

© 2011 Ryan Paul Haggerty

ELICITING CONTINUOUS PROBABILITY DISTRIBUTIONS

BY

RYAN PAUL HAGGERTY

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Systems and Entrepreneurial Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Adviser:

Associate Professor Ali E. Abbas

ABSTRACT

The process of eliciting a representative probability distribution from a judge is fundamental to decision making under uncertainty. This work summarizes several aspects of eliciting continuous probability distributions. Through experimentation involving 103 judges the method of assessing single variable distributions by fixed variable or fixed probability method is compared. Slight, but consistent, superiority is shown by the fixed variable method. The use of iso-probability contours, a novel methodology for assessing joint probability distributions, is tested through experimentation. The ability of this method to assess a judges belief about the correlation of two variables is characterized. Further recommendations for the use of iso-probability contours are provided. The variable weighting function is introduced. This novel tool is designed to represent a subjects true belief about a continuous probability distribution as explained by cumulative prospect theory. Operating similarly to the probability weighting function, the variable weighting function allows an assessor to gain a judges belief about a specific variable value without requiring multiple assessments. Functional form of the variable weighting function is explored and relevant parameter values are provided.

To my wife and parents, for their love and support.

ACKNOWLEDGMENTS

I would like to thank all of my collaborators, especially Professor David V. Budescu, Gillian Wyman, Hsiu-Ting Yu, Jamie Marcus, Yuhong (Rola) Gu, and Ryan Mulligan. Particular gratitude is due to Professor Ali E. Abbas, who introduced me about decision analysis and a whole new way of thinking. His guidance has been extremely valuable. Thanks to Professor Ronald A. Howard for granting access to his students in the decision analysis class at Stanford University.

On a personal note, I would like to thank my wife who stuck by me through the perils of graduate school and my parents who have always shown complete love and encouragement.

The experiments presented in this dissertation were supported by National Science Foundation under award SES 06-20008.

Reprinted by permission, Abbas, A. E., Budescu, D. V., Yu, H.-., Haggerty, R., A comparison of two probability encoding methods: Fixed probability vs. fixed variable values, *Decision Analysis*, volume 5, number 4, (2008). Copyright (2008), the Institute for Operations Research and the Management Sciences, 7240 Parkway Drive, Suite 300, Hanover, Maryland 21076 USA

TABLE OF CONTENTS

| | |
|---|----|
| CHAPTER 1 INTRODUCTION | 1 |
| 1.1 Motivation | 1 |
| 1.2 Objectives | 5 |
| CHAPTER 2 FIXED PROBABILITY VS. FIXED VARIABLE ASSESSMENT | 7 |
| 2.1 Methods of Assessment | 7 |
| 2.2 Experimental Methodology | 10 |
| 2.3 Results | 17 |
| 2.4 Summary, Conclusions, and Recommendations | 27 |
| CHAPTER 3 ISO-PROBABILITY CONTOUR ASSESSMENT | 31 |
| 3.1 Summary of Assessment | 31 |
| 3.2 Observations | 35 |
| 3.3 Results | 36 |
| 3.4 Conclusions | 40 |
| 3.5 Possible Future Directions | 40 |
| CHAPTER 4 VARIABLE WEIGHTING FUNCTION | 44 |
| 4.1 Introduction | 44 |
| 4.2 Methods | 54 |
| 4.3 Results | 59 |
| 4.4 Conclusion and Recommendations | 64 |
| REFERENCES | 67 |
| AUTHOR'S BIOGRAPHY | 71 |

CHAPTER 1

INTRODUCTION

1.1 Motivation

Our decision are only as good as the information they are based on. In order to make better decisions we need to get new information. This can be done by gaining better a understanding of the distribution of variables that perspective outcomes rely on. This is often done by inquiring from an expert. A person might look up the weather in the news paper before planning a picnic, or a stock trader may look to an analysts opinion of a stocks target price before buying. Decision makers often look to experts opinions, but do not always get the full opinion of the expert. Most often an expert will provide a single number that represents the average expected value for the variable. This quantity is generally more useful than any other single number, but it does not provide a complete understanding of the experts understanding. In order to gain complete understanding from the expert a representative probability distribution of all the variable's possible values must be assessed.

The process of assessing a representative probability distribution function is not always straight forward. Many techniques have been developed in order to accurately depict an experts opinion about a variable. Perhaps the most intuitive method constructs the probability distribution based on the assessed moments. [1, 2, 3] A step further uses moments and quartiles to construct the distribution using the maximum entropy approach. [4, 5] Some work has been done to determine the quality of the probability distribution estimates. [6, ?, 7, 8, 9, 10] A more commonly used method involves assuming a functional form of the distribution and then using judgments in order to assess the parameters of the distribution.[11] Each method

has different advantages and can also lead to artifacts that are not representative of the subjects true beliefs. The central theme of this work is to investigate methods of attaining representative probability distributions of random variables.

1.1.1 Probability Encoding Methods

Eliciting probability, the degree of a subject's belief about a proposition, can be a barrier to the integration of decision analysis into modern day industry. Cognitive and motivational barriers make quantifying ones beliefs difficult and can be particularly challenging in an organizationally complex atmosphere. In this work methods of elicitation of quantiles to describe the probability distribution of a continuous random variable is investigated. The quantile of a probability distribution function can be described by $P[x \leq X] = p$, where the probability of a variable, x , being less than or equal to a value, X , is p , the boundary of the quantile of interest. There are three types of probability encoding methods to determine a relationship between X and p , as described by Spetzler and von Holstein [6]. In the first method, fixed probability (FP) assessment, the value p is held constant and X is assessed. Fixed variable assessment (FV) requires X to be held constant, while p is assessed. The third method is a mixture of the prior two.

Each quantile can be assessed a series of binary choices, during each of which the subject is asked to choose between two gambles. In one gamble a clear probability of success is presented. In the other the subject is asked to bet on whether the variable of interest will have an outcome less than X . In the first gamble it is important that the subject be clear on their chance of success. This is often accomplished by using a probability wheel, a two color wheel with an arrow that can be spun. The arrow landing on color one marks success, therefore the ratio of the wheel made up by color one is the probability of success. If a FV assessment is used, the value of X is constant, while the ratio of the two colors can be adjusted on the wheel, changing the probability of success. Once the subject is

indifferent between the two gambles, the probability of success on the wheel is assumed to be equivalent to the probability that the variable of interest will be less than X . If the FP method of assessment is used the probability of success on the wheel is held constant, while successive choices about the variable gamble are made, by adjusting the value of X . Again, when indifference is reached the probability of the wheel is assumed equivalent to the value of the variable being less than X . A series of assessed quantiles can represent the cumulative distribution representing the belief of the subject.

Variants of both the FV and FP methods are used in industry. The choice of assessment method is generally made due to industry standard with little to no consideration of the effect of the method on the distribution assessed. There is no rigorous study as to the effects of each method, even though the choice of method is generally accepted as effecting the outcome distribution. Chapter 2 of this dissertation compares the two methods of assessing quantiles and discusses the effects of the method used on consistency and accuracy.

1.1.2 Iso-Probability Contours

Often assessed variables are assumed to be independent of one another. This greatly simplifies the process of eliciting the variables of interest and using them in calculated decisions. This simplification can often lead to errors, especially if strong correlation exists between the variables. In order to correctly analyze the value of uncertain prospects the joint probability distribution function between the relevant set of random variables must be assessed.

In developing a representation of a subject's belief about a joint probability it is unreasonable to assume the subject can quantify the correlation between the random variables being assessed. Thus a procedure involving simple questions which the subject can be expected to answer is desired. For this reason iso-probability contours, a set of points with the same cumulative probability, has been developed as a way to describe a joint probability distribution.[12] These contours reduce the joint probability distribution into two

dimensional curves. This provides a more intuitive method to assess and to utilize the joint distribution in future calculations. An iso-probability contour relating the variables x and y is assessed by finding a series of gambles in which the subject is indifferent between the following gambles

Gamble 1: $x \leq X_1$

Gamble 2: $x \leq X_2$ and $y \leq Y_2$

All points (X_2, Y_2) will exist on the same contour as (X_1, ∞) . The cumulative probability that relates the points of the contour can be assessed by simply finding the probability that $x \leq X_1$, according to the subjects belief. By assuming a functional form to the distribution, a contour can then be related to the correlation of the two variables. Chapter 3 of this dissertation deals with the implementation of iso-probability contours in assessing a subjects belief about the correlation of two variables. The ease of use and effectiveness of this method is discussed in detail.

1.1.3 Variable Weighting Function

Even when a probability distribution function can be assessed that is representative of the belief of a subject, there still may be cognitive biases that will persuade a decision maker to behave differently then would be dictated by normative decision making. Cumulative prospect theory, for which Kahneman received the Nobel Memorial Prize in Economics in 2002, provides an explanation as to why sub optimal decision making is a common occurrence. In prospect theory a decision maker views value relative to an anchored state of wealth, always judging a prospect as a gain or a loss from that state. This causes the decision maker to overweight great losses or gains. This manifests in the way the decision maker utilizes a probability distribution.

The difference in the belief of the decision maker and the way the behave can be described by the probability weighting function. This function illustrates the weight that a decision

maker places on each probability and can then describe way the decision maker will act facing a given uncertainty. [13] The probability weighting function provides a large step in utilizing cumulative prospect theory. It allows the adjusted probability of a constant variable value to be determined directly, however, the adjustment of the variable value cannot be directly obtained for a given probability. For this purpose the variable weighting function is introduced in chapter 4. The form of the variable weighting function as well as some of its properties are presented.

1.2 Objectives

Each section of this dissertation is meant to assist in assessing a probability distribution from experts. This is done from three major directions each pertaining to the better assessment of probability distribution functions of continuous variable. The first contribution emphasized the choice of assessment method. The current state of knowledge is expanded through experiments, leading to better understanding of how the method of assessment affects the assessed distribution function. More specifically an experiment was conducted where both fixed probability and fixed variable methods of assessment were utilized for each of 103 subjects. The resulting assessments were analyzed with respect to monotonicity, accuracy, and precision. In addition the rate of assessment and participant preference were also considered. Insight into why a method may be preferred is discussed and recommendations regarding when each method should be used are provided.

The second major direction explores and develops the practices of an unused assessment tool, iso-probability contours. Several subjects witnessed two correlated variables, and then participated in an assessment to construct representative iso-probability contours related to the variables. The subjects ability to reproduce the true contours was examined. In addition, the use of the contours to extract information about the true correlation of the variables is investigated and compared to other methods of assessing the correlation of variables.

The final section contributes to the use of assessed distribution functions. This is through the development of the variable weighting function. The existence of a variable weighting function will make utilizing an assessed distribution function much more intuitive and less time consuming for many applications. This allows for greater overall ease of use of the assessed probability distribution function. The variable weighting function was constructed for several subjects. The general properties of the variable weighting function are discussed, and the various functional forms that can be used to represent the variable weighting function are proposed. The parameters of the most useful functional forms will be derived for the most common variable weighting functions as determined from extensive previous work on probability weighting functions. Furthermore, a method for determining the parameters of the variable weighting function for any random variable and subject, given the parameters of the probability weighting function, was also derived.

Each of these contributions provides practical insights into probability distribution assessment of continuous random variables. The advice provided in each section works toward the goal of consistent and accurate assessments. Reaching this goal would be a major step in the complete automation of decision processes and would remove a barrier that often keeps decision analysis methods out of industry standards.

CHAPTER 2

FIXED PROBABILITY VS. FIXED VARIABLE ASSESSMENT

2.1 Methods of Assessment

Assessment of a decision makers belief about a continuous random variable is central to normative decision making. Constructing a representative probability distribution function is not only challenging, but is also often a barrier to successfully integrating decision analysis into industry. There is a variety of literature that analyses the steps involved in eliciting a probability distribution and evaluating the quality of the estimates.[6, ?, 7, 8, 9, 10, 4, 5] Many of these use quantiles and/or moments to construct the distribution. The main focus of this chapter is to determine the advantages of different methods of using quantiles to construct probability distribution functions which are representative of the beliefs of a subject.

The quantile of a probability distribution function can be described by $P[x \leq X] = p$, where the probability of a variable, x , being less than or equal to a value, X , is p , the boundary of the quantile of interest. There are three types of probability encoding methods to determine a relationship between X and p , as described by Spetzler and von Holstein [6]. In the first method, fixed probability (FP) assessment, the value p is held constant and X is assessed. Fixed variable assessment (FV) requires X to be held constant, while p is assessed. The third method is a mixture of the prior two.

2.1.1 Fixed Probability

A fixed probability (FP) assessment is generally performed through a series of binary choices. A probability generating device, (eg. dice roll, coin toss, probability wheel) is fixed such that there is a known probability of success, p . The subject is asked to choose between a gamble involving the probability generating device or a gamble where success is achieved if the variable of interest, x , has a value less than X . If the subject chooses the probability generating device the choice is repeated with a larger value of X . If the subject chooses the second gamble the next choice has a smaller value of X . This is repeated until the subject is indifferent between the two choices.

The most commonly chosen quantiles of the cumulative distribution for the FP method are the median ($p = 0.5$), and the quartiles ($p = 0.25$ and 0.75) [14, 15]. Variants of this approach are widely used in practice when experts provide their High, Base, and Low values (0.1, 0.5, 0.9) for a variable of interest to construct decision trees or to conduct sensitivity using tornado diagrams [16, 17, 18, 19]. In some cases, it may be necessary to assess as many as 5 quantiles (e.g., 0.10, 0.25, 0.50, 0.75, and 0.90), or even 7 quantiles (e.g., 0.01, 0.10, 0.25, 0.50, 0.75, 0.90, and 0.99) [20]. Alpert and Raiffa [21] report, however, that subjects perform poorly when judging the extreme quantiles.

Variants of the FP paradigm are also used in practice, where judges are asked to provide their intervals of a variable of interest [21, 22]. For example, when asking judges for a 50% interval for a quantity (say the price of a stock a year from now), the judge is asked to produce the two quartiles of the variable such that it is equally likely that the variable will fall between them, or outside the interval. This approach has enjoyed popularity in the psychological literature, since it can be used to illustrate the judges alleged overconfidence.[23] Soll and Klayman [24] suggest that the precision of the encoding is improved if, instead of judging $X\%$ intervals, judges are asked to estimate separately its end points (i.e., the $(100 - X)/2$ and the $(100 + X)/2$ quantiles).

2.1.2 Fixed Variable

Fixed variable (FV) assessments also require a set of binary choices leading to a point of indifference. The assessor will fix a variable value, X , rather than a probability. The subject will choose between a gamble using the probability generating device with a probability, p , of success or a gamble that x will be less than the value X . If the subject chooses the first gamble, the probability, p , is reduced for the next choice. On the other hand, if the subject chooses the gamble involving the value of x , the probability, p , will be increased for the next choice. Once indifference is reached the value of p is recorded as the probability that $x < X$. For this type of assessment it is beneficial to have a probability generating device that is not limited to discrete values of probability, such as a probability wheel.

For FV assessments the value of X chosen by the assessor depends on the application, since the assessor must have some knowledge of the expected range before selecting a value X . This is of course the most efficient way to proceed if the end goal is the probability that x will be less than some value. Common applications included competitive bidding situations and assessing the lower and upper bounds of dose-response function curves.[?, 25, 26]

2.1.3 Comparing Methods

Despite the common use of both FP and FV methods, no prior comprehensive comparisons have been evaluated between these methods. The comparisons that are available have employed between-judge experiments that make direct comparison difficult, since different subjects can have different beliefs about a variable.[24, 27, 28] The major focus of this chapter is to present an experiment designed to elaborate on this comparison of FP and FV methods. The methods are judged on the ability of the subjects to give self consistent assessments, the accuracy of the assessments, the preference of the subjects, and the ease and speed which of the assessment. The work in this chapter was previously published by in the Journal of Decision Analysis[29], much of the text and figures in this chapter are reprinted

here with the permission of the publisher.

2.2 Experimental Methodology

Probability assessment is both the subject of purely theoretical research (by psychologists, statisticians, decision theorists) and the “bread and butter” of practitioners of decision analysis. Although both use similar practices they have different emphases and, therefore, approach the work in slightly different ways. For example, decision analysts prefer to make the elicitation task as easy and convenient for the judges as possible. This facilitates the interaction with the judges, minimizes the assessment time, and helps generate data of higher quality and consistency that is necessary for the solution of the problem at hand. On the other hand, theoretical researchers strive for unbiased and neutral designs that are optimal for uncovering the basic principles underlying the judgments, and for characterizing the performance of the various methods. This experiment was conducted in this latter spirit. For example, we elicit the various quantiles in isolation, and in (partially constrained) random order, without allowing the judges direct access to their previous judgments, and we repeat the same number of assessments using two encoding methods. Therefore the results we report are relevant to researchers and also provide a useful benchmark for practitioners.

2.2.1 Subjects

103 students enrolled in the Decision Analyses classes at Stanford University and the University of Illinois volunteered to participate in this experiment. The participants included 71 men and 32 women, whose average age was 26.7 years with a standard deviation of 5.9. Most students majored in Management Science and Engineering and they had all been exposed to probability encoding in class lectures.

2.2.2 Procedure

The experiment was conducted online. After logging on to the site and reading the informed consent form judges answered a few demographic questions. Next they were allowed to choose to assess either

1. the closing value of the Dow Jones Industrial Average on December 12th 2006
2. the high temperature in Palo Alto on December 12th 2006

(students at the University of Illinois were instructed to judge only the DOWJ values). Judges could choose the units Fahrenheit or Celsius for the temperature assessment, but for the purpose of the analysis all temperatures were converted to Celsius centigrade (we found no significant differences between the two sets of judges). A subset of the participants was chosen at random and provided with a chart of historical data for their variable of choice (see Figure 2.1). Table 2.1 shows the number of judges in each condition.

Table 2.1: Number of judges in each experimental condition

| | No chart | With chart | Total |
|--------------------------|----------|------------|-------|
| Dow Jones index | 11 | 20 | 31 |
| Temperature (Celsius) | 13 | 17 | 30 |
| Temperature (Fahrenheit) | 22 | 20 | 42 |
| Total | 46 | 57 | 103 |

The experiment started a week before the target date. Judges could access the site at any time until December 11, 2006. After choosing their preferred variable, judges were asked to provide lower and upper bounds for that variable. Then they performed the two elicitation tasks (FP and FV). The order of the assessment methods was determined randomly. Judges operated at their own pace, but completed all their judgments in one session that lasted between 30 and 45 minutes. Judges were presented with a sequence of binary choices regarding a hypothetical \$20 lottery (see Figure 2.2 for an example): The deal on the left displays a

setting on the probability wheel. Judges win the prize if we spin the wheel and the arrow lands on orange. The deal on the right displays a certain value of the variable of interest. Judges win the prize if the value of the variable of interest is less than the displayed value.

If one chooses the deal on the left, one of two things could happen, on the next screen the comparison would be between

For FP method the same percentage of orange on the wheel and higher value of the variable of interest, say 30 degrees

For FV method the same value of the variable and a lower orange setting on the wheel, say 25% Orange

On the other hand, if they chose the deal on the right, the two possible comparisons on the next screen would be between

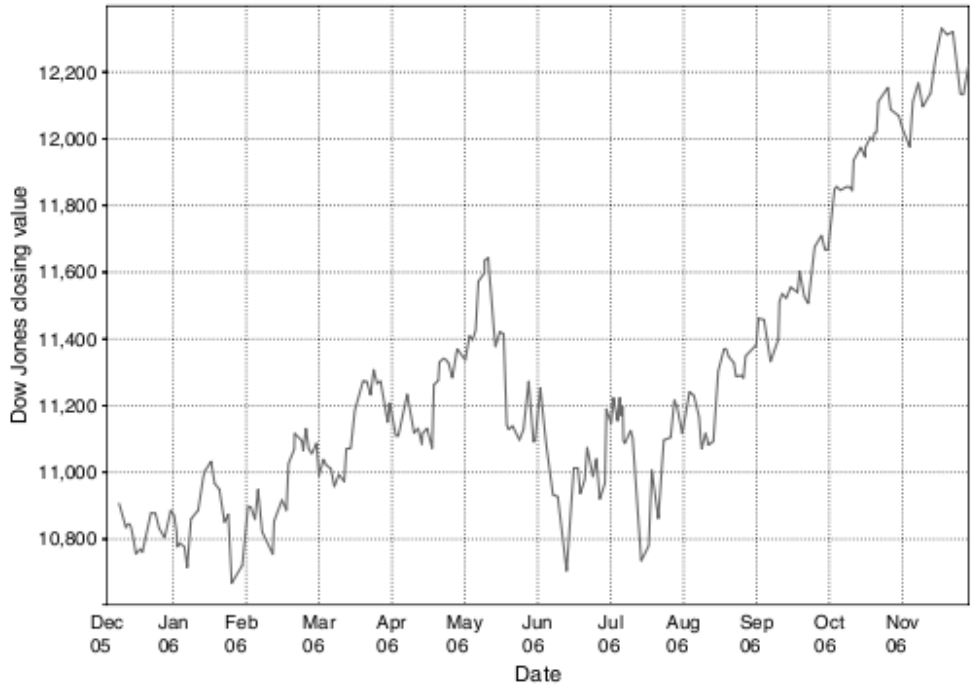
For FP method the same wheel and a lower value of the variable of interest, say 20 degrees

For FV method the same value of the variable and a higher orange setting on the wheel, say 75% Orange

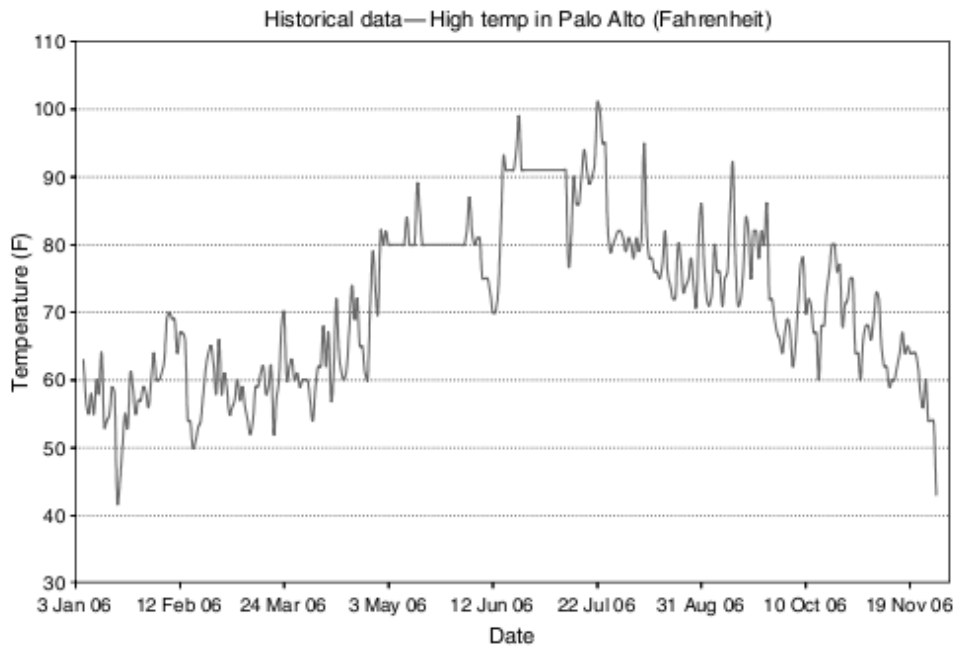
The next value was determined by a halving algorithm (see section 2.2.3). The process stopped when the range of values was below a narrow (predetermined) threshold, or when the judges expressed indifference between the two deals. At that point the judges were asked to confirm their decisions and a new series of choices with a new fixed probability or fixed variable value was initiated. In all our analyses we treat the midpoints of the ranges elicited as the judges judgments.

Judges completed 10 series of judgments to determine five fractiles using each method. For the FP method the fixed probabilities were 5%, 25%, 50%, 75% and 95%; and for the FV method the values were set at the 5%, 25%, 50%, 75% and 95% of the range of variable values specified by the judges (extended by 20%). Note that while all judges used the same five probabilities in the FP method, the actual value used in the FV case varied across

individuals, as a function of the range they specified. In both methods the first point elicited was the central one (i.e., the median for FP, and the midpoint of the range for FV). The other 4 points were presented in one of several predetermined orders that were counterbalanced across judges. The judges did not have direct access to their previous judgments when making their choices. After completing the assessments the judges were asked a series of questions to evaluate the process and compare the two methods in terms of their ease and comfort level.



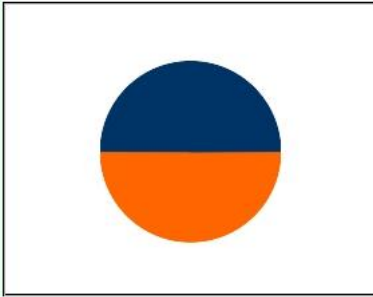
(a) Dow Jones



(b) Palo Alto Temperature

Figure 2.1: The historical charts shown to select judges during their assessment.

You win \$20 if



we spin the wheel and it lands
on **orange**.

<25 degrees C

the high temperature (in
Celsius) in Palo Alto on
Tuesday December 12th 2006
is less than 25.

Choose

Indifferent

Choose

Figure 2.2: Example of an elicitation screen.

2.2.3 Elicitation Algorithm

A range is a certain interval (lower and upper bounds) that bracket the value we are interested in. The range is reduced through a sequence of questions (choice options) and answers (choices). Two types of range reductions were done in the experiment. One was the range of the fraction (percentage) of the orange section of the probability wheel and the other was the range of the target values. The first case reduces to a special case of the second case when the variable range is of size one hundred. As such there was only one reduction algorithm in the experiment. The reduction algorithm consists of halving the range with each question. This approach provides the maximum reduction in the entropy of the range if we believe that a value is uniformly distributed across the range (for further information on halving algorithms, see Abbas [4]). There were two stopping conditions:

1. The user was indifferent between the two deals in the question. This indicates the range should be reduced to this point.
2. The range interval is reduced to less than 3 units

The minimum number of questions to reach a stopping condition is one, while the maximum number of questions is $\log_2(Range) - 1$.

Confirmation questions were also asked a consistency check. The number of questions asked varied depending on the stopping condition of the above algorithm. If the stopping condition was indifference only one confirmation question was asked. If the stopping occurred because the range was narrow enough, two confirmation questions were asked (at the upper and lower values of the range). Taking the confirmation questions into account, the maximum number of questions asked for a given point on the marginal distribution was $\log_2(Range)+1$. For the special case of the fixed quantity this means the maximum number of questions is 7 when indifferent and 8 otherwise.

2.3 Results

2.3.1 Monotonicity of Judgments

Recall that under each method the five points were elicited independently, without visible record of the previous points, and in no obvious order. Thus the first question is whether the judges judgments are monotonic and whether there is a better degree of monotonicity with one method vs. the other. Monotonicity is satisfied if, for every pair of points, X_i and X_j , we observe:

$$X_i > X_j \leftrightarrow F(X_i) > F(X_j) \quad (2.1)$$

To determine the degree to which this condition is met, we calculated the Kendall τ_b rank correlation coefficient for judgments based on (1) midpoints of fractiles estimated with FV method, (2) midpoints of fractiles estimated with FP method, and (3) a combination of both estimates of FV and FP methods. For any sample of size n there are $\binom{n}{2} = n(n-1)/2$ distinct pairs. In our case there are $n = 5$ fractiles (X_i) and there cumulative probabilities $F(X_i)$, defining 10 pairs for each method. Let C be the number of pairs that are concordant (i.e., satisfy the condition in Equation 2.1), and D be the number of pairs that are discordant (i.e., violate the condition in Equation 2.1). Kendall's τ_b is the difference between the proportion of concordant pairs and the proportion of discordant pairs. Formally:

$$\tau_b = \frac{\sum_{i=1}^{n-1} \sum_{j=1}^n \text{sgn}(X_i - X_j) \text{sgn}(F(X_i) - F(X_j))}{\binom{n}{2}} = \frac{C - D}{\binom{n}{2}} = \frac{C - D}{C + D} \quad (2.2)$$

where sgn is the sign function. In the presence of ties the numerator of the fomula is $\sqrt{(C + D + T_x)(C + D + T_y)}$ where T_x is the number of pairs with ties on X (but not Y), and T_y is the number of pairs with ties on Y (but not on X).

Kendall's τ_b is a nonparametric measure that does not depend on the domain of assessments, the scales used, or their range. It is therefore convenient for comparing the two encoding methods. It ranges from -1 (all pairs are discordant) to 1 (all pairs are concordant),

and it is 0 when there are equal numbers of concordant and discordant pairs. Figure 2.3 shows the cumulative distributions of the Kendalls τ_b values for the two encoding methods. It also includes the values calculated using the combined assessments from both methods ($n = 10$ defining 45 pairs). The values based on the judgments with FV are slightly higher than those for the FP and, not surprisingly, both are superior to the joint set since it includes a larger number of points (the FP and the FV assessments). However, it is reassuring that the monotonicity of the combined assessments is not much lower than the results obtained for each method separately. This result also provides some insights into the goodness of the repeatability of the assessments obtained using two different methods.

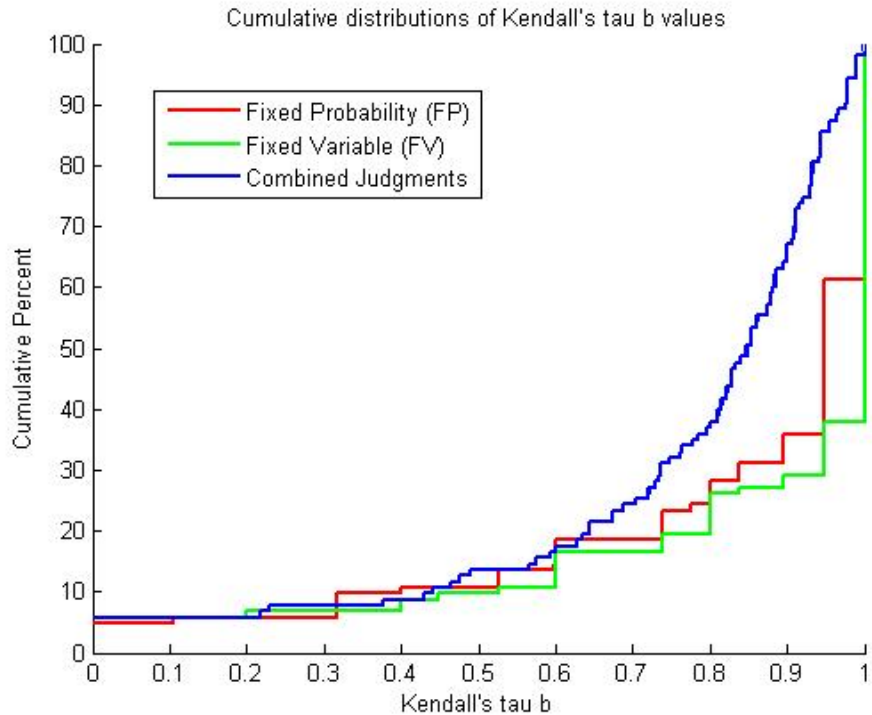


Figure 2.3: Cumulative Percentages of Kendall's τ_b for Fixed Probability Assessments, Fixed Variable Assessments and the Union of the Two Assessments.

Table 2.2 summarizes the medians of Kendalls τ_b values. The first two rows describe the monotonicity of the 5 fractiles elicited within each method (FP and FV) separately, and the third row measures the monotonicity of all 10 fractiles elicited by the two methods

combined. The last row in each panel summarizes the degree of order consistency between the two methods (that is, the degree to which judgments elicited with one method, are ordinarily consistent with those obtained with the other one). Although it is, clearly, lower than the monotonicity achieved within each method separately, it is quite high indicating almost 90% level of rank agreement. These values confirm the impressions from the figure, and highlight the impressive level of monotonicity achieved by the judges with either method, and for the combined assessments.

Table 2.2: Monotonicity of the judgments for each method

| Median Rank Correlation, τ_b | Temp. | Dow Jones | All Assessments |
|-----------------------------------|-------|-----------|-----------------|
| Fixed Probability (FP) | 0.95 | 0.95 | 0.95 |
| Fixed Variable (FV) | 1.00 | 1.00 | 1.00 |
| Combined Points (FP & FV) | 0.86 | 0.85 | 0.85 |
| Cross Methods (FP & FV) | 0.80 | 0.74 | 0.77 |

One third of the judges had identical Kendalls τ_b for both methods. However, the majority of judges (45%) had higher rank correlation coefficients in the FV assessment, and only a minority (22%) was more monotonic in the FP assessment. This difference is significant by a sign test ($Z = 1.81$; $p < .05$ one sided). Interestingly, judges who were monotonic in one method, were more likely to be monotonic in the other method as well: The correlation between the two (within judge) measures of monotonicity is 0.51 ($p < 0.05$), and it is consistent across the two variables.

2.3.2 Fitting the Judgments

Often quantiles are used to construct the full continuous probability distribution function. This can be done by fitting the quantiles to a functional form. [11, 30] A popular choice for this form is the Beta distribution, which can reproduce a wide variety of shapes. This form is also popular among Bayesian statisticians because it is a conjugate prior to the binomial

and Bernoulli distribution. Hughes and Madden [31] provide a comprehensive survey of the different methods used to construct a Beta distribution using location statistics or quantile assessments.

We fitted Beta distributions to the midpoints of the fractiles estimated with the FP and FV methods, separately. The Beta is a continuous two-parameter distribution defined over a given bounded range. Its density is given by:

$$Beta(\alpha, \beta, a, b, x) = \frac{(x - a)^{\alpha-1}(b - x)^{\beta-1}}{\int_a^b (x - a)^{\alpha-1}(b - x)^{\beta-1} dx} \quad (2.3)$$

where a, b are the lower and upper bounds of the domain (respectively), and α, β are the two parameters of the Beta distribution. Of course, Beta is not the only distribution that can be used to model bounded variables[32, 33], but it is frequently used as a prior distribution in Bayesian analysis. We make no claim of superiority or exclusivity for the Beta, but use it to illustrate the results and to facilitate the comparison of the two methods in a meaningful fashion.

We used Matlabs “fminsearch” function to minimize the squared residuals and estimate the two shape parameters ($\hat{\alpha}, \hat{\beta}$) within the range (lower and upper bounds, a and b) defined by the judges judgments. Note that the optimization takes different forms, depending on what is being minimized. When using fixed probabilities, $p_i, i = 1, \dots, 5$, we minimized sums of squared deviations in the metric of the variables ($X = \text{Temperature or Dow Jones}$):

$$\min_{\hat{\alpha}, \hat{\beta}} \sum_{i=1}^5 (X_i - \hat{X}_i)^2 \quad (2.4)$$

where $\hat{X}_i = BetaInverse(p_i, \hat{\alpha}, \hat{\beta}, a, b)$ and X_i is the elicited value of the variable corresponding to a cumulative probability, p_i . In the case of fixed values, $V_i, i = 1, \dots, 5$, we minimized

total squared deviations in the metric of cumulative probabilities:

$$\min_{\hat{\alpha}, \hat{\beta}} \sum_{i=1}^5 (p_i - \hat{p}_i)^2 \quad (2.5)$$

where $\hat{p}_i = \text{Beta}(X_i, \hat{\alpha}, \hat{\beta}, a, b)$ and p_i is the elicited value of the cumulative probability corresponding to the variable value, X_i .

The moments of the Beta distribution are simple functions of the two parameters α and β as well as the upper and lower bounds, a and b . More specifically, the mean, μ , and variance, σ^2 , are given by

$$\mu = \frac{\alpha b + \beta a}{\alpha + \beta}, \sigma^2 = \frac{\alpha \beta (b - a)^2}{(\alpha + \beta)^2 (\alpha + \beta + 1)} \quad (2.6)$$

Figure 2.4 presents examples of some of the better fits obtained from the experiment. The legend of the figures list the shape parameters (α and β), and the upper and lower bounds (a and b). Table 2.3 summarizes the means and standard deviations of the distributions of the various variables for each elicitation method. The top panel uses all 103 judges, and the bottom panel presents results based only on those judges ($n = 74$) who displayed high levels of monotonicity ($\tau_b \geq 0.8$) with either assessment method (FP and FV). The values are reasonable (the temperature on the target date was $15^\circ\text{C} = 59^\circ\text{F}$, and the DOW JONES closed at 12,316, and are quite similar in the two methods.)

Table 2.3: Average means and standard deviations of the fitted beta distributions

| Method | Temp | | Dow Jones | |
|-----------------------------------|--------------|-------------|-------------|------------|
| | Mean | St. Dev. | Mean | St. Dev. |
| FP (X residuals) | 14.99 (4.16) | 3.58 (2.19) | 12523 (623) | 747 (1470) |
| FV ($F(x)$ residuals) | 15.36 (3.90) | 5.19 (2.74) | 12427 (393) | 864 (1648) |
| Only cases with $\tau_b \geq 0.8$ | | | | |
| FP (X residuals) | 15.03 (3.79) | 4.12 (1.98) | 12385 (488) | 494 (1110) |
| FV ($F(x)$ residuals) | 15.15 (3.94) | 5.08 (2.44) | 12967 (351) | 615 (1238) |

The moments of the two distributions fitted for each judge are compared and the number

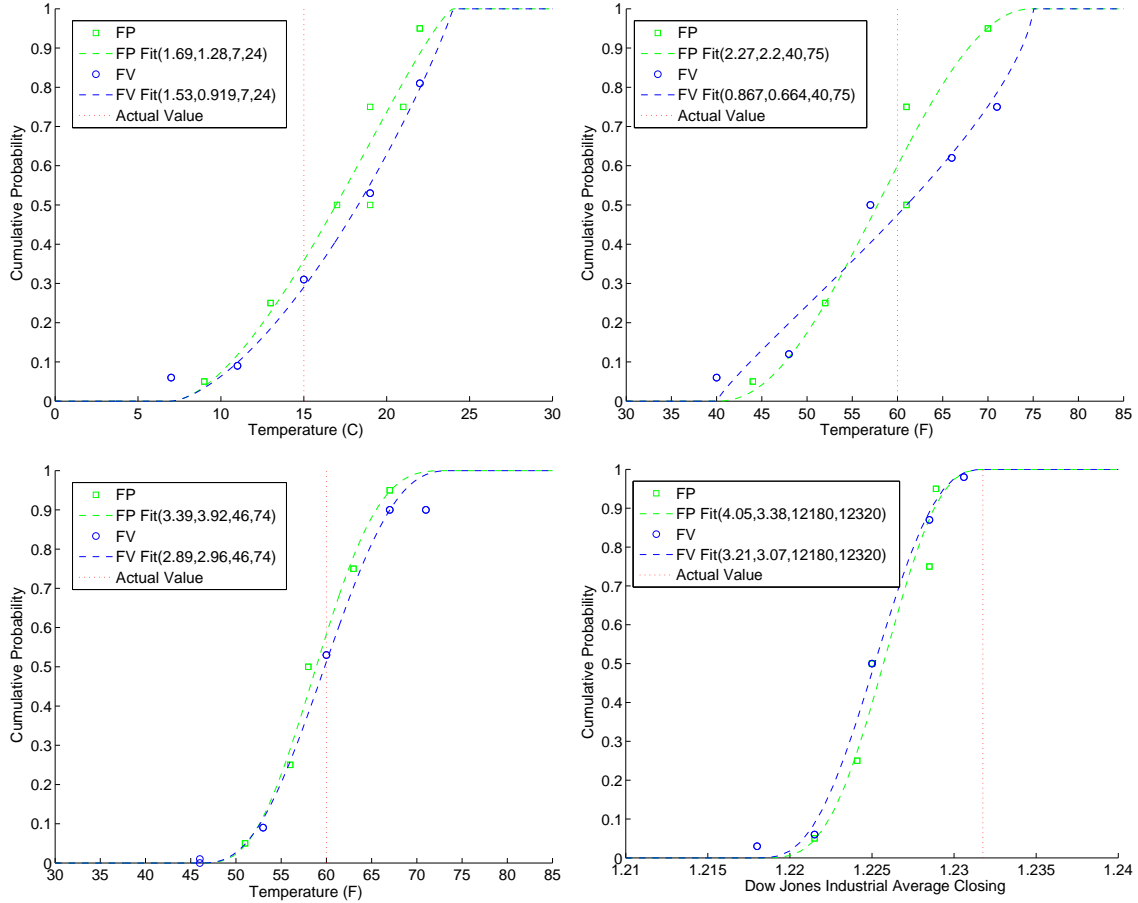


Figure 2.4: Subjects assessments and beta fits.

of cases (i.e., judges) where one method had a higher mean or variance are counted (note, that these are all within-judge comparisons), see Table 2.4. There are only small differences in the fitted means (and almost equal splits of judges with higher/lower means under each method, with 28/44 for temperature and 15/16 for Dow Jones). On the other hand, for 85% of the judges (59 judges for temperature and 29 judges for Dow Jones), the variance of the distribution extracted from the FV is higher than its counterpart based on FP.

The goodness of fit of the solutions as measured by the RMSE (Root Mean Squared Error) for each method and variable (averaged across all judges) are summarized in Table 2.5, along with the count of cases where one method outperformed the other. As indicated earlier, when fitting the distributions we minimized different types of residuals ($F(X)$ in the FV

method, and X in the FP method). In the latter case, we normalized the values relative to the range stated by the judges, so all the RMSEs are in the 0 – 1 range, and can be compared meaningfully. The two methods fit equally well and there is no clear advantage to one method over the other.

2.3.3 Accuracy of the Judgments

In this section we address the question of how well the probability distributions provided by the various judges under the two methods fared with the historical record of the temperatures in Palo Alto. Given that temperatures at a particular location are (a) relatively stable over time and (b) vary only negligibly within a week, we constructed the distribution of the temperatures in Palo Alto on December 12 ± 3 days (i.e., December 9-15) based on the data recorded between 1955 and 2007 (We obtained 345 data points for this location and dates at <http://www.wunderground.com/history/airport/KPAO/2007/12/15/>

Table 2.4: Compared moments of the FP and FV fits for each judge

| | Temp. | Dow Jones |
|--------------------------------------|-------|-----------|
| Mean(FP)-Mean(FV) | -0.37 | 59 |
| Mean(FP)-Mean(FV) | 1.28 | 153 |
| # positive diffs./ # negative diffs. | 28/44 | 15/16 |
| SD(FP)/SD(FV) | 0.75 | 0.59 |
| # ratios > 1/# ratios < 1 | 13/59 | 2/29 |

Table 2.5: Comparison of goodness of fit for distributions fitted with the two methods

| | Temp. | Dow Jones | Overall |
|--------------------------------------|--------|-----------|---------|
| RMSE(FP) | 0.072 | 0.081 | 0.075 |
| RMSE(FV) | 0.078 | 0.098 | 0.084 |
| RMSE(FP)-RMSE(FV) | -0.006 | -0.017 | -0.009 |
| RMSE(FP)-RMSE(FV) | 0.056 | 0.072 | 0.061 |
| # positive diffs./ # negative diffs. | 35/37 | 17/14 | 52/51 |

DailyHistory.html). As a measure of proximity, we used the Kolmogorov-Smirnov (KS) statistic (the maximal absolute difference between the estimated cumulative distributions and the historical distribution of temperatures) for the FV and the FP judgments. On average, the KS scores of the FV estimates are smaller than their FP counterparts (mean difference = 0.064), and they are significantly closer ($t(71) = 2.37, p < 0.05$) to the historical distribution. This pattern also holds for a small (but statistically significant) majority of judges (56%; $Z = 2.39 ; p < 0.05$) of the judges. Thus, the distributions extracted from the FV method fit the historical data slightly better.

The same pattern is observed when we compare the sum of squared differences between the FV estimates and the historical data, and their counterparts based on the FP estimates. They are lower (Mean difference = 0.102, $(t(71) = 1.94, p = 0.06)$). This also holds for a small, but significant, majority of the individual judges (58%, $Z = 2.82 ; p < 0.05$).

Whereas the previous sections analyzed the quality of the judgments extracted by the two methods, in the next analyses we focus on a comparison of the methods in terms of the judges performance and perceptions. More specifically we ask whether the judges find one method easier to use by analyzing both objective and subjective measures.

2.3.4 Reaching Indifference

Our elicitation procedure yields upper and lower bounds for each of the fractiles obtained with either encoding method (recall that the elicitation procedure terminated when the difference between the upper and lower bounds was below a specified low threshold, or when judges clicked Indifferent). Upper and lower bounds may also appear in practice if there is not enough time to reach indifference or if the information available is too vague and prevents one from identifying a precise indifference point[?]. Table 2.6 summarizes the proportion of cases where judges actually converged to a single point. There is an impressive level of convergence (for example, 78% of judges expressed indifference for the assessments of Dow

Jones values with FP). However, the FV method induces higher percentages of indifferences for temperature. This difference is statistically significant for the temperature ($p < 0.05$ by a sign test). The FP method yielded a higher percentage of point estimates for the Dow Jones, but this difference is not significant.

Table 2.6: Proportion of cases where judges converged to a point

| | Temp. | Dow Jones |
|-------------------|--------|-----------|
| Fixed Probability | 60.00% | 78.07% |
| Fixed Variable | 83.05% | 74.84% |

2.3.5 Response Time

Are judgments faster (and, presumably, easier) in one of the two methods? Judges used different numbers of questions to reach the upper- and lower-bound for various fractiles. Thus, we compared the average response times per question in the various conditions. The mean response times of the judges (across all 5 series) were analyzed in a 3 way ANOVA with two between judges factors (units and presence of chart) and one within judge factor (elicitation method). These means are presented in Table 2.7.

The effect of the elicitation method is significant ($F(1, 99) = 13.04, p < 0.05$) with the FV method inducing faster responses. The presence of the chart slowed down the response time by about half a second (6.86 vs. 6.22). This difference was also significant ($F(1, 99) = 5.62, p < 0.05$).

2.3.6 Perceptions and Preferences of the Judges

At the conclusion of the experiment, judges were asked to compare the two methods along three dimensions using 7-point scales. For the purpose of this analysis we collapsed these ratings into 3 coarser categories: The midpoint of the scale (4) is interpreted as indicative

of indifference/neutrality between the methods and all responses on one side of the scale (1-3 and 5-7) were classified as favoring one of the methods. A clear majority (64%) of the respondents thought that the FV elicitation method is simpler and more natural. Table 2.8 shows also a clear preference for this elicitation method. The results are identical for both variables and with/without the benefits of historical charts.

Are the judges preferences for a method reflected in the quality of their judgments? Table 2.9 cross-tabulates the judges preferences for a method and the method in which they had higher Kendalls τ_b . Among those who prefer the FP method (top row), there is no difference between the monotonicity under the two methods but, remarkably, those who prefer the FV elicitation (or were indifferent between the two methods) the preferred method induces, indeed, higher levels of monotonicity in a clear majority of the cases.

Table 2.7: Mean judgment time (in seconds) as a function of the target variable, presence of the chart, and the elicitation method

| | Chart | N | Fixed Probability | | Fixed Variable | | Difference | |
|-----------|-------|-----|-------------------|--------|----------------|--------|------------|--------|
| | | | Mean | SD | Mean | SD | Mean | SD |
| Dow Jones | N | 11 | 8.97 | (3.00) | 3.92 | (0.67) | 5.05 | (3.06) |
| Dow Jones | Y | 20 | 7.41 | (1.20) | 6.82 | (0.91) | 0.59 | (0.90) |
| Temp. | N | 35 | 6.79 | (0.77) | 5.23 | (0.69) | 1.55 | (0.59) |
| Temp. | Y | 37 | 7.00 | (0.60) | 6.21 | (0.46) | 0.78 | (0.73) |
| Overall | | 103 | 7.22 | (0.51) | 5.75 | (0.35) | 1.46 | (0.50) |

Table 2.8: Preference for elicitation method in the various groups

| | Chart | N | Which Method Do You Prefer? | | |
|-----------|-------|-----|-----------------------------|-------|-----|
| | | | FP | Equal | FV |
| Dow Jones | N | 11 | 27% | 9% | 64% |
| Dow Jones | Y | 20 | 30% | 10% | 60% |
| Temp. | N | 35 | 14% | 9% | 77% |
| Temp. | Y | 37 | 27% | 11% | 62% |
| Combined | N | 46 | 17% | 9% | 74% |
| Combined | Y | 57 | 28% | 11% | 61% |
| Overall | | 103 | 23% | 10% | 67% |

2.4 Summary, Conclusions, and Recommendations

We used a web-based system based on simple binary choices to elicit fractiles of probability distributions. Our main goal was to compare two competing methods, Fixed Probability and Fixed Variable values. All the assessments were made in real time, one fractile at a time, so the judges could not see their previous judgments. The FV and FP methods were successful: The judges reported no major problems, and provided high quality monotonic, reasonable and meaningful judgments that were consistent across the two methods.

The results of our experiment show that the two methods were practically indistinguishable in many ways (e.g., the means, and the goodness of fit, of the Beta distributions based on the FP and FV judgments). We did find, however, several systematic differences between the two methods, and these differences point to a slight superiority of the FV method. To recapitulate, we found that the judges were able to make these judgments faster, and were more likely to reach full indifferences (rather than establishing narrow intervals) with the FV assessments. It is not surprising that the majority of the judges express a clear preference for this method in the post-experimental evaluations. Convenience and ease of use do not guarantee quality, so it is reassuring that the FV method also resulted in judgments with higher levels of monotonicity, and matched slightly better the historical distribution of the target variable. The distributions based on the FV method had higher variances than their counterparts based on FP. Given the recurring concern that subjective probabilities are too narrow (reflecting overconfidence), we view this as a positive feature of the method.

Table 2.9: Relationship between preferred method and monotonicity of judgments

| Method Preferred | Higher Monotonicity | | | |
|------------------|---------------------|-------|----|-------|
| | FV | Equal | FP | Total |
| FP | 7 | 10 | 7 | 24 |
| Indifferent | 9 | 1 | 0 | 10 |
| FV | 30 | 23 | 16 | 69 |
| Total | 46 | 34 | 23 | 103 |

We believe that two factors can explain the superiority of the FV method in our study. The first is, simply, the nature of the response scale (probabilities) which is (a) bounded by 0 (impossibility) and 1 (certainty), and (b) universal in the sense that it applies to all events, and is independent of the measurement units of the particular variables. The second factor is the fact that people face more often, and are more familiar with, problems that resemble the FV judgments. These judgments match the single-event format of most decisions (under risk) problems that we encounter daily. We are likely to answer (to ourselves, or to others) questions regarding the likelihood that certain future events will exceed predetermined thresholds. For example, we may need to judge how likely it is that (a) the temperature will be above 30, (b) the rain will last more than 1 hour, (c) ones blood pressure will be below the threshold requiring medication for hypertension, (d) ones childs SAT score will be above the admission cutoff of her favorite college, etc. To answer such questions one relies on his/her life-long experience with these variables and specific cues about the particular target (what I know about todays weather, or my childs abilities). On the other hand, we rarely need to estimate the required level of a certain event to reach a certain probability. Questions such as (a) how hot it should be next Sunday, to exceed 90% of all summer days, or (b) how high should my childs SAT score be to place her in the top 15% of the applicants to her favorite college, etc. are more complex since they require more knowledge about other possible outcomes.

2.4.1 Some Practical Recommendations

Recall that in our study the various fractiles were elicited in isolation using binary comparisons with no direct access to previous assessments. This restriction makes perfect sense in a research setting, but could be relaxed in a decision analysis. It is safe to assume that when judges have access to their previous assessments, and events are presented in a systematic fashion (e.g., in ascending or in descending order) the performance would be improved in

all respects (e.g., monotonicity, speed, level of convergence). Thus, our results provide some sense of the lower bounds for the monotonicity and accuracy of the FP and FV methods in decision analysis.

There are two, somewhat surprising, findings in our study. About half of the judges in our sample had access to historical charts of the relevant variables but, for the most part, this extra information did not make a difference the quality of their judgments and the distributions extracted from them was, essentially, identical to the other half who did not have access to this aid. The simplest explanation is that we observe a “ceiling effect”. Recall that (a) we allowed judges to select which distribution they preferred to assess (and in the case of temperature, to choose their units), and (b) both variables were familiar (and “experienced” on a daily basis) to begin with. In other words, it is likely that judges selected to judge the variables about which they were most knowledgeable, so there was not much information in the charts that was not available to them anyway! We hypothesize that more information in the form of past results would have been beneficial if they were to judge variables with which they were less familiar (say temperatures and stock market outcomes in other countries). This has to be verified in future work, but we recommend (tentatively) having such information in the system, and allowing the judge to determine whether he/she wants to access it while making the judgments.

We observed that judges who took longer times to make the judgments were not necessarily more consistent than those who answered faster (there was no correlation between time to answer and global monotonicity). Judges were not instructed to answer quickly, and were not offered any incentives for slower/faster response rates. It makes sense to assume that they answered at the rate that was most convenient and natural for them and we recommend (tentatively) not imposing time constraints and allowing judges to respond at their preferred rate.

The theoretical literature indicates that it is possible to fit Beta distributions based on as few as two points, but in many applications of decision analysis the norm is three fractiles[19].

We achieved satisfactory fits with 5 points in both methods. In sensitivity analyses (not reported here) we found that the fitting procedure was highly robust to significant changes in the domain (up to $\pm 20\%$) of the fitted Beta distribution. Our results indicate that the removal of the end point fractiles from the elicitation led to the highest change in variance, and the removal of the mid fractile led to the lowest change, and Budescu and Du[22] documented the differential pattern of mis-calibration of the subjective 90% probability intervals. In light of these results we recommend asking judges to (a) estimate the range of the target variable, (b) encourage them to be generous in this task and consider all feasible values of the variable, and (c) elicit at least 5 fractiles.

Our method relied on a self-terminating series of binary questions that identify narrower ranges on every step. Ideally, this series of questions ends when the judge declares his/her indifference between the wheel and the deal that depends on the target quantity. In our algorithm for the FV method (that used up to 7 consistent answers) this ideal was achieved in over 80% of the cases overall, and the ranges identified in cases where convergence to indifference was not achieved were quite narrow (under 0.02 overall). Algorithms that terminate before 7 questions would lead to fewer indifferences and wider ranges.

2.4.2 Future Research

Although the results of this study favor the FV method we recognize that their generalizability should be re-examined in future studies using different variables and judges, including populations of acknowledged experts. An additional factor that should be studied is the robustness of our findings under various changes to the algorithm we employed here. For example, future work should test whether the results hold if the original range of values is pre-determined by the experimenter (rather than being selected by the judge), and the iterative sequence of preferences is replaced by a more direct equivalence judgment.

CHAPTER 3

ISO-PROBABILITY CONTOUR ASSESSMENT

3.1 Summary of Assessment

An isoprobability contour is defined by a set of points in \mathbb{R}^n each of which has a constant probability that n random variables of interest are less than the point, in the respective dimension, is constant.[12] The isoprobability contour of two random variables according to the beliefs of 21 subjects were assessed. Each subject went through a learning session, where they became familiar with the distribution and correlation of the random variables. The subjects belief about the distribution of each independent variable was assessed, followed by the assessment of two isoprobability contours. The subjects belief about the correlation of the variables was also assessed using more traditional methods (eg. providing an estimated correlation coefficient). Each assessed isoprobability contour was fit to the ideal contour described by a constant correlation coefficient. Comparing the fitted correlation coefficient to the correlation coefficient that was estimated by the subject provides a method analyzing the accuracy of the subjects assessed contour.

3.1.1 Subject Learning

Each subject was allowed to use a computer program that generated pairs of values for random variables X and Y . The distribution of the random variables were set by the administrator by manipulating parameters such as the mean and standard deviation of each variables as well as the correlation between the variables. For this experiment the random

variables were set to have a constant distribution with a mean value of 50 and a standard deviation of 10. The correlation of X and Y was set to be either 0.9 or 0.1, in order to see the effects of both high and low correlation. At each iteration of the learning process the computer generated a pair of X, Y values. The value of either X or Y was shown to the subject, who was then asked to estimate the value of the variable they were not shown. There were 50 iterations of the process for each learning session. At the end of the process the subjects were allowed to see a scatter plot of all of the X, Y values that were generated and the statistics of each distribution (ie. the mean, variance and range of the distribution).

3.1.2 Marginal and Iso-probability Contour Assessment

Following the learning process the subjects were asked binary questions in order to assess their understanding about the variables they were shown. First the subjects belief about the marginal distribution of both X and Y was acquired. Using a probability wheel as a reference point of probability, the subjects were asked to choose between spinning the probability wheel or gambling on the value of the random variable of interest being less than or greater than some value. The majority of these assessments were done in fixed variable mode, by fixing the random variable gamble and then adjusting the probability on the wheel until the subject was indifferent. At least two points for each assessment were done using fixed probability assessment mode, where the probability on the wheel is fixed and the value of the variable in the gamble is adjusted until the subject was indifferent. These points were set such that the wheel was a 50% probability and a 25% probability. Such a method directly probed points on the subjects belief about the cumulative distribution of the variable. [29, 6] The marginal assessment was continued until the assessor was convinced an adequate distribution could be fit. In the case that the subject provided inconsistencies in their answers such that the cumulative distribution would have a negative slope, the subject was made aware of the inconsistency and allowed to change their answers if they deemed it

necessary to be true to their belief about the variable.

After the marginal distributions were finalized the iso-probability contours were assessed for constant probability of 50% and 25%. The subject was again asked to choose between two binary gambles. One choice was a gamble that a single variable was less than some value. This value was chosen by the fixed probability assessed points acquired during the marginal assessment indicating the subjects belief about the cumulative distribution at 50% or 25% probability for corresponding iso-probability contour. The second choice was a gamble that an X, Y pair would have the characteristics such that $X < X^*$ and $Y < Y^*$. The value of either X^* or Y^* was held constant while the other was varied until the subject had no preference between the two gambles. The assessment of each iso-probability contour was continued until the assessor found that the contours shape could be defined by the assessed points. Between 5 and 7 points were acquired for each subject. As with the marginal distributions, the subject was made aware of any inconsistencies in their assessment and was allowed to change their answers if they felt the change would better represent their understanding of the correlation of the variables. Such inconsistencies include a sequence of preferred gambles that were deterministically less valuable than the gambles they were preferred over. Another inconsistency that was monitored was the relative positions of the two contours; the contours should not cross.

3.1.3 Comparative Assessments

Following the assessment of iso-probability contours, the subject was asked to provide their belief about the relationship between X and Y using other methods. The subject was first asked to provide the value of the correlation coefficient between X and Y . In some cases the assessor described what the meaning of the correlation coefficient was in the case that the subject was not familiar with the definition. The subject was then asked what the probability was of $X < \bar{X}$ given that $Y < \bar{Y}$. The subject was also asked what the probability was

that $X > \bar{X}$ given that $Y > \bar{Y}$. For the last two questions the subject was asked to imagine that two more points were generated using the same distribution, X_1, Y_1 and X_2, Y_2 . The subjects were asked to provide the probability that $X_1 < X_2$ given that $Y_1 < Y_2$ and the probability that $Y_1 < Y_2$ given that $X_1 < X_2$.

3.1.4 Subject Payment

The subjects were paid based on their score during the learning section up to \$20. As this section was not meant to be difficult, the pay out was between \$18 and \$20 per subject. An additional payment was made based on the out come of a gamble based on the assessed 50% iso-probability contour. The subject was shown their assessed contour. A single point they assessed was then shown to the subject, that can be describe by the points X^*, Y^* . The subject was asked if they thought that the next two points generated by the same distribution would satisfy the condition $X < X^*$ and $Y < Y^*$. They were allowed to choose yes or no separately for each of the two points. After they had chosen the points were generated and the subject was given an additional \$10 for each of the points they choose correctly. The total payout to each subject ranged from \$20 to \$40.

3.1.5 Assessment Details

A total of 21 subjects were used in the experiment. All subjects were students at the University of Illinois at Urbana-Champaign. They Major area of studies of the subjects varied but most were of engineering background. For 12 of the subjects the correlation coefficient describing the relationship between the random variables X and Y was set at 0.9 and for 9 of the subjects the correlation coefficient was set to 0.1. Every subjects belief about the marginal distribution of X and Y was assessed along with the iso-probability contours with constant probabilities 0.5 and 0.25.

3.2 Observations

3.2.1 Comments by Subjects During Assessment

During the assessment several qualitative comments were made by the subjects about the assessment. Many of these provide insight into how the subjects felt about the assessment process and how the assessment might be done more easily. Here some of the relevant comments made are paraphrased.

- One subject said that they would like to see a plot of the data during the assessment stage.
- A subject felt that there was a disconnect between guessing the numbers in the learning section and the guessing of joint probabilities.
- Subject said that betting on the wheel was more fun than betting on the variable.
- Five subjects asked to redo the learning session after the assessment of the marginals and contours had begun.
- When asked those subjects that did multiple learning sessions said it was useful.

3.2.2 Qualitative Observations

There were many occurrences during the assessments that provided interesting insight into the feelings of the subjects and also the motivations behind their assessments. Several of these observations are written below.

- Many of the subjects seemed tired by the end of the assessment.
- In one case the subject seemed to be concerned with what answer would please the assessor.

- A few of the subjects liked to have scratch paper to work with during the assessment.
- Two of the subjects were given a paper with ranges of x and y values on them. The subjects tended to evenly distribute the probability between even ranges of values. This made the correlation attained from their assessments largely negative.

3.3 Results

3.3.1 Marginal Distributions

The assessed points of the marginal distributions were fit using a normal distributions in Matlab to attain the full cumulative distributions of the individual random variables according to each subjects beliefs. These fits were preformed using a least squares technique. They provided the best fit mean and standard deviation of each subject. Some of the subjects believed the distribution of X and Y were the same, in this case all of the points were fit to a single distribution for both variables. The mean and standard deviation for each subject is given in table 3.1. A few of the marginal fits are given in Figure 3.1

3.3.2 Isoprobability Contours

In order to characterize the correlation coefficient using the isoprobability contour, the data must be normalized by the subjects belief about the mean and standard deviation of the marginal distributions. This process and the relationship between the normalized isoprobability contours are described in detail in Abbas et al. [12] The fits to the contours were done for both the case that the distribution of X and Y were considered to be the same (ie. have the same marginal distribution) and for the case where the marginal distributions were different for the two random variables. Once the data was normalized it was fit using a least squares program in Matlab. Some examples of these fits are given in Figure 3.2, where the distributions of X and Y were considered separately. The Table 3.2 shows the values for

Table 3.1: The assessments of the marginal distributions were fit with a normal distribution to get the mean and standard deviation of each variable

| Subject | X | | Y | | Combined $X&Y$ | |
|---------|-------|----------|-------|----------|----------------|----------|
| | Mean | St. Dev. | Mean | St. Dev. | Mean | St. Dev. |
| 1 | 50.05 | 19.49 | 49.86 | 21.13 | 50.03 | 20.01 |
| 2 | 47.06 | 17.30 | 45.90 | 25.14 | 46.56 | 20.76 |
| 3 | – | – | – | – | 51.49 | 28.77 |
| 4 | 51.90 | 22.52 | 38.0 | 6.24 | 50.94 | 23.62 |
| 5 | 56.52 | 18.17 | 43.63 | 11.38 | 54.87 | 21.84 |
| 6 | – | – | – | – | 51.52 | 17.60 |
| 7 | 49.58 | 15.85 | 49.09 | 18.79 | 49.77 | 17.52 |
| 8 | – | – | – | – | 52.67 | 16.05 |
| 9 | – | – | – | – | 49.34 | 18.00 |
| 10 | 52.07 | 13.25 | 50.91 | 15.55 | 51.36 | 14.50 |
| 11 | – | – | – | – | 44.34 | 17.36 |
| 12 | – | – | – | – | 50.45 | 17.74 |
| 13 | 52.46 | 17.76 | 55.98 | 23.96 | 54.63 | 21.83 |
| 14 | 50.33 | 16.49 | 48.10 | 18.08 | 49.28 | 17.39 |
| 15 | 50.55 | 14.32 | 49.16 | 50.22 | 50.22 | 16.42 |
| 16 | 50.22 | 12.58 | 46.43 | 14.48 | 49.31 | 14.36 |
| 17 | 51.96 | 22.45 | 38.00 | 12.23 | 50.78 | 24.56 |
| 18 | 49.67 | 17.18 | 44.26 | 13.91 | 47.42 | 16.48 |
| 19 | – | – | – | – | 48.29 | 15.88 |
| 20 | 47.72 | 14.38 | 48.19 | 14.72 | 47.94 | 14.51 |
| 21 | – | – | – | – | 52.10 | 20.74 |

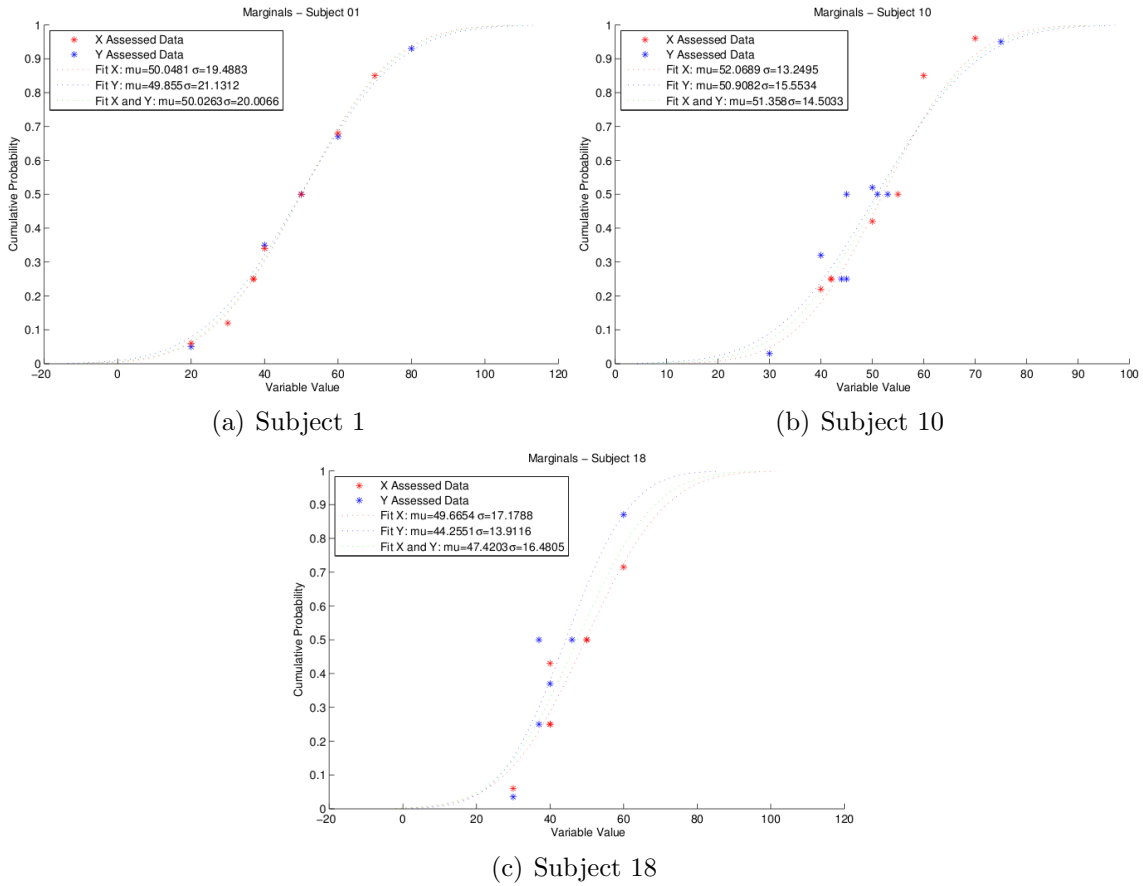


Figure 3.1: A sampling of the marginal assessment data and fits

the correlation coefficient as fit by the Matlab program for each subject. To compare the results between the two contours for each subject the results are shown with the Correlation resulting from the 50% contour in the horizontal axis and the 25% contour in the vertical axis in Figures 3.3 and 3.4, with separated and combined X and Y marginals, respectively.

When the correlation fit to the 50% and 25% contours are looked at together, it was found that these two values were correlated. Meaning that a subject that had reported a large correlation through the 50% contour tended to report a high correlation in the assessed 25% contour. When X and Y marginals are combined, the correlation between the two assessed values is 41.1% with a 6.4% probability of the no correlation hypothesis. When X and Y marginals are fit separately, the correlation between the 50% and 25% contours is 66.0%

Table 3.2: Correlation coefficients resulting from the fit to the iso-probability contours

| Subject | Actual Corr. | Combined $X&Y$ Contours | | Separate $X&Y$ Contours | |
|---------|--------------|-------------------------|------------|-------------------------|------------|
| | | $P = 0.5$ | $P = 0.25$ | $P = 0.5$ | $P = 0.25$ |
| 1 | 0.9 | -0.27 | 0.71 | -0.24 | 0.67 |
| 2 | 0.9 | 0.84 | -0.23 | 0.84 | -0.17 |
| 3 | 0.9 | 0.77 | -0.14 | 0.77 | -0.14 |
| 4 | 0.9 | -0.99 | -0.64 | -0.99 | -0.99 |
| 5 | 0.9 | 0.62 | -0.29 | -0.99 | -0.99 |
| 6 | 0.9 | 0.68 | 0.22 | 0.68 | 0.22 |
| 7 | 0.9 | 0.63 | 0.56 | 0.54 | 0.48 |
| 8 | 0.9 | 0.99 | 0.99 | 0.99 | 0.99 |
| 9 | 0.9 | 0.52 | 0.39 | 0.52 | 0.39 |
| 10 | 0.9 | -0.78 | 0.66 | -0.62 | 0.68 |
| 11 | 0.9 | -0.17 | 0.08 | -0.17 | 0.08 |
| 12 | 0.9 | 0.72 | 0.72 | 0.72 | 0.72 |
| 13 | 0.1 | 0.99 | 0.28 | 0.99 | 0.24 |
| 14 | 0.1 | -0.56 | -0.45 | -0.64 | -0.48 |
| 15 | 0.1 | -0.33 | -0.25 | -0.3 | -0.37 |
| 16 | 0.1 | -0.99 | -0.47 | -0.99 | -0.83 |
| 17 | 0.1 | 0.4 | 0.06 | -0.99 | -0.9 |
| 18 | 0.1 | -0.09 | -0.3 | -0.79 | -0.55 |
| 19 | 0.1 | -0.99 | -0.35 | -0.99 | -0.35 |
| 20 | 0.1 | -0.47 | 0.58 | -0.44 | 0.58 |
| 21 | 0.1 | 0.21 | 0.31 | 0.21 | 0.31 |

with only a 0.1% chance of satisfying the no correlation hypothesis.

3.4 Conclusions

Although the overall principle behind iso-probability curves seemed to work, the assessment of these curves proved to be rather difficult. The correlation between the assessed 50% contour and the 25% contour shows that there is some recognition of the concept of correlation in the answers given by the subjects. The mind set of the subject seemed to play the largest role in the success of their assessment. Overall the subjects tended to give lower correlations through there assessments than the actual correlation of the variables. This could be due to a belief that the correlation values were lower than in actuality, or could be due to the assessment technique causing distorting the believed correlations. The subjects seemed to have a hard time relating the learning number game to the assessment and thus had a difficult time relating the probabilities of different bets.

3.5 Possible Future Directions

The inability of the subjects to related the learned variable behavior to the gambles of the assessment may be remedied by using values that the subjects can relate to. This can be done by having some back story describing what the variables represent. This, of course, comes with the negative attribute that the subjects may have some preconceived notion of how the variables in the back story should behave, but this can be minimized by choosing variables that the subject cannot know anything about, for instance the stock price of two imaginary companies.

The difficulty of the subjects can also be minimized by using the same subjects for multiple assessments with variables of different correlations. This would be helpful in determining if this assessment technique becomes better with practice.

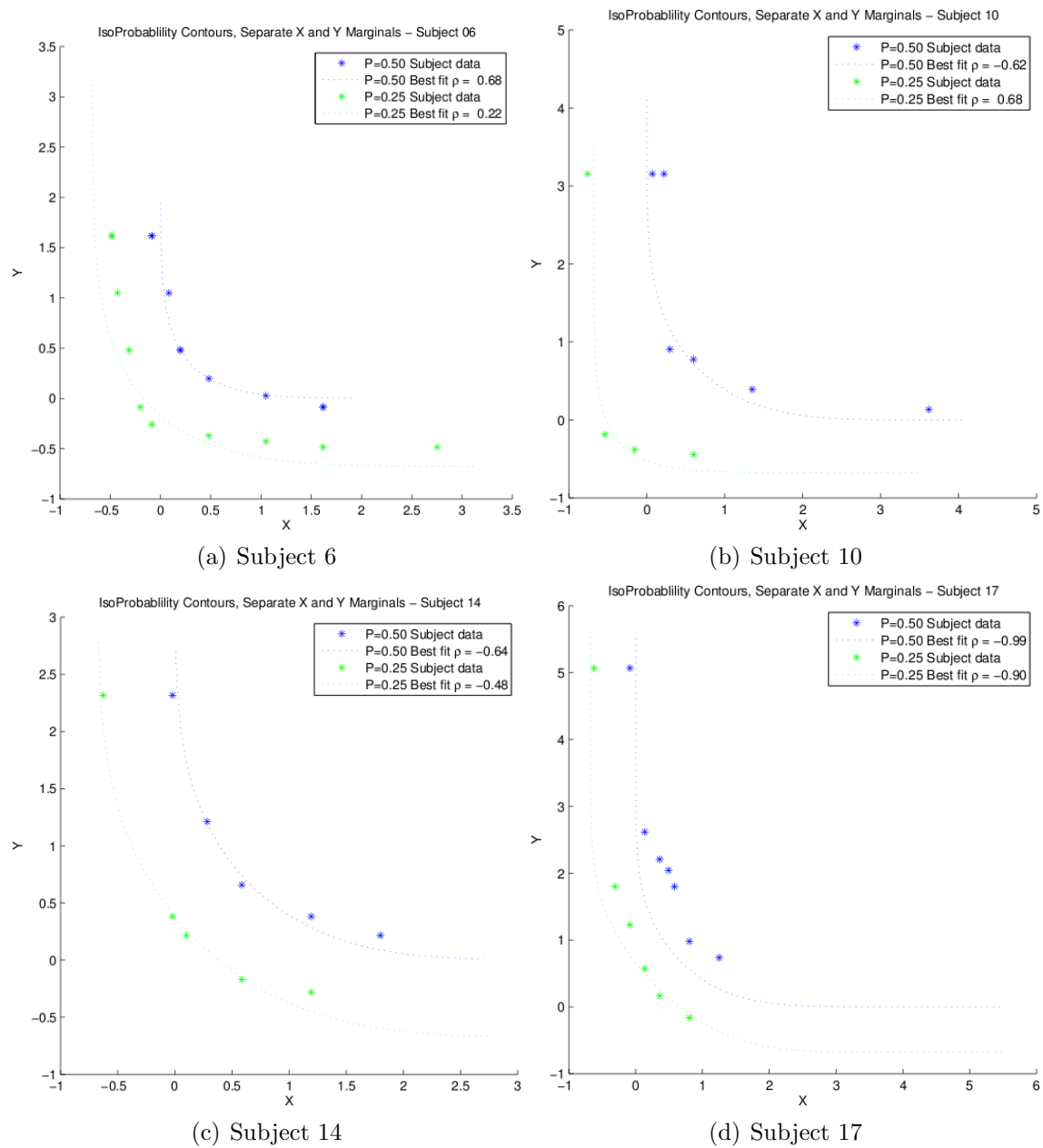


Figure 3.2: A sampling of the fits to normalized assessments of the iso-probability contours

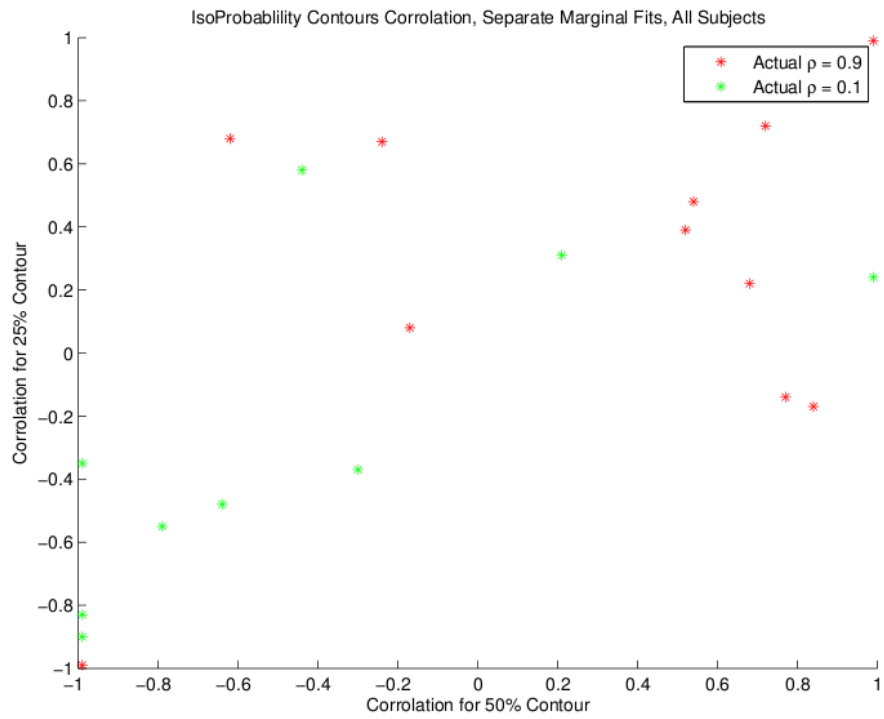


Figure 3.3: Correlation values attained from a fit to both assessed contours using separated margins

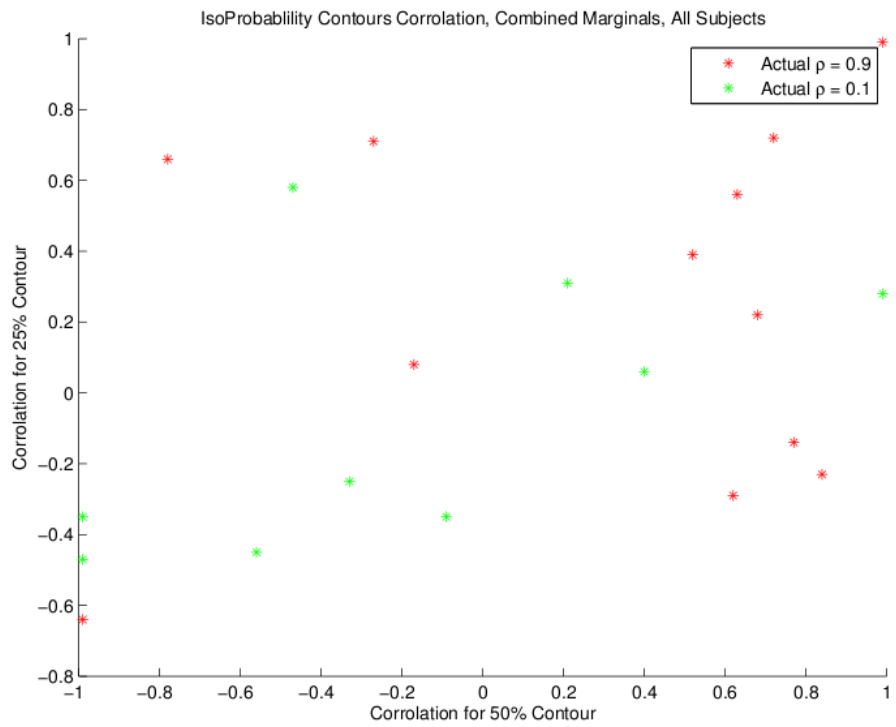


Figure 3.4: Correlation values attained from a fit to both assessed contours using combined margins

CHAPTER 4

VARIABLE WEIGHTING FUNCTION

4.1 Introduction

4.1.1 Motivation

The probability weighting function has been found to describe the nonlinear weights a decision maker places on uncertainties when facing a choice. The function manipulates the belief of a subject regarding the outcome of a random variable to produce the weighting the subject uses to make decisions for which the outcome relies on that variable. In this way the assessed cumulative distribution function describing a subjects belief about a random variable can be converted into the a cumulative weighting the subject would put on that variable. Since the input to the probability weighing function is probability and the output is the weighting, the determination of a value of the random variable that would produce a specific weighting is an iterative procedure. By instead using a variable weighting function in place of the probability weighting function, this procedure can be completed in a single step. This work will use the current state of knowledge of the probability weighting function in order to determine possible variable weighting functions. In order to facilitate this process and to determine the characteristics of the variable weighting function the properties various options for the functional form of the variable weighting function will be explored. Experimentally assessed continuous probability functions will be used to determine the validity of different options.

4.1.2 Background

Cumulative Prospect Theory

Prospect theory was developed in 1979 by Kahneman and Tversky [34] to describe suboptimal decision making observed in real life situations. Although expected utility theory had dominated the area of normative decision making, many common decisions were systematically found to contradict the axioms of expected utility. One example of this is the “Allais Paradox”, [35] which was derived to show that the same person can be both risk seeking and risk adverse for the same decision framed differently. This is analogous to the idea that the same person might purchase insurance (risk adverse behavior) and buy a lottery ticket (risk seeking behavior). In prospect theory a decision maker assesses value in terms of gains or losses from an anchored point of wealth, or a status quo.

In order to incorporate these ideas into expected utility theory Quiggin proposed the use of decision weights in place of probabilities, which account for the weight that a decision maker places on a specific outcome. [36] These ideas were the basis for rank dependent expected utility theory, originally named anticipated utility theory by Quiggin. This model had the advantage of being able to accommodate a large number of outcomes by transforming the entire cumulative probability function, rather than just individual probabilities as in prospect theory. In 1992, Tversky and Kahneman modified prospect theory, in light of these advancements, to develop cumulative prospect theory, for which Kahneman received the Nobel Memorial Prize in Economics in 2002.

Cumulative prospect theory recognizes both the modified value of a prospect based on its relationship to the reference state of the decision maker as well as the nonlinearity in weighting of probabilities by the decision maker. In the model described by Tversky and Kahneman the decision maker observes two natural boundaries in the spectrum of uncertainty, complete certainty and impossibility. [13] The decision maker overweights deviation from these endpoints. For example a small increase in probability, say 0.1, is considered

more important to the decision maker when increasing from 0.1 to 0.2 then and increase in probability from 0.3 to 0.4. An accurate description of the probability weighting phenomena is necessary to accurately describe the choices of a real decision maker.

Probability Weighting Function

In order to account for the nonlinearity in probability a weighting function, $w(p)$ was incorporated into cumulative prospect theory, where p is the understood probability on the interval $[0,1]$ and $w(p)$ is the weighting that the value associated with that uncertainty is given. Understanding the form of the probability weighting function allows greater understanding of how a decision maker views the uncertainties associated with each decision.

Based on the goal of the probability function it is easy to understand why the general shape of the probability weighting function is inverse-S-shaped. Approaching zero probability the weighting function is no less than the probability, $w(p) \geq p$, and the function is concave. Approaching a probability of one the converse is true, $w(p) \leq p$, and the function is convex. The slope of the probability weighting function is greater at the end points, which describes the diminishing sensitivity, termed by Tversky and Kahneman, as the probability moves away from complete certainty or complete impossibility.

Gonzalez and Wu have presented a physiological interpretation of the probability weighting function.[37] Two important properties of the probability weighting function describe the characteristics of the decision maker, the curvature and elevation of the function. The curvature indicates the degree to which the decision maker overweights changes in probabilities near the endpoints of the function, when $p = 0$ or 1 . The discriminability of the decision maker is defined as the amount to which a change in probability effects the weighted probability, or the slope of the probability weighting function. Gonzalez and Wu use the example of two decision makers that represent the extremes in discriminability. The first has a probability function that is described by a step function $w(p) = 0$ and 1 when $p = 0$ and 1 , respectively, and is 0.5 every where else. The second has a linear

probability weighting function, $w(p) = p$. In the first case the subject treats all uncertain events, that are not completely certain or impossible, equally and cannot discern between their likelihood. The second decision maker is sensitive to all probability changes equally and is indifferent between probability changes across the spectrum of probability. This decision maker would act according to expected utility theory, always maximizing their expected utility. In addition to curvature, the elevation of the probability weighting function can provide some insight into the psychology of the decision maker. When a decision maker faces a gamble based on an uncertainty they are more attracted to the gamble the larger the value of the weighting function is at the respective probability. The attractiveness at a given probability is associated to the value of that weighting function at that probability. In the extreme examples discussed above, the decision maker with the step function and the decision maker with a linear weighting function are equally attracted to gambles of probability 0.5. At lower probabilities the decision maker with a step function is more attracted to the gamble and at higher probabilities the opposite is true. It is not necessarily the case that $w(0.5) = 0.5$, and in fact is not often the case. For example, a decision maker with a probability weighting function that has a step form where $w(p) = 0.4$ for all probabilities inside the end points. When compared to the decision previously described step function, this decision maker's probability weighting function has the same discriminability, but is always less attracted to gambles with positive outcomes that are uncertain.

Several studies have sought out the functional form of the probability weighting function.[38, 39] Several functions have been investigated and are generally based on one or two parameter values, which have been estimated by experimental studies involving real gambles.[13, 38, 39, 40, 41, 42] The probability weighting function is found to have similar features in each study. These features are described by Prelec as having the general behavior predicted by cumulative prospect theory. At low probabilities the weighting functions are concave and have values greater than the linear weighting function, while at high probabilities it is convex and has is always less than the linear weighting function. In addition Prelec found that there

were some unpredicted characteristics of the weighting function, in particular the asymmetry of the weighting function. In the studies addressed all the weighting functions crossed the linear weighting function below the middle, at approximately one third probability. (i.e. $w(1/3) \approx 1/3$) [43]

The simplest functions that have been used to characterize the phenomenon of probability weighting are dependent on a single parameter. One of the more commonly used single parameter functions, originally introduced by Tversky and Kahneman in 1992 (TK-1)[13] is shown in Table 4.1 along with all the probability weighting functions examined in this work. The value for the parameter γ has been experimentally determined for the common decision maker to be in the range of 0.56 and 0.91.[13, 44, 40, 41, 38] Another one parameter function that is often used was introduced by Prelec (Pr-1).[43] The parameter defining the shape of this function, γ , has a range of 0.41 to 1 in experimental studies.[38, 40, 45] One advantage of this function is that it always crosses the linear weighting function at $p = 1/e \approx 0.37$, which is in line with the observed characteristics of the probability weighting function. Of course this also means that there is no adjustment for the elevation of the function. Prelec also suggested a similar, two parameter function (Pr-2) in the same work.[43] The second parameter allows the elevation to be adjusted of the function to be adjusted, although this parameter affects the curvature as well. Typical values for the parameters have been acquired by Bleichrodt and Pinto, $\delta = 1.08$ and $\gamma = 0.53$. [38] The most thoroughly studied two parameter function was introduced by Goldstein and Einhorn (GE-2).[46] Many investigators have reported values for the parameters of this function. The range of γ has been reported from 0.44 to 0.84, and the reported range of δ is 0.65 to 1.35. One of the more desirable attributes of the Goldstein and Einhorn function is that γ controls primarily the curvature of the function, while δ controls the elevation.

The two parameter functions have been found to fit the experimental data better than one parameter functions, although the improved fit is only slightly better when it data is aggregated. The simplicity of one parameter functions merits their use in certain circum-

stances, though the two parameter function is used more commonly due to the ability to fit more types of decision makers as individuals.

The weighting function differs for uncertainty related to positive prospects and negative prospects, denoted $w^+(p)$ and $w^-(p)$, respectively.[13] The most distinct difference between weighting functions for positive and negative prospects is the elevation. While the weighting function for positive prospects nearly always crosses the linear weighting function between $p = 0.3$ and 0.4 , the weighting function for negative prospects tends to cross at $p \approx 0.6$. As pointed out by Booij et al, this displays the pessimism of the decision maker and indicates that they prefer to not gamble. Other factors that may effect the probability weighting function include the magnitude of the outcomes [47] (contradicted by findings of Booij et al. [48]) and the type of effect invoked by the prospects [49].Characteristics of the decision maker can also effect the probability weighting function, such as gender [50, 48], age [48], or level of education [48, 45, 51].

A study by Van de Kuilen and Wakker showed that the parameters of the weighting function can be interpreted differently by different methods of assessment.[42] A more specific observation by Goeree et al. found in an experimental study bases on bidding, that the type of decision can affect function. [52]

4.1.3 Variable Weighting Function

The usefulness of the probability weighting function, w_p , is in adjusting a cumulative distribution, $F(v)$, to the corresponding weighted cumulative distribution, $F_t(v)$, such that $F_t(v) = w_p(F(v))$. The function w_p maps from a point in the original distribution, (v, p) , to a point in the weighted distribution, (v, p_w) , at a constant variable value, v . One can imagine a function that can provide the same weighed cumulative distribution, $F_t(v)$, by mapping from a point in the original distribution, (v, p) , to a point at constant probability in the weighted distribution, (v_p, p) . This function is called the variable weighting function,

w_v . It allows the variable value to be assessed and weighted directly for a given probability, such that $p = F_t(w_v(v)) = F(v)$. The variable weighting function is described thoroughly in a work in preparation by Abbas, Budescu, Haggerty, and Wu.[53]

Definition

The following is a summary of the definition provided by Abbas, Budescu, Haggerty, and Wu.[53] Assume a probability weighting function, $w_p(p)$ that is continuous and strictly monotonic such that the inverse function, w_p^{-1} , exists. A continuous cumulative probability distribution, $F(v)$, can be weighted by the probability weighting function to provide a weighted cumulative distribution,

$$F_t(v) = w_p(F(v)) \quad (4.1)$$

The variable weighing function, $w_v(v)$ is a method for obtaining the same weighted distribution, $F_t(v)$, by modifying the variable value at a constant probability.

Proposition 1. *For every continuous and strictly monotonic probability weighting function, $w_p(p)$, operating on a continuous and strictly monotonic cumulative distribution function, $F(v)$, there exists a unique equivalent variable weighting function, $w_v(v)$, satisfying $w_v(v) = F^{-1}(w_p(F(v)))$.*

Proof. Let (v_0, p) be a point on the assessed cumulative distribution function, $F(v)$. The probability weighting function modifies the probability, $p = F(v_0)$, to obtain the weighted probability, $p_w = w_p(p)$, such that (v_0, p_w) is a point on the weighted cumulative probability distribution function, $F_t(v)$.

$$F_t(v_0) = p_w = w_p(p) = w_p(F(v_0)) \quad (4.2)$$

If the value p is held constant the variable value, v_0 , can be adjusted by the variable weighting function, such that $v_w = w_v(v_0)$, where the point (v_w, p) lies on the weighted

distribution, $F_t(v)$. Hence, $F(v_0)$ and $F_t(v_w)$ correspond to the same cumulative probability, p . *i.e.*

$$F_t(v_w) = F(v_0) = p \quad (4.3)$$

where

$$v_w = w_v(v_0) \quad (4.4)$$

From Equation 4.3

$$v_w = F_t^{-1}(p) = F_t^{-1}(F(v_0)) \quad (4.5)$$

From Equations 4.4 and 4.5

$$v_w = w_v(v_0) = F_t^{-1}(p) \quad (4.6)$$

Take the inverse function of both sides of Equation 4.2,

$$F_t^{-1}(t) = F^{-1}(w_p^{-1}(t)) \quad (4.7)$$

From Equations 4.6 and 4.7

$$v_w = w_v(v_0) = F_t^{-1}(p) = F^{-1}(w_p^{-1}(p)) = F^{-1}(w_p^{-1}(F(v_0))) \quad (4.8)$$

Hence

$$w_v(v) = F^{-1}(w_p^{-1}(F(v))) \quad (4.9)$$

□

Such variable weighting function can be used to determine the adjusted value of a variable due to the effect of probability weighting as described by cumulative prospect theory and can reduce the number of assessed points on a cumulative distribution in order to determine the corresponding variable value to a given probability.

Characteristics

At the point at which the probability weighting function intersects the linear weighting function, the cumulative distribution is unaltered by weighting. At this point, where $w(p) = p$, there is a corresponding unaltered weighting point at which the variable weighting function will intersect the linear weighting function, $w(\hat{v}) = \hat{v}$. This implies that the variable weighting function has a characteristic of elevation similar to that of the probability weighting function. The generally accepted inverse-S-shape of the probability weighting function (noted exception in studies by Goeree et al. as well as Van de Kuilen and Wakker [52, 42]) indicates that there is characteristic curvature in the variable weighting function. The domain of the variable weighting function can be assigned to two sets, separated by the unaltered weighting point. Normalized variable values less than the unaltered weighting point should reduce the variable value, $w(\hat{v}) < \hat{v}$, while in domain set of values above the unaltered weighting point the opposite is expected, $w(\hat{v}) > \hat{v}$. This will produce an S-shaped function.

Proposition 2. *If the probability weighting function is inverse S-shaped and the cumulative distribution function is continuous and strictly increasing, then the variable weighting function must be S-shaped.*

Proof. Let $p = F(v)$ be the probability to be weighted, $p_w = w_p(p)$ is the weighted probability, and p^* be the inflection point of the probability weighting function, where $0 < p^* < 1$. For an inverse S-shaped probability weighting function, when $p^* > p > 0$, w_p is concave, and when $1 > p > p^*$, w_p is convex. We can examine these two sections of w_p separately.

When $p^* > p > 0$, w_p *i.e.*, the concave part of the probability weighting function, $w_p(p)$, where

$$p_w > p \tag{4.10}$$

Since w_p^{-1}, F^{-1} are strictly increasing, applying w_p^{-1} to both sides of Equation 4.10 results in

$$w_p^{-1}(p_w) > w_p^{-1}(p) \tag{4.11}$$

and applying F^{-1} to both sides of Equation 4.11 results in

$$F^{-1}(w_p^{-1}(p_w)) > F^{-1}(w_p^{-1}(p)) \quad (4.12)$$

Where $p_w = w_p(p)$,

$$F^{-1}(w_p^{-1}(p_w)) = F^{-1}(w_p^{-1}(w_p(p))) = F^{-1}(p) = v \quad (4.13)$$

And

$$F^{-1}(w_p^{-1}(p)) = F^{-1}(w_p^{-1}(F(v))) = w_v(v) \quad (4.14)$$

From Equation 4.12,

$$x > w_v(x) \quad (4.15)$$

When $1 > p > p^*$, w_p *i.e.*, the convex part of the probability weighting function, $w_p(p)$, where

$$p_w < p \quad (4.16)$$

Since w_p^{-1}, F^{-1} are strictly increasing, applying w_p^{-1} to both sides of Equation 4.16 results in

$$w_p^{-1}(p_w) < w_p^{-1}(p) \quad (4.17)$$

and applying F^{-1} to both sides of Equation 4.17 results in

$$F^{-1}(p) = F^{-1}(w_p^{-1}(p_w)) < F^{-1}(w_p^{-1}(p)) \quad (4.18)$$

From inserting the results from Equations 4.13 and 4.14 into Equation 4.18 results in

$$x < w_v(x) \quad (4.19)$$

From these two cases we can see that for, $v^* = F^{-1}(p^*)$, when $v^* > v > v_{min}$, $v > w_v(v)$ and when $v_{max} > v > v^*$, $v < w_v(v)$. Thus the variable weighting function, $w_v = F^{-1}(w_p^{-1}(F(v)))$, is S-shaped (first convex, then concave) given the probability function is inverse S-shaped and the cumulative distribution function is continuous and strictly monotonic.

□

In order to determine the typical shape of this function a functional form will be assumed. The input variable, v , and the output, $w(v)$, of the variable weighting function will be normalized over the range of possible values for the variable. The normalized input, \hat{v} , will be set by $\hat{v} = \frac{v-v_{min}}{v_{max}-v_{min}}$ and the output, $w(\hat{v})$, can be used to extract the weighted variable, v_w , by $v_w = (v_{max} - v_{min})w(\hat{v}) + v_{min}$, where v_{min} and v_{max} are the minimum and maximum possible values of v , as determined by the decision maker. This will insure that the range and domain of the function is $[0, 1]$, just as in the probability weighting function.

4.2 Methods

4.2.1 Experimentally Assessed Cumulative Probability Distributions

Experimental data collected from 103 subjects concerning their belief about the distribution of a random variable was used to obtain typical variable weighting functions. The subjects used binary decisions to determine the point at which they were indifferent between two bets, one on a probability wheel and the other concerning the value of the random variable of interest. During half of the assessed points the probability wheel was fixed and the maximum value of the random variable was adjusted till indifference (ie. a fixed probability assessment). For the rest of the assessment the maximum quantity of the random variable was fixed while the value of the probability wheel was adjusted till the subject was indifferent between the two bets (ie. a fixed variable assessment). The random variable used in the study was either

the temperature in Palo Alto or the value of the DOW Jones at closing, on a day two weeks after the assessment. More information on the experimental data is published elsewhere.[29] For the following calculations the data assessed from the subjects by both fixed probability and fixed variable method was combined. The assessed cumulative distributions those that were not monotonically increasing were removed from this study based on the value of Kendall's τ once the fixed probability data and the fixed variable data was combined. Of the 103 subjects, 64 produced distributions with the combined methods of assessment with a Kendall's $\tau > 0.8$.

4.2.2 Fitting Procedure

The cumulative distributions assessed by the subjects were fit with a beta distribution. Using this fit distribution, the weighted probability distribution was obtained using an assumed probability weighting function. Each of the cumulative probability values assessed provides an assessed variable value and the weighted variable value for constant cumulative probability was extracted from the curve derived from the assumed probability weighing function. The form of the variable weighting function was assumed so that the parameters of the form could be fit to these points. The fit was chosen such as to minimize the the squared residuals.

4.2.3 Choice of functional form

Probability Weighting Function

The extensive research into the probability weighting function has allowed many researchers to create functions that represent the averaged probability weighting function across a population. Drawing from this research four of the commonly used functional forms of the probability weighting function have been used. For this work they will be abbreviated as TK-1, Pr-1, Pr-2 and GE-2. These functions are shown in Table 4.1. The one parameter functions, TK-1 and Pr-1, depend on the parameter γ . While the two parameter functions,

Pr-2 and GE-2, depend on γ and δ . The median value of these parameters from several studies was used to generate four probability weighting functions. These parameter values are given in Table 4.2. It is expected that the two parameter functions will outperform the one parameter forms due to a larger degree of freedom, but the simplicity of a one parameter form merits its use. This study intends to observe the implications of using a one parameter form instead of a two parameter form. The characteristics of each of the functions used with the assumed parameters can be seen in Figure 4.1. The functions have each of the characteristics expected of the probability weighting function, although the Prelec two parameter function (Pr-2) is slightly different from the rest. Each of the one parameter functions (TK-1 and Pr-1) and the linear in log odds function (GE-2) cross the linear weighting function at nearly the same probability ($p \approx 0.37$). Whereas the Pr-2 function crosses the linear weighting function at a lower probability ($p \approx 0.3$). In addition the Pr-2 function has lower weighting of higher probabilities than the other functions, which have approximately the same weighting at all probabilities.

Variable Weighting Function

Due to the expected shape and characteristics of the variable weighting function, the same functions that are used to describe the probability weighting function are good candidates for the functional form of the variable weighting function. In addition to these functions, the Kumaraswamy distribution will be used.[54] This distribution was chosen for its S-shaped cumulative distribution as well as its bounded domain. Each of the functions used to fit the variable weighting function are shown in Table 4.1.

4.2.4 Parameter Sensitivity

In reality the probability weighting function can be different for every individual. The sensitivity of the parameters of the variable weighting function are likely to be related to the

Table 4.1: The various functional forms used to represent the probability weighting function or the variable weighting functions. The input z was used in place of probability, p , or normalized variable, \hat{v}

| Abbreviated Name | Weighting Function | Source |
|------------------|--|--------------------------------|
| TK-1 | $w(z) = \frac{z^\gamma}{(z^\gamma + (1-z)^\gamma)^{\frac{1}{\gamma}}}$ | Tversky and Kahneman 1992[13] |
| Pr-1 | $w(z) = \exp(-(-\ln z)^\gamma)$ | Prelec 1998[43] |
| Pr-2 | $w(z) = \exp(-\delta(-\ln z)^\gamma)$ | Prelec 1998[43] |
| GE-2 | $w(z) = \frac{\delta z^\gamma}{\delta z^\gamma + (1-z)^\gamma}$ | Goldstein and Einhorn 1987[46] |
| Ku-2 | $w(z) = 1 - (1 - z^\delta)^\gamma$ | Kumaraswamy 1980[54] |

Table 4.2: The parameters used to represent a typical probability weighting function was determined by taking the median of the parameter values from several studies.

| Abbreviated Name | δ | γ |
|------------------|----------|----------|
| TK-1 | | 0.71 |
| Pr-1 | | 0.64 |
| Pr-2 | 1.08 | 0.53 |
| GE-2 | 0.82 | 0.60 |

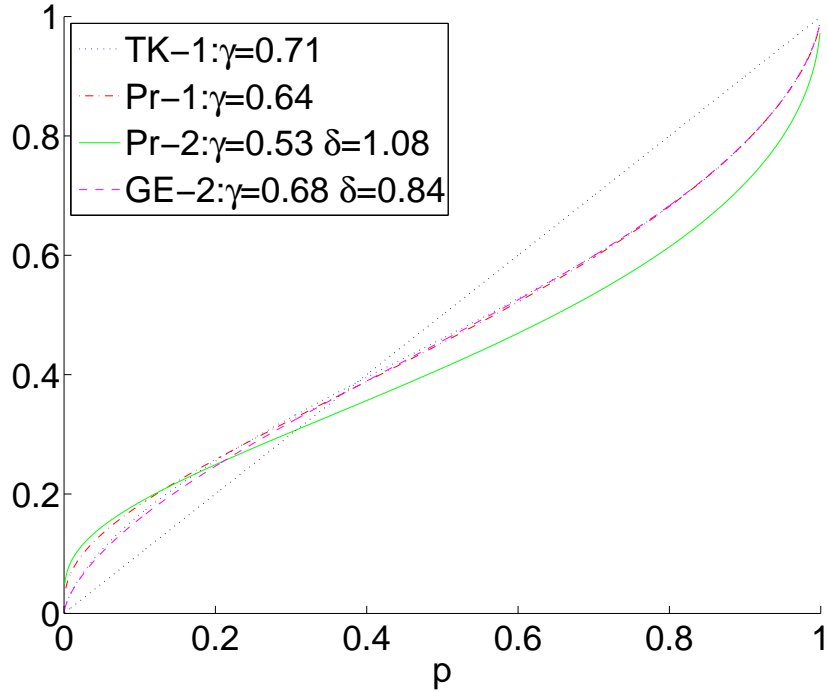


Figure 4.1: Each of the commonly used functional forms of the probability weighting function are shown with the assumed parameters.

curvature and elevation of an individual's probability weighting function. In order to test this and learn more about the intrinsic effects on the variable weighting function the probability weighting function was varied. The linear in log odds form, GE-2, of the probability weighting and variable weighting functions was used, due to the independent control of curvature, by varying γ , and elevation, by varying δ . The input parameters, into the probability weighting function were varied through the entire domain $\gamma = 0.3$ to 0.9 and $\delta = 0.3$ to 1.5 . This range was chosen to represent the range of values found to fit the probability weighting functions of individuals in a study by Gonzalez and Wu in 1999.[37] For each combination of parameters in the probability weighting function the each of the subject's cumulative distribution functions was used to fit the corresponding variable weighting function. The average parameters for the best fit variable weighting function were determined as a function of the input parameters for the probability weighting function.

4.3 Results

4.3.1 Probability Weighted Cumulative Distribution Functions

Each of the subjects cumulative distribution functions was adjusted to account for probability weighting using typical functions given in section 4.2.3. Some examples of the resulting functions are shown in Figure 4.2. The Prelec two parameter function (Pr-2) tended to differ from the other weighted probabilities, as might be expected from the differences observed when compared to the other functions as discussed in section 4.2.3.

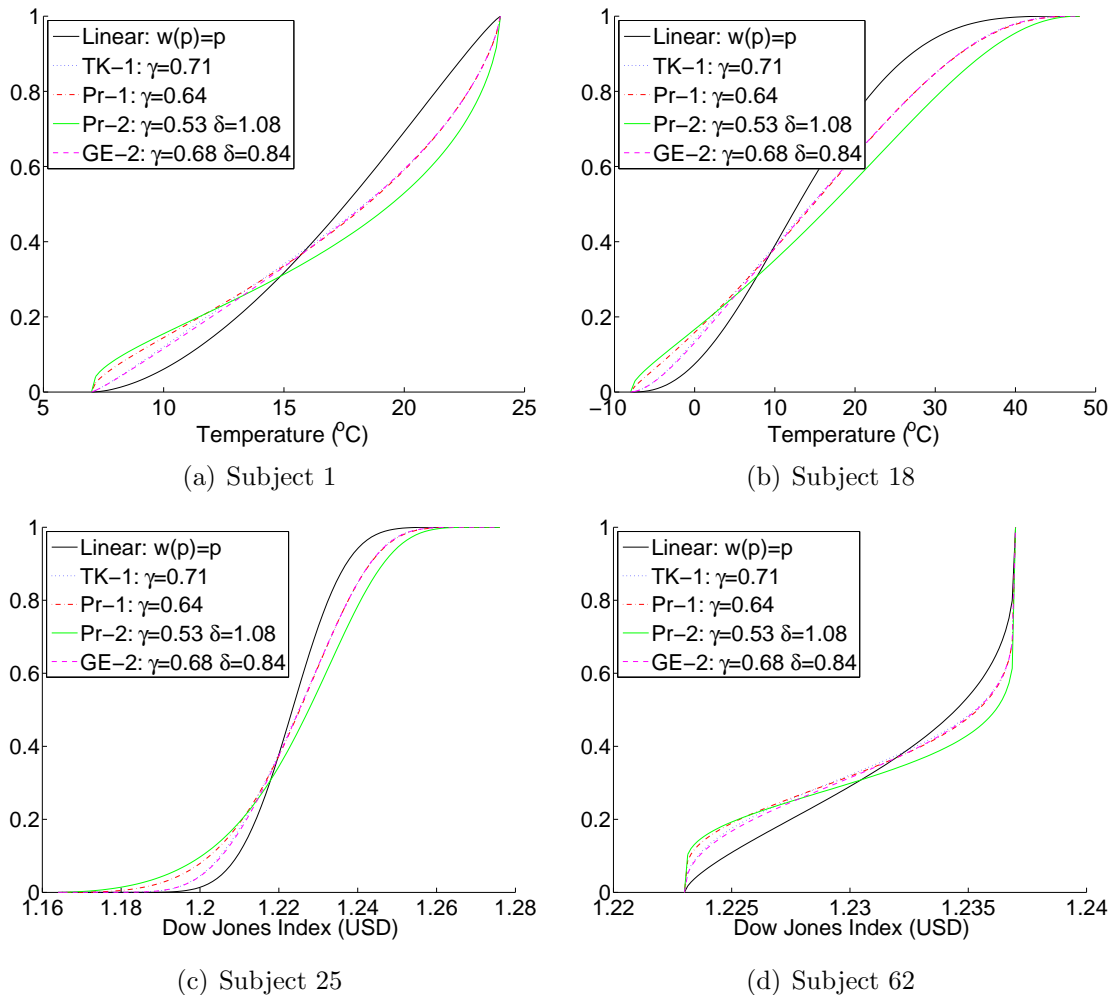


Figure 4.2: A sampling of the assessed cumulative distribution functions adjusted for probability weighting.

4.3.2 Error in Fitting Variable Weighted Cumulative Distribution Functions

The derived weighted distribution functions were used to determine the variable weighting function parameters. Each of the five variable weighting function were fit to each of the four probability weighting functions in order to compare their practical uses. The best fit function was determined by the minimizing the sum of the squared residuals. The characteristic error of each subject's fit was determined by taking the square root of the average squared residual and is then described as a percentage of the entire range of the variable assessed. The average characteristic error percentage for each combination of probability weighting function and the variable weighting function is given in Table 4.3. One parameter variable weighting functions show significantly larger errors than the two parameter functions. In particular the the function proposed by Tversky and Kahneman as a probability weighting function, TK-1, does a poor job when used as a variable weighting function. The average characteristic error is greater than 4%. Leaving the Prelec function, Pr-1, as the best performing one parameter variable weighting function. The average characteristic error does not provide a clear segregation of performance of the two parameter variable weighting functions. The error depends on the form used to simulate the probability weighting function.

Table 4.3: The average characteristic fitting error is given as a percentage of the variable range assessed.

| | Variable Weighting Function | | | | |
|--------|-----------------------------|----------|----------|----------|----------|
| $w(p)$ | TK-1 | Pr-1 | Pr-2 | GE-2 | Ku-2 |
| TK-1 | 4.1(1.9) | 2.7(1.5) | 0.6(0.3) | 0.3(0.1) | 0.4(0.1) |
| Pr-1 | 4.5(1.9) | 3.0(1.7) | 0.2(0.1) | 0.6(0.2) | 0.8(0.3) |
| Pr-2 | 7.6(3.2) | 3.7(2.8) | 0.3(0.2) | 0.8(0.3) | 1.1(0.3) |
| GE-2 | 4.1(1.9) | 2.4(1.4) | 0.5(0.3) | 0.2(0.1) | 0.3(0.1) |

In order to further assess the best fitting form of the variable weighting function, the error of each subject was compared for different combinations of probability weighting function

and variable weighting functional form. A tally of the best fitting one parameter function and the best fitting two parameter variable weighting function was kept for each of the fitted probability weighting functions. The one parameter forms were separated from the two parameter forms because they always had a larger error, which provides a much less interesting comparison. The percentage of occurrences where each function performed the best amongst the variable weighting function with the same number of parameters is given in Table 4.4.

As indicated by the lower average error, the Prelec one parameter function, Pr-1, had the largest number of best fit occurrences amongst the one parameter functions. It performed the best for at least two thirds of the subjects for each of the simulated probability weighting functions. The two parameter variable weighting functions gave mixed results. The Prelec two parameter function, Pr-2, was clearly the best fit when the probability weighting function was either of the Prelec functions, Pr-1 or Pr-2. When the other probability weighting functions, TK-1 or GE-2, were used the linear in log odds variable weighting function, GE-2, fit the weighted probabilities best for the most subjects.

Table 4.4: The percentage of the 64 subjects for which each one or two parameter variable weighting functional form fit best is given when compared to the other forms with the same number of parameters.

| | Variable Weighting Function | | | | |
|--------|-----------------------------|------|------|------|------|
| $w(p)$ | TK-1 | Pr-1 | Pr-2 | GE-2 | Ku-2 |
| TK-1 | 31.8 | 67.2 | 12.5 | 71.9 | 15.6 |
| Pr-1 | 32.8 | 67.2 | 96.9 | 1.6 | 1.6 |
| Pr-2 | 18.8 | 81.3 | 96.9 | 3.1 | 0 |
| GE-2 | 28.1 | 71.9 | 9.4 | 68.8 | 21.9 |

4.3.3 Best Fit Parameters of the Average Variable Weighting Functions

A sampling of some of the variable weighting function fits has been shown in Figure 4.3. Table 4.5 displays the averages of the parameter values for each of the variable weighting functions

for each possible combination of probability weighting function. The average parameter values can be utilized as typical parameters for the variable weighting function. Each of the variable weighting functions has been shown in Figure 4.4. The each of these forms share similar characteristics, in particular, they each take on an s-shape, and cross the linear weighting function. The two parameter functions, which have full flexibility to change the elevation of the variable weighting function independently of the curvature, all cross the linear weighting function at a normalized variable value, \hat{v} , between 0.44 and 0.48.

Table 4.5: The parameters of the variable weighting functions given the probability weighting function and a form of the variable weighting function.

| | | Variable Weighting Function | |
|--------|--|-----------------------------|------------|
| | | TK-1 | Pr-1 |
| $w(p)$ | | γ | γ |
| TK-1 | | 1.18(0.15) | 1.30(0.19) |
| Pr-1 | | 1.21(0.16) | 1.32(0.20) |
| Pr-2 | | 1.17(0.17) | 1.56(0.32) |
| GE-2 | | 1.19(0.14) | 1.30(0.18) |

| Variable Weighting Function | | | | | | | |
|-----------------------------|------------|------------|------------|------------|------------|------------|--|
| | | Pr-2 | | GE-2 | | Ku-2 | |
| $w(p)$ | δ | γ | δ | γ | δ | γ | |
| TK-1 | 1.19(0.30) | 1.48(0.03) | 1.08(0.34) | 1.48(0.03) | 1.77(0.17) | 2.03(0.18) | |
| Pr-1 | 1.24(0.43) | 1.54(0.03) | 1.08(0.36) | 1.53(0.05) | 1.85(0.25) | 2.15(0.14) | |
| Pr-2 | 1.31(0.86) | 1.85(0.06) | 1.38(0.75) | 1.79(0.09) | 2.18(0.40) | 3.12(0.33) | |
| GE-2 | 1.15(0.28) | 1.46(0.03) | 1.11(0.33) | 1.45(0.03) | 1.70(0.16) | 1.96(0.17) | |

4.3.4 Sensitivity of Parameters in Individual's Variable Weighting Function

The sensitivity of the variable weighting function to the characteristics of the probability weighting function were investigated using the linear in log odds form, GE-2, of the probability weighting function. The varied parameters of the probability weighting function were

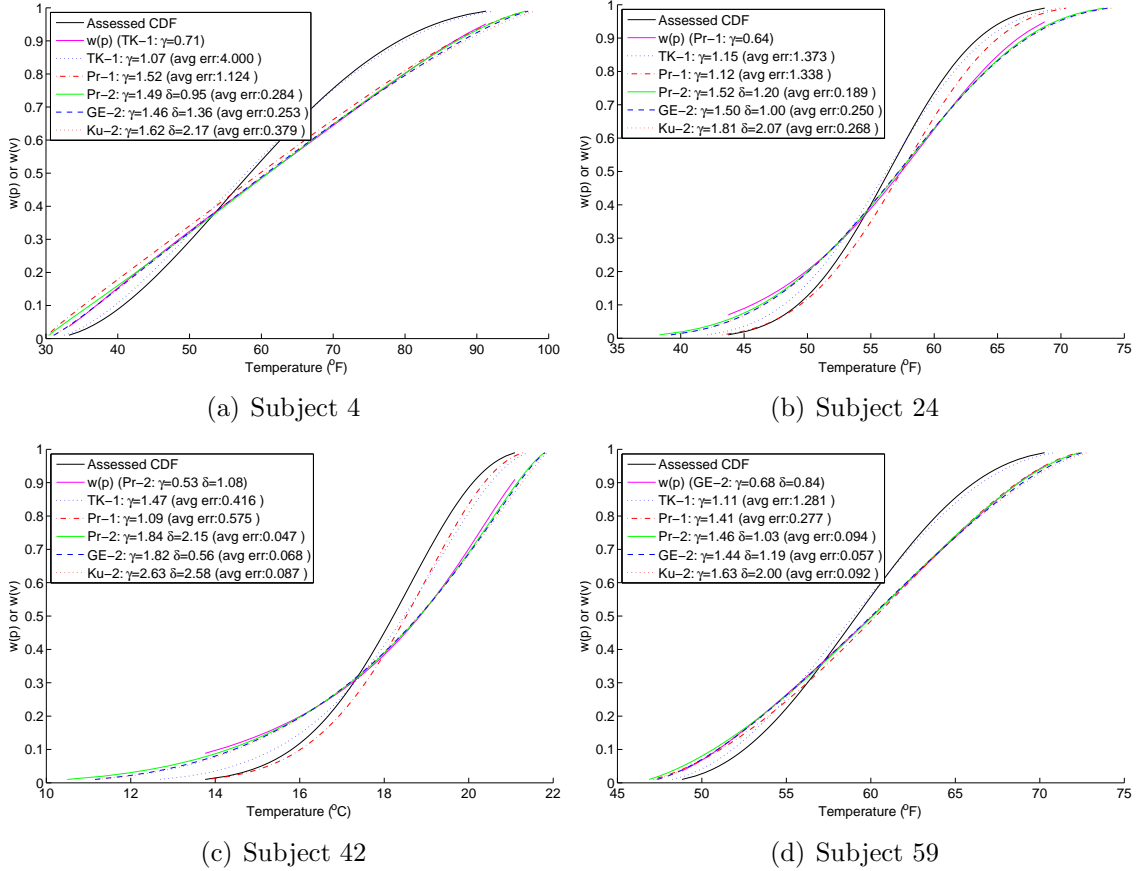


Figure 4.3: Variable weighted cumulative distributions fitted to probability weighted distributions for a sampling of subjects and forms for the probability weighting function.

used in conjunction with the subject data, in the same procedure as above, to produce the average parameters of best fit variable weighting functions. The results are shown in Figure 4.5. The γ parameter of the variable weighting function has very little dependence on the δ value of the probability weighting function. In other words, the curvature of the variable weighting function is almost completely determined by the curvature of the probability weighting function. The δ parameter of the variable weighting function, on the other hand, is a function of both the δ and γ parameters of the probability weighting function. As either the γ parameter or the δ parameter of the probability weighting function decreases the δ parameter of the variable weighting function increases at an exponential rate.

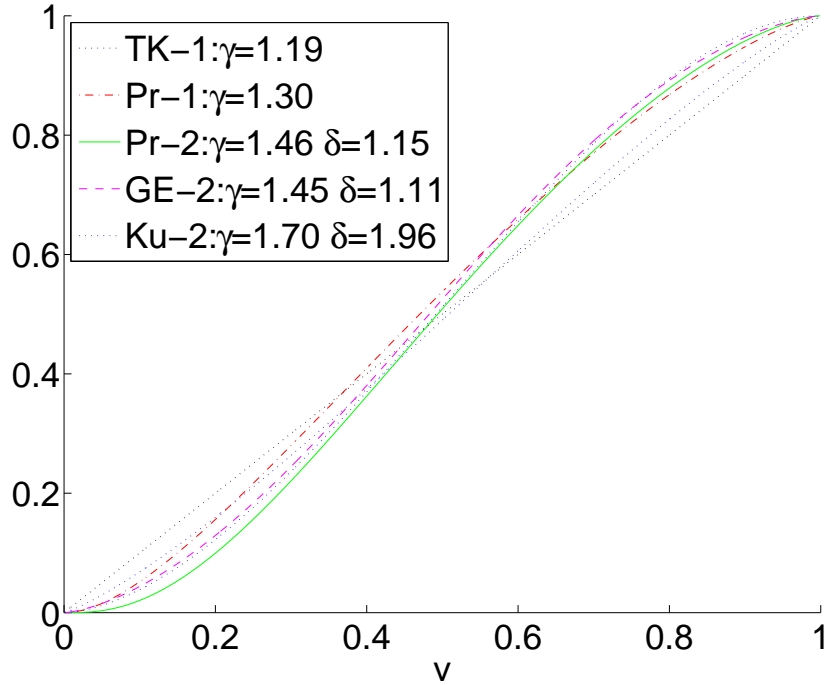


Figure 4.4: The variable weighting function is shown in each of the discussed forms based on the fit when the probability weighting function is assumed to be linear in log odds, GE-2.

4.4 Conclusion and Recommendations

The understanding of a subjects belief about a random variable can be greatly benefited by cumulative prospect theory. Understanding the properties of the probability weighting function provides some insight into the psychology of a decision maker and explains the discrepancy between how decision makers often act and how expected utility theory dictates they should act. In the same way, a greater understanding of how decision makers treat continuous random variables at a constant probability will allow further evaluation of the actions of decision makers. This article has described some of the key features of the variable weighting function. Functional forms of the variable weighting function have been discussed along with their ability to replicate the information of a standard probability weighting function. Furthermore, the characteristics of the variable weighting function have been shown in relation to the properties of the corresponding probability weighting function.

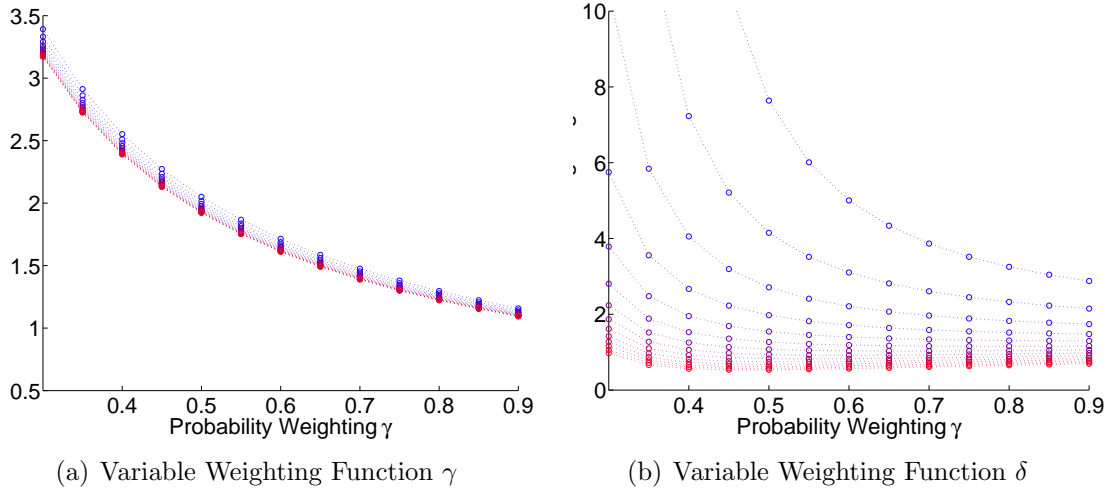


Figure 4.5: Parameters of the variable weighting function (GE-2 form) are given as a function of the parameters of the probability weighting function (GE-2 form). Each line represents a different δ of the probability weighting function, from $\delta = 0.3$, blue to $\delta = 1.5$, red.

The one parameter weighting functions were found to have significantly more error when used to describe variable weighting, due to the lack of degrees of freedom. Of these functions the Prelec function, Pr-1, was found to be the best fit. For the two parameter functions the fit depends on the form of the probability weighting function that the variable weighting function is attempting to replicate. The Prelec two parameter function, Pr-2, was found best represent the variable weighting of probability weighting functions in the forms, Pr-1 and Pr-2. The linear in log odds form, GE-2, of variable weighting fit the other functional forms best and is the most commonly used probability weighting functional form.

The sensitivity of of the variable weighting function to changes in the parameters of the probability weighting function was determined using the linear in log odds form to describe both functions. For a known probability weighting functions the results in Figure 4.5 can provide the parameters that could be used to describe the variable weighting function of the same individual. The curvature of the variable weighting function was found to be dependent on the curvature of the probability weighting function and independent of its elevation. These characteristics of the variable weighting function can be used when assessing a subjects beliefs

about a continuous random variable and are particularly useful when information regarding a fixed probability is desired.

REFERENCES

- [1] J. J. Moder and E. G. Rodgers. Judgment estimate of the moments of pert type distributions. *Management Science*, 18:76–83, 1968.
- [2] C. Perry and I. D. Greig. Estimating the mean and variance of subjective distributions in pert and decision analysis. *Management Science*, 21:1477–1480, 1975.
- [3] J. E. Smith. Moment methods for decision analysis. *Management Science*, 39:340–358, 1993.
- [4] A. E. Abbas. Entropy methods for univariate distributions in decision analysis. In *AIP Conference Proceedings*, volume 659 of *Proceedings of the 22nd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, pages 339–349, 2002.
- [5] A. E. Abbas. Entropy methods for joint distributions in decision analysis. *Engineering Management, IEEE Transactions on*, 53(1):146–159, 2006.
- [6] C. S. Spetzler and C.-A. S. Stael von Holstein. Probability encoding in decision analysis. *Management Science*, 22(3):340–358, November 1975.
- [7] W. Edwards and D. Winterfeldt. Public values in risk debates. *Risk Analysis*, 7(2):141–158, 1987.
- [8] D. von Winterfeldt and W. Edwards. *Decision Analysis and Behavioral Research*. Cambridge University Press, Cambridge, UK, 1987.
- [9] M. W. Merkhofer. Quantifying judgmental uncertainty: Methodology, experiences, and insights. *IEEE Transactions on Systems, Man and Cybernetics*, 17:741–752, October 1987. ACM ID: 40902.
- [10] A. O’Hagan, C. E. Buck, A. Daneshkhah, J. R. Eiser, P. H. Garthwaite, D. J. Jenkinson, J. E. Oakley, and T. Rakow. *Uncertain Judgements: Eliciting Experts’ Probabilities*. Wiley, September 2006.
- [11] D. V. Lindley. The probability approach to the treatment of uncertainty in artificial intelligence and expert systems. *Statistical Science*, 2(1):17–24, February 1987.
- [12] A. E. Abbas, D. V. Budescu, and Y. Gu. Assessing joint distributions with isoprobability countours. *Management Science*, 56(6):997–1011, June 2010.

- [13] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323, 1992.
- [14] S. C. Hora, J. A. Hora, and N. G. Dodd. Assessment of probability distributions for continuous random variables: A comparison of the bisection and fixed value methods. *Organizational Behavior and Human Decision Processes*, 51(1):133–155, February 1992.
- [15] S. Lichtenstein, B. Fischhoff, and L. D. Phillips. *Calibration of probabilities: the state of the art to 1980*. Judgment under Uncertainty: Heuristics and Biases. Cambridge University Press, Cambridge, UK, 1982.
- [16] R. A. Howard. *The evolution of decision analysis*. The Principles and Applications of Decision Analysis. Strategic Decisions Group, Menlo Park, CA, 1983.
- [17] R. A. Howard. Decision analysis: Practice and promise. *Management Science*, 14:679–695, 1988.
- [18] D. L. Keefer. Three-point approximations for continuous random variables. *Management Science*, 29:595–609, 1983.
- [19] P. McNamee and J. Celona. *Decision analysis for the professional*. SmartOrg Inc., Menlo Park, CA, 2001.
- [20] A. H. L. Lau, H. S. Lau, and Y. Zhang. A simple and logical alternative for making pert time estimates. *IIE Transactions*, 28:183–192, 1996.
- [21] M. Alpert and H. Raiffa. *A progress report on the training of probability assessors*. Judgment under Uncertainty: Heuristics and Biases. Cambridge University Press, Cambridge, UK, 1982.
- [22] D. V. Budescu and N. Du. The coherence and consistency of investors probability judgments. *Management Science*, 53:1731–1744, 2007.
- [23] J. E. Russo and P. J. H. Schoemaker. Managing overconfidence. *Sloan Management Review*, 33:7–17, 1992.
- [24] J. B. Soll and J. Klayman. Overconfidence in interval estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2):299–314, March 2004.
- [25] E. C. Carpen, E. V. Clapp, and W. Campbell. Competitive bidding in high risk situations. *Journal of Petroleum Technology*, 23:641–653, 1971.
- [26] H. Gates. Bidding strategies and probabilities. *Journal of Construction Division*, 93:75–107, 1967.
- [27] P. Juslin, P. Wennerholm, and H. Olsson. Format dependence in subjective probability calibration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(4):1038–1052, July 1999.

- [28] J. Klayman. Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, 79(3):216–247, 1999.
- [29] A. E. Abbas, D. V. Budescu, H.-T. Yu, and R. P. Haggerty. A comparison of two probability encoding methods: Fixed probability vs. fixed variable values. *Decision Analysis*, 5(4):190–202, 2008.
- [30] H. Raiffa and R. Schlaifer. *Applied Statistical Decision Theory*. Grad. School Bus. Admin., Harvard Univ., Boston, MA, 1961.
- [31] G. Hughes and L. V. Madden. Some methods for eliciting expert knowledge of plant disease epidemics and their application in cluster sampling for disease incidence. *Crop Protection*, 21(3):203–215, April 2002.
- [32] S. Kotz and J. R. van Dorp. *Beyond Beta: Other continuous families of distributions with bounded support and applications*. World Scientific Press, Singapore, 2004.
- [33] J. R. van Dorp, S. Rambaud, and R. H. Pleguezuelo J. Prez. An elicitation procedure for the generalized trapezoidal distribution with a uniform central stage. *Decision Analysis*, 4(3):156–166, 2007.
- [34] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291, 1979.
- [35] M. Allais. Le comportement de l’Homme rationnel devant le risque: Critique des postulats et axiomes de l’Ecole americaine. *Econometrica*, 21(4):503–546, October 1953.
- [36] John Quiggin. A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4):323–343, December 1982.
- [37] R. Gonzalez and G. Wu. On the shape of the probability weighting function,. *Cognitive Psychology*, 38(1):129–166, February 1999.
- [38] H. Bleichrodt and J. L. Pinto. A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science*, 46(11):1485–1496, November 2000.
- [39] H. P. Stott. Cumulative prospect theory’s functional menagerie. *Journal of Risk and Uncertainty*, 32(2):101–130, 2006.
- [40] G. Wu and R. Gonzalez. Curvature of the probability weighting function. *Management Science*, 42(12):1676–1690, December 1996.
- [41] M. Abdellaoui. Parameter-free elicitation of utility and probability weighting functions. *Management Science*, 46(11):1497–1512, 2000.
- [42] G. van de Kuilen and P. P. Wakker. The midweight method to measure attitudes towards risk and ambiguity. Rotterdam, The Netherlands: Econometric Institute, Erasmus University, March 2009.

- [43] D. Prelec. The probability weighting function. *Econometrica*, 66(3):497–527, May 1998.
- [44] C. F. Camerer and T.-H. Ho. Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, 8(2):167–196, 1994.
- [45] B. Donkers, B. Melenberg, and A. van Soest. Estimating risk attitudes using lotteries : a large sample approach. *Journal of Risk and Uncertainty*, 22(2):165–195, 1999.
- [46] W. M. Goldstein and H. J. Einhorn. Expression theory and the preference reversal phenomena. *Psychological Review*, 94(2):236–254, April 1987.
- [47] N. Etchart-Vincent. Is probability weighting sensitive to the magnitude of consequences? an experimental investigation on losses. *Journal of Risk and Uncertainty*, 28(3):217–235, 2004.
- [48] A. S. Booij, B. M. S. Praag, and G. van de Kuilen. A parametric analysis of prospect theory's functionals for the general population. *Theory and Decision*, 68(1-2):115–148, 2009.
- [49] Y Rottenstreich and C K Hsee. Money, kisses, and electric shocks: on the affective psychology of risk. *Psychological Science: A Journal of the American Psychological Society*, 12(3):185–190, May 2001.
- [50] H. Fehr-Duda, M. Gennaro, and R. Schubert. Gender, financial risk, and probability weights. *Theory and Decision*, 60(2-3):283–313, 2006.
- [51] T. J Dohmen, A. Falk, D. Huffman, J. Schupp, U. Sunde, and G. G Wagner. Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. In *IZA Discussion Paper Series*, volume 1730, Bonn, Germany, September 2005.
- [52] J. K. Goeree, C. A. Holt, and T. R. Palfrey. Quantal response equilibrium and over-bidding in Private-Value auctions. *Journal of Economic Theory*, 104(1):247–272, May 2002.
- [53] A. E. Abbas, D. V. Budescu, L. Wu, and R. P. Haggerty. On the shape of the variable weighting function. Work in Progress, 2011.
- [54] P. Kumaraswamy. A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46(1-2):79–88, March 1980.

AUTHOR'S BIOGRAPHY

Ryan P. Haggerty Graduated from Bellarmine College Preparatory in 2000. He went on to get a Bachelor's degree in Chemical Engineering from University of California at Santa Barbara. He is currently a Ph.D. candidate at University of Illinois, studying high temperature ceramics in the field of Materials Science and Engineering.