2013

Effects of Different Methods of Aggregation of Probabilities on the R&d Investment Portfolio for Optimal Emissions Abatement: An Empirical Evaluation

Olaitan P. Olaleye
University of Massachusetts Amherst

EFFECTS OF DIFFERENT METHODS OF AGGREGATION OF PROBABILITIES ON THE R&D INVESTMENT PORTFOLIO FOR OPTIMAL EMISSIONS ABATEMENT: AN EMPIRICAL EVALUATION

A Thesis Presented

by

OLAITAN OLALEYE

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH

February 2013
INDUSTRIAL ENGINEERING AND OPERATIONS RESEARCH

EFFECTS OF DIFFERENT METHODS OF AGGREGATION OF PROBABILITIES ON THE R&D INVESTMENT PORTFOLIO FOR OPTIMAL EMISSIONS ABATEMENT: AN EMPIRICAL EVALUATION

A Thesis Presented

By

OLAITAN OLALEYE

Approved as to style and content by:

_____

Erin Baker, Chairperson

_____

Jenna Marquard, Member

_____

Senay Solak, Member

_____

Donald Fisher, Head

Mechanical and Industrial Engineering Department

ACKNOWLEDGEMENTS

I would like to sincerely express my sincere gratitude to everyone who extended support towards the delivery of this MS thesis.

My utmost dedication for this work is to my late mother for her care and support. In the same stead, I will like to recognize the invaluable support and patience of my wife and my family who always offered motivation and encouragements during the times this task seemed unachievable.

I will also like to emphasize the role of my advisor, Dr. Erin Baker, who really has been the singular most important contributor to this work and my education here at the university. She has been not only an excellent academic advisor but also a mentor and guide in all aspects. Special thanks also to my friends and research laboratory mates for the time and support towards making my Masters a wonderful learning experience.

ABSTRACT

EFFECTS OF DIFFERENT METHODS OF AGGREGATION OF PROBABILITIES ON THE R&D INVESTMENT PORTFOLIO FOR OPTIMAL EMISSIONS ABATEMENT: AN EMPIRICAL EVALUATION

FEBRUARY 2013

OLAITAN OLALEYE, B.SC., UNIVERSITY OF LAGOS

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Erin Baker

This thesis examines two possible orders of combining multiple experts in elicitations with multiple de-composed events: Should experts be combined early or later in the decision process? This thesis is in conjunction with the paper (Baker & Olaleye, 2012) where we show that it is best to combine experts early as later combination leads to a systematic error. We conduct a simulation to more fully flesh out the theoretical model. We also conduct a theoretical analysis aimed at determining how significantly these two methods differ. We find that all results are in accordance with the theory but combining experts later might lead to less error in some cases due to randomness.

We then conduct an empirical evaluation of the two methods using data from a previous study. We show that the experts exhibit some form of correlation. The impact of using the two methods of combining experts is then evaluated using an optimal R&D investment portfolio model. We find that the elicitation inputs have a significant effect on the outcome of the optimal portfolio and that there is an advantage from combining experts early.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

CHAPTER 1

INTRODUCTION

1.1 Thesis Statement

In an earlier paper, (Baker & Olaleye, 2012) show theoretically that experts' opinions are better suited to earlier rather than later aggregation due to dependency and self-correlation. This thesis examines the practical impact of the different methods of combining expert opinions, using work on the optimal energy R&D portfolio as an example.

1.2 Background

When assessing the probability of occurrence of a rather complex event, it is generally better to divide such an event up into bits and then to assess the chance of occurrence of the individual bits. A specific example of this is a pharmaceutical company's choice of investment in the development of a new drug for a medical condition that is expected to be prevalent in a decade. It is best to first evaluate the odds of the medical condition being present in a decade, the possibility of the drug overcoming challenges at the various research and development levels, the possibility of developing the drug before a competitor and much more. These events should be independent or at least conditionally independent. In other words, it is preferable in most cases to decompose the event into independent events and combine them later to force the experts to think critically and improve decision making.

The next step in the process of assessing probabilities of complex events is to solicit opinions on the likelihood of these decomposed events i.e. experts are 'elicited'. Innate cognitive biases of human experts can sometimes lead to biases and errors in the estimation of these events.

To increase the information value of these elicitations, the general consensus is to use multiple experts [e.g. (Clemen & Winkler, 2004)]. Hence not only are there multiple decomposed events but also multiple experts with multiple innate biases. The interaction of the decomposition nature of most evaluations and the subjective nature of human experts usually results in different outcomes of the evaluations, depending on how the experts are combined. Of the two general classes of possible combination of multiple experts: mathematical and behavioral, we consider only mathematical combination as behavioral combination requires some level of subjectivity on the part of the decision analyst. Studies have also shown that of the various mathematical combination methods, simple mathematical un-weighted averaging is usually best when trading off accuracy and complexity [(Clements & Harvey, 2009), (Clemen & Winkler, 1999), (Cooke R. M., 1991)], as such we consider only simple mathematical un-weighted averaging.

For simple mathematical averaging, the order in which the experts and the de-composed events are re-composed matters, as rarely do we obtain the same outcome when different orders of combination are used. We examine two of those orders, combining the experts for each of the de-composed events before re-composition, Method I (early aggregation), or re-composing the de-composed events for each expert before combining experts, Method II (later aggregation). We rely on theory from (Baker & Olaleye, 2012) to show that Method I should always be used as it results in less error, on average.

1.3 Objectives

In this section, we discuss the objectives of this study and our method of approach. Our motivation comes from a recent paper (Baker & Olaleye, 2012), which shows that it is best to combine experts early rather than later, especially when individual experts seem to show some degree of optimism or pessimism across all events.

We compare both orders of aggregating experts: using both a simulation and data from an empirical study. We initially develop a Monte Carlo simulation of experts with particular characteristics and examine the (outcome of elicitations) when either early or later combination of the experts is used. We also theoretically evaluate the likelihood that early combination is better than later.

We then empirically examine expert elicitations from previous studies by Baker et al ((2009)(2009) (2008)(2011)). We first examine if experts show characteristics such as optimism and pessimism, and then make predictions about the combined probabilities based on our knowledge of the theory. We examine if our predictions are right and if the insights from the theory in (Baker & Olaleye, 2012) support the empirical data.

Finally, we examine what effect the use of the two aggregation methods will have on a real decision problem. We use the resulting probabilities from the two methods of combination as inputs to an optimal portfolio model and assess the impact of the changes. The loss from aggregating later is also determined. A detailed discussion of our approach is given below.

1.3.1 Monte Carlo Simulation

To empirically examine the theoretical postulations, we develop a simple model of multiple experts eliciting the probability of success of some technological constraints needed for the development

of a technology. In this model, experts are combined using either earlier or later aggregation across the technological constraints. A Monte Carlo simulation of this model is then used to examine the characteristics of both methods of combination of experts.

1.3.2 Theoretical Evaluation Of The Dominance Of Early Aggregation

We next carry out a theoretical estimation of the dominance of one order of aggregation over the other. As the exact distribution of the errors from both aggregation orders is unknown, we assume a normal distribution for the errors and then estimate the probability that the error from early combination is less than that from later.

1.3.3 Expert Classification

We use data from an empirical study to examine the theoretical postulations. We classify the experts from the study in terms of their inherent characteristics i.e. optimism or pessimism. Experts are classified in two ways. One is to determine if the experts individually exhibit a tendency to be generally optimistic or pessimistic in their estimates. We term this as being “self-correlated”. Two, we examine if the experts are correlated to each other; we term this as “cross-correlation”.

1.3.4 Projection Based On Expert Characteristics

Based on our classification, we infer, based on the theory, what patterns the resulting probabilities show when experts are combined early or later. In cases where majority of the individual experts show a high degree of optimism or pessimism, we expect that combining experts later will lead to an increased end point relative to combining the experts early.

1.3.5 Comparing Projections And Actual Data

To examine if the effects are as theoretically predicted we recombined the experts using the two orders of combination i.e. earlier and later. Using the results from combining experts later as a base reference (used in a previous study (Peng, 2010)), we compare these with the probabilities from

early combination and identify if the pattern shown is as predicted by theory. In essence we examine if the end probabilities of success will be less when the experts are combined early when they show a high level of self-correlation.

1.3.6 Optimal Portfolio Model

One of the goals for this study was to examine the impact of using the different orders of aggregation of expert opinions on the R&D investment portfolio. Thus we extend previous work by (Baker & Solak, 2011) where experts' opinions, which were combined later, were used as input to an optimal energy portfolio model. We do this by replacing these inputs with resulting probabilities from early combination of the experts. We then examine the change in the optimal portfolio of future energy technologies and the total societal cost when the different methods of combination of experts' opinions are used.

1.3.7 Policy Implications From Use Of The Various Methods Of Combination

In conclusion of this thesis, we examine the implications resulting from the use of the various orders of combination. Hence, we discuss the policy implications resulting from the use of the two orders of combining experts' opinion based on the changes in the resulting optimal portfolio. We also examine the societal cost with respect to the theoretical correct method (early combination). We emphasize the significance of carrying out sensitivity analysis on the end probabilities. We therefore rely on the previous work by Baker et al. [(2008), (2009), (2009), (Baker & Peng, 2012), (Baker & Solak, 2011)] as the basis for this work.

The remainder of the paper is structured as follows. Chapter 2 presents a literature review, including the need for expert elicitations, the impact of bias between and within experts, and the resulting effect of correlation between experts. Chapter 3 gives an overview of the preceding

research that this work builds on, including a brief discussion of the various possible combination methods, the theoretical foundations from (Baker & Olaleye, 2012), how correlation results in the superiority of one order of aggregation and a brief discussion on the use of the elicited probabilities to parameterize the marginal abatement curve. Chapter 4 presents a detailed description of three variants of the Monte Carlo simulation of the theoretical model and also discusses the results from the simulation model. Chapter 5 details the theoretical estimation of the dominance of either order under basic assumptions. Chapter 6 outlines the results and implications from the use of the different aggregation orders. Chapter 7 presents the output from the optimal portfolio model and discusses the policy implications. Chapter 8 gives the summary and conclusion. The appendix and reference sections conclude this paper.

CHAPTER 2

LITERATURE REVIEW

In this section, we discuss some related literature to the different topics central to our work. Section 2.1 gives an overview of the use of experts to estimate the probability of events. The discussion entails why expert elicitations rather statistical analysis is needed or used, the resulting issue of de-composition and re-composition of expert probabilities, and the various biases that affect these elicitations. Section 2.2 discusses the different cognitive biases that exist in expert elicitations. Section 2.3 presents the role of correlation between and within experts, and the resulting impacts from such a relationship.

2.1 Expert Elicitations - Importance

2.1.1 Importance And Appropriateness Of Expert Elicitations

The development of future energy technologies represents one of the major sources of uncertainty in modeling the impact of climate change. We would like to obtain a best estimate of such uncertain events, development of novel impact technologies, before these estimates are used as base inputs to optimal portfolio models. Consequently the decision analyst faces an immediate choice of either using statistical analysis to derive the probabilities of such events or using assessments from human sources that possess some degree of expertise in the field. While statistical projection has the advantage of consistency and lack of subjectivity, it is particularly limited in fields with no historical trends and a high degree of uncertainty ((Makridakis, 1986)). Expert elicitations on the other hand seem more applicable to fields such as climate change and energy resource development where there exists a 'rich diversity of opinions' (Morgan & Keith,

1995). Thus, a practice is to use human experts as the complex nature of these events can be extremely hard to accurately model statistically. As such, some studies rely on human experts to provide best estimates for these events (Mosleh, Bier, & Apostolakis, 1987), (Clemen & Winkler, 1999)). However expert elicitations are subject to subjectivity and bias on the part of the experts leading to 'systematic and predictable errors' (Tversky & Kahneman, 1974)).

After the decision is made to estimate odds using expert elicitations, the immediate challenge faced is determining the number of experts to use for the elicitations. While it is generally beneficial to have more than one expert (multiple experts with varying opinions) for obvious advantages of a better forecast and more information value ((Clemen & Winkler, 2004), (Clemen & Winkler, 1999)), the value of additional experts increases at a diminishing rate ((Budescu & Rantilla, 2000)) and the optimal number of experts is considered to be anywhere between 3-6 experts (Clemen, 1986) depending on how similar the experts are.

2.1.2 Decomposition And Recombination

On deciding which estimation procedure (expert elicitation or statistical analysis) is to be used, a choice exists between estimating the probability of the single event i.e. direct assessment or decomposing such an event into the multiple independent sub-events. Arguments exist for the use of either of the methods, direct assessment or decomposition, depending on the degree of vagueness of the event whose likelihood is to be estimated. Some authors have advocated for direct assessment citing limitations such as the increased requirements for decomposition and possible error in the modeling of the full space of the event to be decomposed (Armstong, Denniston, & Gordon, 1975). Proponents for decomposition, (Raiffa, 1968), (Kleinmuntz, Ravinder, & and Dyer, 1986), (Cooke M. R., 1986) and (Clemen & Winkler, 1999) all show that when faced

with the problem of assessing the odds of occurrence of a rather complex event, it is best to decompose the event into sub-events which are easier to understand and assess. Given the degree of vagueness of the event (probability of development of novel non incremental technologies) considered in this study, it was considered best to decompose to enable better decision making.

While experts elicitations are generally considered to be very good estimates, a resulting problem ensues when the elicitation from the multiple experts are to be combined, as it has been shown that seldom do any two orders of aggregating these sub events result in the same main event probability (Mosleh(1988)). An example of such discrepancy is shown in the next section.

Alongside this problem is the issue of subjectivity faced by the experts being human themselves. Various articles have shown that humans' decision making heuristics (Representativeness, Availability, Anchoring and Inadequate Adjustment et al.) will also have serious consequence on the end results ((Kleinmuntz, Ravinder, & and Dyer, 1986), Mosleh (1988), (Baker & Olaleye, 2012)) depending on the order of combination of these experts. The minimization of these errors from de-composing and re-composing of multiple events by multiple experts has been the aim of much research. A review of the literature shows that it is usually agreed that a simple mathematical average of mutually exclusive independent sub-events is best given the tradeoff between computational complexity and accuracy (e.g. (Clemen & Winkler, 1999), (Cooke R. M., 1991), (Clements & Harvey, 2009)) unless the particular event elicited lends itself to a particular method. Hence we limit the scope of this study to simple mathematical equally weighted averaging.

2.2 Bias

In this section we discuss the literature on some of the biases that show up in expert elicitations. These include over-confidence, pessimism or optimism, representativeness, availability, anchoring

and adjustment, recency, confirmation, and default among others ((Tversky & Kahneman, 1974)). As these biases are likely to be found in individual expert's estimates, these may lead to correlation within experts (2.3.1).

2.2.1 Over-Confidence

Over-Confidence is a much studied area of human decision making because of the dramatic impact it has on the decision process, "No problem in judgment and decision making is more prevalent and more potentially catastrophic than overconfidence" (Plous & S., 1993). Top experts, by the nature of their expertise, can be very confident and tend to be very extremely opinionated (Heath & Tversky, 1991). A review of the literature seems to show a direct relation between self-rated expertise and overconfidence in judgments of widely speculative events (e.g. (Gunther, 2004), (Cesarini, Sandewall, & Johannesson, 2006)). To address this issue, approaches include using experts with varying range of expertise (e.g. (Budescu & Rantilla, 2000), (Gunther, 2004)), use of experts with varying areas of expertise (e.g. (Kuhnert, Hayes, Martin, & McBride, 2009)), use of low probability/high consequence events to de-bias (e.g. (Li, Li, Chen, Bai, & Ren, 2010)), using a feedback method for experts to revise their estimates (e.g. (Winkler, 1991)), framing of the elicitations in ways to de-bias the experts and many more de-biasing techniques. While over-confidence will always be present, the use of these de-biasing techniques will aid in reducing the error in estimations.

2.2.2 Availability

Human experts are also subject to the availability heuristic. This stems from the experts erroneously estimating the likelihood of an event to be assessed based on how retrievable some events are in memory (Tversky & Kahneman, 1974). Factors such as salience, imaginability, illusory

correlation and familiarity can therefore lead to this heuristic being a bad means of assessing the frequency and probability of events (e.g. (Tversky & Kahneman, 1974), (Frey & Rubin, 1991)). As such, an expert who conducts active research in a field of study, if elicited on the probability of breakthrough of a technology, it is possible that such expert estimates these probabilities by recalling mostly instances where such research was successful. To limit this effect, the decision analyst is urged to limit experts as much as possible to their area of main expertise (e.g. (Kynn, 2008)), use feedback methods and also use differing range of expert with different expertise levels (Kuhnert, Hayes, Martin, & McBride, 2009).

2.2.3 Representativeness

The representative heuristic is used by people in estimating the probability of random events based on a sample. As the focus is on the similarity of the sample to the estimated random event; if it shares more characteristics with the estimated random event, this leads to errors as it is not influenced by several factors that should affect the judgement of probability ((Kahneman & Tversky, 1972)(1974), (1982)). One of such factors is the neglect of the prior probability of outcomes; the expert might ignore the base rate frequency of the sampled event. Other factors are neglect of the sample size, misconceptions of the inherent randomness, illusion of validity and predictability neglect ((Kahneman & Tversky, 1972), (1982)).

2.2.4 Optimism And Pessimism

People, experts in particular, are very susceptible to the optimism bias, as we often underestimate the probability of negative events happening to us and overestimate the likelihood of experiencing positive events (Shepperd, Carroll, Grace, & Terry, 2002), (Weinstein, 1980). Similarly, the overestimation of negativity is due to the pessimism bias. Several cognitive factors can lead to

these biases some of which include representativeness, sole focus on oneself and erroneously comparing the difference of the compared sample to the evaluated event to the difference in risk. This tendency of individual experts to exhibit optimism and pessimism leads to positive correlation within experts, the resulting correlation within experts has a significant effect on the overall endpoint of the elicitation process (Baker & Olaleye, 2012).

2.2.5 Anchoring And Insufficient Adjustment

Anchoring and In-sufficient adjustment is also a byproduct of the over confidence heuristic where by experts start by anchoring their assessment to an internally determined estimate but do not adjust this properly as the other sub-events of the elicitation turn out (e.g. (Frey & Rubin, 1991), (Eroglu & Keely L Croxton, 2010)). (Eroglu & Keely L Croxton, 2010) also show that the anchoring effect might be due to personality traits such as conscientiousness and extraversion. An essential feedback process to address this bias is advocated where the expert is not overly overwhelmed by initial estimates (Kuhnert, Hayes, Martin, & McBride, 2009).

2.3 Correlation

- Correlations and Dependence in assessments

Any pair of randomly occurring variables which exhibit some form of a linear relationship are said to be correlated. Correlation can also be considered as a sort of covariance between two standardized variables (David, 1979). Since correlation is just a test for linear dependence, it is possible for the correlation between two variables to be zero and for the variables to still exhibit

some form of non-linear dependence (Embrechts, McNeil, & Straumann, 1999). While this is a much studied area of research, the general implication of correlation or dependence between variables is that they allow for projections or predictions to be made due to the underlying relationship between the variables; an example would be electricity consumption and time of day or cotton yield and rainfall (Ezekiel & Fox, 1959). It is also important to note that correlation does not imply causation or vice versa.

2.3.1 Self-Correlation

Auto-Correlation is defined as the correlation of a variable sequence to itself; as such it is a measure of the correlation between two values of the same variable at different times of a cycle (Box & Jenkins, 1976), (Weisstein, 2005)). In the context of this thesis, Self-Correlation of an individual expert is defined as when elicitations from a single expert on two independent events are correlated to one another (Baker & Olaleye, 2012). This can be viewed as some sort of inherent bias of an expert which leads to non-randomness of the experts' assessments.

All the biases discussed in the previous chapter 2.2 are possible causes of self-correlation.

Over-Confidence in experts, due to their high level of expertise, can result in expert bias (Heath & Tversky, 1991). This can lead to self-correlation within individual experts.

Since the degree of recall of similar events used to estimate the to-be assessed probability are inherent to individual experts(Tversky & Kahneman, 1974), the availability heuristic also systemically affects the decision making process of the experts, possibly leading to self-correlation within individual experts.

The use of the Representativeness heuristic can also lead to self-correlation within experts as the experts consistently ignore the base prior rate and are insensitive to the sample size (Tversky & Kahneman, 1974).

Optimism or pessimism is the most obvious causes of self-correlation in experts (Baker & Olaleye, 2012)(Cooke M. R., 1986). An expert who is optimistic will likely be much more correlated than one who is unbiased.

2.3.2 Cross-Correlation

Cross-Correlation is the measure of the correlation or similarity between two randomly occurring variables (Park L et al. (2008)). We hence consider Cross-correlation to be the correlation between the assessments about a single event from different experts. Different statistical tools can be used to measure this effect including the Pearson Product-Moment correlation coefficient, Spearman's rank correlation coefficient and various other tools in robust statistics. Some of the several possible cognitive causes of cross-correlation in elicitations are discussed below.

2.3.2.A Knowledge Based Bias

Knowledge based bias results in cross-correlation when several of the experts used in the elicitation study rely on a small size or single source of knowledge as their information source (Miller, Forthcoming), (Booker & Meyer, 1988). An instance is when the main source of information on a subject is a single report from a national laboratory; all other reports are usually versions of the main report.

2.3.2.B Similar Research Area

A similar cause of inter expert correlation in elicitations occurs when the experts selected are from very similar research fields, this leads to very high homogeneity of the opinions (Gilbert, 1994).

2.3.2.C Recency

Recency can also lead to several experts providing very similar opinions (Booker & Meyer, 1988). In eliciting chances of technological breakthroughs, the effect of recent technological breakthroughs may have an effect on how the experts perceive the elicitation.

2.3.2.D Framing

The framing and order of elicitations can also lead to correlations between experts as minute changes in the framing and order of elicitations can result in general shifts of preferences of the experts (Tversky & Kahneman, 1981). Aside generally change in preferences, the framing and order of elicitations can also cause the experts to violate rationality axioms (Tversky & Kahneman, 1981).

CHAPTER 3

BACKGROUND RESEARCH

To aid in better understanding of this paper, we carry out a brief review of the background research that this work builds on. We discuss the two aggregation orders relevant to this work in section 3.1. A review of the underlying background theory from (Baker & Olaleye, 2012) and its implications is carried out in section 3.2. A brief overview of the preceding research related to the use of the expert elicitations in the optimal portfolio model, the technologies considered in the optimal portfolio and their tree structures is carried out in Section 3.3. Section 3.4 introduces the R&D portfolio model to be used in the study and Section 3.4.2 gives a brief discussion on the role of the elicited probabilities in the R&D model.

3.1 Description And Definition Of The Three Possible Aggregation Orders

In this section, we discuss the two orders of aggregation considered in this paper. We focus only on simple mathematical equally weighted averaging. Method I refers to averaging the experts' elicitations at the sub-events level (earlier). Method II refers to averaging the expert elicitations at the de-composed sub events have been re-composed for each expert (later).

Let's take as an example an event ABC, decomposed into exclusive multiplicative events A, B and C. To assess the probability of occurrence of the event ABC i.e. P(ABC), we can decide to elicit opinions on the odds of the three decomposed sub-events A, B and C from two experts i and j.

Method I aggregates the experts at the sub-event level resulting in

$[\frac{(P_i(A)+P_j(A))}{2}]*[\frac{(P_i(B)+P_j(B))}{2}]*[\frac{(P_i(C)+P_j(C))}{2}]$ while Method II is

$\frac{[P_i(A)*P_i(B)*P_i(C)]+[P_j(A)*P_j(B)*P_j(C)]}{2}$. A graphical representation of Method I and II with two experts and 5 sub-events is shown in Figure 1.

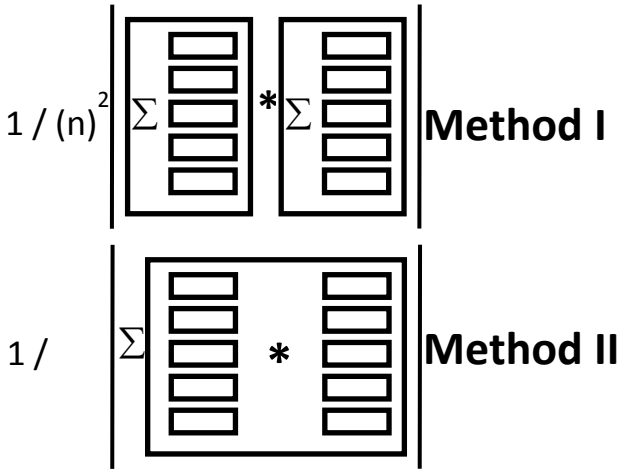


Figure 1: Methods of Simple Average Aggregation

Baker & Peng (2010) use Method II as the aggregation order for the elicitation. Here we propose to use the two aggregation orders and then compare the defective aggregation order, Method I, to Method II.

Specifically with respect to the nature of the elicitation done for this study, Method I refers to averaging of the technological success probabilities for each hurdle level across all experts before aggregating all the hurdles (technological constraints) for the final probability of success of the technologies. This implies taking the average of all the experts for each hurdle before the hurdles

are then re-composed together to get the event probability. Method II refers to when the hurdles are combined first for each expert before the probabilities from the experts are now averaged.

To show the discrepancy between the two methods, we give the underlying example using the basic model given in the beginning of this of this section. We assume that the probability of the events are given as in table 2 below, we note that the difference between Method I and II is over 54%.

$P_i(A)$	$P_i(B)$	$P_i(C)$	$P_j(A)$	$P_j(B)$	$P_j(C)$
0.8	0.4	0.9	0.3	0.2	0.3
Met 1	$[\frac{(P_i(A)+P_j(A))}{2}]*[\frac{(P_i(B)+P_j(B))}{2}]*[\frac{(P_i(C)+P_j(C))}{2}]$		$\frac{.8+.3}{2}*\frac{.4+.2}{2}*\frac{.9+.3}{2}$		**0.099**
Met 2	$\frac{[P_i(A)*P_i(B)*P_i(C)]+[P_j(A)*P_j(B)*P_j(C)]}{2}$		$\frac{[.8*.4*.9]+[.3*.2*.3]}{2}$		**0.153**

Table 1: Example to show discrepancy between methods

3.2 Review Of Theory Baker & Olaleye

A review of the "Combining Probabilities" paper by (Baker & Olaleye, 2012) is discussed in this section. The paper theoretically examines the effect of using the two orders of aggregating expert opinions (earlier and later) based on the experts' characteristics. We show in the paper that it is generally best to combine experts opinions earlier rather later as later combination leads to a higher self-correlation error. We also show that even without self-correlation within individual

experts, the variance of the error from later aggregation is larger. A brief summary of the paper is given below.

We define independent events to be indexed by i. These events are elicited by experts, indexed by j. Each elicited probability (hurdle or breakthrough technology level) is defined to be an additive combination of the true probability $\mathbf{p_i}$ "degree of belief of the entire community if all biases and errors could be avoided", $\boldsymbol{\varepsilon_{ij}}$ "errors correlated across experts", $\boldsymbol{\mu_{ij}}$ "errors correlated across hurdles or sub-events" and $\mathbf{d_{ij}}$ "independent idiosyncratic errors" (Baker & Olaleye, 2012).

$$\boldsymbol{q_{ij} = p_i + \varepsilon_{ij} + \mu_{ij} + \delta_{ij}} \qquad 1$$

Let $\boldsymbol{\rho}$ represent the correlation between the $\boldsymbol{\mu_{ij}}$and $\boldsymbol{\mu_{i'j}}$ and $\boldsymbol{\sigma^2}$represent the variance of $\boldsymbol{\mu_{ij}}$.

The resulting error and variance from both methods are shown in Table 2 where $\boldsymbol{\rho}$ represents the correlation between the $\boldsymbol{\mu_{ij}}$ and $\boldsymbol{\mu_{i'j}}$ and $\boldsymbol{\sigma^2}$represents the variance of $\boldsymbol{\mu_{ij}}$.

Method	Mean of Expected Value	Mean of Expected Error	Variance [error within events]	Variance[error across experts]
Method I	$\boldsymbol{E\left[q_{ii'}{}^{(n)}\right]}$ $\boldsymbol{= p_i p_{i'} + \frac{\rho\sigma^2}{n}}$	$\boldsymbol{\frac{\rho\sigma^2}{n}}$	$\boldsymbol{var[\mu_{ii'}]}$ $\boldsymbol{= \frac{\sigma^4}{n^2}(1+\rho^2)}$	$\boldsymbol{var\left[\varepsilon_i{}^{(n)}\varepsilon_{i'}{}^{(n)}\right]}$ $\boldsymbol{= \frac{1}{n^2}(1+(n-1)\rho_\varepsilon)^2\sigma_\varepsilon{}^4}$
Method II	$\boldsymbol{E\left[\tilde{q}_{ii'}{}^{(n)}\right]}$ $\boldsymbol{= p_i p_{i'} + \rho\sigma^2}$	$\boldsymbol{\rho\sigma^2}$	$\boldsymbol{var[\tilde{\mu}_{ii'}]}$ $\boldsymbol{= \frac{\sigma^4}{n}(1+\rho^2)}$	$\boldsymbol{var[\tilde{\varepsilon}_{ii'}]}$ $\boldsymbol{= \frac{1}{n}(1+(n-1)\rho_\varepsilon{}^2)\sigma_\varepsilon{}^4}$

Table 2: Theoretical Comparisons of Method I and II (Baker & Olaleye (2012))

We can observe that the mean of the expected error from Method II is n times greater than Method I, where n is the number of experts. This is a result of the self-correlation $\boldsymbol{\rho}$ within experts which is not zero. If the self-correlation $\boldsymbol{\rho}$ is zero then the mean of the expected error is the same for both methods. Also the variance of the error correlated within events and across experts from Method I is also shown to be lower than those of Method II.

This shows that the variance of Method II is always greater than that of Method I even in the absence of correlation i.e. even when $\boldsymbol{\rho}$ is zero.

3.3 Review Of Background Research On The Technologies Elicited

In this section, a discussion of the methodology of the selection of the energy technologies used in the energy portfolio model is carried out. In making policy recommendations on funding of R&D in energy source technologies that can readily combat climate change, (Baker & Solak, 2011) focused on some key criteria for climate damages abatement reduction. These criteria include electricity generation potential, uncertainty in success of the specific technology and having a considerably large resource base. Thus three technologies were considered in the portfolio: Solar, Carbon Capture and Storage (CCS) and Nuclear. These technologies are discussed briefly below. Figures 3 and 4 gives a brief summary of the classification of the sub-technologies, the technology success levels and the funding levels considered in the study.

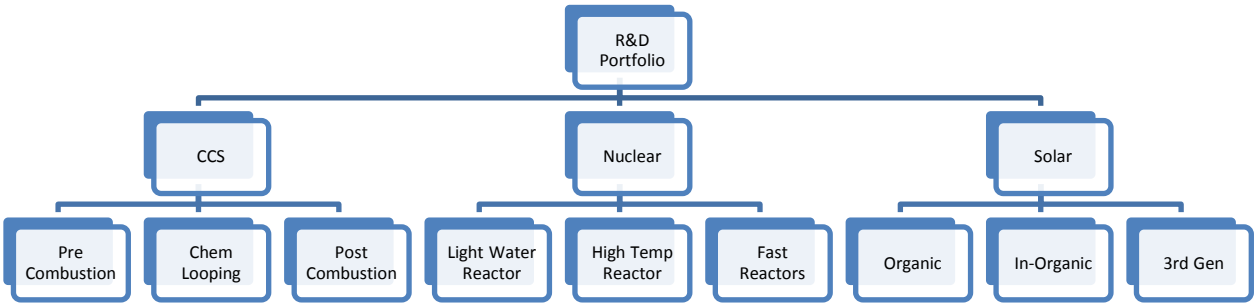


Figure 2: Portfolio Technologies and Sub Technologies considered

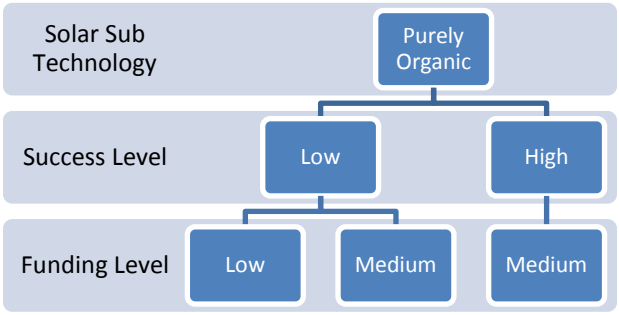


Figure 3: Success Level and Funding Level in Purely Organic Solar Sub-Technology

Solar Photovoltaic: The three sub-technologies elicited were Organic Photovoltaic, In-Organic Photovoltaic and Third generation solar cell technologies; see (Baker, Chon, & Keisler, 2009) for a detailed discussion of the technological hurdles and tree structure. These sub technologies were elicited by three experts. It is important to note that the most cost efficient sub-technology will take the whole solar energy share.

Nuclear Energy: The three sub-technologies considered here were Advanced Light Water Reactors, High Temperature Reactors and Fast Reactors. (Baker, Chon, & and Keisler, 2008) provides a very detailed description of the technological hurdles and the tree structure.

Carbon Capture and Storage: Please refer to (Baker, Chon, & and Keisler, 2009) for a detailed description of the technological hurdles and the tree structure for the three sub-technologies; Pre-Combustion, Post-Combustion and Chemical Looping technologies.

(Peng, 2010) gives a detailed review of the optimal portfolio model used. (Baker & Shittu, 2005) and (Baker & Solak, 2011) give a detailed rationale for the choice of these technologies.

3.4 Optimal Energy Portfolio Model

This section discusses the nature of the optimal portfolio model problem and the role of the elicited probabilities in the model.

3.4.1 Optimal Portfolio Model

In this section we discuss briefly the specific model used in (Baker & Peng, 2012) and (Baker & Solak, 2011), which we build on, and use later, in this thesis. The objective of the model is to minimize the expected net social cost which is the sum of the expected damages and the cost of abatement. The exact form of the formulation is given below.

$$\min_{X} E_{\alpha,Z}\left(\min_{\mu}[C(\mu:\alpha(X))+ZD(\mu)]\right) \quad (2)$$

Subject to

$$\sum_{i}\sum_{j}\sum_{k} f_{ijk}x_{ijk} \le B \quad (3)$$

$$\sum_{k} x_{ijk} \le 1 \;\forall i,j \quad (4)$$

Where $C(\mu;\vec{\alpha})$ represents the cost of abatement after investment in a portfolio of technologies [5]. The cost is dependent on the state of the portfolio invested in $\vec{\alpha}$, and the amount of abatement

μ∈[0,1], the abatement cost at 50% abatement($c(50\%)$), and the abatement cost before technical change C(μ) [6]. The damage function is represented by D(μ) [7] and is multiplied by a random multiplier to represent the uncertain nature of the possible damages from climate change. Indices i, j and k index the technology, sub-technology and funding level respectively. X_{ijk} denotes a binary variable indicating if sub-technology j is funded at level k, f_{ijk} denotes the amount of funding at this level and where M is a parameter of the model. This formulation [(2)-(4)] is to obtain the optimal portfolio that minimizes the expected social cost. Where

$$C(\mu:\alpha)=\prod_i(1-\alpha_i)[c(\mu)-hc(50\%)\mu] \quad (5)$$

$$C(\mu)=b_0\mu^{b_1} \quad (6)$$

$$D(\mu)=M_0(S-M_1\mu)^2 \quad (7)$$

For better sensitivity analysis, (Peng, 2010) developed a slight modification to the above formulation including the R&D budget cost in the objective function of the formulation. An opportunity cost β (1, 2, 4 or 8) is used for the research and development cost. The formulation is given below

$$\underset{X}{Min}\ \beta\sum_i\sum_j\sum_k F_{ijk}X_{ijk}+E_{\alpha,Z}\left(\min_{\mu}[C(\mu:\alpha(X))+ZD(\mu)]\right) \quad (8)$$

$$\text{Subject to } \sum_k X_{ijk}\le 1\ \forall i,j \quad (9)$$

3.4.2 Role Of The Elicited Probabilities In The Optimal Portfolio Model

In this section we discuss the role of the elicited probabilities in the optimal portfolio model. This is to emphasize its importance on the optimal portfolio output. The expert elicitations for all the nine sub technologies in the three technologies considered, (Solar Photovoltaic, Nuclear and Carbon Capture &Storage), serve as input to condition the expectation of each of the possible realizations α of the objective function [ ]. Due to the uncertain and random two stage decision nature of the climate problem; all possible realizations of the investment portfolio for the nine sub technologies must be considered. Then by independence of the sub-technology elicited probabilities, the probability of each scenario is simply the multiplication of the nine (j) elicited probabilities depending on the funding level (k) and success level (l).

10

$$P = \prod_{i,j,k,l} p_{ijk,l} \qquad \forall j \tag{10}$$

The resulting probability P of each scenario (example [0, 0, 1, 1, 1, 0, 1, 0, 1]) is then used to appropriately weight each of the resulting random expectations. As such we expect the elicited probabilities to have significant impact on the model output.

A detailed summary of the interaction of the model components from the parameterization of the marginal abatement curve, the calibration to the Dice 2007 model and the derivation of the base portfolio is given in (Baker & Solak, Under Review).

CHAPTER 4

MONTE CARLO SIMULATION

In this chapter, we discuss our work on the development of a Monte Carlo simulation of the mathematical formulation of the problem. We proceed by basing our simulation on the underlying theoretical model in [1]. We start by modeling all the individual components of each sub event and then aggregate these sub events to obtain the joint event probability from using both orders of aggregation. A discussion on the simulation model used is described in this chapter.

There are three variants of the simulation model. Model 1 (Base Un-truncated) is a simulation exactly like the theoretical model described in Section 3. A problem, however, exists that the resulting joint event probability might not be limited to [0, 1]. This is due to the nature of how we model the ε_{ij}, μ_{ij} and δ_{ij}errors, discussed in detail below. To address this problem i.e. limit the end probability to [0, 1], we develop two additional variants of the base model; a log odds variant and another variant where the probabilities are truncated at 0 and 1. The other two variants (Model 2 and Model 3) are minor adjustments to Model 1. As such in the sections below, Model 1 is described in detail and only the difference from Model 1 is discussed for Models 2 and 3. Section 4.1 outlines the need for the simulation model, Section 4.2 describes the simulation of the Base Un-truncated model, Section 4.3 discusses the Log Odds Model, the Truncated variant is briefly discussed in Section 4.4 and analysis of the results obtained discussed in Section 4.5.

4.1 Need for Simulation

In this section, we discuss several of the reasons why a Monte Carlo simulation of the mathematical formulation of the problem is required. Some of the advantages of having the simulation are given below.

4.1.1 Evaluation Of Different Aspects Of The Formulation

One of the main benefits of doing a simulation of the problem is that we can consider different variants of the same problem including those which do not have the deficiency (the resulting probability q_{ij} [1] not been limited between 0 and 1) as the theoretical formulation. As such we can both model the exact same problem and also different variants where we can limit the resulting probability q_{ij} between [0, 1]. The simulation, when the resulting probability is bounded [0, 1], therefore gives a better representation of the problem being modeled (4.3).

4.1.2 Examination Of Different Properties Of The Distribution Of The Errors

Another major advantage of having a simulation model is to enable better analysis of the distribution of the errors from both orders of combination. We can examine in detail and graphically the distribution of the errors from these methods. This enables us to evaluate several characteristics of the errors, some of which includes the mean, variance and the absolute mean of error. We can do this for different combinations of standard deviations and correlation coefficients. We can also measure the probability that the size of the absolute error will be larger when either order of combination is used. The graphical output from the simulation model also ensures we can better characterize either aggregation order's performance.

4.2 Model 1 -Base Model (Un-Truncated [0, 1])

The simulation discussed here follows exactly as the theoretical model in [1]. A description of how p_{ij}, ε_{ij}, μ_{ij}andδ_{ij} from [1] are modeled for the base model is given below.

For simplicity and since any space can be modeled by the intersection of any two events, we model the main event as decomposed into two mutually exclusive events, these two mutually exclusive events are elicited from 5 experts, see Figure 4. Therefore n=5 and $\boldsymbol{q_{ij}}$'s =2.

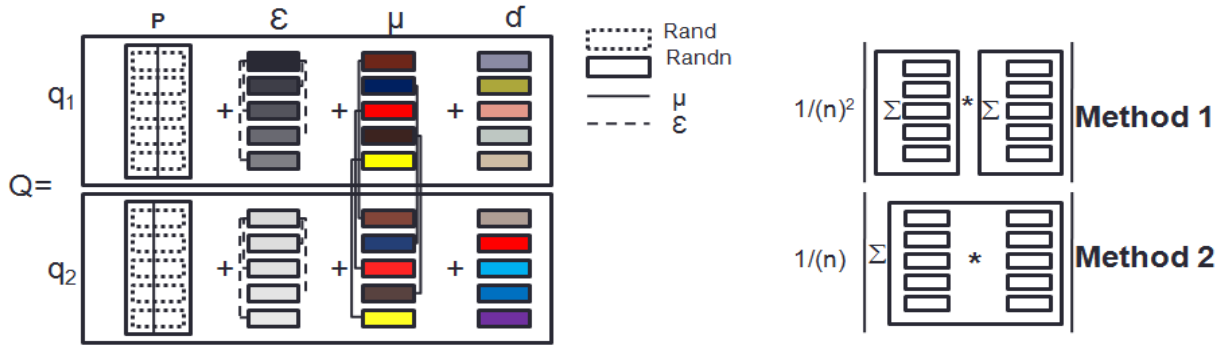


Figure 4: Graphical description of the Model and Simulation

4.2.1 Generating Random Errors And Random Probabilities

We describe the modeling of each generic true probability or error probability in this section. The entire model [1] can be observed to be composed of either a 'true probability' $\boldsymbol{p_i}$or an 'errorprobability'$\boldsymbol{\varepsilon_{ij}, \mu_{ij}\ or\ \delta_{ij}}$.

True Probability $\boldsymbol{p_i}$: As we expect the true probability of an event to be a value between [0,1], we model the Random probabilities as a random number between [0, 1] using the 'rand' function in Matlab.

Error Probability: We model random errors as a random variable with mean zero. The 'randn' function in Matlab is used which gives a standard normal distribution with an expected value of zero.

Based on this representation, we construct a simple model consisting of two sub-events (technological hurdles) and five experts as shown in Figure 4. For simplicity, the experts are modeled each time as a [5X1] column for the true probability or the errors. This results in four ($\boldsymbol{p_i}$, $\boldsymbol{\varepsilon_{ij}}$, $\boldsymbol{\mu_{ij}}$ $\boldsymbol{or}$ $\boldsymbol{\delta_{ij}}$) [5X1] columns for each of the two sub-events. A detailed description is given in the subsequent sections.

4.2.1.A Generating The True Probability [p]

In modeling the true probability of an event to be elicited by all the experts, we expect the true probability of that event to be the same for all the experts. This means that without any bias or misjudgment all the experts eliciting a sub-event should have the same true probability. As discussed above in Section 4.2.1, the random true probabilities are modeled as a random number [0, 1]. Then to ensure that this true probability is constant across the experts, we simply multiply the single generated true probability by a [5X1] matrix of one's. This is done for each sub event. The actual modeling is described here; a single random variable is generated for both hurdles i.e.

$$\boldsymbol{p_i} = \boldsymbol{rand}() \text{ and } \boldsymbol{p_j} = \boldsymbol{rand}() \qquad \textbf{11}$$

Therefore the first sub-event True Probability is modeled as **p1=$\boldsymbol{p_i}$*([1 1 1 1 1]^t)**. 12

The other sub-event True Probability is modeled as **p2=$\boldsymbol{p_j}$*([1 1 1 1 1]^t)**. 13

4.2.2 Generating The Errors Terms

We discuss the generation of the errors in this section. The errors include the self-correlation, cross-correlation and idiosyncratic errors. To appropriately model the errors, we need to condition the generated errors to the parameters representing the correlation and standard deviation. We model these parameters as inputs to the model and let $\boldsymbol{\rho^{\mu}}$represent the correlation coefficient while $\boldsymbol{\sigma^{\mu}}$ represents the standard deviation for the errors correlated within experts [$\boldsymbol{\mu}$]. $\boldsymbol{\rho^{\varepsilon i}}$ represents the correlation coefficient and $\boldsymbol{\sigma^{\varepsilon}}$ represent the standard deviation for the errors correlated across experts [$\boldsymbol{\varepsilon}$]. We also let $\boldsymbol{\sigma^{\delta}}$ represent the standard deviation for the idiosyncratic errors [$\boldsymbol{\delta}$]. The modeling of each of the three errors is discussed in detail below.

4.2.2.A Generating The Errors Correlated Within The Expert [μ]

Here we describe the modeling of the self-correlation errors. While it is expected that the 'true event probability' is the same for all the experts in each sub-event, we do not expect the errors to be the same. We basically want to model an error for each individual expert which is correlated within the expert for all the sub events. As such we first generate two sets of [5X1] random standard normal variables.

rv1[μ] = randn(5, 1)

rv2[μ]=randn(5, 1) 14

As our aim is to correlate each individual expert (see Column 3 in Figure 2) to itself, hence, the first column **rv1[μ]** representing the first sub-event is conditioned to mean zero and input standard deviation $\boldsymbol{\sigma^{\mu}}$for self correlated errors.

$$\mu 1=\mu rv1[\mu] = 0 + \sigma^{\mu} * rv1[\mu] \quad (15)$$

The elicited probability from each expert of the first column representing the first sub-event is then correlated to the elicited probability from the same expert in the second sub-event as in Figure 4. The resulting expert is conditioned to mean 0 and the input standard deviation for self correlated errors σ^{μ} [16].

$$\mu 2=\mu rv2[\mu] = \{0 + \sigma^{\mu} *(\rho^{\mu}*rv1[\mu](1,1) + sqrt(1 - (\rho^{\mu})^2)rv2[\mu](1,1));$$

$$0 + \sigma^{\mu} *(\rho^{\mu}*rv1[\mu](2,1) + sqrt(1 - (\rho^{\mu})^2)rv2[\mu](2,1));$$

$$0 + \sigma^{\mu} *(\rho^{\mu}*rv1[\mu](3,1) + sqrt(1 - (\rho^{\mu})^2)rv2[\mu](3,1));$$

$$0 + \sigma^{\mu} *(\rho^{\mu}*rv1[\mu](4,1) + sqrt(1 - (\rho^{\mu})^2)rv2[\mu](4,1));$$

$$0 + \sigma^{\mu} *(\rho^{\mu}*rv1[\mu](5,1) + sqrt(1 - (\rho^{\mu})^2)rv2[\mu](5,1))\} \quad (16)$$

This uses the relation that y = s*(p*u+sqrt(1-p^2)*v)+m where 'u' is the random variable been correlated to the random variable 'v' while 'p', 's' and 'm' are the correlation, standard deviation and mean respectively. Each Individual expert is as such now self-correlated.

4.2.2.B Generating The Errors Correlated Across Experts [ε]

Cross correlated errors are errors that are correlated across all the experts within a sub-event or hurdle. Again here we do not expect each individual expert to have the same cross correlated error, so we model the cross correlated errors as different for each expert. As generating a

sequence of random variables that are cross correlated while maintaining their independence is difficult, we model this by correlating each of the experts elicited probability to a different unused expert or support variable elicited probability (we call this 'support variable'). The new correlation between the elicited experts elicited probability is then the square of the correlation of the individual experts elicited probability to the unused variable or expert (support variable) elicited probability.

We generate a base random error for each sub-event to represent the support variables to which each of the individual experts are to be correlated to. These support variable errors are modeled as usual as mean zero Standard Normal variables with the randn function in Matlab **ru1[Ɛ]** and **ru2[Ɛ]**.

As this does not represent an expert, we do not condition with a mean or a standard deviation. These results in

$$\textbf{ru1[Ɛ] = randn()}$$

$$\textbf{ru2[Ɛ] = randn()} \qquad \textbf{17}$$

All five experts elicited probability are then correlated to the support variable's elicited probability for each of the two sub events and conditioned again with mean 0 and input standard deviation for cross correlated errors $\boldsymbol{\sigma}^{\varepsilon}$. We use the same relation y = s*(p*u+sqrt(1-p^2)*v)+m but using individual variables of the matrix here.

These results in the sub-event 1 across correlated errors for five experts elicited probability [5X1] as

$$\textbf{eƐru1[Ɛ]=\{0 + } \boldsymbol{\sigma}^{\varepsilon} \textbf{ *(} \boldsymbol{\rho}^{\varepsilon} \textbf{*ru1[}\varepsilon\textbf{] + sqrt(1 – (} \boldsymbol{\rho}^{\varepsilon}\textbf{)}^{\mathbf{2}}\textbf{)*randn());}$$

$$0 + \sigma^{\varepsilon} * (\rho^{\varepsilon} * ru1[\varepsilon] + sqrt(1 - (\rho^{\varepsilon})^2) * randn());$$

$$0 + \sigma^{\varepsilon} * (\rho^{\varepsilon} * ru1[\varepsilon] + sqrt(1 - (\rho^{\varepsilon})^2) * randn());$$

$$0 + \sigma^{\varepsilon} * (\rho^{\varepsilon} * ru1[\varepsilon] + sqrt(1 - (\rho^{\varepsilon})^2) * randn());$$

$$0 + \sigma^{\varepsilon} * (\rho^{\varepsilon} * ru1[\varepsilon] + sqrt(1 - (\rho^{\varepsilon})^2) * randn())\} \quad 18$$

And the sub-event/hurdle 2 cross correlated errors **Ɛ2** is

$$e\text{Ɛ}ru2[\text{Ɛ}] = \{0 + \sigma^{\varepsilon} * (\rho^{\varepsilon} * ru2[\varepsilon] + sqrt(1 - (\rho^{\varepsilon})^2) * randn());$$

$$0 + \sigma^{\varepsilon} * (\rho^{\varepsilon} * ru2[\varepsilon] + sqrt(1 - (\rho^{\varepsilon})^2) * randn());$$

$$0 + \sigma^{\varepsilon} * (\rho^{\varepsilon} * ru2[\varepsilon] + sqrt(1 - (\rho^{\varepsilon})^2) * randn());$$

$$0 + \sigma^{\varepsilon} * (\rho^{\varepsilon} * ru2[\varepsilon] + sqrt(1 - (\rho^{\varepsilon})^2) * randn())\} \quad 19$$

To now obtain the desired input correlation between the experts elicited probability, this is simply the square root of the correlation between the elicited probability from individual experts and the support variable.

As such $$\rho^{\varepsilon i} = (\rho^{\varepsilon})^{1/2}$$

4.2.2.C Generating The Idiosyncratic Errors[δ]

The modeling of the idiosyncratic errors is discussed here. Idiosyncratic errors are errors that are totally random without any inherent correlation either within or across experts, therefore it is expected that the idiosyncratic error will be different for each expert as they are totally random and not correlated Figure 4. As such we model these as random Standard Normal and condition on the input standard deviation for the idiosyncratic error $\boldsymbol{\sigma}^{\text{d}}$

$$\textbf{d1= drv1[d]}=0 + \boldsymbol{\sigma}^{\text{d}} \textbf{*randn(5, 1)}$$

$$\textbf{d2= drv2[d]}=0 + \boldsymbol{\sigma}^{\text{d}} \textbf{*randn(5, 1)} \qquad \textbf{20}$$

4.2.3 Generating The Resulting Sub-Event Probability [q]

Here we discuss the addition of the true probability and the errors. All the terms that compose the resultant error are then added for each sub-event/hurdle

$$\boldsymbol{q1} = \boldsymbol{p1} + \boldsymbol{\varepsilon 1} + \boldsymbol{\mu 1} + \boldsymbol{\delta 1}$$

$$\boldsymbol{q2} = \boldsymbol{p2} + \boldsymbol{\varepsilon 2} + \boldsymbol{\mu 2} + \boldsymbol{\delta 2} \qquad \textbf{21}$$

It should be noted that the resultant probabilities **q1** and **q2** are [5X1] columns consisting of the resultant probabilities of each expert for each hurdle.

4.2.4 Overall Errors From Both Methods

On obtaining the $\boldsymbol{q_{ij}}$'s for each hurdles, we now need to combine the experts using the two possible aggregation orders. The true probability from eliciting both sub-events/hurdles is simply

$$P = p1 * p2 \qquad 22$$

Method I: Since Method I implies aggregating early, the probability of each of the two events 1 and 2 are calculated by first averaging across the experts. The probability of the joint event is then the product of the two averaged sub event components (see Figure 1 and Figure 2). Hence the resulting probability from using Method I is

$$M1 = ((sum(q1))/n) * ((sum(q2))/n) \qquad \text{where n is 5} \qquad 23$$

The error from Method I is then

$$eM1 = M1 - P \qquad 24$$

Method II: In Method II, aggregation is later; the probability of the joint event is calculated for each expert before the experts are then averaged (see Figure 1 and Figure 2). The resulting probability from using Method II is

$$M2 = (sum(q1.*q2))/n) \qquad \text{where n is 5} \qquad 25$$

The error from Method II is then

$$eM2 = M2 - P \qquad 26$$

4.2.5 Monte Carlo Sampling

The Monte Carlo simulation is discussed here. To do a Monte Carlo simulation, we use two Matlab 'm' files; the first m-file gives a function that outputs the true Probability (say P), the error from Method I(eM1)and the error from Method II (eM2) for a single run. The second m-file does a

Monte Carlo sampling using the outputs from the first m-file as inputs. We use a sample rate of 1,000,000.

We ran the model using the different combinations of the input correlation and standard deviation parameters Table 3. In all the cases simulated, the results obtained were consistent with the theoretical formulation; some new results were shown in Table 4 and discussed in Section 4.2.6.

$\boldsymbol{\rho_{\mu}}$	-0.8	0	0.8	1
$\boldsymbol{\sigma_{\mu}}$	0.2	0.5		
$\boldsymbol{\rho_{\varepsilon}}$	0	0.2	0.5	0.8
$\boldsymbol{\sigma_{\varepsilon}}$	0	0.2	0.8	
$\boldsymbol{\sigma_{\delta}}$	0	0.2	0.8	

Table 3: Simulation Parameters for Base Model (77 Combinations Simulated)

4.2.6 Results

We discuss the results from the Base simulation model in this section. Figure 5 shows the no risk case, where the input parameters (Standard deviation and variance) for the idiosyncratic and cross correlated errors were zero. It gives the ratio of the expected error of Method II to Method I. We see that Method II truly does not correct for self-correlation between the experts and the difference between the two methods increases as the self-correlation increases.

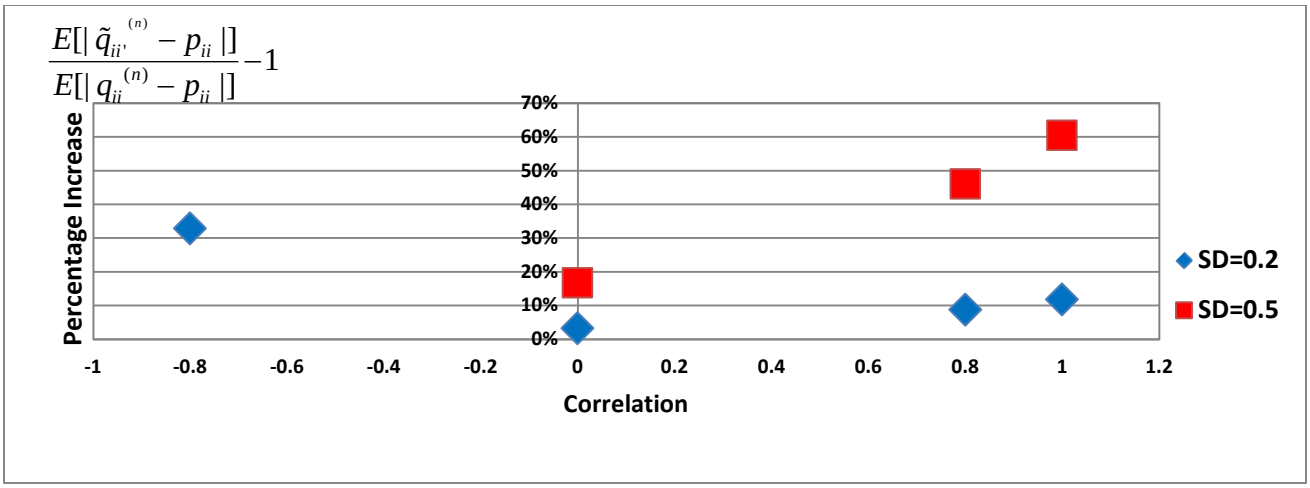


Figure 5:Ratio of the expected error in Method II to Method I

ρ corr	σ sd	Mean(Q-P)		Var (Q-P)		Mean |Q-P|		Pr{|Qii-P|<|Qi-P|}
		Mt I	Mt II	Mt II	Mt II	Mt I	Mt II	
0	0.2	0	0	0.0054	0.0057	0.0549	0.0567	0.4554
0	0.5	0.0002	0.0003	0.0358	0.0458	0.1396	0.163	0.4
0.8	0.2	0.0065	0.0321	0.0086	0.0091	0.0684	0.0744	0.4216
0.8	0.5	0.0398	0.2	0.0574	0.074	0.1721	0.2513	0.3142
1	0.2	0.0079	0.0399	0.0094	0.0099	0.0712	0.0796	0.4035
1	0.5	0.0501	0.2501	0.0634	0.0835	0.1798	0.2886	0.2718
-0.8	0.2	-0.0064	-0.032	0.0022	0.0027	0.035	0.0465	0.329

Table 4: Results from Simulation of Base model

Table 4 gives the results for a subset (Ɛ and ɗ parameters are set to zero) of the parameter combination given in Table 3. Method I showed dominance to Method II in all the cases considered as the absolute error was lower in Method I than Method II. The mean of the error from Method II was also of order n=5 greater than that of Method I. The variance of the error can also be noticed to always be higher in Method II. The last column in Table 4 gives the probability than the absolute error in Method II is greater than that of Method I in the 1,000,000 runs, we note that this also favors Method I. As discussed above, it is obvious that probabilities not within the range [0, 1] are possible. To address this, two modifications to the above model are developed.

4.3 Model 2: Log Odds Adjusted

The Log Odds variant of the model is discussed in this section. The only change to the above model is that after the random true probabilities $\boldsymbol{p_i}$ are generated, they are log-odds transformed.

From the generation of the single random variable for each sub-event

$$\boldsymbol{p_i = rand()} \text{ and } \boldsymbol{p_j = rand()} \quad \textbf{27}$$

Both are then log-odds transformed $\boldsymbol{h_i = \log\left[\frac{p_i}{1- p_i}\right]}$ **28**

The true probability for all the experts for sub-event 1 is then **h1= $\boldsymbol{h_i}$*([1 1 1 1 1]^t)**.

Everything proceeds as in the base model until when the resulting probability for each sub-event has been obtained, the sum $\boldsymbol{q1}$ is then re-transformed.

$$\boldsymbol{q1 = h1 + \varepsilon 1 + \mu 1 + \delta 1} \quad \textbf{29}$$

$$\boldsymbol{qL1 = \left[\frac{\exp(q1)}{1+q1}\right]} \quad \textbf{30}$$

4.2.3.A Results Log Odds Model

Table 5 gives the combination of parameters used for the Log Odds simulation. We chose the standard deviations to be relatively large compared to the Base Model, this is due to the Log Odds transformation.

ρ_μ	-0.8	0	0.8	1
σ_μ	5	10		
ρ_ε	0	0.2	0.5	
σ_ε	0	5	10	
σ_δ	0	5	10	

Table 5: Simulation Parameters Log Odds Model (49 Combinations Simulated)

The results obtained Table 6 from this model were similar to the base model and are not discussed.

		Mean(Q-P)		Var (Q-P)		Mean |Q-P|		Pr{|Qii-P|<|Qi-P|}
Corr	sd	Mt 1	Mt 2	Mt 1	Mt 2	Mt 1	Mt 2	
0	5	0.0001	0.0001	0.0372	0.0414	0.1476	0.1563	0.4475
0	10	0	0	0.0537	0.0606	0.1799	0.192	0.4449
0.8	5	0.0221	0.1101	0.0483	0.0499	0.1688	0.2003	0.3532
0.8	10	0.0272	0.1354	0.0677	0.0703	0.2021	0.2425	0.354
1	5	0.0295	0.1474	0.0521	0.0516	0.1755	0.2201	0.3198
1	10	0.039	0.1947	0.0738	0.0719	0.2116	0.2738	0.3099
-0.8	5	-0.0221	-0.1101	0.0263	0.0279	0.1246	0.1413	0.4349

Table 6: Log Odds Simulation Results

4.4 Model 3: Truncated Variant

To limit the resulting probabilities to be between [0, 1] we develop this third variant Truncated Variant. We discuss only the difference between the base model and the truncated model. Here, the resultant probability $\boldsymbol{q's}$ are truncated between [0, 1] for each individual run.

$$\boldsymbol{q1} = \boldsymbol{min}(\boldsymbol{max}\big((\boldsymbol{h1} + \boldsymbol{\varepsilon 1} + \boldsymbol{\mu 1} + \boldsymbol{\delta 1}), ([\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}])\big), ([\mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1}]))$$

The results we obtain from this simulation were also similar to that of the base Model 1.

4.5 Results Analysis

In this section we discuss the results from the simulation runs. The mean, variance and absolute mean (Table 6) are generally in favour of Method I as theoretically predicted. For the probability which shows the dominance of Method I [$\mathbf{Pr\{|Q_{ii}-P|>|Q_i-P|\}}$], we note that sometimes the error from Method II is lesser as the probability range obtained is between (0.27-0.46). This can be readily attributed to randomness in the simulation. This is because the resulting error term from the use of both terms is quite small in comparism to the simulated true random probabilities and additive random errors. Thus variation due to randomness in these simulated probability and additive errors

leads to a very significant number of cases where the error from Method II is lesser. As the cases where the error from Method II is lesser are totally random, it will always be preferable to use Method I as the desired method of aggregation.

CHAPTER 5

THEORETICAL PREDICTION OF THE DOMINANCE OF METHOD I

To obtain a sense of the degree of theoretical dominance of Method I, we estimate the probability that the error from Method I will be lesser than the error from Method II. While the simulation results from the previous chapter (Table 4 and Table 6) gives this dominance from a modeling perspective, to further corroborate the simulation results we present a theoretical estimation of this dominance.

The means and variances of the errors from (Baker & Olaleye, 2012) Table 2 provide a ready starting point for this evaluation. As the distribution of the errors is not readily known, we assume a normal distribution for the errors from both methods. The errors from both orders of aggregation are expected to be very highly correlated.

5.1 Probability That Method I Error Is Lesser Than Method II Error

Let $e(q_{ii'})$ and $e(\tilde{q}_{ii'})$ represent the error term for Method I and II respectively. Also we let $\mho_I$, $\mho_{II}$, $\varpi_I{}^2$ and $\varpi_{II}{}^2$ represent the error mean and error variance from Methods I and II respectively.

Hence, from (Baker & Olaleye, 2012) $\mho_I = E\left(e(q_{ii'})\right) = \frac{\rho\sigma^2}{n}$ and $\mho_{II} = E\left(e(\tilde{q}_{ii'})\right) = \rho\sigma^2$. From Table 2 and (Baker & Olaleye, 2012), we have that $\varpi_I{}^2 \approx \mathrm{var}\left(\mu_{ii'}\right) = \frac{\sigma^4}{n^2}\left(1+\rho^2\right)$ and

$\varpi_{II}{}^2 \approx \mathrm{var}\left(\tilde{\mu}_{ii'}\right) = \frac{\sigma^4}{n}\left(1+\rho^2\right)$.

We assume that the errors are normally distributed with the given means and variances.

Therefore $e(q_{ii'}) \sim N(\mho_I, \varpi_I^2)$ and $e(\tilde{q}_{ii'}) \sim N(\mho_{II}, \varpi_{II}^2)$. 31

Hence $P\left(e(\tilde{q}_{ii'}) > e(q_{ii'})\right) = P\left(e(\tilde{q}_{ii'}) - e(q_{ii'}) > 0\right) = 1 - P\left(e(\tilde{q}_{ii'}) - e(q_{ii'}) \le 0\right)$ 32

Given our normality assumption, we know that $[e(\tilde{q}_{ii'}) - e(q_{ii'})]$ is normally distributed with:

Mean $\mho: E\left([e(\tilde{q}_{ii'}) - e(q_{ii'})]\right) = \mho_{II} - \mho_I = \rho\sigma^2\left[\frac{n-1}{n}\right]$ 33

Variance $\varpi^2 = Var\left([e(\tilde{q}_{ii'}) - e(q_{ii'})]\right) = \varpi_I^2 + \varpi_{II}^2 - 2\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})}\varpi_I\varpi_{II}$ 34

Where $\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})}$ is the correlation between both errors. $\mho$ and ϖ are the mean and standard deviation for the difference in the errors of Method II and Method I $[e(\tilde{q}_{ii'}) - e(q_{ii'})]$.

Hence from Equations 31 and 33,

$$\frac{[e(\tilde{q}_{ii'}) - e(q_{ii'})] - \mho}{\varpi} \sim N(0,1) \quad 35$$

Therefore $P\left([e(\tilde{q}_{ii'}) - e(q_{ii'})] \le 0\right) = P\left[\frac{[e(\tilde{q}_{ii'}) - e(q_{ii'})] - \mho}{\varpi} \le \frac{0 - \mho}{\varpi}\right] = \Phi\left(\frac{-\mho}{\varpi}\right)$ 36

Where Φ is the standard normal cumulative distribution function of the N(0,1)distribution.

Therefore $P\left(e(\tilde{q}_{ii'}) \le e(q_{ii'})\right) = \Phi\left(\frac{-\mho}{\varpi}\right)$ 37

In the next section, we evaluate $\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})}$ from 34.

5.1.1 Correlation Between The Errors $\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})}$

The correlation is by definition the covariance divided by the product of the standard deviation.

$$\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})} = Corr\left(e(\tilde{q}_{ii'}), e(q_{ii'})\right) = \frac{\text{cov}(e(\tilde{q}_{ii'}), e(q_{ii'}))}{\varpi_I \varpi_{II}} = \frac{E\left[(e(q_{ii'}) - \mho_I)(e(\tilde{q}_{ii'}) - \mho_{II})\right]}{\varpi_I \varpi_{II}} \quad 38$$

From (Baker & Olaleye, 2012), $e(q_{ii'})$ and $e(\tilde{q}_{ii'})$, the error from Method I and II is given as

$$e(q_{ii'}) = \mu_{ii'} = \mu_{i^{(n)}} \mu_{i'^{(n)}} = \frac{1}{n}\sum_{j=1}^{n} \mu_{ij} \frac{1}{n}\sum_{j=1}^{n} \mu_{i'j} \quad 39$$

$$e(\tilde{q}_{ii'}) = \tilde{\mu}_{ii'} = \frac{1}{n}\sum_{j=1}^{n} \mu_{ij}\mu_{i'j} \quad 40$$

From Table 2, we know $\mho_I = \frac{\rho\sigma^2}{n}$, $\mho_{II} = \rho\sigma^2$, $\varpi_I = \sqrt{\frac{\sigma^4(1+\rho^2)}{n^2}}$ and $\varpi_{II} = \sqrt{\frac{\sigma^4(1+\rho^2)}{n}}$.

Substituting these terms into 38, we obtain

$$\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})} = \frac{E\left[(\frac{1}{n}\sum_{j=1}^{n} \mu_{ij} \frac{1}{n}\sum_{j=1}^{n} \mu_{i'j} - \frac{\rho\sigma^2}{n})(\frac{1}{n}\sum_{j=1}^{n} \mu_{ij}\mu_{i'j} - \rho\sigma^2)\right]}{\sqrt{\frac{\sigma^4(1+\rho^2)}{n^2}} * \sqrt{\frac{\sigma^4(1+\rho^2)}{n}}} \quad 41$$

Taking out $\frac{1}{n}$ from $(\frac{1}{n}\sum_{j=1}^{n} \mu_{ij} \frac{1}{n}\sum_{j=1}^{n} \mu_{i'j} - \frac{\rho\sigma^2}{n})$ and simplifying the denominator gives

$$\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})} = \frac{n^{3/2}}{\sqrt{\sigma^8(1+\rho^2)^2}} * \frac{1}{n} * E\left[(\frac{1}{n}\sum_{j=1}^{n} \mu_{ij} \sum_{j=1}^{n} \mu_{i'j} - \rho\sigma^2)(\frac{1}{n}\sum_{j=1}^{n} \mu_{ij}\mu_{i'j} - \rho\sigma^2)\right] \quad 42$$

Multiplying out we get

$$\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})} = \frac{n^{1/2}}{\sigma^4(1+\rho^2)} * E\left[(\frac{1}{n^2}\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\mu_{ij}\mu_{i'j} - \frac{\rho\sigma^2}{n}\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j} - \frac{\rho\sigma^2}{n}\sum_{j=1}^{n}\mu_{ij}\mu_{i'j} + \rho^2\sigma^4) \right] \quad 43$$

Since $E[\mu_{ij}\mu_{ij'}]_{j\neq j'} = 0$, expanding we obtain

$$\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})} = \frac{n^{1/2}}{\sigma^4(1+\rho^2)} * E\left[\frac{1}{n^2}\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\mu_{ij}\mu_{i'j} - (\frac{2\rho\sigma^2}{n}\sum_{j=1}^{n}\mu_{ij}\mu_{i'j}) + \rho^2\sigma^4 \right]$$

From (Baker & Olaleye, 2012) we know that the expected value of $\sum_{j=1}^{n}\mu_{ij}\mu_{ij'} = n\rho\sigma^2$, hence

$$\rho_{e(q_{ii'}),e(\tilde{q}_{ii'})} = \frac{n^{1/2}}{\sigma^4(1+\rho^2)} * [E\left[\frac{1}{n^2}\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\mu_{ij}\mu_{i'j} \right] - \rho^2\sigma^4] \quad 44$$

Evaluating $E\left[\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\mu_{ij}\mu_{i'j} \right]$ from 44,

We use the relation, $\mu_{i'j} = \rho\mu_{ij} + \sqrt{1-\rho^2}x_{i'j}$ i.e. $\mu_{i'j}$ results from the correlation of an independent variable $x_{i'j}$ with μ_{ij} (Baker & Olaleye, 2012).

$$E\left[\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\mu_{ij}\mu_{i'j} \right] = E\left[\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\mu_{ij}(\rho\mu_{ij} + \sqrt{1-\rho^2}x_{i'j}) \right]$$

$$= E\left[\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\rho\mu_{ij}^{\ 2} + \sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sqrt{1-\rho^2}\sum_{j=1}^{n}\mu_{ij}x_{i'j} \right] \quad 45$$

$$E\left[\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\rho\mu_{ij}^{\,2}+\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sqrt{1-\rho^2}\sum_{j=1}^{n}\mu_{ij}x_{i'j}\right]$$

$$=E\left[\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}(\rho\mu_{ij}+\sqrt{1-\rho^2}x_{i'j})\sum_{j=1}^{n}\rho\mu_{ij}^{\,2}+\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}(\rho\mu_{ij}+\sqrt{1-\rho^2}x_{i'j})\sqrt{1-\rho^2}\sum_{j=1}^{n}\mu_{ij}x_{i'j}\right] \quad (46)$$

$$=E\left[\begin{array}{l}\rho^2(\sum_{j=1}^{n}\mu_{ij})^2\sum_{j=1}^{n}\mu_{ij}^{\,2}+\rho\sqrt{1-\rho^2}\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}x_{i'j}\sum_{j=1}^{n}\mu_{ij}^{\,2}+\rho\sqrt{1-\rho^2}(\sum_{j=1}^{n}\mu_{ij})^2\sum_{j=1}^{n}\mu_{ij}x_{i'j}\\ +(1-\rho^2)\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}x_{i'j}\sum_{j=1}^{n}\mu_{ij}x_{i'j}\end{array}\right]$$

Since $E[\mu_{ij}\mu_{ij'}]_{j\neq j'}=0$, expanding we obtain

$$E\left[\begin{array}{l}\rho^2(\sum_{j=1}^{n}\mu_{ij})^2\sum_{j=1}^{n}\mu_{ij}^{\,2}+\rho\sqrt{1-\rho^2}\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}x_{i'j}\sum_{j=1}^{n}\mu_{ij}^{\,2}+\rho\sqrt{1-\rho^2}(\sum_{j=1}^{n}\mu_{ij})^2\sum_{j=1}^{n}\mu_{ij}x_{i'j}\\ +(1-\rho^2)\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}x_{i'j}\sum_{j=1}^{n}\mu_{ij}x_{i'j}\end{array}\right]$$

$$=E\left[\rho^2\left(\sum_{j=1}^{n}\mu_{ij}^{\,2}\right)^2+\rho\sqrt{1-\rho^2}\sum_{j=1}^{n}x_{i'j}\sum_{j=1}^{n}\mu_{ij}^{\,3}+\rho\sqrt{1-\rho^2}\sum_{j=1}^{n}\mu_{ij}^{\,3}\sum_{j=1}^{n}x_{i'j}+(1-\rho^2)\sum_{j=1}^{n}\mu_{ij}^{\,2}x_{i'j}^{\,2}\right] \quad (47)$$

Simplifying the $\left(\sum_{j=1}^{n}\mu_{ij}^{\,2}\right)^2$ term in 47 results in

$$\left(\sum_{j=1}^{n}\mu_{ij}^{\,2}\right)^2=\left(\mu_{i1}^{\,2}+\mu_{i2}^{\,2}+.....+\mu_{in}^{\,2}\right)^2=\sum_{j=1}^{n}\mu_{ij}^{\,4}+\mu_{i1}^{\,2}\sum_{j=2}^{n}\mu_{ij}^{\,2}+..+\mu_{in}^{\,2}\sum_{j=1}^{n-1}\mu_{ij}^{\,2} \quad (48)$$

$$=\sum_{j=1}^{n}\mu_{ij}^{\,4}+\sum_{j=1}^{n}\mu_{ij}^{\,2}(\sum_{k\neq j}^{n-1}\mu_{ik}^{\,2})$$

Assuming each μ_{ij} is normally distributed, from the moments of a normal distribution

$$E[(X-\mu)^p]=\begin{cases}0 & when\ p-odd\\ \sigma^p(p-1)!! & when\ p-even\end{cases}$$

Which gives $E[\mu_{ij}^{\,4}]=3\sigma^4$, $E[\mu_{ij}^{\,3}]=0$ and $E[\mu_{ij}^{\,2}]=\sigma^2$. Substituting in 47 gives

$$E\left[\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\mu_{ij}\mu_{i'j}\right]=3n\rho^2\sigma^4+n(n-1)\rho^2\sigma^4+(2\rho\sqrt{1-\rho^2}x_{ij}*0)+n(1-\rho^2)\sigma^4$$
$$=n^2\rho^2\sigma^4+n\rho^2\sigma^4+n\sigma^4 \qquad 49$$

Substituting $E\left[\sum_{j=1}^{n}\mu_{ij}\sum_{j=1}^{n}\mu_{i'j}\sum_{j=1}^{n}\mu_{ij}\mu_{i'j}\right]$ back into 44 gives

---

$$\rho_{e(\tilde{q}_{ii'}),e(q_{ii'})}=\frac{n^{1/2}}{\sigma^4(1+\rho^2)}[\frac{1}{n^2}(n^2\rho^2\sigma^4+n\rho^2\sigma^4+n\sigma^4)-\rho^2\sigma^4]=\frac{n^{1/2}}{n^2(1+\rho^2)}[n^2\rho^2+n\rho^2+n-n^2\rho^2]$$
$$=\frac{n^{1/2}}{n^2(1+\rho^2)}[n(1+\rho^2)] \qquad 50$$

This simplifies to $\rho_{e(\tilde{q}_{ii'}),e(q_{ii'})}=\frac{1}{n^{1/2}}$

51

We find that the correlation between the errors from the Methods are independent of the

correlation and standard deviation between the μ's (ρ and σ).

5.1.2 Checks And Parameterization

As a test, we evaluate the value of the correlation between the errors when the number of experts

elicited is 1. We expect the errors to be 1.

We substitute for n=1 ➔ $CorrX,Y(n=1)=\frac{1}{n^{1/2}}=1$

The test confirms this.

Parameterization with number of experts = 5

Let $n = 5$ ➔ $\rho_{e(\tilde{q}_{ii'}),e(q_{ii'})} = Corr(\tilde{q}_{ii'}, q_{ii'})(n = 5) = \frac{1}{(5^{1/2})} \approx 0.45$

Therefore, the correlation between the errors of Methods I and II when the number of experts is 5, under the normality assumption is 0.45. This shows there is significant correlation between the errors from both methods.

Figure 6 gives the parameterization of the correlation between errors for different numbers of experts used.

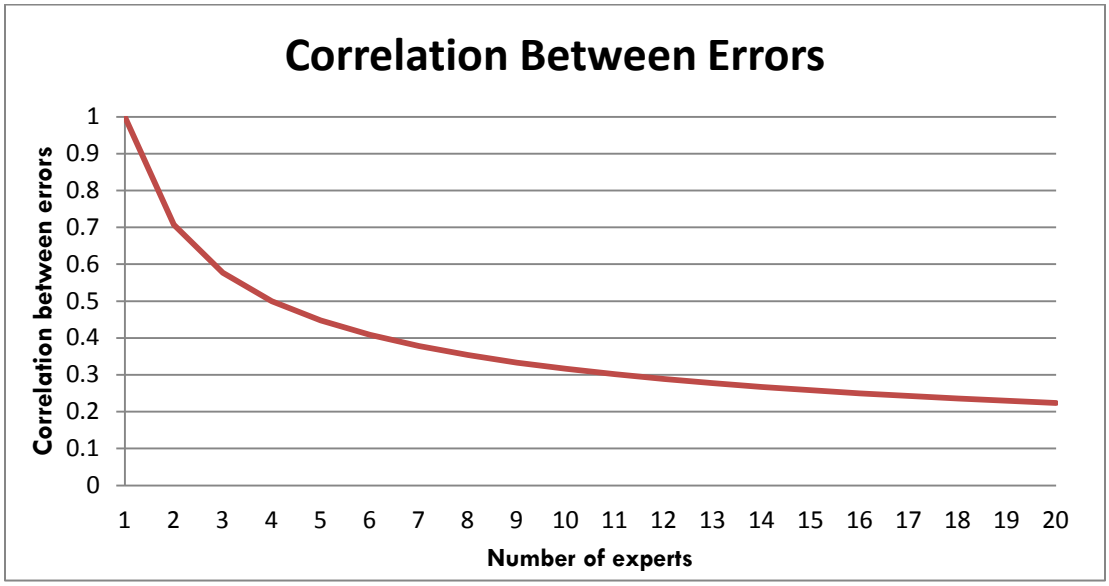


Figure 6: Correlation between the errors of the two methods as a function of the number of experts

5.2 Probability That The Error From Method I Is Lesser Under Normality Assumption

In this sub section, we calculate the probability that the error from Method I is lesser using the previously obtained correlation between the errors Equation 51.

The variance 34 results in

$$\varpi^2 = \frac{\sigma^4\left(1+\rho^2\right)}{n}\left[\frac{1+n}{n}\right] - \left(2*\sqrt{\frac{\sigma^4\left(1+\rho^2\right)}{n^2}*\frac{\sigma^4\left(1+\rho^2\right)}{n}}*\left(\frac{1}{n^{1/2}}\right)\right) \quad 52$$

This simplifies to $$\varpi = \frac{\sigma^2}{n}\sqrt{\left(1+\rho^2\right)(1+n)-2\left(1+\rho^2\right)} = \frac{\sigma^2}{n}\sqrt{\left(1+\rho^2\right)(n-1)} \quad 53$$

$$\frac{\mu}{\varpi} = (\rho\sigma^2\left[\frac{n-1}{n}\right]) / (\frac{\sigma^2}{n}\sqrt{\left(1+\rho^2\right)(n-1)}) \quad 54$$

Taking out the σ^2 and n term and dividing gives

$$\frac{\mho}{\varpi} = \frac{\rho(n-1)}{\sqrt{\left(1+\rho^2\right)(n-1)}} = \rho\sqrt{\frac{n-1}{1+\rho^2}} \quad 55$$

Substituting equation 55 in Equation 37 we find that

$$P(e(\tilde{q}_{ii'}) \le e(q_{ii'})) = \Phi\left(\frac{-\mho}{\varpi}\right) = \Phi\left[-\rho\sqrt{\frac{n-1}{1+\rho^2}}\right] \quad 56$$

Therefore $P(e(q_{ii'}) \le e(\tilde{q}_{ii'})) = 1-\Phi\left[-\rho\sqrt{\frac{n-1}{1+\rho^2}}\right]$

As we are really interested in the difference in the absolute values of both errors, we transform into

$$P(|e(q_{ii'})| \le |e(\tilde{q}_{ii'})|) = \left[P(e(q_{ii'}) \le e(\tilde{q}_{ii'}), e(q_{ii'}) > 0)\right] + \left[P(e(q_{ii'}) > e(\tilde{q}_{ii'}), e(q_{ii'}) \le 0)\right]$$

5.2.1 Numerical Evaluation Of The Probability That Method I Error Is Less

For the numerical evaluation, Figure 7 gives the Probability that the absolute error from Method I is less than that of the absolute error from Method II. We note that the result shows the absolute error from Method I is very likely to be less than that from Method II. This is in accordance with our work so far.

Correlation between μ's(N=5)	1	0.8	0.6	0.4	0.2	0	-0.2	-0.4	-0.6	-0.8	-1
$\Pr(\lvert errMetI \rvert \leq \lvert errMetII \rvert)$ **(N=5)**	0.72	0.68	0.64	0.58	0.52	0.50	0.52	0.58	0.64	0.68	0.72

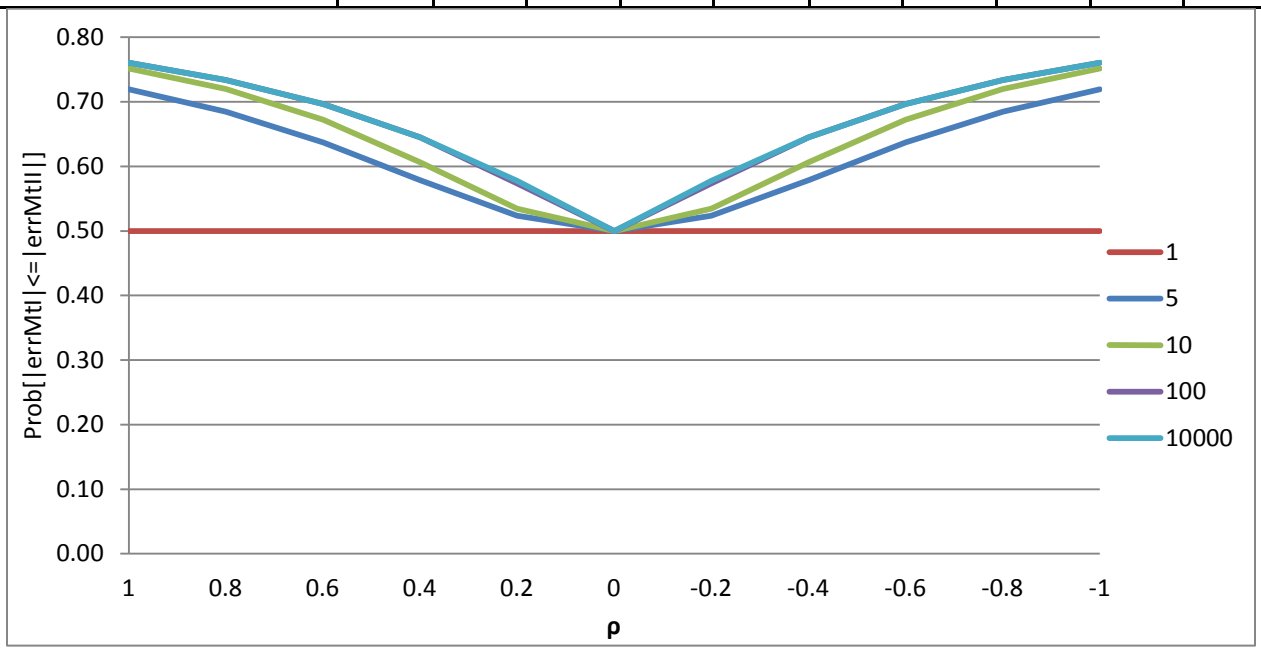


Figure 7: Probability [Absolute-Error-Method I ≤ Absolute-Error-Method II] $P(\lvert e(q_{ii'}) \rvert \leq \lvert e(\tilde{q}_{ii'}) \rvert)$ **with N experts**

5.3 Comparing With Simulation Results

The result generally conforms to the simulation results. The exception is that for the simulation results, the correlation between errors is dependent on the correlation and standard deviation between the μ's. Here, however the correlation between errors is only dependent on the number of experts. This is likely due to our assumption that the errors are from a normal distribution. We note also that when the number of experts is one or when the self-correlation (ρ) is zero, both methods perform equally.

CHAPTER 6

EMPIRICAL EXAMPLE: EXPERT AGGREGATION

In this chapter, we use actual expert elicitation data from the previous study by ((Baker, Chon, & and Keisler, 2008), (2009), (2009)) to examine if correlation in experts leads to the predicted difference when the two orders of aggregation are used. We examine the intrinsic characteristics of the experts. Individual expert's responses are initially analyzed to examine if they exhibit self-correlation in their judgments. Next we examine the level of cross-correlation among the experts. We will then predict how these patterns will impact the actual aggregated probabilities under the different methods.

We then combine the individual expert elicitations using both aggregation orders. The projections from the theory and the empirical outputs from both orders of aggregation are then compared to check if the results are as expected. We finally examine if the result from the two aggregation orders are significantly different to warrant a reexamination of the optimal portfolio using the theoretically superior aggregation method (i.e. combining early).

6.1 Expert's Characterization And Classification

While (Baker & Peng, 2012) noted the elicited experts' optimism and pessimism, we extend this study by carrying out a more elaborate study detailed below. For each of the sub-technologies considered, we classify the elicitation results in terms of Self-correlation and Cross-correlation. The tables in Appendix A give all the expert elicitation across all the hurdles for all the technologies considered in this study. Figure 2 gives an overview of the nine sub technologies that make the

three technologies while Figure 3 gives an example of how the success and funding levels compose the sub technologies using the Purely Organic Solar sub technology as an example.

To provide a clear representation of the technological hurdles, funding levels, success levels and sub technologies, we use the Nuclear technology shown in Table 7 as an example. The technological hurdles to success levels (e.g. Efficiency, Deep Burn Rate and Capital<1000) make up each of the funding success level (e.g. Low funding High Temperature High Success level). The funding levels (e.g. Low, Medium and High) compose the Sub-technologies Success levels (e.g. High Temperature High Success level). Each of the three technologies (Nuclear, CCS and Solar) are composed of three sub-technologies each Figure 2.

		NUCLEAR: High Success Probability Distribution																	
		LWR			HIGH TEMP REACTORS									FAST REACTORS					
		LOW	MED	HIGH	LOW			MED			HIGH			LOW	MED	HIGH			
					Eff	DBR	C<1000	Eff	DBR	C<1000	Eff	DBR	C<1000						
	Exp 1	0.2	0.4	0.75	0.4	0.15	0.1	0.8	0.65	0.5	0.9	0.7	0.7	0	0	0			
	Exp 2	0.05	0.2	0.4	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.005	0.02	0.05			
	Exp 3	0.1	0.25	0.35	0.3	N/A	0.1	0.6	N/A	0.2	0.8	N/A	0.2	0	0	0.1			
NUCLEAR	Exp 4	0.5	0.5	0.9	0	0.5	0.25	0.5	0.8	0.2	0.6	0.9	0.3	0	0	0.5			
		NUCLEAR: Low Success Probability Distribution																	
		LWR			HIGH TEMP REACTORS									FAST REACTORS					
		LOW	MED	HIGH	LOW			MED			HIGH			LOW		MED		HIGH	
					Eff	DBR	1000<C<1500	Eff	DBR	1000<C<1500	Eff	DBR	1000<C<1500						
	Exp 1	0.2	0.4	0.75	0.4	0.15	0.3	0.8	0.65	0.2	0.9	0.7	0.15	0	0	0	0	0	0
	Exp 2	0.05	0.2	0.4	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.005	0.005	0.02	0.8	0.05	0.9
	Exp 3	0.1	0.25	0.35	0.3	N/A	0.15	0.6	N/A	0.3	0.8	N/A	0.5	0	0.3	0	0.4	0.1	0.6
	Exp 4	0.5	0.5	0.9	0	0.5	0.25	0.5	0.8	0.2	0.6	0.9	0.2	0	0	0	0.1	0.5	0.9

Table 7: Nuclear Elicitation Hurdle

To examine each of the technologies for self-correlation, we classify the technologies according to the number of highly self correlated experts. We also discuss the sub technologies under these technologies where the experts do not follow the trend observed at the technology level. We do the same in examining for cross-correlation. We will discuss each technology (Solar, CCS then Nuclear) in turn, since there is a different group of experts for each technology.

6.1.1 Solar Photovoltaic

For Solar Photovoltaic, three experts were considered across three or four technological hurdles for each of the funding success levels which make up the sub-technologies. The base elicitation probabilities used are from (Baker, Chon, & Keisler, 2009). We first present the self-correlation characteristics then discuss the Cross-correlations in the same pattern.

6.1.1.A Correlation Within Experts

Here we examine the solar experts for self-correlation using the dispersion from the sub-technology hurdle mean. By measuring the dispersion of the individual expert probability from the hurdle mean probability (mean of the experts at each sub-technological hurdle), we can readily examine if the expert is generally more biased compared to the other experts at the same sub-technology hurdle. If an expert's probabilities are always consistently less than the hurdle mean, this implies that the expert is generally pessimistic compared to the other experts. Also always being above the mean implies optimism. An expert that shows any of these two characteristics is classified as highly self correlated.

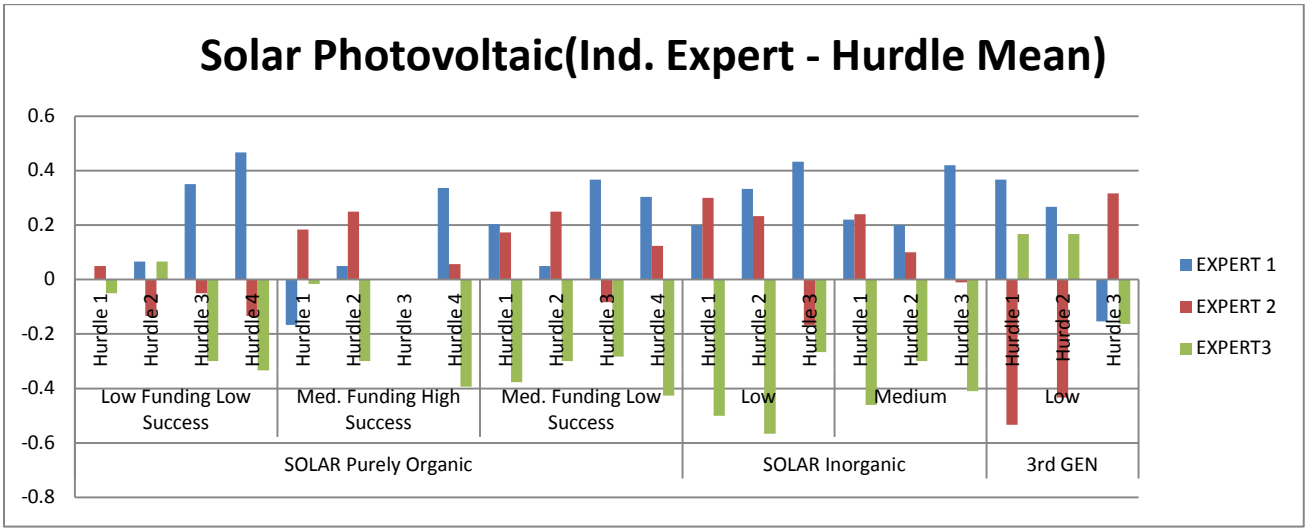


Figure 8: Difference of Individual expert opinion from Hurdle mean-SOLAR

TECHNOLOGY LEVEL (SOLAR PHOTO-VOLTAIC)

To classify the experts at the Solar technology level, we rely on the individual dispersion from the hurdle mean (assumed true probability) Figure 8. We observe that expert 1 is generally very optimistic, been below the hurdle mean in only 2 of the 21 hurdles that constitute this technology level. Expert 3 is also very pessimistic across the Solar technology level, seen by been above the hurdle mean in only 3 of the 21 hurdles. Unlike the other two experts, expert 2 does not show such strong bias been below the mean in 8 of the 21 hurdles and also pessimistic in 2 of the 6 funding success levels, resulting in slight optimism. In summary, we classify the solar technology level as consisting of two very highly self correlated experts and one not highly self correlated expert.

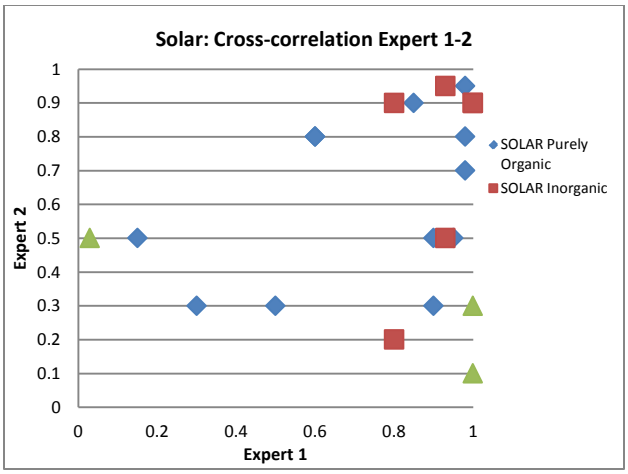
These patterns are consistent for all the sub technologies that compose the Solar technology except the 3rd Gen sub technology were the experts did not show any consistent pattern.

6.1.1.B Correlation Across Experts

In this section, we examine the Solar experts for cross-correlation. To measure this effect we rely on two measures; graphical charts showing correlation of an expert to another [Figure 9] and the Pearson product moment correlation coefficient [57] which is given by [Table 8].

$$Correl(X,Y) = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2 \sum (y-\bar{y})^2}}$$

57



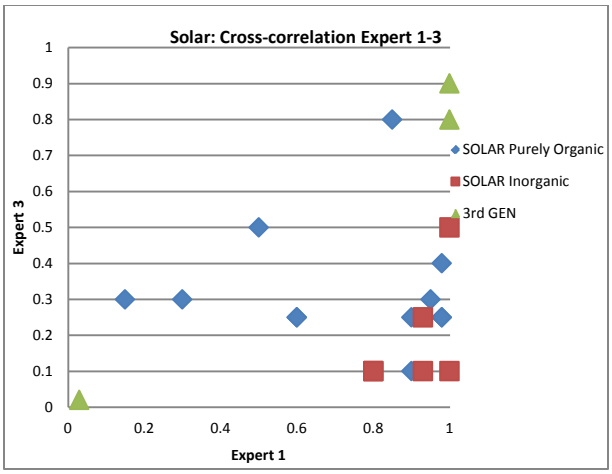


Figure 9: Cross-correlation Expert (1 & 2) Solar **Figure 10: Cross-correlation Expert (1 & 3) Solar**

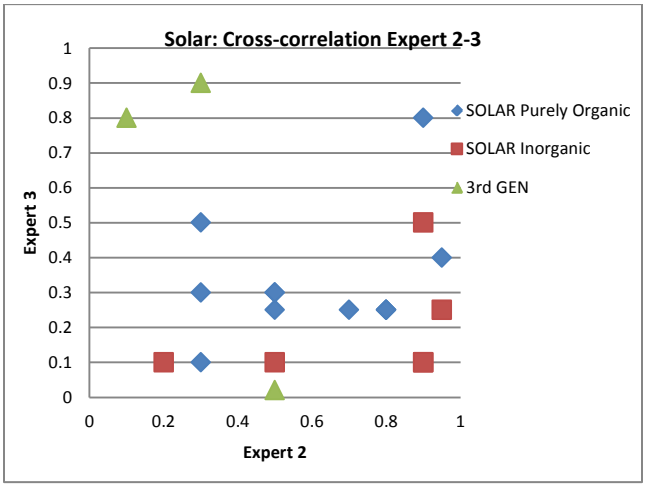


Figure 11: Cross-correlation between Expert (2 and 3) Solar

	SOLAR TECHNOLOGY
Expert 1 – 2	0.23
Expert 1 – 3	0.26
Expert 2 – 3	-0.16

Table 8: Correlation Coefficient Solar

TECHNOLOGY LEVEL – SOLAR

A look at the correlation charts (Figure 9 -Figure 11) and Table 8shows that the experts are very independent across the technology and do not show any significant level of cross-correlation.

6.1.2 Carbon Capture And Storage

Three sub-technologies were considered under CCS; Pre-Combustion, Chemical Looping and Post Combustion. Three experts (experts' 1, 2 and 3) were elicited across 3 technological hurdles in Pre-Combustion, 2 experts (experts' 1 and 3) across 5 hurdles in Chemical Looping and 4 experts across 4 hurdles in Post Combustion.

6.1.2.A Correlation Within Experts (Self-Correlation)

Using the dispersion from the hurdle mean as in the solar evaluation, we examine the experts for self-correlation Figure 12: Dispersion from the mean - CCS. Due to the experts varying area of expertise, only two experts completed all the technological hurdles for CCS.

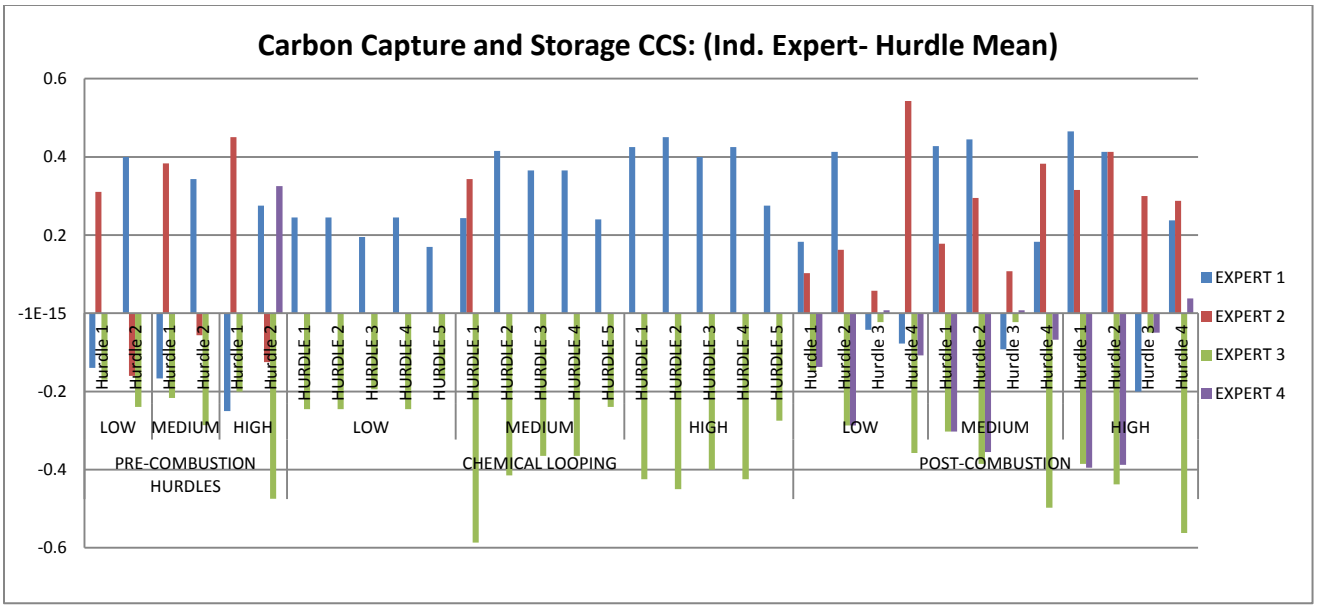


Figure 12: Dispersion from the mean - CCS

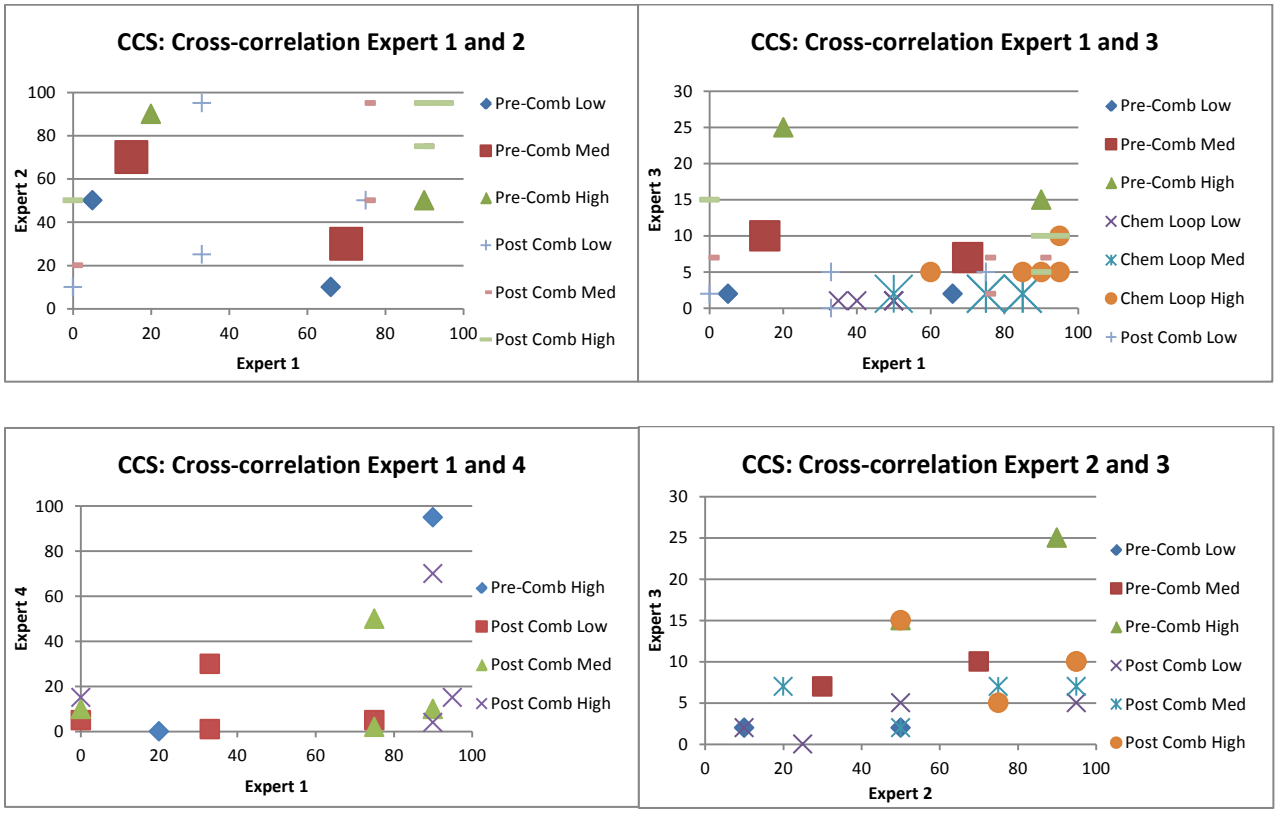
In CCS, only Experts' 1 and 3 answered all the elicitation hurdles. Expert 4 only responded for the Post Combustion sub-technology. While Expert 2 provided estimates for the Pre Combustion and Post Combustion sub-technologies, and for only one technological hurdle in the Chemical Looping sub technology.

Expert 1 was generally optimistic across the technology (consistently above the hurdle mean) while Expert 3 was pessimistic for all the hurdles considered in CCS. Expert 2 is generally an optimist while Expert 4 is pessimist.

Aberrations: The experts showed the characteristics described above for the all the sub technologies (Chemical Looping, Post Combustion and Pre Combustion). Therefore no aberrations were noted at the sub technology level.

6.1.2.B Correlation Across Experts

To examine for cross-correlation across the experts, we utilize graphical charts showing correlation of an expert to another [Figure 13(a-f): Cross-correlation within experts CCS] and the Pearson product moment correlation coefficient given by [Table 9]. As not all the experts provided responses for all the technology hurdles, we run into a problem of small statistical size for these experts. Only for Experts 1 and 3 do we not experience this limitation.



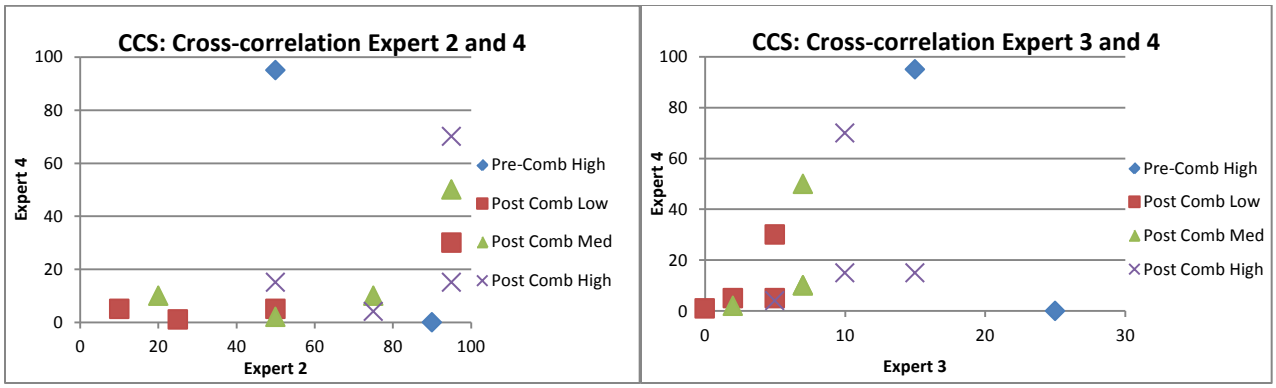


Figure 13(a-f): Cross-correlation within experts CCS

	SUB-TECHNOLOGY LEVEL			CCS
	PRE-COMBUSTION	CHEMICAL LOOPING	POST-COMBUSTION	
Expert 1 - 2	-0.59		0.70	
Expert 1 - 3	-0.05	0.68	0.03	-0.06
Expert 1 - 4			0.26	
Expert 2 - 3	0.79		0.42	
Expert 2 - 4			0.64	
Expert 3 - 4			0.43	

Table 9: Cross-correlation CCS

For the two experts who elicit all the technological hurdles in CCS (Experts 1 and 3), no significant cross-correlation can be noticed between them (-0.06). While the cross-correlation between all other experts seems to range low to moderate, we do not classify this due to low statistical size.

6.1.3 Nuclear Technology

Three sub-technologies were evaluated under Nuclear; Light Water Reactors, High Temperature Reactors and Fast Reactors. As Light Water Reactors and Fast Reactors are composed of only one

technological hurdle, we do not analyze these as no resulting change will occur due to self or cross-correlations of experts. The sub-technology we are then concerned with, High Temperature Reactor, was accessed at two success levels (high and low). Three experts (1, 3 and 4) were used in the elicitation of the funding levels in this sub-technology, while both experts 1 and 4 answered all elicitations under High Temperature Reactor, expert 3 did not respond to the Deep Burn Rate hurdle leading to some limitations in consideration of the results.

6.1.3.A CORRELATION WITHIN EXPERTS

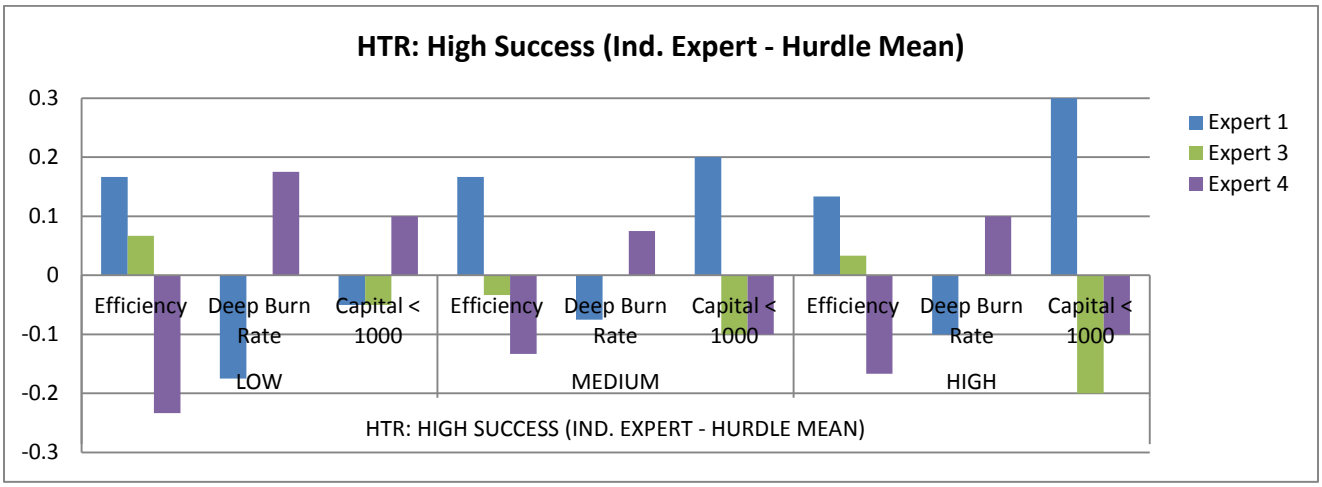


Figure 14: Dispersion from the hurdle mean; High Success: High Temperature Reactor Nuclear

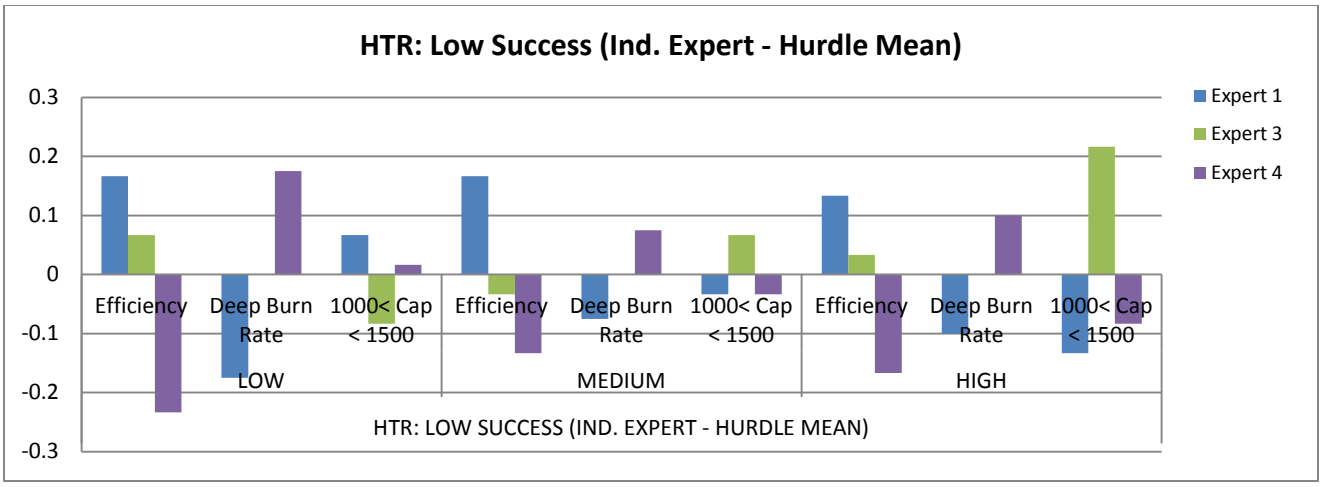


Figure 15: Dispersion from the hurdle mean; Low Success: High Temperature Reactor Nuclear

Examination of Figure 14 and Figure 15 shows no particular optimism or pessimism among any individual expert in Nuclear. Hence we don't notice high self-correlation within any of the experts.

6.1.3.B Correlation Across Experts (Cross-Correlation)

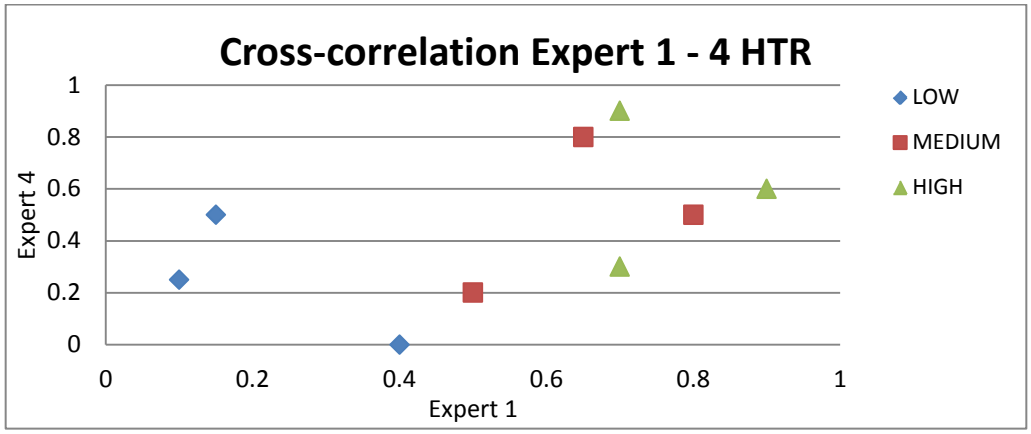


Figure 16: Cross-correlation between Expert 1 and 4 HTR

In this section, we examine the experts for cross-correlation. Of the three experts who provided elicitations, only Experts 1 and 4 responded across all the technological hurdles for High

Temperature Reactor. Examination of Figure 16 shows a limited to moderate cross-correlation. A statistical test for cross-correlation gives 0.58.

6.1.4 Summary

The table (Table 10) below summarizes the assessed self-correlation noted within individual experts for all the sub-technologies and technologies elicited in the study. The shaded cells indicate sub-technologies with a majority of highly self correlated experts.

Tech	Sub-Tech	Experts	High Self Correlated Experts	Specific Expert
SOLAR	Organic	3	2	1 and 3
	In-Organic	3	3	1, 2 and 3
	3rd Gen	3	0	
CCS	Pre-Comb	3	1	3
	Chem. Loop	2	2	1 and 3
	Post-Comb	4	3	2, 3 and 4
NUCLEAR	LWR	N/A	N/A	
	HTR	3	0	
	FR	N/A	N/A	

Table 10: Summary of assessed Self-correlation

6.2 Rank Independence Classification Of Experts

In addition to the previous classification of experts (6.1), we examine the different technologies (Solar, CCS and Nuclear) for high self-correlation using the concept of rank independence (Cooke M. R., 1986). We use a hypothesis test to examine the experts in this section.

6.2.1 Rank Independence Definition

When a group of experts provide an estimate of a quantity, we can rank them from highest to lowest. An expert would be considered “rank independent’ if the ordinal rank of the expert is random and the expert does not have a particular tendency to be in any particular position about

the median rank. On the other hand, an expert is *not* rank independent if they have a tendency to always be toward the top of the ranking, or toward the bottom of the ranking.

6.2.2 Hypothesis Test For Rank Independence

The aim of this hypothesis test is to examine if the individual experts exhibit rank independence. From the analyzed data, we expect most of the CCS and Solar experts not to show rank independence (because they tend to be optimistic or pessimistic) and the Nuclear experts to be rank independent.

- Null Hypothesis: The null hypothesis is that the experts are independent in rank.
- Alternate Hypothesis: Is that the experts are not independent in rank.

We assume the elicited probabilities are from a normal distribution. We observe the same procedure as in (Cooke M. R., 1986). The elicited probabilities are ranked in ascending order; the median rank is then obtained for each of the sets of elicited probabilities (where a set represents a particular technological hurdle).We then compare the individual experts' rank orderings with the base population ordering i.e. we determine whether the experts are consistently above or below the median rank. Hence, we find the probability that the Null hypothesis (expert elicitations are rank independent) can be refuted at a particular significance level.

6.2.2.A Hypothesis

Hypothesis: The aim of the hypothesis is to determine if individual experts have a tendency to always be above the median rank (optimistic) or to always be below the median rank (pessimistic) of the elicited probabilities. Here we provide some definitions.

X^{i}_{jk} is a binary indicator of the elicited probability of an expert i for the j technological hurdle elicited, at the k funding level. We assume there are "I" experts, "J" technological hurdles and "K" funding levels. Note that the number of technological hurdles might actually differ from the number of hurdles elicited. This is because experts, due to expertise and other personal reasons, might choose not to provide success probabilities for all the technological hurdles.

X^{i}_{jk} = 1 if $X^{i}_{jk} > Median(X^{1}_{jk}, X^{2}_{jk}....X^{I}_{jk})$ and 0 otherwise. (Note that the technological hurdle elicited j and the funding level k are the same for the experts i).

- μ_k Population mean: $\mu_k = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} X^{i}_{jk}}{N_k}$. Where N_k is the number of X^{i}_{jk}'s examined: it is $N_k \leq (I * J)$ depending on whether some experts do not respond to all the elicitations. Thus, μ_k is the mean of the X^{i}_{jk}'s for each funding level across all the experts i and all hurdles j. It is the percentage of the probabilities that are greater than the median at that investment level. Given three experts, we would expect this to be about 33%. However, it might not be 33% due to ties at the median level (Cooke M. R., 1986).
- σ_k: The population standard deviation: $\sigma_k = \sqrt{\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\left(X^{i}_{jk} - \mu_k\right)^2}{N_k - 1}}$. σ_k is the standard deviation of the X^{i}_{jk} for each k funding level.
- These two parameters, μ_k and σ_k, define the base population.

- $\tilde{x}_{ik}$: Sample mean: $\tilde{x}_{ik} = \frac{\sum_{j=1}^{J} X^{i}_{jk}}{n_{ik}}$. n_{ik} is the number of the probabilities elicited by expert i at the k funding level. Note that $n_{ik} \leq J$ as an expert might choose not to answer all elicitations. $\tilde{x}_{ik}$ gives the percentage of times that an expert is above the median, over all of the technological hurdles elicited. It is the mean of the X^{i}_{jk} 's for expert i at funding level k. For example, if $\tilde{x}_{ik} = 1$ this means expert i is very optimistic and always gives higher success probabilities than the other experts.
- n_{ik} : Sample size. As defined above, this is the number of the probabilities elicited by expert i at the k funding level i.e. the number of technological hurdles j elicited by an expert i at funding level k. Again, we note that $n_{ik} \leq J$ as an expert might choose not to answer all elicitations.
- s_{ik} : The sample standard deviation: $s_{ik} = \sqrt{\frac{\sum_{j=1}^{o} \left(X^{i}_{jk} - \tilde{x}_{ik} \right)^2}{n_{ik} - 1}}$. s_{ik} is the standard deviation of the X^{i}_{jk} for each expert i at each k funding level.

6.2.2.B Hypothesis Parameters

- Parameter of interest: $\tilde{x}_{ik}$. We are interested in determining how each individual expert's propensity to be above the median $(\tilde{x}_{ik})$ deviates from the population mean (μ_k). Ideally this should be $\begin{matrix} \tilde{x}_{ik} = 0.5 if\,(I = even) \\ \tilde{x}_{ik} \to 0.5 if\,(I = odd) \end{matrix}$ (i.e. about 33% for 3 experts or 50% for 4 experts);but

non-independence of the experts results in ties at the median leading to different values of $\tilde{x}_{ik}$ (Cooke M. R., 1986).

- Null Hypothesis: $H_0 : \tilde{x}_{ik} = \mu_k$. The null hypothesis is that mean of the X^{i}_{jk} for expert i is the same as that for all the experts for a funding level k .
- Alternative Hypothesis: $H_a : \tilde{x}_{ik} \neq \mu_k$.
- Due to the small sample size $n_{ik} < 30$, we use the One-Sample T-Test.

$$t_{ik} = \frac{(\tilde{x}_{ik} - \mu_k)}{(s_k / \sqrt{n_{ik}})}$$ (Devore, 2004) 58

$n_{ik} - 1$ gives the degree of freedom df .

- Null Hypothesis Rejection Region: Using a 10% significance level, the definition of the alternative hypothesis H_a implies that we require a two tailed test with rejection region either $t \geq t_{0.1}$ or $t \leq t_{0.1}$ (Devore, 2004).

6.2.3 Implications

We fail to reject the Null hypothesis $\tilde{x}_{ik} = \mu_k$ if the resulting one sample test value is within the 10% significance level $t_{-0.1}$. This occurs when the absolute value of the calculated T-test statistic is less than or equal to that of the 10% significance level, $t_{ik} \leq t_{0.1}$ and $t_{ik} \geq t_{-0.1}$. In this case, we cannot reject Null Hypothesis and the expert is rank independent. This means that the individual expert's propensity to be above the median elicitations is not significantly different from the group of expert's propensity to do the same. Therefore the probabilities given by the expert are relatively independent for that funding level i.e. the expert does not have any tendency to be very optimistic or pessimistic.

We reject the null hypothesis if the calculated one sample T-test is outside the 10% significance level $t_{-0.1}$. This occurs when the absolute value of the calculated T-Test statistic is greater than that of the 10% significance level, $(t_{ik} > t_{0.1}) or (t_{ik} < t_{-0.1})$. This implies that we reject the Null Hypothesis and the expert is not rank independent. This means that the individual expert's propensity to be above the median elicitation is significantly different from the group of experts'. This implies that the probabilities given by the expert for the funding level are significantly dependent likely due to the expert's optimism or pessimism.

6.2.3.1 Considerations

In this sub section, we note that we can only consider 3 funding success levels combinations per technology as a maximum. This is despite that there are 3 sub-technologies per the 3 funding levels in a technology, therefore there are numerically 9 funding success levels combination per technology. This is because the success probabilities for the three funding levels of a sub-technology should be correlated. This implies that we can only consider 3 funding success levels per sub technology as a maximum.

6.2.4 Methodology: Solar Low Funding Low Success Sub-Technology As An Example

We go through our methodology using the low funding low endpoint Solar Organic sub-technology as an example Table 11. The Number of experts I =3 and K =1.

1. We obtain the X^{i}_{jk} 's for the $k = 1$ funding level by ranking the elicited probabilities.

As discussed in 6.2.2.A Hypothesis, X^{i}_{jk} takes the value 1 when if the probability is above the median for the hurdle. Table 11 gives an example for the Solar low success low funding Level. We do this for all the sub-technologies at the different funding levels.

FUNDING LEVEL	LOW FUNDING ($15M) LOW ENDPOINT (EFF. 15%)								
	EX 1	**EX 2**	**EX 3**				**EX 1**	**EX 2**	**EX 3**
	(%)	**(%)**	**(%)**		Median		X_{11}^{1}	X_{11}^{2}	X_{11}^{3}
Probability of Efficiency	0.85	0.9	0.8		0.85		0	1	0
Probability of Stability	0.5	0.3	0.5		0.5		0	0	0
Probability of deposition cost	0.9	0.5	0.25		0.5		1	0	0
Probability of Indium Substitute	0.9	0.3	0.1		0.3		1	0	0
Average Probability	**0.34425**	**0.0405**	**0.01**						

Table 11: Solar Low Success Low Funding sub-technology Rank Independence Example

2. To obtain the population mean μ_k, the $\mu_k = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} X_{jk}^{i}}{N_k}$, of the all the elicited probabilities for a funding level k (not just one sub-technology shown in Table 11) is obtained. Table 12 gives the X_{jk}^{i}**'s** and the t-test statistic for the Solar low funding level. μ_k is the percentage of times the experts is above the median. In the case shown below, the population mean is 9/30=0.3. The $\tilde{x}_{ik}$, s_{ik}, n_{ik} for each expert i is also obtained.

	Expert 1 X_{j1}^{1}'s	Expert 2 X_{j1}^{2}'s	Expert 3 X_{j1}^{3}'s
Solar Organic Low	**0**	**1**	**0**
	0	**0**	**0**
	1	**0**	**0**
	1	**0**	**0**
Solar Inorganic Low	**0**	**1**	**0**
	1	**0**	**0**
	1	**0**	**0**
3rd Gen	**1**	**0**	**0**
	1	**0**	**0**
	0	**1**	**0**
Sample mean $\tilde{x}_{ik}$	$\tilde{x}_{1k}=\frac{6}{10}=0.6$	$\tilde{x}_{2k}=\frac{3}{10}=0.3$	$\tilde{x}_{3k}=\frac{0}{10}=0$
Sample Standard Deviation s_{ik}	0.52	0.48	0
Sample size n_{ik}	$n_{1k}=10$	$n_{2k}=10$	$n_{3k}=10$
Population Mean	**Total cells** N_k **=30,** $\mu_k=\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}X_{jk}^{i}}{N_k}$ **, Pop. Mean** μ_k **=9/30=0.3**		

Table 12: Solar Low Success Low Funding Methodology Parameterization Example

3. We then calculate the T-statistic for expert i (t_{ik}), Equation . This is compared 58

with the $(t_{ik}(0.1))$ 10% significant T-test value from a statistic table using the degree of freedom $df=n_{ik}-1$. If the absolute value of t_{ik} is greater than $t_{ik}(0.1)$ then the hypothesis is rejected and

the probabilities from the expert adjudged to be not independent. If the $t_{ik} \le t_{ik}(0.1)$ then we fail to reject the hypothesis and the probabilities from expert i are relatively independent.

6.2.5 Results From The Rank Independence Classification

The results from the hypothesis tests are listed below

6.2.6 CCS

Low funding level $(k=1)$	Expert 1 $(j=1)$	Expert 2 $(j=2)$	Expert 3 $(j=3)$	Expert 4 $(j=4)$
Population Mean μ_l	$\mu_1 = 0.47$			
Sample size n_{i1}	$n_{11} = 11$	$n_{21} = 6$	$n_{31} = 11$	$n_{41} = 4$
Sample mean $\tilde{x}_{i1}$	$\tilde{x}_{11} = 0.82$	$\tilde{x}_{21} = 0.83$	$\tilde{x}_{31} = 0$	$\tilde{x}_{41} = 0.25$
Sample Stand Deviation s_{i1}	$s_{11} = 0.405$	$s_{21} = 0.408$	$s_{31} = 0$	$s_{41} = 0.5$
One Sample T-Test (t_{i1})	$t_{11} = 2.865$	$t_{21} = 2.187$	$t_{31} = -\infty$	$t_{41} = -0.875$
$t_{i1}(0.1)$	$t^*_{11}(0.1) = 1.812$	$t^*_{21}(0.1) = 2.015$	$t^*_{31}(0.1) = 1.812$	$t^*_{41}(0.1) = 2.353$
Result	Non Ind.	Non Ind.	Non Ind.	Ind.

Table 13: Rank Independence; CCS Low Funding

KEY: Ind.➔Independent Non Ind.➔ Not Independent

Note: Some experts choose not to access some of the hurdles hence the different degrees of freedom $df = n_{ik} - 1$ for different experts. This may lead to different values of $t_{i1}(0.1)$.

Medium funding level $(k=2)$	Expert 1 $(j=1)$	Expert 2 $(j=2)$	Expert 3 $(j=3)$	Expert 4 $(j=4)$
Population Mean μ_k	$\mu_2 = 0.45$			
Sample size n_{i2}	$n_{12} = 11$	$n_{22} = 7$	$n_{32} = 11$	$n_{42} = 4$
Sample mean $\tilde{x}_{i2}$	$\tilde{x}_{12} = 0.73$	$\tilde{x}_{22} = 0.86$	$\tilde{x}_{32} = 0$	$\tilde{x}_{42} = 0.25$
Sample Stand. Deviation s_{i2}	$s_{12} = 0.467$	$s_{22} = 0.378$	$s_{32} = 0$	$s_{42} = 0.5$
One Sample T-Test (t_{i2})	$t_{12} = 1.936$	$t_{22} = 2.818$	$t_{32} = -\infty$	$t_{42} = -0.818$
$t_{i2}(0.1)$	$t^*_{12}(0.1) = 1.812$	$t^*_{22}(0.1) = 1.943$	$t^*_{32}(0.1) = 1.812$	$t^*_{42}(0.1) = 2.353$
Result	Non Ind.	Non Ind.	Non Ind.	Ind.

Table 14: Rank Independence; CCS Medium Funding

High funding level ($k=3$)	Expert 1 ($j=1$)	Expert 2 ($j=2$)	Expert 3 ($j=3$)	Expert 4 ($j=4$)
Population Mean μ_k	$\mu_3 = 0.44$			
Sample size n_{i3}	$n_{13} = 11$	$n_{23} = 6$	$n_{33} = 11$	$n_{43} = 4$
Sample mean $\tilde{x}_{i3}$	$\tilde{x}_{13} = 0.82$	$\tilde{x}_{23} = 0.83$	$\tilde{x}_{33} = 0$	$\tilde{x}_{43} = 0$
Sample Stand Deviation s_{i3}	$s_{13} = 0.404$	$s_{23} = 0.408$	$s_{33} = 0$	$s_{43} = 0$
One Sample T-Test (t_{i3})	$t_{13} = 3.1212$	$t_{23} = 2.375$	$t_{33} = -\infty$	$t_{43} = -\infty$
$t_{i3}(0.1)$	$t^*_{13}(0.1) = 1.812$	$t^*_{23}(0.1) = 2.015$	$t^*_{33}(0.1) = 1.812$	$t^*_{43}(0.1) = 2.353$
Result	Non Ind.	Non Ind.	Non Ind.	Non Ind.

Table 15: Rank Independence; CCS High Funding

Here, the Null hypothesis (Independence) can be refuted for Experts1, 2 and 3 in all 3 funding levels. While the Null Hypothesis is rejected in the High funding level for the 4th expert. Hence CCS experts show a high degree of non-independence which results from the inherent high self-correlation due to optimism and pessimism of the experts (3 out of 4 experts show some degree of self-correlation). We note that for the 4th expert the degree of freedom ($df = n_{1k} - 1 = 3$) is also very small. Thus we expect significant changes when Method I and Method II are used for the CCS elicitations.

6.2.7 NUCLEAR

Low funding level $(k=1)$	Expert 1 $(j=1)$	Expert 2 $(j=2)$	Expert 3 $(j=3)$	Expert 4 $(j=4)$
Population Mean μ_k	$\mu_1 = 0.34$			
Sample size n_{i1}	$n_{11} = 12$	$n_{12} = 6$	$n_{13} = 11$	$n_{14} = 12$
Sample mean $\tilde{x}_{i1}$	$\tilde{x}_{11} = 0.25$	$\tilde{x}_{12} = 0.83$	$\tilde{x}_{13} = 0.09$	$\tilde{x}_{14} = 0.42$
Sample Stand. Deviation s_{i1}	$s_{11} = 0.452$	$s_{21} = 0.408$	$s_{31} = 0.301$	$s_{41} = 0.515$
One Sample T-Test (t_{i1})	$t_{11} = -0.701$	$t_{21} = 2.951$	$t_{31} = -2.756$	$t_{41} = 0.506$
$t_{i1}(0.1)$	$t^*_{11}(0.1) = 1.796$	$t^*_{21}(0.1) = 2.015$	$t^*_{31}(0.1) = 1.812$	$t^*_{41}(0.1) = 1.796$
Result	Ind.	Non Ind.	Non Ind.	Ind.

Table 16: Rank Independence; Nuclear Low Funding

Medium funding level $(k=2)$	Expert 1 $(j=1)$	Expert 2 $(j=2)$	Expert 3 $(j=3)$	Expert 4 $(j=4)$
Population Mean μ_k	$\mu_2 = 0.37$			
Sample size n_{i2}	$n_{12} = 12$	$n_{22} = 6$	$n_{32} = 11$	$n_{42} = 12$
Sample mean $\tilde{x}_{i2}$	$\tilde{x}_{12} = 0.25$	$\tilde{x}_{22} = 0.83$	$\tilde{x}_{32} = 0.18$	$\tilde{x}_{42} = 0.42$
Sample Stand. Deviation s_{i2}	$s_{12} = 0.452$	$s_{22} = 0.408$	$s_{32} = 0.404$	$s_{42} = 0.515$
One Sample T-Test (t_{i2})	$t_{12} = -0.887$	$t_{22} = 2.805$	$t_{32} = -1.509$	$t_{42} = 0.342$
$t_{i2}(0.1)$	$t^*_{12}(0.1) = 1.796$	$t^*_{22}(0.1) = 2.015$	$t^*_{32}(0.1) = 1.812$	$t^*_{42}(0.1) = 1.796$
Result	Ind.	Non Ind.	Ind.	Ind.

Table 17: Rank Independence; Nuclear Medium Funding

High funding level $(k=3)$	Expert 1 $(j=1)$	Expert 2 $(j=2)$	Expert 3 $(j=3)$	Expert 4 $(j=4)$
Population Mean μ_k	$\mu_3 = 0.41$			
Sample size n_{i3}	$n_{13} = 12$	$n_{23} = 6$	$n_{33} = 11$	$n_{43} = 12$
Sample mean $\tilde{x}_{i3}$	$\tilde{x}_{13} = 0.33$	$\tilde{x}_{23} = 0.50$	$\tilde{x}_{33} = 0.18$	$\tilde{x}_{43} = 0.67$
Sample Stand. Deviation s_{i3}	$s_{13} = 0.492$	$s_{23} = 0.548$	$s_{33} = 0.404$	$s_{43} = 0.492$
One Sample T-Test (t_{i3})	$t_{13} = -0.572$	$t_{23} = 0.382$	$t_{33} = -1.909$	$t_{43} = 1.77$
$t_{i3}(0.1)$	$t^*_{13}(0.1) = 1.796$	$t^*_{23}(0.1) = 2.015$	$t^*_{33}(0.1) = 1.812$	$t^*_{43}(0.1) = 1.796$
Result	Ind.	Ind.	Non Ind.	Ind.

Table 18: Rank Independence; Nuclear High Funding

For the nuclear technology, we fail to reject the Null hypothesis (Independence) for all the funding levels for experts 1 and 4, this implies that the probabilities from the two experts are independent. For experts 2 and 3, we reject the Null hypothesis for 2 of the 3 funding levels, hence the probabilities from these experts are not independent due to the self-correlation within the experts. We note however that the statistical significance (degree of freedom $df = n_{1k} - 1 = 5$) for expert 2 is small.

6.2.8 SOLAR

The results from the hypothesis test for the Solar sub-technologies are given below

Low Funding [A] $(k=1)$	Expert 1 $(j=1)$	Expert 2 $(j=2)$	Expert 3 $(j=3)$
Population Mean μ_k	$\mu_1 = 0.303$		
Sample size n_{i1}	$n_{11} = 10$	$n_{12} = 10$	$n_{13} = 10$
Sample mean $\tilde{x}_{i1}$	$\tilde{x}_{11} = 0.44$	$\tilde{x}_{12} = 0.23$	$\tilde{x}_{13} = 0$
Sample Standard Deviation s_{i1}	$s_{11} = 0.516$	$s_{21} = 0.483$	$s_{31} = 0$
One Sample T-Test (t_{i1})	$t_{11} = 0.823$	$t_{21} = -0.473$	$t_{31} = -\infty$
$t_{i1}(0.1)$	$t^*_{11}(0.1) = 1.833$	$t^*_{21}(0.1) = 1.833$	$t^*_{31}(0.1) = 1.833$
Result	Ind.	Ind.	Non Ind.

Table 19: Rank Independence; Solar Low Funding

Medium Funding [B] $(k=2)$	Expert 1 $(j=1)$	Expert 2 $(j=2)$	Expert 3 $(j=3)$
Population Mean μ_k	$\mu_2 = 0.3$		
Sample size n_{i2}	$n_{12} = 11$	$n_{22} = 11$	$n_{32} = 11$
Sample mean $\tilde{x}_{i2}$	$\tilde{x}_{12} = 0.55$	$\tilde{x}_{22} = 0.36$	$\tilde{x}_{32} = 0\,0\%$
Sample Standard Deviation s_{i2}	$s_{12} = 0.522$	$s_{22} = 0.505$	$s_{32} = 0$
One Sample T-Test (t_{i2})	$t_{12} = 1.559$	$t_{22} = 0.418$	$t_{32} = -\infty$
$t_{i2}(0.1)$	$t^*_{12}(0.1) = 1.812$	$t^*_{22}(0.1) = 1.812$	$t^*_{32}(0.1) = 1.812$
Result	Ind.	Ind.	Non Ind.

Table 20: Rank Independence; Solar Medium Funding

For the Solar technology, we reject the Null independence hypothesis for Expert 3 across all the funding levels and fail to reject the hypothesis for experts 1 and 2. Low funding Level ^{A} consists of low funding low success organic level and the two others low funding levels for In-Organic and 3rd Gen. Medium funding Level ^{B} consists of the high success medium funding organic level, the low success medium funding organic level and the Medium funding In-Organic level.

6.2.9 Conclusion

The results are generally in accordance with the previous classification Table 10. We show that three experts (1, 2 and 3) are not rank independent in CCS i.e. all the three individual experts exhibit a high degree of self-correlation. For Solar technology, we find that one of the experts, Expert 3, also exhibits high self correlation. The experts in Nuclear are shown to be relatively independent in their estimations. Thus, in accordance with the expert characterization, chapter 6.1, we expect very significant changes in CCS when Method I is used instead of Method II and moderate changes in the Solar and Nuclear sub-technologies.

6.3 Expectations Based On Baker & Olaleye 2012 Theory

- Expectations based on expert characteristics

In this section we predict the change in the end probabilities based on the inherent self-correlation of the experts when either method of aggregation is used. We expect from the theory that since Method II (combining experts later) does not correct for the self-correlation of the experts, then in most cases where there exists high self-correlation within individual experts, the resulting probability from Method II should be larger than that from Method I (Table 2). This is because the

mean of the expected error and the variance of the correlation errors are larger by a factor n (number of experts) in Method II.

From Table 21, the shaded cells indicate the probabilities aggregated using Method I. We expect that probabilities from Method II (combining experts later) will be higher than those from Method I (earlier combination) in the Organic, In-Organic, Chemical Looping and Post Combustion sub-technologies. We also expect the same effect in the Medium and High Funding High Success HTR and the High Funding Low Success HTR hurdle levels.

6.4 Data: Recombination Of Probabilities Using Method I

The elicitations are recombined using Method I and the results shown below (Table 21). The shaded cells Method I probabilities while the characters in bold represent the numerically higher of the two methods.

6.4.1 New Probabilities

	Carbon Capture and Storage						**Nuclear High Success**		**Nuclear Low Success**		**Solar High Success**						**Solar Low Success**	
	Pre-Comb		**Chem Loop**		**Post-Comb**		**HTR**		**HTR**		**Organic**		**In-Organic**		**3rd Gen**		**Organic**	
	Mt I	**Mt II**	**Mt I**	**Mt II**	**Mt I**	**Mt II**	**Mt I**	**Mt II**	**Mt I**	**Mt II**	**Mt I**	**Mt II**	**Mt I**	**Mt II**	**Mt I**	**Mt II**	**Mt I**	**Mt II**
High	**0.23**	0.22	0.16	**0.42**	**0.93**	0.79	0.25	**0.30**	**0.17**	0.10								
Med	0.11	0.11	0.14	**0.30**	**0.86**	0.70	0.14	**0.17**	**0.11**	0.09	0.03	**0.04**	0.29	**0.43**			0.13	**0.25**
Low	**0.05**	0.03	0.02	**0.08**	**0.68**	0.59	**0.01**	0.00	**0.02**	0.01			0.15	**0.27**	**0.09**	0.02	0.09	**0.13**

Table 21: Aggregated probabilities Method I and 2

We note from Table 21 that out of the 21 levels where the probabilities were aggregated, Method I and II resulted in an increase in ten of the 21 levels each. The later chapters discuss if these outputs are in accordance with our theoretical projections.

6.4.2 Examination Of Empirical Results With Theoretical Projections

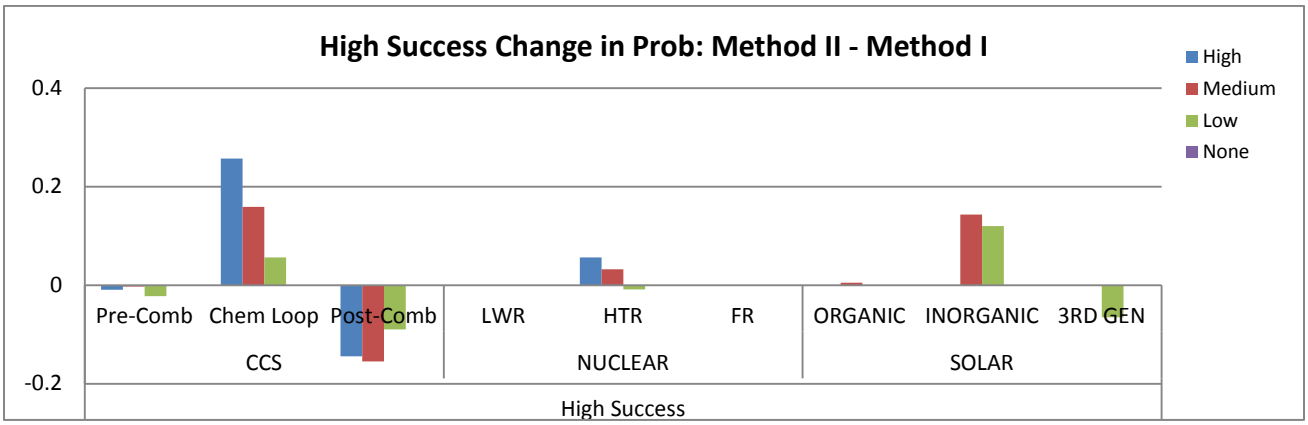


Figure 17: Difference in high success end probabilities between both aggregation methods

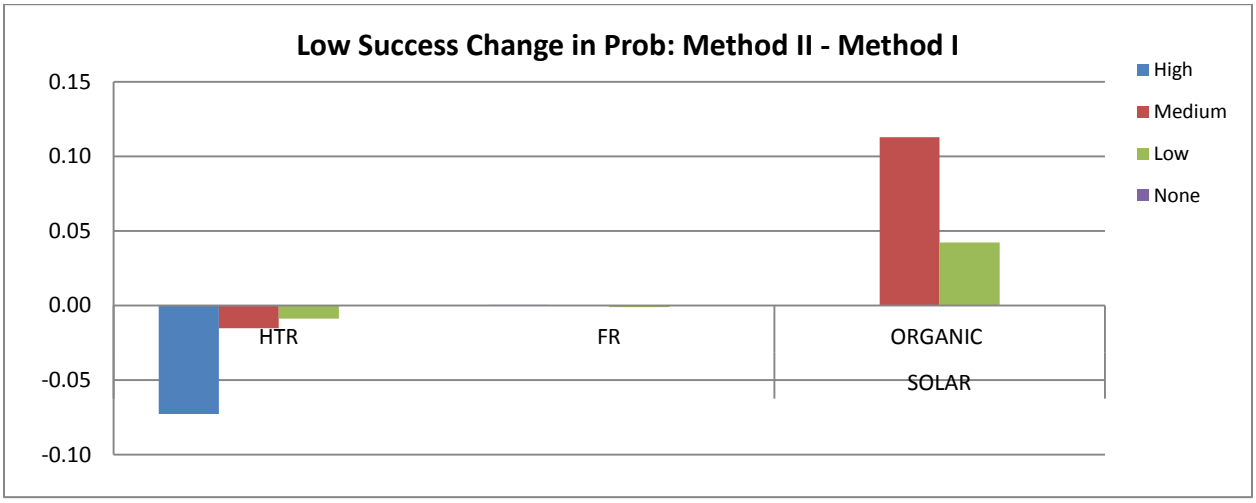


Figure 18: Difference in low success end probabilities between both aggregation methods

In this section, we summarize the results of the use of both aggregation orders (Figure 17 and Figure 18) and compare it to the projections made (Table 10). All the sub-technologies, except Post Combustion, which had a significant number of experts showing a high degree of self-correlation (CCS-Chemical Looping, Solar-Inorganic and Solar-Organic), resulted in a lower end point when

Method I was used relative to Method II. These patterns are expected and are in accordance with the underlying theory.

For the remaining technology sub-categories where experts did not exhibit high self-correlation; CCS-Pre Combustion and Solar-3rd Gen. Both of these sub-technologies resulted in a higher end probability when Method I was used relative to Method II. Also for the Nuclear-High Temperature Reactor sub-category, in which a few of the funding levels had experts with some degree of self-correlation and no correlation in some other funding levels, this resulted in a varied result depending on the funding level.

In summary our results indicate that high self-correlation within experts leads to a very substantial increase in the end success probability when experts are combined later rather than earlier. We noticed this for all the sub-technologies with significant number of experts with high self-correlation except Post Combustion.

6.4.3 Infer Based On Theory To Try To Identify What Is Going On

We observe from the recalculated probabilities (Table 21: Aggregated probabilities Method I and 2) that self-correlation has a significant impact. This is expected due to Method II not correcting for the self-correlation of experts (Baker & Olaleye, 2012). A similar factor that might cause the same effect is the redundancy of having similar experts (Kleinmuntz, Ravinder, & and Dyer, 1986), this was not noticed in this study as most of the technology sub-category were evenly divided between optimists, pessimists and fairly independent experts. Also Cross-correlation between experts was insignificant indicating no redundancy among experts.

6.4.4 Significance Of Difference When Either Method Of Aggregation Is Used

While it is obvious that combining experts earlier is superior to later aggregation, the primary concern in this study is if the difference is indeed significant enough to warrant a re-estimation of the optimal portfolio of energy technologies. From Figure 19 and Figure 20, it seems clear that the method of aggregation indeed has a significant effect on the end point probabilities in majority of the technologies. Thus the changes are significant enough to warrant a re-estimation of the optimal portfolio using the Method I probabilities as inputs to the portfolio optimization model.

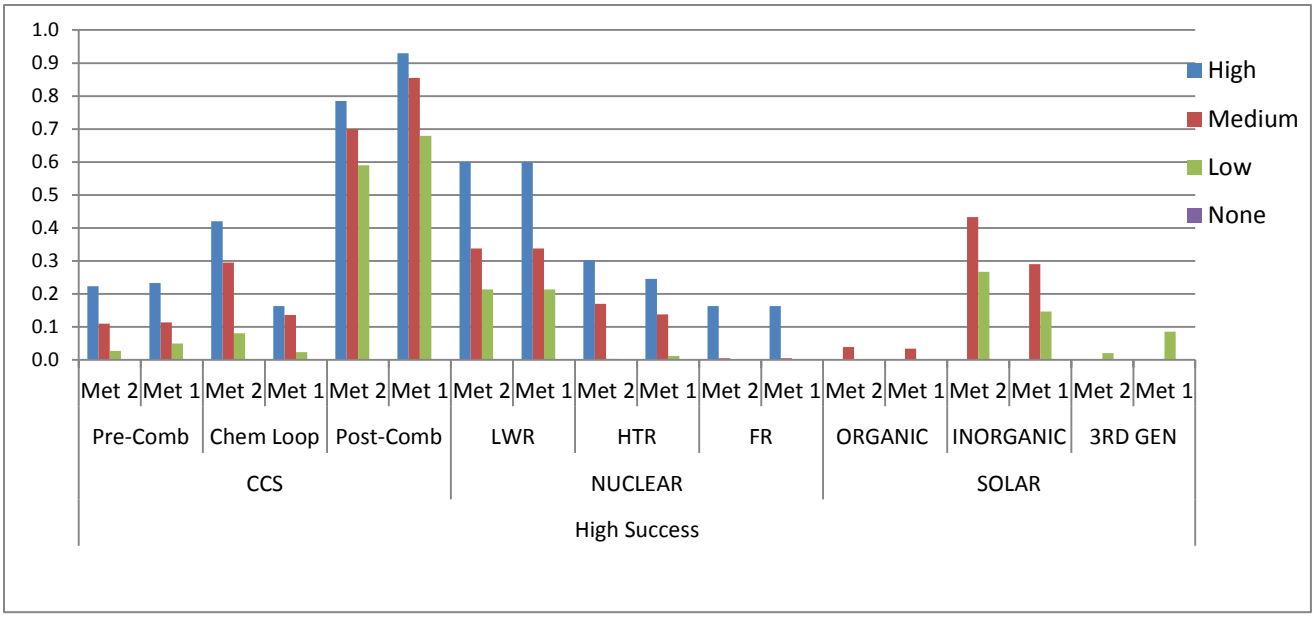


Figure 19: Significance of difference – High Success between both methods

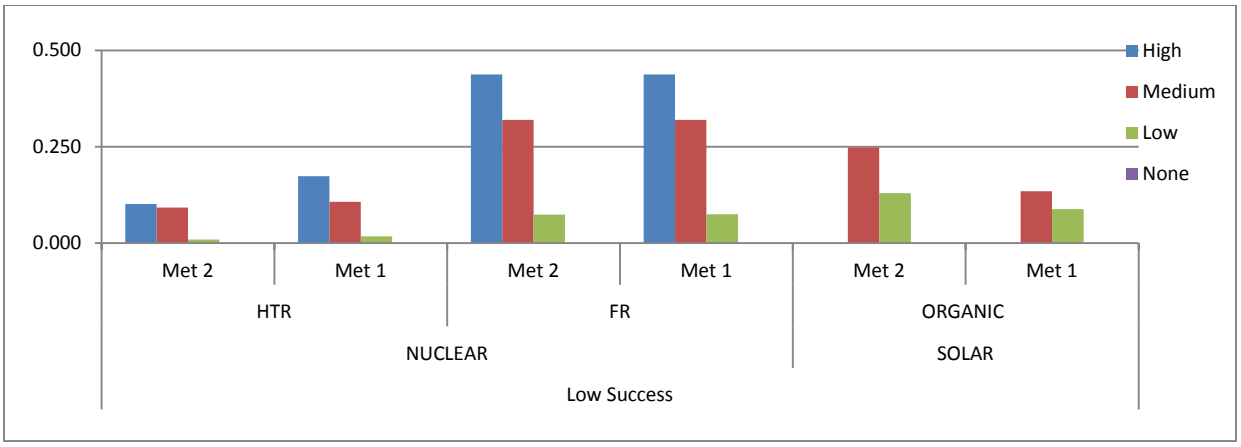


Figure 20: Significance of difference between both methods - Low Success

6.4.5 Issues And Concerns

A couple of immediate questions and assumptions will be discussed herewith as regards to the elicitation process.

6.3.5.A Tree Structure

Here the tree structures of the sub-technologies are examined to further understand the nature of the aggregation error to be expected. As the theory in (Baker & Olaleye, 2012) is based on tree structures with intersecting multiplicative joint hurdles, we only discuss in this section, the technologies with tree structures which depart from this i.e. non-standard tree structures. An examination of all the elicitations carried out shows that only the Carbon Capture Chemical Looping and Post Combustion sub-technologies exhibits a non standard structure. The tree structure and resulting implications are discussed here.

Chemical Looping CCS: This consists of the union of two events which are composed of the intersection of several events as seen by the tree structure in Figure 21. Then it is expected that this is the same as the case with the intersecting events, hence the results should be as expected for standard tree structure.

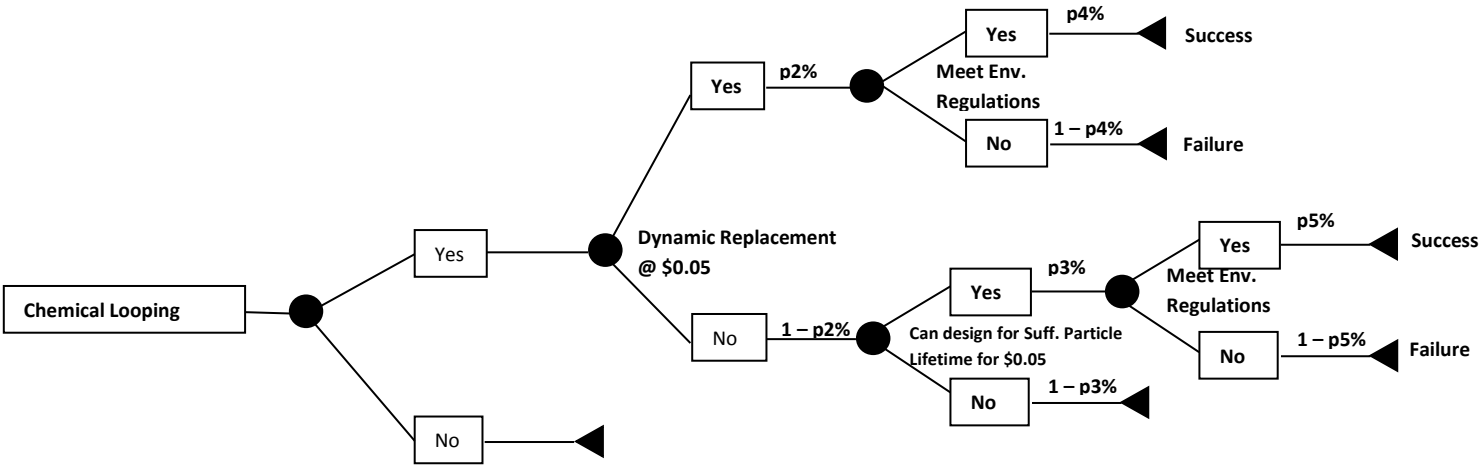


Figure 21: CCS Sample Tree Structure

Post Combustion: This consist of dependent events where the end point success across the level is dependent on the consecutive success or failure of each hurdle i.e. one hurdle say Cryogenic Methods work is dependent on the failure of another hurdle say Molten Solvent(Figure 22). Thus the elicitation proceeds with the assessment of an event been conditional on the probability that another event does not happen, therefore "the analysis will be the same except that the sign of the error will be opposite" (Baker & Olaleye, 2012). This explains why the Method II (combining later) probabilities seemed to be less than Method I despite most of the constituting experts having high self-correlation (Table 10 and Figure 12).

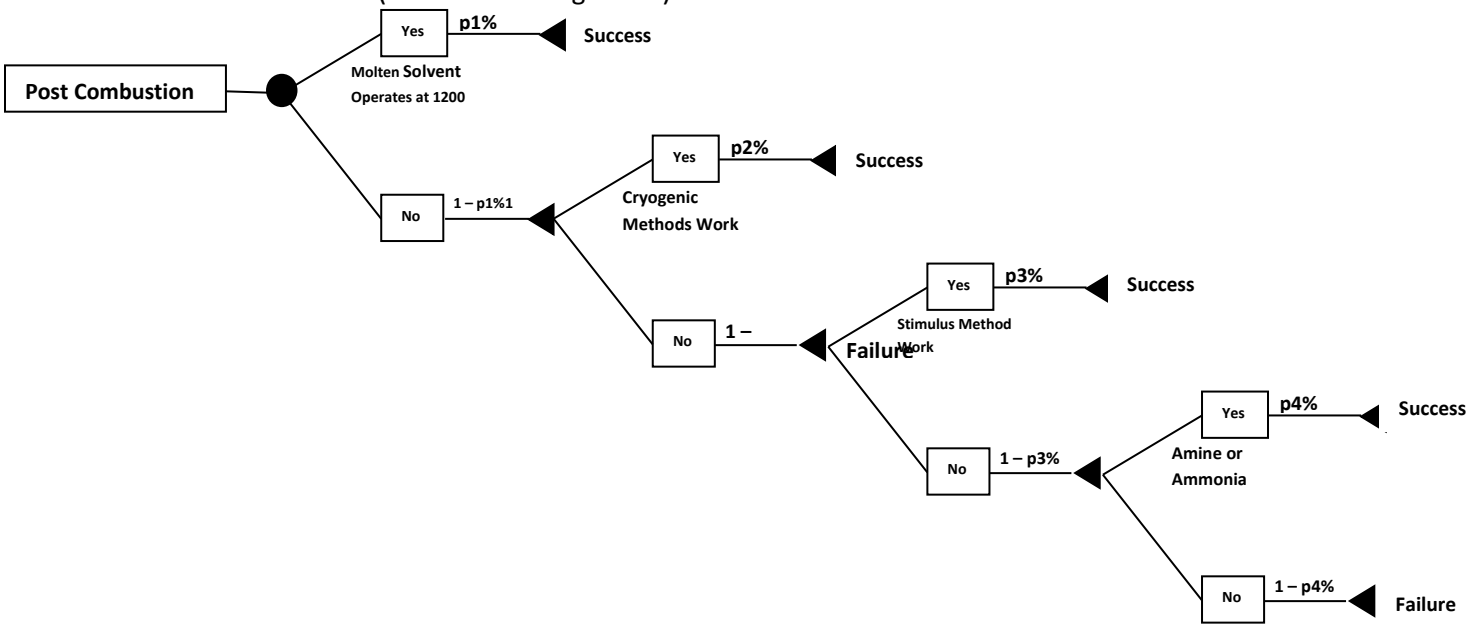


Figure 22: Post Combustion Tree Structure CCS

6.3.5.B Statistical Significance

This is a major concern due to the few number of hurdles assessed for the sub-technologies. This explains why the expert classification is done at the technology level rather than at the funding or sub technology level.

CHAPTER 7

PORTFOLIO OUTPUT

7.1 Energy Portfolio Model Using Method I

Using the same portfolio optimization model used in (Baker & Solak, 2011), the new endpoint success probabilities from Method I aggregation are used as inputs to the two stage optimal portfolio model and the results are discussed below.

7.1.1 New Optimal R&D Portfolio

The new optimal R&D investment portfolio is shown below in Table 22.

	New (Method I) OPTIMAL PORTFOLIO												
	CCS				NUCLEAR				SOLAR				SOCIETAL
Budget (mill $)	Pre-Comb	CCS	Post-Comb	SUM	LWR	HTR	FR	SUM	ORGANIC	IN-ORGANIC	3RD GEN	SUM	COST (tril $)
200	39	56	52	147	0	0	0	0	0	39	0	39	13.4239
400	0	0	52	52	346	0	0	346	0	0	0	0	13.1564
600	0	19	224	243	346	0	0	346	0	0	0	0	13.098
800	39	56	224	319	346	0	0	346	0	77	0	77	13.0674
1000	0	56	519	575	346	0	0	346	0	77	0	77	13.0513
1200	154	56	519	729	346	0	0	346	0	77	0	77	13.0455
1500	386	56	519	961	346	0	0	346	116	77	0	193	13.0376
2000	386	56	519	961	346	0	0	346	116	77	386	579	13.0328
3000	386	56	519	961	346	1544	0	1890	0	77	0	77	13.0039
4000	386	56	519	961	346	1544	0	1890	116	77	386	579	12.9973
5000	386	56	519	961	346	3089	0	3435	116	77	386	579	12.9706
10000	386	56	519	961	346	3089	4633	8068	116	77	386	579	12.9578
20000	386	56	519	961	346	3089	15443	18878	0	77	0	77	12.935

Table 22: New Optimal Portfolio

7.1.2 Change In Optimal Portfolio

In this section, we discuss the change in the optimal portfolio. Table 23 gives the differences in the Optimal Portfolio between Methods I and II. We only display the budget levels where a difference exists between both methods portfolio. The optimal investment portfolio is different in 4 of the 13 budget scenarios considered while the total societal cost increases in all but one of the budget levels, with respect to Method I. As Table 23 shows the figures relative to Method II, we observe that at the 200 and 600 million dollars budget level, Method I favors investment in CCS at the expense of Solar technology. We also note that there was no change in the Nuclear portfolio.

We reiterate again that it is best to always aggregate experts early (Method I), most especially in cases such as ours where the experts seem to show high self-correlation. We thus discuss the loss and policy implications from the use of Method II (later aggregation) in the next section.

	CHANGE IN OPTIMAL PORTFOLIO [METHOD II – METHOD I]							
	CCS				SOLAR			
Budget (mill $)	Pre-Comb	Chem.-loop	Post-Comb	Sum	Organic	In-Organic	3rd Gen	Sum
200	-39			-39		38		38
600	39	37	-172	-96		77		77
4000					714		-386	328
10000					714		-386	328
Max Investment	386	56	519	961	830	77	386	907

Table 23: Change in Optimal Portfolio

7.2 Policy Implications

Budget ($ mill)	Method I($ mill)	Method II($ mill)	Diff ($ mill)
200	13,423,900	13,425,300	1,400
600	13,098,000	13,111,300	13,300
4000	12,997,300	12,999,200	1,900
10000	12,957,800	12,959,700	1,900

Table 24: Total Cost: Policy Implications - Method II Optimal Portfolio using Method I probabilities

Here we examine the societal cost and impact of using the defective method of aggregation (Method II). To do this, we evaluate the optimal portfolio from Method II using Method I probabilities. Table 24 gives the societal cost of evaluating both methods' optimal portfolio using Method I's probabilities. We note that the costs are most substantial in the $600 million budget level. But more importantly we note that relative to the total societal costs which are tens of trillions of dollars, the changes are very insignificant ranging from 0.1% to 0.01%.

7.3 Welfare Maximizing Portfolio

We discuss the work on the alternative R&D portfolio model, Equation [8], for the overall welfare maximizing portfolio where the opportunity cost of the R&D cost is used in place of the budget constraint in this chapter. As discussed earlier in the previous chapter (3.4.1), the objective is to determine the optimal R&D investment portfolio and the accompanying total social cost considering the different assumptions of the opportunity cost of allocating the R&D funding.

Noubara, 2012 (DjimAdoumbaye, 2012) developed a similar greedy algorithm to (Peng, 2010) in solving for the optimal total social cost. This greedy algorithm is readily applicable here as the complexity of the problem makes the genetic algorithm incapable of solving the problem (Peng, 2010). As the greedy algorithm is an approximation algorithm it is not guaranteed to always give the optimal solution, however (Peng, 2010) show that for this particular data set, the resulting solutions will be the same as those from both the genetic algorithm and the stochastic

programming versions. We discuss only the difference in the optimal portfolio and expected social cost when Method I is used Table 25.

Risk	Opportunity Cost	Nuclear		Solar			Total R&D Cost (mil $)	Total ESC (tril $)
		HTR	FR	Organic	Inorganic	3rd Gen		
Base	1	High	High	Medium	Medium	Low	21132	12.9468
	2	High	High	Medium	Medium	Low	21132	12.968
	4	High	None	Low	Medium	Low	4975	12.9905
	8	High	None	Low	Medium	Low	4975	13.0104
Low risk	1	High	High	Medium	Medium	Low	21132	13.8801
	2	High	High	Medium	Medium	Low	21132	13.9013
	4	High	None	Medium	Medium	Low	5689	13.9334
	8	High	None	Low	Medium	Low	4975	13.9538
Medium	1	High	High	Medium	Medium	Low	21132	11.8583
	2	High	High	Medium	Medium	Low	21132	11.8794
	4	High	High	Medium	Medium	Low	21132	11.9217
	8	High	None	Low	Medium	Low	4975	11.9466
High	1	High	Medium	Medium	Medium	Low	10322	10.339
	2	High	None	Low	Medium	Low	4975	10.3447
	4	High	None	Low	Medium	Low	4975	10.3547
	8	Medium	None	None	Medium	None	2928	10.3702
All Excluded Sub-technologies are funded at the HIGH investment level								

Table 25: Welfare Maximizing Portfolio (Method I).

Table 25 gives the funding levels recommended for the different sub-technologies at the different climate damages risk levels (Base, Low risk, Medium and High risk) and opportunity cost of the R&D budgets (1, 2, 4 and 8) when Method I is used. The first column shows the different climate damages risk levels considered (Base, Low risk, Medium and High risk). The second column indicates the various assumed values of the opportunity cost of the R&D budgets (1, 2, 4 and 8). Columns three to seven gives the recommended funding levels. The Total R&D cost and the expected social cost of the portfolios are also given in the last two columns. Table 26 gives the

difference between both Method I and II of the recommended funding level and total R&D cost. As an example, the second to the last column in Table 26 gives the difference (Method II-I) in the R&D funding and the total R&D cost for the High damages case when an R&D budget opportunity cost of 8 is assumed. While Table 25 gives the recommended funding in terms of levels (High, Medium, Low), Table 26 gives the difference in actual $ cost.

From Table 26, Method I will typically result in more investment in R&D, inferring that Method II generally leads to under investment in R&D. The exceptions are the base risk opportunity cost 4 scenario and the Medium risk opportunity cost 8 scenario. This is due to over investment in R&D for Organic Solar in both cases, which is as a result of the self-correlation between the experts which lead to a lower endpoint success probability for Method I. The very significant difference in the success probabilities for 3rd Gen (0.2 for Method II and 0.9 for Method I) ensures that Method I almost always recommends 3rd Gen inclusion in the optimal portfolio while Method II doesn't due to its low success probability when later aggregation is used.

For the degree of significance of the results, we note that when Method I is used, only in 12 of the possible 144 sub-technologies do we observe any change. Also we observe that using Method II would likely lead to underinvestment in R&D for 3rd Gen by 386 million dollars, see second to last column.

Risk	Opportunity Cost	Nuclear	Solar		Total R&D Cost ($mill)
		FR ($mill)	Organic ($mill)	3rd Gen ($mill)	
Base	1	0	0	0	0
	2	0	0	0	0
	4	0	714	-386	328
	8	0	0	-386	-386
Low risk	1	0	0	0	0
	2	0	0	0	0
	4	0	0	-386	-386
	8	0	0	-386	-386
Medium	1	0	0	0	0
	2	0	0	0	0
	4	-10810	0	-386	-11196
	8	0	714	-386	328
High	1	0	0	-386	-386
	2	0	0	-386	-386
	4	0	0	-386	-386
	8	0	0	0	0
There are no changes in the excluded sub-technologies					

Table 26: Changes in the Welfare Maximizing Portfolio (Method II - I)

7.4 Discussion And Policy Implications

We discuss the different inferences from the results from the portfolio outputs in this section.

7.4.1 Optimal R&D Budget

From Table 26 we can determine the relative value of the improvements from the increase in each R&D budget level; the results are given in Table 27. The incremental benefits from the increase in the R&D budgets seems to tail off from the 1,000 ($ million) budget level. The return from the additional 200 ($ million) from the 1,000 budget level is -29, compared to -1337, -292, -153, -80 for the 0, 200, 400, 600 and 800 R&D budget levels respectively.

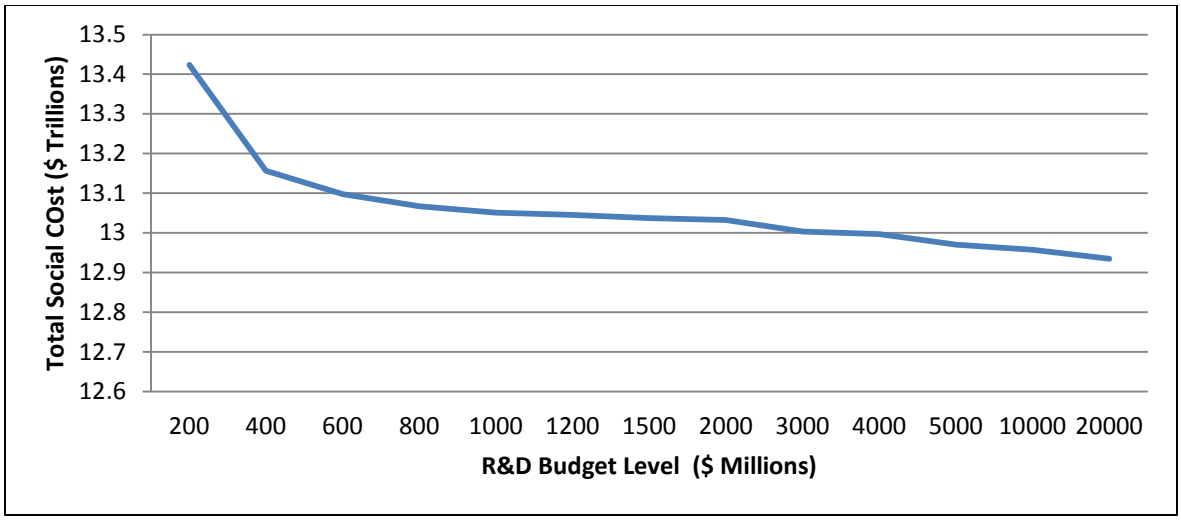


Table 27: R&D cost to reduction in total social cost benefit

7.4.2 Relationship Between The Input Technology Success Probabilities And The Portfolio Outputs

To observe how the changes in the input expert probabilities from the use of the new aggregation order (Method I) impact the optimal portfolio, we observe the percentage change in Figure 23 and Figure 24 compared with the change in the optimal portfolio Table 23. We note that all the sub-technologies outputs follow the trend inferred by the input probabilities. When the aggregated probability using Method II is higher, the optimal portfolio investment from using Method II is always at least as much as from Method I for each sub-technology.

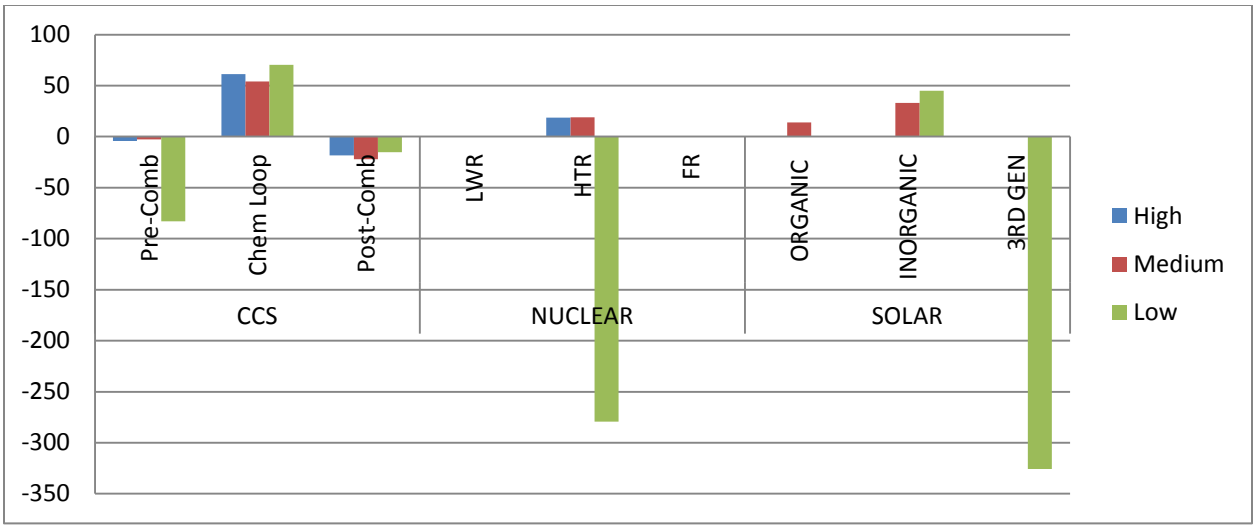


Figure 23: Percentage Change in High Success Probability wrt Met II [((Met II - Met I)/Met II)*100]

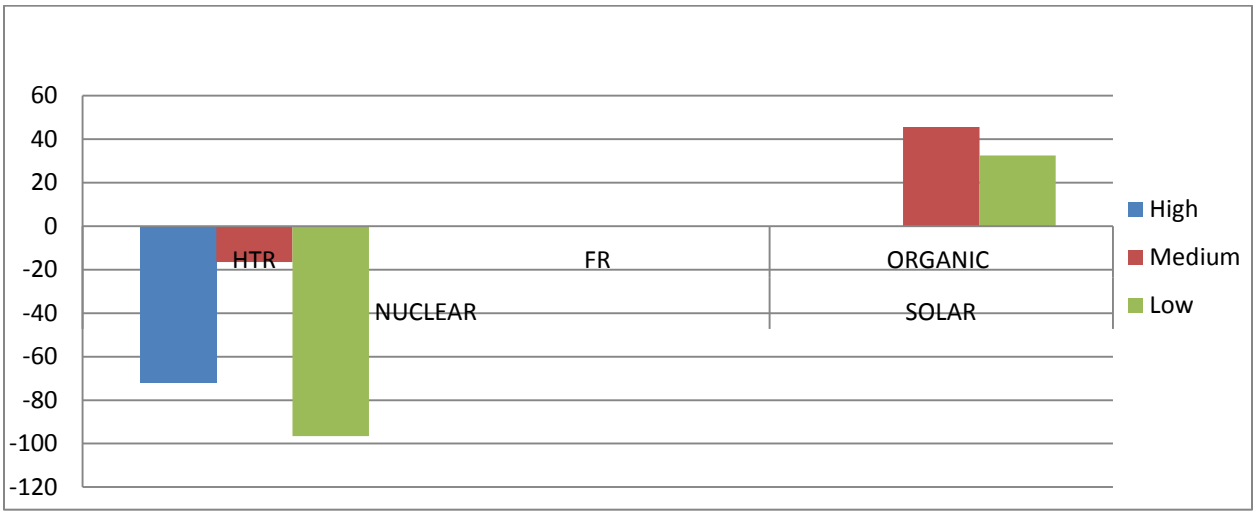


Figure 24: Percentage Change in Low Success Probability wrt Met II [((Met II - Met I)/Met II)*100]

7.4.3 Post Proposal Insights

In this section, we discuss some of the post proposal questions and insights related to the model outputs.

7.4.2.A Significant Increase In The Total Societal Cost For The 600 Million Budget Level

From Table 23, it can observed that the 600 million budget level shows a very significant increase in the total societal cost, this is due to the significantly different resulting optimal portfolios, from using the two orders of aggregation at the $600 million budget level Table 21.

7.4.2.B Why Method I Always Seems To Favor Inclusion Of CCS To Solar

As discussed in the previous section, the changes in the optimal portfolio outputs Table 23 are a function of the inherent self-correlation within experts in the input elicitation probabilities. An evaluation of Figure 23 and Figure 24 shows this. When the input sub-technology probabilities for Chemical Looping, In-organic Solar and Organic Solar are higher when Method II is used, with respect to Method I (due to less self-correlation for Method II), this leads to an increase in the investment allocation to these sub-technologies in Method II, compared to Method I. We also note the opposite for the Pre-Combustion, Post Combustion and 3rd Gen sub-technologies, where the input sub-technology probabilities are lesser for Method II with respect to Method I. As expected, we observe a decrease in the investment in these sub-technologies in Method II compared to Method I.

As such, the input probabilities are higher for Organic and Inorganic Solar and lower for Pre-combustion and Post combustion when Method II is used relative to Method I. Therefore, Method II will favor the inclusion of Solar to CCS while Method I will favor the inclusion of CCS with respect to Solar.

CHAPTER 8

SUMMARY AND CONCLUSIONS

This thesis evaluates the impact of the use of different aggregation methods with respect to the optimal R&D investment portfolio. We conduct simulations of actual human experts eliciting de-composed sub-events. We find that it is always best to aggregate experts early. We show that it is possible in some cases that aggregating later lead to less error, but this is due to randomness. We carry out a theoretical estimation of the likelihood that aggregating earlier is better under the condition of normality of the error distribution. Our results show that early aggregation leads to less errors as the number of expert increase and as the self correlation between experts increase.

Using an empirical study, we show that a high degree of self-correlation within experts is present in some sub-technologies in the elicitation. We combine these probabilities from the experts using the two aggregation orders. We show that later aggregation leads erroneously to a larger success probability when the experts are highly self correlated, because later aggregation does not correct for the self-correlation error of experts.

The resulting probabilities from the elicitation are applied to an optimal portfolio decision problem. We determine the optimal portfolios from using both aggregation orders. We evaluate the loss and policy implications to the society of aggregating the experts later rather than earlier. We find that while the quantifiable impact to the society is very large in proportion to the total societal cost of climate change abatement, the impact is low.

APPENDIX A

ELICITATION TABLES

	SOLAR Purely Organic												SOLAR Inorganic			
	Low Funding Low Success				Med. Funding High Success				Med. Funding Low Success				Low			
	Hurdle 1	Hurdle 2	Hurdle 3	Hurdle 4	Hurdle 1	Hurdle 2	Hurdle 3	Hurdle 4	Hurdle 1	Hurdle 2	Hurdle 3	Hurdle 4	Hurdle 1	Hurdle 2	Hurdle 3	Hurdle 1
EXPERT 1	0.85	0.5	0.9	0.9	0.15	0.6	0.3	0.98	0.98	0.6	0.95	0.98	0.8	1	0.8	0.93
EXPERT 2	0.9	0.3	0.5	0.3	0.5	0.8	0.3	0.7	0.95	0.8	0.5	0.8	0.9	0.9	0.2	0.95
EXPERT 3	0.8	0.5	0.25	0.1	0.3	0.25	0.3	0.25	0.4	0.25	0.3	0.25	0.1	0.1	0.1	0.25
Average	0.85	0.433333	0.55	0.433333	0.316667	0.55	0.3	0.643333	0.776667	0.55	0.583333	0.676667	0.6	0.666667	0.366667	0.71

Table 28: Solar Elicitation Table

	PRE-COMBUSTION HURDLES						CHEMICAL LOOPING									
	LOW		MEDIUM		HIGH		LOW					MEDIUM				
	Hurdle 1	Hurdle 2	Hurdle 1	Hurdle 2	Hurdle 1	Hurdle 2	Hurdle 1	Hurdle 2	Hurdle 3	Hurdle 4	Hurdle 5	Hurdle 1	Hurdle 2	Hurdle 3	Hurdle 4	Hurdle 5
EXPERT 1	5	66	15	70	20	90	50	50	40	50	35	85	85	75	75	50
EXPERT 2	50	10	70	30	90	50	N/A	N/A	N/A	N/A	N/A	95	N/A	N/A	N/A	N/A
EXPERT 3	2	2	10	7	25	15	1	1	1	1	1	2	2	2	2	2
EXPERT 4	N/A	N/A	N/A	N/A	N/A	95	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Mean	19	26	31.66667	35.66667	45	62.5	25.5	25.5	20.5	25.5	18	60.66667	43.5	38.5	38.5	26

Table 29: Carbon Capture and Storage Elicitation Table

APPENDIX B

BAKER & OLALEYE 2012 SIMULATION CODE

```
function [P, ABSERR1, ABSERR2] = errsimnmamn
% Program to obtain the normal random variable obtained by adding differently correlated rv

Mc=-0.8;
% Correlation of the Self Correlated Errors

Ms=0.1;
% Standard Deviation of the Self Correlated Errors

Ecc=0.5;
% Correlation of the Cross Correlated Errors

Ec=sqrt(Ecc);
% Correlation Coefficient of all experts to a random variable

Es=0.5;
% Standard Deviation of the Cross Correlated Errors

Ds=0.2;
% Standard Deviation of the Independent Errors

P1=rand();
% First Standard Random Number between 0 and 1

P2=rand();
% Second Standard Random Number between 0 and 1

eP1=P1*[1;1;1;1;1];
% Original Probability of the Five Experts for the First Hurdle

eP2=P2*[1;1;1;1;1];
% Original Probability of the Five Experts for the Second Hurdle

ruE1=randn();
% Random Variable with Mean Zero 'Support Variable'

% rvE1=0 + Es*ruE1;
% Random Variable ruE1 with mean 0 and STDEV Es

eE1=[ 0 + Es*(Ec*ruE1 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE1 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE1 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE1 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE1 + sqrt(1 - Ec^2)*randn())];
% eE1=[rvE1; 0 + Es*(Ec*ruE1 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE1 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE1 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE1 + sqrt(1 - Ec^2)*randn())];
% 5 Random Variables Correlated Across Experts Representing Hurdle 1
```

```
ruE2=randn();
% Random Variable with Mean Zero 'Support Variable'

% rvE2=0 + Es*ruE2;
% Random Variable ruE2 with mean 0 and STDEV Es

eE2=[ 0 + Es*(Ec*ruE2 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE2 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE2 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE2 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE2 + sqrt(1 - Ec^2)*randn())];
% eE2=[rvE2; 0 + Es*(Ec*ruE2 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE2 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE2 + sqrt(1 - Ec^2)*randn()); 0 + Es*(Ec*ruE2 + sqrt(1 - Ec^2)*randn())];
% 5 Random Variables Correlated Across Experts Representing Hurdle 2

rvM1=randn(5,1);
% 5 by 1 matrix of Random Variables

rvM2=randn(5,1);
% 5 by 1 matrix of Random Variables

eM1=0 + Ms*rvM1;
% First Matrix Representing the Self Correlated Errors, Not Correlated

eM2=[ 0 + Ms*(Mc*rvM1(1,1) + sqrt(1 - Mc^2)*rvM2(1,1)); 0 + Ms*(Mc*rvM1(2,1) + sqrt(1 - Mc^2)*rvM2(2,1)); 0 + Ms*(Mc*rvM1(3,1) + sqrt(1 - Mc^2)*rvM2(3,1)); 0 + Ms*(Mc*rvM1(4,1) + sqrt(1 - Mc^2)*rvM2(4,1)); 0 + Ms*(Mc*rvM1(5,1) + sqrt(1 - Mc^2)*rvM2(5,1))];
% Easy Second Matrix Representing the Self Correlated Errors, Correlated

% eM2=0 + Ms*(Mc*rvM1 + sqrt(1 - Mc^2)*rvM2);
% Second Matrix Representing the Self Correlated Errors, Correlated

eD1=0 + Ds*randn(5,1);
% First Matrix (Hurdle 1) of Independent Errors With Mean Zero and STDEV Ds

eD2=0 + Ds*randn(5,1);
% Second Matrix (Hurdle 1) of Independent Errors With Mean Zero and STDEV Ds

Q1= eP1 + eE1 + eM1 + eD1;
% Sum Across for the First Hurdle

Q2= eP2 + eE2 + eM2 + eD2;
% Sum Across for the Second Hurdle

Q=((sum(Q1))/5)*((sum(Q2))/5);
% Sumproduct of both Hurdles (Method 1)(Avg Across Hurdles before
% Aggregating)[Average Experts First]

R=(sum(Q1.*Q2))/5;
% Products Across Rows of both Experts(Method 2) (Avg Across Experts before
% Aggregating)[Average Experts Later]
```

```
P=P1*P2;
% Product of Both Experts Probabilities

ABSERR1=Q - P;
% Error From (Method 1)(Avg Across Hurdles before Aggregating)

ABSERR2=R - P;
% Error From (Method 2)(Avg Across Experts before Aggregating)
end


NumTrials = 1000000;
P = zeros(NumTrials,1);
ABSERR1 = zeros(NumTrials,1);
ABSERR2 = zeros(NumTrials,1);
for Trial = 1 : NumTrials
  [P(Trial), ABSERR1(Trial), ABSERR2(Trial)] = errsimnmamn;
end
avgP = mean(P(1:end));
avg1 = mean(ABSERR1(1:end));
avg2 = mean(ABSERR2(1:end));
G = [avg1 avg2];
avgAbsP = mean(abs(P(1:end)));
avgAbs1 = mean(abs(ABSERR1(1:end)));
avgAbs2 = mean(abs(ABSERR2(1:end)));
Gabs = [avgAbs1 avgAbs2];
varP = var(P(1:end));
var1 = var(ABSERR1(1:end));
var2 = var(ABSERR2(1:end));
GV = [var1 var2];
% FindEva=find(abs(ABSERR2(1:end))<abs(ABSERR1(1:end)));
% ProbQP=(numel(FindEva))/NumTrials;
ProbQP=(numel(find(abs(ABSERR2(1:end))<abs(ABSERR1(1:end)))))/NumTrials;
% ProbQP=(numel(find(abs(ABSERR1(1:end))<abs(ABSERR2(1:end)))))/NumTrials;
% ProbQP=(numel(find((abs(ABSERR2(1:end))<abs(ABSERR1(1:end)))>1e-15)))/NumTrials;
Prob1=(numel(find(abs(ABSERR1(1:end))>.05)))/NumTrials;
Prob2=(numel(find(abs(ABSERR2(1:end))>.05)))/NumTrials;
RESULT=[G, GV, 0, 0, Gabs, Prob1, Prob2, ProbQP];
% x = 1:NumTrials;
% ABS1=[abs(ABSERR1(1:1000000,1)),abs(ABSERR2(1:1000000,1))];
% hist(ABS1); figure(gcf)
% title('Histogram Of Absolute Errors |Q-P|');
% xlabel('Error')
% ylabel('Frequency')
```

REFERENCES

Ansolabehere, e. a. (2003). *The future of Nuclear energy: An interdisciplinary MIT study.*

Armstong, S. J., Denniston, W. B., & Gordon, M. M. (1975). The use of the Decomposition Principle in making Judgements. *Organizational Behavior and Human Performance , 14* (2), 257-263.

Baker, E., & Adu-Bonnah, K. (2008). Investment in Risky R&D Programs in the Face of Climate Uncertainty. *Energy Economics , 30*, 465-486.

Baker, E., & Olaleye, O. (2012). Combining Experts: Decomposition and Aggregation Order. *Risk Analysis* .

Baker, E., & Peng, Y. (2012). The value of better information on technology R&D projects in response to climate change. *Environmental Modeling and Assessment , 17* (1), 107-121.

Baker, E., & Shittu, E. (2005). Profit-maximizing R&D in response to a random carbon tax. *Resource and Energy Economics , 28*, 160-180.

Baker, E., & Solak, S. (2011). Climate Change and Optimal Energy Technology R&D Policy. *European Journal of Operations Research , 213*, 442 – 454.

Baker, E., & Solak, S. (Under Review). Optimal Climate Change Policy: R&D Investments and Abatement under uncertainty. *Production and Operations Management* .

Baker, E., Chon, H., & and Keisler, J. (2008). *Advanced Nuclear Power: Combining Economic Analysis with Expert Elicitations to Inform Climate Policy.* Technical report.

Baker, E., Chon, H., & and Keisler, J. (2009). Carbon Capture and Storage: Combining Expert Elicitations to Inform Climate Policy. *Climatic Change , 96* (3), 379.

Baker, E., Chon, H., & Keisler, J. (2009). Advanced Solar R&D: Applying Expert Elicitations to Inform Climate Policy. *Energy Economics , 31*, 37-49.

Booker, J., & Meyer, M. (1988). Sources of and effects of interexpert correlation: An empirical study. *IEEE Transactions on systems, man and cybernetics , 18* (1), 135-142.

Box, G. E., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control.* Holden-Day.

Budescu, D., & Rantilla, K. A. (2000). Confidence in aggregation of expert opinion. *Acta Psychologica* .

Cesarini, D., Sandewall, O., & Johannesson, M. (2006). Confidence interval estimation tasks and the economics of overconfidence. *Journal of Economic Behavior & Organization , 61*, 453-470.

Clemen, R. (1986). Calibration and Aggregation of probabilities. *Management Science* .

Clemen, R., & Winkler, R. (1999). Combining Probability Distributions from Experts in Risk Analysis. *Risk Analysis , 19* (2).

Clemen, R., & Winkler, R. (2004). Multiple Experts vs. Multiple Methods: Combining Correlation Assessments. *Decision Analysis* , 167-176.

Clements, M. P., & Harvey, D. I. (2009). Combining probability forecasts. *International journal of forecasting , 27*, 208-223.

Cooke, M. R. (1986). Problems with empirical Bayes. *Risk Analysis , 6* (3), 269-74.

Cooke, R. M. (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science.* Oxford university press.

David, K. (1979). *Correlation and causality* (1st ed.). John Wiley.

de Palma, A., & Picard, N. (2009). Behaviour Under Uncertainty. *The Expanding Sphere of Travel Behaviour Research: Selected Papers from the 11th International Conference on Travel Behaviour Research* (p. 423). In Kitamura, Ryūichi; Yoshii, Toshio; Yamamoto, Toshiyuki: Emerald Group Publishing.

Devore, J. L. (2004). *Probability and statistics for engineering and the sciences.* California: Thomson Brooks/Cole.

DjimAdoumbaye, N. (2012). *Solar Grid Integration.* Amherst: University of Massachusetts, Amherst.

Embrechts, P., McNeil, A., & Straumann, D. (1999). Correlation and dependency in risk management: properties and pitfalls. *Preprint, ETH Zurich* .

Eroglu, C., & Keely L Croxton, K. L. (2010). Biases in judgmental adjustments of statistical forecasts: The role of individual differences. *IJF. Elsevier.* .

Excel, M. (2007). Statistics, Ztest.

Ezekiel, M., & Fox, K. (1959). *Methods of correlation and regression analysis: Linear and Curvilinear.* New York: Wiley.

Feiveson, H., Mian, Z., Ramana, M., & Von Hippel, F. (2011). International Panel on Fissile Materials: Managing Spent Fuel from Nuclear Power Reactors. .

Frey, H., & Rubin, a. E. (1991). *Development and Application of a Probabilistic Evaluation Method for Advanced Process Technologies, 3 Topical Report (Draft).* METC.

Green, M. (2001). Third generation photovoltaic: Ultra-high conversion efficiency at low cost. *Progress in Photovoltaic's: Research and Applications*, (pp. 123-135).

Gunther, T. (2004). The over-optimism among experts in assessment and foresight. *Technological Forecasting and Social Change , 71* (4), 341-363.

Harrison, M. J. (1977). Independence and calibration in decision analysis. *Management Science , 24* (3), 320-328.

Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty , 4*, 5–28.

Kahneman, D., & Tversky, A. (1972). Subjective probability: Adjustment of representativeness. *Cognitive Psychologica , 3*, 430–454.

Kahneman, D., & Tversky, A. (1982). The psychology of preferences. *Scientific American , 246*, 160-173.

Kleinmuntz, D., Ravinder, H., & and Dyer, J. (1986). Reliability of subjective probabilities obtained through decomposition. *Journal of Behavioral Decision Making* .

Kuhnert, P., Hayes, K., Martin, T., & McBride, M. (2009). Expert opinion in statistical models. *World IMAC Congress.*

Kynn, M. (2008). The heuristics and biases bias in expert elicitation. *Royal Statistical Society* .

Li, S., Li, J.-Z., Chen, Y.-W., Bai, X.-W., & Ren, X.-P. (2010). Can Over confidence be debiased by low-probability High-consequence events. *Risk Analysis* .

Makridakis, S. (1986). The art and science of forecasting: An assessment and future directions. *International Journal of Forecasting , 2* (1).

Metz, B., Davidson, O., Coninck, H. C., Loos, M., & (eds.), a. L. (2005). *IPCC, 2005: IPCC Special Report on Carbon Dioxide Capture and Storage. .*]. Cambridge University Pres.

Miller, B. (Forthcoming). When Is Consensus Knowledge Based? *Synthese* .

Morgan, M. G., & Keith, D. W. (1995). Subjective Judgments By Climate Experts. *Environmental Science & Technology , 29*, A468-A476.

Mosleh, A. (1992). On decomposition and aggregation error in estimation: Some basic principles and examples. *Risk Analysis , 12* (2), 203-214.

Mosleh, A., Bier, V. M., & Apostolakis, G. (1987). *Methods for the elicitation and use of expert opinion in risk assessment: Phase 1, A critical evaluation and directions for future research. Technical report.* NRC.

Peng, Y. (2010). *A stochastic R&D portfolio model under climate uncertainty.* Amherst: University of Massachusetts, Amherst.

Plous, & S. (1993). *The Psychology of Judgment and Decision Making.* New York: McGraw-Hill.

Raiffa, H. (1968). *Decision Analysis: Introductory Lectures on Choices Under Uncertainty.* Random House.

Shepperd, J., Carroll, P., Grace, J., & Terry, M. (2002). Exploring the causes of comparative optimism. *Psychologica Belgica, 42* , 65-98.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty Heuristics and Bias. *Risk Analysis* .

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the pcysology of choice. *Science , 211* (4481), 453-458.

Weinstein, N. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology , 39* (5), 806-820.

Weisstein, E. (2005). *Correlation Coefficient.* Retrieved 2012, from Mathword: http://mathword.wolfram.com/CorrelationCoefficient.html

Winkler, R. (1991). Ambiguity, Probability, Preference, and Decision Analysis. *Journal of Risk and Uncertainty , 4* (3), 285-297.