Project 001(D) Alternative Jet Fuel Supply Chain Analysis

The Pennsylvania State University

Project Lead Investigator

Saurabh Bansal
Associate Professor of Supply Chain Management
Department of Supply Chain and Information Systems
The Pennsylvania State University
405 Business
University Park, PA 16802
814-863-3727
sub32@psu.edu

University Participants

The Pennsylvania State University (Penn State)

- PI: Saurabh Bansal, Associate Professor of Supply Chain Management.
- PI: Lara Fowler, Senior Lecturer, Penn State Law School; Assistant Director, Penn State Institutes of Energy and the Environment
- PI: Ekrem Korkut, Penn State Law School

The Washington State University (WSU)

- Kristin Brandt, Staff Engineer

University of Tennessee

- Tim Rials, Associate Dean Ag Research
- Burt English, Professor of Agricultural and Resource Economics

Project Funding Level

FAA Funding: $200,000
Matching, Penn State: $200,000
Total Funding: $400,000

Investigation Team

Task 1.3.1 (Lead: Bansal; supported by Brandt and English): risk-reward profit sharing modeling for first facilities.
Task 1.3.2 (Lead: Bansal; supported by Brandt and English): additional quantification of risk and uncertainties in supply chains (foundational part of Task above).
Task 1.3.3 (Lead: Bansal; supported by Brandt and English): supply chain risk analysis tools for farmer adoption.
Task 1.4.1 (Lead: Fowler; supported by Korkut): national survey of current and proposed state and federal programs that monetize ecosystem services.
Task 1.4.3 (Lead: Fowler; supported by Korkut): support in stakeholder engagement efforts.

Project Overview

The project focuses on developing a qualitative and quantitative understanding of factors that can help the establishment of biofuel supply chains aimed at supplying alternative jet fuels. Efforts are being made to establish these supply chains. However, many of these efforts are challenged because of a lack of clarity regarding the incentives that stakeholders would require to engage in these supply chains and devote their resources to invest in the facilities required for these supply chains. To this end, the project has two goals:

1. Develop proforma cash flows that represent the financial status of various participants in biofuel supply chains for

alternative jet fuels to inform a transparent risk-sharing tool, and
2. Understand the policy landscape that exists in various parts of the U.S. to encourage these supply chains and identify further policy initiatives that may be needed.

Task 1.3.1 – Risk-Reward Profit Sharing Modeling for First Facilities

The Pennsylvania State University

Objective

Develop a transparent risk-sharing tool to provide all partners with an understanding of the cash flows and risks faced by all supply chain partners.

Research Approach

We first collected a large number of risk-sharing tools that have been proposed in the supply chain literature. Subsequently, we narrowed the list down to 9–12 mechanisms. We created an Excel-based framework in which the cash flows of all supply chain partners are modeled by using the numbers from the techno-economic analyses developed by WSU. This framework incorporates the risk sharing mechanisms.

Milestone

We developed the Excel models for four realistic configurations by using data from techno-economic analysis models from WSU.

Major Accomplishments

We developed an Excel-based framework showing the cash flows of four key stakeholders of alternative jet fuel supply chains: farmers, preprocessors, refineries, and airlines. The framework shows various risk-sharing contracts that each of the stakeholders can extend to others, as well as the financial burden or opportunity associated with these mechanisms. The framework also shows the government's financial burden of supporting these mechanisms. The framework is developed for four levels of refinery capacities. Overall, this framework can be used as a decision support tool by various stakeholders to determine whether to engage in alternative jet biofuel supply chains and negotiate with each other.

Publications

We anticipate publishing a paper based on combined work from the last year and the coming year.

Outreach Efforts

The tool has been presented and discussed at three ASCENT advisory committee meetings.

Awards

None

Student Involvement

None

Plans for Next Period

We were planning on running laboratory studies with graduate students. However, the behavioral research lab was closed at Penn State due to the COVID-19 situation. We will run these studies when students are back to campus. We would be able to run these studies only when students are able to interact with each other in a simulated negotiation environment. We will provide the tool and a training in use of the tool to project sponsor.

Task 1.3.2 – Additional Quantification of Risk and Uncertainties in Supply Chains (Foundational Part of Task Above)

The Pennsylvania State University

Objective

Develop methods to rely on expert judgments to quantify uncertainties associated with biofuel supply chains.

Research Approach

We developed a new econometric approach to quantify probability distributions of uncertain quantities such as yield or demand when a panel of experts provides judgments regarding the most-likely values. This approach exploits the well-known theory of generalized least squares in statistics for the context in which historical data are available to calibrate expert judgments or when these data are not available.

Milestones

We have described the method in two manuscripts provided as attachments with this ASCENT annual report. In the first manuscript, *Using Subjective Probability Distributions to Support Supply Chain Decisions for Innovative Agribusiness Products*, we develop a two-stage procedure to calibrate expert judgements for the distribution of biofuel uncertainties, such as the uncertain yield of new varieties of oil seeds, demand, or selling price. In the first step of the procedure, we calibrate the expert judgements by using historical data. Specifically, we use prior judgments provided by experts and compare them with actual realizations (such as predicted yield versus actual yield) to determine the frequency with which each expert over- or underestimated the uncertainty, e.g., Expert 1 underestimated the yield 60% of the time, but Expert 2 underestimated the yield 90% of the time. In the second manuscript, *Optimal Aggregation of Individual Judgmental Forecasts to Support Decision Making in a R&D Program*, we use this information to determine the optimal way to aggregate the experts' judgments to determine the mean and standard deviation of the probability distributions. In the second manuscript, we develop a new optimization protocol to determine the optimal acreage for growing specific crops, by taking into account the estimated mean and standard deviation as well as incorporating the variability in these estimates.

Major Accomplishments

Theoretical development and a numerical study have demonstrated the promise of this approach.

Publications

One paper has been accepted. The second paper is finished.

Outreach Efforts

N/A

Awards

None

Student Involvement

None

Plans for Next Period

The second paper has been submitted for review. It will be sent out for a publication during this period.

Task 1.3.3 – Supply Chain Risk Analysis Tools for Farmer Adoption

The Pennsylvania State University

Objectives

Understand farmers' risk preferences over a long duration and how these preferences affect their decisions to grow crops that can support alternative jet fuel supply chains

Research Approach
We surveyed farmers to understand their risk preferences over extended durations. Specifically, we showed them sample yield ranges over extended periods and asked them to estimate the lowest equivalent guaranteed yield that they would be willing to accept given the uncertain yields. We used these responses for statistical analyses.

Milestones
We have completed the survey and finished a manuscript based on the survey.

Major Accomplishments
We compiled data from 43 farmers in central Pennsylvania regarding their preferences given the uncertain yields from their land. The results quantify the loss of value that farmers attribute to an uncertain yield. The reported results are for both 1-year and 10-year horizons. For the 10-year horizon, we also report results with an initial yield buildup, as is the case with most biofuel crops. The key takeaways from this study are that: (a) farmers' valuation of a new crop decreases acutely as the uncertainty in yield increases, and (b) the initial build-up period of low yields can be a large deterrent to farmers' adopting new crops for the purpose of supporting biofuels.

Publications
The paper was finished and was provided to the sponsor.

Outreach Efforts
N/A

Awards
None

Student Involvement
None

Plans for Next Period
The results in the first version of the paper revealed something interesting: when faced with uncertain yields, say from x to y, farmers were willing to swap their output for a consistent output at levels that were lower than x. This finding was surprising at first. However, the research team has recently found prior research in economics documenting similar behavior. We would like to collect more data during the year to bolster the manuscript, pending a resolution to the COVID-19 situation.

Task 1.4.1 – National Survey of Current and Proposed State and Federal Programs that Monetize Ecosystem Services
The Pennsylvania State University

Objective
Conduct a survey and summarize current and proposed state and federal programs to monetize ecosystem services.

Research Approach
This Task builds on and continues the work done under ASCENT Project 01, Task 8.1, which focused on the biomass and water quality benefits to the Chesapeake Bay watershed. Under this Task, we examined the biofuel law and policy landscape of the Pacific Northwest and Southeast regions, as well as the state of Hawaii. We also researched federal biofuel law and policy. We have had a change in personnel working on this project. Lara Fowler remains the lead; however, Gaby Gilbeau left the project in August 2018, and Ekrem Korkut joined the project during the fall of 2018.

Milestones
We have captured this research in three region-specific white papers describing the biofuel law and policy incentives, and the ecosystem service drivers for the subregions. In addition, we added another U.S.-level white paper to the list of tasks.

- Project 01A, Tasks 3.1, the Pacific Northwest.

- Project 01B, Task 3.2, Hawaii.
- Project 01E, the Southeast.

Copies of these documents are available online:

- Western U.S. policy paper (with a focus on Washington State): https://psu.box.com/s/l9ektkcr8lk10gjqu93l4jmm9djmnmhf
- Southeast policy paper (with a focus on Tennessee): https://psu.box.com/s/iyeowdfo0447t4ya8dl5md2zu5un48u6
- Hawaii policy paper: https://psu.box.com/s/92a7tl19tpphg69t4ff12t9d4rdshgq1
- Federal level white paper: https://psu.app.box.com/file/629416796137?s=5r15l1xg8yeg1nnms1nfjx023p3wzkfu
- Poster: https://psu.box.com/s/20ugtneqsmu8ufrjrahos87hp47dk2zm

Major Accomplishments

We have captured this research in three regional white papers describing the biofuel law and policy incentives. In addition, we have researched and finished drafting a document summarizing aviation and biofuel at the national level in the U. S. As part of this, we have examined how legal and policy drivers from other parts of the world are affecting U.S. incentives.

Publications

The white papers have been sent to ASCENT leads for review and comment (including Nate Brown and Michael Wolcott); comments on the federal white paper have been addressed and incorporated.

We are working on turning these papers into publications for the *Frontiers in Energy* special edition. In addition, we have circulated the white papers to ASCENT team members for their background and information.

Outreach Efforts

Lara Fowler and Ekrem Korkut created and shared a poster for the September 29, 2020 annual meeting. This poster is linked above and addresses the federal, state, regional and international aspects of aviation biofuel law and policy.

Awards

None

Student Involvement

Ekrem Korkut continues to be a full-time student at the Penn State School of International Affairs. He has continued to work on the ASCENT project as a part-time research assistant while conducting his studies.

Plans for Next Period

As noted above, we are turning the existing white papers into published papers (at least one policy related piece for the *Frontiers in Energy* special issue) and planning on an additional review at the state/regional level. In addition, we are working with other ASCENT team members on law and policy research questions they have identified, including how landfill regulations shape opportunities in Hawaii and other related topics.

Task 1.4.3 – Help Support Stakeholder Engagement Efforts

The Pennsylvania State University

Objective

Facilitate dialogue among producers, industry, government, and other affected stakeholders.

Research Approach

Our work under this objective focused on stakeholder engagement and facilitation of effective dialogue to help bridge the gaps among producers, industry, government, and other affected stakeholders. This role supports other team members' needs.

Milestone
These efforts supported the stakeholder engagement efforts led by other teams, including but not limited to the regional partners identified in ASCENT Project 01, Tasks 3.1, 3.2, and 3.3.

Major Accomplishments
This set of tasks has been more limited, with no major accomplishments to date. We have continued to participate in discussions and calls related to potential stakeholder engagement needs.

Publications
N/A

Outreach Efforts
N/A

Awards
None

Student Involvement
None

Plans for Next Period
Future work under this objective will include presenting to the project partners on facilitation skills and tactics. Additional support for regional projects will be offered as needed for facilitation and stakeholder engagement sessions as the regional projects move to the deployment stage.

Using Subjective Probability Distributions to Support Supply Chain Decisions for Innovative Agribusiness Products *

Saurabh Bansal†, Tianyang Wang ‡

November 19, 2019

Abstract

This research is motivated by a research problem in agribusiness R&D environments where a decision maker such as a farmer or a firm makes investment/selection decisions for a portfolio of innovative products for nascent supply chains and relies on experts to quantify key uncertainties for these products in the form of their probability distributions. Previous literature suggests that the subjective distributions obtained from domain experts contain judgmental errors. But no clear guidelines exist for incorporating these subjective distributions in managerial decisions. In this paper we (i) develop prescriptive tools to make these decisions using subjective probability distributions with judgmental errors, and (ii) identify the benefits of incorporating judgmental errors present in subjective probability distributions into decision making. We first develop a hierarchical uncertainty model where the outer level models the judgemental errors in the expert's judgments and the inner level models the supply chain resource allocation problem conditioned on magnitude of judgmental errors. To solve this hierarchal model, we develop a copula based dependent decision tree approach. The approach is efficient and permits scaling up to a large number of uncertainties present in a typical supply chain with a portfolio of products. Numerical results for representative industry data in the agribusiness domain show that ignoring the uncertainty in subjective probability distributions can lead to a (i) loss of profit for a farmer by 2-5%, and (ii) incorrect R&D portfolio selection decisions by as much as 24%.

Keywords: R&D process, New Product Development, Supply Chain Management, Human Element

*This research was funded by the U.S. Federal Aviation Administration Office of Environment and Energy through ASCENT, the FAA Center of Excellence for Alternative Jet Fuels and the Environment, project 12447380 through FAA Award Number 13-C-AJFE-PSU-031 under the supervision of Nate Brown and Kristin Lewis. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the FAA.

†The Pennsylvania State University. Email: sub32@psu.edu

‡Colorado State University. Email: Tianyang.Wang@colostate.edu¿

1 Introduction

The human element – human thoughts, actions, and decisions – plays an important role in the development of new supply chains. This element is instrumental in idea generation and evaluation, and managing the transition of an idea into a product, and these aspects have been studied extensively in the past. This paper focuses on an aspect that has received relatively less attention in this literature – managing supply chains for portfolios of new products when the key uncertainties associated with new products are obtained from experts in the form of *subjective* probability distributions.

Subjective probability distributions are obtained from experts and are used for forecasting/decision making when sufficient historical data are not available for constructing statistical distributions. The lack of historical data is especially common in industries engaged in developing a new class of products. Our contextual focus is on one such business – biofuel supply chains – that is in nascent stage. Biofuel supply chains use specifically developed crops such as Miscanthus, Camelina etc to provide biomass that is converted into fuel in specifically designed refineries using customized chemical processes. A critical challenge in this industry to develop new varieties of crops that will provide higher energy density while also being able to thrive in local weather and geographical conditions. These new varieties are developed by teams of plant biology experts. A frequent issue faced by these teams pertains to selecting which varieties to pursue based on limited test information. Teams seldom have have the time and resources to collect extensive data to build systematic models to identify patterns in previously successful varieties and then identify other potentially high performance varieties by exploiting these patterns. More often, the teams rely on their judgment and intuition to determine the likely distribution of the yield of new varieties. Based on this information and the impact of likely yields on supply chains they select a set of varieties to pursue further.

The prior literature discusses two approaches for estimating subjective distributions in such situations. In the first approach the firm seeks point forecasts from a number of experts or managers (we will use the words managers and experts interchangeably in the rest of the paper). The average of these forecasts provides the mean of the distribution; the variance of the distribution is proportional to the variance of the point forecasts obtained from managers (Gaur et al. 2007). But

this approach is applicable only when a large number of managers with the contextual knowledge are available to provide judgments. In contrast in many instances as few as one or two experts may be available to provide judgments for the probability distributions and when that happens this theory is not applicable.

A second approach discussed in the literature focuses on eliciting probability distributions from single or more experts, using various protocols. These protocols include obtaining direct estimates of probability distributions, i.e., asking an expert directly to specify the mean and variance or standard deviation such as those discussed in Clemen et al. (2000); and indirect methods in which moments of probability distributions are deduced from quantile judgments, e.g. weighted linear functions of the quantile judgments of 5th, 50th, and 95th quantiles, such as discussed in Bansal et al. (2017). This literature recognizes that the experts are not perfect – their direct or indirect judgmental estimates for means and standard deviations are not always equal to the unknown true values – and that these subjective estimates are subject to judgmental errors comprised of systematic biases and random inconsistencies or noise.

But this literature does not address how subjective distributions obtained from one or more experts should be modeled in order to make supply chain decisions for portfolios of new products. In this paper we develop a mathematical approach to accomplish this task. Specifically, we focus on three questions:

1. How can the judgmental errors in subjective distributions (as characterized by their mean and standard deviation) obtained from an individual expert be quantified?

2. Once quantified, how can these judgmental errors be incorporated into decision making for supply chains in a tractable and efficient manner?

3. What are the benefits of accounting for judgmental errors during decision making?

To answer the first question we discuss two approaches to quantify the judgmental errors present in the moment judgments provided by experts for probability distributions. This quantification separates the bias – the systematic component of the judgmental errors – from the noise. To answer the second question we formulate a general supply chain portfolio resource allocation decision problem that explicitly models the quantified bias and noise in moment judgments, and then develop

an efficient Copula based solution to the formulation in order to provide a real time decision support. We first explain the need to incorporate the bias and noise in moment judgments into portfolio decision making, followed by technical challenges of solving such a formulation. Intuition suggests that the bias and noise present in the moment judgments provided by an expert should not be ignored during decision making for a portfolio involving multiple products. Ignoring the bias systematically leads to an underestimation or overestimation of the mean and standard deviation, which would systematically misestimate the cash flows from a product in the portfolio. The noise is the random error component in expert's judgmental errors, and is modeled as a random variable with the mean of 0 with an associated distribution. Ignoring the noise leads to an underestimation of the uncertainty in the portfolio, which leads to an incorrect portfolio selection/allocation decision as well as underestimating the uncertainty in the cash flows from the portfolio selected.

But no clear modeling approaches have been documented in the literature for incorporating biased and noisy moment judgments into decision making. Accordingly, we first develop a hierarchial formulation to model the portfolio management problem with judgmental errors in moment judgments. In this hierarchical formulation the outer level represents the judgmental errors in moment judgments for yield distributions. The inner level represents the supply chain resource allocation problem conditioned on the the magnitude of the judgmental errors.

This explicit incorporation of judgmental errors in moment judgments increases the complexity and problem size of portfolio selection/allocation problems. For instance, consider a supply chain in which a firm invests in growing a portfolio of 20 varieties of crops for the purposes of biofuel. The presence of judgmental errors in moment estimates for the yield distributions adds further 40 uncertainties (one distribution each for the judgmental error in the estimate of the mean and standard deviation for the yield distribution of each seed) to the already present 20 uncertainties in the portfolio. The optimal resource allocation decision for these varieties needs to be made over all possible scenarios of judgmental errors and then for all possible yield scenarios that come off of each error scenario. The correlations between the judgmental errors for the mean and standard deviations, if present, must also be incorporated appropriately.

At the same time portfolio decisions are typically made in managerial discussions that seek to accommodate specific preferences of various functions, e.g. marketing, finance, operations, of firms (Jones et al. 2001, Bansal and Nagarajan 2017). To provide a decision support to such discussions

in a real time mode, approaches that can solve portfolio optimization problems efficiently and provide a quick what-if analysis are desirable. Accordingly we develop a copula based dependent decision tree approach to solve the hierarchial formulation. The approach is precise and efficient in evaluating the cash flows from various projects in a portfolio, and permits scaling up to a large number of uncertainties as is needed for managing multi-asset portfolios. Finally, numerical results for representative industry data show that ignoring the uncertainty in subjective probability distributions can lead to a (i) loss of profit for a farmer by 2-5%, and (ii) incorrect R&D portfolio selection decisions by as much as 24%.This loss of profit stems from, among other sources, a systematically incorrect allocation of resources between a portfolio of crops.

The remainder of this article is organized as follows. We discuss related literature in the rest of this Section. In Section 2, we present a portfolio selection problem and then discuss two models. The first model ignores the judgmental uncertainty in the moment estimates of the distributions, and the second model incorporates this uncertainty. We then highlight the additional complexity in the second model. In Section 3, we discuss the solution. In Section 4, we present a numerical analysis to quantify the benefit of incorporating judgmental errors into the firm's portfolio decision. Section 5 contains a summary.

1.1 Literature Review

The process of developing new products relies heavily on human element, and the existing literature has focused on three aspects of managers' behavior and subjective decisions during the new product development process. The first aspect is on managing the idea generation process. Existing literature investigates the characteristics of individuals and teams that lead to better ideas for new products, environments that stimulate high-quality idea generation, biases in designers and innovators during short-listing designs for a further development, and the processes that lead to innovative designs verses derivative designs (see, e.g., Toubia 2006). The second stream of literature focuses on managing the process of transforming an idea or a design into a marketable product once the idea/design has been finalized. This literature addresses how managers use judgments to allocate time between planning and implementation activities, and between development and testing activities (see, e.g., Annacchino 2003).

A third stream of literature, to which we contribute to, focuses on estimating uncertainties

associated with new products to make business decisions to manage the new products. The common theme in this literature is that often relevant historical data do not exist for new products to quantify various business uncertainties and a firm must rely on managers' subjective judgments to estimate the uncertainties. Gaur et al. (2007) focus on the case when multiple experts are available for estimating the distributions of the uncertainties. They suggest that the firm collects a point forecast from each manager and treats the forecasts obtained as *iid* realizations from the underlying population. It follows from the properties of sampling distributions that the mean of the point forecasts is a good predictor of the population mean. Further the variance in the point forecasts is directly proportional to the true variance; to determine the appropriate proportionality constant they use a calibration exercise in which the same team of experts provides points forecasts for a number of distributions for which historical data exist. A regression of the true variances of these distributions on the forecasts' variance provides the proportionality constant. This approach requires two inputs to be successful — multiple managers (in order to be able to compute forecasts' variances), and calibration distributions (in order to determine the proportionality constant). But this approach does not perform well when only a small number of experts – such as two or three – are available to provide judgments.

Recent developments by Baker and Solak (2011) and Kettunen and Salo (2017) adopt a different approach that can use judgmental estimates from a small number of experts to make portfolio decisions. Baker and Solak (2011) study the problem of selecting a portfolio of unproven technologies to invest in. They directly elicit a number of points on the probability distributions of the efficiencies of various technologies in the future from an expert. Then they use this discretized elicited distribution in a stochastic program to determine the optimal investment levels. Our work diverges from Baker and Solak (2011) in various aspects. First, we quantify the judgmental errors present in the subjective distributions and make the portfolio selection decision while accounting for judgmental errors. Second, we identify specific ways in which ignoring judgmental errors leads to suboptimal decision making in portfolio decision making. In our context of integrated project selection and production planning, we find that the suboptimality can be significant. Finally, we focus on eliciting the moments of distributions that enables firms to make nuanced decisions for portfolio selection.

Kettunen and Salo (2017) focus on project portfolio selection and show that when experts

provide subjective judgments for project valuations, the errors in the valuations could lead to a mis-specification of the risk measure of the portfolio. They then suggest a simulation based calibration based approach to correctly evaluate the selected portfolio. In contrast, we focus on portfolio selection itself in the presence of judgmental errors, and develop insights for the impact of ignoring judgmental errors on the resource allocation decision.

2 Problem Description

2.1 Portfolio Selection and Management Under Uncertainty: Preliminary Model and Assumptions

Consider a portfolio of *i=1,2,...,N* products of a supply chain with associated random variables $\boldsymbol{X} = [X_1, X_2, \ldots, X_N]^{\mathrm{T}}$ with each random variable X_i corresponding to a unique asset i. These random variables could be for uncertain demand or supply or price or a combination of these factors. The distributions of $\boldsymbol{X}$ are estimated using expert judgment. The distribution of X_i has two parameters $\boldsymbol{\theta}_i = [\theta_{i1}, \theta_{i2}]$ with the pdf denoted as $f(x_i; \boldsymbol{\theta}_i)$. The vector of parameters $\boldsymbol{\theta}_i$ is denoted as $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N]$. Consistent with Lindley (1987) we assume that the functional form of the pdf $f(x_i; \boldsymbol{\theta}_i)$ is known from the business context. Other uncertainties with available statistical distributions from historical data may also be present in the decision problem, and they can be easily modeled within our framework.

The portfolio management problem has a span of two periods. In Period 1, the decision maker chooses the investment levels for the assets, $\mathbf{Q} = [q_1, q_2, ..., q_N]^{\mathrm{T}}$, incurring a per unit cost $\mathbf{C} = [c_1, c_2, ..., c_N]^{\mathrm{T}}$, i.e. at a total cost of $\sum_i c_i q_i$. We do not make any specific assumption about the domain of $\mathbf{Q}$ except that these decision variables take strictly non-negative values, as is the norm in portfolio selection/allocation problems. Each q_i can take binary values to represent Yes/No investment decisions or continuous values to represent varying levels of investments such as amount of inventory purchased or capacity used for production. This decision is subject to multiple constraints collectively denoted as $\mathbf{g}(\mathbf{Q}) \leq \mathbf{L}$ where $\mathbf{g}(\mathbf{Q}) = [g_1(\mathbf{Q}), g_2(\mathbf{Q}), ..., g_M(\mathbf{Q})]^{\mathrm{T}}$ and $\mathbf{L} = [l_1, l_2, ..., l_M]^{\mathrm{T}}$. These constraints could correspond to resource limits on investments in projects of specific categories, relationships between various investments such as choosing a specific number of projects of a specific class, and total capital investment limits. In Period 2, the decision

maker observes the uncertainty realizations $\mathbf{x} = [x_1, x_2, ..., x_N]^{\mathrm{T}}$ and obtains the payoff $\Pi(\mathbf{Q}, \mathbf{x})$. His two-stage expected profit maximization problem is stated as

$$\max_{\mathbf{Q} \geq \mathbf{0}} \quad -\sum_i c_i q_i + \mathbb{E}[\Pi(\mathbf{Q}, \mathbf{x})] \Bigg) \tag{1}$$

$$\text{s.t. } \mathbf{g}(\mathbf{Q}) \leq \mathbf{L} \tag{2}$$

We focus on the case where the payoff $\Pi(\mathbf{Q}, \mathbf{x})$ is separable in the outcome of each uncertainty X_i.

We next discuss obtaining expert judgments for the parameters $\boldsymbol{\theta}_i$. In the rest of the development, we will assume that $f(x_i; \boldsymbol{\theta}_i)$ is a location-scale distribution. The family of Location-scale distributions includes some common distributions such as the Uniform distribution, Normal distribution, and the Gumble distributions that are used in a variety of operations and supply chain contexts. In this framework the parameters $\boldsymbol{\theta}_i = [\theta_{i1}, \theta_{i2}]$ represent, without a loss of generality, the mean and standard deviation, respectively, of the uncertainty of the random variable X_i (Casella and Berger 2002).

2.2 Expert Input for Probability Distributions

In the absence of historical data, the parameters $\boldsymbol{\theta}_i$ are obtained using expert judgments. We discuss here two frequentist approaches to obtain these parameters, (i) the direct approach, and (ii) the indirect approach. These approaches have the same underlying principle as Cooke (1991). Cook suggests asking experts for the probability of occurrence of a specific set of events and then compare these probability judgments against the observed true values to calibrate the experts. We extend this approach to calibrate the expert's judgments for the mean and standard deviation using the direct and indirect approach. In both approaches the expert's judgments are compared against actual data for a series of available distributions and the bias and noise in the judgments is quantified. The information generated during this quantification process then provides the optimal way to process the expert's judgments for those distributions that do not have prior data.

2.2.1 Direct Approach

In the direct approach the expert provides direct estimates of moments $\hat{\theta}_{rj}; j = 1, 2$ for a number of distributions $r = 1, 2, \ldots, R$ for which historical data are available. The actual values θ_{rj} are regressed on the judgments $\hat{\theta}_{rj}$ to obtain the regression lines:

$$\theta_{rj} = a_j + b_j \hat{\theta}_{rj}; j = 1, 2; r = 1, 2, ..., R$$

The standard errors of the two regressions are denoted as $\sigma^2_{se,j}; j = 1, 2$, and the correlation between the fitted residuals for *j=1,2* is denoted as ρ^θ. This calibration information is used for estimating future distributions $i = 1, 2, ..., N$ that do not have prior data, as follows. We first an obtain expert's judgments $\hat{\theta}_{ij}$. For these judgments the marginal distribution of parameter θ_{ij}, $\phi_{ij}(\theta_{ij}; \alpha_{ij}, \beta_{ij})$ with hyper parameters α_{ij}, β_{ij}, has the mean $E[\hat{\theta}_{ij}] = a + b\hat{\theta}_{ij}$ and the standard deviation equal to the standard error $\sigma^2_{se,j}; j = 1, 2$. The correlation between the marginal distributions is equal to ρ^θ.

To determine the parametric form of the marginal distributions $\phi_{ij}(\theta_{ij}; \alpha_{ij}, \beta_{ij})$, we determine the best fitting distributions for the errors in estimates $\hat{\theta}_{rj}$ in the calibration set of R distributions, $e^\theta_{rj} = a + b\hat{\theta}_{rj} - \theta_{rj}$. Moment matching provides the parameter values $\alpha_{ij}, \beta_{ij}; j = 1, 2$. For example, if the marginal distributions ϕ_{i1}, ϕ_{i2} are normal, then $\alpha_{ij} = a_j + b_j\hat{\theta}_{ij}$ and $\beta_{ij} = \sigma_{se,j}$. The joint distribution of parameters $\boldsymbol{\theta}_i = [\theta_{i1}, \theta_{i2}]$ is denoted as $\phi_i(\boldsymbol{\theta}_i; \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \rho^\theta)$. Finally, the distributions for two products, say i and k, $\boldsymbol{\phi}_i(\boldsymbol{\theta}_i; \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \rho^\theta)$ and $\boldsymbol{\phi}_k(\boldsymbol{\theta}_k; \boldsymbol{\alpha}_k, \boldsymbol{\beta}_k, \rho^\theta)$ respectively for $i \neq k$ are mutually independent. For notational brevity, we denote the joint distributions $\boldsymbol{\phi}_i(\boldsymbol{\theta}_i; \boldsymbol{\alpha}_i, \boldsymbol{\beta}_i, \rho^\theta)$ as $\boldsymbol{\phi}_i(\boldsymbol{\theta}_i)$.

2.2.2 Indirect Approach

In the indirect approach, estimates for quantiles are used to deduce distribution parameters. Bansal et al. (2017) recently showed that if an expert provides unbiased estimates of quantiles $\hat{v}_k; k = 1, \ldots, m$ corresponding to a set of probabilities p_k and standardized values z_k of a location scale distribution, and the variance-covariance matrix of the judgmental errors in these percentile judgments is $\boldsymbol{\Omega}$, then the estimates of $\theta_j; j = 1, 2$ are obtained as $\hat{\theta}_j = \mathbf{W}_j[\hat{v}_1 \ldots \hat{v}_m]^{\mathrm{T}}$ where

$$\mathbf{W}_j^{\mathrm{T}} = \mathbf{a}_j^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\boldsymbol{\Omega}^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\boldsymbol{\Omega}^{-1} \tag{3}$$

where the matrix $\mathbf{Z} = \begin{bmatrix} 1 & 1 & 1 \\ z_1 & z_2 & z_3 \end{bmatrix}$, $\mathbf{a}_1^{\mathrm{T}} = [1, 0]$ and $\mathbf{a}_2^{\mathrm{T}} = [0, 1]$. The variances and covariances of estimates are equal to $\mathrm{Var}[\hat{\theta}_j] = \mathbf{a}_j^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{Z})^{-1}\mathbf{a}_j$ and $\mathrm{Covar}[\hat{\theta}_1\hat{\theta}_2] = \mathbf{a}_1^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{Z})^{-1}\mathbf{a}_2$.

We next operationalize this result in our context. Consider a number of calibration distributions $r = 1, 2, \ldots, R$ for which historical data exist and true parameter values θ_{rj} are known. For each distribution r, the expert provides his judgments for M quantiles $\hat{v}_{rm}; m = 1, 2, 3..., M$. The quantile judgments are compared with the true values v_{rm} to obtain the error,

$$e_{rm} = \hat{v}_{rm} - v_{rm} \; ; m = 1, 2, ..., M; r = 1, 2, ..., R \tag{4}$$

Subsequently, the bias in each quantile is estimated as

$$\delta_m = \sum_r e_{rm}/R \; ; m = 1, 2, ..., M \tag{5}$$

Using this bias, the unbiased errors in the estimation of the M quantiles are obtained as

$$e^u_{rm} = \hat{v}_{rm} - \delta_m; m = 1, 2, ..., M; r = 1, 2, ..., R \tag{6}$$

From these M streams of errors, a $M \times M$ variance-covariance matrix $\mathbf{\Omega}$ is obtained, which then provides the weights $\mathbf{W}_j^{\mathrm{T}}$ in (3). Once this calibration is complete, the expert's percentile judgments for new distributions are processed as follows. The expert provides his quantile judgments $[\hat{v}_{i1}, \hat{v}_{i2}, ...\hat{v}_{iM}]$ for the distribution of product i. Then the marginal distribution of parameter θ_{ij}, $\phi_{ij}(\theta_{ij}; \alpha_{ij}, \beta_{ij})$ with hyper parameters α_{ij}, β_{ij}, has the mean $E[\hat{\theta}_{ij}] = \mathbf{W}_j[\hat{v}_{i1} - \delta_1, \hat{v}_{i2} - \delta_2, ..., \hat{v}_{iM} - \delta_M]^{\mathrm{T}}$, the variance $\mathrm{Var}[\hat{\theta}_{ij}] = \mathbf{a}_j^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{Z})^{-1}\mathbf{a}_j$, and $\mathrm{Covar}[\hat{\theta}_{i1}\hat{\theta}_{i2}] = \mathbf{a}_1^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{Z})^{-1}\mathbf{a}_2$. To determine the parametric form of the marginal distributions, we determine the best fitting distributions for the errors in estimates $\hat{\theta}_{rj}$ in the calibration set, $e^{\theta}_{rj} = \mathbf{W}_j[\hat{v}_{r1} - \delta_1, \hat{v}_{r2} - \delta_2, ..., \hat{v}_{rM} - \delta_M]^{\mathrm{T}} - \theta_{rj}$. Moment matching provides the parameter values $\alpha_{ij}, \beta_{ij}; j = 1, 2$, similar to the direct elicitation.

In the remainder of the article we will not differentiate between whether the distributions of parameters $\boldsymbol{\phi}_i(\boldsymbol{\theta}_i)$ were obtained using the direct or the indirect approach.

3 Incorporating Subjective Distributions into Supply Chain Decision Making

We now focus on supply chain decision making for a portfolio of products when the parameters $\boldsymbol{\theta}_i$ are obtained through subjective judgmental estimates and have the distribution $\boldsymbol{\phi}_i(\boldsymbol{\theta}_i)$. We first setup a benchmark case.

3.1 Known Parameter Portfolio Problem

Under traditional optimization framework, the firm's problem (1)–(2) is one of maximizing the expected profit from the portfolio of products over the possible realizations of the uncertainty $\boldsymbol{X}$ assuming that the parameters $\boldsymbol{\theta}$ are known deterministically, and it is formulated as:

$$\text{KPP: } \max_{\mathbf{Q}\geq 0} \Pi^k(\mathbf{Q}) = -\sum_{i=1}^{n} c_i q_i + \int_{x_n} \dots \int_{x_1} \Pi(\mathbf{Q}, \mathbf{x}) f(\mathbf{x}; \boldsymbol{\theta}) dx_1 dx_2 \dots dx_n \tag{7}$$

$$\text{s.t. } \mathbf{g}(\mathbf{Q}) \leq \mathbf{L} \tag{8}$$

The optimal solution of this problem is denoted as $\mathbf{Q}^{k*}$. In this Known Parameter Problem (KPP), typically the mean values of the distributions of uncertain $\boldsymbol{\theta}$ are used. This setup ignores the judgmental errors in the estimates of parameters $\boldsymbol{\theta}$. Nevertheless, it serves as a useful benchmark for quantifying the benefit from incorporating judgmental errors present in the estimates of $\boldsymbol{\theta}$ into portfolio decision making.

3.2 Hierarchical Formulation for Unknown Parameter Portfolio Problem

We next develop the formulation that explicitly incorporates judgmental errors in the estimates of $\boldsymbol{\theta}$. Given the uncertainty in parameters $\boldsymbol{\theta}$ with distributions $\boldsymbol{\phi}_i(\boldsymbol{\theta}_i)$, the firm should seek to make the optimal decision $\mathbf{Q}^*$ over all possible values of $\boldsymbol{\theta}$ and the ensuing uncertainty in $\boldsymbol{X}$ with the pdf $f(x_i; \boldsymbol{\theta}_i)$ for a specific value of $\boldsymbol{\theta}_i$. We formulate this dynamic using a hierarchical formulation where in hierarchy level 2 or the inner level, expected payoff over all possibilities of $\boldsymbol{X}$ given specific values of $\boldsymbol{\theta}$ is considered, as $\int_{x_n} \dots \int_{x_1} \Pi(\mathbf{Q}, \mathbf{x}) f(\mathbf{x}; \boldsymbol{\theta}) dx_1 dx_2 \dots dx_n$. Hierarchy level 1 or the outer level considers all possibilities of $\boldsymbol{\theta}$ for the expected payoff from Hierarchy 2. The following formulation

represents this hierarchical optimization framework for the model (1)–(2):

$$\text{UPP: } \max_{\mathbf{Q}\geq 0} \Pi^u(\mathbf{Q}) = -\sum_{i=1}^{n} c_i q_i + \int_{\boldsymbol{\theta}_n} \dots \int_{\boldsymbol{\theta}_1} \left(\int_{x_n} \dots \int_{x_1} \Pi(\mathbf{Q}, \mathbf{x}) f(\mathbf{x}; \boldsymbol{\theta}) dx_1 dx_2 \dots dx_n \right) \boldsymbol{\phi}_1(\boldsymbol{\theta}_1) \dots \boldsymbol{\phi}_n(\boldsymbol{\theta}_n) d\boldsymbol{\theta}_1 \dots d\boldsymbol{\theta}_n \quad (9)$$

$$\text{s.t. } \mathbf{g}(\mathbf{Q}) \leq \mathbf{L} \quad (10)$$

The optimal solution of this problem is denoted as $\mathbf{Q}^{u*}$. Note that this formulation distinguishes between two sources of uncertainty. The parameters $\boldsymbol{\theta}$ are uncertain because the expert's judgments for these parameters/moments have judgmental errors. There exist true values of these parameters but they remain unknown in both Period 1 and Period 2. This is the epistemic uncertainty in the model. The decision maker only knows the probability distributions $\boldsymbol{\phi}_i(\boldsymbol{\theta}_i)$ over these true values, from the calibration exercise with the expert. In contrast, the values of realizations $\mathbf{x}$ for payoff variables $\boldsymbol{X}$ are unknown in Period 1 because they have not occurred yet, but they are known in period 2 when the payoff $\Pi(\mathbf{Q}, \mathbf{x})$ is observed. Hence the uncertainty in $\boldsymbol{X}$ is an aleatory uncertainty. We call this formulation (9) – (10) as the unknown parameter problem (UPP).

3.3 Structural Comparison of KPP and UPP

In the remainder of this paper, we assume that KPP is concave and has a unique solution. This assumption helps us focus exclusively on the complexity introduced by an imperfect knowledge of $\boldsymbol{\theta}$. We first establish a result that is immediate from the definitions of KPP and UPP.

Proposition 1 *If KPP is concave, then UPP is also concave.*

This result implies that if we find a solution to the UPP, it must be the only solution to the problem. This property lends some smoothness characteristics such as $\lim_{Var(\hat{\theta}_{i1}\to 0, \hat{\theta}_{i2}\to 0)} \mathbf{Q}^{u*} = \mathbf{Q}^{k*}$, and $\lim_{Var(\hat{\theta}_{i1}\to 0, \hat{\theta}_{i2}\to 0)} \Pi^u(\mathbf{Q}^{u*}) = \Pi^k(\mathbf{Q}^{k*})$. The next result establishes some natural relationships between the expected profits from these systems.

Proposition 2 $\Pi^k(\mathbf{Q}^{k*}) > \Pi^u(\mathbf{Q}^{u*}) > \Pi^u(\mathbf{Q}^{k*})$

The proposition has two implications. First, ignoring the parametric uncertainty always leads to optimistic estimates of expected profit, as $\Pi^k(\mathbf{Q}^{k*}) > \Pi^u(\mathbf{Q}^{u*})$. In other words, a disappointment

or loss will be incurred on average by ignoring the uncertainty in expert's judgments. This expected loss is equal to

$$\Delta_o = \Pi^k(\mathbf{Q}^{k*}) - \Pi^u(\mathbf{Q}^{k*}) \tag{11}$$

where the subscript o denotes the bias due to the optimism of the error ignoring approach. Equivalently, this amount is also equal to the value of incorporating the judgmental noise present in the estimates of $\boldsymbol{\theta}$ in the decision making process.

Proposition 3 *Let V_a denote the variance of the estimate of moment/parameter $a \in \{\theta_{i,j}\}$ obtained from an expert. Then $\frac{\partial \Delta_o}{\partial V_a} \geq 0$.*

This result states that as the expert becomes less precise, i.e., the noise in his estimates increases the average disappointment from ignoring the noise increases, or alternatively, the value of including judgmental errors into the portfolio allocation decision increases.

3.4 Challenges in Solving UPP: Illustrative Example

In general the UPP problem (9)–(10) is analytically intractable. To appreciate the challenges in solving this problem we focus on an illustrative example from the biofuel sector. Suppose that a large holding farmer or cooperative decides to grow various crops indexed as *j=1,2,...,M* for biofuel usage. The production decision problem for the portfolio of crops is as follows. In period 1, the farmer decides the area q_i (number of acres) to use to grow the crop i, incurring a cost of c_i per unit area. The production yield X_i per unit area (bags of hybrid seeds obtained per acre) for hybrid i is uncertain. The farmer makes the acreage decision subject to the availability of appropriate land for the crops in quantities $L_1, L_2, ..., L_M$, i.e. the constraints in (2) are written as:

$$\mathbf{g}(\mathbf{Q}) \leq \mathbf{L} \doteq \mathbf{Y}\mathbf{Q}^{\mathrm{T}} \leq \mathbf{L} \tag{12}$$

In period 2 the farmer observes the yield x_i and the output $q_i x_i$ and tries to meet a known demand D_i. The selling price per unit is s_i per unit. In period 2, the firm's profit is equal to:

$$\Pi(\mathbf{Q}, \mathbf{x}) = \sum_i s_i \min(q_i x_i, D_i) \tag{13}$$

The UPP formulation (9)–(10) in this context is non-trivial to solve due to analytical intractability and computational issues. To illustrate this, consider the simplest case of portfolio of size 1 where the farmer produces only one crop for biofuel and the allocation constraint $\mathbf{g}(\mathbf{Q}) \leq \mathbf{L}$ is not present. Using (13) as the payoff function the KPP problem (7)–(8) can be written as $\max_q -cq + E_{X_1} \min[qX_1, D]$ or alternately as $\max_q -cq + \int_{D/q}^{\infty} D\phi_1 dx_1 + \int_{-\infty}^{D/q} qx_1\phi_1 dx_1$. This formulation is analytically tractable for a number of commonly used probability distributions for X_1. However its UPP form $\max_q -cq + E_{\boldsymbol{\theta}_1}\left(E_{X_1} \min[qX_1, D]\right)$ is intractable. More specifically, we can write the UPP formulation as

$$\max_q -cq + \int_{\theta_{12}} \int_{\theta_{11}} E_{X_1} \min[qX_1, D]\boldsymbol{\phi}_1(\boldsymbol{\theta}_1) d\theta_{11} d\theta_{12} \tag{14}$$

The marginal distributions $\phi_{1j}(\theta_{1j}; \alpha_{1j}, \beta_{1j})$ depend on the data generated during calibration and may belong to different families in which case the joint distribution $\boldsymbol{\phi}_1(\boldsymbol{\theta}_1; \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1)$ in (14) will be analytically intractable. As a result, the integration operation in the formulation (14) above will be intractable. Even if the joint distribution $\boldsymbol{\phi}_1(\boldsymbol{\theta}_1; \boldsymbol{\alpha}_1, \boldsymbol{\beta}_1)$ can be expressed analytically, the integration operation in (14) is likely to be intractable except for the uniform distribution.

For a single crop, the problem above could still be explored numerically using Monte Carlo simulations using a stochastic linear programming approach. However, such approaches are not practical for portfolio management problems. The simulation error present in the solutions of these approaches increases with the number of uncertainties and decision variables. In the UPP problem, there are n decision variables, and three uncertainties for each decision variable (uncertainty in X_i, and the uncertainty in $\theta_{i1}, \theta_{i,2}$). Therefore to obtain a high level of accuracy for the portfolio decision, one would need a large number of random realizations for each of the $3n$ uncertainties in the portfolio. For a portfolio of 20 or more uncertainties as is typically the case in biofuel supply chains, one would need to simulate 60 uncertainties. Simulating several hundred or thousand scenarios of each of these uncertainties would lead to a formulation of a considerable size. The solution time of such formulations would be substantial and would not support a real time decision aid needed for the portfolio decision making.

To address these practical issues, we develop a discretization-based approach to solve UPP within a decision tree wherein continuous uncertainties are discretized into a limited number of

outcomes. From a technical perspective, decision trees with discrete outcomes of uncertainties are similar to simulation based methods. Monte Carlo methods essentially approximate the values of complicated functions by a weighted sum of a large number of simulated scenarios. The scenarios are drawn randomly and assigned equal probability weights. In contrast, in a decision tree framework a "clever" selection of very few discrete points and probability weights reduces the computational burden without sacrificing accuracy. This efficiency benefit is significant for portfolio models where a large number of uncertainties need to be modeled. Decision tree models with such discretizations have been shown to perform well when compared to Monte Carlo simulations. Bickel et al. (2011) and Clemen and Reilly (1999) show that the quality of these discretizations is equivalent to the several tens of thousands of simulated draws. Finally, from a managerial perspective, the use of discretization based decision trees aids in understanding of the problem by defining distinct scenarios, a characteristic not completely shared by simulation methods which are often viewed as "black boxes" by decision makers.

4 Discretization Based Approach to Solving UPP

We adopt a two-step approach to solve UPP. In Step 1, we use Copulas to create a discrete decision tree version of UPP. This discretization leads to a intuitively appealing representation and a tractable formulation. This formulation is solved in Step 2 for the optimal solution $\mathbf{Q}^{u*}$ using a standard nonlinear solver.

4.1 Step 1: Use Copulas to Discretize the Parameter Space

We discretize the distributions in the UPP formulation based on its hierarchical structure. Specifically, for each seed i we create a state of nature tree discretizing the outer level of distributions of parameters $\boldsymbol{\theta}$, followed by modeling the cash flows Π.

This tree starts with the discretization of uncertainty in parameter θ_{i1} as a set of three possible outcomes – "up", "middle" and "down" (the motivation for the section of three point over more or fewer points is discussed shortly). Similarly for the second uncertainty θ_{i2}, a set of three subsequent states representing three contingent outcomes follows each realization of the preceding state of the parameter θ_{i1}. The final set of nine outcomes fully represent a discrete approximation of the joint

distribution of the two uncertainties of θ_{i1}, θ_{i2}. More specifically, we first discretize the distribution of θ_{i1} into a three point distribution with values $\theta_{i1k}; k = 1, 2, 3$ for "up", "middle" and "down" respectively with probability values p_{i1k}. Then at each of three points $\theta_{i1k}; k = 1, 2, 3$ we discretize the conditional distribution of $\theta_{i2}|\theta_{i1k}$ into three values $\theta_{i2k}|\theta_{i11}, \theta_{i2k}|\theta_{i12}, \theta_{i2k}|\theta_{i13}; k = 1, 2, 3$ with corresponding probability values $p_{i2k}|\theta_{i11}, p_{i2k}|\theta_{i12}, p_{i2k}|\theta_{i13}; k = 1, 2, 3$. Therefore there are a total of 9 points at the end of this discretization. At each end note, we then determine the expected payoff Π. Figure 1 shows this discretization for the uncertainties in $\theta_{i,1}, \theta_{i,2}$ for the payoff for the seed problem $\Pi = E_{X_i} \min[qX_i, D_i]$ for deterministic demand.

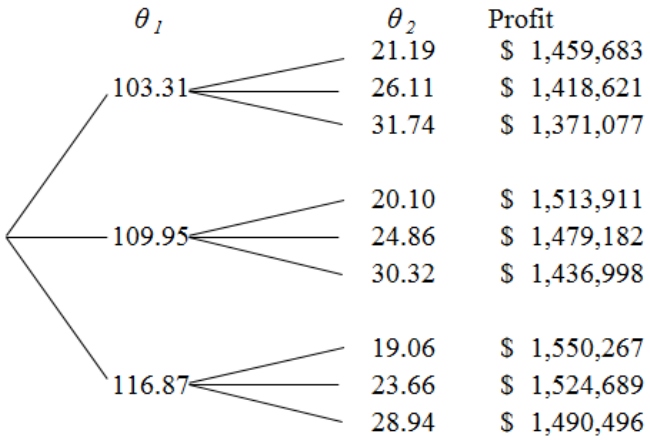

Figure 1: Representative Normal-Copulas dependent tree for one crop.

Two items are necessary to obtain this discretization, (i) an appropriate discretization scheme for the distribution of θ_{i1}, i.e., specifications of $\theta_{i1k}; k = 1, 2, 3$ with the corresponding probability values $p_{i1k}; k = 1, 2, 3$, and (ii) conditional distribution of θ_{i2} given specific values θ_{i1k} of θ_{i1}, i.e. specifications of $\theta_{i2k}|\theta_{i11}, \theta_{i2k}|\theta_{i12}, \theta_{i2k}|\theta_{i13}; k = 1, 2, 3$ with corresponding probability values $p_{i2k}|\theta_{i11}, p_{i2k}|\theta_{i12}, p_{i2k}|\theta_{i13}; k = 1, 2, 3$. We discuss both issues in §4.1.1 and §4.1.2 respectively.

4.1.1 Discretizations to Use for the distribution of θ_{i1}

The distribution of parameter θ_{i1} is obtained from calibrating expert's judgments as discussed in Section 3, and can be of any parametric family. One could use a Gaussian Quadrature to obtain a n point discrete approximation that will match the first *2n-1* moments of this distribution. Empirically, three point discretizations have been found to provide good approximations of cash flows

(see, e.g., Keefer and Bodily 1983, Wang and Dyer 2012). Our selection of three point discretization is based on this literature. This literature also suggests using a constant set of pre-optimized probabilities (0.185, 0.630, and 0.185) for the realizations ("up", "middle" and "down") corresponding to the 5^{th}, 50^{th}, and 95^{th} percentiles of the continuous conditional distributions. It follows that $[\theta_{i11}, \theta_{i12}, \theta_{i13}] = [\Phi^{-1}_{\theta_{1i}}(0.05), \Phi^{-1}_{\theta_{1i}}(0.50), \Phi^{-1}_{\theta_{1i}}(0.95)]$ and $[p_{i11}, p_{i12}, p_{i13}] = [1/6, 4/6, 1/6]$. Note that these weights also correspond to the PERT calculations in project management literature to determine the expected value of activity duration distributions using the largest, most likely and smallest values respectively. Therefore these weights also have the advantage of being familiar to managers.

4.1.2 Construction of Conditional Distributions and Their Discretizations

The marginal distribution of the first uncertainty of θ_{i1} determines its percentile realizations $\theta_{i1k}; k = 1, 2, 3$. We next focus on determining the three percentiles of the conditional distributions of the random variables $\theta_{i2k}|\theta_{i11}, \theta_{i2k}|\theta_{i12}, \theta_{i2k}|\theta_{i13}; k = 1, 2, 3$. Consider the case when we calibrated the expert and found the distributions for the means and standard deviations both to be correlated and belong to the Gamma family of distributions. For the Gamma distribution, there is no known tractable functional form for bivariate correlated distributions. In the absence of this joint bivariate form, the conditional distribution also does not exist in a tractable form. This problem becomes more severe when the distributions of the mean and standard deviation belong to different families. To obtain the conditional distributions for such instances, we adopt a Copula based approach. We first provide a general development and then discuss its application to our context.

A Copula uses the information available for marginal distributions and correlations as input and provides conditional distributions as output. A Copula is a joint distribution of random variables $\Phi(Y_1, ..., Y_n) = C(\Phi_1(Y_1), ..., \Phi_n(Y_n))$ to be expressed as a function $C(.)$ of the marginal distributions $\Phi(Y_1), \ldots, \Phi(Y_n)$, and fully captures the dependence structure among the uncertainties through the choice of the copula function $C(.)$ and the dependence measure. See Frees and Valdez (1998), Nelson (1999), and Embrechts et al. (2001) for an excellent introduction and review on copulas.

We discuss here the use of Normal Copula. This Copula is useful when marginal distributions do not have fat tails, which is commonly the case for operational uncertainties such as uncertain demand or supply. A n-dimensional Normal copula C is given by $C(u_1, \ldots, u_n) =$

$\Phi_\Sigma(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n))$. It is derived from a multivariate normal cumulative distribution function Φ_Σ with mean zero and correlation matrix Σ, by transforming the Uniform random variables $(u_1, \ldots, u_n)$ on [0,1] by taking the inverse of the standard Normal distribution function Φ. With the information of marginal and assessed pairwise correlations, the normal copula-based dependent decision tree for variables $(Y_1, \ldots, Y_n)$ is created as follows.

We first construct a discrete approximation for the unconditional Uniform variable u_1, and then recursively compute the dependent Uniform variables $u_k (k = 2, \ldots, n)$, conditioning on each point realizations of the previous discrete approximations for $(u_1, \ldots, u_{k-1})$. Thus, u_n can be written as

$$u_n = \Phi(A_{n1}\Phi^{-1}(\alpha_1) + \ldots + A_{n(n-1)}\Phi^{-1}(\alpha_{n-1}) + A_{nn}\Phi^{-1}(\alpha_n)) \tag{15}$$

where A_{nj} is the (n, j) element of the lower triangular matrix A obtained during Cholesky factorization of the covariance matrix Σ as $\Sigma = AA^T$ and α_i is the pre-determined (optimally chosen) percentiles of the conditional distribution $u_i|u_1,\ldots,u_{i-1}$.

After the calculation of u_i, we combine the marginal information and transform them to obtain the discrete approximations to the original uncertainties. The discrete approximations of Y_i are obtained by applying the inverse of the target marginal distribution function for each realization of u_i, i.e., $Y_i = \Phi_i^{-1}(u_i)$. Similarly, we create the contingent tree for each successive node until we generate the complete multivariate standard decision tree for $(Y_1, \ldots, Y_n)$.

4.1.3 Illustrative Example

We now illustrate this approach in the biofuel context. Our goal is to model the dependence structure of the two parameter uncertainties $\theta_{i,1}, \theta_{i,2}$ for crop i. We use a set of parameters from Bansal et al. (2017) for this illustration. They calibrated an expert for new varieties of a commercial crop. For one of the varieties that mean of the yield distribution had the average value of 110 units per acre and standard deviation of 4.12 units. The standard deviation of the yield distribution was uncertain with an average of 25 units with a standard deviation of 3.2 units. Furthermore while the yield distribution itself was Normal, the marginal distributions for the parameters had a Gamma distributions with a correlation of -0.15. To create the probability tree for the standardized Uniform variables, we first generate the discretization for $u_{\theta_{i,1}}$. It is a three point discrete approximation

Table 1: Decomposed Lower Triangular Cholesky Matrix

Seed	θ_1	θ_2
θ_1	1	0
θ_2	-0.2334	0.9724

for the standard Normal distribution with probabilities 0.185, 0.630, and 0.185 assigned to the percentiles 0.05, 0.5 and 0.95. The three discretized possible outcomes of the distribution of $\theta_{i,1}$ are therefore 103.31, 109.95, and 116.87, the inverse at the 5^{th}, 50^{th}, and 95^{th} percentile of the Gamma distribution describing $\theta_{i,1}$, respectively.

Due to the dependence between $\theta_{i,1}$, and $\theta_{i,2}$, the discrete approximations for $\theta_{i,2}$ are contingent on the outcomes of $\theta_{i,1}$. We first apply the Cholesky factorization to decompose the correlation matrix into a lower triangular Cholesky matrix shown in Table 1 to assist the calculation of the dependent uniform variables. The dependent uniform u_2 given the outcomes of $u_{\theta_{1,1}}$. Using formula (15) for the bivariate case, we can calculate u_2 as follows:

$$u_2 = \Phi(-0.2334\Phi^{-1}(\alpha_1) + 0.9724\Phi^{-1}(\alpha_2)) \tag{16}$$

For instance, when the outcome of $u_{\theta_{i,1}}$ is 0.5 (i.e., the 50^{th} percentile), the conditional distribution for $u_{\theta_{i,2}}|u_{\theta_{i,1}} = 0.5$ is calculated for the 5^{th}, 50^{th}, and 95^{th} percentiles using (16), yielding the three contingent outcomes of $u_{\theta_{i,2}}$ to be 0.0548, 0.5, and 0.9451, respectively. Therefore, the three discrete outcomes of $\theta_{i,2}$ given $\theta_{i,1} = 109.95$ are therefore 20.10, 24.86 and 30.32, the inverse of the conditional distribution for $\theta_{i,2}$ at the 5^{th}, 50^{th}, and 95^{th}, respectively. There are nine ($3^2 = 9$) possible states taking into account the dependent pair of risks ($\theta_{i,1}$, $\theta_{i,2}$).

4.2 Step 2: Reformulation UPP using Tree and Optimizing over the Tree

Building on the parameter uncertainties described by the dependent state of nature tree, we model the firm decisions by characterizing cash flows for each crop contingent on each possible parameter realization, while considering various constraints. Given the investment opportunities and risk exposures, each crop's decision making process is modeled by a tree. We then consolidate the multiple crop-level decisions to obtain expected profit cash flows from the crop portfolio. With the copulas-based dependent tree depicting the yield distributions and therefore profit for each

seed, we can optimize the resources allocation across the crop portfolio:

Proposition 4 *Using the discretized tree, the UPP formulation can be approximated as*

$$APP: \max_{\mathbf{Q}\geq 0} \Pi^u(\mathbf{Q}) = -\sum_{i=1}^{n} c_i q_i + \sum_{k_1}\sum_{k_2}\left(r_{k_1} r_{k2} \int_{x_n} \dots \int_{x_1} \Pi(\mathbf{Q},\mathbf{x}) f(\mathbf{x};\theta_{1k_1},\theta_{2k_1k_2}) dx_1 dx_2 \dots dx_n\right) \tag{17}$$

$$s.t.\ g(\mathbf{Q}) \leq \mathbf{L} \tag{18}$$

where $r_{k_1} r_{k2}$ represent the probability weights for the discretization. Finally, we establish that a unique solution will be obtained using this formulation.

Proposition 5 *If UPP is concave, APP in (17)-(18) will have a unique solution.*

This result formally finishes the technical development. We next discuss a numerical illustration for typical industry parameters.

5 Benefits from Incorporating Judgmental Errors into Portfolio Decision Making: Illustrative Numerical Study

We now present an illustrative example to identify the benefit from explicitly incorporating parameter uncertainty while making supply chain decisions as compared to when this uncertainty is ignored. We will focus on two problems. The first problem in where a farmer decides how much area to allocate from his land to a biofuel crop. In the second problem we focus on the ranking and selection problem faced by a firm that has a number of crop-variety candidates for a future development and it selects the top x% crops based on say the 25th quantile of the yield distribution.

5.1 Farmer's land allocation problem

In this problem the farmer has an area of Q acres that he will split between a conventional crop and a biofuel crop, in amounts q_1 and q_2. It follows that $q_1 + q_2 \leq Q$. The farmer also has a firm order of units D_1 and D_2 from the conventional crop and biofuel crop respectively. The tilling cost per acre is c_1 and c_2 respectively, and the price per unit of output is p_1 and p_2 respectively. The uncertain yield from the two crops are denoted by random variables Y_1 and Y_2

respectively. Consistent with the prevalent situation in the agribusiness industry, we consider that credible parameters of the distribution of yield uncertainty for Y_1 are available from prior data, but the parameters for the distribution of Y_2 are obtained from expert judgment and therefore are subject to an uncertainty. We assume both yields to be normally distributed consistent with prior agribusiness literature, and used the same parametric uncertainty as discussed earlier from Bansal et al. (2017).

We first specify the known parameter problem in which we assume that the parameters for Y_2 are equal to the mean estimates. The problem is stated as:

$$\text{KPP:} \max_{\mathbf{Q}\geq 0} \Pi^k(\mathbf{Q}) = -\sum_{i=1}^{n} c_i q_i + p_1 \mathbf{E}_{Y_1}[min(q_1 Y_1, D_1)] + p_2 \mathbf{E}_{Y_2}[min(q_2 Y_2, D_2)] \tag{19}$$

$$\text{s.t. } q_1 + q_2 \leq Q \tag{20}$$

We note that for normally distributed yields, this problem is solved easily by expanding the last two terms and then using a standard gradient based solution. For example, the second term is written as:

$$\mathbf{E}_{Y_1}[min(q_1 Y_1, D_1)] = \int_{y_1 \leq D_1/q_1} q_1 y_1 f(y_1) dy_1 + \int_{y_1 > D_1/q_1} D_1 f(y_1) dy_1.$$

Using the properties of the normal distribution we can rewrite this term as:

$$\mathbf{E}_{Y_1}[min(q_1 Y_1, D_1)] = q_1 \left(-\sigma_{y_1}^2 \Phi(\frac{D_1}{q_1}) + \mu_{y_1} \phi(\frac{D_1}{q_1}) \right) + D_1 \left(1 - \Phi(\frac{D_1}{q_1}) \right).$$

By substituting this expression for the conventional crop and a similar expression for the biofuel crop in the KPP formulation, we obtain a tractable expression of the expected profit which can be optimized using standard optimization routines.

We next specify the unknown parameter problem as follows:

$$\text{UPP:} \max_{\mathbf{Q}^u\geq 0} \Pi^u(\mathbf{Q^u}) = -\sum_{i=1}^{n} c_i q_i^u + \mathbf{E}_{Y_1}[min(q_1^u Y_1, D_1)] + \int_{\theta_{21}} \int_{\theta_{22}} \mathbf{E}_{Y_2}[min(q_2^u Y_2, D_2)] f(\theta_{21}, \theta_{22}) d\theta_{21} d\theta_{22} \tag{21}$$

$$\text{s.t. } q_1^u + q_2^u \leq Q \tag{22}$$

To solve the UPP, we discretized the problem focusing the second yield uncertainty. Specifically, Figure 1 illustrates the constructed normal-copulas dependent tree for Seed 1. In the constructed tree, the first uncertainty is a gamma distribution for the yield's mean and the second uncertainty is another gamma distribution for the yield's standard deviation. The uncertainties evolve in the tree structure as a sequence of dependent uncertainties. The conditional relationship between the two uncertainties θ_{21} and θ_{22} is easy to see in the tree. For example, if the outcome for θ_{11} is high, then the conditional gamma probability distribution for θ_{22} tends to have lower values, reflecting the negative correlation between these two variables. We optimized the problem (21)-(22) using the tree shown in Figure 1 for a large number of combinations for costs c_i, prices p_i, and yield distribution of the commercial crop. .

To determine the benefit from using the KPP formulation over the UPP formulation, we determined the percentage optimality gap for each scenario as follows:

$$\text{OptGap} = \frac{\Pi^u(\mathbf{Q^{u*}}) - \Pi^u(\mathbf{Q}^*)}{\Pi^u(\mathbf{Q^{u*}})}\%100$$

The numerator in this expression is equal to the difference between the expected profit at inventory levels provided by the UPP formulation $\Pi^u(\mathbf{Q^{u*}})$ and the KPP formulation $\Pi^u(\mathbf{Q}^*)$, when the true parameters are not known. For a wide variety of parameters for biofuel crops and conventional crops, we found this difference to vary between 2-5%.

5.2 R&D variety selection problem

In this problem a R&D team has a number of candidate varieties *j=1,2,...,m* of biofuel crops to pursue. The team's objective is to identify the top x% of the varieties. The selection criterion is based on a pessimistic return for the farmer. Specifically, the criterion is equal to the 25th or similarly low quantile of the crop times the selling price of this crop based on its moisture and other content. Mathematically, the performance of variety m is equal to

$$\text{Perf}_m = p_m \times \text{25th Quantile of Yield}$$

The uncertain yield from each variety m is denoted as Y_m, with parameters $\boldsymbol{\theta}_m$ which are uncertain and have probability distributions as well. We next describe how the team would make these evaluations in KPP and UPP mode.

In the KPP mode, the team would treat the parameters $\boldsymbol{\theta}_m$ as known and would estimate the performance as:

$$\text{Perf}_m^k = p_m(\theta_{m1} + z_{0.25}\theta_{m2}) \tag{23}$$

In the UPP mode, the team would treat the parameters $\boldsymbol{\theta}_m$ as unknown and would estimate the performance as:

$$\text{Perf}_m^u = \int_{\theta_{m1}} \int_{\theta_{m2}} p_m(\theta_{m1} + z_{0.25}\theta_{m2}) f(\theta_{m1}, \theta_{m2}) d\theta_{m1} d\theta_{m2} \tag{24}$$

As before we assume that θ_{m1}, θ_{m2} have a marginal gamma distribution with a rank order correlation of -0.15. We omit the details for brevity here, but note that we can use the Copula based approach to approximate this performance as:

$$\text{Perf}_m^A = \sum_{k_1} \sum_{k2} p_m r_{k_1} r_{k_2} (\theta_{m1k_1} + z_{0.25}\theta_{m2k_2}) \tag{25}$$

We used the parameters from Bansal et al. (2017) for 24 distributions for crops for this analysis. for each distribution, we determined the value of the 25th quantile, the 15th quantile, and the 5th quantile using KPP and UPP, and then ranked ordered the distributions for selection. To determine the benefit from using the KPP formulation over the UPP formulation, we determined the percentage of crops that were incorrectly selected by KPP in a portfolio of top x% of the varieties. We found that at x$\sim$ 10%, there is no difference between the two portfolios. However as the fraction x increased as would be the case when the team would choose a higher fraction of crops to pursue, the error rate increases. At x $\sim$50%, the error was highest at 24% and it progressively reduced from that point on. Overall, this analysis suggests that the benefit from incorporating the judgmental errors in the portfolio selection decision is unimodal, and that it is substantial when the fraction of selected candidates it neither too high nor too small.

6 Summary

In this paper we develop an Copula based approach to explicitly model the uncertainty associated with subjective probability distributions provided by domain experts for new business uncertainties in problems involving decision making under uncertainty. Our domain focus is in the context of biofuels. This industry is in nascent stage. There is a growing emphasis on developing new varieties of crops that can be used to provide biomass to convert to biofuel. Unfortunately prior data are not sufficient to develop statistical models to assist in selecting newer varieties. This industry relies on expert judgements to construct distributions of yields from various varieties to identify varieties to pursue further and to recommend to farmers to grow in their fields. We show a farmer who makes a decision to invest in a portfolio of such crops can suffer a loss of profit by 2-5% by not incorporating the uncertainty in subjective probability distributions for uncertain yields. The implications for more severe for R&D activities for which ignoring the uncertainty in yield distribution estimates can lead to an incorrect selection of varieties for develop by as much as 24%.

References

Marc Annacchino. *New product development: from initial idea to product management*. Elsevier, 2003.

Erin Baker and Senay Solak. Optimal climate change policy: R&d investments and abatement under uncertainty. *Production and Operations Management*, 2011.

Saurabh Bansal and Mahesh Nagarajan. Product portfolio management with production flexibility in agribusiness. *Operations Research*, 65(4):914–930, 2017.

Saurabh Bansal, Genaro J Gutierrez, and John R Keiser. Using experts noisy quantile judgments to quantify risks: Theory and application to agribusiness. *Operations Research*, 65(5):1115–1130, 2017.

J Eric Bickel, Larry W Lake, John Lehman, et al. Discretization, simulation, and swanson's (inaccurate) mean. *SPE Economics & Management*, 3(03), 2011.

George Casella and Roger L Berger. *Statistical inference*, volume 2. Duxbury Pacific Grove, CA, 2002.

Robert T Clemen and Terence Reilly. Correlations and copulas for decision and risk analysis. *Management Science*, 45(2):208–224, 1999.

Robert T Clemen, Gregory W Fischer, and Robert L Winkler. Assessing dependence: Some experimental results. *Management Science*, 46(8):1100–1115, 2000.

Roger M Cooke. Experts in uncertainty: opinion and subjective probability in science. 1991.

Paul Embrechts, Filip Lindskog, and Alexander McNeil. Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich*, 2001.

Edward W Frees and Emiliano A Valdez. Understanding relationships using copulas. *North American actuarial journal*, 2(1):1–25, 1998.

Vishal Gaur, Saravanan Kesavan, Ananth Raman, and Marshall L Fisher. Estimating demand uncertainty using judgmental forecasts. *Manufacturing & Service Operations Management*, 9(4):480–491, 2007.

Philip C Jones, Timothy J Lowe, Rodney D Traub, and Greg Kegler. Matching supply and demand: The value of a second chance in producing hybrid seed corn. *Manufacturing & Service Operations Management*, 3(2):122–137, 2001.

D.L. Keefer and S.E. Bodily. Three-point approximations for continuous random variables. *Management Science*, 29(5):595–609, 1983.

Janne Kettunen and Ahti Salo. Estimation of downside risks in project portfolio selection. *Production and Operations Management*, 26(10):1839–1853, 2017.

Dennis V Lindley. Using expert advice on a skew judgmental distribution. *Operations research*, 35(5): 716–721, 1987.

RB Nelson. An introduction to copulas lecture notes in statistics 139 springer verlag new york. 1999.

Olivier Toubia. Idea generation, creativity, and incentives. *Marketing Science*, 25(5):411–425, 2006.

Tianyang Wang and James S Dyer. A copulas-based approach to modeling dependence in decision trees. *Operations Research*, 60(1):225–242, 2012.

Submitted to *Operations Research*
manuscript

Optimal Aggregation of Individual Judgmental Forecasts to Support Decision Making in a R&D Program

Saurabh Bansal
The Pennsylvania State University, State College, 16803, sub32@psu.edu

Genaro J. Gutierrez
The University of Texas at Austin, Austin, 78712, genaro@austin.utexas.edu

This paper is based on managing agribusiness research and development (R&D) activities in a multi-year, multi-institution industrial research program conducted by the Federal Aviation Authority (FAA) with the objective of developing new varieties of perennial energy-grasses and technologies that can provide biomass based jet fuel. Such research programs tend to be organized in several teams with experts from different domains who collectively pursue development of new grass varieties and technologies. Typically experts in each team provide individual point forecasts for the energy-density of candidate varieties. The individual forecasts tend to differ, and there is often a need to aggregate these point judgments into measures of the potential and the risk for the energy-density of each grass variety. There was also an awareness that (i) collectively team members brought both complementary or substitutive perspectives to the teams, and (ii) some experts were better at estimation than others. The literature does not provide a systematic approach to aggregate multiple point forecasts into actionable signals while explicitly accounting for these expertise related factors. In this paper we develop a new characterization of multiple point forecasts provided by experts, and use it in an optimization framework to deduce actionable signals including the mean, standard deviation, or a combination of the two for probability distributions. This framework consists of three steps: (i) calibrate experts' point forecasts to determine which quantile they provide on average, when asked for forecasts, (ii) quantify the precision in the experts' forecasts around their average quantile, and (iii) use this calibration information in an optimization framework to deduce the signals of interest. We also show that precision and accuracy in expert judgments are complementary in terms of their informativeness.

Key words: Agribusiness R&D, Risk-return tradeoff, Probability distribution.

1. Introduction and research focus

In research and development (R&D) program management there is a fundamental trade-off between the risk and return associated with various candidate projects. The risk is typically measured as the variability in the potential of a project and the return is measured as the expected (average) potential of the project. The risk-return trade-off is operationalized in several ways in R&D program management: Resource allocation decisions are often made based on average potential less a fraction of the uncertainty in the potential of candidate projects in the spirit of the traditional Markowitz model (Chopra et al. 2011); portfolio selection decisions in stage gate research pipelines are made based on the probability of exceeding a performance threshold (Kettunen and Salo 2017); efforts to collect more information for candidate projects in a portfolio are driven by perceived uncertainty in the potential of the projects (Jaafari 2001). Project selection decisions are central to R&D management, and a quantification of the uncertainties associated with various candidate projects is necessary to make these decisions. Yet in R&D programs it is typical to explore new products or processes for which historical data are not available to quantify the uncertainties. Hess (1993) discusses R&D efforts at multiple firms in this environment, observing that "Selecting research and development projects is always difficult because data for assessment are seldom available".

In the absence of historical data, firms need to rely on experienced managers or domain-experts to evaluate the potential of candidate projects. Domain-experts often provide only point forecasts or judgments for the most likely potential of a project. Yet in applications such as R&D management, the need to "produce predictive distributions is compelling" in order to make risk-return trade-offs (Geweke and Amisano 2011). *When provided with point forecasts by multiple experts, how can a manager deduce the risk and average return for a candidate project and determine the action she should take for R&D management?* – in this paper we develop an analytical framework to answer this question. We next describe a academic-industry research program that motivated this paper and illustrative field-data collected in this multi-year effort to further explain the forecasting and decision making problem.

1.1. Agricultural R&D to plan for regional sustainability and role of domain expertise

Since 2016, the FAA has sponsored an industry-building research program in the northeastern U.S. The program seeks to boost regional energy generation in the Northeastern U.S. by developing high-energy-density perennial grasses – grasses that release high amounts of heat when burned – that can be used translated into jet fuel. This R&D program is driven by several socio-economic and sustainability needs. A number of regional pockets in the U.S. have marginal lands that cannot

support cultivation of traditional food or cash crops. In many cases the limited fertility of these lands is due to geological factors such as rocky terrain. In other cases these lands are the sites of mineral/coal mines whose top soil was removed for mining activities but the mines are defunct now. Perennial grasses offer a source of income to the owners of these otherwise barren-lands. Some specific perennial grasses thrive well in marginal lands and, once dried, provide fuel that landowners can sell to energy-plants. At a macro-level these grasses contribute towards making the US self-reliant for energy, a goal that the federal government actively pursues. Perennial grasses also reduce the use of fossil fuels for energy-generation and contribute towards a reduction in CO_2 emissions. The program has nearly one dozen participating academic institutions industry partners.

The R&D activities at the program of interest to us focus on developing better varieties of these grasses. Research-scientists teams are tasked with developing varieties that provide higher energy-yields in specific climate and geographical conditions in the Northeast US. Each team is comprised of experts from commercial firms, industry research firms, and academic research and outreach divisions. To develop new varieties, each team starts with a broad range of parent-strains (varieties with unique genetic structure), develops new hybrid strains by cross-pollination, and then progressively explores hybrids of hybrids. The selection of specific hybrids over others during this progressive selection is based on objective criteria including the average energy-yield, the variance in the yield, and specific quantiles say 75th or 25th quantiles of yields.

Unfortunately credible predictive models are not available to provide estimates of performance of new hybrids on these criteria, for several reasons. These grasses are not an established part of animal or food chain. Few scientific efforts have been made in the past to collect quantitative information about parent strains or to develop hybrids, and only scant data are available for the energy-yields of various strains. Efforts made to use regressions on genetic markers to predict yields have been unsuccessful due to data sparsity issues. As a result there is a need to rely on expertise of domain experts to estimate yield distributions. To further validate the need to rely on expert-judgments for developing new varieties of perennial grasses, we conducted an anonymous survey of nearly one hundred researchers associated with biofuel supply chains. We asked the participants to identify the major risks associated with developing biofuels based supply chains. For each risk identified by a participant, we further asked (i) whether the participant believed that the risk was internal or external to biofuels, and (ii) whether data were available to quantify the risk or if one would need to rely on expert-opinion or both. We also asked for other information e.g., research focus of the expert (variety development, combustion etc), but we will restrict the discussion here to the context of variety development. The results of this survey are shown in Figures 1 and 2. Figure 1 shows that nearly one-third of participants believed that developing better varieties was

a major source of risk to the research program (the third bar from the bottom). Furthermore, as shown in Figure 2 a majority of these researchers believed that additional data alone would not be sufficient in the near future to support this R&D activity and that expert-judgments would be necessary to support this activity.

Figure 1 Survey results for the identification major risks associated with biofuel supply chains

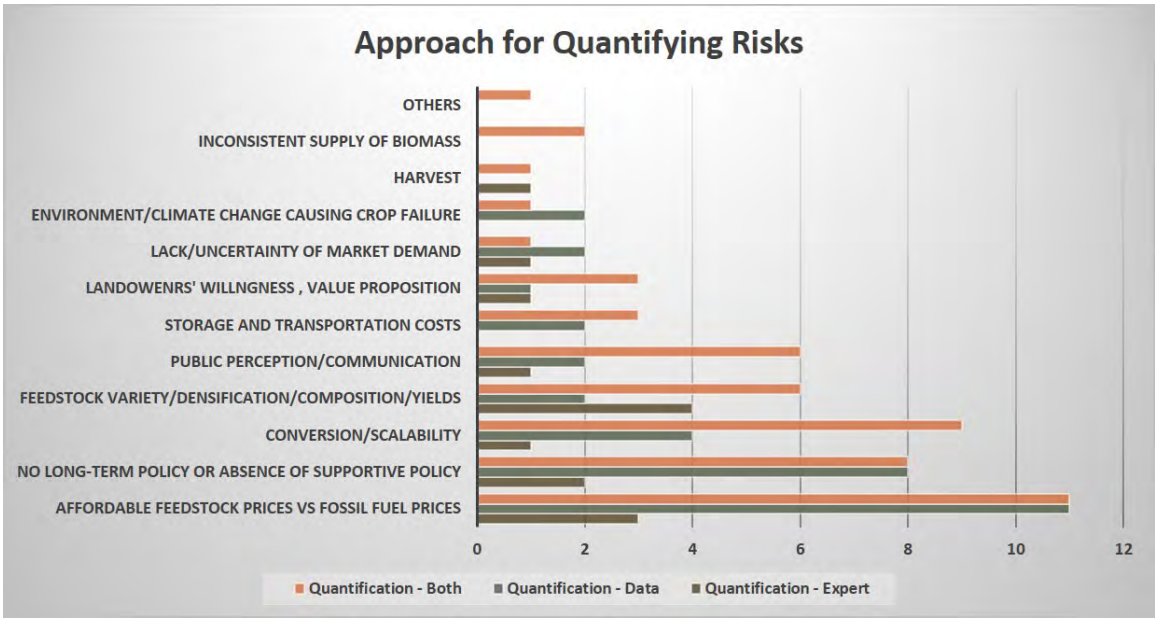

Figure 2 Survey results for quantification of major risks associated with biofuel supply chains

In this domain-expertise driven estimation the team members typically provide point forecasts (their judgments for most likely values) for the yields of new hybrids. A key challenge then is then

Hybrid	Prediction by Expert 1	Prediction by Expert 2	Prediction by Expert 3	Mean from Data	StdDev from Data
XXXX01	9.7	15	13.5	11	2.77
XXXX02	20.5	34.8	25.2	24	7.34
XXXX03	23.8	34.2	26.1	25	6.3
XXXX04	26.8	31	24.8	24	4.75
XXXX05	18.6	22.4	17.6	17	3.67
XXXX06	26.9	34.2	26.9	26	5.61
XXXX07	10	16.8	12	11	3.96
XXXX08	7.3	12.8	8.2	8	1.44
XXXX09	13.9	22.3	16.7	16	4.32
XXXX10	11	23.5	13.7	13	4.21
XXXX11	19	32.7	22.5	21	6.42
XXXX12	12.1	22.8	14.7	14	4.03
XXXX13	15.8	29.8	23.7	18	4.53
XXXX14	9.3	13.1	10.3	10	2.16

Table 1 Sample data from three experts in a team.

to aggregate these point forecasts into a metric that could be used to compare various alternatives. The research teams are typically composed of personnel in different fields (e.g., field trial experts, bio-statisticians etc.) who bring different technical perspectives; and personnel with different expertise levels (e.g., experienced industry-scientist, relatively new post-doctoral researcher etc). It was desirable to account explicitly for these individual differences when preparing aggregate metrics. For example, columns 2, 3, and 4 in Table 1 show the judgments provided by three experts in a team for two dozen hybrids. Columns 4 and 5 show the summary statistics for actual data collected at a later date from multiple crops of these hybrid strains. We focus on using this information to inform aggregation of the expert-provided point judgments for portfolio selection and investment decisions for future hybrids.

1.2. Our Research Focus and Contributions to Theory and Practice

In this paper we first develop an algorithmic approach to deduce actionable signals in the form of the mean or standard deviation or a combination of the two obtained from point judgments provided by multiple experts when experts' forecasts can be calibrated using data such as the one shown in Table 1. A key feature of this approach is that it characterizes the judgments of experts on two-dimensions, (i) average location on the probability distribution curve, and (ii) precision centered around this location. This characterization also enables us to develop an objective measure of information contained in judgments that is consistent with the existing theory of order statistics. We then use this characterization into an optimization model to aggregate point judgments for future uncertainties and obtain an actionable signal. This approach naturally leads to a protocol to select a subset of experts from many, when this selection must be made due to cost or other reasons. We also establish analytically that the information content in an expert's judgments depends both

on the bias and the noise in his judgments, and further, these two dimensions are complementary. Specifically, an expert whose judgments on average are in the vicinity of the mode of a distribution must also provide these judgments with a higher consistency in order to be as informative as an expert who on average provides judgments in a tail.

The rest of the article is organized as follows. In Section 1.3, we discuss the relevant literature and our contributions. In Sections 2 – 4 we discuss the model, solution, structural properties, and a special case. In Section 5, we discuss model extensions; in Section 6 we conclude with a summary and directions for future research.

1.3. Literature Review and Overview of Our Approach

Three streams of research are relevant to our work: (i) individual expert/managerial forecasting, (ii) connecting point forecasts from multiple experts to probability distributions when direct historical information is not available for the uncertainties of interest, and (iii) operations research developments for agribusiness. The first stream of work has focused on behavioral issues in forecasting when a single expert provides point forecasts or judgments. Prior research in this domain has documented a number of biases and inconsistencies present in expert judgments (see e.g., Sterman 1989, Bazerman and Moore 2008). This literature shows that the biases and inconsistencies in expert-judgments tend to be individual specific, i.e., all experts are not alike - some experts are more biased than others and some experts are more consistent in their judgment than others (see e.g., Agnew 2006). Our focus in this paper is on quantifying this bias and consistency in expert-judgments in group settings and then aggregating them to deduce optimal actions.

The idea of using multiple point judgments to deduce a distribution has been explored in two different ways. Barron and Stuerke (1998) adopt an approach that is agnostic of individual differences. Specifically, they showed empirically that when multiple financial analysts provide point forecasts for return of a stock, (i) the average of these forecasts is a good estimate of the mean of the probability distribution of the stock-return, and (ii) the variability measured as variance or standard deviation in the point forecasts is directly proportional to the true variance or standard deviation of the probability distribution. The proportionality constant for the variance is estimated in a calibration process. In this calibration, analysts provide point forecasts for multiple known probability distributions (similar to the data in Table 1), and the variance in the point-forecasts for each distribution is calculated (point-variances). Subsequently the true variances of the distributions are regressed on the calculated point-variances. For the regression specification with no intercept, the regression coefficient provides the proportionality constant. This model has been adopted in several domains, including operations management (Gaur et al. 2007), and finance (see

e.g., Barron et al. 1998, Johnson 2004), among others. In contrast to this body of work we develop a bottom-up analytical approach in which we focus on characterizing individual judgments. This characterization then suggests how to optimally aggregate the judgments. This focus enable us to relax some strong assumptions made in this literature. Specifically, this literature assumes that the experts provide judgments that are independent and identically distributed (Barron and Stuerke 1998). However, empirical data show that experts' judgments can be correlated (Budescu et al. 1997) and that the quality of judgments by some experts is markedly superior to others' (see e.g., Budescu and Chen 2014, Davis-Stober et al. 2015). These developments also require the presence of a large number of experts for statistical validity. However in many business situations, only three or four experts are available to provide forecasts.

Our analytical development addresses these issues. Specifically, we first use the calibration data available for each expert to identify the quantile each individual expert provides on average when asked for the most likely value, and the variability in his judgment around this quantile. We then determine the variance-covariance matrix for the variabilities in judgments for all experts. The diagonal elements of this matrix (variance) is used as a proxy for the reliability of each expert's forecasts. The off-diagonal elements quantify the correlations between the point forecasts provided by various experts. We then use this information in an optimization framework to estimate an actionable signal as a function of the individual judgments provided by the experts. This approach is based on the theory of judgmental errors discussed in prior decision analysis literature (e.g., Ravinder et al. 1988) and is grounded in statistical theory, specifically on order statistics (e.g., Lloyd 1952). Using the estimate of the average quantile provided by an expert and its variability we obtain an objective quantification of expertise which then enables us compare experts and quantify their individual contribution to the group's aggregated forecast.

In the second approach individual differences in experts' quality are accounted for. Bates and Granger (1969) consider the problem where a decision maker wants to combine two unbiased judgments when the standard deviations of the errors in both judgments are available. Winkler (1981) extend this model to more than two experts. Both articles (and the ones that follow this paradigm) assume that the error in the judgments has been quantified by determining the differences between an expert's historical forecasts and actual realizations for the same underlying uncertainty. This focus leads to a limitation where an expert's errors for say multiple products with different demand magnitudes (e.g., a product with demand in 100s while the other one in 1000s) cannot be used for quantifying judgmental errors. In contrast to this framework, our model considers that complete historical distributions are available to calibrate an expert's judgments, i.e., we can compare an expert's point judgments against complete distributions. We leverage this information to transform the judgmental errors present in point judgments into values on a standardized location-scale

distribution (e.g., on the standardized z-scale for the normal distribution). This standardization enables us to do two things. First it enables us to quantify the bias in an expert's judgments as the average quantile she provides and the noise as the variability in her judgments around this quantile. Second, we can now use the judgmental errors for multiple products using a milder assumption that an expert uses a stable mental process to provide probability judgments for multiple quantities. We also use this framework to show analytical results such as the one in Proposition 2 where we establish that expert accuracy and precision are complementary in nature.

A small but growing body of literature has explored the use of calibration information to improve decision making. For example, Cooke (1991), Bansal et al. (2017) and Kettunen and Salo (2017) discuss approaches to use historical information for calibration and then use this information to improve forecasts for future quantities. Our paper is closest to Kettunen and Salo (2017) with some important differences in methodology and context which make their results not applicable to our problem. Specifically, Kettunen and Salo (2017) study project selection decisions based on quantile-estimates of cash flows obtained from a Monte-Carlo simulation model, and show that pure random variations can lead to systemic biases in this selection process. They suggest using calibration data to adjust the quantile values obtained in simulation models to avoid this bias. In contrast, in our context experts provide point forecasts and these forecasts need to be aggregated to deduce a specific quantile or alternatively the mean and standard deviation of the uncertainty for project selection decisions. Our characterization of individual judgments and optimization model used for optimal aggregation is also different in our problem context.

Finally, a large body of literature exists on using operations research approaches to manage agribusinesses. See Weintraub and Romero (2006), Bjørndal et al. (2012) for latest reviews and syntheses of these developments. The discussion in Weintraub and Romero (2006) suggests that risk and uncertainty are important aspects of agribusinesses and a large body of work focuses on production planning models under risk, but the issue of estimating risks especially for agribusiness R&D has not received only scarce attention. Our work addresses this gap.

2. Model Preliminaries

In Section 2.1, we discuss the problem context, information available, notation and assumptions. In Section 2.2, we discuss the functional form of actionable signals considered in our development.

2.1. Problem Context and Notation

We consider $j = 1, 2, ..., m$ experts who provide point forecasts $\hat{\mathbf{y}} = [\hat{y}_1, \ldots, \hat{y}_m]^{\mathrm{T}}$ for an uncertain quantity or random variable Y. The business objective is to use the estimates $\hat{\mathbf{y}}$ and determine

the value of the signal or function $g(\mu, \sigma)$, where μ, σ are the mean and standard deviation of distribution of Y respectively. The signal could be of any one of multiple forms based on context driven decision criteria. A common practice in the agribusiness domain to rank order varieties of a crop based on specific quantiles of yield distributions and select the top varieties in the list (see, e.g., Batur and Choobineh (2010) and Kettunen and Salo (2017) for other applications of this criterion in R&D management). Second, consistent with financial portfolio theory, the selection decision could also be made based on a linear combination of the mean yield and the (negative of) yield uncertainty. Two special cases of this paradigm are the ones where decisions are made based on only mean and only standard deviation. We capture these multiple criteria using the following form of the signal:

$$g(\mu, \sigma) = \alpha_1 \mu + \alpha_2 \sigma \tag{1}$$

Suppose a team has decided that it will rank order varieties based on the yield values corresponding to probability p or select varieties for which the p^{th} quantile of energy-density yield would exceed a threshold. Then, for the widely used normal distribution for energy-yield, the decision criterion is $q = \mu + z\sigma$ where $z = F^{-1}(p; 0, 1)$ and $F(\cdot; 0, 1)$ is the standardized CDF the energy-density yield, i.e., $\alpha_1 = 1, \alpha_2 = z$. The second criterion discussed above for weighing mean and standard deviation is represented using $\alpha_1 \mu - \alpha_2 \sigma$ where α_1 and α_2 are the importance weights assigned to the mean and standard deviation of the uncertainty. The two special cases for making decision based on mean only and standard deviation only are obtained using $(\alpha_1 \neq 0, \alpha_2 = 0)$ and $(\alpha_1 = 0, \alpha_2 \neq 0)$ respectively.

In the immediate development, we assume that the underlying uncertainty has a location-scale distribution. This focus is consistent with an established body of literature including Barron and Stuerke (1998), among others. This focus is (i) not overly restrictive from practice perspective (ii) provides analytical tractability and insights to the problem we study, and (iii) can be extended to other distributions. Location-scale distributions include some frequently used distributions in operations literature and practice such as the Normal distribution, Logistic distribution, and Gumbel distribution, and as such is consistent with prior literature on R&D management such as Kettunen and Salo (2017). For these distributions, our immediate technical development provides closed form solutions for aggregating point judgments. In Section 5 we extend the technical development to distributions that are obtained by a monotone transformation of location-scale distributions, e.g., the Johnson family of distributions that are obtained from transformations of the normal distribution.

The mean and standard deviations of location-scale distributions are linear functions of the location and scale parameters μ_1 and μ_2 respectively. For notational ease in the remainder of the

paper we will focus on estimating the form $g(\mu_1, \mu_2) = \alpha_1\mu_1 + \alpha_2\mu_2$. Consistent with prior literature (e.g., Cooke (1991)), we first consider on the case where data are available to calibrate the point judgments provided by individual experts, similar to Table 1. Specifically we address the case where a team of experts was asked to provide individual point forecasts $\mathbf{X}$ with elements $\hat{x}_{ij}$ for each expert j and calibration distribution for uncertainty i, for $i = 1, 2, ..., n$. Calibration distributions are empirical distributions for different uncertainties from similar contexts that are available from prior data collection or analysis. The probability distribution function for calibration distribution i is denoted as $F_i(\cdot; \boldsymbol{\mu}_i)$ where $\boldsymbol{\mu}_i = [\mu_{i1}, \mu_{i2}]^{\mathrm{T}}$ are respectively the location and scale parameters of the distribution. The vector of these distributions is denoted as $\mathbf{F}$. The experts' point forecasts will be compared with the historical frequency of observed values; the information generated in this calibration process, $(\mathbf{X}, \mathbf{F})$ will be used to process the experts' forecasts $\hat{\mathbf{y}}$ for the new random variable of interest, Y, for which historical data do not exist. In Section 5 we extend the results to the case when these calibration distributions are not available.

2.2. Actionable Signals as Functions of Judgments and Outline of Development

We seek to obtain the estimate $g(\hat{\mu}_1, \hat{\mu}_2)$ of $g(\mu_1, \mu_2)$ as an aggregation-function of the m forecasts $\hat{\mathbf{y}}$ and calibration data $(\mathbf{X}, \mathbf{F})$,

$$g(\hat{\mu}_1, \hat{\mu}_2) = h(\hat{\mathbf{y}}; \mathbf{X}, \mathbf{F}) \tag{2}$$

In the development that follows we first transform the calibration data; define an optimality criterion; identify the optimal functional form $h(\hat{\mathbf{y}}; \mathbf{X}, \mathbf{F})$ of the signal ; and then estimate it. In order to do this, we characterize the point judgments provided by each expert using two quantities – the average of the quantile corresponding to the judgments he provides, and a measurement of the consistency (or alternatively the spread) of his judgments. The details of this step are in Section 3. In Section 4, we use this characterization in a least squares minimization framework. We first show in Proposition 3 that for these transformed judgments a linear functional form $g(\hat{\mu}_1, \hat{\mu}_2) = \mathbf{w}^{\mathrm{T}}\hat{\mathbf{y}}$ provides unbiased estimates of $g(\mu_1, \mu_2)$ with lowest variance, where the vector $\mathbf{w} \equiv [w_1, w_2, ..., w_m]^{\mathrm{T}}$ represents importance weights for the judgments, and then provide the weights in Proposition 4. This development has several practical merits. First, it is consistent with the classical signal processing theory in which multiple signals are aggregated using a weighted average to constitute an actionable signal (see e.g., Bates and Granger 1969). Furthermore the weights directly depend on the bias and consistency of the experts' forecasts for the calibration distributions, as deduced from the calibration data $(\mathbf{X}, \mathbf{F})$, directly linking experts' performance on calibration distributions with the weights assigned to their future judgments.

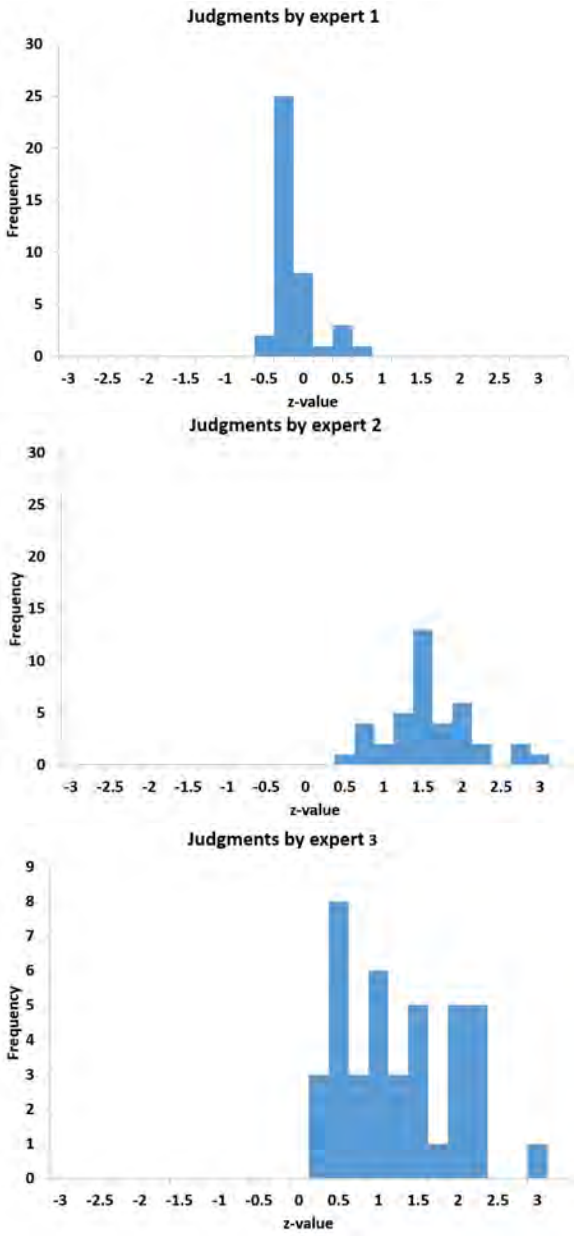

Figure 3 Transformed judgments by three experts

3. Characterizing Point Forecasts from Multiple Experts and Structural Properties

In Section 3.1, we use the calibration data to characterize experts' forecasts as an average quantile plus noise. In Section 3.2, we use this characterization to derive a metric of the predictive value for each expert's estimates.

3.1. Procedure for Characterizing Forecasts

Figure 3 provides intuition for our characterization process. The figure uses the data from Table 1 for the forecasts provided by three experts for the calibration distributions. In the first panel, we transformed the n judgments provided by Expert 1, $\hat{x}_{i1}$ for $i = 1, 2, ..., n$, into standardized z values as $\hat{z}_{i1} = F_i^{-1}\big(F_i(\hat{x}_{i1}; \boldsymbol{\mu}_i); 0, 1\big)$, and then created a histogram of these standardized values. The x-axis is defined in terms of the standardized z-value, and extends on $\mathbb{R}$. The y-axis shows the frequency count. The second and third panels show these histograms for the judgments provided by experts 2 and 3 respectively. This transformation of responses into the z-domain has several advantages. The z-domain is a scale-free domain and allows a comparison of the experts' point forecasts across all calibration distributions. For example, the three panels show that Expert 1 is more likely to provide judgments that are close to the location parameter of a distribution (i.e., near z-value of 0), whereas Experts 2 and 3 tend to provide judgments in the right tail (i.e. $z > 0$) on average. We can further conclude that the judgments of Expert 1 tends to be more consistent as they are dispersed in a small range, and the judgments of Experts 2 and 3 are relatively less consistent as they have a larger dispersion.

Mathematically, the point forecasts provided by each expert j on all calibration distributions $i = 1, 2, ..., n$ are translated into the corresponding standardized z-values, as

$$\hat{z}_{ij} = F_i^{-1}\big(F_i(\hat{x}_{ij}; \boldsymbol{\mu}_i); 0, 1\big) \tag{3}$$

We then take an average of all these standardized z-values for each expert j to identify the quantile she provides, on average:

$$\hat{\bar{z}}_j = \frac{\sum_i \hat{z}_{ij}}{n} \tag{4}$$

This calculation provides a direct characterization of each expert's average behavior in the probability domain. Specifically, we can compute the probability $\hat{\bar{p}}_j = F(\hat{\bar{z}}_j; 0, 1)$ that on the average each expert provides. For example, for the specific data in Figure 3, one can perform the calculations above and deduce that Expert 1 provides the 39th quantile on average (with average z-score of -0.29), Expert 2 provides the 98th quantile on average (with average z-score of 2.00), and Expert 3 provides the 80th quantile on average (with average z-score of 0.83).

To characterize the consistency with which each expert provides her average quantile, we assume an additive error model $z(\hat{x}_j) = z(p_j) + \epsilon$ as suggested in Ravinder et al. (1988) and Wallsten and Budescu (1983) for modeling judgmental errors, and we estimate the variance-covariance matrix Ω for the judgments for all experts around their average quantile $\hat{\bar{z}}_j$:

$$\Omega = \begin{bmatrix} \omega_{11} & \dots & \omega_{1m} \\ \vdots & \vdots & \vdots \\ \omega_{m1} & \dots & \omega_{mm} \end{bmatrix} \tag{5}$$

where the individual elements are

$$\omega_{jj} = \frac{\sum_{i=1}^{n}(\hat{z}_{ij} - \hat{\bar{z}}_j)^2}{n-1};\ j = 1, 2, ..., m \tag{6}$$

$$\omega_{kj} = \frac{\sum_{i=1}^{n}(\hat{z}_{ij} - \hat{\bar{z}}_j)(\hat{z}_{ik} - \hat{\bar{z}}_k)}{n-1};\ k, j = 1, 2, ..., m;\ k \neq j \tag{7}$$

Each diagonal element $\omega_{jj};\ j = 1, 2, ..., m$ in the matrix denotes the variance in the judgments for expert j around the quantile $\hat{\bar{z}}_j$. The non-diagonal elements $\omega_{kj}; k \neq j$ denote the covariances between the noise in the estimates of experts k and j around their average quantiles, $\hat{\bar{z}}_k$ and $\hat{\bar{z}}_j$ respectively. When the matrix Ω is a diagonal matrix, it implies that the experts' variations around their central quantile are mutually independent. For more general matrices, the matrix captures correlations present between pairwise judgments. For the specific data in Figure 3 for three experts, this matrix is obtained as $\Omega = \begin{bmatrix} 0.09 & -0.07 & 0.04 \\ -0.07 & 0.55 & 0.29 \\ -0.04 & 0.29 & 1.11 \end{bmatrix}$. Based on this matrix we conclude that the judgments of Experts 2 and 3 are mildly correlated, while the judgments of Expert 1 show only a small correlation with the judgments of Experts 2 and 3.

The above calibration procedure provides information about the average forecast behavior of every expert and the variability in his judgments. We next provide a statistical grounding to this characterization to measure the quality of the judgments of each expert.

3.2. Quantification of Expertise into Equivalent Sample Size and Complementarity of Bias and Precision in Point Judgments

We start by noting a useful result for the quantiles of sampling distributions. The sample variance of quantile x corresponding to probability p for a sample size n is equal to $Var(x) = p(1-p)/\big(n[f(F^{-1}(p))]^2\big)$ where f is the pdf with parameters μ_1, μ_2 and F is the corresponding cdf (Stuart and Ord 1994). This result states that the sample variance is higher for tail quantiles (i.e., for small or large values of p) and is lower for quantiles in the center of the distribution. Intuitively tail realizations from a parent population are less frequent in a sample and therefore its estimation has a higher variability. In contrast, realizations from the central part of the distribution, say the mode, tend to be more common in a random sample and hence its estimation has lower variability. We next adopt this result for our context and specialize it for location-scale distributions. From the definition of location-scale distributions the quantile $x = \mu_1 + z\mu_2$ where z is the quantile corresponding to probability p for the standardized pdf $f_s(\cdot)$ and standardized cdf $F_s(\cdot) = F(\cdot; 0, 1)$ with $\mu_1 = 0, \mu_2 = 1$. It follows after some algebra that $Var(z) = p(1-p)/\big(n\sigma^2[f(F^{-1}(p))]^2\big)$. Now using the property $f = f_s/\mu_2$, we obtain that the variance of the z value corresponding to probability p obtained from a sample of size n is equal to $Var(z) = p(1-p)/\big(n[f_s(z)]^2\big)$.

In our development in the previous section, we calibrated each Expert j and quantified the variance in his judgments for the z-value corresponding to probability $\hat{\bar{p}}_j$, as ω_{jj}. By equating this variance with the corresponding quantile sampling variance, we can quantify the expert's consistency to be equivalent with a sample of N_j iid observations:

Proposition 1 (Equivalence of Individual Judgments with Random Sample Size) *The precision of Expert j in providing the judgments for the $\hat{\bar{p}}_j$th quantile is equal to the precision of the estimate obtained of the same quantile obtained from an iid sample of size N_j, where*

$$N_j = \frac{\hat{\bar{p}}_j(1-\hat{\bar{p}}_j)}{\omega_{jj}\left(f_s(\hat{\bar{z}}_j)\right)^2} \tag{8}$$

A salient feature of the equivalent sample size is that it accounts for both accuracy and the precision of the expert judgments. The accuracy is specified in terms of the average quantile $\hat{\bar{p}}_j$ of the expert's judgments and the precision information is captured by the variability ω_{jj} with which the experts provides his judgments. This result has several implications. First, we can develop an objective ranking of multiple experts in terms of their informativeness. Specifically, if the equivalent sample sizes of two experts are N_1 and N_2 with $N_1 > N_2$, then it follows that *accounting for the average quantile these experts provide*, Expert 1's judgments provide more information.

Second, we can use this result to draw isoquant curves for equivalent sample size, i.e., we can deduce a continuum of combinations of ω_{jj} and $\hat{\bar{p}}_j$ such that at each of these combinations the expert's judgments provide a constant level of information. Specifically, using the relationship between probability and z-scores $\hat{\bar{z}}_j = F_s^{-1}(\hat{\bar{p}}_j)$ we can rewrite the expression $N_j = \hat{\bar{p}}_j(1-\hat{\bar{p}}_j)/\left(\omega_{jj}\left(f_s(\hat{\bar{z}}_j)\right)^2\right)$ as $\omega_{jj} = \hat{\bar{p}}_j(1-\hat{\bar{p}}_j)/\left(N_j\left(f_s(F_s^{-1}(\hat{\bar{p}}_j))\right)^2\right)$. Observe that the R.H.S. of this expression is a function of only $\hat{\bar{p}}_j$ for a given value of N_j. It follows that for any average quantile $\hat{\bar{p}}_j$ of the judgments there exists a variability ω_{jj} such that the information in the judgments as measured in terms of the equivalent sample size remains the same. Figure 4 shows the isoquant curve for the judgments of Expert 1. The x-axis shows the average quantile and the y-axis shows the variability in judgments. Specifically, Expert 1 provides judgments for the 39th quantile with a variability of 0.09 units in the z-domain, with the equivalent sample size of 18 data points. His equivalent sample size would have remained the same if he had provided judgments at any other point on the curve, e.g., for the 85th quantile on average with variability level of 0.13.

Third, this development suggests that the bias in expert judgments and precision in the judgments are complementary characteristics.

Proposition 2 (Complementarity of Bias and Precision in Judgments) *Let the function $\omega_{jj} = g(\hat{\bar{p}}_j)$ denote the variance in expert js forecasts such that the equivalent sample size N_j is constant. Then, $g(\hat{\bar{p}}_j)$ has a U-shape with the minimum value at the mode i.e., at z where $f_s'(\hat{\bar{z}}_j) = 0$.*

Figure 4 shows this result pictorially for the judgments of Expert 1 for a normal distribution. The isoquant curve attains its lowest value at an intermediate point (at cumulative probability value of 0.5) and then increases in both directions from this point. This trend shows that the accuracy in individual forecasts as measured by the average quantile provided by an expert and the precision in the judgments are complementary. If an expert provides judgments on average that are close to the mode of a distribution then he must also provide these judgments with a greater level of consistency to remain informative at a specific level, as compared to the case when he provides judgments on average that are in either tail. The curvature of the isoquant curve depends on the underlying probability distribution of the uncertainty. Specifically, to see this, we first rearrange the equation $N_j = \frac{\hat{\bar{p}}_j(1-\hat{\bar{p}}_j)}{\omega_{jj}\left(f_s(\hat{\bar{z}}_j)\right)^2}$ as the function $\omega_{jj} = \frac{\hat{\bar{p}}_j(1-\hat{\bar{p}}_j)}{N_j\left(f_s(\hat{\bar{z}}_j)\right)^2} = \frac{f_s(\hat{\bar{z}}_j)(1-f_s(\hat{\bar{z}}_j))}{N_j\left(f_s(\hat{\bar{z}}_j)\right)^2}$. Next we differentiate the function to obtain its gradient as $\frac{\partial \omega_{jj}}{\partial \hat{\bar{z}}_j} = \frac{f_s'(\hat{\bar{z}}_j) - 2f_s'(\hat{\bar{z}}_j)f_s(\hat{\bar{z}}_j)}{N_j\left(f_s(\hat{\bar{z}}_j)\right)^2} - \frac{2f_s'(\hat{\bar{z}}_j)f_s(\hat{\bar{z}}_j)(1-f_s(\hat{\bar{z}}_j))}{N_j\left(f_s(\hat{\bar{z}}_j)\right)^3}$ which on simplification reduces to $\frac{\partial \omega_{jj}}{\partial \hat{\bar{z}}_j} = -\frac{f_s'(\hat{\bar{z}}_j)}{N_j\left(f_s(\hat{\bar{z}}_j)\right)^2}$. Since $f_s(\hat{\bar{z}}_j) > 0$ the minimum is achieved at $f_s'(\hat{\bar{z}}_j) = 0$. This analysis shows that both the curvature of the isoquant curve as well as the point where the minimum is reached varies by the distribution family. As shown in Figure 2, when the underlying distribution is symmetric the gradient $\frac{\partial \omega_{jj}}{\partial \hat{\bar{z}}_j} = \frac{f_s'(\hat{\bar{z}}_j)}{N_j\left(f_s(\hat{\bar{z}}_j)\right)^2}$ is symmetric with a minimum at $p = 0.5$. This result is especially applicable to the R&D environments where project outcomes and their estimates are typically normally distributed (Loch and Kavadias 2002, Smith and Winkler 2006, Kettunen and Salo 2017). In contrast when the distribution is left skewed as shown in Figure 5 for a skew-normal distribution, the minimum value is reached at $p > 0.5$. This happens because the the mode is reached in the right side region of the pdf and the expert needs to provide his judgments with a greater accuracy in this region. Figure 6 for a right skewed distribution is also interpreted similarly.

The equivalent sample size N_j plays an important role in selecting a set of experts for the purposes of aggregating their judgments to deduce the function $g(\hat{\mu}_1, \hat{\mu}_2)$, which we focus on next.

4. Aggregation Problem and Results for Group Composition

In Section 4.1 we focus on estimating the function $g(\hat{\mu}_1, \hat{\mu}_2)$ for Y for aggregating individual judgments $\hat{\mathbf{y}}$. In Section 4.2 we analyze the relationships between individual expertise and the expertise of a group of experts, and in Section 4.3, we discuss optimal team selection.

4.1. Deducing an Actionable Signal Using All Judgments

Consistent with a large body of literature on deducing distribution parameters from discrete information for judgments (e.g., Keefer and Bodily 1983, Keefer and Verdini 1993, Johnson 1998),

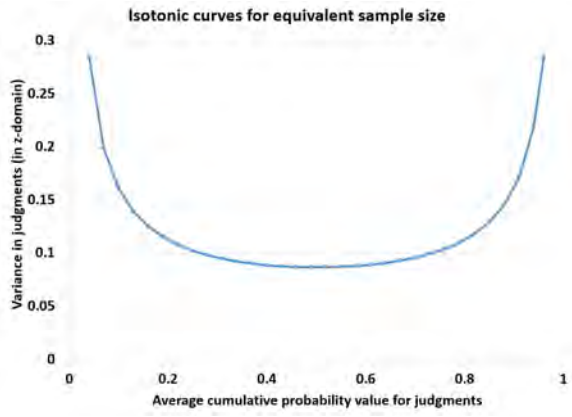

Figure 4 Isoquant curve for the normal distribution. Note that the u-shape will prevail for all symmetric distributions.

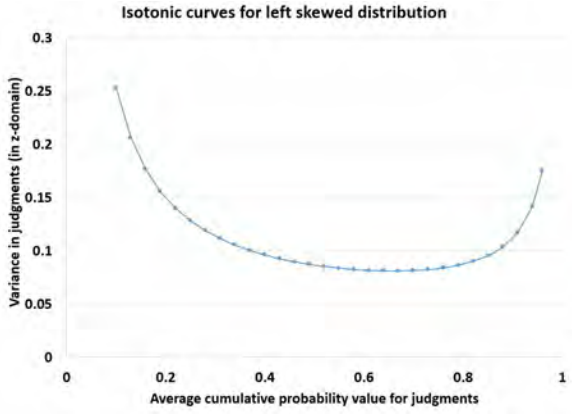

Figure 5 Isoquant curve for a left tailed skew-normal distribution with skewness factor -0.2.

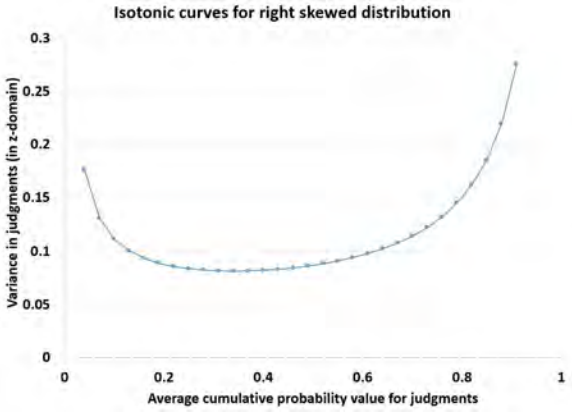

Figure 6 Isoquant curve for a right tailed skew-normal distribution with skewness factor 0.2.

we focus on obtaining an estimate of $g(\mu_1, \mu_2)$ with a minimum variance. To this end, we first note that using the calibration information for expert-judgments generated in Section 3.1 we can write

the judgments provided by each expert j as $\hat{y}_j = \mu_1 + \hat{\bar{z}}_j\mu_2 + \epsilon$ with $\mathbf{E}[\epsilon_j] = 0$. Due to this linear specification of quantiles in terms of distribution parameters, the optimal functional form $h(\cdot)$ for the estimation $g(\hat{\mu}_1, \hat{\mu}_2) = h(\hat{\mathbf{y}}; \mathbf{X}, \mathbf{F})$ is given by a linear function of the vector of judgments $\hat{\mathbf{y}}$.

Proposition 3 *For location scale distributions, the minimal variance unbiased value of the function $g(\mu_1, \mu_2)$ is obtained by a weighted linear function of judgments $\hat{\mathbf{y}}$ as $g(\hat{\mu}_1, \hat{\mu}_2) = \mathbf{w}^{\mathrm{T}}\hat{\mathbf{y}}$ where $\mathbf{w} = [w_1, w_2, ..., w_m]^{\mathrm{T}}$ is the vector of weights given to each expert's judgment.*

This result shows that a simple weighted average of judgments provides an actionable signal with desirable properties (unbiasedness and minimum variance). Additionally, a weighted average of judgments is intuitively appealing to industry managers/experts and it is easy to implement in practice. The detailed proof is in Appendix. In the proof we exploit properties of the least squares framework to establish that the estimates $\hat{\mu}_1, \hat{\mu}_2$ of the two parameters are obtained as weighted average of the observations $\hat{\mathbf{y}}$, and then exploit the linearity of $g(\mu_1, \mu_2)$ in μ_1 and μ_2 to establish that $g(\mu_1, \mu_2)$ is also estimable as a weighted average of $\hat{\mathbf{y}}$. We note that our analysis is prima-facia similar to Lloyd (1952) who show that the moments of a location-scale distributions are expressed as unique linear combinations of order statistics of these distributions. However there is an important difference in our context. In the development in Lloyd (1952) the variance-covariance matrix for the order statistics is a structural property of the distribution family and is estimated using Monte Carlo simulation or from analytical properties of distributions of order statistics, whereas in our case we use quantile judgments to estimates the parameters when the variance-covariance matrix is estimated separately in a calibration exercise of expert-judgments. As a result while the optimal weights $\mathbf{w}^*$ in Lloyd (1952) are constant in our context they depend on the quality of expert judgments. Next, we determine these optimal weights.

Specifically, we first setup the problem for the estimation of parameters $\boldsymbol{\mu}$ and from these estimates we will deduce the weights for estimating $g(\mu_1, \mu_2)$. While the variance minimization problem is standard in the literature, we next provide the complete specification of the optimization formulation for the sake of completeness as well as to deduce properties of the optimal weights. We first rewrite the vector of judgments $\hat{\mathbf{y}} = [\hat{y}_1, \hat{y}_2, .., \hat{y}_m]^{\mathrm{T}}$ as $\hat{\mathbf{y}} = \mathbf{Z}\boldsymbol{\mu} + \boldsymbol{\epsilon}$ where $\mathbf{Z}$ is m$\times$ 2 matrix with columns $[1, 1, ..., 1]$ and $[\bar{z}_1, \bar{z}_2, ..., \bar{z}_m]$ and $\boldsymbol{\mu} = [\mu_1, \mu_2]^{\mathrm{T}}$. The calibration exercise discussed in Section 3 has provided us with the values of $\bar{z}_j$ as well as the variance-covariance matrix of $\boldsymbol{\epsilon}$, Ω. We next specify the following optimization problem for minimizing the estimation variance:

$$\min_{\boldsymbol{\mu}} (\hat{\mathbf{y}} - \mathbf{Z}\boldsymbol{\mu})^{\mathrm{T}}\Omega^{-1}(\hat{\mathbf{y}} - \mathbf{Z}\boldsymbol{\mu}) \tag{9}$$

Proposition 4 (Optimal Weights for Experts' Judgments) *The weights that provide unbiased minimum variance estimate of* $g(\mu_1, \mu_2)$ *as* $g(\hat{\mu}_1, \hat{\mu}_2) = \mathbf{w}^{\mathrm{T}}\hat{\mathbf{y}}$ *are given by* $\mathbf{w}^* = \mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\Omega^{-1}$ *with* $\mathbf{a}^{\mathrm{T}} = [\alpha_1 \;\; \alpha_2]$.

Using Proposition 4 we can identify three factors that determine the weights assigned to various experts. The first factor is the quantile each expert provides on average as well as the quantiles provided by other experts. This information is captured in matrix $\mathbf{Z}$. The second factor is the variability in an expert's judgments for his average quantile as quantified in the i^{th} diagonal element of the variance-covariance matrix as well as the relative magnitude of the variability in other experts' judgments. This information is captured in the variance-covariance matrix Ω. The third factor is the values of α_1 and α_2 in the function g which determine what quantity is to be estimated using the point judgments provided by the experts, where $\mathbf{a}^{\mathrm{T}} = [1 \;\; 0]$ for the estimation of the mean, $\mathbf{a}^{\mathrm{T}} = [0 \;\; 1]$ for the estimation of standard deviation, and $\mathbf{a}^{\mathrm{T}} = [1 \;\; z]$ for the estimation of quantile corresponding to the z-value. Finally, the property that the optimal weights provide an unbiased estimate of $g(\mu_1, \mu_2)$ leads to an important structural property of the weights. Specifically, the unbiased property of the estimate $g(\hat{\mu}_1, \hat{\mu}_2) = \mathbf{w}^{*\mathrm{T}}\hat{\mathbf{y}}$ implies that $\mathbf{E}[\mathbf{w}^{*\mathrm{T}}\hat{\mathbf{y}}] = \mathbf{a}^{\mathrm{T}}\boldsymbol{\mu}$. Now substituting $\mathbf{E}[\hat{y}_j] = \mu_1 + z_j\mu_2$ in vector notation we obtain $\mathbf{Z}^{\mathrm{T}}\mathbf{w}^* = \mathbf{a}$, which on further simplification shows that the optimal weights for individual judgments add up to the weight on the mean, i.e., $\sum_{i=1}^{m} w_i^* = \alpha_1$. We can further rewrite this equation as $\sum_{i=1}^{m} w_i'^* = 1$ where $w_i' = \frac{w_i}{\alpha_1}$. This equation implies that the weights assigned to experts' judgments can be normalized to adding to 1, and therefore each expert's judgment can be thought of contributing to a total weight of 1, a feature that is easy to explain to practitioners. This framework can also be used to deduce multiple signals. Specifically, suppose that a decision maker is interested in determining two quantities $g_1 = \alpha_{11}\mu_1 + \alpha_{21}\mu_2$ and $g_2 = \alpha_{12}\mu_1 + \alpha_{22}\mu_2$. One can show after some algebra that these two quantities have individual estimation variances of $[\alpha_{11}, \alpha_{21}](\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}[\alpha_{11}, \alpha_{21}]^{\mathrm{T}}$ and $[\alpha_{21}, \alpha_{22}](\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}[\alpha_{21}, \alpha_{22}]^{\mathrm{T}}$; and with covariance $[\alpha_{11}, \alpha_{12}](\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}[\alpha_{21}, \alpha_{22}]^{\mathrm{T}}$. The intuition for estimation of the covariance is that the underlying judgmental errors are the same, and hence the two quantities estimated based on the errors are correlated.

4.2. Impact of Team Composition on Performance

We now discuss how the characteristics of individual experts affect the group performance. In order to do this we consider the case when the expert judgments are aggregated to estimate a specific

quantile for probability p, i.e., $\alpha_1 = 1$ and α_2 corresponds to probability p. One can show after some calculations that the variance of the quantile estimate $g(\hat{\mu}_1, \hat{\mu}_2) = \hat{\mu}_1 + \alpha_2 \hat{\mu}_2$ is equal to

$$\text{Var}(g(\hat{\mu}_1, \hat{\mu}_2)) = [1\,0](\mathbf{Z}^{\text{T}}\Omega^{-1}\mathbf{Z})^{-1}[1\,0]^{\text{T}} + [0\,\alpha_2](\mathbf{Z}^{\text{T}}\Omega^{-1}\mathbf{Z})^{-1}[0\,\alpha_2]^{\text{T}} + 2[1\,0](\mathbf{Z}^{\text{T}}\Omega^{-1}\mathbf{Z})^{-1}[0\,\alpha_2]^{\text{T}} \quad (10)$$

where the first two terms are the variances in $\hat{\mu}_1$ and $\hat{\mu}_2$ respectively, and the third term is twice the covariance between $\hat{\mu}_1$ and $\hat{\mu}_2$. The first two terms in the R.H.S. are non-negative. The sign of the third is undetermined and depends on the value of α_2 which can be positive or negative values as well as on $\mathbf{Z}$ and Ω. It follows that the characteristics of individual experts (as captured by $\mathbf{Z}$ and Ω) and the quantiles to be elicited together influence whether a specific team would provide judgments for a specific quantile with higher or lower precision. Specifically, a team would either provide more precise judgments for either the left tail quantile or the right tail quantile (for a positive and negative value of the third term respectively). For a more nuanced understanding, we focus on the case when Ω is a diagonal matrix, i.e., when the expert judgments are mutually independent.

Proposition 5 *Consider Ω to be a diagonal matrix. When each $\hat{\bar{z}}_j < 0$ the aggregate estimate of a quantile in the right tail with $\alpha_2 > 0$ has a larger variance than a quantile in the left tail and vice-verse otherwise.*

This result shows that when all experts are biased in the same direction, i.e., they on average provide judgments in the same tail, then a quantile deduced in the same tail from their point judgments has a lower variance as compared to the case when their judgments are used to deduce a quantile in the other tail. For example, say three experts provide judgments, on average, for the 23rd, 45th, and 38th quantile. Then the estimate obtained for the 15th quantile will be more precise as compared to the estimate for the 85th quantile. By extension, if a large pool of experts is available and it is desirable to deduce multiple quantiles with a constant precision, one should select the experts such that the term $[1\,0](\mathbf{Z}^{\text{T}}\Omega^{-1}\mathbf{Z})^{-1}[0\,\alpha_2]^{\text{T}}$ is either equal to or close to zero. This analysis provides a statistical backing to the notion that a diversity in judgments is valuable in teams. Specifically, expert selection should be informed by the correlation between the judgments of various experts as well as the average quantiles estimated by these experts.

4.3. Selecting the Optimal Set of Experts

We next consider the problem of selecting a set of $K < m$ experts from the pool of m experts. This focus is relevant for situations in which the budget restrictions or work load restrictions require that experts are rationed among multiple teams or that only a limited number of experts are

engaged. Let $\mathbf{V} = [v_1 \;\; v_2 \;\; \ldots \; v_m]$ denote a vector of binary decision variables denoting whether or not experts $j = 1, 2, ..., m$ are selected in the group for future elicitations. We would like to determine the optimal composition of the group. To measure optimality, we continue to use the objective function used in (9). Specifically, using optimal weights $\mathbf{w}^* = \mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\Omega^{-1}$ for m experts, the estimation variance is equal $\mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}\mathbf{a}$. Excluding a subset of experts will lead to an increase in this variance. Accordingly we seek to determine the subset of exactly K experts that should be included in the group while minimizing the estimation variance.

We use the big-M method of implied penalty to modify the variance-covariance matrix Ω for this purpose. Specifically, we define:

$$\Omega^B = \begin{bmatrix} v_1\omega_{11} + (1-v_1)M & \ldots & \omega_{1m} \\ \vdots & \vdots & \vdots \\ \omega_{m1} & \ldots & v_m\omega_{mm} + (1-v_m)M \end{bmatrix} \tag{11}$$

With this matrix, the problem of selecting K experts is specified as follows.

$$\min_{\mathbf{V}} [\alpha_1 \;\; \alpha_2](\mathbf{Z}^{\mathrm{T}}(\Omega^B)^{-1}\mathbf{Z})^{-1}[\alpha_1 \;\; \alpha_2]^{\mathrm{T}} \tag{12}$$

$$\text{s.t. } v_j \in \{0, 1\} \tag{13}$$

$$\sum_{j=1}^{m} v_j = K \tag{14}$$

This formulation works as follows. Consider the problem in which we seek to reduce the group size by one, i.e., $K = m - 1$ and in the candidate solution $v_1 = 0$, i.e., expert 1 is to be removed from the group. Substituting the value $v_1 = 0$ in the matrix (11) we see that $\omega_{11}^B = M$. A large value of M ensures that the variance in the estimates of this expert are now effectively very large, and moreover, the estimates of this expert are approximately independent from those of the other experts. As a result, the optimal weight for this expert would be equal to zero, replicating the situation when the expert is not present in the group.

In general the presence of binary variables implies that the problem cannot be solved in polynomial time. When the group of experts is small even a complete enumeration is usually practical. The selection of K experts among a set of m experts can be done in ${}^nC_K$ ways. As an example for selecting say 4 out 6 experts, this leads to a ${}^6C_4 = 15$ combinations and the estimation variance for each combination can be computed efficiently. When the number of combinations is large, we have found three heuristics to perform well. These heuristics exploit the results developed earlier in Proposition 2 for the equivalent sample size of each expert. In the first heuristic we rank order experts based on the individual equivalent sample size determined using Proposition 1 and select

the top K experts. This heuristic is easy to implement, nevertheless, it has the drawback that it ignores the quantiles each expert estimates as well as the correlations between the judgments of the experts. The second and third heuristics incorporate these correlations in the selection process. In the second heuristic we start with the group size of zero. In the first step, we include the expert with the largest equivalent sample size. Next we independently consider the remaining experts and determine the decrease in the estimation variance when each of the expert is added. We then select and add the expert whose addition to the group would lead to the largest decrease in variance, and repeat this process until the group has K experts. In the third heuristic, we first consider the complete set of experts. Next we consider individual experts and determine the increase in estimation variance if the expert were removed from the set. We then select the expert whose removal from the group would lead to the lowest increase in the estimation variance, and continue this process until the group has K experts. For the sake of brevity we will not discus numerical results for the performance of these heuristics.

5. Extensions

We next extend our technical development to the case when each expert provides multiple judgments in Section 5.1, and in Section 5.2 we discuss the case when limited historical data are available to calibrate an expert.

5.1. Combining Distribution Assessments

A key feature of our development is that experts' judgments and the errors in these judgments are quantified in the scale free domain. We next discuss how this quantification provides the ability to combine the distributions implied by multiple judgments provided by more than one expert (we thank an anonymous reviewer for suggesting this development). Specifically, consider that each expert j responds to *r=1,2,...R* identical questions for an uncertain quantity or random variable Y and provides point judgments $\hat{\mathbf{y}}_j = [\hat{y}_{j1}, \ldots, \hat{y}_{jR}]^{\mathrm{T}}$. As an example, the expert could provide responses for the pessimistic, most likely, and optimistic values for the uncertainty. We assume for notational ease that each expert answers the same R questions but this assumption can be relaxed in a straight forward manner. The experts also provide these judgments for calibration distributions *i=1,2,...,n* as values $\hat{x}_{ijr}$. These calibration data are first transformed into the corresponding standardized z-values, as $\hat{z}_{ijr} = F_i^{-1}\big(F_i(\hat{x}_{ijr}; \boldsymbol{\mu}_i); 0, 1\big)$. The average quantile an expert j provides in response to each of the *r=1,2,...,R* questions is determined in the standardized scale as $\hat{\bar{z}}_{jr} = \dfrac{\sum_i \hat{z}_{ijr}}{n}$.

We next seek to quantify the variability in the judgments provided by expert j for the quantile $\hat{\bar{p}}_j$. In the absence of complete information for calibration distributions, we adopt a bootstrap approach. Specifically, let $\hat{\mathbf{x}}_{kj} = [\hat{x}_{k1j}, \hat{x}_{k2j}, \ldots, \hat{x}_{knj}]$ denote bootstrap samples $k = 1, 2, ..., K$ of size n. For each sample we determine the average quantile implied by the expert's judgments, as $\hat{p}_{kj}^{\hat{}} = \dfrac{\sum_{l}^{N} \mathbb{1}(\hat{x}_{klj} > x_{lj})}{n}$, $j = 1, 2..., m$. We then translate this information into z-domain, as $\hat{z}_{kj}^{\hat{}} = \dfrac{F^{-1}(\hat{p_{kj}^{\hat{}}}; 0, 1)}{n}$, and we quantify the uncertainty in the quantile judgments, as $Var(\hat{z}_j) = \dfrac{\sum_{k=1}^{K}(\hat{z}_{kj} - \bar{z}_{kj})^2}{K-1}$, where $\bar{z}_{kj} = \dfrac{\sum_{k=1}^{K} \hat{z}_{kj}}{K}$, and covariance between two experts j_1, j_2 as $Cov(\hat{z}_{j_1}, \hat{z}_{j_2}) = \dfrac{\sum_{k=1}^{K}(\hat{z}_{kj_1} - \bar{z}_{kj_1})(\hat{z}_{kj_2} - \bar{z}_{kj_2})}{K-1}$. From these covariances, we estimate the matrix Ω and then use the optimal weights developed in Proposition 4.

5.2. Absence of calibration distributions and non location-scale distributions

We now extend the results to two settings. In the first setting complete calibration distributions are not available and instead we only have historical data for the judgmental forecasts provided by individual experts and the realized values. We first operationalize these data to deduce the quantile provided by each expert using a frequency count. Specifically, for each expert j, we first determine the frequency with which her judgments $\hat{x}_{ij}$ exceed the observed values x_{ij} as $\sum_i \mathbb{1}(\hat{x}_{ij} > x_{ij})$. From this information we deduce the quantile she provides on average, as $\hat{\bar{p}}_j^{\hat{}} = \dfrac{\sum_i \mathbb{1}(\hat{x}_{ij} > x_{ij})}{n}$. For example when $\hat{\bar{p}}_j^{\hat{}} = 0.4$, it means that the expert provides us with the 40th quantile judgments, on average. From this quantity, we deduce the quantile in the standardized z-value as $\hat{\bar{z}}_j^{\hat{}} = F^{-1}(\hat{\bar{p}}_j^{\hat{}}; 0, 1)$.

We next seek to quantify the variability in the judgments provided by expert j for the quantile $\hat{\bar{p}}_j^{\hat{}}$. In the absence of complete information for calibration distributions, we adopt a bootstrap approach. Specifically, let $\hat{\mathbf{x}}_{kj} = [\hat{x}_{k1j}, \hat{x}_{k2j}, \ldots, \hat{x}_{knj}]$ denote bootstrap samples $k = 1, 2, ..., K$ of size n. For each sample we determine the average quantile implied by the expert's judgments, as $\hat{p}_{kj}^{\hat{}} = \dfrac{\sum_{l}^{N} \mathbb{1}(\hat{x}_{klj} > x_{lj})}{n}$, $j = 1, 2..., m$. We then translate this information into z-domain, as $\hat{z}_{kj}^{\hat{}} = \dfrac{F^{-1}(\hat{p_{kj}^{\hat{}}}; 0, 1)}{n}$, and we quantify the uncertainty in the quantile judgments, as $Var(\hat{z}_j) = \dfrac{\sum_{k=1}^{K}(\hat{z}_{kj} - \bar{z}_{kj})^2}{K-1}$, where $\bar{z}_{kj} = \dfrac{\sum_{k=1}^{K} \hat{z}_{kj}}{K}$, and covariance between two experts j_1, j_2 as $Cov(\hat{z}_{j_1}, \hat{z}_{j_2}) =$

$\frac{\sum_{k=1}^{K}(\hat{z}_{kj_1}-\bar{z}_{kj_1})(\hat{z}_{kj_2}-\bar{z}_{kj_2})}{K-1}$. From these covariances, we estimate the matrix Ω and then use the optimal weights developed in Proposition 4. The second extension is for the case when the underlying uncertainty does not have a location-scale distribution. For the sake of brevity we provide the details in Appendix B.

6. Summary and Future Research

This paper addresses a frequently faced problem in R&D management: multiple experts provide point judgments for the most likely potential of new candidate projects, and one needs to aggregate these point judgments to deduce the average potential as well as the uncertainty in the project-potential. These two quantities are required to make risk-return trade-offs during portfolio selection and resource allocation decisions. In distributed R&D settings experts often differ in their experience and intuition which translates into differences in individual judgments for project-potential. It is important to incorporate these individual-specific differences into mechanisms to deduce the average potential of a project, its variability, or a combination of these two quantities. Accordingly, we developed a two-step algorithmic approach. In Step 1 we characterize the judgments provided by individual experts as the average quantile that their judgments correspond to and the variability in this quantile. Then we use an optimization framework to deduce the mean, standard deviation or a combination of these two quantities for a probability distribution. These estimates are then used to make portfolio selection decisions.

This research can be extended to answers several questions in the context of R&D management. First, in multi-year projects experts typically learn from prior experiences and their judgments are expected to improve over time. Simultaneously it is likely that the new products being considered are more complex and the heuristics developed by the experts on previous products may not perform well for the future products considered in the R&D program. The net impact of these two effects on the quality of judgments is not clear. Future empirical research should focus on this net outcome by calibrating judgments of experts over a multi-period horizon in a learning environment.

Second, in several instances a prior distribution on the performance of a new product may be available and this prior may need to be updated using the data generated during expert-calibration. Future research should develop algorithms to support this Bayesian updating. Specifically, the estimates obtained using the approach developed in this paper may not have a conjugate prior. Furthermore in several instances it may be of interest to obtain estimates for multiple quantities simultaneously, for example the mean and the standard deviation. These estimates, in general, are

correlated. Furthermore based on calibration data they may also have distributions with different parametric families. Traditional conjugate prior-posterior based framework for Bayesian updating will not be adequate and one would need to resort to Monte Carlo simulation methods. Future research should focus on fleshing out details of this simultaneous updating of multiple estimates.

References

Agnew JR (2006) Do behavioral biases vary across individuals? evidence from individual level 401 (k) data. *Journal of Financial and Quantitative Analysis* 41(4):939–962.

Bansal S, Gutierrez GJ, Keiser JR (2017) Using experts noisy quantile judgments to quantify risks: Theory and application to agribusiness. *Operations Research* 65(5):1115–1130.

Barron OE, Kim O, Lim SC, Stevens DE (1998) Using analysts' forecasts to measure properties of analysts' information environment. *Accounting Review* 421–433.

Barron OE, Stuerke PS (1998) Dispersion in analysts' earnings forecasts as a measure of uncertainty. *Journal of Accounting, Auditing & Finance* 13(3):245–270.

Bates JM, Granger CW (1969) The combination of forecasts. *Journal of the Operational Research Society* 20(4):451–468.

Batur D, Choobineh F (2010) A quantile-based approach to system selection. *European Journal of Operational Research* 202(3):764–772.

Bazerman MH, Moore DA (2008) Judgment in managerial decision making .

Bjørndal T, Herrero I, Newman A, Romero C, Weintraub A (2012) Operations research in the natural resource industry. *International Transactions in Operational Research* 19(1-2):39–62.

Budescu DV, Chen E (2014) Identifying expertise to extract the wisdom of crowds. *Management Science* 61(2):267–280.

Budescu DV, Erev I, Wallsten TS (1997) On the importance of random error in the study of probability judgment. part i: New theoretical developments. *Journal of Behavioral Decision Making* 10(3):157–171.

Chopra VK, Ziemba WT, et al. (2011) The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Kelly Capital Growth Investment Criterion: Theory and Practice* 3:249.

Cooke RM (1991) Experts in uncertainty: opinion and subjective probability in science .

Davis-Stober CP, Budescu DV, Broomell SB, Dana J (2015) The composition of optimally wise crowds. *Decision Analysis* 12(3):130–143.

Gaur V, Kesavan S, Raman A, Fisher ML (2007) Estimating demand uncertainty using judgmental forecasts. *Manufacturing & Service Operations Management* 9(4):480–491.

Geweke J, Amisano G (2011) Optimal prediction pools. *Journal of Econometrics* 164(1):130–141.

Greene WH (2003) *Econometric analysis* (Pearson Education India).

Hess SW (1993) Swinging on the branch of a tree: project selection applications. *Interfaces* 23(6):5–12.

Jaafari A (2001) Management of risks, uncertainties and opportunities on projects: time for a fundamental shift. *International journal of project management* 19(2):89–101.

Johnson D (1998) The robustness of mean and variance approximations in risk analysis. *Journal of the Operational Research Society* 49(3):253–262.

Johnson NL (1949) Systems of frequency curves generated by methods of translation. *Biometrika* 36(1/2):pp. 149–176, ISSN 00063444.

Johnson TC (2004) Forecast dispersion and the cross section of expected returns. *The Journal of Finance* 59(5):1957–1978.

Keefer D, Bodily S (1983) Three-point approximations for continuous random variables. *Management Science* 29(5):595–609.

Keefer DL, Verdini WA (1993) Better estimation of pert activity time parameters. *Management Science* 39(9):1086–1091.

Kettunen J, Salo A (2017) Estimation of downside risks in project portfolio selection. *Production and Operations Management* .

Lloyd E (1952) Least-squares estimation of location and scale parameters using order statistics. *Biometrika* 39(1/2):88–95.

Loch CH, Kavadias S (2002) Dynamic portfolio selection of npd programs using marginal returns. *Management Science* 48(10):1227–1241.

Ravinder H, Kleinmuntz D, Dyer J (1988) The reliability of subjective probabilities obtained through decomposition. *Management Science* 34(2):186–199.

Smith JE, Winkler RL (2006) The optimizers curse: Skepticism and postdecision surprise in decision analysis. *Management Science* 52(3):311–322.

Sterman JD (1989) Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management science* 35(3):321–339.

Stuart A, Ord JK (1994) Kendall's advanced theory of statistics. vol. i. distribution theory. *Arnold, London* .

Wallsten TS, Budescu DV (1983) State of the art ? encoding subjective probabilities: A psychological and psychometric review. *Management Science* 29(2):151–173.

Weintraub A, Romero C (2006) Operations research models and the management of agricultural and forestry resources: a review and comparison. *Interfaces* 36(5):446–457.

Winkler RL (1981) Combining probability distributions from dependent information sources. *Management Science* 27(4):479–488.

7. Appendix A: Proofs

Proof of Proposition 1: The variance of sample quantile for probability $\hat{\hat{p}}_j$ is equal to

$$\frac{\hat{\hat{p}}_j(1-\hat{\hat{p}}_j)}{n(f_s(\hat{\hat{z}}_j))^2}.$$

The variance in the judgments of expert j in the z-domain is equal to w_{jj}. Next we equate $w_{jj} = \frac{\hat{\hat{p}}_j(1-\hat{\hat{p}}_j)}{n(f_s(\hat{\hat{z}}_j))^2}$. Interchanging the places of n and w_{jj} we obtain $n = \frac{\hat{\hat{p}}_j(1-\hat{\hat{p}}_j)}{w_{jj}(f_s(\hat{\hat{z}}_j))^2}$. Replacing n with N_j we obtain the expression in the proposition. □

Proof of Proposition 2: We show that the function ω_{jj} is convex with a minimum at $f_s'(\hat{\hat{z}}_j) = 0$. To this end we first rewrite the function as $\omega_{jj} = \dfrac{\hat{\hat{p}}_j(1-\hat{\hat{p}}_j)}{N_j\left(f_s(\hat{\hat{z}}_j)\right)^2} = \dfrac{f_s(\hat{\hat{z}}_j)(1-f_s(\hat{\hat{z}}_j))}{N_j\left(f_s(\hat{\hat{z}}_j)\right)^2}$. Next we differentiate the function to obtain its gradient as $\dfrac{\partial \omega_{jj}}{\partial \hat{\hat{z}}_j} = \dfrac{f_s'(\hat{\hat{z}}_j) - 2f_s'(\hat{\hat{z}}_j)f_s(\hat{\hat{z}}_j)}{N_j\left(f_s(\hat{\hat{z}}_j)\right)^2} - \dfrac{2f_s'(\hat{\hat{z}}_j)f_s(\hat{\hat{z}}_j)(1-f_s(\hat{\hat{z}}_j))}{N_j\left(f_s(\hat{\hat{z}}_j)\right)^3}$ which on simplification reduces to $\dfrac{\partial \omega_{jj}}{\partial \hat{\hat{z}}_j} = -\dfrac{f_s'(\hat{\hat{z}}_j)}{N_j\left(f_s(\hat{\hat{z}}_j)\right)^2}$. The only feasible solution to this equation is $f_s'(\hat{\hat{z}}_j) = 0$, i.e., at the mode. Next, we show that at this value the second derivative is positive. To this end, we first note that $\dfrac{\partial^2 \omega_{jj}}{\partial \hat{\hat{z}}_j^2} = -\dfrac{f_s''(\hat{\hat{z}}_j)}{N_j\left(f_s(\hat{\hat{z}}_j)\right)^2} + 2\dfrac{f_s'(\hat{\hat{z}}_j)}{N_j\left(f_s(\hat{\hat{z}}_j)\right)^3}$. Substituting $f_s'(\hat{\hat{z}}_j) = 0$, we obtain $\dfrac{\partial^2 \omega_{jj}}{\partial \hat{\hat{z}}_j^2} = -\dfrac{f_s''(\hat{\hat{z}}_j)}{N_j\left(f_s(\hat{\hat{z}}_j)\right)^2}$. Next, we observe that $f_s'(\hat{\hat{z}}_j) = 0$ at the mode of the distribution. Furthermore, at the mode, the slope changes from positive to negative. It follows that $\dfrac{\partial^2 \omega_{jj}}{\partial \hat{\hat{z}}_j^2} > 0$ at the mode.

Proof of Proposition 3: We start by noting that the judgments provided by expert j can be stated as $x_j = \mu_1 + z_j\mu_2 + \epsilon$. We next adopt the approach used by Lloyd (1952) for the rest of the proof. These judgments have expectations $E[x_j] = \mu_1 + z_j\mu_2$ that are linear functions of the two parameters. Furthermore, the variances and the covariances in the z-transformed judgments are available to us from the calibration exercise. The transformation means that the variance-covariance matrix we calculated is a scaled version of the variance-covariance matrix of expert judgments. As a result the general form of the least-squares theorem is applicable, and a linear function of the judgments provide unbiased estimates of μ_1, μ_2 with minimum variance. □

Proof of Proposition 4: The proof is in two steps. In the first step we obtain the estimates of μ_1 and μ_2 by using $\mathbf{a}^{\mathrm{T}} = [1\,0]$ and $\mathbf{a}^{\mathrm{T}} = [0\,1]$ respectively. The derivation of the solution to the problem

$$\min_{\boldsymbol{\mu}} (\hat{\mathbf{y}} - \mathbf{Z}\boldsymbol{\mu})^{\mathrm{T}}\Omega^{-1}(\hat{\mathbf{y}} - \mathbf{Z}\boldsymbol{\mu})$$

is available in a number of texts including Greene (2003); this solution is obtained as $\mu_j = \mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\hat{\mathbf{y}}$ with an appropriate value of the vector $\mathbf{a}$. It follows that the estimate μ_j is a weighted average of judgments $\hat{\mathbf{y}}$ where the weights are equal to $\mathbf{w}^* = \mathbf{a}^{\mathrm{T}}(\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}\mathbf{Z}^{\mathrm{T}}\Omega^{-1}$.

In Step 2 we note that we sought out to estimate the function

$$g(\mu_1, \mu_2) = \alpha_1\mu_1 + \alpha_2\mu_2.$$

From the properties of linear estimation it follows that we can obtain the estimate $\hat{g}(\mu_1, \mu_2)$ as $\hat{g}(\mu_1, \mu_2) = \alpha_1\hat{\mu}_1 + \alpha_2\hat{\mu}_2$. We can further rewrite the R.H.S. as $[\alpha_1\ 0][\mu_1\ , \mu_2]^{\mathrm{T}} + [0, \alpha_2][\mu_1\ , \mu_2]^{\mathrm{T}} = [\alpha_1,\ \alpha_2][\mu_1\ , \mu_2]^{\mathrm{T}}$. It follows that the optimal weights for the judgments are obtained after substituting $[\alpha_1,\ \alpha_2]$ for $\mathbf{a}^{\mathrm{T}}$. □

Proof of Proposition 5: We start with the observation made in Section 4.2.1 that the variance of the quantile estimate $g(\hat{\mu}_1, \hat{\mu}_2) = \hat{\mu}_1 + \alpha_2\hat{\mu}_2$ is equal to

$$\mathrm{Var}(g(\hat{\mu}_1, \hat{\mu}_2)) = [1\ 0](\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}[1\ 0]^{\mathrm{T}} + [0\ \alpha_2](\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}[0\ \alpha_2]^{\mathrm{T}} + 2[1\ 0](\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}[0\ \alpha_2]^{\mathrm{T}} \quad (15)$$

The first two terms are always positive and we focus on the third term. For m experts with a diagonal matrix Ω, this last term is obtained as

$$[1\ 0](\mathbf{Z}^{\mathrm{T}}\Omega^{-1}\mathbf{Z})^{-1}[0\ \alpha_2]^{\mathrm{T}} = -\alpha_2 \frac{\sum_{j=1}^{m} \hat{\bar{z}}_j \sigma_1^2\sigma_2^2...\sigma_{j-1}^2\sigma_{j+1}^2....\sigma_m^2}{(\hat{\bar{z}}_1 - \hat{\bar{z}}_2)^2\sigma_3^2\sigma_4^2...\sigma_m^2 + (\hat{\bar{z}}_1 - \hat{\bar{z}}_3)^2\sigma_2^2\sigma_4^2...\sigma_m^2 + ...\binom{m}{2}\text{terms}} \quad (16)$$

The denominator is always positive, and we focus on the numerator. When $\alpha_2 < 0$, i.e., when a left tail quantile is to be deduced and for each expert $\hat{\bar{z}}_j < 0$ the numerator and hence the complete term is negative. It follows that the variance in (27) is lower. In contrast, the variance will be higher when a right tail $\alpha_2 > 0$ is to be deduced. A similar analysis is straightforward for the case when for each expert $\hat{\bar{z}}_j > 0$. □

8. Appendix B: Extension to non location-scale distributions

The technical development discussed so far for location–scale distributions also enables us to estimate the parameters of Johnson distributions used to model probability distributions with a greater degree of flexibility. The key connection between location–scale distribution and Johnson distributions is that random variables X with Johnson distributions with parameters $\boldsymbol{\theta}$ result in a normal variable Y with parameters $\boldsymbol{\theta}$ after a non-linear monotonic transformation, g – i.e., $g(X) = Y$ (Johnson 1949). For example, if X is a lognormal random variable (type 2 Johnson variable), then $Y = ln(X)$ is a normal random variable.

In the absence of elicitation errors, the estimation of the parameters of X is straightforward. For any $x \in \mathbb{R}$ we have $\Pr\{X \le x\} = \Pr\{g(X) \le g(x)\}$ where $g(x)$ is non-decreasing in x; hence

the p_i-quantiles of X and $g(X)$, denoted respectively as $x(p_i)$ and $x^g(p_i)$, satisfy the relationship $g(x(p_i)) = x^g(p_i)$. Therefore, one could simply take the inverse g^{-1} of quantile estimates $\hat{x}_i$ to transform them on the underlying normal distribution, and then estimate the parameters of this underlying normal distribution using the results developed earlier. However, this process does not carry over to when expert's estimates have errors. As an example for the log-normal distribution, any elicited quantile satisfies $ln(\hat{x}_i) = ln(x_i + \epsilon_i)$, but clearly $ln(\hat{x}_i) \neq ln(x_i) + ln(\epsilon_i)$. However, we can approximate $g(X)$ using the second–order Taylor series expansion of $g(x_i)$ about $\hat{x}_i = x_i + \epsilon_i$ as:

$$g(x_i) \approx g(\hat{x}_i) + g'(\hat{x}_i)(x_i - \hat{x}_i) + \frac{g''(\hat{x}_i)}{2}(x_i - \hat{x}_i)^2. \tag{17}$$

Since $g(x_i)$ follows a normal distribution with parameters $\boldsymbol{\theta}$, it can be expressed as $g(x_i) = \theta_1 + \theta_2 z_i$. Combining this with (17), we obtain an approximate model for the estimation of parameters $\boldsymbol{\theta}$ for $g(X)$ as

$$g(\hat{x}_i) \approx \theta_1 + \theta_2 z_i + e_i^g, \tag{18}$$

where the transformed elicitation errors $e_i^g = \hat{x}_i^g - x_i^g$ are obtained from (17) as $e_i^g = g(\hat{x}_i) - g(x_i) = g(\hat{x}_i) - (\theta_1 + \theta_2 z_i) \approx g'(\hat{x}_i)\epsilon_i + \frac{g''(\hat{x}_i)}{2}\epsilon_i^2$. The second term is a bias that comes from the non-linear transformation of the estimated quantile of X to $g(X)$. For any given elicited quantile $\hat{x}_i$, this bias is denoted as $B(\hat{x}_i) \equiv \mathrm{E}\big[\, e_i^g \,\big] = \frac{1}{2} g''(\hat{x}_i)\sigma_{\epsilon_i}^2$. Substituting this bias in (18), we obtain the following model to estimate the parameters θ_1 and θ_2 of the normal pdf of $g(X)$:

$$g(\hat{x}_i) + B(\hat{x}_i) = \theta_1 + \theta_2 z_i + e_i^g. \tag{19}$$

The variance-covariance matrix $\widehat{\Omega}$ for the errors in (19) is determined as follows:

$$\begin{aligned}
\widehat{\Omega}_{ii} &= \mathrm{E}\big[\, \big(e_i^g - \mathrm{E}\big[\, e_i^g \,\big]\big)^2 \,\big] = \mathrm{E}\big[\, \big(g'(\hat{x}_i)\epsilon_i + \frac{1}{2}g''(\hat{x}_i)\epsilon_i^2 - \frac{1}{2}g''(\hat{x}_i)\sigma^2\big)^2 \,\big] \\
\widehat{\Omega}_{ij} &= \mathrm{E}\big[\, \big(e_i^g - \mathrm{E}\big[\, e_i^g \,\big]\big)\big(e_j^g - \mathrm{E}\big[\, e_j^g \,\big]\big) \,\big] \\
&= \mathrm{E}\big[\, \big(g'(\hat{x}_i)\epsilon_i + \frac{1}{2}g''(\hat{x}_i)\epsilon_i^2 - \frac{1}{2}g''(\hat{x}_i)\sigma^2\big)\big(g'(\hat{x}_j)\epsilon_j + \frac{1}{2}g''(\hat{x}_j)\epsilon_j^2 - \frac{1}{2}g''(\hat{x}_j)\sigma^2\big) \,\big]
\end{aligned}$$

Since $g(x)$ transformations for Johnson distributions are tractable, each term in the expressions for $\widehat{\Omega}_{ii}$ and $\widehat{\Omega}_{ij}$ admits algebraic simplification. Once the matrix $\widehat{\Omega}$ is determined, the estimates $\hat{\boldsymbol{\theta}}$ are obtained by using Proposition 4, replacing Ω with $\widehat{\Omega}$.