The impact of expert selection, elicitation design and R&D assumptions on experts' estimates of the future costs of photovoltaics

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

Abstract

Expert elicitation of future energy technology costs can improve energy policy design by explicitly characterizing uncertainty in future costs. However, the recent surge in the use of the expert elicitation methodology raises questions about the reliability and comparability of their results. In this paper, we standardize disparate expert elicitation data from five EU and US studies, involving 66 experts, of the future costs of photovoltaic (PV). We show that in-person elicitations are associated with more optimistic 2030 PV cost estimates but have little impact on the uncertainty range. Expert selection (affiliation type and nationality) also affects results, although these effects are less significant. In contrast with previous research on expert elicitations in nuclear power, EU experts and private sector experts in our sample are more optimistic about future PV costs than their US and non-private sector counterparts. We suggest that this difference may be due to availability heuristics related to the greater deployment of solar in the EU and the direct exposure of private sector experts to declining costs. As expected, higher R&D investment is associated with lower future costs and greater uncertainty about those costs, although with a diminishing effect in both cases.

Key Words: Photovoltaic costs, energy R&D, expert elicitation, survey design, heuristics

* Corresponding Author:
Elena Verdolini
Fondazione Eni Enrico Mattei and CMCC
Corso Magenta 63, 20123 Milano, Italy
elena.verdolini@feem.it

[1] Harvard University
[2] Univ. of Wisconsin-Madison
[3] Fondazione Eni Enrico Mattei and CMCC

1 Introduction

The inherent uncertainty surrounding the future cost and performance of energy technologies has important consequences for energy policy and for decisions about energy infrastructure investment, subsidy policies, mitigating health outcomes, and the like. To meet the world's energy challenges, policy makers will have to accelerate innovation in low-carbon and efficient energy technologies. Hence, the interest regarding what governments can do in this respect is only growing. Public investment in research and development (R&D) is an important tool that governments have at their disposal (Cohen and Noll, 1991). However, retrospective analyses show how difficult it is to predict the future of the energy system and to assess the impact of public investment on energy costs (Craig et al. 2002). A key difficulty is designing R&D policies that are robust to the uncertainty around the impact of such policies on future energy costs. Hence, knowledge about the range of possible cost outcomes and their associated probabilities greatly helps in informing policy makers and in evaluating the effectiveness and robustness of potential energy policies.

Expert elicitations have been increasingly used to fill this gap and gather experts' opinion on the range of possible future energy costs for a number of reasons. First, they allow analysts to account for the fact that the future may not look like the past, in particular in technology innovation, i.e., that learning curves (e.g. Junginger et al. 2005, Söderholm and Klaassen 2007) and factor decomposition (Nemet 2006, McNerney et al. 2011) may not be good predictors of future change. Second, experts have information that may not be available elsewhere due to their deep knowledge of the technologies. Finally, expert elicitations allow an explicit characterization of uncertainty, namely they can provide not only a range of possible outcomes but also their associated probabilities. Indeed, a 2007 report from the National Research Council recommends that the U.S. Department of Energy begin to use expert elicitation for their RD&D allocation decisions, to explicitly characterize probabilistic estimates of the outcomes of RD&D investments (NRC 2007).

The recent surge in the use of expert elicitation to collect probabilistic information on future energy technology costs raises however the fundamental issue of comparing the elicited metric across different studies and different technological options. In this paper, we collect and standardize data from these different elicitations. We focus on expert elicitations on solar power from photovoltaics (PV), a promising low carbon energy technology which involves no fuel costs, minimal operating costs, and the potential for very low manufacturing costs.[4] Due to these characteristics and potential, governments around the world have prioritized making solar PV competitive with other fossil power generation alternatives.[5] Since 2007, five research groups in the United States and Europe have conducted probabilistic expert elicitations on future costs of solar PV with the aim of informing the communities of policy makers and energy modelers (see Table 1). Using a variety of methods, experts, and policy scenarios, these groups have separately gathered a wealth of data regarding the expected impact of R&D investments on future costs ranges and their associated probabilities.

The standardization and data analysis we carry out is beneficial in many respects. First, it improves the knowledge about possible future technological pathways and performance.

---

[4] Crystalline silicon solar module costs came down by a factor of 70 between 1970 and 2010 and by almost a factor of 2 between 2010 and 2013.

[5] Countries that have most prominently supported solar PV are, for example, the United States (Sunshot Program), the European Union (SET-Plan), Germany (Erneubare-Energien-Gesetz), Japan (2012 feed-in-tariff for renewable energy), and China (the National Energy Administration's goal of incentivizing 14 GW in 2014).

Elicitation studies are expensive and time consuming to conduct, for analysts as well as subjects. Hence, the number of experts involved is generally low and the technologies included in each survey may be a selection of the possible technological paths that researchers could consider. Yet the lack of comparability means that policy makers and analysts often use information from just one study, and therefore do not benefit from the whole set of information available. For instance, grouped studies would arguably provide more reliable estimates, simply by including a larger sample of the population of experts. Our first contribution is therefore to collect data from all the most recent probabilistic expert elicitations on PV and standardize it to make it comparable conditional on a set of clearly specified assumptions. This arguably represents the most consistent and comprehensive representation of expert opinion on the influence of R&D funding on future PV costs. As such, it has potential to inform the research and the policy communities beyond the analysis we carry out in this paper.

Second, the collected and standardized data are used to inform and improve expert elicitation methods and outcomes. Since elicitation protocols can vary in design, collecting data from different survey is a starting point to study whether differences in protocol design and expert sampling have an impact on the elicited metrics. With this analysis, we further contribute to the science of better understanding how to design expert elicitations by exploring whether various characteristics of survey design, such as elicitation method and expert selection, are statistically significant predictors of central estimates and uncertainty ranges. A previous effort along this lines was carried out in Anadon et al. (2013) for nuclear technologies. Here, we extend the analysis to the case of solar PV and present insights which can be used to (1) complement with quantitative contributions the qualitative prescriptions on optimal elicitation protocol design and expert selection (O'Hagan et al. 2006) and (2) test whether the results for nuclear technologies carry over to the case of solar PV and whether, in fact, they can be considered generally applicable.

Finally, as in Anadon et al. (2013), we quantify the average expected impact of government research, development, and demonstration (RD&D) investments on expected PV costs in 2030 the by pooling data from a larger number of experts. These are helpful insights for decision makers, who confront a wide variety of decisions regarding which types of policies to implement, the level of public investment, the timing of policies, but also how to prioritize among technological pathways within solar PV as well as between solar PV and other energy technologies.

This paper is thus motivated by the potential to improve the design of future elicitations by exploring how elicitation design choices affect the elicited outcomes, as well as policy decisions by allowing the use of a larger set of elicited technology costs.

The rest of the paper is organized as follows. Section 2 describes the data, summarizes the standardization procedure and presents the meta-regression set up, including details on the dependent and independent variables of interest. Section 3 presents the empirical results and Section 4 concludes with relevant research and policy implications emerging from this study. The supplementary information (SI) contains further details and results.

2 Materials and Methods

2.1 Description of expert elicitations

We use individual participant data from 5 expert elicitations conducted between 2007 and 2011 on the future costs of solar PV. Note that we only include probabilistic expert elicitations in our exercise, namely those that ask experts to assess a range of percentiles. We intentionally exclude other types of forecasts, such as central estimates and ranges with no probabilities attached, for 3 reasons: 1) those estimates do not typically conduct assessments conditional on both BAU and non-BAU R&D expenditures, 2) they do not include a process of de-biasing which is central to the expert elicitation methodology, and 3) they cannot be used to explore the effect of protocol and expert characteristics on different points of the cost distribution. Table 1 summarizes the main characteristics of each study. For a more detailed description of each elicitation study, please refer to the original articles. Here, we summarize the aspects of the studies which are relevant for the current analysis.

The 5 elicitations provide variation along several dimensions. In particular, three of the elicitations were conducted in person, and two were conducted online. Three of the elicitations were published in the peer-reviewed literature and two of the elicitations were published as reports. Four of the expert elicitations were based in the United States and consulted U.S. experts, and one of them consulted European experts. All elicitations but one included experts from academia, the private sector, and public institutions. It is this heterogeneity which allows us to explore the effects of survey design and expert selection features, as well as various public R&D funding scenarios, on the elicited PV costs in the sample. The external validity of these results depends on how one interprets our sample. While we include all structured elicitations of the cost of PV conducted in the past 5 years, we include only 66 experts, which is presumably a subset of an unknown population of people who are experts on the future of PV and adept at thinking in terms of probabilities and policy conditions. That population may not be much larger than 66, but we are conservative in our conclusions and use them mainly to inform future research (Cooke, et al. 2014).

In addition, each protocol focused on different aspects of future solar PV costs. UMass asked questions about the probability that specific technical and cost goals would be met by 2050, while all other surveys asked experts to provide the 10^{th}, 50^{th}, and 90^{th} percentile cost and performance estimates by 2030. Moreover, the FEEM elicitation focused on levelized cost of electricity (LCOE) given specific insolation and discount rate assumptions, which were provided to the experts and on which the experts agreed. Conversely, the Harvard study asked about several components of solar costs: module capital cost, module efficiency, module lifetime, inverter cost, inverter efficiency, inverter lifetime, other materials cost, installation labor, overhead cost, operation and maintenance expense. Finally, CMU and UMass focused on module costs and conversion efficiency and NearZero focused only on module costs.

Elicitations also differ in their assumptions about future public R&D investment. The UMass, FEEM, and Harvard studies explicitly asked questions conditional on three different public R&D investment levels. The NearZero elicitation asked experts to provide cost estimates consistent with a business-as-usual (BAU) global R&D funding scenario. Finally, CMU asked questions under a BAU R&D funding level and a much larger funding level (10 times the BAU level) coupled with two different solar deployment scenarios.

[Table 1 around here]

2.2 Standardization of expert elicitation data

The first issue confronting us with respect to these different studies was the fact that estimates were not directly comparable. As explained above, the different groups focused on different metrics, with some groups eliciting the LCOE while other focusing on the different components of costs (such as module capital cost, module efficiency, etc.) or different times in the future. The first step we undertook was to convert estimates into a consistent metric—to standardize estimates.

Due to the lack of details on specific cost components for the FEEM study we chose LCOE in $2010 as the metric of interest. FEEM had asked experts to make LCOE estimates assuming particular values of key metrics like discount rate and solar insolation. In the standardization process, we used individual cost components provided by experts in the Harvard, CMU, and NearZero, along with the FEEM assumptions for discount rate and solar insolation to calculate the LCOE. UMass researchers converted experts' estimates of module cost for 2050 to LCOE estimates for 2030 relying on the same assumptions about the impact of time on technological change, about the discount rate and about the insolation assumed by the FEEM team (as described in Baker et al. 2014). We then further standardized the elicited data from the different studies into the $2010 LCOE, as described in detail in the SI.

Since all studies focused on probabilistic cost elicitation, the standardization process allows to compare a number of different cost estimates. The first is the 50th percentile, or central estimate, provided by experts (P50). The second is the 10^{th} percentile cost estimate (P10), which we interpret as the value of LCOE associated with a "best-case scenario", or breakthrough technology development. Third, the 90^{th} percentile cost estimate (P90) is the highest cost estimate, and can be thought of as the "worst case scenario" in terms of future technology performance. Finally, measures of the range of uncertainty (Urange) associated with such cost estimates can also be computed. We focus here on a measure of experts' "normalized" uncertainty around future costs, which is defined as the difference between the 10^{th} and 90^{th} percentile of each expert's estimate divided by their most likely expected cost (P50). Hence, Urange=(P90-P10)/P50.

In addition to converting the cost variables, which are the dependent variables in this study, we also coded a number of other key variables for each of the study in the sample. First, as already mentioned, most studies elicited cost data under clearly defined R&D investment scenarios, which differ across studies (see Table 1). Asking experts to provide costs estimates under varying R&D scenario is instrumental in understanding the impact of public research investment on technologies future performance. The different R&D scenarios can be easily compared in absolute value across studies, but focusing on exact assumed budget amounts could be misleading for two main reasons. First, 40% of the observations in our sample are not associated with dollar amounts of R&D spending due to the way in which the elicitation protocol was designed. Second, even experts typically rely on heuristics when making estimates (Kahneman, 2011). Even though each study provided experts with detailed background information on historical levels of public R&D, they still may find some difficulty in thinking about specific investment levels, and instead may use these levels to think about the outcomes of worst-case and best-case investment scenarios. If this is indeed the case, then the cost estimates for different R&D levels would not necessarily reflect full range of R&D levels and their effects. We therefore chose to focus on a categorical definition of the R&D investments, and coded R&D amounts into three bins indicating “Low”, “Medium”, and “High” investment. Such binning of R&D values might be a closer representation of the experts’ thinking than the actual levels they were basing their estimates upon. The details on the categorization of R&D investment into the different bins are included in the SI, Table S3. We then explore the robustness of results with the use of the continuous R&D investment variable

The process of standardization allows comparing the insights on the relationship between R&D investments and elicited costs in the different studies in a qualitative manner. Figure 1 shows the full range of elicitation results used in the analysis grouped under the low, mid and high R&D investment scenarios. The three panels in Figure 1 have different numbers of experts, with the “low” level being the most populated one and the “mid” being the least populated one. This is due to the fact that (a) the CMU elicitation only populated the low and high R&D categories; and (b) the NearZero elicitation only covered a low (or BAU) R&D investment level.

[Figure 1 around here]

Figure 1 shows that including a greater number of observations generally results in a greater range of outcomes, with the range of P50 estimates decreasing from the low, to the high, to the mid R&D scenarios, partly due to the decreasing number of experts and studies. Variation between P50 estimates and P10 and P90 estimates are different across experts. Ordering the observations by increasing P50 does not result in P10 and P90 observations that are also ordered. This means that the 10^{th} and 90^{th} percentile estimates of each expert are to a large extent expert specific and suggests that any quantitative analysis should include expert fixed effects to appropriately control for variation among experts.

Moreover, in the highest R&D scenario the P50 curve is shifted downward and it has a smaller slope as compared to the low R&D scenario. This indicates that experts believe that increased R&D in will decrease PV technology costs. Finally, the impact of higher R&D investment on the range of expected costs for each expert (namely, the difference between the 90^{th} and the 10^{th} percentile) is less clear in this descriptive framework.

This qualitative data analysis does not however take into account that variables other than the different R&D scenarios might be affecting cost estimates. Specific choices in the design of the elicitation protocol or the selection of an “optimist” group of expert, are recognized as resulting in estimates that are biased upward or downward (see O’Hagan et al. 2006 and Meyer and Booker 2001). Such potential biases are worth investigating to ensure that insights from expert elicitations feed into efficient and cost-effective energy policy. Anadon et al. (2013) focused on nuclear technologies. Here, the analysis and results are extended to the case of solar PV. The higher diversity in elicitation design in the case of PV allows to focus on some elicitation design characteristics that were only marginally considered in the nuclear case due to the higher homogeneity of the sample. Moreover, it is not clear to what extent the results presented in Anadon et al. (2013) are technology specific. We shed light on this through the analysis presented in the next Section.

2.3 Empirical approach

The surveys included in our sample provide information on other three categories of variables that might affect expert's costs and uncertainty ranges in addition to R&D scenarios, as suggested by the literature on expert elicitation design (O'Hagan et al. 2006 and Meyer and Booker 2001). The first category, *elicitation design variables*, includes three variables of interest: in person (denoting an elicitation that was conducted in person) vs. online (denoting an elicitation that relied on an online tool); published (denoting an elicitation that was published in a peer-reviewed journal) vs. unpublished (denoting an elicitation that was reported in a non-peer reviewed journal); and the year an estimate was made (the year in which the expert elicitation was conducted). Even though previous studies have looked at the differences in the design of elicitation protocols, highlighting in particular the importance of the expert selection phase (Raiffa 1968; Keeney and Winterfeldt 1991; Meyer and Booker 1991; Phillips 1999; Clemen and Reilly 2001; Walls and Quigley 2001), to our knowledge the impact of these variables has been evaluated empirically only in Anadon et al. (2013) with a focus on nuclear technologies.

The second relevant category of explanatory variable defines the *market in which the technology competes* (residential, commercial, utility). Electricity from solar PV generally competes with electricity produced by other sources, sometimes known as "grid" electricity. But the price of grid electricity varies considerably depending on who is buying it, whether retail or commercial customers. Similarly, the scale of production, and thus costs, can differ considerably whether at the single-digit kilowatt scale of residences, tens of kilowatts for commercial installations, and even thousands of kilowatts for utility scale. In the SI we show a cross-tab of technology type and market to show that, although the categories are not well balanced, the two variables are not well correlated.

The third category encompasses *expert selection variables*. Studies suggest that selecting a diverse pool of experts can help avoiding anchoring to a usually conservative reference point (Meyer and Booker 2001). There are two relevant dimensions of expert background that can be coded for the elicitation in our sample. US denotes experts based in the United States, as opposed to experts based in the EU. *Academia* denotes experts working in universities, compared with *private* sector denoting experts working companies, and *public*, denoting experts working in public institutions, such as national laboratories and regulatory bodies.

Table 2 summarizes the descriptive statistics of the dataset, which include 310 observations from 66 individual experts and provides information on the availability of data in the different studies. Since not all studies specified R&D amounts or market segments, observations are lower for those variables. Considering the different R&D investment categories, 54% of the estimates are conditional on low investment levels, 17% on medium levels, and 29% on high levels. The NearZero study does not specify R&D levels (although it is largely consistent with a BAU R&D level). In terms of expert affiliation: 33% of the estimates come from experts in academia, 31% from experts in the private sector, and 36% from experts in public institutions. Only 10% of the experts were based in the European Union; the rest were from the U.S. Only 2 out of five studies ask experts to provide cost estimates conditional on specific deployment scenarios (some CMU and NearZero observations). Finally, market characteristics can be coded only for a subset of studies (namely, FEEM, Harvard and UMass).

[Table 2 around here]

To provide more quantitative insights on the factors affecting cost estimates we apply a meta-analysis approach to the standardized data. We regress each of the cost variables of interest (namely, P50, P10 and Urange) on the set of control variables described above. In this way, we study the relationship between costs estimates and the three other groups of variables using individual data points and in a multivariate setting. We thus isolate two different components of the cost reductions implicit in each expert's estimate: (1) the variation arising from differences in experts' views about future technological improvement and (2) the variation arising from differences in expert and survey characteristics.

With respect to the *R&D investment levels* we first use dummy variables to indicate three generalized R&D scenarios (low, medium, high) as explained above to assess the extent to which experts expect reductions in future costs as a consequence of higher public R&D investment. Second, we use annual investment levels in constant dollar amounts to check whether the definition of the R&D variable has an impact on the empirical results. Third, we add a squared term to the annual investment levels to account for the possibility of decreasing marginal returns to R&D investments. In this framework, the coefficient associated with the R&D variable describes the average effect *ceteris paribus* of R&D spending on different measures of costs across the whole sample. Conversely, the coefficients associated with variables describing expert and survey characteristics indicate whether expert selection or specific studies design affect the estimates in any particular direction.

Our base specification reads as follows:

$$\ln(Y) = \alpha + \boldsymbol{S}\,\beta + \mathbf{T}\,\gamma + \boldsymbol{E}\,\delta + \boldsymbol{R}\,\theta + \vartheta_i + \varepsilon$$

Where Y is either P50, P10, P90, Urange or future costs normalized by current (2010) costs (normP50). The vector of study characteristics (S) includes a dummy variable equal to one if the survey was conducted in person and the year in which the elicitation was carried out. The vector of technology variables (T) includes dummy variables for type of technology or alternatively market segment. The vector of expert characteristics (E) includes dummy variables indicating the expert was from academia or the public sector, with private sector being the reference category. It also includes a dummy variable indicating whether the expert worked in the European Union, with US experts being the default category. The vector of R&D variables (R) includes either the dollar amount and its square or two dummy variables indicating medium and high funding, with business-as-usual R&D funding (low) being the reference. ϑ_i are experts' dummies in our fixed effect framework.

When performing our empirical analysis we confront three issues. First, given that our sample is composed of repeated observations for the same subjects, we need to control for the serial correlation in the error term. We can address this issue and control for the heterogeneity across different experts in two ways, namely using a fixed effects or a random effects model. We choose the former under the assumption that the experts' characteristics which are not observable are constant within experts across the different observations. This is the most reasonable assumption given that all elicited cost values were collected from various experts within a very short span of time (a few hours).

Second, due to the collinearity between published and in person studies (as apparent in Table 1) we focus on the in person variable. The motivations behind this choice are that: (a) as already pointed out, the mode of administering a survey is of great interest for researchers involved in expert elicitations (O'Hagan 2006), who are continuously facing the trade-off between in person interviews, which increase the interaction between experts and researchers, and well-designed online protocols, which allow researchers to more cost-effectively reach a wider pool of experts; and (b) conversations with researchers involved in the unpublished elicitations suggest that the collinearity between published and in person is due to spurious correlation: the reasons for not seeking publication in the peer-reviewed literature are unrelated to characteristics of the estimates, unlike what has been observed in the health field (McGauran et al. 2010).

Third, given that certain variables are only available for a subset of studies, we estimate models on samples of different sizes and confine non-core results to the SI. Our main specification includes 211 observations: this sample includes all observations with binned R&D variables excluding those with specific deployment scenarios (part of CMU's data and all of NearZero's). The models using the continuous R&D variable further exclude the remainder of the CMU observations, resulting in 141 observations. An additional model including all available data but controlling for the presence of assumptions on deployment is presented in the SI. Furthermore, we present results on the impact of technology variables in the SI.

3 Discussion of results

Here we present our findings on the effects of expert selection, elicitation design and R&D investment on the different cost variables of interest.

Table 3 focuses on P50. The differences between Models are in the way in which the R&D variable is coded. Models 1-3 include binned R&D variable, with Model 1 representing our base regression. Model 2 excludes fixed effects and is presented for comparison purposes. Model 3 uses normP50, the normalized P50, as the dependent variable. Models 4 and 5 use R&D levels (hence the smaller sample size), with Model 5 including a square term to explore decreasing marginal returns to R&D.

Table 4 includes results for P10 (Models 1-3) and for P90 (Model 4). Model 1 is our base regression, relying on binned R&D variables. Models 2 and 3 use R&D levels, without and with an additional square term, respectively. Model 4 is on P90.

Table 5 presents results for Urange. Models 1 and 2 include the results using binned R&D variables—Model 1 with, and Model 2 without, expert fixed effects. Models 3 and 4 use R&D levels in dollars, the latter adding a squared R&D term.

In the SI we include additional results including technology and market variables and deployment for P50 and Urange.

Overall, we find that the inclusion of expert fixed effects, survey design, expert selection, and R&D investment level variables accounts for about two thirds of the observed variation in expert estimates of P50 and for about three quarters of the variation in Urange. Summarizing our results, we show that in-person elicitations are associated with more optimistic 2030 PV cost estimates but have little impact on the uncertainty range. Expert selection (affiliation type and nationality) also affects results, although these effects are less significant. In contrast with previous research on expert elicitations in nuclear power, EU experts and private sector experts in our sample are more optimistic about future PV costs than their US and non-private sector counterparts. Finally, as expected, higher R&D investment is associated with lower future costs and greater uncertainty about those costs, although with a diminishing effect in both cases. We discuss and comment these results in detail below.

[Table 3 around here]

[Table 4 around here]

[Table 5 around here]

3.1 Effects of expert selection

Relationship between expert selection and central estimates

Expert selection has mixed effects on experts' central estimates (P50, Table 3). The coefficient associated with the EU dummy variable is negative in all specifications, suggesting that these experts are more optimistic than their U.S. counterparts as on average their elicited cost values are lower. The p-value reaches however acceptable levels of significance only in Model 4 which uses the continuous R&D variable and assumes diminishing marginal returns to investment, while it is just higher than 0.10 in Models 1 and 3 using binned variables, making the coefficient slightly insignificant. Similar results are presented also in the additional models in the SI.

These results are in contrast with those of Anadon et al. (2013) for nuclear technologies, which indicated that U.S. experts were consistently more optimistic than EU experts on central estimates of future nuclear costs. Evidence provided here thus suggests that depending on the energy technology under consideration, geographical differences might be associated with more optimistic or more pessimistic estimates. Even if not strongly supported in our models, a higher confidence of EU experts in solar technologies could indeed be plausible for a number of reasons. For instance, during the years the elicitations were being conducted (2007-11), governments in Europe subsidized the adoption of solar power much more intensively than did governments in the United States. Hence, solar PV deployment was dramatically different in the two regions. While in 2000 cumulative solar TWh installed were comparable, by 2012 the EU had surpassed the U.S. by more than an order of magnitude (BP 2013). Experts may have been influenced by the growth of the PV industry in their local markets, and thus the availability heuristics in both regions would differ (Tversky and Kahneman, 1974; Kahneman, 2011). Another possible explanation is that EU experts were the only ones that were asked about levelized cost of electricity (LCOE) directly, while the U.S. experts were asked about other variables than were used to calculate LCOE *ex post*. It is possible that the conversion process introduced a bias that made the calculations of LCOEs from U.S. expert estimates more pessimistic—namely, if U.S. experts had been asked about LCOE they may have given different and perhaps more optimistic estimates than those we obtained from the standardization process. It is also possible that asking the separate questions on the core components of a given technology pushes experts to think more carefully and more conservatively about future technology performance (Morgan 2014).

The effects of variables characterizing the background of each expert are less consistent across model specifications. Academic and public sector dummies are associated with positive coefficients (hence, higher P50 cost estimates) in Models 1-3, Table 3. This suggests that private sector experts are more optimistic about future solar PV costs than academic ones, and especially more optimistic than public sector experts. However, the coefficient is statistically significant only in the case of public sector experts.

These insights are again at odds with those in Anadon et al. (2013) for nuclear technologies, in which private sector experts emerge as the most pessimistic. One possible explanation of this difference is that industry experts are more familiar with recent construction than are public sector and academic experts. In nuclear, industry experience in 2007-11 would have created a heightened awareness of the recent challenges, delays, and escalating costs, whereas in solar, industry experience would have heightened awareness of rapidly falling costs and expanding markets, partly as a result of the greater public acceptance of solar PV. Thus, the availability heuristics (Kahneman, 2011) of private experts would have been different than that of other experts.

Note however that the pessimism of public sector experts is reversed when one third of the elicitation results are dropped due to missing R&D levels (Models 4 and 5, Table 3). This change is in fact entirely attributable to the shrinking in sample size, and not to the different coding of the R&D variables (binned versus continuous), as shown in the SI.

Relationship between expert selection and low cost outcomes (P10)

Focusing on the low cost outcomes, EU experts are consistently more optimistic than their U.S. counterparts. Results from Models 1-3 in Table 4 suggest that EU estimates are around one fifth lower and statistically significant in all models.

Regarding experts' backgrounds, we also find similar results to those presented above for P10. In Model 1, both academics and public sector experts emerge as more pessimistic about 10th percentile costs than private experts. The coefficients associated with these two variables are positive, larger than for P50 and significant, hence indicating that elicited costs for these experts are on average higher. However, as in the P50 case, coefficients change sign when performing the analysis on the reduced sample (Models 2-3), while still remaining significant.

Model 4 shows that, unlike the P50 and P10 results, expert background does not affect P90 estimates. The coefficients associated with the dummy variables for EU, academia, and public are not statistically different from zero. This might be because expert background plays a smaller role for pessimistic (high-cost) outcomes because there is more consensus about them. PV experts may agree that the worst-case scenario for future solar costs are current costs. This consensus is in contrast to nuclear power, which has experienced costs escalating over time in industrialized countries.

Relationship between expert selection and the uncertainty range

We find that expert selection has a much higher impact on Urange, namely the confidence of experts in their responses (Table 5). This is particularly true for variables indicating experts' background. Indeed, results seem to point to European experts being less confident than their

US counterparts. This is consistent across model specifications, but it is only significant in Model 4 of Table 5, in which the quadratic term is included. Additional specifications in the SI are in line with what is discussed above: the sign of this EU coefficient remains consistent, but not statistically significant, with the exception of the model that controls for deployment.

When using the categorical R&D variables, academic and public sector experts are associated with significantly smaller uncertainty ranges than private sector experts, meaning that they are more confident in their estimates. This effect is however reversed in Models 4 and 5, in which one third of the observations are dropped because they are missing continuous R&D expenditure variables. Once again, the switch in coefficients is not attributable to the different coding of the R&D variable, rather to the drop in observations between the two models. Consistent results with those presented here for both the larger and the smaller sample are presented in the SI.

3.2 Effects of study characteristics

Relationship between survey design and central estimates

Elicitations conducted in-person were associated with more optimistic responses about central estimates (P50) than elicitations conducted online. This effect is consistent and highly significant across all five specifications in Table 3. It is also robust to the alternative specifications shown in the SI. In the base specification (Model 1), experts interviewed in person gave average P50 estimates that were roughly between 50 percent and 70 percent lower than those gathered on-line. The SI includes results in which we drop the experts from the UMASS study. In-person remains negative and significant. One caveat to this interpretation is that since this variable was collinear with the published variable this result may also capture the differences in estimates between published and unpublished studies. However, as mentioned above, our discussions with the authors does not give us any specific reason to believe that the decision to not publish a study was based on the level of the included estimates. Hence, the collinearity between these two variables seems to be due to spurious correlation. Finally, we find that the year in which an elicitation was conducted has no consistent effect in expected future costs. This result is noteworthy considering that between 2007 and 2011, when the elicitations were carried out, solar panel costs were decreasing dramatically.

Relationship between survey design and low and high cost outcomes (P10 and P90)

As in the case of P50, in person interviews are also associated with consistently lower P10 and P90 estimates, but the effect of elicitation year is not significant.

Relationship between survey design and the uncertainty range

Study design has a less significant effect on Urange estimates. Year of elicitation is negative and significant in the specification using the binned R&D variables and based on the larger sample, suggesting that as time passes, elicitations include estimates whose range of uncertainty is narrower. The coefficient however switches sign in the models based on the smaller sample. The effect of in-person on uncertainty range is negative, but insignificant, when analyzing the smaller sample, either in Table 5 in the Text or in the specifications presented in the SI. Its sign changes when the studies without R&D dollar levels are dropped.

3.3 Technology and market characteristics

We explore the role of technology characteristics on cost estimates and uncertainty ranges and present the results in the SI. Controlling for the use of solar power in a residential, commercial, or utility scale context had a statistically significant impact on future P50 LCOE costs, with utility scale solar power expected to be around 40 percent cheaper than residential scale solar power. This difference is roughly in line with the current difference between wholesale and retail power purchase prices at midday when solar would be used. Conversely, market characteristics do not have significant effects on uncertainty range.

Some studies also accounted for differences in the types of PV on which the experts were to make predictions. We explore this in the SI by adding to the regression binary variables for advanced PV designs. This generally does not change the effects of expert characteristics, survey design, or R&D on central estimates. We observe similar lack of effect for Urange, although, in that case, adding sub-technology characteristics does remove the significance of expert affiliation (Model 2 Table S6).

3.4 Effects of R&D

Relationship between R&D scenarios and the central estimate

R&D investment has a consistently significant effect on median costs: the higher R&D investment, the lower the cost estimates (Models 1-4). These results are robust to alternative specifications included in the SI. Specifically, the medium (high) R&D scenario is associated with expected costs that are on average roughly 35 (40) percent lower than the low funding scenario (Model 1). This suggests that increasing R&D funding from low to mid has a greater impact on costs than increasing R&D funding from mid to high, implying some sort of diminishing marginal returns to R&D investment. In many in-person interviews and written submissions, experts seemed quite aware of the potential for decreasing returns at higher R&D levels, especially due to the availability of trained scientists and engineers, as well as problems that might not be resolvable in the laboratory, such as grid congestion and intermittence. Using the continuous R&D variable, results suggest that a 1% increase in investment lowers expected cost by 0.26%. The hypothesis of diminishing marginal returns is however not supported by the model using a continuous R&D variable and including a squared term, since the coefficient on the square term is positive but not statistically significant.

Relationship between R&D scenario and low and high cost outcomes (P10 and P90)

Higher R&D investment not only affects the median outcome, but also the probability of breakthrough, as measured by the P10 estimates (Table 4). The effects of R&D on the lowest cost outcomes (P10) are similar to those for P50, although the effects are slightly larger. Also in this case, diminishing marginal returns are not confirmed when using the continuous R&D variable. In the case of P90, the effect is also strongly significant, but of a smaller magnitude when compared to P50 and P10. This suggests that R&D has an impact on the whole distribution of costs; it shifts the distribution of experts' predictions lower but also expands it.

Relationship between R&D scenario and the uncertainty range

Higher R&D generally has a positive coefficient in the Urange specifications, meaning that the range of uncertainty increases in the higher R&D scenarios. Hence, increasing the level of R&D with which experts are confronted in the elicitation reduces their confidence (i.e. increases

Urange). This is confirmed in most specifications, both on the small sample or those presented in the SI. This could be due to a number of reasons. For example, medium and high R&D scenarios might mean that funding is also devoted to sub-technologies, which are newer and/or more risky, or that higher total investment allows for inclusion of more of the riskier R&D, resulting in an increase in the uncertainty around future central estimates. An alternative explanation is that this could also result from experts facing significantly different (higher) R&D scenarios from the business-as-usual might have more difficulty in fully projecting costs. The effect of the medium R&D scenario is larger than that of high R&D in almost every case. One can see in Table 4 that, relative to high R&D, medium R&D has a comparatively larger effect on P10 than it does on P90 (using the difference in coefficients in model 4 and coefficients in model 1 for High and Mid R&D, respectively). The breadth of technological pathways available in High R&D may improve outcomes in the high cost outcome, thus reducing the Urange.

In the SI, we present additional specifications as robustness checks. First, we added P50 as an independent variable to explain Urange. The associated coefficient is negative and strongly significant, suggesting that a lower, more optimistic, median elicitation, is associated with a greater normalized uncertainty range (with less confidence). Note that this P50 effect reduces some of the effects of R&D. Some of this effect is difficult to separate since we know that R&D is reducing P50. But one possible interpretation is that R&D is shifting the entire distribution to lower costs; once that effect is accounted for with P50 as an independent variable, the R&D effects on Urange are quite similar.

Additional explanatory variables

Furthermore, in the SI we present models that add a “deployment” variable to the regressors and include all Near Zero and CMU observations in the sample. In these studies experts were instructed to make assumptions about how much PV is deployed and to consider how this would impact costs. This allowed them to account for learning by doing and scale effects in their estimates, which makes them difficult to compare to the other studies. As a first look we find results that are generally intuitive; deployment: (1) has a negative effect on P50 costs, (2) has a negative effect on Urange, and (3) does not seem to reduce the effect of R&D in P50. These results are strongly significant and suggest there would be value in future elicitations on R&D that control for scale and learning by doing, which empirical studies have shown to be important but difficult to disentangle from R&D.

4 Conclusion and Policy Implications

A number of probabilistic expert elicitations for solar PV technologies were carried out by researchers both in the U.S. and in Europe between 2007 and 2011. These studies differ in survey and expert characteristics, in the sub-technologies considered, and in the level of R&D investment with which the experts are confronted. In this paper, we collect, standardize and analysis individual expert data from such expert elicitations. We contribute to the literature by (1) providing standardized estimates of future PV costs for 65 experts that could be used as inputs to support policy decision; (2) measuring the impact of survey protocol and expert selection in solar PV elicitation outcomes and how these differ by technology by comparing them with results in nuclear in Anadon et al. (2013); and (3) estimating the average impact of R&D on future solar PV costs.

Our results have implications for the design of future elicitations, for understanding the effects of R&D, and perhaps for understanding the relationship between an expert's context its impact on future costs estimates. The merits of this approach can be highlighted in at least three respects.

First, standardizing the data and making it more comparable allows to fully exploit the knowledge of a wider sample of experts than any study could reach. This arguably helps in getting a better representation of the full spectrum of estimates with respect on the future development of technology costs and hence in developing cost-effective policies to support PV specifically and renewable energy more in general. The collection of standardized data on elicited future costs we provide can potentially be used to inform both policy makers and researchers.

Second, standardized data can be used to test whether differences in protocol design and expert selection have impact on the elicited costs, as suggested by the expert elicitation literature, and whether such differences are the same across different energy technologies. Using a meta-regression approach we show that choices in protocol design give rise to consistently higher or lower estimates. In the case of PV technologies, in person interviews are associated with more optimistic estimates (lower P50, P10, and P90). This finding seems to contradict some of the qualitative literature on expert elicitation protocol design, which holds that in person interviews allow the researcher to push experts to think about all possible technological bottlenecks (and hence, would likely result in higher elicited costs). This result is also in contrast with the preliminary evidence for in person interviews found in Anadon et al. (2013) in the case of nuclear technologies. We suggest this effect is highly technology-specific and sample-dependent.

For example, in nuclear, the elicitations carried out in person appear to have stimulated the expert to consider unforeseen problems. In contrast, such discussions in solar seem to have encouraged the expert to consider new technology breakthroughs or possibilities for surmounting bottlenecks. The effect of in person interviews is less clear in the case of uncertainty ranges as the direction of the effect depends on which studies are included. Finally, according to our analysis the year in which an elicitation was conducted had no significant impact on future cost estimates.

Along these lines, we provide an indication that an expert's optimism and level of confidence about future solar PV costs also depends on their geographic location and sectoral background, although these effects are not statistically significant across all specifications and their direction is largely dependent on sample size. Results seem to point to EU experts being more optimistic than their U.S. counterparts for PV (but statistically significant only in the case of low-cost outcomes, P10). In some specifications, private sector experts emerge as more optimistic than academic and public sector experts. These findings, stand in contrast with those of Anadon et al. (2013) for nuclear power, in which private experts were associated with higher cost estimates.

That expert selection and survey design matter and that results for solar PV differ from those for nuclear give rise to important implications for researchers and policy maker alike, pointing to the importance of conducting inclusive elicitations when it comes to expert selection and perhaps of using different elicitation approaches to ensure that results truly account for the uncertainty in the field and result in data that are less (upward or downward) biased. The importance of selection might also be informed in future work through differential weighting of experts, for example by their level of expertise on solar or by their skill in making forecasts. We were not able to weigh experts since only a small subset of the elicitation studies we used included indicators of expertise. Future research would benefit from recent work reviewing the advantages of weighting methods (Cooke et al., 2014), as well as concerns about the reliability of choosing weights (Bolger and Rowe, 2014). Finally, those interested in using the results from Anadon et al. (2013) or this paper to debias results of other elicitations in other technologies should be very cautious.

Third, we use the standardized data to study the effects of assumptions about R&D investments on future costs and performance, as well as their impact on the variability of these estimates (uncertainty). We show that R&D investment lowers elicited costs, but that experts have larger uncertainty ranges at higher R&D investment levels, indicating that with more funding the realm of what is possible for experts in terms of solar PV technology development generally increases. The positive impact of R&D affects both central estimates (P50) and extreme cost estimates (P10 and P90, respectively). This finding is robust and statistically significant across all specifications, suggesting that our results could be used to model the average impact of R&D on future costs based on insights from all solar PV elicitations available to date. Using the coefficients associated with the R&D variables in the case of P10, P50 and P90 would allow researchers to model the average impact of R&D on costs in probabilistic terms. However, these coefficients still represent an average effect across expert opinion, cleaned from expert fixed effects and from the effect of other cost variables. Thus, if a very optimistic expert ends up being right, using the average values in economic models (as is the case in Bosetti et al., 2014) would still result in a pessimistic range of estimates.

We also show that for a subsample of observations, estimates that included assumptions about technology deployment had a significant impact on costs, indicating the importance of deployment and of specifying assumptions clearly. Deployment, as discussed above in terms of recent market experience, may also be affecting how private sector and other experts view the future. Incorporating aspects of adoption in expert elicitation exercises is important to fully characterizing the range of possible outcomes, and thus to inform the policies that depend on them. A challenge in this respect is that the most knowledgeable experts on technical aspects of technology systems may not be knowledgeable about the impact that production-related improvements could have on future technology costs. One possible solution would be to combine elicitations of R&D effects with historically derived estimates for returns to scale and learning by doing, as in Nemet and Baker (2009). This hybrid approach provides a way to incorporate these multiple mechanisms of technological changes without over-relying on expert judgment in areas where knowledge is weak (Morgan, 2014). Alternatively, complementing individual elicitations with subsequent expert workshops, as in Anadon et al. (2012), might enable aggregating judgment among experts with differing, but overlapping, areas of expertise.

Summarizing, several implications for future work and analysis emerge from this paper. For the research community, our results show that diligence is needed in the selection of experts and the design of elicitations in future studies. They also point out the need for careful interpretation of elicitation results and consideration of alternative methodologies by which the reliability of elicitation results can be improved. In this respect, estimating the coefficient associated with the R&D variables using a meta-regression approach potentially represent improved inputs to energy system models that characterize future technology and policy outcomes probabilistically, as well as for probabilistic policy analyses specific to solar. Results of expert elicitations are already being used as inputs to energy economic models, for example in the papers in this Special Issue. In future work, parameter values such as the cost reductions attributable to R&D, could be adjusted, or de-biased, to account for expert selection and study design effects. Researchers should also consider how to include information from the multitude of forecasts of future costs that do not include attached probabilities. Expert weighting may be more helpful in such a context because they may substitute for the processes of de-biasing experts and helping them think through the full range of probabilities that may be lacking in those studies.

Acknowledgements

The authors would like to thank three anonymous referees for the insightful comments which helped improve the paper. Laura Diaz Anadon acknowledges funding from the Science, Technology, and Public Policy program at the Harvard Kennedy School and grants from the Doris Duke Charitable Foundation and BP to the Energy Technology Innovation Policy research group. She also thanks Stephen Elliott for his work on the LCOE calculation for the Harvard data. Elena Verdolini gratefully acknowledges funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 308481 (ENTRACTE) "Economic iNsTRuments to Achieve Climate Targets in Europe." Gregory Nemet acknowledges support from the Wisconsin Alumni Research Foundation (WARF).

Figures

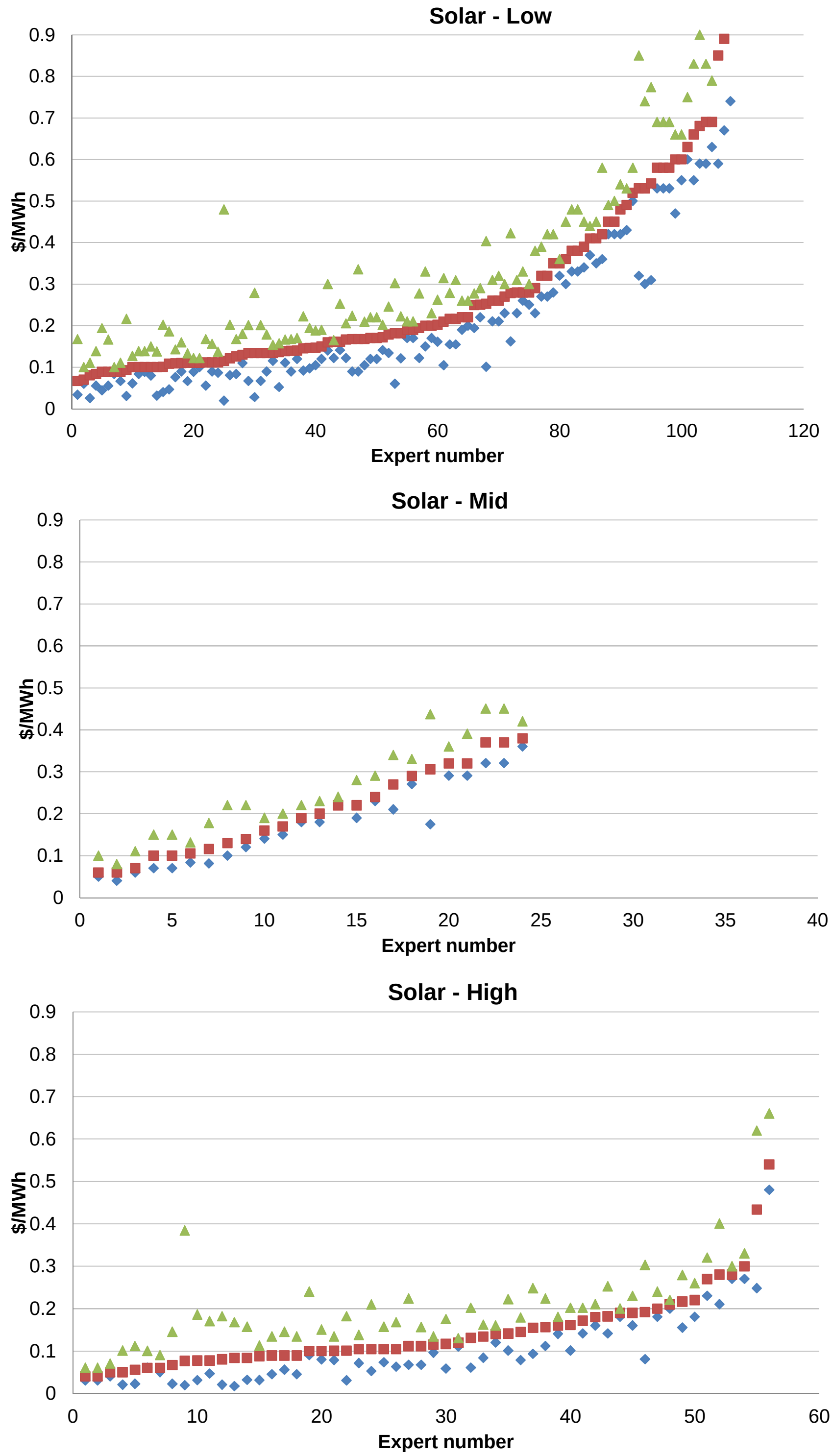

Figure 1. Standardized solar LCOE data from the 5 solar PV expert elicitations grouped by R&D investment level. Top panel: Low R&D investment level; Central panel: Medium R&D investment level; Bottom panel: High R&D investment level. The blue diamonds, red squares and green triangles correspond, respectively, to the 10^{th}, 50^{th}, and 90^{th} percentiles.

Tables

Table 1. Summary of elicitation studies.

Study Code	Group name	Source	# of experts	Aca-demic	Private	Public	year of elicitation	In-person	Pub-lished
1	Baker*	Baker, Chon, and Keisler (2009)	3	3	0	0	2007	1	1
2	Harvard	Anadon et al. (2011)	11	1	6	4	2010	0	0
3	FEEM**	Bosetti et al. (2012)	13	5	4	4	2011	1	1
4	NearZero	Near Zero (2012)	21	11	6	4	2011	0	0
5	Curtwright***	Curtwright et al. (2008)	18	9	4	8	2008	1	1

* University of Massachusetts Amherst

** Fondazione Eni Enrico Mattei

*** Carnegie Mellon University

*** Some experts in this study take more than one affiliation.

Table 2: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Dependent variable, 2030 levelized energy costs ($/kWh)					
Median (p50)	310	0.17	0.15	0.03	1.02
p50/($p50_{2010}$)	310	0.31	0.29	0.05	1.88
(p90-p10)/p50	310	0.89	0.70	0.09	4.75
logP50	310	-2.03	0.69	-3.50	0.02
lognormP50	310	-1.44	0.70	-2.91	0.63
logUrange	310	-0.38	0.73	-2.40	1.56
Investment					
R&D($m)	141	875	1823.20	25.51	10000
logRD	141	5.72	1.22	3.24	9.21
logRD square	141	34.17	16.02	10.49	84.83
RD High	310	0.29	0.46	0	1
RD Mid	310	0.16	0.37	0	1
RD Low	310	0.55	0.50	0	1
Deployment	310	0.32	0.47	0	1
Technology characteristics					
CPV	310	0.06	0.24	0	1
PV	310	0.61	0.49	0	1
Novel PV	310	0.12	0.33	0	1
Thinfilm	310	0.21	0.41	0	1
Market characteristics					
commercial	141	0.23	0.42	0	1
residential	141	0.27	0.45	0	1
utility	141	0.50	0.50	0	1
Expert characteristics					
academia	310	0.35	0.48	0	1
private	310	0.33	0.47	0	1
public	310	0.32	0.47	0	1
EU	310	0.12	0.32	0	1
Study characteristics					
Published	310	0.60	0.49	0	1
Inperson	310	0.60	0.49	0	1
Yearestimatemade	310	2009	1	2007	2011

Table 3: Regression results on the central estimate (Y=P50) of the solar levelized cost of electricity under different models.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	logP50	logP50	lognormP50	logP50	logP50
High R&D	-0.514***	-0.538***	-0.514***		
	[0]	[4.98e-08]	[0]		
Mid R&D	-0.418***	-0.416***	-0.418***		
	[3.36e-09]	[3.12e-05]	[3.38e-09]		
Year estimate made	0.0133	-0.0943*	0.0133	0.0366	0.0428
	[0.806]	[0.0660]	[0.806]	[0.410]	[0.551]
Inperson	-0.772***	-0.813***	-0.978***	-1.135***	-1.237***
	[0.00381]	[0]	[3.81e-05]	[3.66e-06]	[8.37e-07]
academia	0.135	0.0540	0.135	-0.0557	-0.187***
	[0.358]	[0.536]	[0.358]	[0.318]	[0.000117]
public	0.279*	0.208**	0.279*	-0.0158	-0.147***
	[0.0558]	[0.0113]	[0.0558]	[0.464]	[0]
EU	-0.164	-0.000689	-0.164	-0.164***	-0.0323
	[0.121]	[0.996]	[0.121]	[0.000841]	[0.398]
logRD				-0.260***	-0.627**
				[8.50e-10]	[0.0352]
logRD2					0.0278
					[0.218]
Expert FEs	Yes	No	Yes	Yes	Yes
Observations	211	211	211	141	141
Adjusted R-squared	0.634	0.438	0.633	0.646	0.646

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

Table 4. Regression results on low-cost (Y=P10) and high-cost (Y=P90) outcomes for the solar levelized cost of electricity.

	(1)	(2)	(3)	(4)
VARIABLES	logP10	logP10	logP10	logP90
High R&D	-0.540***			-0.444***
	[3.30e-10]			[2.94e-10]
Mid R&D	-0.473***			-0.353***
	[8.01e-10]			[2.04e-08]
logRD		-0.266***	-0.720**	
		[3.89e-10]	[0.0303]	
logRD2			0.0344	
			[0.171]	
Year estimate made	0.106	-0.0408	-0.00203	-0.0191
	[0.131]	[0.451]	[0.981]	[0.699]
Inperson	-0.538*	-1.445***	-1.477***	-0.900***
	[0.0684]	[9.83e-08]	[1.08e-07]	[0.000913]
academia	0.473**	-0.398***	-0.436***	0.00679
	[0.0189]	[0.000918]	[0.000344]	[0.957]
public	0.727***	-0.412***	-0.449***	0.101
	[0.000628]	[0]	[9.08e-10]	[0.434]
EU	-0.243*	-0.243***	-0.205***	-0.156
	[0.0791]	[0.000237]	[0.00818]	[0.137]
Expert FE	Yes	Yes	Yes	Yes
Observations	211	141	141	211
Adj. R-squared	0.733	0.700	0.702	0.594

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

Table 5. Regression results on the uncertainty range (Y=(P90-P10)/P50) of the solar levelized cost of electricity under different models.

VARIABLES	(1) logUrange	(2) logUrange	(3) logUrange	(4) logUrange
High R&D	0.138**	0.134		
	[0.0260]	[0.188]		
Mid R&D	0.207***	0.0852		
	[0.000561]	[0.437]		
Year estimate made	-0.103*	-0.211**	0.223***	0.125**
	[0.0602]	[0.0122]	[0.000105]	[0.0264]
Inperson	-0.243	0.602***	0.883***	0.637***
	[0.220]	[0.000464]	[2.23e-05]	[0.000878]
academia	-0.326**	-0.189*	1.014***	0.684***
	[0.0138]	[0.0921]	[0]	[0]
public	-0.439***	-0.261***	1.125***	0.796***
	[0.00136]	[0.00393]	[0]	[0]
EU	0.169	0.415	0.169	0.499***
	[0.376]	[0.119]	[0.424]	[0.00885]
logRD			0.0537*	0.197
			[0.0703]	[0.331]
logRD2				-0.0108
				[0.484]
Expert FEs	Yes	No	Yes	Yes
Observations	211	211	141	141
Adjusted R-squared	0.791	0.395	0.739	0.737

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

References

Anadon, L., 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, 08/2013 volume 8(3).

Anadon, L., Bunn, M., Chan, G., Chan, M., Jones, C., Kempener, R., Lee, A., Logar, N. & Narayanamurti, V., 2011. Transforming U.S. Energy Innovation. Cambridge, Mass.: Report for Energy Technology Innovation Policy research group, Belfer Center for Science and International Affairs, Harvard Kennedy School.

Baker, E., Chon, H. & Keisler, J., 2009. Advanced Solar R&D: Combining Economic Analysis with Expert Elicitations to Inform Climate Policy. Energy Economics 31:S37-S49.

Baker, E., Bosetti, B., Anadon, L. D., Henrion, M., Aleluia Reis, L., 2014. Future Costs of Key Low-Carbon Energy Technologies: Harmonization and Aggregation of Energy Technology Expert Elicitation Data. Energy Policy, Revise & Resubmit.

Bistline, John (2014). Energy technology expert elicitations: An application to natural gas turbine efficiencies, Technological Forecasting and Social Change, 86.

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(C):308-317.

Bosetti, V., Marangoni, G., Borgonovo, E. Anadon, L. D., Barron R., McJeon, H. C., Politis, S. & Friley, P. (f2014), Sensitivity to Technology Costs: A Multi-model comparison analysis. Energy Policy, forthcoming

Bolger, F. & Rowe, G., 2014. "The Aggregation of Expert Judgment: Do Good Things Come to Those Who Weight?" Risk Analysis: in press.

BP Statistical Review of World Energy 2014, http://www.bp.com/en/global/corporate/about-bp/energy-economics/statistical-review-of-world-energy/statistical-review-downloads.html

Clemen, R. T. and T. Reilly (2001). Making hard decisions with Decision Tools. Pacific Grove, CA, Duxbury

Craig, P. P., Gadgil, A. & Koomey, J. G., 2002. What Can History Teach Us: A Retrospective Examination of Long-Term Energy Forecasts for the United States. Annual Review of Energy and Environment 27: 83-118.

Cohen, Linda and Roger Noll (1991). "The Technology Pork Barrel," Brookings.

Cooke, R. M., M. E. Wittmann, D. M. Lodge, J. D. Rothlisberger, E. S. Rutherford, H. Zhang and D. M. Mason, 2014. Out-of-sample validation for structured expert judgment of Asian carp establishment in Lake Erie. Integrated Environmental Assessment and Management: in press.

Curtright, A. E., Granger Morgan, G. M. , Keith, D. W., 2008. Expert Assessments of Future Photovoltaic Technologies. Environmental Science and Technology 42(24):9031-9038

Grubler, A. & Wilson, C., 2013. Energy Technology Innovation: Learning from Historical Successes and Failures. Cambridge, Cambridge University Press.

Junginger, M., Faaij, A. & Turkenburg, W.C., 2005. Global Experience Curves for Wind Farms. Energy Policy 33:133–150.

Kahneman, Daniel, *Thinking, Fast and Slow*. 2011. MacMillan, USA, NY.

Keeney, R. L. and D. V. Winterfeldt, 1991. Eliciting probabilities from experts in complex technical problems. Transactions on Engineering Management Vol. 38: 191-201.

McGauran, N., Wieseler, B., Kreis, J., Schüler, Y., Kölsch, H. & Kaiser, T. 2010, Reporting bias in medical research - a narrative review, Trials 2010, 11:37

McNerney J., Farmer J. D. , Trancik J. E., Historical Costs of Coal-Fired Electricity and Implications for the Future, Energy Policy 39: 3042-3054

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

Morgan, M. G., 2014. Use (and abuse) of Expert Elicitation in Support of Decision Making for Public Policy. Proceedings of the National Academy of Sciences: 201319946

Near Zero, 2012. How Low Will Photovoltaic Prices Go? An Expert Discussion. Mason Inman, 6 December 2012 (available at: http://www.nearzero.org/reports/pv-learning)

Nemet, G. F., 2006. Beyond the Learning Curve: Factors Influencing Cost Reductions in Photovoltaics. Energy Policy 34(17): 3218-3232

Nemet, G. F. & Baker, E., 2009. Demand Subsidies Versus R&D: Comparing the Uncertain Impacts of Policy on a Pre-commercial Low-carbon Energy Technology. The Energy Journal 30(4): 49-80.

NRC (National Research Council), 2007. Prospective Evaluation of Applied Energy Research and Development at DOE (phase two). National Academies Press, Washington D.C., United States: National Research Council.

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

Phillips, L. D. (1999). Group elicitation of probability distributions: are many heads better than one?, Kluwer Academic Publishers.

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

Söderholm, P. & Klaassen, G., 2007. Wind Power in Europe: a Simultaneous Innovation–Diffusion Model. Environmental and Resource Economics 36(2):163–190.

Tversky, A., Kahneman, D. 1974. Judgment under Uncertainty: Heuristics and Biases. Science 185(4157): 1124-1131.

Walls, L. and J. Quigley (2001). "Building prior distributions to support Bayesian reliability growth modelling using expert judgement." Reliability Engineering and System Safety 74(2): 117-128.