideas.repec.org/s/fem/femwpa.html
http://www.econis.eu/LNG=EN/FAM?PPN=505954494
http://ageconsearch.umn.edu/handle/35978
http://www.bepress.com/feem/
http://labs.jstor.org/sustainability/

NOTE DI LAVORO PUBLISHED IN 2016

ET	1.2016	Maria Berrittella, Carmelo Provenzano: An Empirical Analysis of the Public Spending Decomposition on Organized Crime
MITP	2.2016	Santiago J. Rubio: Sharing R&D Investments in Breakthrough Technologies to Control Climate Change
MITP	3.2016	W. Brock, A. Xepapadeas: Spatial Heat Transport, Polar Amplification and Climate Change Policy
ET	4.2016	Filippo Belloc: Employee Representation Legislations and Innovation
EIA	5.2016	Leonid V. Sorokin, Gérard Mondello: Sea Level Rise, Radical Uncertainties and Decision-Maker's Liability: the European Coastal Airports Case
ESP	6.2016	Beatriz Martínez, Hipòlit Torró: Anatomy of Risk Premium in UK Natural Gas Futures
ET	7.2016	Mary Zaki: Access to Short-term Credit and Consumption Smoothing within the Paycycle
MITP	8.2016	Simone Borghesi, Andrea Flori: EU ETS Facets in the Net: How Account Types Influence the Structure of the System
MITP	9.2016	Alice Favero, Robert Mendelsohn, Brent Sohngen: Carbon Storage and Bioenergy: Using Forests for Climate Mitigation
EIA	10.2016	David García-León: Adapting to Climate Change: an Analysis under Uncertainty
ESP	11.2016	Simone Tagliapietra: Exploring the Potential for Energy Efficiency in Turkey
MITP	12.2016	Gabriel Chan, Carlo Carraro, Ottmar Edenhofer, Charles Kolstad, Robert Stavins: Reforming the IPCC's Assessment of Climate Change Economics
MITP	13.2016	Kenneth Gillingham, William Nordhaus, David Anthoff, Valentina Bosetti, Haewon McJeon, Geoffrey Blanford, Peter Christenn, John Reilly, Paul Sztorc: Modeling Uncertainty in Climate Change: A Multi-Model Comparison
ET	14.2016	Paolo M. Panteghini, Sergio Vergalli: Accelerated Depreciation, Default Risk and Investment Decisions
ET	15.2016	Jean J. Gabszewicz, Marco A. Marini, Ornella Tarola: Vertical Differentiation and Collusion: Cannibalization or Proliferation?
EIA	16.2016	Enrica De Cian, Ian Sue Wing: Global Energy Demand in a Warming Climate
ESP	17.2016	Niaz Bashiri Behmiri, Matteo Manera, Marcella Nicolini: Understanding Dynamic Conditional Correlations between Commodities Futures Markets
MITP	18.2016	Marinella Davide, Paola Vesco: Alternative Approaches for Rating INDCs: a Comparative Analysis
MITP	19.2016	W. Brock, A. Xepapadeas: Climate Change Policy under Polar Amplification
ET	20.2019	Alberto Pench: A Note on Pollution Regulation With Asymmetric Information
EIA	21.2019	Anil Markandya, Enrica De Cian, Laurent Drouet, Josué M. Polanco-Martìnez, Francesco Bosello: Building Uncertainty into the Adaptation Cost Estimation in Integrated Assessment Models
MITP	22.2016	Laura Diaz Anadon, Erin Baker, Valentina Bosetti, Lara Aleluia Reis: Too Early to Pick Winners: Disagreement across Experts Implies the Need to Diversify R&D Investment
ESP	23.2016	Claudio Morana: Macroeconomic and Financial Effects of Oil Price Shocks: Evidence for the Euro Area
EIA	24.2016	Wei Jin, ZhongXiang Zhang: China's Pursuit of Environmentally Sustainable Development: Harnessing the New Engine of Technological Innovation
EIA	25.2016	Doruk İriş, Alessandro Tavoni: Tipping Points and Loss Aversion in International Environmental Agreements
ET	26.2016	Doruk İriş, Jungmin Lee, Alessandro Tavoni: Delegation and Public Pressure in a Threshold Public Goods Game: Theory and Experimental Evidence
EIA	27.2016	Stefan P. Schleicher, Angela Köppl, Alexander Zeitlberger: Extending the EU Commission's Proposal for a Reform of the EU Emissions Trading System
EIA	28.2016	Tomas Ekvall, Martin Hirschnitz-Garbers, Fabio Eboli, Aleksander Sniegocki: A Systemic Approach to the Development of a Policy Mix for Material Resource Efficiency
EIA	29.2016	Silvia Santato, Jaroslav Mysiak, Carlos Dionisio Pérez-Blanco: The Water Abstraction License Regime in Italy: A Case for Reform?
MITP	30.2016	Carolyn Fischer: Strategic Subsidies for Green Goods
MITP	31.2016	Carolyn Fischer: Environmental Protection for Sale: Strategic Green Industrial Policy and Climate Finance
ET	32.2016	Fabio Sabatini, Francesco Sarracino: Keeping up with the e-Joneses: Do Online Social Networks Raise Social Comparisons?
MITP	33.2016	Aurora D'Aprile: Advances and Slowdowns in Carbon Capture and Storage Technology Development
EIA	34.2016	Francesco Bosello, Marinella Davide, Isabella Alloisio: Economic Implications of EU Mitigation Policies: Domestic and International Effects
MITP	35.2016	Shouro Dasgupta, Enrica De Cian, and Elena Verdolini: The Political Economy of Energy Innovation

MITP	36.2016	Roberta Distante, Elena Verdolini, Massimo Tavoni: Distributional and Welfare Impacts of Renewable Subsidies in Italy
MITP	37.2016	Loic Berger, Valentina Bosetti: Ellsberg Re-revisited: An Experiment Disentangling Model Uncertainty and Risk Aversion
EIA	38.2016	Valentina Giannini, Alessio Bellucci, Silvia Torresan: Sharing Skills and Needs between Providers and Users of Climate Information to Create Climate Services: Lessons from the Northern Adriatic Case Study
EIA	39.2016	Andrea Bigano, Aleksander Śniegocki, Jacopo Zotti: Policies for a more Dematerialized EU Economy. Theoretical Underpinnings, Political Context and Expected Feasibility
ET	40.2016	Henry Tulkens: COP 21 and Economic Theory: Taking Stock
MITP	41.2016	Shouro Dasgupta, Enrica De Cian: Institutions and the Environment: Existing Evidence and Future Directions
MITP	42.2016	Johannes Emmerling, Laurent Drouet, Lara Aleluia Reis, Michela Bevione, Loic Berger, Valentina Bosetti, Samuel Carrara, Enrica De Cian, Gauthier De Maere D'Aertrycke, Tom Longden, Maurizio Malpede, Giacomo Marangoni, Fabio Sferra, Massimo Tavoni, Jan Witajewski-Baltvilks, Petr Havlik: The WITCH 2016 Model - Documentation and Implementation of the Shared Socioeconomic Pathways
MITP	43.2016	Stefano Carattini, Alessandro Tavoni: How Green are Economists?
ET	44.2016	Marco Di Cintio, Sucharita Ghosh, Emanuele Grassi: Firm Employment Growth, R&D Expenditures and Exports
ESP	45.2016	Nicola Cantore, Patrick Nussbaumer, Max Wei, Daniel Kammen: Energy Efficiency in Africa: A Framework to Evaluate Employment Generation and Cost-effectiveness
MITP	46.2016	Erin Baker, Valentina Bosetti, Ahti Salo: Finding Common Ground when Experts Disagree: Belief Dominance over Portfolios of Alternatives
MITP	47.2016	Elena Verdolini, Laura Diaz Anadón, Erin Baker, Valentina Bosetti, Lara Aleluia Reis: The Future Prospects of Energy Technologies: Insights from Expert Elicitations
MITP	48.2016	Francesco Vona, Giovanni Marin, Davide Consoli: Measures, Drivers and Effects of Green Employment: Evidence from US Local Labor Markets, 2006-2014
ET	49.2016	Thomas Longden: The Regularity and Irregularity of Travel: an Analysis of the Consistency of Travel Times Associated with Subsistence, Maintenance and Discretionary Activities
MITP	50.2016	Dipak Dasgupta, Etienne Espagne, Jean-Charles Hourcade, Irving Minzer, Seyni Nafo, Baptiste Perissin-Fabert, Nick Robins, Alfredo Sirkis: Did the Paris Agreement Plant the Seeds of a Climate Consistent International Financial Regime?
MITP	51.2016	Elena Verdolini, Francesco Vona, David Popp: Bridging the Gap: Do Fast Reacting Fossil Technologies Facilitate Renewable Energy Diffusion?
MITP	52.2016	Johannes Emmerling, Vassiliki Manoussi, Anastasios Xepapadeas: Climate Engineering under Deep Uncertainty and Heterogeneity
MITP	53.2016	Matthew R. Sisco, Valentina Bosetti, Elke U. Weber: Do Extreme Weather Events Generate Attention to Climate Change?
MITP	54.2016	David Anthoff, Johannes Emmerling: Inequality and the Social Cost of Carbon
MITP	55.2016	Matthew Adler, David Anthoff, Valentina Bosetti, Greg Garner, Klaus Keller, Nicolas Treich: Priority for the Worse Off and the Social Cost of Carbon
EIA	56.2016	Luca Di Corato, Michele Moretto, Sergio Vergalli: Deforestation Rate in the Long-run: the Case of Brazil
ET	57.2016	Carlo Drago: Exploring the Community Structure of Complex Networks
ET	58.2016	Guiomar Martín-Herrán, Santiago J. Rubio: The Strategic Use of Abatement by a Polluting Monopoly
ET	59.2016	Philip Ushchev, Yves Zenou: Price Competition in Product Variety Networks
MITP	60.2016	Marina Bertolini, Chiara D'Alpaos, Michele Moretto: Investing in Photovoltaics: Timing, Plant Sizing and Smart Grids Flexibility
MITP	61.2016	Gregory F. Nemet, Laura Diaz Anadon, Elena Verdolini: Quantifying the Effects of Expert Selection and Elicitation Design on Experts' Confidence in their Judgments about Future Energy Technologies