Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Optimal investment and the ambiguous aggregation of expert opinions

Stergios Athanassoglou, Valentina Bosetti, Gauthier de Maere d'Aertryckey

Working Paper n. 468

This Version: August, 2012

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

The opinions expressed in the working papers are those of the authors alone, and not those of the Institute, which takes non institutional policy position, nor those of CEPR, NBER or Università Bocconi.

Optimal investment and the ambiguous aggregation of expert opinions*

Stergios Athanassoglou, Valentina Bosetti, Gauthier de Maere d'Aertrycke†

August 2012

Abstract

How should a decision-maker assess the potential of an investment when a group of experts provides strongly divergent estimates on its expected payoff? To address this question, we propose and analyze a variant of the well-studied α-maxmin model in decision theory. In our framework, and consistent to the paper's empirical focus on R&D investment, experts' subjective probability distributions are allowed to be action-dependent. In addition, the decision maker constrains the sets of priors to be considered in accordance with ethical considerations and/or operational protocols. Using tools from convex and conic optimization, we are able to establish a number of analytical results including a closed-form expression of our model's value function, a thorough investigation of its differentiability properties, and necessary conditions for optimal investment. We apply our framework to original data from a recent expert elicitation survey on solar technology. The analysis suggests that more aggressive investment in solar technology R&D is likely to yield significant dividends even, or rather especially, after taking ambiguity into account.

Keywords: expert aggregation; ambiguity; α-maxmin; second-order cone programming; renewable energy R&D

---

*We are grateful to Michela Catenacci and Giulia Fiorese for help with data collection. We also thank seminar participants at the Fondazione Eni Enrico Mattei (FEEM) and the "Workshop on the Role of Risk and Uncertainty in Energy and Climate Economics", held at ETH Zurich, for useful comments. The research leading to these results has received funding from the European Research Council under the European Commission's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement #240895, project ICARUS (Innovation for Climate Change Mitigation: a Study of Energy R&D, its Uncertain Effectiveness and Spillovers).

†Athanassoglou: Corresponding Author. European Commission Joint Research Center, Econometrics and Applied Statistics Unit, athanassoglou@gmail.com. Bosetti: FEEM, CMCC, and Bocconi University Economics Department, valentina.bosetti@feem.it. De Maere d'Aertrycke: FEEM and CMCC, gauthier.demaere@feem.it.

1 Introduction

Motivation & sketch of the main idea. Suppose a decision maker is called to determine optimal R&D investment in a breakthrough technology. Since the technology is completely new and untested, there are no historical data that may provide guidance on the potential effectiveness of R&D. Thus, a group of experts with broad experience in related ventures is assembled to provide their assessment of different investment options.

For concreteness, assume the decision maker is contemplating two investment scenarios, r_1 and r_2, and consults 5 experts whose judgments are independent of each other. Each expert $n \in \{1, 2, ..., 5\}$ provides two probability distribution functions (pdfs) $\{\pi_n(\cdot|r_1), \pi_n(\cdot|r_2)\}$ on the future payoff of the technology, conditional on the chosen level of investment. Suppose, for the sake of providing a simple example, that all these pdfs are Normal $\mathcal{N}(\mu, \sigma^2)$ densities. Table 1 lists their expected values and variances:

n	$\pi_n(\cdot\|r_1)$	$\pi_n(\cdot\|r_2)$
1	$\mathcal{N}(3, 3)$	$\mathcal{N}(5.5, 2)$
2	$\mathcal{N}(2, .5)$	$\mathcal{N}(2.25, 0)$
3	$\mathcal{N}(1, 1)$	$\mathcal{N}(1.5, .5)$
4	$\mathcal{N}(5, 1)$	$\mathcal{N}(2.75, 1)$
5	$\mathcal{N}(4, .5)$	$\mathcal{N}(2.5, 3)$

Table 1: An illustrative example.

Clearly, expert opinions vary widely under both r_1 and r_2 and it is not clear which one is better. How can a decision maker assimilate this information and ultimately choose between the two alternatives?

The standard way of tackling this problem is to aggregate over experts in some fashion, compute the resulting aggregate distributions $\pi(\cdot|r_1)$ and $\pi(\cdot|r_2)$, and use them to compare r_1 and r_2. The simplest instance of this practice would assign equal weight to each expert and linearly aggregate the pdfs. Indeed, there is a rich, primarily statistical and management-science, literature that studies the many different ways such aggregations may be performed. The overarching goal of this line of research is to derive a unique probability distribution encapsulating expert beliefs. In their comprehensive surveys, Clemen and Winkler [8, 9] broadly distinguish between (i) mathematical approaches and (ii) behavioral approaches. Mathematical approaches use the individual pdfs to construct a single probability distribution in two basic ways: either through axiomatically-justified mathematical formulas of aggregation, or, where possible, through Bayesian statistical methods that pay particular attention to issues of dependence and bias. Conversely, behavioral approaches are more qualitative in nature and involve the direct repeated interaction between experts in order to reach consensus on a single "group" estimate.

In contrast to this literature, we suppose that the above setting is one in which the quest for a unique probability distribution summarizing expert opinion is generically intractable. This could be due to a variety of reasons: (a) the proposed venture on which expert opinion is sought may be unprecedented and exceedingly complex; (b) experts may bring to the table different backgrounds, skills, and visions that are not readily comparable, (c) the decision maker may not have the time and resources to study expert opinions on a deeper level in order to then combine them, (d) he may need to make a one-shot investment now, so that he will not be able to learn over time the track record and quality of an expert's judgment, (e) he may simply want to simultaneously explore the consequences of many different aggregation schemes in a systematic fashion.

We do not claim that the above modeling assumptions are the norm, nor that they always call for different approaches than the ones surveyed by Clemen and Winkler. Rather, we merely argue that they are not completely implausible and that, when present, may significantly complicate the computation of a unique distribution. As a result, an alternative modeling framework for addressing such situations may be sought in the literature on decision making under ambiguity.[1] In contrast to the Bayesian setting, in which probabilities are assigned to events via a unique Bayesian prior, decision-theoretic models of ambiguity are designed to address situations in which a decision maker is unable to assign precise probabilistic structure to physical and economic phenomena. This, we assume, is the environment our decision maker finds himself in.

Acknowledging his inability to objectively assign weights to experts, the decision maker wishes to examine the implications of many different aggregation schemes. However, he wants to do so in a systematic and transparent fashion. To this end, suppose that he is willing to consider all possible aggregation schemes subject to a constraint that *no single agent be granted weight more than a level* $\tilde{\boldsymbol{b}}$. The latter could be a normative or operational desideratum in the sense that it provides a check on the influence of any single expert. Consequently, the decision maker grants that the payoff of r_i can be distributed according to any distribution satisfying $\sum_{n=1}^{5} p_n \pi_n(\cdot|r_i)$ for a collection of non-negative p_n's such that $\sum_{n=1}^{5} p_n = 1$ and $\max_n p_n \leq \tilde{b}$. The parameter $\tilde{b}$ can range from 1/5, implying that all five agents must be granted equal weight, to 1, which means that the decision maker takes into account all possible aggregation schemes, including ones that assign zero weight to all but one expert. Alternatively, when $\tilde{b} = 1/2$, the above requirement could be considered as the adoption of a sort of "no-dictatorship" clause.

Having parametrized the set of aggregation schemes under consideration in this way, the decision maker wishes to explore the range of possible outcomes that may occur as a result of choosing r_1 or r_2. To this end, the left panel of Figure 1 plots best- and worst-case expected payoffs over the

[1]See Gilboa and Marinacci [13] for a comprehensive recent survey of this literature.

sets of priors implied by $\tilde{b}$, in the manner described above. The plot makes clear that for values of $\tilde{b}$ less than about 0.32, r_1 has higher best- and worst-case payoffs than r_2, while this relation is reversed for $\tilde{b}$ greater than about 0.71. In between 0.32 and 0.71, r_1 dominates r_2 in the best case, and is dominated by it in the worst case.

The range between the best and worst-case curves encapsulates the range of possible outcomes of r_1 and r_2, subject to varying levels of $\tilde{b}$. What has not been provided yet is a criterion for helping decide between the two investment options. To this end, we assume that the decision maker's preferences are captured by *a convex combination of the best- and worst- case expected payoffs*. That is, given a choice $\tilde{b}$, the utility he assigns to r_1 or r_2 is equal to α ($\in [0,1]$) times the worst-case, plus $(1-\alpha)$ times the best-case expected payoffs, over the set of priors implied by the bound $\tilde{b}$. This seems like a reasonable way to rank alternatives, given a general ignorance over how to aggregate experts.[2] The right panel of Figure 1 demonstrates the investment option which dominates in different regions of $(\tilde{b}, \alpha)$ space. The red region denotes dominance by r_1, whereas the blue by r_2. In agreement with our earlier observations, we see that for $\tilde{b} \leq .32$ ($\geq .71$), r_1 (r_2) will always be preferred to r_2 (r_1), independently of α. For $\tilde{b} \in (.32, 0.71)$ the value of α matters. For example, when $\tilde{b} = 1/2$ we see that r_1 (r_2) is preferred as long as α is roughly less (greater) than 0.5. The implications of the decision maker's possible choices of $\tilde{b}$ and α have been clearly laid out in these graphs.

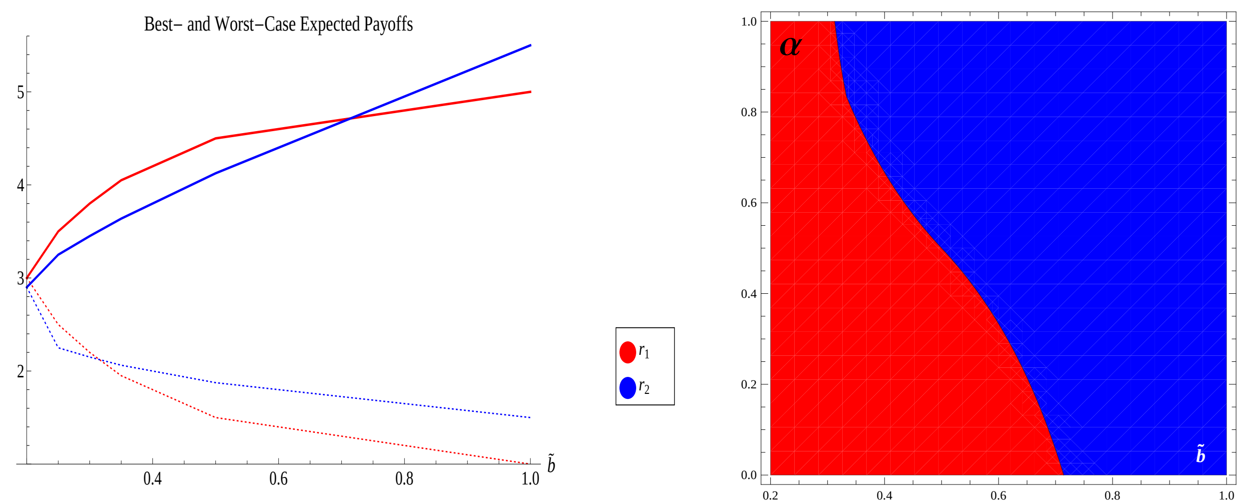

Figure 1: Applying a variant of the α-maxmin model to the example of Table 1.

[2]The axiomatic foundations of this choice are well-studied and go back to the work of Arrow and Hurwicz [2] in the 1970s. We elaborate in the next subsection.

Our contribution. The decision-theoretic model we introduce and analyze in this paper is an extended formal treatment of the above intuitive argument. In our setting, a decision maker elicits the judgment of a set of experts on the effect of R&D investment on the future cost of a technology.[3] Levels of R&D investment r affect the decision maker's problem in two ways: (a) they alter experts' subjective probability distributions on the technology's future cost and (b) they are arguments of a utility function that measures the investment's cost-effectiveness as a function of R&D expenditure and the resulting technological improvement. As an initial benchmark, our framework posits equal-weight linear aggregation over experts' divergent probability distributions. Subsequently, it considers enlargements of the set of possible aggregation schemes by parameterizing over their maximum distance, measured via the $\boldsymbol{l^2}$- norm, with respect to the benchmark equal-weight aggregation. This distance is referred to as *aggregation ambiguity*. It can be naturally interpreted as a bound on the total weight that can be assigned to any group of experts, thus modifying and generalizing the considerations introduced in the previous example with its emphasis on groups of single agents through parameter $\tilde{b}$. Its value should be objectively assigned by the decision maker, in accordance with operational or ethical desiderata. Next, our model computes the best-and worst-case expected outcomes of a given level of R&D investment, subject to the feasible set of distributions that is implied by assigned levels of aggregation ambiguity. Finally, the decision maker's preferences are captured by a convex combination of the best and worst-case expected outcomes. The relative weights placed on the worst and best cases represent his *ambiguity attitude*, the parameter α in the previous example.

Our model nests in a straightforward manner pure averaging and pure best/worst-case optimization. In addition, its simple structure allows for precise analytical insights. Using results from convex and conic optimization (Alizadeh and Goldfarb [1]), we are able to prove differentiability with respect to aggregation ambiguity and provide a simple closed-form expression for our value function and its optimizing arguments. As the optimization problem we are concerned with is a simple instance of second-order cone programming that is likely to appear in other contexts, these results may possibly be of more general interest. We proceed to investigate the value function's differentiability in R&D investment and, where applicable, provide a closed-form expression for this derivative. This can subsequently be used to obtain a necessary condition for optimal R&D investment. We conclude the paper's theoretical section by arguing that, while non-differentiability of the value function with respect to investment is in principle possible, it may not be often encountered in practice.

The paper's empirical section applies our model to original data from the ICARUS project

[3] While our formal model is general and can be applied to other contexts of decision making under uncertainty, we adopt the terminology of "technology" and "R&D investment" for concreteness.

(Bosetti et al. [4]), a recent expert elicitation survey on the potential of European Union R&D investment in alternative energy technologies.[4] As an initial step, we use the collected data of the survey to construct experts' subjective probability distributions on the future cost of solar energy conditional on R&D investment. Subsequently, we employ an integrated assessment model (Bosetti et al. [5]) to calculate the benefits of R&D investment (in the form of lower future solar-electricity costs) and use these estimates to inform our assesment of the relevant R&D alternatives. The application of our theoretical model to these data suggests that ambiguity plays an important role in assessing the potential of solar technology. Our analysis allows us to (cautiously) draw two policy implications: (1) that a doubling of EU investment in solar technology R&D is likely to yield significant dividends, even after taking ambiguity into account, (2) that a 50% increase in investment will likely always be less preferred to either a 100% increase, or maintaining the status quo.

Relation to the literature. The literature on belief aggregation and decision making under ambiguity is huge and spans a number of different disciplines including economics, statistics, operations research, psychology, and philosophy, among others. Thus, in the following remarks we focus purely on and describe the contributions which are directly relevant to the formal model we introduce.

Our framework is a variation of the α-maxmin model that has been studied extensively in the economic-theory literature beginning with Arrow and Hurwicz [2]. Later contributions by Gilboa and Schmeidler [14] (whose seminal paper dealt with the pure maxmin model), Ghirardato et al. [12], Chateauneuf et al. [7], and Eichberger et al. [10] provided axiomatic treatments of similar models in which a decision maker's actions are modeled by Savage acts [20], i.e. functions from a state space to a space of consequences. The model presented herein is not a strict application of this framework. This is because its decision variables are not functions but real numbers, representing levels of R&D investment, that enter the value function as arguments of both (a) a utility function measuring the technology's payoff as well as (b) the set of priors that the decision-maker is taking into account when performing his best- and worst-case analysis. This latter element of action-dependent subjective beliefs is non-standard in the decision-theoretic literature. Jaffray [16] had first introduced a similar notion with an α-maxmin model based on non-additive belief functions, while later Ghirardato [11] analyzed a model in which acts map from states to sets of consequences. More recently, Olszewski [18] studied the α-maxmin model in a related setting in which decision makers are called to choose between sets of lotteries over which the maximum and the minimum payoffs are subsequently computed. Moreover, Viero [21] axiomatized the α-maxmin model in a

[4]For more information see www.icarus-project.org.

setting in which acts map from states to sets of lotteries, and thus can be viewed as a generalization of the model of Olszewski.

Evidently, the model we study in this paper may be categorized as an α-maxmin model with action-dependent subjective probabilities. However, in contrast to the aforementioned papers, our formal setting is considerably less abstract and we do not pursue axiomatic analyses. Instead, we take the α-maxmin model's axiomatic foundations as given and introduce aggregation ambiguity as a novel (and, we hope, meaningful) model parameter defining the set of priors to be considered. Subsequently, we devote a great deal of attention to the derivation and differentiability properties of the α-maxmin value function. Correspondingly, the mathematical machinery we employ is also quite different than that of the more fundamental literature.

We see the primary virtues of this approach as being those of intuitiveness and practicality. Allowing for precise analytics and straightforward interpretation, our work aims to extend and operationalize the insights of the deeper contributions of the literature on the α-maxmin model to realistic decision-making settings. Indeed, the model we propose is an outgrowth of the need to develop a tractable theoretical framework to accommodate expert opinions gathered by the aforementioned ICARUS expert elicitation project (see Section 4).

Paper outline. The structure of the paper is as follows. Section 2 introduces the formal model, while Section 3 analyzes its theoretical properties. Section 4 illustrates the theoretical results with original data from the ICARUS expert elicitation survey on solar technology. Section 5 provides brief concluding remarks and directions for future research. All mathematical proofs, remaining tables and figures, as well as non-essential supplementary information are collected in an Appendix.

2 Model Description

Consider a set $\mathcal{N}$ of experts indexed by $n = 1, 2, ..., N$. R&D investment is denoted by a variable $r \in \mathcal{R}$ and the technology's cost by $c \in \mathcal{C}$, where $\mathcal{R}$ and $\mathcal{C}$ are subsets of real numbers. An expert n's probability distribution of the future cost of technology given investment r is captured by a random variable having a probability distribution function (pdf)

$$\pi_n(c|r). \tag{1}$$

Note that the decision variables of our model (R&D investment) directly affect experts' subjective probability distributions of the technology's cost. This means that our setting is not amenable to standard decision-theoretic frameworks going back to Savage [20].

Expert beliefs over the economic potential of R&D investment may, and usually do, vary significantly. The question thus naturally arises: How do we make sense of this divergence when studying optimal R&D investment? In the absence of data that could lend greater credibility to one expert over another, one straightforward way would be to simply aggregate over all pdfs π_n as given by Eq. (1), so that we obtain an "aggregate" joint pdf $\bar{\pi}$, where

$$\bar{\pi}(c|r) = \sum_{n=1}^{N} \frac{1}{N} \pi_n(c|r). \tag{2}$$

This approach inherently assumes that each and every expert is equally likely to represent reality, and makes use of simple linear aggregation. However, a great deal of information may be lost in such an averaging-out process, especially when there are big differences among experts.

We thus move beyond simple averaging. In our framework each expert n's pdf $\pi_n(c|r)$ is weighted by the decision maker through a second-order probability p_n. The set of admissible second-order distributions $\mathbf{p}$ depends on the amount of ambiguity the decision maker is willing to take into account when aggregating across experts, and in particular on how "far" he is prepared to stray from equal-weight aggregation. Specifically, we consider the set of second-order distributions $\mathrm{P}(b)$ over a set of N experts, parametrized by $b \in \left[0, \frac{N-1}{N}\right]$ where

$$\mathrm{P}(b) = \left\{ \mathbf{p} \in \Re^N : \quad \mathbf{p} \geq \mathbf{0}, \ \sum_{n=1}^{N} p_n = 1, \ \sum_{n=1}^{N} \left(p_n - \frac{1}{N} \right)^2 \leq b \right\}. \tag{3}$$

Here, the set $\mathrm{P}(b)$ captures the uncertainty of the decision-maker's aggregation protocol. Thus, we refer to parameter b it as *aggregation ambiguity.* Letting $\boldsymbol{e_N}$ denote a unit vector of dimension N, we see that distributions $\mathbf{p}$ belonging to $\mathrm{P}(b)$ satisfy $||\mathbf{p} - \frac{\boldsymbol{e_N}}{N}||_2 \leq \sqrt{b}$, where $|| \cdot ||_2$ denotes the L_2-norm. Setting $b = 0$ implies complete certainty and adoption of the equal-weight singleton, while $b = \frac{N-1}{N}$ complete ambiguity over the set of all possible second-order distributions.[5] (We discuss the interpretation and implications of different choices of b shortly.)

Weighting the expert pdfs (1) under all aggregation schemes belonging in $\mathrm{P}(b)$ induces the following set of priors

$$\Pi(b, r) = \left\{ \sum_{n=1}^{N} p_n(b) \pi_n(\cdot|r) : \quad \mathbf{p} \in \mathrm{P}(b) \right\} \tag{4}$$

governing the future cost of the technology conditional on R&D investment r. Thus, holding r fixed, an increase in b implies an expansion of the set of priors a decision maker is willing to consider.

Now, define the real-valued function

$$u\,(c, r) : \mathcal{C} \times \mathcal{R} \mapsto \Re,$$

[5]The latter statement holds in light of the fact that values of $b > \frac{N-1}{N}$ cannot enlarge the feasible set. This is because the maximizers of $\sum_{n=1}^{N} \left(p_n - \frac{1}{N} \right)^2$ over the set of probability vectors concentrate all probability mass on one expert, leading to an aggregation ambiguity of $\left(1 - \frac{1}{N}\right)^2 + (N-1) \cdot \left(\frac{1}{N}\right)^2 = \frac{N-1}{N}$.

as representing the utility of R&D investment r, under cost realization c. Given investment r, utility u, and the set of second-order distributions $\mathrm{P}(b)$ introduced in (3), we can calculate the best- and worst-case expected outcomes associated with r, given aggregation ambiguity b. This provides a measure of the spread, as measured by utility u, between the worst and best-cases, given a "willingness" to stray from the benchmark equal-weight distribution that is constrained by b. More formally, we consider the functions

$$V_{max}(r|b) = \max_{\pi \in \Pi(b,r)} \int_{\mathcal{C}} u(c,r)\, \mathrm{d}\pi(c) \tag{5}$$

$$V_{min}(r|b) = \min_{\pi \in \Pi(b,r)} \int_{\mathcal{C}} u(c,r)\, \mathrm{d}\pi(c). \tag{6}$$

The functions (5)-(6) fix a level of aggregation ambiguity b and subsequently focus on the best and worst cases. As such they capture extreme attitudes towards uncertainty in aggregation. To express more nuanced decision-maker preferences we consider the following value function

$$V(r|b,\alpha) = \alpha \cdot V_{min}(r|b) + (1-\alpha) \cdot V_{max}(r|b) \;\; \alpha \in [0,1], \tag{7}$$

representing a convex combination of the worst- and best-cases. The parameter α above captures the decision maker's *ambiguity attitude*. It measures his degree of pessimism given aggregation ambiguity b: the greater (smaller) α is, the more (less) weight is placed on the worst-case scenario. Given values for b and α, Eq. (7) operates as an objective function when searching for optimal investment decisions r.

What do different choices of b imply? We provide a straightforward interpretation of an ambiguity level b in our model (the interpretation of α is clear). Consider the benchmark equal-weight aggregation $\frac{1}{N}\,\boldsymbol{e_N}$. Now take a set of experts $\widehat{\mathcal{N}}$ of cardinality $\widehat{N}$ and begin increasing the collective second-order probability attached to their pdfs. The convex structure of the feasible set $\mathrm{P}(b)$ enables us to provide a tight upper bound on the maximum total second-order probability that can be placed on this set of experts, as a function of b and $\widehat{N}$ (we denote $\left|\widehat{\mathcal{N}}\right| = \widehat{N}$):

$$\max_{\mathbf{p} \in \mathrm{P}(b)} \sum_{n \in \widehat{\mathcal{N}}} p_n = \min\left\{ \frac{\widehat{N}}{N} + \widehat{N}\sqrt{\frac{N-\widehat{N}}{\widehat{N}N}b},\ 1 \right\}. \tag{8}$$

Extending this logic to any subset of experts, we have the following holding:

$$\mathbf{p} \in \mathrm{P}(b) \;\;\Leftrightarrow\;\; \left\{ \mathbf{p} \geq \mathbf{0},\ \sum_{n=1}^{N} p_n = 1,\ \sum_{n \in \widehat{\mathcal{N}}} p_n \leq \min\left\{ \frac{\widehat{N}}{N} + \widehat{N}\sqrt{\frac{N-\widehat{N}}{\widehat{N}N}b},\ 1 \right\},\ \text{for all } \widehat{\mathcal{N}} \subseteq \mathcal{N} \right\}. \tag{9}$$

As mentioned before, the value of b should be objectively assigned by the decision maker in accordance with ethical or operational desiderata. For instance, and again recalling the motivating example of Section 1, choices of b which seem natural are ones that ensure that no single agent is granted weight greater than 1/2. Applying this requirement to Eq. (8) would mean that we restrict ourselves to choices of b satisfying

$$b \leq \frac{(N-2)^2}{4N(N-1)}. \tag{10}$$

Yet, while the above restriction is reasonable, we remain agnostic with regard to the determination of a universally-preferred choice of b, and envision different values to be appropriate in different contexts. Indeed, when the stakes are very high and disagreement very acute, it may be more important to set a high value for b and allow for assigning a lot of weight to extreme opinions, than to respect no-dictatorship considerations leading to bounds such as (10).

3 Theoretical Results

In this section we focus on the optimization problems (5) and (6) and analyze the behavior of value function $V(r|b, \alpha)$, as we vary ambiguity levels b and R&D investment r. Using results from convex optimization we are able to compute this function and establish its differentiability in b (everywhere except at a single point). Differentiablity with respect to r is more subtle and we use the results of Milgrom and Segal [17] to provide ranges of b and α for which it holds. However, we conclude the section by arguing that, while non-differentiability of the value function with respect to investment is in principle possible, it may not be often encountered in practice. Where applicable, we ease notation in the following manner:[6]

$$\begin{aligned}
u_n(r) &\equiv \int_{\mathcal{C}} u(c,r) \mathrm{d}\pi_n(c|r), \\
V_{max}(r,b) &\equiv V_{max}(r|b) = \max_{\pi \in \Pi(b,r)} \int_{\mathcal{C}} u(c,r)\, \mathrm{d}\pi(c) = \max_{\mathbf{p} \in \mathrm{P}(b)} \sum_{n=1}^{N} p_n u_n(r) & (11) \\
V_{min}(r,b) &\equiv V_{min}(r|b) = \min_{\pi \in \Pi(b,r)} \int_{\mathcal{C}} u(c,r)\, \mathrm{d}\pi(c) = \min_{\mathbf{p} \in \mathrm{P}(b)} \sum_{n=1}^{N} p_n u_n(r). & (12)
\end{aligned}$$

Eqs. (11) and (12) are valid by the linearity of the expectation operator under mixtures of distributions.

The remainder of the Section is divided into two parts. Section 3.1 studies optimization problems (11) and (12) under a fixed value of r. Our objective here is to study the effect of varying levels of b. Section 3.2 uses the results derived in 3.1 to investigate the differentiability of the value function $V(r|b, \alpha)$ and provide necessary conditions for optimal R&D investment.

[6] While b is a parameter, we will abuse notation and, throughout Section 3.1, consider it a variable.

3.1 Varying b under fixed r

In this subsection, we focus our attention on the maximization problem (11) as the results and proofs for (12) are completely symmetric. This is because $V_{min}(r,b) = -\max_{\mathbf{p}\in \mathrm{P}(b)} \sum_{n=1}^{N} p_n(-u_n(r))$. Furthermore, while we keep r fixed, we do not suppress it as an argument in order to stress the results' dependence on chosen levels of R&D investment.

We begin by proving a few straightforward properties regarding continuity, monotonicity, and concavity of $V_{max}(r,b)$.

Proposition 1 *Fix $r \in \mathcal{R}$. The function $V_{max}(r,b)$ defined in Eq. (11) is increasing, concave, and continuous in b.*

Before we state our next result we need to introduce additional notation. First, let $\mathcal{N}_k(r)$ denote the set of experts sharing the k'th order statistic of $\{u_1(r), u_2(r), ..., u_N(r)\}$. There are a total of $N(r)$ such sets where, depending on the problem instance, $N(r)$ can be any number in $\{1, 2, ..., N\}$, and we define $N_k(r) = |\mathcal{N}_k(r)|$. For instance, and with apologies for the clunky notation, $\mathcal{N}_{N(r)}(r)$ denotes the set of agents sharing the maximum of $\{u_1(r), u_2(r), ..., u_N(r)\}$. Furthermore, let $\mathcal{N}_k^+(r) = \bigcup_{i=k}^{N} \mathcal{N}_i(r)$, $\mathcal{N}_k^-(r) = \bigcup_{i=1}^{k} \mathcal{N}_i(r)$ and $N_k^+(r) = \left|\mathcal{N}_k^+(r)\right|$, $N_k^-(r) = \left|\mathcal{N}_k^-(r)\right|$. Our model structure enables us to easily show the following Lemma.

Lemma 1 *Fix $r \in \mathcal{R}$ and consider the optimization problem (11). Define ambiguity level $b^*_{max}(r) \equiv \frac{1}{N_{N(r)}(r)} - \frac{1}{N}$. $V_{max}(r,b)$ is strictly increasing in $b \in [0, b^*_{max}(r)]$ and equal to $\max_{n\in\mathcal{N}} u_n(r)$ in $b \in [b^*_{max}(r), \frac{N-1}{N}]$.*

Lemma 1 suggests that $b^*_{max}(r)$ is an important threshold. It represents the level of aggregation ambiguity above which the set $\mathrm{P}(b)$ allows for the maximum expert estimate to be attained as an objective function value of (11). Our next result establishes that for levels of ambiguity smaller than this extreme value, the optimal solutions of problem (11) will be unique and bind the quadratic ambiguity constraint associated with set $\mathrm{P}(b)$.

Proposition 2 *Fix $r \in \mathcal{R}$. Suppose $b \in [0, b^*_{max}(r)]$ and consider the maximization problem (11). There exists a unique optimal solution $\mathbf{p^{max}}(r,b)$ and it must satisfy the quadratic constraint of set (3) with equality. For $b \in (b^*_{max}(r), \frac{N-1}{N}]$ all probability vectors $\mathbf{p^{max}}(r,b)$ satisfying $p_n^{max}(r,b) = 0$ for $n \notin \mathcal{N}_{N(r)}(r)$ and $\sum_{n\in\mathcal{N}_{N(r)}} \left(p_n^{max}(r,b) - \frac{1}{N}\right)^2 \leq b - \frac{N^-_{N(r)-1}(r)}{N^2}$ will be optimal solutions of (11).*

We are now ready to prove the paper's first main result. Theorem 1 establishes that function $V_{max}(r,b)$ is differentiable with respect to b everywhere on $\left(0, \frac{N-1}{N}\right)$ except at the point $b^*_{max}(r)$. Moreover, it formalizes a straightforward monotonicity property of the optimal solutions of (11) and (12) that is essential to the derivation of the value function pursued in Theorem 2. In proving Theorem 1 we make extensive use of results from conic optimization, in particular the duality theory of second-order cone programming (see Alizadeh and Goldfarb [1]).

Theorem 1 *Fix $r \in \mathcal{R}$.*

(a) The function $V_{max}(r, b)$ is differentiable with respect to b everywhere on $b \in \left(0, \frac{N-1}{N}\right)$ except $b^*_{max}(r)$.

(b) Let $\mathbf{p^{max}}(r, b)$ denote an optimal solution of $V_{max}(r, b)$. The following levels of aggregation ambiguity

$$b_k^{max}(r) \quad = \quad \min\left\{\hat{b}: \ \forall b \geq \hat{b}, \ \textit{ we have } \ p_n^{max}(r, b) = 0 \ \textit{ for all } \ n \in \mathcal{N}_k(r)\right\}, \tag{13}$$

where $k \in \{1, 2, ..., N(r) - 1\}$, are well-defined and strictly increasing in k.

Part (a) of Theorem 1 shows that, keeping r fixed, $V_{max}(r, b)$ is a smooth function of b, except for a single kink at the level of aggregation ambiguity at which the maximum estimate can be obtained. Part (b) implies that $b_k^{max}(r)$ can be interpreted in the following way: it denotes the threshold level of ambiguity such that, for all b greater than or equal to it, at optimality no probability mass is ever allocated to experts belonging in $\mathcal{N}_k^-(r)$ (i.e., having a $u_n(r)$ that is less than or equal to the k'th order statistic of $\{u_1(r), u_2(r), ..., u_N(r)\}$). Thus, when b exceeds this level, one can safely disregard experts in $N_k^-(r)$. While the existence and monotonicity of these ambiguity thresholds (13) may be intuitive, their proofs are relatively involved and require insights from conic duality [1].

Having established differentiability with respect to b, we go on to provide a set of differential equations that $V_{max}(r, b)$ must satisfy. These differential equations will prove valuable in its subsequent derivation.

Proposition 3 *Fix $r \in \mathcal{R}$ and let $\mathbf{p^{max}}(r, b)$ denote the unique optimal solution of maximization problem (11) as a function of $b \in [0, b^*_{max}(r)]$. Suppose expert n_k satisfies $n_k \in \mathcal{N}_k(r)$. Consider $b_k^{max}(r)$ defined in Eq. (13). $V_{max}(r, b)$ satisfies the following differential equation:*

$$2\frac{\partial}{\partial b}V_{max}(r, b)\left(p_{n_k}^{max}(r, b) - \frac{1}{N} - b\right) = u_{n_k}(r) - V_{max}(r, b), \ \ b \in (0, b_k^{max}(r)). \tag{14}$$

Before presenting the paper's second main result, let $u_{(k)}(r)$ denote the k'th order statistic of $\{u_1(r), u_2(r), ..., u_N(r)\}$, where $k = 1, 2, ..., N(r)$. Now, define the following quantities:

$$\overline{u}_k(r)^+ \quad = \quad \frac{\sum_{n \in \mathcal{N}_k^+(r)} u_n(r)}{N_k^+(r)}, \tag{15}$$

$$d_k(r)^+ \quad = \quad N_{k+1}^+\left(u_{(k)}(r) - \overline{u}_{k+1}(r)^+\right)^2, \tag{16}$$

where $k \in \{1, 2, ..., N(r)\}$ (we set $d^+_{N(r)} \equiv 0$).

The term $\overline{u}_k(r)^+$ is simply an average of the values of the set $\{u_1(r), u_2(r), ..., u_N(r)\}$ that are greater than or equal to its k'th order statistic. The term $d_k(r)^+$ measures the total dispersion between the k'th expected payoff and the average of those greater than it, adjusted for size of the latter group.

Now, we use the above quantities to define

$$b^+_0(r) \equiv 0, \;\; b^+_{N(r)}(r) \equiv \frac{N-1}{N}, \tag{17}$$

$$b^+_k(r) = \frac{N^-_k(r)}{N^+_{k+1}(r)N} + \frac{\sum_{l=k+1}^{N(r)-1} d_l(r)^+}{N^+_{k+1}(r)d_k(r)^+}\,\frac{N_k(r)}{N^+_k(r)}, \;\; k \in \{1, 2, ..., N(r)-1\}. \tag{18}$$

Straightforward algebra yields the following monotonicity properties, where $k \in \{1, 2, ..., N(r)-1\}$:[7]

$$\begin{aligned} N^+_{k+1}(r)d_k(r)^+(r) &> N^+_{k+2}(r)d_{k+1}(r)^+, \\ b^+_k(r) &< b^+_{k+1}(r). \end{aligned}$$

We are now ready to state our second main result and provide a closed-form expression for $V_{max}(r, b)$. To prove the following Theorem, we explicitly solve the systems of differential equations established in Proposition 3.

Theorem 2 *Fix $r \in \mathcal{R}$. Consider the optimization problem (11) and the vectors $(\overline{\boldsymbol{u}}^+(r), \boldsymbol{d}^+(r), \boldsymbol{b}^+(r))$ defined in Eqs. (15)-(16)-(17)-(18). The vector $\boldsymbol{b}^+(r)$ satisfies*

$$b^+_k(r) = b^{max}_k(r) \quad k \in \{1, 2, ..., N(r)-1\}$$

where $b^{max}_k(r)$ is defined in Eq. (13). The function $V_{max}(r, b)$ equals

$$V_{max}(r, b) = \overline{u}_k(r)^+ + \sqrt{\left(b - \frac{N^-_{k-1}(r)}{N^+_k(r)N}\right)\sum_{l=k}^{N(r)-1} d^+_l(r)}, \;\; b \in \left[b^+_{k-1}(r), b^+_k(r)\right), \tag{19}$$

where $k = 1, 2, ..., N(r)$.

Theorem 2 shows that, keeping r fixed, $V_{max}(r, b)$ will be a concatenation of appropriately-specified square-root-like functions (when $k = 1$, we set $\frac{N^-_{k-1}(r)}{N^+_k(r)N} \equiv 0$).[8] These concatenations occur

[7] Details available upon request.

[8] These results are consistent with the more general analysis of Section 4.2 in Iyengar [15]. However, Iyengar uses different arguments and does not prove differentiability in b, nor does he derive and interpret differential equations and a precise formula for V_{max} and its optimal solution $\mathbf{p^{max}}$ (we provide the latter in Corollary 1). Instead, his analysis is concerned with determining the complexity of calculating an optimal solution of (12).

at levels of ambiguity $\boldsymbol{b^{max}}$ which are interpreted by Eq. (13), and can be computed explicitly through Eq. (18). The curvature of these functions is driven by the dispersion of experts' expected estimates, as captured by the quantities $d_k(r)^+$ of Eqs (16).

As we can see from Eq. (8), when $k \geq 2$, the fraction $\frac{N^-_{k-1}(r)}{N^+_k(r)N}$ in the square root represents the minimum level of b at which it becomes possible to assign zero weight to all experts in $\mathcal{N}^-_{k-1}(r)$. Hence, the fact that $b^{max}_{k-1}(r) = b^+_{k-1}(r) > \frac{N^-_{k-1}(r)}{N^+_k(r)N}$ for all k (see Eq. (18)), ensures that Eq. (19) is well-defined.

We can now combine the various results we have established to characterize the optimal solution of (11).

Corollary 1 *Consider $r \in \mathcal{R}$ and $b \in \left[0, \frac{N-1}{N}\right]$. Suppose first that $b < b^+_{N(r)-1}(r)$. Consider any expert $n_l \in \mathcal{N}_l(r)$ for some $l \in \{1, 2, ..., N(r)\}$. There exists a unique optimal solution $\mathbf{p^{max}}(r, b)$ and it satisfies*

$$p^{max}_{n_l}(r, b) \quad = \quad \begin{cases} \frac{1}{N^+_k(r)} + \left(u_{n_l}(r) - \overline{u}_k(r)^+\right) \frac{\sqrt{b - \frac{N^-_{k-1}(r)}{NN^+_k(r)}}}{\sqrt{\sum_{i=k}^{N(r)-1} d^+_i(r)}} & b \in \left[b^+_{k-1}(r), b^+_k(r)\right),\ k = 1, 2, ..., l \\ 0 & b \in \left[b^+_l(r), b^+_{N(r)-1}(r)\right). \end{cases}$$

Now suppose $b \geq b^+_{N(r)-1}(r)$. Here, by Proposition 2 all vectors $\mathbf{p^{max}}(r, b)$ satisfying $p^{max}_n(r, b) = 0$ for $n \notin \mathcal{N}_{N(r)}(r)$ and $\sum_{n \in \mathcal{N}_{N(r)}} \left(p^{max}_n(r, b) - \frac{1}{N}\right)^2 \leq b - \frac{N^-_{N(r)-1}(r)}{N^2}$ will be optimal. This set is a singleton at $b = b^+_{N(r)-1}(r)$.

Corollary 1 provides succinct expressions for the optimal expert probabilities given investment r and aggregation ambiguity b.

3.2 Necessary conditions for optimal R&D investment r

We now shift the focus of our analysis to investigate the differentiability of value function $V(r|b, \alpha)$, given by Eq. (7), with respect to r.[9] Here, the picture is considerably more subtle. We use the results of Milgrom and Segal [17] to state the following Theorem.

Theorem 3 *Fix $b \in \left[0, \frac{N-1}{N}\right]$ and $\alpha \in [0, 1]$ and consider the value function $V(r|b, \alpha)$ given by Eq. (7). Assume $\mathcal{R} = [r_m, r_M] \subset \Re$ and that the functions $u_n(r)$ are continuously differentiable on $\mathcal{R}$ for all $n \in \mathcal{N}$. Let $\mathbf{P^{max}}(r, b)$ and $\mathbf{P^{min}}(r, b)$ denote the sets of optimal solutions of problems (11) and (12) respectively, as given by Corollary 1 and its equivalent statement for the minimization*

[9] We now return to considering b as a parameter of the value function.

problem. The function $V(\cdot|b,\alpha):\ \mathcal{R}\to\Re$ is differentiable at $r_0\in(r_m,r_M)$ if and only if the set

$$\left\{\sum_{n=1}^{N}\left(\alpha p_n^{min}(r_0,b)+(1-\alpha)p_n^{max}(r_0,b)\right)\frac{\mathrm{d}}{\mathrm{d}r}u_n(r_0)\right\}$$

is a singleton. In that case,

$$\frac{\mathrm{d}V}{\mathrm{d}r}(r_0|b,\alpha) \quad = \quad \sum_{n=1}^{N}\left(\alpha p_n^{min}(r_0,b)+(1-\alpha)p_n^{max}(r_0,b)\right)\frac{\mathrm{d}}{\mathrm{d}r}u_n(r_0), \tag{20}$$

for all pairs of optimal $\mathbf{p^{min}}(r_0,b)\in\mathbf{P^{min}}(r_0,b)$ and $\mathbf{p^{max}}(r_0,b)\in\mathbf{P^{max}}(r_0,b)$.

Theorem 3 in combination with Proposition 2 allows us to establish the differentiability of the value function $V(r|b,\alpha)$ at a point $r=r_0$ for a range of b and α.

Corollary 2 *Suppose the assumptions of Theorem 3 hold. The function $V(\cdot|b,\alpha):\mathcal{R}\to\Re$ is differentiable at $r_0\in(r_m,r_M)$ for all $b\in[0,\min\{b^*_{max}(r_0|\alpha),b^*_{min}(r_0|\alpha)\}]$ where $b^*_{max}(r|\alpha)=b^*_{max}(r)$ if $\alpha<1$ and $\frac{N-1}{N}$ otherwise and $b^*_{min}(r|\alpha)=b^*_{min}(r)$ if $\alpha>0$ and $\frac{N-1}{N}$ otherwise. The derivative is given by Eq. (20) where $\mathbf{p^{max}}(r_0,b)$ and $\mathbf{p^{min}}(r_0,b)$ are uniquely defined by Corollary 1.*

Conversely, Theorem 3 also suggests that it is possible for the function $V(r|b,\alpha)$ to be non-differentiable at a value r_0 for a nontrivial range of b and α. This non-differentiability is due to the fact that for $b>b^*_{max}(r_0)$ ($b^*_{min}(r_0)$) optimization problem $\max_{p\in P(b)}\sum_{n=1}^{N}p_nu_n(r_0)$ ($\min_{p\in P(b)}\sum_{n=1}^{N}p_nu_n(r_0)$) admits multiple optimal solutions. Consequently, within this range of (b,α), a derivative at $r=r_0$ will generally fail to exist. Proposition 4 formalizes this observation.

Proposition 4 *Suppose the conditions of Theorem 3 hold and consider $r_0\in(r_m,r_M)$.*

(a) Suppose there exist two experts n_1 and n_2 satisfying $\frac{\mathrm{d}u_{n_1}}{\mathrm{d}r}(r_0)\neq\frac{\mathrm{d}u_{n_2}}{\mathrm{d}r}(r_0)$ and $n_1,n_2\in\mathcal{N}_{N(r_0)}(r_0)$ ($\mathcal{N}_1(r_0)$). Then the function $V_{max}(r|b)$ ($V_{min}(r|b)$) is not differentiable at $r=r_0$ for all $b>b^*_{max}(r_0)$ ($b>b^*_{min}(r_0)$).

(b) Suppose $\alpha\in(0,1)$. If there exist experts n_1 and n_2 satisfying $\frac{\mathrm{d}u_{n_1}}{\mathrm{d}r}(r_0)\neq\frac{\mathrm{d}u_{n_2}}{\mathrm{d}r}(r_0)$ such that $n_1,n_2\in\mathcal{N}_1(r_0)$ then the function $V(r|b,\alpha)$ is not differentiable at $r=r_0$ for all $b>b^*_{min}(r_0)$. If there exist experts n_3 and n_4 satisfying $\frac{\mathrm{d}u_{n_3}}{\mathrm{d}r}(r_0)\neq\frac{\mathrm{d}u_{n_4}}{\mathrm{d}r}(r_0)$ such that $n_3,n_4\in\mathcal{N}_{N(r_0)}(r_0)$ then the function $V(r|b,\alpha)$ is not differentiable at $r=r_0$ for all $b>b^*_{max}(r_0)$.

In instances described by Proposition 4, it is clear that one cannot use first-order conditions to establish the potential optimality of an R&D investment r_0. This non-differentiability is an unsatisfying, though not entirely unexpected, consequence of the maxmin nature of our model.

It stems from the fact that beyond a certain level of aggregation ambiguity there may exist a multiplicity of aggregation schemes that combine to yield the absolute maximum and minimum payoffs. Nevertheless, in the remainder of this section, we suggest that such non-smoothness issues may not, to a significant degree, encountered in practice.

Addressing non-differentiability. Given Corollary 2 and Proposition 4, we would like to narrow the range of b and α over which $V(r|b,\alpha)$ would fail to be differentiable. For this purpose, we provide a plausible lower bound on problematic ranges of b with the following informal argument. Consider carefully the continuously differentiable functions $u_n(r) = \int_{c\in\mathcal{C}} u(c,r)\mathrm{d}\pi_n(c|r)$. Since subjective probability distributions will generally differ across experts, then, assuming the domain $\mathcal{C}$ is moderately large, it is unlikely that at any point r_0 we will have more than 2 experts sharing the same expected payoff, including the maximum and minimum values of $\{u_1(r_0), u_2(r_0), ..., u_N(r_0)\}$. Therefore, it is likely that $N_1(r) \leq 2$ and $N_{N(r)}(r) \leq 2$ for all $r \in \mathcal{R}$. This observation leads to the following bound

$$\min_{r\in\mathcal{R}} \min\{b^*_{max}(r), b^*_{min}(r)\} \geq \frac{1}{2} - \frac{1}{N} = \frac{N-2}{2N},$$

so that Corollary 2 implies that $V(\cdot|b,\alpha)$ will be, at the very least, everywhere differentiable for any choice of $\alpha \in [0,1]$ and $b \leq \frac{N-2}{2N}$. If we assume that decision makers are constrained in their maximal choice of b due to, say, the no-dictatorship requirements leading to bounds such as those of Eq. (10), then potentially problematic ranges of $b \geq \frac{N-2}{N}$ are less likely to be considered and the negative result of Proposition 4 loses its bite.

Related to the above, Corollary 2 implies the diffentiability of $V(\cdot|b,\alpha)$ at all $r_0 \in (r_m, r_M)$, for all b and α, for an important special case: that in which there exists a pair of experts that are consistently the most optimistic and pessimistic across all levels of R&D. If surveyed experts have different backgrounds, such consistently optimistic and pessimistic opinions may occur.

Corollary 3 *Suppose the conditions of Theorem 3 hold and there exist two experts n_1 and n_2 such that $u_{n_1}(r) > u_n(r)$ for all $n \neq n_1$ and $u_{n_2}(r) < u_n(r)$ for all $n \neq n_2$, for all $r \in \mathcal{R}$. In other words, experts n_1 and n_2 are consistently the most optimistic and pessimistic across all levels of R&D. Then Theorem 3 implies that $V(r|b,\alpha)$ is differentiable at all $r_0 \in (r_m, r_M)$ for all choices of b and α, with its derivative given by Eq. (20).*

Hence, for problem instances satisfying the conditions of Corollary 3, first-order conditions that make use of Eq. (20) may be always invoked to solve for the maximizer of the value function $V(r|b,\alpha)$, regardless of the values of b and α. To be sure, these conditions will be necessary, though not in general sufficient, for optimality.

4 Empirical Application to Solar-Technology R&D

We base the empirical application of our paper to original data collected by the ICARUS survey, an expert elicitation on the potential of solar technologies. During the course of 2010-2011, the ICARUS survey collected expert judgments on future costs and technological barriers of different Photovoltaic (PV) and Concentrated Solar Power (CSP) technologies.[10] Sixteen leading European experts from academia, the private sector, and international institutions took part in the survey. The elicitation collected probabilistic information on (1) the year-2030 expected cost of the technologies; (2) the role of public European Union R&D investments in affecting those costs; and (3) the potential for the deployment of these technologies (both in OECD and non-OECD countries). We refer readers interested in the general findings of the survey to Bosetti et al. [4] and we focus here on the data on future costs as they form the basis of our analysis.

Current 5-year EU R&D investment in solar technology is estimated at 165 million US dollars. The ICARUS study elicited the probabilistic estimates of the 16 experts on the 2030 solar electricity cost (2005 c$/kWh) under three future *Scenarios*: (1) keeping current levels of R&D constant until 2030, (2) increasing them by 50%, and (3) increasing them by 100%. Coherent responses were obtained from 14 out of the 16 experts so the analysis that follows focuses solely on them. We used linear interpolation of the survey's collected data (generally 3-6 points of each expert's cumulative distribution function(cdf) conditional on R&D investment) to compute a pdf for each expert $n \in \{1, 2, ..., 14\}$, given the three relevant levels of R&D investment denoted by $r \in \{r_1, r_2, r_3\}$ (here r_i refers to Scenario i).[11] These pdfs represent experts' subjective probability distributions of the cost of technology as denoted in Eq. (1). Figure 2 plots the corresponding cdfs as well as the cdf that the aggregate pdf (2) leads to, under all three Scenarios.

[FIGURE 2 here]

As one can see in Figure 2 there is considerable disagreement between experts over the potential of solar technology. This disagreement is particularly acute under Scenario 1, and diminishes as R&D levels increase. Nonetheless, the breakthrough nature of innovation and the need to cross certain firm cost thresholds, means that ambiguity in expert estimates remains an important concern, even under Scenario 3. This will become apparent in the analysis to follow.

We measure the utility of an investment via its net payoff. Denoting the benefit associated to

[10] The survey is part of a 3-year ERC-funded project on innovation in carbon-free technologies (ICARUS - Innovation for Climate chAnge mitigation: a study of energy R&D, its Uncertain effectiveness and Spillovers www.icarus-project.org).

[11] Please refer to section A2 of the Appendix for more information on how expert pdfs were constructed from the ICARUS survey data.

a technology cost c by the function $B(c)$ and the opportunity cost of an investment r by $O(r)$, this is given by the following utility function:

$$u(c, r) = B(c) - O(r). \tag{21}$$

The next section describes how we provide numerical values for $B(\cdot)$ and $O(\cdot)$.

Quantifying benefits and opportunity costs of solar technology R&D. Expected benefits of solar technology R&D investments are quantified via a general equilibrium intertemporal model that can account for a range of macro-economic feedbacks and interactions. These include the effects of energy and climate change policies, the competition for innovation resources with other power technologies, the effect of growth, as well as a number of other factors.[12] To capture the long-term nature of such investments, the integrated assessment model is run over the time horizon 2105-2100 in 5-year time periods for the whole range of exogenously-imposed possible 2030-costs of solar power that we are considering. Subsequently, simulation results are compared to the benchmark case in which the cost of solar power is so high that the technology is not competitive with alternative production modes. For each possible 2030 solar-power cost, the benefit to the European Union is quantified by the discounted EU-consumption improvement over the entire time-horizon 2005-2100 with respect to the case where solar technology is not competitive. Table 2 summarizes the results.

[TABLE 2 here]

Three important assumptions are at the basis of the numbers reported in Table 1. First, as the survey concentrated on public EU R&D investment and the effects of increasing it, we disregard spillovers and technological transfers to the rest of the world and consider only the consumption improvement for Europe. Second, we evaluate the benefit of alternative 2030 costs of solar power assuming that no carbon policy is in place and that no special constraints on other technologies are imposed (e.g., a partial ban on nuclear technology). Third, we discount cash flows using a 3% discount rate. Although our choice is well in the range of discount rates adopted for large scale public projects, it is important to note that the cost threshold for positive payoffs is robust for a wide range of more myopic discount rate values. Our assumptions all err on the side of being conservative about the potential payoffs of solar-technology R&D.

We now explain how we calculate the costs of solar R&D investment. Given an R&D investment r, we assume that actual R&D spending is fixed at r during the period 2005-2030, in line with the

[12] The analysis is carried out using the World Induced Technical Change Hybrid (WITCH) model (Bosetti et al. [5]), an energy-economy-climate model that has been used extensively for economic analysis of climate change policies. See www.witchmodel.org for a list of applications and papers.

survey questions. After 2030 we assume that spending drops to half its initial value, i.e. $r/2$, and remains at that level until 2100. This drop occurs because we assume that post-2030 funds represent the government's commitment to maintain the technological gains achieved by 2030. We now derive the discounted opportunity cost of such expenditure streams of solar-technology R&D spending. In doing so we follow Popp [19] and assume that, at every time period, this opportunity cost is equal to 4 times the original investment. Thus, in our model the opportunity cost of a level of R&D investment r is given by the net present value of the stream $O(t)$ where $O(t) = 4 \cdot r$ for $t = 1, 2, ..., 6$ and $2 \cdot r$ for $t = 7, 8, ..., 20$ (once again we use a 3% discount rate). Table 3 summarizes these results for the three R&D Scenarios that the ICARUS survey focused on.

[TABLE 3 here]

Application of the decision-theoretic framework. We now extend our analysis to explicitly account for aggregation ambiguity and adopt the decision-theoretic model introduced in Section 2.[13] Our objective is to compare the three R&D Scenarios, and we do not consider optimizing over a continuous R&D domain $\mathcal{R}$. We make this choice primarily because we wish to keep the applied section brief and pursue more in-depth empirical analysis in future work.[14]

Figure 3 plots $V_{max}(r, b^2)$ and $V_{min}(r, b^2)$ over $b \in \left[0, \sqrt{\frac{13}{14}}\right] \approx [0, .96]$ for the three Scenarios. The parametrization b^2 is adopted since it allows us to (a) dampen the curvature of the original functions as given by Theorem 2 and (b) interpret the parameter b as a bound on the Euclidean distance of admissible aggregation schemes with respect to the benchmark equal-weight aggregation.

[FIGURE 3 here]

Focusing first on Scenario 1, we note that pure aggregation of expert opinion (corresponding to $b = 0$) yields a payoff of approximately $\$1.36 \times 10^9$. We observe that the worst-case payoff drops to about $\$-3.4 \times 10^9$ at $b \approx .25$ at which point it largely stops being sensitive to changes in b, slowly asymptoting to its minimum value of $\$-3.67 \times 10^9$; in contrast, the best-case one increases steadily to a maximum value of $\$22.7 \times 10^9$ at the maximum level of $b = .96$. Under Scenario 2, the payoff under zero ambiguity is equal to $\$7.8 \times 10^9$. Subequently, we see that the worst-case payoff drops to 0 at $b = 0.15$, at which point it keeps decreasing at a smaller rate until it practically reaches its minimum value of $\$-5.5 \times 10^9$ at $b \approx .55$. Conversely, the best-case payoff rises steadily to about $\$32 \times 10^9$ for $b \approx .55$ at which point it continues to rise at a much smaller rate until it reaches a maximum value of $\$33.3 \times 10^9$ at $b = .96$. Thus for Scenario 2, aggregation uncertainty becomes

[13] All simulations are performed in Mathematica.

[14] Indeed, constructing plausible approximations of experts' $u_n(r)$ functions over an interesting range of r will likely require further engagement with the experts.

largely unimportant once b reaches the threshold of 0.55. Under Scenario 3 the unambiguous payoff is around $\$20 \times 10^9$, significantly higher than both other Scenarios. The worst-case payoff drops relatively smoothly to a minimum value of $\$-7.35 \times 10^9$ for $b = .96$, while the best-case one rises at a comparatively higher rate to $\$70.9 \times 10^9$.

It is clear that aggregation ambiguity is important under Scenario 3, for both the worst- and best-case payoffs, significantly more so than under Scenarios 1 and 2. This fact is interesting in light of Figure 1, which shows that experts' pdfs are much more dispersed under Scenarios 1 and 2 than they are under 3. The reason behind this seemingly unexpected result is straightforward. As Table 1 suggests, expected payoffs of R&D investment are very sensitive at low cost values, i.e., less than 8c\$/kWh. The more aggressive investment of Scenario 3 has a greater effect on these lower cost values, and therefore its best- and worst-case payoffs are in turn more sensitive to changes in b.

We now consider the effect of ambiguity attitude on the decision maker's problem. Figure 4 plots the value function $V\left(r|b^2,\alpha\right)$ given by Eq. (7) for all three investment Scenarios, over all levels of aggregation ambiguity and a decision-maker's attitude toward it. This allows policy makers to visualize the effects of the three R&D investment decisions over the entire range of possible ambiguity levels and ambiguity attitudes. As we expect from Figure 3, Scenario 3 fares much better than both 1 and 2 over a very wide range of b and α, and is much more sensitive to changes in both.

[FIGURE 4 here]

Figure 5 goes a step further and compares the three R&D Scenarios for all possible combinations of b and α. Following the color scheme of Figure 3, a region's color corresponds to the Scenario that performs the best within it, while the bold numbers within regions denote the relative order of the three Scenarios within this range of (b, α) (e.g., an expression "321" means Scenario 2 dominates 1, and Scenario 3 dominates both 2 and 1).

[FIGURE 5 here]

Figure 5 makes clear that Scenario 3 dominates 1 and 2 for an extremely wide range of combinations of b and α. Conversely, Scenario 1 is the best option for a combination of very high b and α. Somewhat surprisingly, we see that Scenario 2 is dominated by either 1 or 3 for all possible combinations of b and α and thus will never be chosen by a decision maker whose preferences are captured by Eq. (7). Thus, on the basis of the presented data, it is clear that policy makers should opt for the most aggressive R&D investment, unless they are both (a) open to ignoring a very large set of surveyed experts (b) extremely concerned about the possibility of worst-case failure. Moreover, assuming all three options are readily implementable, they can safely disregard the middle-range R&D investment implied by Scenario 2.

5 Conclusions and Directions for Future Work

Structured expert surveys can play an important role in assessing the potential of uncertain investments. If designed well, they may be able to capture in a transparent and objective way subjective probabilities that can subsequently be used as scientific data in the decision maker's deliberations.

Yet, gathered information can vary substantially across experts. In particular, if the elicitation is designed correctly it should exactly aim at covering all prevailing "visions" about the specific investment. The different backgrounds and perspectives that experts bring to the elicitation process imply that collected subjective probability distributions will, more often than not, span a wide and potentially confusing spectrum.

Condensing all of the problem's uncertainty into one single average probability distribution may, especially in cases where standard aggregation methods cannot be readily applied, conceal important information and yield policy recommendations that are not robust. To deal with this issue, we proposed and analyzed a novel decision-theoretic framework inspired by the well-studied α-maxmin model. In line with the paper's focus on R&D investment, decision variables in our model affect experts' subjective probability distributions of the future cost-effectiveness of an investment. We applied our framework to original data from a recent expert elicitation survey on solar technology. The analysis suggested that more aggressive investment in solar technology R&D is likely to yield substantial benefits even after ambiguity over expert opinion has been taken into account.

Our work suggests several fruitful avenues for future research. A particularly challenging one would be to extend the model to take into account meaningful nonlinear functions of higher moments of the considered mixture distributions. For example, settings where V_{max} would be equal to an expression like $\max_{\pi \in \Pi(b,r)} \frac{\mathbf{E}_{\pi}[u(r)]}{\sqrt{\mathbf{Var}_{\pi}[u(r)]}}$ and V_{min} its minimization analogue. This change would introduce nonlinearities that significantly complicate the analysis of Section 3. Alternatively, one could keep the current framework and take into account issues of dependence and bias across experts by modifying the constraint implied by the set $\mathrm{P}(b)$. On the applied front, extensions of this work would delve deeper into the ICARUS survey data to obtain more data points and ultimately solve for optimal R&D investment over a continuous domain $\mathcal{R}$.

Appendix

A1: Proofs

To ease notation, in our proofs we suppress dependence on R&D investment r except where necessary.

Proposition 1. We prove the Proposition for $V_{max}(b)$ (the argument for $V_{min}(b)$ is analogous). That $V_{max}(b)$ is increasing in b follows by definition. Consider the optimization problems given by the right-hand-side of Eq. (11) for $b_1 \in [0, \frac{N-1}{N}]$ and $b_2 \geq b_1$ and denote their optimal solutions by $\mathbf{p^{max}}(b_1)$ and $\mathbf{p^{max}}(b_2)$ respectively. By feasibility we may note the following:

$$\sum_{n=1}^{N}\left(p_n^{max}(b_1)-\frac{1}{N}\right)^2 \leq b_1, \quad \sum_{n=1}^{N}\left(p_n^{max}(b_2)-\frac{1}{N}\right)^2 \leq b_2. \tag{22}$$

Consider a convex combination of b_1 and b_2 given by $b(\lambda)=\lambda b_1+(1-\lambda)b_2$ for some $\lambda \in [0,1]$ and the optimization problem

$$V_{max}(b(\lambda)) = \max_{\mathbf{p}\in \mathrm{P}(b(\lambda))} \sum_{n=1}^{N} p_n m_n. \tag{23}$$

To prove concavity of V_{max} in b it suffices to show that

$$V_{max}(b(\lambda)) \geq \lambda V_{max}(b_1) + (1-\lambda)V_{max}(b_2).$$

To this end, consider the probability vector given by

$$\mathbf{p}(\lambda) = \lambda \mathbf{p^{max}}(b_1) + (1-\lambda)\mathbf{p^{max}}(b_2).$$

By feasibility of $\mathbf{p^{max}}(b_1)$ and $\mathbf{p^{max}}(b_2)$ we immediately deduce that $\mathbf{p}(\lambda) \geq \mathbf{0}$ and that $\sum_{n=1}^{N} p_n(\lambda) = 1$. Now we may write

$$\begin{aligned}
\sum_{n=1}^{N}\left(p_n(\lambda)-\frac{1}{N}\right)^2 &= \sum_{n=1}^{N}\left(\lambda\left(p_n^{max}(b_1)-\frac{1}{N}\right)+(1-\lambda)\left(p_n^{max}(b_2)-\frac{1}{N}\right)\right)^2 \\
&\overset{\text{triangle ineq.}}{\leq} \left[\lambda\left(\sum_{n=1}^{N}\left(p_n^{max}(b_1)-\frac{1}{N}\right)^2\right)^{\frac{1}{2}}+(1-\lambda)\left(\sum_{n=1}^{N}\left(p_n^{max}(b_2)-\frac{1}{N}\right)^2\right)^{\frac{1}{2}}\right]^2 \\
&\overset{(22)}{\leq} \left[\lambda\sqrt{b_1}+(1-\lambda)\sqrt{b_2}\right]^2 \leq \left[\sqrt{\lambda b_1+(1-\lambda)b_2}\right]^2 = b(\lambda).
\end{aligned} \tag{24}$$

By Eq. (24) and the observations immediately preceding it we can conclude that $\mathbf{p}(\lambda)$ is feasible for optimization problem (23). Thus we may write

$$\begin{aligned}
V_{max}(b(\lambda)) \geq \sum_{n=1}^{N} p(\lambda)_n u_n &= \lambda\sum_{n=1}^{N} p_n^{max}(b_1)u_n + (1-\lambda)\sum_{n=1}^{N} p_n^{max}(b_2)u_n \\
&= \lambda V_{max}(b_1) + (1-\lambda)V_{max}(b_2),
\end{aligned}$$

where the last equality follows from the assumed optimality of $\mathbf{p^{max}}(b_1)$ and $\mathbf{p^{max}}(b_2)$. We now proceed to show continuity. By concavity $V_{max}(b)$ will be continuous on the open interval $(0, \frac{N-1}{N})$ so we need only consider the endpoints 0 and $\frac{N-1}{N}$. Since $V_{max}(b)$ is increasing in b we must have $\lim_{b\to(\frac{N-1}{N})^-} V_{max}(b) \leq V_{max}(\frac{N-1}{N})$. However, if $\lim_{b\to(\frac{N-1}{N})^-} V_{max}(b) < V_{max}(\frac{N-1}{N})$ then we reach a contradiction if we apply concavity to $(N-1)/N$ and other values of b.

To prove continuity at $b = 0$ consider an $\epsilon > 0$. Now let $\delta > 0$ and write

$$\begin{aligned}
|V_{max}(\delta) - V_{max}(0)| &= V_{max}(\delta) - V_{max}(0) = \sum_{n=1}^{N} \left(p_n^{max}(\delta) - \frac{1}{N}\right) u_n \\
&\leq \max_{n\in\mathcal{N}} |u_n| \sum_{n=1}^{N} \left|p_n^{max}(\delta) - \frac{1}{N}\right| \\
&\overset{\text{Hölder's ineq.}}{\leq} \max_{n\in\mathcal{N}} |u_n| \left[\sum_{n=1}^{N} \left(p_n^{max}(\delta) - \frac{1}{N}\right)^2\right]^{\frac{1}{2}} \leq \max_{n\in\mathcal{N}} |u_n| \sqrt{\delta}.
\end{aligned}$$

Thus, any choice of $0 < \delta < \frac{\epsilon^2}{(\max_{n\in\mathcal{N}} |u_n|)^2}$ will ensure that $|V_{max}(\delta) - V_{max}(0)| < \epsilon$, completing the proof. ■

Lemma 1. The function $V_{max}(b)$ is bounded above by u_n for any $n \in \mathcal{N}_{N(r)}$. This upper bound is attained by a probability vector $\mathbf{p}$ if and only if it satisfies

$$\sum_{n\in\widehat{\mathcal{N}}} p_n = 1, \quad \text{for some } \widehat{\mathcal{N}} \subseteq \mathcal{N}_{N(r)}$$

Consider a subset $\widehat{\mathcal{N}} \subseteq \mathcal{N}_{N(r)}$, with cardinality $\widehat{N}$. Eq. (8) implies that the value of b at which it first becomes possible to assign probability 1 to subset $\widehat{\mathcal{N}}$ is given by

$$b(\widehat{N}; N) = \frac{1}{\widehat{N}} - \frac{1}{N}.$$

The minimizer of $b(\widehat{N}; N)$ over $\widehat{\mathcal{N}} \subseteq \mathcal{N}_{N(r)}$ is the entire set $\mathcal{N}_{N(r)}$, yielding the desired result.

Now consider $b < b^*_{max}$ and the optimal solution $\mathbf{p^{max}}(b)$. As $b < b^*_{max}$ there must exist a $j \neq \mathcal{N}_{N(r)}$ such that $p_j^{max}(b) > 0$. Now consider increasing b by an amount ϵ. For $\delta > 0$ small enough the solution $\tilde{\mathbf{p}}$ which is identical to $\mathbf{p^{max}}(b)$ except that $\tilde{p}_j = p_j^{max}(b) - \delta$ and $\tilde{p}_k = p_k^{max}(b) + \delta$ for some $k \in \mathcal{N}_{N(r)}$ will be feasible and result in a strictly greater objective value, so that $V_{max}(b + \epsilon) > V_{max}(b)$. Equivalent reasoning applies to the V_{min} case. ■

Proposition 2. Suppose first that $b = b^*_{max}$. It is clear here that the unique optimal solution is given by $\mathbf{p^{max}}$ such that $p_n^{max} = 1/N_{N(r)}$ for all $n \in \mathcal{N}_{N(r)}$ and $p_n^{max} = 0$ otherwise. The quadratic ambiguity constraint binds by the definition of b^*_{max}.

Consider now the case $b < b^*_{max}$ and suppose there exists an optimal solution $\mathbf{p^{max}}(b)$ such that the quadratic ambiguity constraint is slack. As $b < b^*_{max}$ there must exist an $j \neq \mathcal{N}_{N(r)}$ such that $p_j^{max}(b) > 0$. For $\epsilon > 0$ small enough the solution $\tilde{\mathbf{p}}^{\mathbf{max}}$ in which $\tilde{p}_j = p_j^{max}(b) - \epsilon$ and $\tilde{p}_k = p_k^{max}(b) + \epsilon$ for some $k \in \mathcal{N}_{N(r)}$ will be feasible and result in a strictly greater objective value, contradicting $\mathbf{p^{max}}(b)$'s optimality. Thus, all optimal solutions must satisfy the quadratic ambiguity constraint with equality.

We now prove uniqueness. Suppose there exist two optimal solutions $\mathbf{p^{max,1}}$ and $\mathbf{p^{max,2}}$. By the preceding argument they must bind the quadratic ambiguity constraint. Consider the set of probability vectors given by their convex combinations

$$\mathbf{p}(\lambda) = \lambda \mathbf{p^{max,1}} + (1-\lambda)\mathbf{p^{max,2}}, \ \ \lambda \in [0,1].$$

For $\lambda \in (0,1)$, $\mathbf{p}(\lambda)$ will satisfy the ambiguity constraint with strict inequality, since:

$$\begin{aligned}
\sum_{n=1}^{N}\left(p_n(\lambda) - \frac{1}{N}\right)^2 &= \sum_{n=1}^{N}\left(\lambda\left(p_n^{max,1} - \frac{1}{N}\right) + (1-\lambda)\left(p_n^{max,2} - \frac{1}{N}\right)\right)^2 \\
&\overset{\text{strict convexity}}{<} \sum_{n=1}^{N}\left[\lambda\left(p_n^{max,1} - \frac{1}{N}\right)^2 + (1-\lambda)\left(p_n^{max,2} - \frac{1}{N}\right)^2\right] \\
&= \lambda\sum_{n=1}^{N}\left(p_n^{max,1} - \frac{1}{N}\right)^2 + (1-\lambda)\sum_{n=1}^{N}\left(p_n^{max,2} - \frac{1}{N}\right)^2 \\
&= \lambda b + (1-\lambda) b = b.
\end{aligned}$$

Thus all solutions $\mathbf{p}(\lambda)$ are feasible. That they are optimal follows trivially by the assumed optimality of $\mathbf{p^{max,1}}, \mathbf{p^{max,2}}$ and the linear objective function of (11). But this is a contradiction as all optimal solutions must satisfy the quadratic ambiguity constraint with equality. The second claim of the Proposition regarding $b > b_{max}$ is trivial. ■

Theorem 1. We prove the result for V_{max}; the argument for V_{min} is analogous. To do so we need to invoke results from conic duality. We begin with part (a). Given $\mathbf{x} = (x_0, \bar{\mathbf{x}}) \in \Re^{n+1}$ we introduce the following notation to denote inclusion in a second-order cone of dimension $n+1$

$$(x_0, \bar{\mathbf{x}}) \in \mathcal{L}^2_{n+1} \ \ \Leftrightarrow x_0 \geq ||\bar{\mathbf{x}}||_2.$$

We follow Alizadeh and Goldfarb [1] to write (11) as a primal conic program $\mathcal{P}(b)$ and introduce its dual $\mathcal{D}(b)$ (for clarity, next to the primal constraints we indicate the corresponding dual variables):

$$
\begin{array}{llll}
\max\limits_{\mathbf{p},\mathbf{q},q_0,\theta} & \sum\limits_{n=1}^{N} u_n p_n & \min\limits_{\mathbf{y},y_0,\boldsymbol{\gamma},\beta_0,\mathbf{z_p},\mathbf{z_q},z_{q0},\mathbf{z}_\theta} & y_0 + \sqrt{b+\frac{1}{N}}\beta_0 \\
\text{s.t.} & -p_n + q_n = 0, \;\; \forall n \in \mathcal{N}, \;\; (y_n) & \text{s.t.} & -y_0 - y_n + z_{pn} = -u_n, \;\; \forall n \in \mathcal{N} \\
& \sum\limits_{n=1}^{N} -p_n = -1, \;\; (y_0) & & y_n + z_{qn} = 0, \;\; \forall n \in \mathcal{N} \\
& & & \gamma_n + z_{\theta n} = 0, \;\; \forall n \in \mathcal{N} \\
\mathcal{P}(b) & \theta_n = 0, \;\; \forall n \in \mathcal{N}, \;\; (\gamma_n) & \mathcal{D}(b) & -\beta_0 + z_{q0} = 0 \\
& q_0 = \sqrt{b+\frac{1}{N}}, \;\; (\beta_0) & & (z_{pn}, z_{\theta n}) \in \mathcal{L}_2^2, \;\; \forall n \in \mathcal{N} \\
& (p_n, \theta_n) \in \mathcal{L}_2^2, \;\; \forall n \in \mathcal{N} & & (z_{q0}, -\mathbf{z_q}) \in \mathcal{L}_{n+1}^2. \\
& (q_0, \mathbf{q}) \in \mathcal{L}_{n+1}^2, & &
\end{array}
$$

Since both the primal and the dual have feasible strictly interior solutions, strong duality holds (see Theorem 13 of [1]). Without loss of generality, we can immediately simplify $\mathcal{D}(b)$ by setting $\mathbf{z}_\theta = \boldsymbol{\gamma} = \mathbf{0}$ and $\mathbf{z_p} \geq \mathbf{0}$. Correspondingly, we can eliminate the variable $\mathbf{z_q}$ by replacing it with $-\mathbf{y}$. Finally, it is evident that at optimality the quadratic constraint of the dual will be binding so that $z_{q0}^* = \beta_0^* = \sqrt{\sum_{n=1}^N (-y_n)^2} = \sqrt{\sum_{n=1}^N y_n^2}$. Collecting all of these observations we may re-write the dual in the following much simpler way:

$$
\begin{aligned}
\mathcal{D}_1(b) = \min_{\mathbf{y},y_0} \quad & y_0 + \sqrt{b+\frac{1}{N}}\sqrt{\sum_{n=1}^{N} y_n^2} \\
\text{s.t.} \quad & -u_n + y_0 + y_n \geq 0, \;\; n = 1, 2, ..., N.
\end{aligned} \tag{25}
$$

Examining (25) we deduce that at optimality $y_n^* = \max(0, u_n - y_0)$. Thus we may simplify the dual even further to an unconstrained optimization problem with just one variable:

$$
\mathcal{D}_2(b) = \min_{y_0} y_0 + \sqrt{b+\frac{1}{N}}\sqrt{\sum_{n=1}^{N} \max(0, u_n - y_0)^2}. \tag{26}
$$

By strong duality the dual optimal objective will be bounded between $\frac{1}{N}\sum_{n=1}^N u_n$ and $u_{(N(r))}$. We immediately see that solutions satisfying $y_0 > u_{(N(r))}$ result in strictly greater objective function values than $y_0 = u_{(N(r))}$, so that we can safely disregard them. Conversely, solutions satisfying $y_0 < 0$ yield

$$
\begin{aligned}
y_0 + \sqrt{b+\frac{1}{N}}\sqrt{\sum_{n=1}^{N} \max(0, u_n - y_0)^2} &= y_0 + \sqrt{b+\frac{1}{N}}\sqrt{\sum_{n=1}^{N} (u_n - y_0)^2} \\
&> y_0 + \sqrt{Nb+1}(|u_{(1)}| - y_0) = \sqrt{Nb+1}|u_{(1)}| + y_0(1 - \sqrt{Nb+1}).
\end{aligned}
$$

Thus, values of $y_0 < \frac{|u_{(N(r)}|}{1-\sqrt{Nb+1}}$ result in a strictly greater objective function value than $y_0 = u_{(N(r))}$ and hence can also be disregarded. With these observations we may rewrite the dual (26) in the following way:

$$\mathcal{D}_3(b) = \min_{y_0 \in \left[\frac{|u_{(N(r))}|}{1-\sqrt{Nb+1}}, u_{(N(r)}\right]} y_0 + \sqrt{b + \frac{1}{N}} \sqrt{\sum_{n=1}^{N} \max(0, u_n - y_0)^2} \tag{27}$$

The domain of $\mathcal{D}_3(b)$ is thus compact, for any $b > 0$. For values of $b \in [0, b^*_{max})$ we know that the optimal solution of the primal will be strictly less than $u_{(N(r))}$. Thus, strong duality implies that for all $b \in (0, b^*_{max})$, any optimal solution $y^*(b)$ must satisfy $y^*(b) < u_{(N(r)-1)}$. However, notice that the objective function of $\mathcal{D}_3$ is strictly convex for $y_0 < u_{(N(r)-1)}$. Thus, we may deduce that when $b \in (0, b^*_{max})$ $\mathcal{D}_3(b)$ admits a unique optimal solution $y^*_0(b)$.

The above observation implies that we can apply Danskin's theorem (see Proposition B.25 in Bertsekas [3]) to conclude that the optimal dual objective value, and therefore by strong duality $V_{max}(b)$ as well, is differentiable at all $b \in (0, b^*_{max})$ and that

$$\frac{\mathrm{d}V_{max}}{\mathrm{d}b}(b) = \frac{\sqrt{\sum_{n=1}^{N} \max(0, u_n - y^*_0(b))^2}}{2\sqrt{b + \frac{1}{N}}}, \quad b \in (0, b^*_{max}). \tag{28}$$

Before we proceed with investigating the endpoints $b = 0$ and b_{max}, we show that $y^*_0(b)$ is strictly increasing in $b \in (0, b^*_{max})$. Consider $b_1 < b_2$ with both belonging in $(0, b^*_{max})$ and their optimal solutions $y^*_0(b_1)$ and $y^*_0(b_2)$. By uniqueness of $y^*_0(b)$ in this range of b we have

$$y^*_0(b_1) + \sqrt{b_1 + \frac{1}{N}} \sqrt{\sum_{n=1}^{N} \max(0, u_n - y^*_0(b_1))^2} < y^*_0(b_2) + \sqrt{b_1 + \frac{1}{N}} \sqrt{\sum_{n=1}^{N} \max(0, u_n - y^*_0(b_2))^2}$$
$$y^*_0(b_2) + \sqrt{b_2 + \frac{1}{N}} \sqrt{\sum_{n=1}^{N} \max(0, u_n - y^*_0(b_2))^2} < y^*_0(b_1) + \sqrt{b_2 + \frac{1}{N}} \sqrt{\sum_{n=1}^{N} \max(0, u_n - y^*_0(b_1))^2}.$$

Summing the above inequalities and rearranging terms yields

$$\left(\sqrt{b_2 + \frac{1}{N}} - \sqrt{b_1 + \frac{1}{N}}\right) \left(\sqrt{\sum_{n=1}^{N} \max(0, u_n - y^*_0(b_1))^2} - \sqrt{\sum_{n=1}^{N} \max(0, u_n - y^*_0(b_2))^2}\right) > 0$$
$$\Rightarrow \sqrt{\sum_{n=1}^{N} \max(0, u_n - y^*_0(b_1))^2} - \sqrt{\sum_{n=1}^{N} \max(0, u_n - y^*_0(b_2))^2} > 0 \Rightarrow y^*_0(b_2) > y^*_0(b_1).$$

We discuss now the differentiability of V_{max} at $b \in \{0, b^*_{max}\}$. At $b = 0$ the domain of (27) is no longer bounded below and therefore we can no longer invoke Danskin's theorem. Consequently, we reason in a different way. By continuity (recall Proposition 1) we must have

$$\lim_{b \to 0^+} V_{max}(b) = \sum_{n=1}^{N} \frac{u_n}{N} \Leftrightarrow \lim_{b \to 0^+} y^*(b) + \sqrt{\frac{1}{N}} \sqrt{\sum_{n=1}^{N} \max(0, u_n - \lim_{b \to 0^+} y^*(b))^2} = \sum_{n=1}^{N} \frac{u_n}{N}$$

The strict monotonicity of $y^*(b)$ and Hölder's inequality imply that $\lim_{b\to 0^+} y^*(b) < 0$. Subsequently, simple algebra obtains:

$$\lim_{b\to 0^+} y^*(b) - \frac{\sum_{n=1}^N u_n}{N} = -\sqrt{\frac{1}{N}\sum_{n=1}^N (u_n - \lim_{b\to 0^+} y^*(b))^2}$$

If we take squares now on both sides and re-apply Hölder's inequality, we see that

$$\lim_{b\to 0^+} y^*(b) = -\infty \quad \Rightarrow \quad \lim_{b\to 0^+} \frac{\mathrm{d}V_{max}}{\mathrm{d}b}(b) = +\infty.$$

Now we consider $b = b^*_{max}$. Note that the optimal solution $y^*_0(b^*_{max})$ is not unique; instead it can take any value in the interval $[u_{(N(r)-1)}, u_{(N(r))}]$. Hence Danskin's theorem implies that the subdifferential of $V_{max}(b)$ at b^*_{max} will consist of all convex combinations of $\frac{\sqrt{N_{(N(r))}}(u_{(N(r))}-u_{(N(r)-1)})}{2\sqrt{b^*_{max}+\frac{1}{N}}}$ and 0.

We now prove part (b). Let us go back to the original primal-dual pair $(\mathcal{P}(b), \mathcal{D}(b))$ and consider a pair of optimal solutions of the primal and dual problems. By Proposition 2 the primal optimal solution $(\mathbf{p}^*(b), \mathbf{q}^*(b), \boldsymbol{\theta}^*(b), q^*_0(b))$ is unique, while our reasoning in part (a) established the uniqueness of the optimal dual variables $(\beta^*_0(b), \mathbf{y}^*(b), y^*_0(b), \mathbf{z}^*_{\mathbf{p}}(b), \mathbf{z}^*_{\mathbf{q}}(b))$. Applying Theorem 16 and part (ii) of the complementarity conditions of Lemma 15 of Alizadeh and Goldfarb [1], we arrive at the following conditions:

$$q^*_0(b) z^*_{qn}(b) + \beta^*_0(b) q^*_n(b) = 0 \Leftrightarrow -\sqrt{b+\frac{1}{N}} y^*_n(b) + \sqrt{\sum_{n=1}^N y^*_n(b)^2} p^*_n(b) = 0, \quad n = 1,2,...,N$$

$$\Leftrightarrow -\sqrt{b+\frac{1}{N}}\ \max(0, u_n - y^*_0(b)) + \sqrt{\sum_{n=1}^N \max(0, u_n - y^*_0(b))^2}\ p^*_n(b) = 0, \quad n = 1,2,...,N \tag{29}$$

When $b < b^*_{max}$, strong duality implies $y^*_0(b) < u_{(N(r)-1)}$ which in turn ensures $\sum_{n=1}^N \max(0, u_n - y^*_0(b))^2 > 0$. As mentioned earlier, when $b = b^*_{max}$ $y^*_0(b)$ can take any value in $[u_{(N(r)-1)}, u_{(N(r))}]$ so we choose one that again yields $\sum_{n=1}^N \max(0, u_n - y^*_0(b^*_{max}))^2 > 0$. Hence, the complementarity conditions (29) yield

$$p^*_n(b) = 0 \Leftrightarrow u_n - y^*_0(b) \leq 0, \quad n = 1,2,...,N. \tag{30}$$

Since $y^*_0(b)$ is strictly increasing in b in $(0, b^*_{max})$ and $\lim_{b\to 0^+} y^*_0(b) = -\infty$ and $\lim_{b\to b^*_{max}-} y^*(b) = u_{(N(r)-1)}$, Eq. (30) implies the existence of a set $\{b_1, b_2, ..., b_{N(r)-1}\}$ such that

$$\begin{aligned} &0 < b_1 < b_2 < .... < b_{N(r)-1} = b^*_{max} \\ &\left\{ \left\{ p^*_n(b) = 0 \ \ \forall n \in \mathcal{N}^-_k \right\} \ \Leftrightarrow \ b \geq b_k \right\}, \ \ \forall k = 1,2,...,N(r)-1. \end{aligned}$$

■

Proposition 3. We prove the result for V_{max}; the argument for V_{min} is analogous. Focusing on optimization problem (11), we introduce Lagrangian multipliers and write the Karush-Kuhn-Tucker (KKT) conditions:

$$u_n - 2\lambda\left(p_n - \frac{1}{N}\right) + \mu + \nu_n = 0, \quad n \in \{1,2,...,N\} \tag{31}$$

$$\lambda\left(\sum_{n=1}^{N}\left(p_n - \frac{1}{N}\right)^2 - b\right) = 0, \quad \lambda \geq 0 \tag{32}$$

$$\sum_{n=1}^{N}\left(p_n - \frac{1}{N}\right)^2 \leq b, \quad \sum_{n=1}^{N} p_n = 1, \quad \mathbf{p} \geq \mathbf{0} \tag{33}$$

$$\nu_n p_n = 0, \quad \nu_n \geq 0, \quad n \in \{1,2,...,N\}. \tag{34}$$

Since our problem is concave with affine equality constraints and satisfies Slater's condition (see section 5.2.3 in [6]), strong duality holds and the KKT conditions (31)-(34) will be necessary and sufficient for both primal and dual optimality. In other words, the duality gap is zero and the vector $(\mathbf{p}^*, \boldsymbol{\nu}^*, \lambda^*, \mu^*)$ satisfies (31)-(34) if and only $\mathbf{p}^*$ and $\lambda^*, \boldsymbol{\nu}^*, \mu^*$ are primal and dual optimal respectively (see section 5.5.3 in [6]).

From Proposition 2 we know that there exists a unique primal optimal solution $\mathbf{p}^*$. By strong duality, the Lagrangean dual problem admits an optimal solution, and we refer to it by $\lambda^*, \boldsymbol{\nu}^*, \mu^*$.[15] Since $V_{max}(b)$ is differentiable (Theorem 1) and strong duality holds we follow Section 5.6.3 in Boyd and Vandenberghe [6] to deduce the following simple relation:

$$\frac{\mathrm{d}}{\mathrm{d}b} V_{max}(b) = \lambda^{max}(b), \quad b \in (0, b^*_{max}). \tag{35}$$

Eq. (35) means that we can now focus on calculating the Lagrange multiplier $\lambda^{max}(b)$. Before we do so we note the following useful identity

$$\begin{aligned}\sum_{n=1}^{N}\left(p_n^{max}(b) - \frac{1}{N}\right)^2 &= \sum_{n=1}^{N} p_n^{max}(b)\left(p_n^{max}(b) - \frac{1}{N}\right) - \frac{1}{N}\sum_{n=1}^{N} p_n^{max}(b) + \sum_{n=1}^{N}\frac{1}{N^2} \\ &= \sum_{n=1}^{N} p_n^{max}(b)\left(p_n^{max}(b) - \frac{1}{N}\right). \end{aligned} \tag{36}$$

Multiplying both sides of Eq. (31) by $p_n^{max}(b)$ and then summing over all $n = 1, 2, .., N$ obtains

$$\begin{aligned} &\sum_{n=1}^{N} u_n p_n^{max}(b) - 2\lambda^{max}(b)\sum_{n=1}^{N} p_n^{max}(b)\left(p_n^{max}(b) - \frac{1}{N}\right) + \mu^{max}(b)\sum_{n=1}^{N} p_n^{max}(b) = 0 \\ \overset{(36)}{\Rightarrow} \quad &\sum_{n=1}^{N} u_n p_n^{max}(b) - 2\lambda^{max}(b)\sum_{n=1}^{N}\left(p_n^{max}(b) - \frac{1}{N}\right)^2 + \mu^{max}(b) = 0 \end{aligned}$$

[15] Note that at this point one can manipulate the KKT conditions (31)-(34) to show that the Lagrangean dual's optimal solution is also unique.

$$\overset{\text{Prop. 2}}{\Rightarrow} \quad \mu^{max}(b) = 2\lambda^{max}(b) \cdot b - \sum_{n=1}^{N} u_n p_n^{max}(b). \tag{37}$$

Now we consider Eq. (31) for expert $n_k \in \mathcal{N}_k$. By part (b) of Theorem 1 we must have $p_{n_k}^{max}(b) > 0$ if and only if $b \in [0, b_k^{max})$. Substituting the value of $\mu^{max}(b)$ obtained in Eq. (37), and applying the complementary slackness condition (34) we obtain

$$\begin{aligned} & u_{n_k} - 2\lambda^{max}(b)\left(p_{n_k}^{max}(b) - \frac{1}{N}\right) = \sum_{n=1}^{N} u_n p_n^{max}(b) - 2\lambda^{max}(b) \cdot b \\ \overset{(35)}{\Rightarrow} \quad & 2\frac{\mathrm{d}}{\mathrm{d}b} V_{max}(b)\left(p_{n_k}^{max}(b) - \frac{1}{N} - b\right) = u_{n_k} - V_{max}(b), \quad b \in (0, b_k^{max}). \end{aligned} \tag{38}$$

■

Theorem 2. We focus on V_{max}; the argument for V_{min} is symmetric. Recall the definition of b_k^{max} of Eq. (13). Consider first $b \in (0, b_1^{max})$ so that $p_n^{max}(b) > 0$ for all $b \in (0, b_1^{max})$ and $n \in \mathcal{N}$. Recalling Proposition 3 and adding Eqs. (38) for all $n \in \mathcal{N}$ yields the following differential equation

$$-2Nb\frac{\mathrm{d}V_{max}(b)}{\mathrm{d}b} = -NV_{max}(b) + \sum_{n \in \mathcal{N}} u_n, \quad b \in (0, b_1^{max}). \tag{39}$$

Solving differential equation (39) leads to the following expression:

$$V_{max}(b) = C_1^{max}\sqrt{b} + \frac{\sum_{n \in \mathcal{N}} u_n}{N}, \quad b \in [0, b_1^{max}), \tag{40}$$

where C_1^{max} is a constant to be determined. Consider now $b \in [b_{k-1}^{max}, b_k^{max})$ for $k \in \{2, 3, ..., N(r) - 1\}$. In this range of b we will have $p_n^{max}(b) > 0$ if and only $n \in \mathcal{N}_k^+$. Adding Eqs. (38) for all such $n \in \mathcal{N}_k^+$ yields the following differential equation

$$2\left(\frac{N_{k-1}^-}{N} - N_k^+ b\right)\frac{\mathrm{d}V_{max}}{\mathrm{d}b} = \sum_{n \in \mathcal{N}_k^+} u_n - N_k^+ V_{max}(b), \quad b \in [b_{k-1}^{max}, b_k^{max}) \tag{41}$$

Solving differential equation (41) gives the following:

$$V_{max}(b) = \overline{u}_k^+ + C_k\sqrt{N_k^+ b - \frac{N_{k-1}^-}{N}}, \quad b \in [b_{k-1}^{max}, b_k^{max}), \tag{42}$$

for $k \in \{2, 3, ..., N(r) - 1\}$, where C_k^{max} is a constant to determined. Finally since $b_{N(r)-1}^{max} = b_{max}^*$ we use Lemma 1 to conclude

$$V_{max}(b) = \max_{n \in \mathcal{N}} u_n, \quad b \in \left[b_{N(r)-1}^{max}, \frac{N-1}{N}\right]. \tag{43}$$

Putting together Eqs. (40), (42), and (43) we see that V_{max} will equal

$$V_{max}(b) = \begin{cases} \frac{\sum_{n\in\mathcal{N}} u_n}{N} + C_1^{max}\sqrt{b} & b \in [0, b_1^{max}) \\ \overline{u}_k^+ + C_k^{max}\sqrt{N_k^+ b - \frac{N_{k-1}^-}{N}} & b \in \left[b_{k-1}^{max}, b_k^{max}\right),\ k = 2, 3, ..., N(r)-1 \\ \max_{n\in\mathcal{N}} u_n & b \in \left[b_{N(r)-1}^{max}, \frac{N-1}{N}\right] \end{cases} \tag{44}$$

for appropriately chosen constants $(C_1^{max}, C_2^{max}, ..., C_{N(r)-1}^{max})$ and $\left(b_1^{max}, b_2^{max}, ..., b_{N(r)-1}^{max}\right)$. By Proposition 1 and Theorem 1, V_{max} is continuous everywhere and differentiable everywhere at $(0, \frac{N-1}{N})$ except b^*_{max}. Thus, the vectors $(C_1^{max}, C_2^{max}, ..., C_{N(r)-1}^{max})$ and $\left(b_1^{max}, b_2^{max}, ..., b_{N(r)-1}^{max}\right)$ must fulfill these criteria of continuity and differentiability and are thus uniquely determined by the following system of nonlinear equations (45)-(52):

<u>Case 1</u>: $N(r) = 2$.

$$\frac{\sum_{n\in\mathcal{N}} u_n}{N} + C_1^{max}\sqrt{b_1^{max}} = \max_{n\in\mathcal{N}} u_n \tag{45}$$

$$b_1^{max} = \frac{1}{N_2^{max}} - \frac{1}{N}. \tag{46}$$

<u>Case 2</u>: $N(r) \geq 3$.

$$\frac{\sum_{n\in\mathcal{N}} u_n}{N} + C_1^{max}\sqrt{b_1^{max}} = \overline{u}_2^+ + C_2^{max}\sqrt{N_2^+ b_1^{max} - \frac{N_1^-}{N}} \tag{47}$$

$$\frac{C_1^{max}}{\sqrt{b_1^{max}}} = \frac{C_2^{max} N_2^+}{\sqrt{N_2^+ b_1^{max} - \frac{N_1^-}{N}}} \tag{48}$$

$$\overline{u}_k^+ + C_k^{max}\sqrt{N_k^+ b_k^{max} - \frac{N_{k-1}^-}{N}} = \overline{u}_{k+1}^+ + C_{k+1}^{max}\sqrt{N_{k+1}^+ b_k^{max} - \frac{N_k^-}{N}}, \quad k = 2, 3, ..., N(r)-2 \tag{49}$$

$$\frac{C_k^{max} N_k^+}{\sqrt{N_k^+ b_k^{max} - \frac{N_{k-1}^-}{N}}} = \frac{C_{k+1}^{max} N_{k+1}^+}{\sqrt{N_{k+1}^+ b_k^{max} - \frac{N_k^-}{N}}}, \quad k = 2, 3, ..., N(r)-2 \tag{50}$$

$$\overline{u}_{N(r)-1}^+ + C_{N(r)-1}^{max}\sqrt{N_{N(r)-1}^+ b_{N(r)-1}^{max} - \frac{N_{N(r)-2}^-}{N}} = \max_{n\in\mathcal{N}} u_n \tag{51}$$

$$b_{N(r)-1}^{max} = \frac{1}{N_{N(r)}^{max}} - \frac{1}{N} \tag{52}$$

It now remains to show that the solution of System (45)-(52) will eventually lead to the expression of the Theorem. To do this we calculate explicitly the C_k^{max} and b_k^{max}'s and then show how applying them to formula (44) yields the desired result.

We begin with Case 1 and $N(r)=2$. That

$$b_1^{max} = \frac{1}{N_2} - \frac{1}{N} \tag{53}$$

is trivially true. Then, Eq. (45) yields

$$C_1^{max} = \sqrt{\frac{N_2}{N_1} \cdot N} \cdot \left(u_{(2)} - \overline{u}_1^+\right) = \sqrt{\frac{N_1}{N_1^+} d_1^+}. \tag{54}$$

We now focus on Case 2 and $N(r) \geq 3$. Once again, we have by definition $b_{N(r)-1}^{max} = \frac{1}{N_{N(r)}} - \frac{1}{N}$, whence Eq. (51) implies

$$C_{N(r)-1}^{max} = \sqrt{\frac{N_{N(r)}}{N_{N(r)-1}}} \cdot \left(u_{(N(r))} - \overline{u}_{N(r)-1}^+\right) = \sqrt{\frac{N_{N(r)-1}}{N_{N(r)-1}^+} d_{N(r)-1}^+}.$$

Focusing on Eq. (50) for $k \in \{2, 3, ..., N(r)-2\}$ and solving for C_k^{max} yields:

$$C_k^{max} = \frac{\sqrt{N_k^+ b_k^{max} - \frac{N_{k-1}^-}{N}}}{N_k^+} \frac{C_{k+1}^{max} N_{k+1}^+}{\sqrt{N_{k+1}^+ b_k^{max} - \frac{N_k^-}{N}}}. \tag{55}$$

Plugging (55) into Eq. (49) we obtain:

$$C_{k+1}^{max} \sqrt{N_{k+1}^+ b_k^{max} - \frac{N_k^-}{N}} \left(1 - \frac{N_{k+1}^+}{N_k^+} \cdot \frac{N_k^+ b_k^{max} - \frac{N_{k-1}^-}{N}}{N_{k+1}^+ b_k^{max} - \frac{N_k^-}{N}}\right) = \bar{u}_k^+ - \bar{u}_{k+1}^+. \tag{56}$$

After some algebra, the left-hand-side of Eq. (56) can be simplified so that:

$$\begin{aligned}
\frac{C_{k+1}^{max} N_{k+1}^+}{N} \frac{\frac{N_{k-1}^-}{N_k^+} - \frac{N_k^-}{N_{k+1}^+}}{\sqrt{N_{k+1}^+ b_k^{max} - \frac{N_k^-}{N}}} &= \frac{\sum_{n \in \mathcal{N}_k^+} u_n}{N_k^+} - \frac{\sum_{n \in \mathcal{N}_{k+1}^+} u_n}{N_{k+1}^+} \\
\Rightarrow \qquad -C_{k+1}^{max} \frac{N_k}{N_k^+ \sqrt{N_{k+1}^+ b_k^{max} - \frac{N_k^-}{N}}} &= \frac{\left(u_{(k)} N_{k+1}^+ - \sum_{n \in \mathcal{N}_{k+1}^+} u_n\right) N_k}{N_k^+ N_{k+1}^+} \\
\Rightarrow \qquad -\frac{C_{k+1}^{max}}{\sqrt{N_{k+1}^+ b_k^{max} - \frac{N_k^-}{N}}} &= u_{(k)} - \overline{u}_{k+1}^+
\end{aligned} \tag{57}$$

Combining Eqs. (50) and (57) obtains for $k = 2, 3, ..., N(r)-2$:

$$C_k^{max} = -\frac{N_{k+1}^+}{N_k^+} \left(u_{(k)} - \overline{u}_{k+1}^+\right) \sqrt{N_k^+ b_k^{max} - \frac{N_{k-1}^-}{N}} \tag{58}$$

$$b_k^{max} = \frac{\left(\frac{C_{k+1}^{max}}{u_{(k)} - \overline{u}_{k+1}^+}\right)^2 + \frac{N_k^-}{N}}{N_{k+1}^+}, \tag{59}$$

which after some simple algebra leads to the following nonhomogeneous linear recursion for the squares of the C_k^{max}'s:

$$(C_k^{max})^2 = \frac{N_{k+1}^+}{N_k^+}\,(C_{k+1}^{max})^2 + \frac{N_k}{N_k^+}\frac{d_k^+}{N_k^+}, \quad k = 2, 3, ..., N(r) - 2. \tag{60}$$

Solving recursion (60) backwards with (previously derived) initial value $C_{N(r)-1}^{max}$, taking square roots, and recalling the positive sign of the C_k^{max}'s, leads to a simple expression for the C_k^{max}'s:

$$C_k^{max} = \sqrt{\frac{1}{N_k^+}\sum_{l=k}^{N(r)-1}\frac{N_l}{N_l^+}d_l^+}, \quad k = 2, 3, ..., N(r) - 1. \tag{61}$$

Applying Eq. (61) to Eq. (59) yields

$$b_k^{max} = \frac{N_k^-}{N_{k+1}^+ N} + \frac{\sum_{l=k+1}^{N(r)-1} d_l^+}{N_{k+1}^+ d_k^+}\,\frac{N_k}{N_k^+}, \quad k = 2, 3, ..., N(r) - 1. \tag{62}$$

Finally plugging C_2^{max} into Eqs. (47)-(48) yields

$$C_1^{max} = \sqrt{\sum_{l=1}^{N(r)-1}\frac{N_l}{N_l^+}d_l^+}, \tag{63}$$

$$b_1^{max} = \frac{N_1^-}{N_2^+ N} + \frac{\sum_{l=2}^{N(r)-1} d_l^+}{N_2^+ d_1^+}\,\frac{N_1}{N_1^+}. \tag{64}$$

Note that Eqs. (63)-(64) are consistent with the results for Case 1 as given by Eqs. (53)-(54). Thus there is no more need to distinguish between Case 1 and 2.

Finally, applying Eqs. (61)-(62)-(63)-(64) to Eq. (44) and performing elementary algebra establishes the result. ■

Corollary 1. The result follows by Propositions 2 and 3 and Theorems 1 and 2. ■

Theorem 3. Here we apply part (iii) of Corollary 4 in Milgrom and Segal [17] to functions $V_{max}(r|b)$ and $V_{min}(r|b)$ (we express the latter as a maximization problem). ■

Corollary 2. Follows by Proposition 2 and Theorem 3. ■

Proposition 4. Follows from Proposition 2 and Theorem 3. ■

Corollary 3. The statement of the Corollary implies that $\mathcal{N}_{N(r)}(r) = \{n_1\}$ and $\mathcal{N}_1(r) = \{n_2\}$ for all $r \in \mathcal{R}$. Hence, $b^*_{max}(r) = b^*_{min}(r) = \frac{N-1}{N}$ for all $r \in [r_m, r_M]$. Applying Corollary 2 establishes the result. ■

A2: Constructing expert pdfs for the three R&D Scenarios from ICARUS survey data

In the ICARUS survey, experts were asked to provide values for the 10th, 50th, and 90th percentile of their distributions for the 2030 cost of solar technology conditional on all three Scenarios. In addition, they were asked to provide values for the probability of this cost being less than or equal to the following three values: 11.3, 5.5, and 3c$/kWh. These "threshold" cost levels correspond to projections of the costs of electricity from fossil fuels or nuclear in 2030. The first (11.27 c$/kWh) corresponds to the 2030 projected cost of electricity from traditional coal power plants in the presence of a specific policy to control CO2 emissions (thus effectively doubling electricity costs from fossil sources). The second threshold cost (5.5 c$/kWh) is the projected cost of electricity from traditional fossil fuels in 2030, without considering any carbon tax. Finally, the third (3 c$/kWh) reflects a situation in which solar power becomes competitive with the levelized cost of electricity from nuclear power.

Asking experts the follow up question on the likelihood of reaching threshold cost targets allowed the survey authors to guard against the cognitive pitfalls associated with direct elicitation of subjective probabilities, to increase the amount of elicited information, and to deepen the discussion with the expert, hence improving their perception of his/her beliefs. In cases where the two sets of answers (percentile values and threshold probabilities) were inconsistent, we contacted the expert in order to obtain coherent estimates. Moreover, we asked all experts to give values for the upper and lower limits of their distribution's support in order to pinpoint the intervals over which their implied probability distributions range.

Such corrected estimates were obtained from 14 out of the original 16 experts, and therefore the analysis of Section 4 focuses solely on them. Among the respondents, not all provided values on the left and right endpoints of their distributions' support. As a result, we deduced between 6 and 8 points of 14 experts' cumulative distribution functions (cdf) of the 2030 cost of solar electricity, given the aforementioned three R&D investment Scenarios. From these points a probability distribution function (pdf) was constructed using linear interpolation in the following way. First of all, and in accordance with the experts' answers, we considered cost levels c lying in [2c$/kWh, 30c$/kWh] and discretized this interval on a scale of 0.5 (30c$/kWh represents an estimate of the technology's current cost). Now, suppose an expert reported the values of his/her cdf F_n at two successive points c_1 and c_2 where $c_2 > c_1$ and gave no further information on cost levels between c_1 and c_2. Assuming right-continuity of F_n we took the probability mass $F_n(c_2) - F_n(c_1)$ to be distributed uniformly among the cost levels $\{c_1 + .5, c_1 + 1, ..., c_2\}$. For experts who did not provide values for the lower limit of their distribution's support we assumed that whatever probability mass remained to be allocated (always less than .1) was distributed uniformly between the smallest argument of the cdf and two cost levels below it. For example, if an expert only indicated that c_l was his y'th percentile and gave no further points of the cdf below this, we assumed that a probability mass of y was distributed evenly across $\{c_l - 1, c_l - .5, c_l\}$. In the case of an unknown upper limit, if an expert only indicated that c_u was his yth percentile and gave no further arguments for the cdf above it, we assumed that a probability mass of $1-y$ was distributed evenly across $\{c_u+.5, c_u+1\}$.

Following this procedure we arrived at probability distribution functions for all 14 experts conditional on all three Scenarios. The implied cumulative distribution functions are depicted in Figure 2.

A3: Tables and Figures not in main text

2030 solar-technology cost c (2005 USc$/kWh)	Benefit $B(c)$ (US$ 10^9)
2	189.90
2.5	170.76
3	151.26
3.5	131.74
4	112.12
4.5	92.29
5	71.47
5.5	50.64
6	29.27
6.5	23.59
7	12.32
7.5	3.67
8	1.76
> 8	0

Table 2: EU discounted consumption improvement as a function of 2030 solar-power cost

R&D Scenario r	Opportunity Cost $O(r)$ (US$ 10^9)
r_1	3.67
r_2	5.51
r_3	7.35

Table 3: Discounted opportunity cost of R&D Scenarios

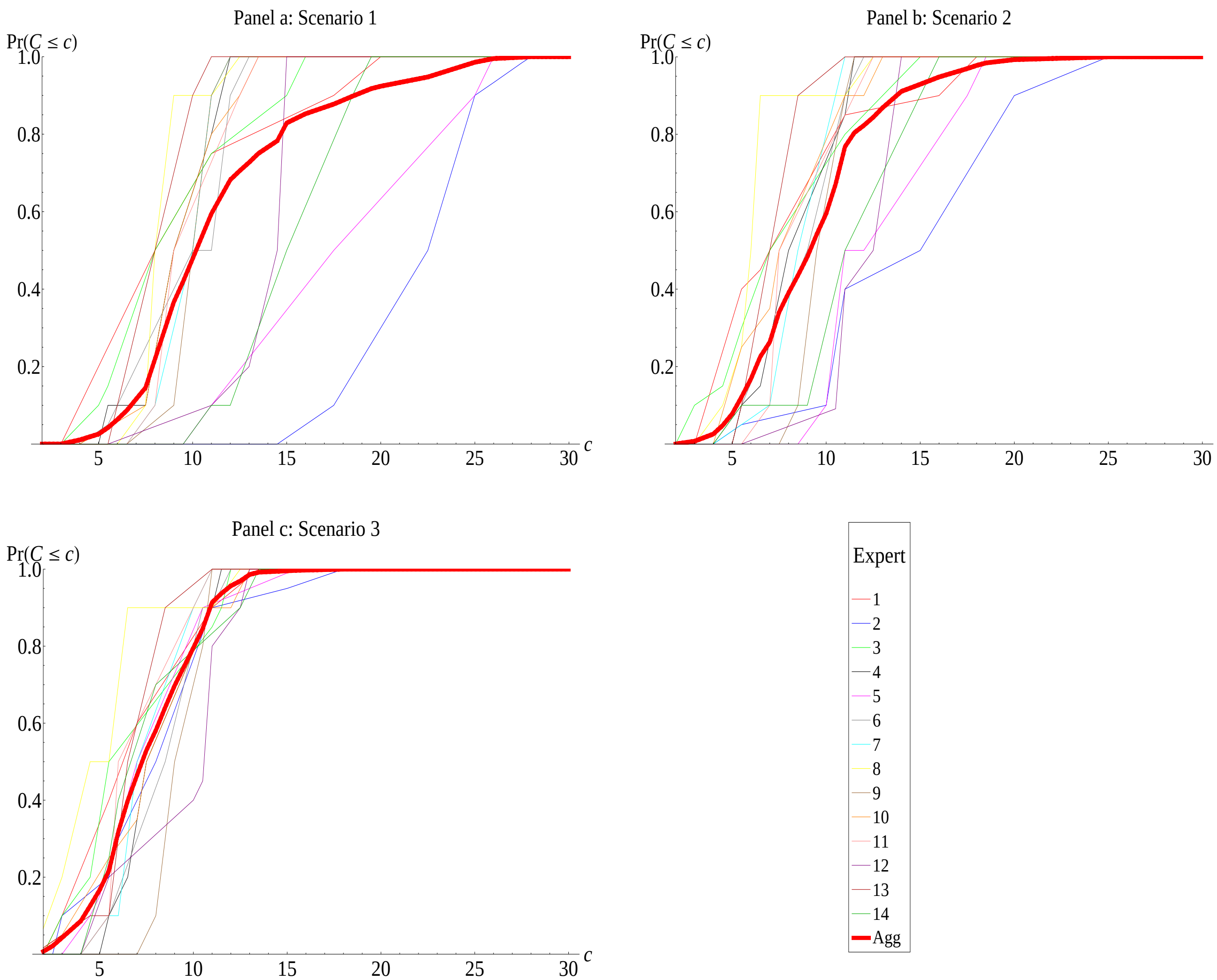

Figure 2: Expert and aggregate cdfs of the 2030 cost of solar technology under the three R&D Scenarios. Recall that the cdf's domain is $\{2, 2.5, ..., 29, 29.5, 30\}$. Cost is measured in 2005 USc\$/kWh.

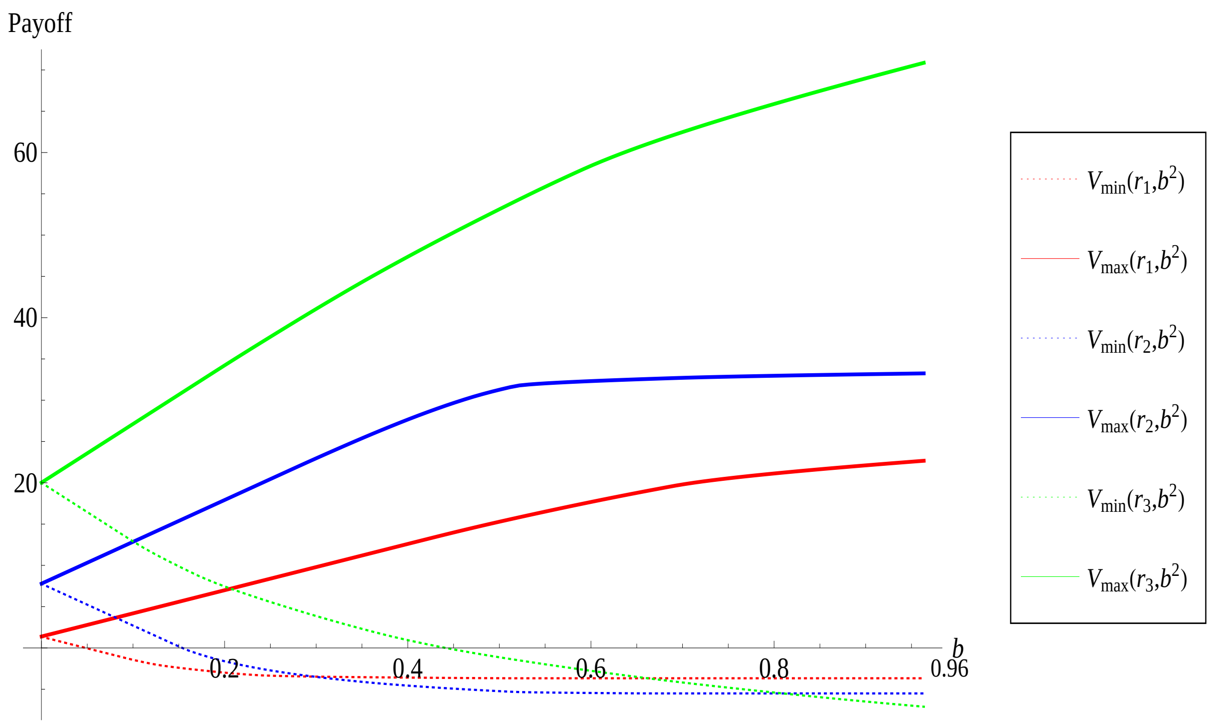

Figure 3: Worst and Best-Case net payoffs (benefits minus opportunity cost) for the three R&D scenarios. Net payoffs are measured in US$$10^9$.

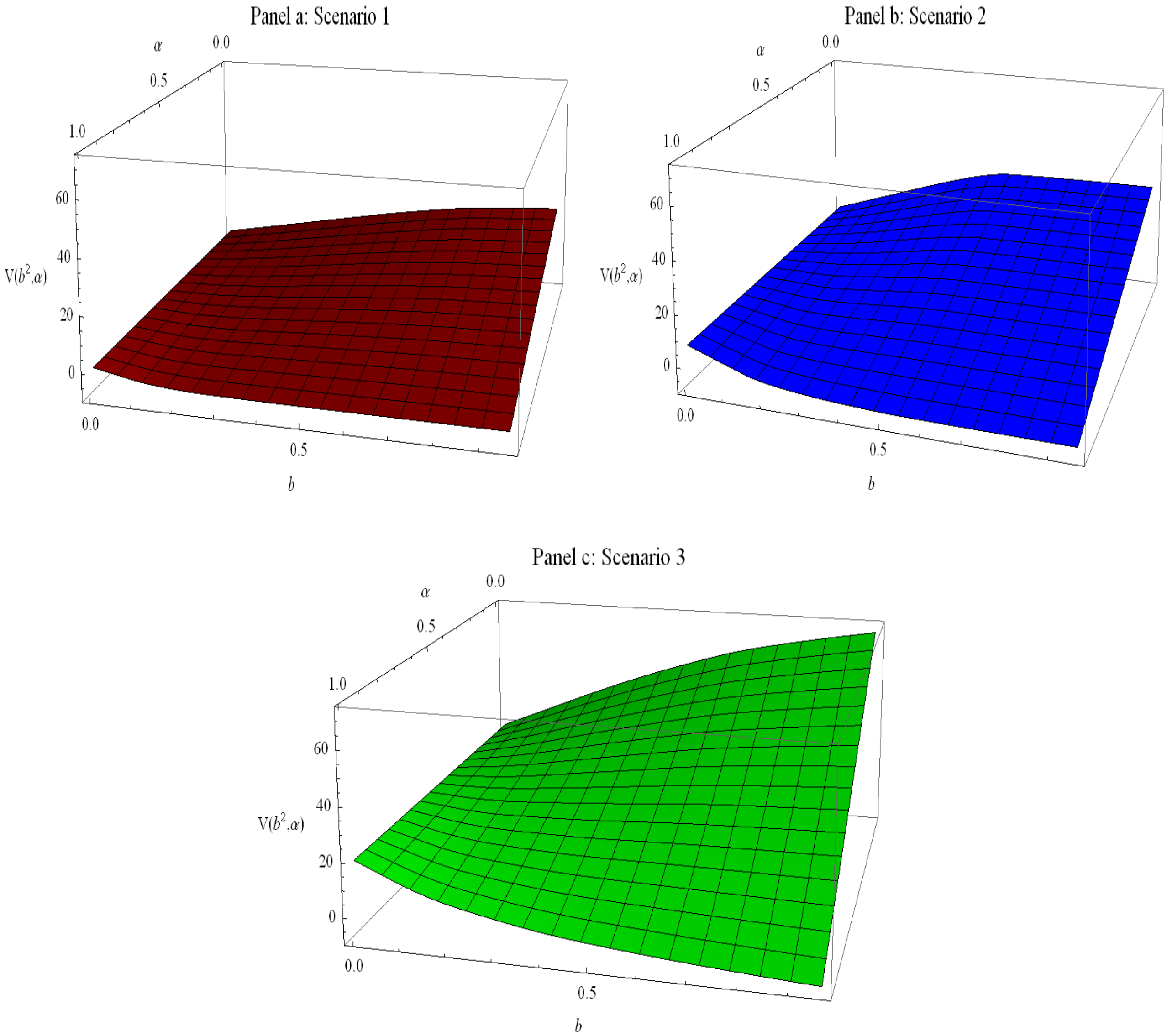

Figure 4: Net payoff (benefits minus opportunity cost) for the three R&D scenarios, as a function of ambiguity b and ambiguity attitude α. Net payoffs are measured in US\$$10^9$.

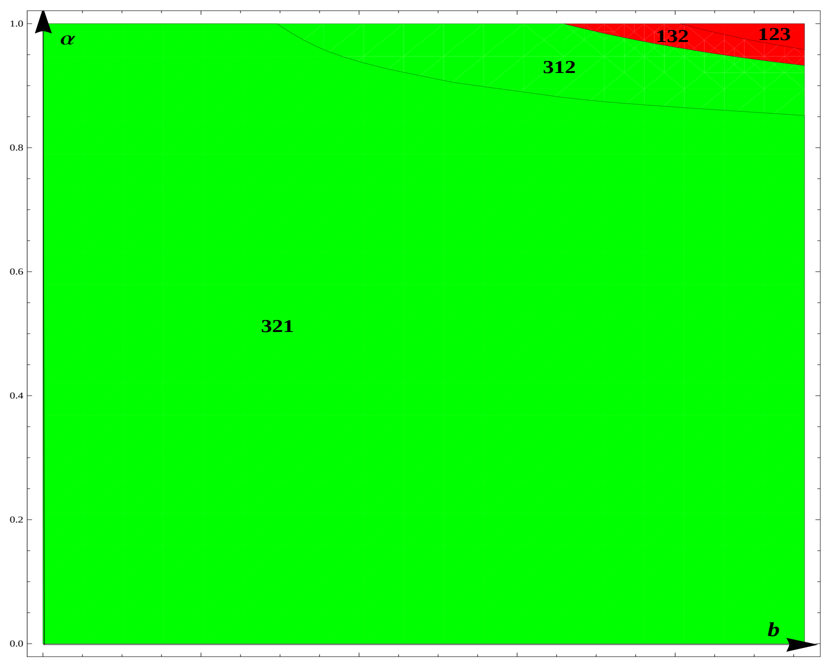

Figure 5: Comparison of the three R&D scenarios over all values of b and α.

References

[1] F. Alizadeh and D. Goldfarb (2003), "Second-order Cone Programming," *Mathematical Programming B*, 95, 3–51.

[2] K. Arrow and L. Hurwicz (1972), "An Optimality Criterion for Decision-Making under Ignorance," in C.F. Carter and L.J. Ford (eds.) *Uncertainty and Expectations in Economics: Essays in Honour of G.L.S. Shackle* (Oxford: Basil Blackwell), 1-11.

[3] D. Bertsekas (1999), *Nonlinear Programming*, Belmont, MA: Athena Scientific.

[4] V. Bosetti, M. Catenacci, G. Fiorese, and E. Verdolini (2011), "SOLAR PV and CSP technologies: Policy recommendations from the ICARUS survey on current state and future development," available at: www.icarus-project.org/wp-content/uploads/2011/06/wp-icarus-web.pdf.

[5] V. Bosetti, C. Carraro, M. Galeotti, E. Massetti and M. Tavoni (2006), "WITCH: A World Induced Technical Change Hybrid Model," *Energy Journal*, 2, 13-38.

[6] S. Boyd and L. Vandenberghe (2008), *Convex Optimization*. New York, NY: Cambridge University Press.

[7] A. Chateauneuf, J. Eichberger, and S. Grant (2007), "Choice under Uncertainty with the Best and Worst in Mind: Neo-additive Capacities," *Journal of Economic Theory*, 137, 538–567.

[8] R. Clemen and R. Winkler (1999), "Combining Probability Distributions from Experts in Risk Analysis," *Risk Analysis*, 19, 187–203.

[9] R. Clemen and R. Winkler (2007), "Aggregating probability distributions," In: W. Edwards, R. Miles, & D. von Winterfeldt (Eds.) Advances in Decision Analysis .(pp. 154-176). Cambridge , UK : Cambridge University Press.

[10] J. Eichberger, S. Grant, D. Kelsey, and G. Koshevoye (2011), "The α-MEU model: A comment," *Journal of Economic Theory*, 146, 1684–1698.

[11] P. Ghirardato (2001), "Coping with Ignorance: Unforeseen Contingencies and Non-Additive Uncertainty," *Economic Theory*, 17, 247–276.

[12] P. Ghirardato, F. Maccheroni, and M. Marinacci (2004), "Differentiating Ambiguity and Ambiguity Attitude," *Journal of Economic Theory*, 118, 133-173.

[13] Gilboa, I. and M. Marinacci (2011), "Ambiguity and the Bayesian Paradigm," Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society, forthcoming.

[14] I. Gilboa and D. Schmeidler (1989), "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153.

[15] G. Iyengar (2005), "Robust Dynamic Programming," *Mathematics of Operations Research*, 30, 257-280.

[16] J.-Y. Jaffray (1989), "Linear Utility Theory for Belief Functions," *Operations Research Letters*, 8, 107–112.

[17] P. Milgrom and I. Segal (2002), "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583–601.

[18] W. Olszewski (2007), "Preferences Over Sets of Lotteries," *Review of Economic Studies*, 74, 567-595.

[19] D. Popp (2004), "ENTICE: Endogenous Technological Change in the DICE Model of Global Warming," *Journal of Environmental Economics and Management*, 48, 742–768.

[20] L. J. Savage (1954), *The Foundations of Statistics*, Wiley, New York.

[21] M.L. Viero (2009), "Exactly What Happens After the Anscombe-Aumann Race? Representing Preferences in Vague Environments," *Economic Theory*, 41, 175–212.