Computational Determination of Coherence of Financial Risk Measure as a Lower Prevision of Imprecise Probability

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for the degree of Master of Science
in the Department of Mathematics and Statistics
University of Saskatchewan
Saskatoon

By
Sudhakar Achath

Permission to Use

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Mathematics & Statistics
142 McLean Hall
106 Wiggins Rd
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5E6

Abstract

This study is about developing some further ideas in imprecise probability models of financial risk measures. A financial risk measure has been interpreted as an upper prevision of imprecise probability, which through the conjugacy relationship can be seen as a lower prevision. The risk measures selected in the study are value-at-risk (VaR) and conditional value-at-risk ($CVaR$). The notion of coherence of risk measures is explained. Stocks that are traded in the financial markets (the risky assets) are seen as the gambles. The study makes a determination through computation from actual assets data whether the risk measure assessments of gambles (assets) are coherent as an imprecise probability. It is observed that coherence of assessments depends on the asset's returns distribution characteristic.

Acknowledgements

I express sincere thanks to my supervisor (mon enseignant vu), Professor Emeritus Mikelis G. Bickis who introduced me with abiding interest to the wide field of imprecise probability, and for his constant support, guidance, and patience I received in preparing this thesis.

I thank Members and Faculty of the Department of Mathematics & Statistics, and the College of Graduate Studies & Research, University of Saskatchewan for providing me with the environment for undertaking able research and for providing financial assistance.

I am very indebted to M. Troffaes & G. de Cooman, *Lower Previsions*, 2014, and T. Augustin, F.P.A. Coolen, G. de Cooman & M. Troffaes (Eds.), *Introduction to Imprecise Probabilities*, 2014 (mes enseignants invisibles). I have unabashedly borrowed and used these two works in this study. My sincere thanks. I think I am now ready to read this *magnum opus* by P. Walley, *Statistical Reasoning with Imprecise Probabilities*, 1991.

I thank Amrita Vishwa Vidyapeetham (University), Coimbatore and Satguru Sri Mata Amritanandamayi Devi who gave me this opportunity to undertake a journey to Canada and study.

I dedicate this thesis to Kula Deivam Manapullikavu Bhagavathy Amman, Palakkad, Kerala, South India from whom I may derive all strength, solace and ability to learn.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 The Meaning of Imprecise Probability

In this thesis, we computationally determine whether financial risk measures computed from assets data obtained from secondary financial markets are coherent from the imprecise probability point of view. Coherence is a logical consequence of rational behavior of individuals who subjectively make assessments about betting for (buying) or betting against (selling) a gamble from some desirable set of gambles. Financial securities in the form of risky assets are the gambles considered here. An important point here is that betting for and betting against are not one and the same, they will not lead to one single assessment, in fact they are one and the same only as a special case. This is the distinguishing feature of a model of imprecise probability. The corresponding two assessments are the lower prevision and the upper prevision.

Rational behavior is first expressed through a set of axioms that an individual is expected to follow in making this betting assessment for each gamble. So if the axioms are violated then the behavior is not rational, and can lead to adversity for the individual. The axioms would lead to theorems that can specify certain conditions to be imposed on assessments that ensure rationality. The two conditions are the notions of consistency and of coherence of assessments. This also implies that it would be possible to know when the assessments are either not consistent or not coherent. Therefore when the assessment is not consistent or not coherent, a method needs to be developed to change these initial assessments so that the new assessments are now coherent and thus conform with rational behavior. This method has already been developed in the imprecise probability literature. This is part of what is called the theory of lower and upper previsions.

Following this method ensures that the assessment is consistent, and having ensured consistency, the assessments are corrected to achieve coherence. This correction is called the natural extension of the assessment. These notions are explained in the chapters that follow.

1.2 Applying an Imprecise Probability Model

Application of imprecise probability models to real-life situations is a growing field. This study attempts an addition to this field. The problem selected is in financial risk measures. Financial risk measurement as an imprecise probability has been well-studied. A risk measure has been interpreted as an upper prevision.

Coherence of risk measures has been a problem that has been examined from the finance as well as the imprecise probability point of view. The two points of view are not one and the same. It has been found that while some risk measures are coherent, there are others that are not coherent. Since coherence is a desirable property, there is an interest in knowing which new measures found or developed are coherent.

Value-at-risk (VaR) is currently being prescribed for and used by financial institutions because of its merit of focusing on only the down-side risk. However they do not satisfy coherence. Therefore, another risk measure conditional value-at-risk ($CVaR$) has been suggested, which overcomes the deficiency of VaR.

To the best of the knowledge of this author, an empirical study to determine coherence of VaR and $CVaR$ from the imprecise probability point of view, and calculated from historical assets data has not been undertaken yet. The issue is of interest because institutions have only historical data available for estimating measures of risk, for better risk management. Therefore we would like to know whether such measures obtained from observed data are coherent or not. This is an objective of the study. Another objective of the study is compute a measure of imprecise probability from historical data and test for coherence. It is observed from this study undertaken that in general, the calculated risk measures are not coherent and therefore need correction from an imprecise probability point of view.

1.3 Organization of this Thesis

The chapters in this study are organized as follows. Chapter 2 gives a brief overview of the literature on imprecise probabilities. Chapter 3 gives definitions that are needed for studying lower previsions. Chapter 4 gives the theory of lower previsions as in Troffaes and Cooman [67], and attempts to give a brief explanation of the proofs used in [67]. Chapter 5 gives the method for applying the theory of lower previsions to calculate consistency and coherence of lower previsions, as given by the two algorithms in Walley et al [76], and implemented as *improb* by Troffaes [66]. Chapter 6 is about application. It initially explains how a risk measure is interpreted by Vicig [70] as an upper and therefore as a lower prevision, and the notion of coherence as discussed in the finance literature. The chapter then discusses data used, calculations done, and results obtained about coherence of the calculated risk measures. Chapter 7 summarizes the study, and suggests directions for further study. References cited in the study are given next. The Appendices show calculations and numerical results obtained, which are given there in the form of Tables and Figures.

Chapter 2

Imprecise Probability – An Overview

Probability can be seen from the point of view of different strands of thought. This is more or less true of imprecise probabilities also, but there is also a unifying notion here, as can be seen from the description below and in later chapters. In the least, we can say that probability has a dual nature.

The discussion in this chapter is organized in the following manner. We start with different interpretations of probability, then some aspects of axiomatizing of subjective probability are considered which may be thought of as yet another interpretation of probability. We then move to a discussion of non-additive probabilities and imprecise probabilities, where we finally focus on lower previsions as our imprecise probability model. Only a very elementary and partial discussion on the wide field of probability is attempted here.

The Stanford Encyclopaedia of Philosophy (SEP [62]) gives six main interpretations of probability. They are, classical probability, logical probability, subjective probability, the frequentist interpretation, the propensity interpretation and the best system interpretation. Starting from sources in SEP and other sources as well, we shall briefly discuss the first five here.

Classical probability originates possibly from Bayes (1763) and from Laplace (1812). Bayes expressed probability as a ratio. A classical probability interpretation is based on the principle of indifference (a terminology coming from Keynes), which was previously called the principle of insufficient reason. The principle of indifference implies a certain symmetry in occurrences, in the sense that all possible events are equally probable to occur. Similarly, the principle of insufficient reason suggests that there is no reason to believe that one event is more or less likely to occur than another. An example is the probability of obtaining some face in an experiment of throwing a six-faced fair die. It is equal and is one-sixth for each face. In this interpretation, evidence is not assumed to be available, or even if it is available, it favours equally all possible outcomes.

The frequency interpretation of probability which is referred to as aleatory probability somewhat generalizes classical probability, in the sense that it can take care of a fair as well as an unfair coin. For example, a weight of 0.6 uniformly (and therefore, assigning equal probabilities) to all occurrences of heads in a finite sequence of tosses of an unfair coin. The relative frequency of occurrence of heads give us a finite frequentist, and similarly for infinite tosses, an infinite frequentist point of view of interpreting probability. The frequentist notion does not look at causes of the phenomenon. The underlying structure (like the physical symmetry of a fair coin or die), i.e., the causes of the phenomena, is not emphasized.

In this notion, A. de Moivre (noted for his work "The Doctrine of Chances" of 1718 [42]) gave the normal approximation to one case of the binomial distribution in 1733. This result was generalized to the central limit theorem by Laplace in 1810. The work of 1718 above mentions relative frequency as probability, and appears to have been the motivator for the development of frequentist inference. Thus frequentist inference is a method for drawing conclusions about unknown but fixed population parameters using frequency of observed actual occurrences of particular events.

Further and in this tradition, is the limiting frequency notion, a relative frequency in the long-run, which was first expressed by J. Venn in 1866. R. von Mises in "Probability, Statistics and Truth" that was first printed in 1919 formalizes this notion further (Hacking [23]). The frequency view here gives a physical notion for probability, in the sense that it is the probability associated with a physical phenomena, like a property, like throwing a six-faced die and obtaining six. In this interpretation, probability is the limit of the relative frequency of an event in a large number of random occurrences [73]. Convergence of an infinite random sequence of relative frequencies to a probability is assumed here.

A notion related to the frequentist interpretation, originating initially from C.S. Peirce and then from Karl Popper, is the propensity interpretation of probability (Hacking [23]). A coin tossing experiment (which can be likened to a chance set-up), if it has probability p of obtaining a head in a single toss, then in repeated tosses, we say that the chance set-up has the tendency or the propensity for the long-run relative frequency of heads to attain probability p. This appears to be very similar to the limiting frequency interpretation of von Mises, but is intended to explain the probability of a single-event.

So frequency and propensity come as physical interpretations of probability. Each one is an aleatory probability. However, they cannot explain unique events, i.e., for experiments that are not repeatable. A volley-ball final game may be played as best of fives, but a soccer final may be played only once. Also, in a chance set-up like tossing a coin to obtain heads, we do not know how many number of tosses are needed to get a "good" approximation to the true probability. So we leave the empirical world and seek another interpretation, viz., the knowledge-based or epistemic view. These are also evidential probabilities, as we will now see.

Epistemic probability does not have a physical interpretation. Within the epistemic view, Fine [13] attributes the development of logical probability to J.M. Keynes (1921), H. Jeffreys (1939), B. Koopman (1940), and R. Carnap (1962). The other epistemic view is that of subjective probability. Kyburg [35] states that early contributions to subjective probability are from Ramsey (1964), de Finetti (1964), Savage (1954), Koopman (1940), and Good (1950, 1962). Ramsey, de Finetti and Savage give a decision-oriented view of subjective probability. Koopman and Good have an intuitionist view of probability.

In logical probability, probability is the logical relation between two propositions, say a hypothesis and an evidence. It is an inference and expresses the degree of partial entailment, which can be thought of as a measure of partial implication. Since this is the probability of hypothesis given evidence, this is taken as a conditional probability. Additionally, Levi [39]) also says that there is an epistemological, and a knowledge-based, degree of partial implication, that is not necessarily numerically determinable, meaning that logical probabilities need not be numerically expressed. Keynes also asserts that probabilities need not have a numerical representation.

Keynes sees this probability judgement as a measure of the degree of objective knowledge possessed (by an individual) in confirming the hypothesis, just as the process of logic is objective. It is not based on individual opinion, in the sense that given the same evidence, different individuals would come to the same probability judgement. Koopman strengthened Keynes' notions by making probability a qualitative conditional probability judgement. By qualitative, one means an order relation on logical probabilities that is decided by an individual. Therefore, this is a comparative conditional probability. By introducing further notions, Koopman transforms this to quantitative comparative conditional probability. This is part of the axiomatization of probability which we shall briefly visit later in this chapter. Carnap sees logical probability as a quantitative measure of the relation between hypothesis and evidence. He calls this a degree of confirmation function, and this is assessed (measured) through reasoning inductively for confirming a hypothesis from evidence. A mathematical discussion of these ideas is not attempted here but can be found in Fine [13].

Subjective probability as a degree of belief based on a state of knowledge, and called a personal probability (coined by Savage) or a Bayesian probability, is a measure of one's own belief or judgement about the chances of occurrence of an event that is held by an individual. It may agree with what can be known from the classical or the frequentist point of view but is not a physical construct. But individuals can differ in their assessments, so one may not find a common measure of belief between any two. These degrees of belief are also called as credences. According to SEP [63], credences are interpreted in terms of a subject's limiting willingness to bet, it can be inferred from the subject's behaviour, and provide its quantitative measure. The subject could revise or update these beliefs based on new knowledge.

Thus in de Finetti, the numerical measure of probability is the betting rate for a gamble (a ratio of the amount an individual stakes his bet for it with respect to the total stake for the gamble) decided by an individual. De Finetti [11] terms these betting rates as previsions, and uses the terms prices, previsions and probability interchangeably. When the random quantity is an event then prevision is a probability. Therefore, when expressed monetarily, the betting rate is a price, it is a prevision when expressed non-monetarily, and a probability when the random quantity is an event. The term prevision is intended as a better description of Kolmogorov's probability. A prevision assessed by the individual is taken as the fair price for buying and selling gambles by him. The individual is willing to take both sides of the bet. This means that for prices below the fair price the individual is disposed to bet in favour, while for prices greater than the fair price the individual is disposed to bet against. The notion of coherence given by de Finetti which is related to rational behaviour of this individual also originates from here. We shall discuss this later.

So we see that betting behaviour reveals the individual's belief or assessment about a gamble, in fact through his stated previsions. So subjective probability models beliefs. The intuitionist view, however, takes a less extreme position. This view maintains that beliefs come from an individual's intuition, they need not be expressed, even numerically, and therefore beliefs need not be revealed through making choices.

De Finetti's works spans from 1926–1983, unfortunately they were known to the English-speaking world only during the later years. De Finetti is considered to have made important contributions in the foundations of probability of the 20th century (Suppes [64]). Ramsey and de Finetti developed the subjective theory of probability in similar ways, but independently. This work is said to have been extended and synthesized by Savage [53]. We shall not pursue this part here further, but shall discuss rational behaviour and its relation to probability.

Our subjective judgements also need to follow a probability calculus in the sense that, if I say that there is at least a 60% chance that I will be in Switzerland next summer, then I cannot say that there is at most a 50% chance that I will not be in Switzerland next summer (it needs to be at most a 40% chance). This consistency in probability judgements is taken as a criterion for rational behaviour of an individual.

While placing bets, de Finetti specifies that the previsions assessed for each bet must be such that the individual avoids sure loss. Incurring sure loss means that the a bet has been placed in a manner such that whatever be the outcome of the gamble, the individual will make a loss. In other words, the individual would need to place bets in such a way that he should not be Dutch booked, and de Finetti terms maintaining this consistency criterion as satisfying the notion of coherence. This applies to all the gambles that are considered. Then, their respective assessed previsions are said to be coherent, and the individual behaviour is deemed rational. This also means satisfying the probability calculus which implies satisfying the Bayes' rule. A discussion of probability calculus and Bayes' rule can be found in Skyrms [58].

Zabell gives a summarized account of probability in the volume edited by Haaparanta [83]. Here Zabell categorizes probabilities in the following way. To start with, probabilities can be either aleatory or epistemic. Aleatory means that they are properties of some physical or chance phenomena. Epistemic means that probability is based on our current knowledge of events. Aleatory probabilities can be either in the form of propensities or in the form of frequencies. Frequency can be either finite or infinite. These are not inductive probabilities in the sense that an inductive argument, starting from some premises, is not used to arrive at a conclusion. On the other hand, epistemic probabilities are inductive inferences. They can be either qualitative (Keynes, Koopman) or quantitative (Carnap, Ramsey, de Finetti). They can also be categorized as being either objective (Keynes, Carnap) or subjective (Koopman, Ramsey, de Finetti), where objective is in the sense that they are either a measure of the rational degree of belief a person should have given their state of knowledge, or subjective in the sense that it is the belief of one person and can vary from person to person. Subjective probability can be either personal or psychological, i.e., they either satisfy certain rationality constraints or they do not. According to him, epistemic probability with all its sub-categories is refereed to as subjective probability.

Expositions on theories of probability can be found in Fine [13], Good [19], Hartigan [26], Jeffreys [29], Jaynes [28], Narens [44], Skyrms [58], and von Plato [74]. Since we have covered interpretations of probability, we shall now focus some attention on axiomatizing probability. The above cited references contain discussions on axiomatization. Further, articles by Narens [43] and Villegas [72] discuss axiomatizing qualitative probability. Fishburn [15] gives a survey on the axioms of subjective probability, while according credit to the pioneering work done by Ramsey, de Finetti, Savage, and Koopman. His article is followed by many critical comments and his rejoinder to them. Berger (in his comments to [15]) indicates that probability can be derived from various axiom systems. In fact, Fine [13] analyzes and discusses different axiom systems that include comparative probability relations, quantitative probability, relative frequency and probability, and subjective probability models. Fine's book is not for the faint-hearted.

The first ideas of axiomatizing probability could have come from Huygens in 1657. There, the fair price of a bet (which is actually an expected value of an event), is taken as a primitive idea to axiomatize. Decisions are based on what we believe and what we want, leading to the notion of probability as one's belief and utility or reward as one's want (Hacking [23]). In the 1930's, the motivation appears to have come from Hilbert's 6th problem in 1900 where among others like physics, probability theory was desired to be written along the lines geometry is written, i.e., axiomatically (von Plato [74]).

Good [19] explains the axiomatic method in general as follows. First there are certain entities. Relations are defined between entities, and axioms are assumptions that further specify the nature of the relations. The set of axioms is minimal, and axioms are expected to be independent of one another. All mathematical results are then derived from them which is called the abstract theory. According to Good there are additional rules for applying the theory. The theory needs to be consistent in the sense that it should not lead to contradiction or unreasonableness. Also, the theory must facilitate applicability to practical situations including some idealized situations. The theory thus leads to a structure for the subject that is under study.

In 1933 in the "Foundations of the Theory of Probability", Kolmogorov [32] gave an axiomatic development of probability based on events, on an algebra of a collection of events taken as sets. The probability that a point is in a set is given by the measure of the set. The three axioms of probability are non-negativity, normalization, and additivity. When the collection of events is a sigma algebra, then probability has the additional axiom of countable additivity. This additional axiom makes probability continuous.

Fine [13] while discussing quantitative probability, mentions that as von Mises had pointed out, the (sigma) algebra of Kolmogorov may contain sets that may not have an interpretable meaning as events and therefore no interpretable probability values for those sets, or that some events that are relevant and are to be present in the algebra, may be absent as sets in the algebra. Two other algebras that are smaller are the Dynkin's lambda- and the pi-system, however they could be over-restrictive. Von Mises constructs a field where from events in it one can calculate or "measure" probability by repeating an experiment a finite number of times. For an understanding of the von Mises field we refer the reader to a discussion in Fine [13]. Further, Fishburn in a rejoinder to comments on his 1986 article [15], says that countable additivity is much like the axiom of choice. If one can do without it then so much the better, or if one cannot avoid it then choose countable additivity to solve the intended probability application on hand.

According to Villegas [72], first de Finetti (1931), and later Koopman (1940) axiomatize qualitative probability. Here, qualitative probability is meant as an order relation between events in terms of "at least as likely as", or "more likely than", or "as likely as" relations. This approach is taken because individuals are more tuned to making qualitative comparisons than making quantitative comparisons. The concern was also about axiomatizing order relations on a Boolean algebra, with an expectation to ground Kolmogorov's theory of probability in terms of preference relations. A probability function defined on the algebra such that order relations are preserved, and mapping to the real line gives a quantitative probability.

In the axiomatic approach to probability, the probability calculus that we have mentioned earlier is captured by a set of axioms that are specifying certain rationality conditions. The axioms also place sufficient restrictions on qualitative probability judgements to be able to provide a numerical representation.

We shall now briefly discuss de Finetti's axiomatization of subjective probability. Essentially, the coherence condition mentioned earlier is implemented through a set of axioms. The common thread is that rational belief is constrained by coherent preference, and binary choices reveal preferences (Seidenfeld's comments to [15]).

The qualitative probability relation is given as "an event A is at least as probable as an event B" relation. The four rationality axioms of de Finetti are weak order (i.e., order relation is asymmetric and transitive), non-triviality, non-negativity, and finite additivity (Fishburn [15]). Asymmetric means that if A is preferred over B, then B cannot be preferred over A. For example, the order relation on real numbers with "greater than" relation is an asymmetric relation. Transitivity means that if A is preferred over B, and B is preferred over C, then A is preferred over C. Non-triviality means that some event is bound to occur as compared to no event occurring.

The coherence condition for the bets is then defined by de Finetti, and the individual chooses the probability for each event such that the coherence condition is not violated. This is the only restriction placed on the choices made by the individual bettor. De Finetti then shows that coherence implies that the probabilities chosen fulfill the axioms of finitely additive probability (von Plato [74]).

De Finetti [11] also shows that coherent probability is an admissible decision based on minimizing a loss function. This discussion on coherence is of interest to us later when we see previsions as a special case of imprecise previsions.

The term imprecise probability was coined by Walley [78]. The notion of an imprecise probability corresponds to a doubt whether our beliefs can be represented by a single real-valued probability function (Stanford Encyclopaedia of Philosophy [63]). This doubt seems to have been formally first expressed by Boole [7]. Boole's probability actually follows from Laplace (Gridgeman [21]). A problem Boole was trying to solve was to determine lower and upper probability bounds of a Boolean function of events, given propositional statements about the events. This was later solved by Hailperin [24] using a linear programming formulation.

The origin of imprecise probability can be be seen in non-additive probabilities. A non-additive probability means that the probability of a union of two disjoint or incompatible events need not be equal to the sum of individual probabilities of the two events. An interesting discussion on non-additive probabilities by Prof. Colin Howson may be found at Howson [27]. Vicig and Seidenfeld [71] give a brief historical note on the origin of imprecise probabilities. Non-additive probabilities first appeared in "Ars Conjectandi" of Jacob Bernoulli in 1713 (Hacking [22], Hampel [25]). Lambert, in his work "Neues Organon" of 1764, discusses several examples where evidence seems to justify only non-additive probabilities (Shafer [57]). Lambert continues Bernoulli's work and gives a rule for combining non-additive probabilities of evidence (Hampel [25], Shafer [55], [56]).

Later on, Lambert's work provided the structural basis for belief functions, and Dempster's rule of combination for belief functions is a generalization of this (Shafer [57]). Belief function, also known as evidence theory or Dempster-Shafer theory, is a model for reasoning with uncertainty and has connections with models of probability and imprecise probability. In this, belief is the degree to which evidence supports hypotheses. The theory shows a way to combine different evidences to give a overall measure of the degree of belief.

Keynes in "A Treatise on Probability" [31] states that probabilities cannot be simply ordered but can be placed between numerical limits in the sense that probability can be compared to zero probability and to one probability [37]. Brady & Arthmar [8] suggest that Keynes used Boole's mathematical logic to develop a non-additive approach to the specification of interval estimates in terms of an upper and a lower probability.

The intuitionist view of Koopman and Good also maintain that preferences can only be partially ordered and therefore expressed through an interval probability, similar to that given by Keynes. Partial ordering may imply that probabilities need not be numerical leading to a qualitative probability, but if numerical values can be assigned then we may find upper and lower probabilities and therefore interval values (Good, 1983). Vicig and Seidenfeld [71], in their note, mention that de Finetti quotes Koopman and Good that giving numerical values for upper and lower probabilities follows previous ideas from Keynes.

Koopman [33] following Keynes axiomatized intuitive probability where the additivity criterion was not required, and represented by a greatest lower bound and a least upper bound, because our probability judgements are only partially ordered. Aumann [5] revised the utility theory of von Neumann-Morgenstern and Savage from a total ordering of preferences to a partial ordering of preferences. Preference ordering is not complete. Accordingly, the probability axioms were revised by Good [20] and Kyburg [37]. Good says that most subjective probabilities are regarded as belonging only to some interval of values, because our qualitative judgements are imperfect and can only be partially ordered. Upper and lower probabilities are obtained as constraints, and a set of axioms similar to that of Koopman are derived.

Kyburg is notable for his contribution to inductive inference and the logic of rational belief (Spielman [61]). For Kyburg, a statistical statement is given in terms of an interval. Jeffrey [30] does not take belief states to be full preference rankings of Boolean algebras, and says that partial rankings characterize belief states that are infinite sets of probability functions.

Suppose that an urn contains 30 black balls and 70 white balls. Then we may say that there is 0.3

probability of drawing (with replacement) a black ball from the urn. Suppose that we say there is a 0.3 probability of a bus strike in Saskatoon next month. Even though the probabilities are the same, we can see that there is less uncertainty in the second statement when compared to the first. Therefore we might want to express this probability in terms of an interval. And if we are totally ignorant about uncertainty of a strike then we might want to express probability as an interval from zero to one.

Fine [14] and Walley & Fine [77] later developed interval-valued probabilities. The lower and upper probabilities were developed from a frequency point of view as observed in some physical or empirical processes ([80]).

Kyburg states that Smith ([59], [60]) was motivated towards interval-valued probability by the desire for a better description of degrees of belief than what subjective probability could do. This was also the motivation in Walley [78]. Smith considers lower and upper betting odds in discussing lower and upper probabilities, and also speaks of a desirable set of gambles (Dempster [9]), Augustin et al [4]). In Levi [38] the probability of a statement is given as an interval. The beliefs of an individual is given by a convex set of probability functions defined over a field of propositions. Levi calls this a credal set. A discussion of the approach by Levi can be found in Kyburg [36].

Ellsberg [12] suggests that we have to move away from total order to understand uncertainty. Probability judgements are not to be seen as just one probability distribution but is to be seen as a set of probability distributions. A similar view is held by Allais [2], who suggests maintaining total order, but extending Bayesian point probabilities to a distribution of probabilities. Ellsberg gives empirical evidence where when given a choice between aversion to risk versus aversion to ambiguity, an individual prefers aversion to ambiguity over aversion to risk. The paradox arises when we use subjective utility theory with additive probability. However when an imprecise probability model is used it is seen in the example that the individual is neutral between the two choices.

A formal theory of imprecise probability starts with previsions. The work of de Finetti in previsions is extended by Willaims [82] in two ways – (1) from precise previsions to imprecise previsions and, (2) from conditional previsions to conditional imprecise previsions. Williams states that an individual may be willing to accept only one side of a bet, i.e., it may not be realistic to say that there is just one fair price. This lead to the concept of lower and upper previsions. Williams, using the notion of acceptability of gambles, states rationality criteria for sets of acceptable gambles, and derives a generalized form of coherence. Conditions for coherence, consistency, and extension are given. This forms the foundation for the theory of lower previsions. An accessible review of William's work can be found in Vicig, Zaffalon and Cozman [69].

The theory of lower previsions is thus a model of epistemic uncertainty. This theory has been comprehensively developed in Walley [78]. In the theory of coherent lower previsions given by Walley there are three models, viz., coherent lower and upper previsions, closed convex sets of linear previsions which are sets of probability measures, and sets of desirable gambles. A summary of Walley's work and the inter-connection of the theory of lower previsions with other theories in given in Miranda [40].

Walley [75] says that sets of desirable gambles or partial preference orderings is the most general form of imprecise probability models, and also the most informative. They uniquely determine upper and lower previsions. In this article, Walley states that in an unpublished manuscript by him, a method is outlined for constructing partial preference orderings from simple judgements.

There is a duality relationship between a model of coherent lower prevision and its corresponding set of linear previsions. This permits a choice between the use of the models, where depending on ease of formulation, some problems are formulated as sets of linear previsions, while others are modelled as coherent lower previsions, therefore Walley advocates the use of both. Also, it is not necessary to adhere to one unified theory of imprecise probability. With suitable restrictions special models are used for specific applications, just as there are special types of probability models for probability theory (Walley [75]).

As an illustration in statistical inference using imprecise probabilities, Walley [79] presents the Imprecise Dirichlet Model (IDM). The prior probability in the IDM is the set of Dirichlet distributions. This is updated with data observed from a multinomial distribution using Bayes' Theorem to find the posterior IDM. Lower and upper probability for prior and posterior is obtained by minimizing/maximizing probability with respect to prior strength (Bernard [6]).

In the same year 1991 that Walley published the book "Statistical Reasoning with Imprecise Probabilities", Kuznetsov published in Russian a book titled "Interval Statistical Models". They had not collaborated with each other, yet it was essentially the same theory of coherent lower and upper previsions (Kozine [34]).

Weichselberger [81] starts with a probability measure that obeys Kolmogorov's axioms, and defines an interval-valued set function for an event. The definition is not restricted to interpretation and enables a generalization of probability measure.

The organization *The Society of Imprecise Probability: Theory and Applications (SIPTA)*, created in 2002 with homepage at www.sipta.org, now coordinates activities in the development of imprecise probability. Its first international symposium (*ISIPTA*) was held in 1999, and the first summer school (*SSIPTA*) was held in 2004.

The study that is conducted here is primarily about the theory of lower previsions that follows as a consequence of rational betting behaviour, and an attempt is made to apply this theory to practice. The application is in finance, in the area of risk measurement. A financial risk measure is defined, it is seen as an assessment of a lower prevision, and the study aims to determine computationally whether the assessments are coherent when they are considered from an imprecise probability point of view. The background and supporting literature and the method for determining coherence are developed in the following chapters. The next chapter attempts to give the motivation and need to define some terms that would introduce us to the subject of lower previsions.

Chapter 3

Some Definitions

In this chapter we start with some definitions of terms that we would come across in this study. All definitions given here are not used in this study, however their presentation here will help to develop later an understanding of the theory of lower previsions. The development that is presented here mostly follows from Troffaes & de Cooman [67].

The main imprecise probability model that is discussed in this study is that of coherent lower previsions which is discussed in the next chapter. That chapter also discusses other imprecise probability models. Therefore, in this chapter we develop some initial ideas that we would need as we progress.

In representing beliefs using models, there is a need to compare gambles and learn about preferences. Therefore, the ideas of partially ordered and totally ordered sets need to be developed. Also, by introducing certain axioms of rational behaviour while evaluating gambles, one would want to reason from them and arrive at their consequences for choice and decision. These consequences of reasoning would appear as conditions for consistency and coherence for the models. These conditions are explained further on in our study.

Similarly, one should be able to evaluate preferences for a combination of two or several gambles, increase or decrease our stake in some particular gamble, or opt for both buying and selling of a set of gambles. An algebra on linear spaces is therefore needed.

When we study a space of gambles, and build imprecise probability models on it, the primary mathematical tool that we would use in the study is that of a function space. A gamble is seen as a real-valued function defined on some possibility space. This space could be finite or infinite. The set of gambles defined on this possibility space can be assumed to be a vector space. Real functions are defined on this vector space of gambles. We shall see later that imprecise probabilities are real functionals. We can see this from the definitions of lower and upper prevision in Chapter 4 (Equations 4.2 and 4.3). Then, properties of real functionals on a vector space transfer to, or are inherited as properties of imprecise probabilities, viz., the lower and upper previsions, and also as lower and upper probabilities. From these properties, we obtain a theory of imprecise probability. Our focus of study would be on lower prevision as our imprecise probability model. Further on in this study, our notion of lower prevision is applied to empirically determine coherence of a risk measure in finance. When we are able to interpret a certain risk measure as a lower prevision, then we are seeing the risk measure as an imprecise probability for it, and our task then is to study coherence of this risk measure from this imprecise probability point of view.

Lattices give additional structure to the space of gambles. A semi-lattice (Definition 3.1) of bounded gambles help us define a sub-class of lower previsions called n-monotone lower previsions (Troffaes et al. [67]). When gambles are considered only as events (we saw in the earlier chapter how a random quantity can be seen as an event), they help us define n-monotone probabilities (Troffaes et al. [67]). Lattices and monotone lower previsions however, are not studied here.

We now define and formalize somewhat these notions in the following.

A possibility space $\mathcal{X}$ is a non-empty finite or infinite set of outcomes or elementary events. It does not contain impossible outcomes. A gamble is a bounded real-valued function f on $\mathcal{X}$, i.e., $f : \mathcal{X} \to \mathbb{R}$. It is the uncertain valuation, reward or utility $f(x)$ from the gamble when $x \in \mathcal{X}$ occurs or is determined.

The gamble f is analogous to a random variable. For example, suppose that in soccer the possibility space of outcomes of a game where team Bayern Munich is matched against team Toronto FC is given by $\mathcal{X} = \{w, l, t\}$, and corresponding to a win, lose, or tie outcome respectively for Bayern M. Suppose that the pay-offs (in million Euros) for Bayern M. from this game are $1.0, -1.0$, and 0.4 respectively. Then $f = \{f(w), f(l), f(t)\}^T = \{1.0, -1.0, 0.4\}^T$ is a gamble with rewards as shown, and f is analogous to a random variable.

The set of all gambles on the possibility space is denoted as $\mathcal{G}(\mathcal{X})$. Unless stated otherwise, we will denote this set as $\mathcal{G}$. So the space $\mathcal{G}$ with the properties of element-wise addition, and scalar multiplication with a real number is a real vector space.

Suppose that $*$ is a binary operation on real numbers. In the case of gambles, for any two gambles f and g in $\mathcal{G}$, then $f * g$ is the gamble in $\mathcal{G}$ defined by

$$(f * g)(x) \coloneqq f(x) * g(x) \text{ for all } x \in \mathcal{X}.$$

This binary operation on two gambles could include addition, subtraction, or multiplication of any two gambles. The binary relation given by $\leq$ (the "less than or equal to" order on real numbers) defined on $\mathcal{G}$ gives a partial ordering of its elements, and it is reflexive, anti-symmetrical, and transitive (Definition 3.2). When g will give a reward at least as much as that of f whatever be the outcome in the possibility space $\mathcal{X}$, then we say that gamble g dominates gamble f. Thus

$$f(x) \leq g(x) \text{ for all } x \in \mathcal{X} \text{ implies that } f \leq g.$$

With respect to this partial order, the supremum and the infimum of two gambles f and g on $\mathcal{X}$ is defined as the point-wise maximum and minimum respectively which is given by

$$(f \vee g)(x) \coloneqq \max\{f(x), g(x)\}$$
$$(f \wedge g)(x) \coloneqq \min\{f(x), g(x)\}.$$

Further, the supremum and infimum of a single gamble f on $\mathcal{X}$ are defined as

$$\sup f \coloneqq \sup\{f(x) \colon x \in \mathcal{X}\} = \min\{a \in \mathbb{R} \colon a \geq f\}$$
$$\inf f \coloneqq \inf\{f(x) \colon x \in \mathcal{X}\} = \max\{a \in \mathbb{R} \colon a \leq f\}$$

The absolute value of a gamble f is given by $|f|(x) \coloneqq |f(x)|$ and the negation of f is $(-f)(x) \coloneqq -f(x)$ for all $x \in \mathcal{X}$. A constant gamble as used above is $a(x) \coloneqq a,\ a \in \mathbb{R}$ for all $x \in \mathcal{X}$. The scalar product of a real number λ and a gamble f is defined as $(\lambda f)(x) \coloneqq \lambda(f(x))$ for all $x \in \mathcal{X}$.

We will use vector inequalities in comparing gambles. We say that,

(1) f is non-negative, if $f(x) \geq 0$ (it is everywhere non-negative)

(2) f is non-positive, if $f(x) \leq 0$ (it is everywhere non-positive)

(3) f is negative, i.e., $f < 0$, if $f(x) \leq 0$ and at least one $f(x) \neq 0$ (it is everywhere non-positive and negative in at least one element)

(4) f is positive, i.e., $f > 0$, if $f(x) \geq 0$ and at least one $f(x) \neq 0$ (it is everywhere non-negative and positive in at least one element).

Further, we characterize the space of gambles in the following way.

Definition 3.1. *A subset $\mathcal{K}$ of $\mathcal{G}$ is:*

(1) Negation-invariant if it is closed under negation, i.e., $-\mathcal{K} \coloneqq \{-f : f \in \mathcal{K}\} \subseteq \mathcal{K}$

(2) A cone if it is closed under non-negative scalar multiplication, i.e., $\lambda\mathcal{K} \coloneqq \{\lambda f : f \in \mathcal{K}\} \subseteq \mathcal{K}$ for all $\lambda \in \mathbb{R}_{\geq 0}$

(3) Convex if it is closed under convex combinations,
i.e., $\lambda\mathcal{K} + (1-\lambda)\mathcal{K} \coloneqq \{\lambda f + (1-\lambda)g : f, g \in \mathcal{K}\} \subseteq \mathcal{K}$ for all $\lambda \in (0,1)$

(4) A convex cone if it is closed under non-negative linear combinations,
i.e., nonneg$(\mathcal{K}) \coloneqq \{\Sigma_{k=1}^{n}\lambda_k f_k : n \in \mathbb{N}, f_k \in \mathcal{K}, \lambda_k \in \mathbb{R}_{\geq 0}\}$

(5) A linear space if closed under addition and scalar multiplication,
i.e., span$(\mathcal{K}) \subseteq \mathcal{K}$, where span$(\mathcal{K}) \coloneqq \{\Sigma_{k=1}^{n}\lambda_k f_k : n \in \mathbb{N}, f_k \in \mathcal{K}, \lambda_k \in \mathbb{R}_{\geq 0}\}$

(6) A $\wedge$ semilattice if it is closed under point-wise minimum and a $\vee$ semilattice if it is closed under point-wise maximum

(7) A lattice if is closed under point-wise minimum and maximum

(8) A linear lattice if it is both a linear space and a lattice.

Negation-invariance helps us to obtain a self-conjugate lower or upper prevision from which we obtain a prevision. This is discussed in Chapter 4. From a given finite set of desirable gambles the cone help us to obtain the complete set of gambles that are as desirable as the initial set. A linear space is also defined as a negation-invariant convex cone. The space of gambles $\mathcal{G}$ is closed under addition and scalar multiplication, and also closed under point-wise minimum and maximum for every pair of gambles as defined earlier, therefore $\mathcal{G}$ is a linear lattice.

The set of all non-negative gambles in $\mathcal{K}$ is denoted by $\mathcal{G}_{\geq 0}(\mathcal{X}) = \{f \in \mathcal{G}_{\geq 0}(\mathcal{X}) : f \geq 0\}$. $\mathcal{G}_{\geq 0}$ is a convex cone.

A partially ordered set is defined as follows.

Definition 3.2. *Let A be a set with a binary relation $\leq$ on it. Then A is a partially ordered set if and only if for all $x, y, z \in A$,* *(1) Reflexive:* $x \leq x$

(2) Transitive: $x \leq y$ *and* $y \leq z$ *implies that* $x \leq z$

(3) Antisymmetric: $x \leq y$ *and* $y \leq x$ *implies that* $x = y$.

We next define a real functional.

Definition 3.3. *A real functional is a real-valued function P defined on $\mathcal{G}$. It is monotone if $f \leq g$ implies that $P(f) \leq P(g)$ for all $f, g \in \mathcal{G}$.*

Properties of a real functional P are given below. These properties are desirable properties to have in them. For example, we will see that upper previsions are sub-additive, and that risk measures can be interpreted as upper previsions. This means that a risk measure can be sub-additive, a desirable property to have, because sub-additivity helps in minimizing risk through portfolio diversification.

Definition 3.4. *A real functional P defined on a convex cone of gambles $\mathcal{G}$ is:*

(1) Positively homogenous if $P(\lambda(f)) = \lambda(P(f))$ for all $f \in \mathcal{G}$ and $\lambda > 0$

(2) Non-negatively homogenous if $P(\lambda(f)) = \lambda(P(f))$ for all $f \in \mathcal{G}$ and $\lambda \geq 0$

(3) Homogenous if $P(\lambda(f)) = \lambda(P(f))$ for all $f \in \mathcal{G}$ and $\lambda \in \mathbb{R}$

(4) Convex if $P(\lambda(f) + (1 - \lambda(g)) \leq \lambda(P(f)) + (1 - \lambda)P(g)$ for all $f, g \in \mathcal{G}$ and $\lambda \in (0, 1)$

(5) Concave if $P(\lambda(f) + (1 - \lambda(g)) \geq \lambda(P(f)) + (1 - \lambda)P(g)$ for all $f, g \in \mathcal{G}$ and $\lambda \in (0, 1)$

If in addition, $\mathcal{G}$ is a linear space, then we call P:

(6) Super-additive if $P(f + g) \geq P(f) + P(g)$ for all $f, g \in \mathcal{G}$

(7) Sub-additive if $P(f + g) \leq P(f) + P(g)$ for all $f, g \in \mathcal{G}$

(8) Additive if $P(f + g) = P(f) + P(g)$ for all $f, g \in \mathcal{G}$

(9) Linear if $P(\lambda f + \mu g) = \lambda P(f) + \mu P(g)$ for all $f, g \in \mathcal{G}$ and $\lambda, \mu \in \mathbb{R}$.

An event is a subset $A \subseteq \mathcal{X}$. In classical probability theory a subset need to be a measurable set in order to be called an event. However in imprecise probability any subset of possibility space can be thought of as an event, since we are only going to define lower and upper previsions on them, just as de Finetti defined previsions.

Definition 3.5. *The indicator function I_A of event A is a gamble such that $I_A(x) = 1$ if $x \in A$, and $I_A(x) = 0$ if $x \notin A$.*

So events can be identified with their indicator functions. This is useful when we consider gambles as taking only two values, i.e., as events that either occur or do not occur. This restriction imposed on the gamble therefore obtain us a sub-class of imprecise probability models, in fact from the class of lower and upper previsions to the class of lower and upper probabilities which is discussed in Chapter 4.

An important theorem used in imprecise probability is the Hahn-Banach Extension Theorem. This theorem is used to establish a relationship between a lower prevision and the dominating set of linear previsions, using a notion of duality between the two. This is known as the lower envelope theorem (Theorem 4.31).

Schechter [54] gives different versions of the Hahn-Banach Theorem. We use a version that involves a linear space. The theorem that uses a linear space is as follows.

Theorem 3.6. *Let V be a linear space and U a linear subspace of V. Let ψ be a concave real functional on V dominated by a real linear functional ρ defined on U, i.e, $\rho(x) \geq \psi(x)$ for all $x \in U$. Then ρ can be extended to a real linear functional φ on V, and that dominates ψ, i.e., $\varphi(x) = \rho(x)$ for all $x \in U$, and $\varphi(x) \geq \psi(x)$ for all $x \in V$.*

The theorem and its proofs are given in Schechter [54] (Sections 12.31, and 12.35 to 12.37).

In summary, this chapter helps us to obtain a beginning idea of imprecise probability. The next chapter (Chapter 4) give us an idea of how the theory of lower prevision is developed. The chapter in its entirety practically follows from Troffeas & de Cooman [67]. In that chapter an attempt is made to also explain the proofs of theorems and propositions given in [67], to give us a simpler understanding of them. For each theorem and proposition, the beginning and end of the explanation of its proof is indicated by a diamond ($\Diamond$) sign. We now move to the theory of lower prevision.

Chapter 4

The Theory of Lower Previsions

In this chapter we study the theory of lower previsions. The purpose of writing this chapter is to gain an understanding for oneself of the theoretical basis of lower previsions, how the theory of lower previsions is developed, and to present the subject matter in such a way that it facilities a more convenient access later to the study of the original works of Walley [78], Troffaes & de Cooman [67], the book edited by Augustin, Coolen, de Cooman & Troffaes [4], and recent research.

The discussion below follows mainly from Chapters 3 & 4 of Troffaes & de Cooman. From the edited book that is cited above, we also discuss in part Chapter 1 written by Quaeghebeur and Chapter 2 written by Miranda & de Cooman. A useful additional reference which gives a quick and broad orientation to the subject matter of imprecise probability is the presentation by Miranda [41] given at 4th SIPTA Summer School, 2010.

We start with sets of desirable gambles which follow certain rationality criteria of behaviour laid down by us, and this would lead us to a theory of lower previsions. This is the approach of Williams [82]. Consideration of this set of desirable gambles will lead us to developing criteria for consistency, coherence, and natural extension of our assessments of a set of desirable gambles as well as buying and selling rates for gambles in this set. This is part of inference. Connection to the notions of prevision and fair price of de Finetti [11] are also shown.

Beliefs can be expressed through alternative models of imprecise probability. Therefore, in addition to developing the theory of coherent lower previsions as above and showing the upper prevision as its conjugate, an ongoing theme in this chapter is to bring out the interplay or correspondence with other models that can also express beliefs. Two such models are coherent sets of desirable gambles, and sets of linear previsions. We discuss these two models in this chapter.

For the axioms, theorems and propositions given in this chapter, an attempt is made in this study to give an explanation of the proofs that are originally there in the works cited above. It is important to note that no new proofs are given in this chapter. Explanation of the proofs for some of the theorems are given. As mentioned earlier, the explanation for the proofs begin and end with a diamond ($\Diamond$) sign. The reader who is not interested in the explanation of the proofs is welcome to skip them, skipping would involve no loss in continuity or in understanding. It is hoped that remarks given before and after the theorems and the proofs would give an adequate understanding.

4.1 Sets of Desirable Gambles

We want to model beliefs of an individual from his observed behaviour. We first say that he has some knowledge or information about the outcomes in the possibility space $\mathcal{X}$. Based on this knowledge and information, he is willing to accept or not accept to transact in certain gambles in $\mathcal{G}$. Specification by the individual of the gambles acceptable to him reveals his beliefs, and we call the specified set of acceptable gambles as the set of desirable gambles $\mathcal{D}$. Initially he does not have the complete set $\mathcal{D}$ but a smaller set. This is explained further a little later. Now suppose that $\mathcal{D}_1$ and $\mathcal{D}_2$ are sets of desirable gambles of two individuals such that $\mathcal{D}_1 \subseteq \mathcal{D}_2$, then the set inclusion implies that the gambles in $\mathcal{D}_1$ are at most as desirable as those in $\mathcal{D}_2$. We say that $\mathcal{D}_1$ is at most as informative as, or at most as committal as $\mathcal{D}_2$.

Rules imposed on an individual for including or not including gambles in $\mathcal{D}$ are given by the following rationality axioms (Troffaes & de Cooman [67]).

Axiom 4.1.

Let $\mathcal{G}$ be the set of all gambles. Then for any $f, g \in \mathcal{G}$

(1) Addition

$f, g \in \mathcal{D}$ implies that $f + g \in \mathcal{D}$

(2) Non-negative Scaling

$f \in \mathcal{D}$ implies that $\lambda f \in \mathcal{D}$ for all $\lambda \geq 0$

(3) Avoiding partial loss

$f < 0$ implies that $f \notin \mathcal{D}$

(4) Accepting partial gain

$f \geq 0$ implies that $f \in \mathcal{D}$

Axiom (1) states that a combination of two desirable gambles is also desirable. Axiom (2) states that for a desirable gamble, a higher or lower stake of it is also desirable. These axioms are part of constructive rationality, which assumes that the rewards from the gambles can be expressed in units of utility which has a linear scale. It effectively answers the question that if some gambles are desirable then are there other gambles that are desirable as well (Augustin et al. [4]).

Axiom (3) states that a gamble that could give a negative reward but never a positive reward whatever be the outcome is not desirable. Axiom (4) states that a gamble that could give a positive reward but never a negative reward whatever be the outcome of the game is desirable. The two rationality axioms effectively answers the question whether there are some gambles that we can include as clearly desirable, and some gambles that we can eliminate as clearly not desirable (Augustin et al. [4]). In the above, Troffaes (also following Williams) assumes that the zero gamble is desirable, since it will not give a negative reward.

A consequence of Axioms (1) and (2) is as follows. Is is about dominance of gambles.

If $f \in \mathcal{D}$ and $g \geq f$, then $g \in \mathcal{D}$. This means that a gamble that dominates a desirable gamble is also desirable. To show this:

$\Diamond$ Suppose that $f \in \mathcal{D}$ and $g \geq f$. So $g - f \geq 0$ and by Axiom (4) $g - f \in \mathcal{D}$. Write gamble g as $g = f + (g - f)$ and by Axiom (1) $g \in \mathcal{D}$. $\Diamond$

We are now ready to define the notion of coherence for sets of desirable gambles.

Definition 4.2. *Any set of desirable gambles $\mathcal{D}$ that satisfies the rationality criteria of avoiding partial loss, accepting partial gain, non-negative scaling and addition of gambles as given in Axiom 4.1 is called coherent. The collection of all coherent sets of desirable gambles on $\mathcal{X}$ is denoted by $\mathbb{D}(\mathcal{X})$ or just $\mathbb{D}$.*

It follows from the axioms that a coherent set of desirable gambles is a convex cone in the linear space of gambles $\mathcal{G}$ (see also Definition 3.1). It includes the non-negative first orthant, and excludes the negative third orthant.

We have mentioned earlier that initially an individual does not have set $\mathcal{D}$. Let us now see how we can obtain a coherent set of desirable gambles starting from an assessment, which is actually a process of inference. Let us call a partial specification of the individual's set of desirable gambles as an assessment $\mathcal{A}$ (this could be a finite set, and $\mathcal{A} \subseteq \mathcal{D}$). These are the gambles that the individual finds acceptable and chosen to transact in. We want to know whether it is possible to extend this partial specification to a coherent set of desirable gambles. The answer is yes. We do this by adding more gambles to $\mathcal{A}$. We want this addition to be minimal so as to add only those gambles that are needed in order to achieve coherence.

Once the agent specifies an assessment of a set of desirable gambles, the construction of a coherent set of desirable gambles will proceed as follows. We first check whether the assessment is consistent (i.e., whether they avoid partial loss), or find conditions when the assessment can be consistent. When an assessment of gambles is consistent, then for inference this set is extended to give a coherent set of desirable gambles. To get the extension, we take the union of gambles in the assessment with that of the gambles in the first orthant, and then take the positive hull (applying constructive rationality). This is called the natural extension of the assessment (and will contain $\mathcal{A}$). The natural extension therefore helps us to infer which gambles are desirable when given an assessment $\mathcal{A}$.

Definition 4.3. *Let $\mathcal{A}$ be an assessment, and consider gambles g in the non-negative orthant of the space of all gambles. Then the natural extension of $\mathcal{A}$ is a set defined by*

$$\mathcal{E}_{\mathcal{A}} \coloneqq \{g + \Sigma_{k=1}^{n} \lambda_k f_k \colon \lambda_k \geq 0, g \geq 0, f_k \in \mathcal{A}, n \in \mathbb{N}\} \tag{4.1}$$

So using Definition 4.3, we form the set $\mathcal{E}_{\mathcal{A}}$ for a given $\mathcal{A}$. We know that $\mathcal{E}_{\mathcal{A}}$ is the smallest set such that $\mathcal{A} \subseteq \mathcal{E}_{\mathcal{A}}$, and satisfies three conditions for coherence, viz., accepting partial gain, non-negative scaling, and addition.

We now want to find an assessment $\mathcal{A}$ that is consistent. Consistency is defined as follows.

Definition 4.4. *A set $\mathcal{A}$ of desirable gambles is called consistent or avoids partial loss if one of the following two conditions is satisfied. The two conditions are equivalent.*

(1) $\mathcal{A}$ is contained in some coherent set of desirable gambles, and the collection $\mathbb{D}$ is non-empty, i.e., $\{\mathcal{D} \in \mathbb{D} | \mathcal{A} \subseteq \mathcal{D}\} \neq 0$.

(2) For all $n \in \mathbb{N}$, and all $\lambda_1, ..., \lambda_n \in \mathbb{R}_{\geq 0}$, and gambles $f_1, ..., f_n \in \mathcal{A}$, we have $\Sigma_{k=1}^{n} \lambda_k f_k \geq 0$.

$\Diamond$ We note that Condition (2) shows that consistency is the same as avoiding partial loss. Consider any $\mathcal{D}$ that contains $\mathcal{A}$. A non-negative linear combination of gambles in $\mathcal{A}$ is in $\mathcal{D}$, by conditions of non-negative-scaling and addition (Definition 4.1). But by condition of avoiding partial loss (Definition 4.1) $g \geq 0$.

To show the converse, notice that the elements of $\mathcal{E}_{\mathcal{A}}$ satisfy the three conditions of coherency. From (2), each gamble $h \in \mathcal{E}_{\mathcal{A}}$ is such that $h \geq 0$, i.e., avoids partial loss, and so forms a coherent set (Definition 4.1). So $\mathcal{A}$ is consistent since $\mathcal{A} \subseteq \mathcal{E}_{\mathcal{A}}$. $\Diamond$

Condition (2) can be taken as our criterion for consistency. So, if $\mathcal{A}$ is consistent then it can be extended to a coherent set of desirable gambles. Among all such extensions we are interested in finding the smallest extension. This is obtained by taking the intersection, as follows.

Proposition 4.5. *Consider a non-empty collection of sets of desirable gambles, $\mathcal{D}_i$, $i \in I$. If each D_i is coherent then $\cap_{i \in I} \mathcal{D}_i$ is coherent.*

$\Diamond$ If each D_i is coherent and non-empty, then their intersection $\cap_{i \in I} \mathcal{D}_i$ is coherent. $\Diamond$

The intersection of all coherent sets of gambles is the smallest set and would be taken as the least committal set. We want to be as conservative as possible, and find the smallest set that is coherent. We may say that a certain property is preserved under intersection. Then we may define closure of $\mathcal{A}$, denoted by $Cl(\mathcal{A})$, as the smallest coherent set that contains $\mathcal{A}$. This is the same as its natural extension $\mathcal{E}_{\mathcal{A}}$. This is explained below.

Definition 4.6. *Let $\mathbb{D}$ be the collection of all coherent sets of desirable gambles that contain $\mathcal{A}$, the closure of $\mathcal{A}$ is defined as, $Cl_{\mathbb{D}}(\mathcal{A}) \coloneqq \cap\{\mathcal{D} \in \mathbb{D} \,|\, \mathcal{A} \subseteq \mathcal{D}\}$.*

Theorem 4.7. *If the set $\mathcal{A}$ of desirable gambles is consistent, then there is a smallest coherent set of desirable gambles that contains $\mathcal{A}$. It is given by,*

$Cl_{\mathbb{D}}(\mathcal{A}) = \mathcal{E}_{\mathcal{A}}$

$= \{g + \Sigma_{k=1}^{n} \lambda_k f_k \colon g \geq 0, \lambda_k \geq 0, f_k \in \mathcal{A}, n \in \mathbb{N}, k = 1, ..., n\}$

$= \{h \in \mathcal{G} \colon h \geq \Sigma_{k=1}^{n} \lambda_k f_k,\ \lambda_k \geq 0, f_k \in \mathcal{A}, n \in \mathbb{N}, k = 1, ..., n\}$.

$\mathcal{E}_{\mathcal{A}}$ is the natural extension of $\mathcal{A}$.

$\Diamond$ Suppose that $\mathcal{A}$ is consistent. Then by Proposition 4.5, $Cl_{\mathbb{D}}(\mathcal{A})$ is a coherent set that contains $\mathcal{A}$. It is contained in every coherent set $\mathcal{D}$ that contains $\mathcal{A}$ and is the smallest.

To show the equality $Cl_{\mathbb{D}}(\mathcal{A}) = \mathcal{E}_{\mathcal{A}}$. From conditions (2)-(4) of Axiom 4.1, $\mathcal{E}_{\mathcal{A}}$ is contained in any coherent set $\mathcal{D}$ that contains $\mathcal{A}$, and so contains $Cl_{\mathbb{D}}(\mathcal{A})$. $\mathcal{E}_{\mathcal{A}}$ is a coherent set of desirable gambles, since each element h satisfies conditions (1)-(4) of Axiom 4.1. $\Diamond$

So in this section we give the axioms of rationality and define coherence of a set of desirable gambles. We then see that given some assessment of gambles we can apply the axioms of rationality to obtain a coherent set of desirable gambles. We also see that we can make this set as less committal as possible.

4.2 Going from a Set of Desirable Gambles to Lower and Upper Previsions

A coherent set of desirable gambles will help us to define lower and the upper prevision. We take the lower prevision of a gamble as the supremum buying rate for a gamble above which the individual cannot be induced to bet for the gamble. The upper prevision is taken as the infimum selling rate below which the individual cannot be induced to bet against the gamble. We obtain both from a coherent set of desirable gambles.

Definition 4.8. *Let $\mathcal{G}$ be set of all gambles on $\mathcal{X}$, and $\mathcal{D} \subseteq \mathcal{G}$ be a coherent set of desirable gambles. Define real functionals*
$\underline{P}(\mathcal{D}) : \mathcal{G} \to \mathbb{R}$ *given by*

$$\underline{P}(\mathcal{D})(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathcal{D}\} \tag{4.2}$$

$\overline{P}(\mathcal{D}) : \mathcal{G} \to \mathbb{R}$ *given by*

$$\overline{P}(\mathcal{D})(f) := \inf\{\mu \in \mathbb{R} : \mu - f \in \mathcal{D}\}. \tag{4.3}$$

The real functionals (see Definitions 3.3 and 3.4) $\underline{P}(\mathcal{D})$ and $\overline{P}(\mathcal{D})$ are called lower prevision and upper prevision respectively, defined for each $f \in \mathcal{D}$.

The following theorem shows the properties for a lower prevision that is defined on a coherent set of desirable gambles as we have above. The three main consequences for lower prevision are that of lower bound, super-additivity and non-negative homogeneity. The remaining three properties can be derived from these, as shown in Troffaes & de Cooman [67].

Theorem 4.9.

Let $\underline{P}(\mathcal{D})$ be the lower prevision associated with a coherent set of desirable gambles $\mathcal{D}$. Then for all bounded gambles f, g in $\mathcal{D}$ and all $\lambda \geq 0$ and $\mu \in \mathbb{R}$ we have,

(1) Lower Bound

$\inf f \leq \underline{P}(\mathcal{D})(f)$

(2) Super-additivity

$\underline{P}(\mathcal{D})(f + g) \geq \underline{P}(\mathcal{D})(f) + \underline{P}(\mathcal{D})(g)$

(3) Homogeneity (non-negative)

$\underline{P}(\mathcal{D})(\lambda f) = \lambda \underline{P}(\mathcal{D})(f)$

(4) Upper Bound

$\underline{P}(\mathcal{D})(f) \leq \sup f$

(5) Constant Gamble

$\underline{P}(\mathcal{D})(\mu) = \mu$

(6) Additivity with a Constant

$\underline{P}(\mathcal{D})(f + \mu) = \underline{P}(\mathcal{D})(f) + \mu$

$\Diamond$ (1). We know that $f - \inf f \geq 0$ and so f is in $\mathcal{D}$ by the axiom of accepting partial gain. It follows from the definition of a lower prevision that $\inf f \leq \underline{P}(\mathcal{D})(f)$ for all $f \in \mathcal{D}$.
(2). Suppose that $f, g \in \mathcal{D}$. Let $\alpha < \underline{P}(\mathcal{D})(f)$ and $\beta < \underline{P}(\mathcal{D})(g)$. Then $f - \alpha \in \mathcal{D}$ and $g - \beta \in \mathcal{D}$, and so $(f+g)-(\alpha+\beta) \in \mathcal{D}$ by the axiom of combination. Then from the definition of lower prevision, it follows that $\underline{P}(f+g) \geq \alpha + \beta$. But $\sup \alpha < \underline{P}(\mathcal{D})(f)$, and $\sup \beta < \underline{P}(\mathcal{D})(g)$. We get $\underline{P}(\mathcal{D})(f+g) \geq \underline{P}(\mathcal{D})(f) + \underline{P}(\mathcal{D})(g)$.
(3). Let $\lambda > 0$. Consider the gamble λf. From Axiom 4.1 (1) we have that if $\lambda f - \mu \in \mathcal{D}$ then $f - \mu/\lambda \in \mathcal{D}$. We apply this to Definition 4.2 for lower prevision to obtain the non-homogeneity result.
(4). From (1), we have $\inf(-f) \leq \underline{P}(\mathcal{D})(-f)$. From (2), $\underline{P}(f) + \underline{P}(\mathcal{D})(-f) \leq \underline{P}(\mathcal{D})(f-f) = \underline{P}(\mathcal{D})(0) = 0$, by (3) and $\lambda = 0$, i.e., $\underline{P}(\mathcal{D})(f) \leq -\underline{P}(\mathcal{D})(-f)$. Therefore we get, $\sup f = -\inf(-f) \geq -\underline{P}(\mathcal{D})(-f) \geq \underline{P}(\mathcal{D})(f)$.
(5). For the constant gamble $\mu \in \mathcal{D}$, we have $\inf \mu = \sup \mu = \mu$. So by (1) and (4), $\underline{P}(\mathcal{D})(\mu) = \mu$.
(6). By (2) and by (5) of lower prevision we have, $\underline{P}(\mathcal{D})(f+\mu) \geq \underline{P}(\mathcal{D})(f)+\mu$. Also by (2), $\underline{P}(\mathcal{D})((f+\mu)-\mu) \geq \underline{P}(\mathcal{D})(f+\mu) + \underline{P}(\mathcal{D})(-\mu)$, i.e., $\underline{P}(\mathcal{D})(f) \geq \underline{P}(\mathcal{D})(f+\mu) - \mu$, by (5). Therefore, $\underline{P}(\mathcal{D})(f+\mu) = \underline{P}(\mathcal{D})(f) + \mu$.
$\Diamond$

We shall see later that these are indeed the properties of a coherent lower prevision. There we also see the connection between a coherent lower prevision and an associated coherent set of desirable gambles (see Definition 4.16 and Theorem 4.20).

Continuing from the theorem above, suppose that we have defined a real functional $\underline{P}$ on $\mathcal{G}$. We ask the question whether it is possible to, or under what conditions can we associate a $\underline{P}$ so defined with some coherent set of desirable gambles $\mathcal{D}$ in $\mathcal{G}$. Note that a set of desirable gambles would be coherent if the Axioms 4.1 are satisfied.

The theorem below states the following. If we have a real functional defined as above and associate it as the lower prevision of a coherent set of desirable gambles, then that functional will satisfy the first three properties given above. On the other hand, if we are given a real functional that satisfies the three properties and associate it with a set of desirable gambles, then that set will be coherent, meaning that the set will satisfy the four axioms of coherence. Here we see the connection between certain properties of lower prevision and the coherence of a set of desirable gambles.

Theorem 4.10. *Suppose that a real functional $\underline{P}$ is defined on $\mathcal{G}$ and associate it such that $\underline{P} = \underline{P}(\mathcal{D})$, then $\mathcal{D} \subseteq \mathcal{G}$ is a coherent set of desirable gambles if and only if $\underline{P}$ satisfies properties (1)-(3) of Theorem 4.9.*

$\Diamond$ The "only if" part is already shown in the theorem above. To show the converse please see Troffaes & de Cooman [67] $\Diamond$

So far we have used a coherent set of desirable gambles and obtained lower and upper previsions. But is it necessary or possible that the individual has or can know about any coherent set of desirable gambles in order to find lower and upper previsions? According to Troffaes & de Cooman [67], it is possible to define lower and upper prevision for a gamble without the knowledge of a coherent set of desirable gambles.

The lower prevision $\underline{P}(f)$ for a gamble f can be thought of as the highest price s such that for any $t < s$, the individual says he accepts to pay t before observing X, if he is guaranteed to receive $f(x)$ when observed that $X = x$. So the individual is saying that he accepts gambles of the type $f - \mu$ for all $\mu < \underline{P}(f)$. The fixed μ can be thought of as some certainty equivalent of a random f though not implied.

Similarly, the upper prevision $\overline{P}(f)$ for a gamble f can be thought of as the lowest price s such that for any $t > s$, the individual says he accepts to receive t before observing X, if he is guaranteed to give $f(x)$ when observed that $X = x$. So the individual is saying that he accepts gambles of the type $\mu - f$ for all $\mu > \overline{P}(f)$. Again, fixed μ can be thought of as some certainty equivalent of a random f.

Here, a lower prevision is a real functional defined on some subset $D(\underline{P})$ of $\mathcal{G}$ called the domain of $\underline{P}$. Similarly, $\overline{P}$ is a real functional defined on $D(\overline{P}) \subseteq \mathcal{G}$. We are saying that the domains of lower and upper prevision are some subsets of $\mathcal{G}$. Later on we would examine extending lower and upper prevision from their respective domains to the set of all gambles.

There is a relationship between lower and upper prevision, it is the property of conjugacy. Conjugacy is obtained from the following. Using the definitions of lower and upper previsions we get, $\underline{P}(-f) = \sup\{x\colon -f - x \in \mathcal{D}\} = -\inf\{-x\colon -f - x \in \mathcal{D}\} = -\inf\{y\colon -f + y \in \mathcal{D}\} = -\overline{P}(f)$. So given a $\underline{P}$ with $D(\underline{P})$ we define its conjugate $\overline{P}$ on $-D(\underline{P}) = \{-f | f \in D(\underline{P})\}$ where buying rate for f is the same as negative of the selling rate for $-f$, i.e., $\underline{P}(f) = -\overline{P}(-f)$. Since they are conjugates of one another, we can work with any one imprecise probability model, either lower prevision or upper prevision (Quaeghebeur [4]). We shall choose lower previsions $\underline{P}$ as the model in our study.

So in this section we see that given a coherent set of desirable gambles one can define lower and upper prevision for each gamble in the set which have certain properties. We also see that an individual can assess lower and upper prevision even without knowledge of a coherent set of gambles, but we do not know whether they are consistent or coherent. Finally, the conjugate relation between lower and upper prevision allow us to choose any one of the two as our imprecise probability model.

4.3 Consistency of Lower Previsions

From the previous sections we see that an individual's assessment could be made not only in terms of certain desirable gambles but also in terms of a lower prevision, meaning that he can instead assess a supremum buying price for each one of them. So we now ask the larger question whether this information will help him in assessing desirability of other gambles, in terms of assessing lower previsions for them. We answer this question by beginning with the notion of consistency of a lower prevision, the first criterion of rational behaviour of an individual. We have defined earlier that an assessment $\mathcal{A}$ is consistent if it avoids partial loss (Definition 4.4).

To start with, we use the ideas of dominating lower previsions and set inclusions. The primary purpose of using this construction to bring out and show the relation between sets of desirable gambles and that of the corresponding previsions assessed for them. This is achieved through using the notions of dominance of one lower prevision over another and the corresponding set inclusions of their corresponding sets of desirable gambles. As mentioned earlier, the theme of correspondence between different imprecise probability models is prominent in this chapter.

We say that lower prevision $\underline{Q}$ is at least as informative as $\underline{P}$ when $\underline{Q}$ dominates $\underline{P}$. When we want to express this in terms of a set inclusion relation then we say that $D(\underline{Q})$ is at least as committal as $D(\underline{P})$. This leads to the following definition.

Definition 4.11. *A lower prevision $\underline{Q}$ dominates a lower prevision $\underline{P}$ if $D(\underline{P}) \subseteq D(\underline{Q})$, and $\underline{P}(f) \leq \underline{Q}(f)$ for every gamble $f \in D(\underline{P})$.*

A consequence of this definition is the implication for their natural extensions. We do not need this result right now but would need it later.

Proposition 4.12. *Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two consistent sets of desirable gambles. If $\mathcal{A}_1 \subseteq \mathcal{A}_2$, then $Cl_{\mathbb{D}}(\mathcal{A}_1) \subseteq Cl_{\mathbb{D}}(\mathcal{A}_2)$ and $\underline{P}(Cl_{\mathbb{D}}(\mathcal{A}_1)) \leq \underline{P}(Cl_{\mathbb{D}}(\mathcal{A}_2))$*

$\Diamond$ This proposition follows from Definitions 4.6 and 4.11. $\Diamond$

To start with, suppose that a lower prevision $\underline{P}$ has been defined for a set of desirable gambles. We want to obtain a subset of this set of gambles whose lower previsions are dominated by $\underline{P}$. We define this set as follows.

$$\mathcal{A}_{\underline{P}} \coloneqq \{f - \mu \,|\, f \in D(\underline{P}),\, \mu < \underline{P}(f)\}. \tag{4.4}$$

Essentially, what we are doing here is going from a given lower prevision to a set of desirable gambles. Our first concern would be about whether this set of desirable gambles is consistent. The following set of definitions bring together implications of the criterion of consistency. Only the first four implications given in Troffaes & de Cooman [67] are stated here.

Definition 4.13. *We say that a lower prevision $\underline{P}$ is consistent or avoids sure loss, if any or all of the following equivalent conditions are satisfied.*

(1) The set of desirable gambles $\mathcal{A}_{\underline{P}}$ is consistent.

(2) There is some coherent set of desirable gambles $\mathcal{D}$ such that $\underline{P}(\mathcal{D})$ dominates $\underline{P}$.

(3) There is some lower prevision $\underline{Q}$ defined on the set $\mathcal{G}$ of all gambles that satisfies properties (1)-(3) of Theorem 4.9 and that dominates $\underline{P}$.

(4) For all $n \in \mathbb{N}$, and gambles $f_1, ..., f_n \in D(\underline{P})$, it is true that $\Sigma_{i=1}^n \underline{P}(f_i) \leq \sup \Sigma_{i=1}^n f_i$.

$\Diamond$ (1) implies (2). Suppose that the set of desirable gambles $\mathcal{A}_{\underline{P}}$ is consistent. Then by Definition 4.4 (1), we have that $\mathcal{A}_{\underline{P}}$ is included in a coherent set $\mathcal{D}$ of desirable gambles. We prove (2) by showing that the lower prevision defined on $\mathcal{D}$ dominates $\underline{P}$.
(2) implies (3). We let $\underline{Q}$ be the lower prevision defined on a coherent set $\mathcal{D} \subseteq \mathcal{G}$ of gambles. This lower prevision dominates $\underline{P}$, and by Theorem 4.9 the three properties are satisfied.
(3) implies (4). Suppose that lower prevision $\underline{Q}$ defined on set $\mathcal{G}$ of all gambles satisfies properties (1)-(3) of Theorem 4.9, and $\underline{Q} \geq \underline{P}$. Consider gambles $f_0, ..., f_n \in D(\underline{P})$. We then use Theorem 4.9 (2) and (4) to obtain the inequality. $\Diamond$

Under the criterion of consistency, the four definitions are bringing together the connection or relationship between four notions, viz., the notion of a consistent lower prevision, a consistent set of desirable gambles, existence of a dominating coherent set of desirable gambles and the corresponding lower prevision that dominates, existence of a lower prevision that satisfies the three properties of a lower prevision defined on a coherent set of desirable gambles, and the notion of avoiding sure loss.

Definition (4) is the one that an individual would use as the criterion for avoiding sure loss given the lower previsions assessed for his set of desirable gambles. In fact, this is the one that we would use later in our computations to check for consistency of lower previsions.

We may note that avoiding partial loss can be taken as the stronger requirement, since avoiding partial loss implies avoiding sure loss. An example of avoiding sure loss would be as follows. If one team is to take part in twenty soccer matches, but knows for certain about the outcomes that it would not win any of them, then avoiding sure loss would be the action or the rational decision taken by the team not to play any match – assuming that losing or obtaining a tie in a game are not desirable outcomes.

4.4 Coherence of Lower Previsions

We continue with our inference problem. We asked earlier that in the inference problem, given assessments about some desirable gambles what can we say for the assessment about other gambles. The assessments that we refer to are now in terms of lower previsions. One could be concerned with assessment about any existing gamble vis-a-vis the assessments already made about other gambles in the set, or it could be about new gamble(s) that the individual is considering. Naturally, they ought to be consistent with each another in some manner, in the sense that assessing a supremum buying price for any gamble should not depend on one's assessments of supremum buying price for other gambles in our desirable set. This is the second step of rationality criteria which we call as the criterion of coherence of assessments of lower previsions. The process of revising or correcting an incoherent assessment of a gamble (if any) to coherence, in order to meet this rationality criterion, is also what we call its natural extension. We discuss these in the following.

We have seen that if the assessed lower prevision $\underline{P}$ is consistent, then the set of desirable gambles $\mathcal{A}_{\underline{P}}$ is also consistent. This set then can be extended to give a coherent set of desirable gambles. The smallest extension is the natural extension $Cl_{\mathbb{D}}(\mathcal{A}_{\underline{P}})$. Recall that this set was defined earlier (see Theorem 4.7 and Equation 4.4). From these we have the following.

$$\begin{aligned} Cl_{\mathbb{D}}(\mathcal{A}_{\underline{P}}) &= \mathcal{E}_{\mathcal{A}_{\underline{P}}} \\ &= \{h \in \mathcal{G} | h \geq \epsilon + \Sigma_{k=1}^{n} \lambda_k (f_k - \underline{P}(f_k)), \epsilon > 0, n \in \mathbb{N}, f_k \in D(\underline{P}), \lambda_k \in \mathbb{R}_{\geq 0}\}. \end{aligned} \tag{4.5}$$

We can think of $f_k - \underline{P}(f_k)$ as the marginal gamble of f_k under $\underline{P}(f_k)$. Under the assumption of consistency, these marginal gambles are also consistent (avoid sure loss). So what we are doing here is adding gambles h in $\mathcal{G}$ that dominate a positive linear combination of these marginal gambles. The $\epsilon > 0$ makes sure that h is strictly desirable over these.

We now say that for this extended coherent set of desirable gambles we can find a lower prevision. In fact we are defining this lower prevision as a real functional on $\mathcal{G}$.

Definition 4.14. *Let $\underline{P}$ be a lower prevision that is consistent (avoids sure loss). Then the lower prevision defined as $\underline{E}_{\underline{P}} \coloneqq \underline{P}(Cl_{\mathbb{D}}(\mathcal{A}_{\underline{P}})) = \underline{P}(\mathcal{E}_{\mathcal{A}_{\underline{P}}})$ is called the natural extension of $\underline{P}$, defined for all $f \in \mathcal{G}$, and is given as,*

$$\begin{aligned} &\underline{E}_{\underline{P}}(f) \\ &\quad = \sup\{\alpha \in \mathcal{R} | f - \alpha \geq \Sigma_{k=1}^{n} \lambda_k (f_k - \underline{P}(f_k)), n \in \mathbb{N}, f_k \in D(\underline{P}), \lambda_k \in \mathbb{R}_{\geq 0}\},\ f \in \mathcal{G}. \end{aligned} \tag{4.6}$$

This equation follows from Equation 4.2 (definition of a lower prevision) and from Equation 4.5 above.

We next consider restricting a given $\underline{P}$ to a smaller domain, and similarly extending $\underline{P}$ to a larger domain. This gives us the notions of a restriction and an extension of a lower prevision. When we restrict or extend a lower prevision we want to know the implications for consistency and coherence. The definition and the implications are given in the following.

Definition 4.15. *Let $\underline{P}$ and $\underline{Q}$ be lower previsions. Then $\underline{P}$ is said to be a restriction of $\underline{Q}$, and $\underline{Q}$ is said to be an extension of $\underline{P}$, whenever $D(\underline{P}) \subseteq D(\underline{Q})$ and $\underline{P}(f) = \underline{Q}(f)$ for all $f \in D(\underline{P})$.*

The implications of coherence of lower previsions are as follows. Only the first four definitions in [67] is stated here.

Definition 4.16. *Let $\underline{P}$ be a lower prevision. Then $\underline{P}$ is coherent if and only if any (and therefore all) of the following hold.*
(1) There is some coherent set of desirable gambles $\mathcal{D}$ such that $\underline{P}(\mathcal{D})$ is an extension of $\underline{P}$.
(2) There is some lower prevision $\underline{Q}$ on $\mathcal{G}$ that satisfies properties (1)-(3) given in Theorem 4.9, that is an extension of $\underline{P}$.
(3) $\underline{P}$ avoids sure loss and it is a restriction of its natural extension $\underline{E}_{\underline{P}}$.
(4) For all $m, n \in \mathbb{N}$ and gambles $f_0, ..., f_m$ in $D(\underline{P})$ it is true that $\Sigma_{i=1}^{n}(f_i - \underline{P}(f_i)) \geq m(f_0 - \underline{P}(f_0))$.

◊ The proofs showing equivalence are fairly lengthy, therefore an explanation for the original proofs is not given here. The reader is directed to see the original proofs given in Troffaes & de Cooman [67]. ◊

Definition (1) states that if a lower prevision is coherent then its extension is also coherent. (2) states that if $\underline{P}$ is coherent then there exists an extension to this in the space of all gambles that satisfies the properties of lower boundedness, super-additivity and non-negative homogeneity. (3) states that a lower prevision that avoids sure loss but is a restriction of its natural extension is coherent.

Definition (4) may be taken as our general definition of coherence. Notice that if we set $m = 0$, then this definition reduces to the criterion for consistency, which is the criterion of avoiding of sure loss. This would mean that there would not be any opportunity for arbitrage. This is explained below.

An arbitrage position or situation is one where there is an opportunity for one to be able to (simultaneously) buy a gamble in one market and sell the same in another market to make a sure gain. If this occurs, then the competitor is incurring a sure loss. If this is so, then the lower prevision cannot be coherent. Therefore coherence implies the absence of any opportunity for arbitrage. We shall visit this point again in Chapter 6 when we compute risk measures from historical data, as we would expect the computed risk measures there to be coherent.

To show Definition (4), suppose that it is not true. Then there exists an $\epsilon > 0$ such that $\Sigma_{i=1}^{n}(f_i - (\underline{P}(f_i) - \epsilon)) < m(f_0 - (\underline{P}(f_0) + \epsilon))$. The gambles on the left-hand side of this inequality are desirable, and it is a linear combination of desirable gambles (Axiom 4.1 (1)). The gamble on the right-hand side strictly dominates this and therefore is also desirable (consequence of Axioms 4.1 (1) and (2)). This means that $\underline{P}(f_0) + \epsilon$ is an acceptable supremum buying rate for gamble f_0 which contradicts our assumption that $\underline{P}(f_0)$ is the supremum buying rate.

So (4) states that while considering the lower prevision setting for gambles, coherence implies that there will be no scope left for relaxing the maximum buying rate that has been set for any in that set. We will later use this criterion in our calculations when we check for coherence of risk measures.

4.5 Linear Previsions and Probability

Before we move on to some properties of coherent lower prevision, we discuss three notions here. One notion is to see a self-conjugate lower prevision as a linear prevision. The second one is to see precise probability as a special case of an imprecise probability. The third notion is to see the connection between linear prevision and a set of desirable gambles.

Our discussion in Section 4.2 on lower and upper previsions as conjugates of one another lead us to the notion of self-conjugacy. If $D(\underline{P}) = D(\overline{P}) = -D(\underline{P})$ then $\underline{P}$ is said to be self-conjugate, where $\underline{P}(f) = \overline{P}(f) = P(f)$ (say). So the upper and lower prevision reduce to one value $P(f)$ which we call the prevision of f. This is the "fair price" of de Finetti [10, 11]. There is no region of indifference (indecision), we have a situation where the bet is now two-sided in the sense that we would buy gamble f for prices $s < P(f)$, and sell gamble f for prices $s > P(f)$. The region of indifference reduces from an interval to a point.

In fact, the linear prevision can be seen as an expectation. We show below in this section that the linear prevision is an expected value when the possibility space $\mathcal{X}$ is finite.

The above lead us to examine the situation when imprecise probability reduces to precise probability. Following Definition 3.5 of an indicator function, we denote the lower prevision $\underline{P}(I_A)$ of an indicator I_A as $\underline{P}(A)$. So for events $A \subseteq \mathcal{X}$, the lower prevision $\underline{P}(I_A)$ is the lower probability $\underline{P}(A)$ of event A, and similarly the upper prevision $\overline{P}(I_A)$ is the upper probability $\overline{P}(A)$ of event A. When $\underline{P}$ is self-conjugate, we have $\underline{P}(A) = \overline{P}(A) = P(A)$. Then $P(A)$ is called the probability of event A. In the case when the domain of prevision P contains only indicator functions I_A along with negations $-I_A$, then we say that P is a probability.

So we have prevision when lower prevision is self-conjugate. It is being defined on a negation-invariant set of gambles (see Definition 3.1 (1)). If the set is not negation-invariant, then it can be extended to become negation-invariant by taking the union $D(\underline{P}) \cup -D(\underline{P})$ such that $P(-f) = P(f)$. Interestingly for the case of previsions, coherence is the same as avoiding sure loss for a gamble. In fact, de Finetti [11] had defined coherence of a prevision as avoiding sure loss as we saw in the earlier chapter. We note this in the following theorem.

We first define a linear prevision as a coherent lower prevision that is self-conjugate.

Definition 4.17. *A lower prevision is a linear or a coherent prevision, if it is self-conjugate and coherent. The set of all linear previsions with domain $\mathcal{K} \subseteq \mathcal{G}$ is denoted as $\mathcal{P}^{\mathcal{K}}$. The set of all linear previsions is denoted as $\mathcal{P}$.*

The theorem below clarifies the connections between the notions of avoiding sure loss and coherence for lower prevision and linear prevision, in fact some equivalences.

Theorem 4.18. *A prevision P is linear if and only if one of the following conditions is satisfied.*

(1) P, as a lower prevision, is coherent.

(2) P, as a lower prevision, avoids sure loss.

(3) For all $n \in \mathbb{N}$, $\lambda_1, ..., \lambda_n \in \mathbb{R}$, and gambles $f_1, ..., f_n \in D(P)$, it is true that $\Sigma_{i=1}^{n} \lambda_i P(f_i) \leq sup(\Sigma_{i=1}^{n} \lambda_i f_i)$.

(4) For all $n \in \mathbb{N}$, and gambles $f_1, ...f_n \in D(P)$, it is true that $\Sigma_{i=1}^{n} P(f_i) \leq sup(\Sigma_{i=1}^{n} f_i)$.

n$\Diamond$ A coherent lower prevision $\underline{P}$, is a coherent prevision P, if it is also self-conjugate. It follows from Definition 4.17 that the coherent prevision P is a linear prevision and conversely also.

(1) implies (2). From Definition 4.16 (3), we know that P, as a lower prevision, avoids sure loss.

(2) implies (3). We know that for any single gamble $f \in D(P)$, in order to avoid sure loss $P(f) \leq sup(f)$. The proof uses the fact that therefore, a net positive linear combination of the supremum values of gambles will dominate the corresponding combination of linear previsions to show that $\Sigma_{i=1}^{n} \lambda_i P(f_i) \leq sup(\Sigma_{i=1}^{n} \lambda_i f_i)$.

(3) implies (4). Let all $\lambda_i = 1$ in (3). Then (4) follows from (3).

(4) implies (1). By self-conjugacy, we know that P is a lower prevision. To show coherence, P needs to satisfy Definition 4.16 (4) of a coherent lower prevision. We can show this by a construction. We select

$n, m \in \mathbb{N}$, and an $f_0 \in D(P)$, and rewrite expression in (3) in the form of the definition. Since the definition is satisfied, P is coherent as a lower prevision. $\Diamond$

In particular, we note that coherence of a prevision assessed for a set of gambles is equivalent to avoiding sure loss in the sense of statement (4) of the above theorem, as was given by de Finetti.

We return to our discussion of probability. A restriction of linear prevision to events only is a finitely additive probability. When linear prevision P is defined on indicator function I_A of event $A \in \mathcal{F}$, where $\mathcal{F}$ is an algebra of subsets of $\mathcal{X}$, then P is a finitely additive probability called a *probability charge*. Then P satisfies the following probability axioms.

Axiom 4.19.

P1. $P(A) \geq 0$ *for all* $A \in \mathcal{F}$

P2. $P(\mathcal{X}) = 1$,

P3. $\cup_{i=1}^{n} A_i \in \mathcal{F}$ *for all* $A_i \in \mathcal{F}$ *and* $A_i \cap A_j = \emptyset$, $i \neq j$.

On the other hand when $\mathcal{X}$ is finite, P is a linear prevision on $\mathcal{G}$ if there is a function $p : \mathcal{X} \to [0, 1]$ with $\Sigma_{x \in \mathcal{X}}\ p(x) = 1$ such that $P(f) = \Sigma_{x \in \mathcal{X}}\ p(x)f(x)$ for all $f \in \mathcal{G}$. This shows the linear prevision as an expectation (Troffaes & de Cooman [67]).

In order to see the connection between linear prevision and desirable gambles, we ask the following question. If we are given a linear prevision $P \in \mathcal{P}$ then which gambles in $\mathcal{G}$ are desirable? We say that a gamble f is non-strictly desirable when its fair price is positive, i.e., $P(f) \geq 0$. If $P \in \mathcal{P}$ is given then,

$$\mathcal{D}_P \coloneqq \mathcal{E}_{A_P}, \quad \text{where} \quad \mathcal{A}_P \coloneqq \{f \in \mathcal{G} \colon P(f) \geq 0\}.$$

This gives us a way of obtaining the set of desirable gambles given a linear prevision as assessment.

A set of linear previsions is also an imprecise probability model. This set is called a credal set. A credal set can form when we are not able to precisely state the linear prevision but can only state it in terms of bounds, a lower and an upper bound. So we again ask the question, given a credal set $\mathcal{M} \subseteq \mathcal{P}$ which gambles in $\mathcal{G}$ are desirable? We form the following criterion for going from credal set to a set of desirable gambles.

$$\mathcal{D}_{\mathcal{M}} \coloneqq \mathcal{E}_{\mathcal{A}_{\mathcal{M}}}, \quad \text{where} \quad \mathcal{A}_{\mathcal{M}} \coloneqq \{f \in \mathcal{G} \colon \text{for all } P \in \mathcal{M},\ P(f) \geq 0\} = \bigcap_{P \in \mathcal{M}} \mathcal{A}_P.$$

Similarly, given a set of desirable gambles $\mathcal{D}$, we can form the credal set as follows.

$$\mathcal{M}_{\mathcal{D}} \coloneqq \bigcap_{f \in \mathcal{D}} \{P \in \mathcal{P} : P(f) \geq 0\}$$

.

So here we have a means of going from a set linear previsions to a set of desirable gambles, and vice-versa. This means that we can express beliefs as either a set of desirable gambles or as a set of linear previsions.

Before we complete this section we establish one more result, i.e., the set of linear previsions (credal set) is closed and convex. The set of all linear previsions is a closed and convex set. For any gamble f in $\mathcal{G}$ the linear prevision $P(f) \geq 0$ is a linear inequality giving a closed half space. The credal set is an intersection of closed and convex sets, therefore it is closed and convex. Credal sets are discussed in further detail in Quaeghebeur [4].

4.6 Some Properties of Coherent Lower and Upper Prevision

We saw in Theorem 4.9 some of the consequences for lower prevision defined for a coherent set desirable gambles. We summarize them here as properties of coherent lower and upper previsions. The properties are useful when we want to reason with gambles. For example, when we are assessing consistency and then coherence of a risk measure we check for boundedness. Similarly, we would like the risk measure to possess the property of sub-additivity when it is interpreted as an upper prevision.

Walley [78] gives and proves these properties. Proofs are also given in Troffaes & de Cooman [67]. Six of the properties are considered here.

Theorem 4.20. *Let $\underline{P}$ be a coherent lower prevision. Let $f, g, a \in \mathcal{G}$, where a is a constant gamble. Let $\lambda \in \mathbb{R}_{\geq 0}$. Then the following statements hold.*

(1) Boundedness:

$\inf f \leq \underline{P}(f) \leq \overline{P}(f) \leq \sup f$

(2) Constant Gamble:

$\underline{P}(a) = \overline{P}(a) = a$

(3) Additivity with a Constant:

$\underline{P}(f + a) = \underline{P}(f) + a$ *and,* $\overline{P}(f + a) = \overline{P}(f) + a$

(4) Monotonicity:

$f \leq g + a$ *implies that* $\underline{P}(f) \leq \underline{P}(g) + a$ *and* $\overline{P}(f) \leq \overline{P}(g) + a$.

(5) Super and Sub-additivity:

$\underline{P}(f) + \underline{P}(g) \leq \underline{P}(f + g)$, *and* $\overline{P}(f) + \overline{P}(g) \geq \overline{P}(f + g)$.

(6) Homogeneity (non-negative):

$\underline{P}(\lambda f) = \lambda \underline{P}(f)$, *and* $\overline{P}(\lambda f) = \lambda \overline{P}(f)$

Property (1) is useful to us for checking boundedness while computing coherence of lower prevision. Property (2) follows for that for a constant gamble. Property (3) is the implication for lower and upper prevision when we add constant gambles while getting an extension of our set of gambles. Similarly, Property (6) is the implication for lower and upper prevision while getting an extension of our set of gambles to achieve coherence. Property (5) admits that imprecise probability is non-additive.

When the lower prevision is coherent and self-conjugate, i.e., when $\underline{P}(f) = \overline{P}(f)$, the above theorem simplifies to giving properties of coherent (linear) previsions, as follows. Notice that prevision now is additive.

Corollary 4.21. *Let P be a coherent prevision. Let $f, g, a \in \mathcal{G}$, where a is a constant gamble. Let $\lambda \in \mathbb{R}_{\geq 0}$. Then the following statements hold.*

(1) Boundedness:

$\inf f \leq P(f) \leq \sup f$

(2) Constant Gamble:

$P(a) = a$

(3) Additivity with a Constant:

$P(f + a) = P(f) + a$

(4) Monotonicity:

$f \leq g + a$ *implies that* $P(f) \leq P(g) + a$

(5) Additivity:

$P(f) + P(g) = P(f + g)$

(6) Homogeneity (non-negative):

$P(\lambda f) = \lambda P(f)$

We have seen earlier that for a given upper prevision $\overline{P}$ and for all $f \in D(\overline{P})$, its corresponding conjugate lower prevision is given by $\underline{P}(f) = -\overline{P}(-f)$, where $-f \in D(\overline{P})$. So from conjugacy, we say that $\overline{P}$ is coherent when its corresponding conjugate lower prevision $\underline{P}$ is coherent. Similarly, we say that $\overline{P}$ avoids sure loss when its conjugate $\underline{P}$ avoids sure loss.

If $D(\underline{P})$ is a linear space, then coherence of $\underline{P}$ implies the following. Here we are asserting the correspondence between coherent lower previsions and the three properties of lower boundedness, positive homogeneity and super-additivity.

Theorem 4.22. *Let $\underline{P}$ be a lower prevision, and $D(\underline{P})$ a linear subspace of $\mathcal{G}$. Let $\lambda > 0$. Then $\underline{P}$ is coherent if and only if the following are true for all $f, g \in D(\underline{P})$.*
(1) $\underline{P}(f) \geq \inf f$.
(2) $\underline{P}(\lambda f) = \lambda \underline{P}(f)$.
(3) $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(f)$.

$\Diamond$ This proof is by Walley [78].
(Only if). Suppose that $\underline{P}$ is coherent. Then properties (1)-(3) above follow from Definition 4.20 (1), (5), and (6) respectively for lower previsions.
(If). Suppose that properties (1)-(3) above are true. We choose arbitrary values of n, m, and gambles $f_0, f_1, ..., f_n \in D(\underline{P})$, and define $f = mf_0, g = \Sigma_{k=1}^{n} f_k$, and $h = f - g$. Then using (1)-(3), we show that lower previsions satisfy the definition of coherence, and therefore are coherent. $\Diamond$

Coherence can be shown also when previsions are defined on a linear space of gambles (Finetti [11], Walley [78]). This is given in the following theorem, again showing the correspondence between coherent prevision and the properties of boundedness and additivity. In fact, additivity implies positive homogeneity.

Theorem 4.23. *Let P be a prevision, and assume that $D(P)$ is a linear subspace of $\mathcal{G}$. Then P is a coherent (linear) prevision if and only if the following conditions are met for all $f, g \in D(\underline{P})$.*

(1) $P(f) \geq \inf f$

(2) $P(f+g) = P(f) + P(g)$

$\Diamond$ (Only if). Suppose that $\underline{P}$ is coherent and $D(P)$ is linear. Then properties (1) and (2) above follow directly from Corollary 4.21 (1) and (5) respectively.
(If). Suppose that the properties (1), (2) are true. We choose arbitrary values of n and gambles $f_0, f_1, ..., f_n \in D(\underline{P})$ a linear space, and define $f = -\Sigma_{k=1}^{n} f_k$. Then $f \in D(\underline{P})$. Then using (1)-(2), we show that lower previsions satisfy the definition of coherence. $\Diamond$

Another property of lower prevision is that the lower envelope of a collection of coherent lower previsions is coherent. This helps us to develop later the idea of a lower envelope of a collection of coherent previsions which is given by the lower envelope theorem. This is given in the following proposition.

Proposition 4.24. *Suppose that Ω is a non-empty collection of lower previsions defined on $\mathcal{K} \subseteq \mathcal{G}$. Define the lower envelope of Ω as $\underline{Q} : \Omega \to \mathbb{R}$ such that $\underline{Q}(f) := \inf_{\underline{P} \in \Omega} \underline{P}(f)$ for all $f \in \mathcal{K}$. Then,*
(1) If some lower previsions in Ω avoid sure loss and $\underline{Q}$ is real-valued, then $\underline{Q}$ avoids sure loss, i.e., if a lower prevision is dominated by a lower prevision that avoids sure loss, then it avoids sure loss as well.
(2) If all lower previsions in Ω are coherent, then $\underline{Q}$ is coherent.

$\Diamond$ (1). This follows from Definition 4.13 (3). Suppose that $\underline{P} \in \Omega$ avoids sure loss. Then there exists a coherent lower prevision $\underline{R}$ on $\mathcal{G}$ such that $\underline{R} \geq \underline{P}$. Since $\underline{P} \geq \underline{Q}$ we have that $\underline{R} \geq \underline{Q}$. This implies that $\underline{Q}$ will also avoid sure loss.
(2). Since all $\underline{P}(f) \geq \inf f$, we have that $\underline{Q}(f) \geq \inf f$. Given that all lower previsions are coherent, we apply Definition 4.16 (4) for coherence to show that $\underline{Q}$ is also coherent. We consider arbitrary values of n, m, and a set consisting a finite number of gambles corresponding to the lower previsions in Ω. We then show that $\underline{Q}$ satisfies the above definition. $\Diamond$

4.7 Natural Extension as Least Committal Extension

From Definition 4.14 we see that if $\underline{P}$ avoids sure loss, then it is possible to obtain its natural extension $\underline{E}_P$ which is coherent. Also, we see that there are coherent lower previsions on $\mathcal{G}$ that dominate $\underline{P}$. In this section, we discuss one more property of natural extension. In particular, we discuss the natural extension as the least committal extension.

The natural extension is the least committal extension of $\underline{P}$ i.e., it is the most conservative extension in the sense that all coherent lower previsions that dominate $\underline{P}$ also dominate $\underline{E}_P$. We bring out this notion by the following theorem.

Theorem 4.25.

(1) If $\underline{P}$ avoids sure loss, then $\underline{E}_P$ is the point-wise smallest coherent lower prevision on $\mathcal{G}$ that dominates $\underline{P}$.

(2) If $\underline{P}$ is coherent, then $\underline{E}_{\underline{P}}$ is the point-wise smallest coherent lower prevision on $\mathcal{G}$ that coincide with $\underline{P}$ on $D(\underline{P})$, i.e., the point-wise smallest coherent extension of $\underline{P}$.

$\Diamond$ (1). Suppose that $\underline{P}$ avoids sure loss. Then by Definition 4.14 of natural extension, $\underline{E}_{\underline{P}}$ is a coherent lower prevision that dominates $\underline{P}$. Let $\underline{Q}$ on $\mathcal{G}$ be some lower prevision that dominates $\underline{P}$. Then by Proposition 4.12 and Definition 4.16 (3), $\underline{Q} = \underline{Q}(Cl_{\mathbb{D}}(\mathcal{A}_{\underline{Q}}))$ dominates $\underline{P} = \underline{P}(Cl_{\mathbb{D}}(\mathcal{A}_{\underline{P}}))$. But $\underline{E}_{\underline{P}} = \underline{P}(Cl_{\mathbb{D}}(\mathcal{A}_{\underline{P}}))$.

(2). By Definition 4.16 (3), $\underline{E}_{\underline{P}}$ is a coherent extension of $\underline{P}$. By (1), it is also the point-wise smallest one. $\Diamond$

The next proposition states that for lower previsions that are dominated, their extensions also are dominated. We need this result later to say that a linear prevision that dominates a lower prevision also dominates its natural extension, while discussing the lower envelope theorem. The dominating lower prevision would have a higher value assigned for the gamble.

Proposition 4.26. *Let $\underline{P}$ and $\underline{Q}$ be lower previsions that avoid sure loss. If $\underline{Q}(f) \geq \underline{P}(f)$ then $\underline{E}_{\underline{Q}}(f) \geq \underline{E}_{\underline{P}}(f)$ for all $f \in \mathcal{G}$.*

$\Diamond$ Given that $\underline{Q}(f) \geq \underline{P}(f)$. We first show that $\mathcal{A}_{\underline{Q}} \supseteq \mathcal{A}_{\underline{P}}$. Suppose that $g \in \mathcal{A}_{\underline{P}}$. Then by the definition of $\mathcal{A}_{\underline{P}}$ given in Equation 4.4, there exist f and μ such that $f \in D(\underline{P})$ and $\mu < \underline{P}(f)$, where $g = f - \mu$. This implies that $f \in D(\underline{Q})$ and $\mu < \underline{Q}(f)$, and so $g \in \mathcal{A}_{\underline{Q}}$ as well. Then from Proposition 4.12, we have the result that $\underline{E}_{\underline{Q}}(f) \geq \underline{E}_{\underline{P}}(f)$ for all $f \in \mathcal{G}$. $\Diamond$

4.8 Lower Prevision in Terms of a Set of Linear Previsions

There is a relationship between coherent lower previsions and sets of linear previsions (Williams [82] and Walley [78]). It is also known as duality. Using this relationship it is possible to state natural extension in terms of dominating linear previsions of a lower prevision. We discuss this connection in this section.

A duality map from lower previsions to sets of linear previsions is defined as follows.

Definition 4.27. *Let $\mathcal{P}$ be the set of linear previsions on $\mathcal{G}$. For a given lower prevision $\underline{P}$ we define the set of all linear previsions on $\mathcal{G}$ that dominate $\underline{P}$ as:*
$\mathcal{L}(\underline{P}) \coloneqq \{Q \in \mathcal{P} :$ for all $f \in D(\underline{P})$ such that $Q(f) \geq \underline{P}(f)\}$. This is the dual model of $\underline{P}$.

Similarly, a duality map from sets of linear previsions to lower previsions is defined as follows.

Definition 4.28. *Let $\mathcal{M}$ be any set of linear previsions on $\mathcal{G}$. A lower prevision on $\mathcal{M}$ is defined as: $\underline{P}(\mathcal{M}(f)) \coloneqq \inf\{Q(f) : Q \in \mathcal{M}\}$ for $f \in \mathcal{G}$. This is the dual model of $\mathcal{M}$. $\mathcal{M}$ is in fact the credal set associated with $\underline{P}$.*

A notable property is that the natural extension of a lower prevision also has the same dual model. This follows from our earlier result that the extension of a dominating lower prevision also is dominating, and is given below.

Proposition 4.29. *Let $\underline{P}$ avoid sure loss, and let $\underline{E}_{\underline{P}}$ be its natural extension. Then for any $P \in \mathcal{P}$, $\underline{P} \leq P$ if and only if $\underline{E}_{\underline{P}} \leq P$, and consequently $\mathcal{L}(\underline{P}) = \mathcal{L}(\underline{E}_{\underline{P}})$.*

$\Diamond$ From Proposition 4.26 it follows that the set of dominating linear previsions are the same for $\underline{P}$ and $\underline{E}_{\underline{P}}$. $\Diamond$

The following lemma states that a linear functional on the space of gambles that dominates a coherent lower prevision is a linear prevision. The lemma helps us establish the connection between linear functional and lower prevision. At this point we may also want to remark that a linear prevision can be defined as a linear real functional on $\mathcal{G}$.

Lemma 4.30. *Let φ be a linear functional on $\mathcal{G}$ that dominates a coherent lower prevision $\underline{P}$ on $\mathcal{G}$. Then φ is a linear prevision.*

$\Diamond$ Suppose that $\varphi \geq \underline{P}$ on $\mathcal{G}$. Given that the lower prevision $\underline{P}$ is coherent, by Theorem 4.22 (1) we have, $\varphi \geq \underline{P} \geq \inf f$. $D(\varphi)$ is a linear subspace of $\mathcal{G}$ and so by Theorem 4.23 (1), P is a coherent (linear) prevision. $\Diamond$

The next theorem is the lower envelope theorem. It allows us to express avoiding sure loss, coherence and natural extension of a lower prevision $\underline{P}$ in terms of its dual model $\mathcal{L}(\underline{P})$. The Hahn-Banach extension theorem is used here to establish the existence of a linear functional on $\mathcal{G}$ that dominates the natural extension of a lower prevision.

Theorem 4.31. *Let $\underline{P}$ be a lower prevision. Then the following statements are true.*
(1) $\underline{P}$ avoids sure loss if and only if $\mathcal{L}(\underline{P})$ is non-empty.
(2) $\underline{P}$ is coherent if and only if it avoids sure loss and
$\underline{P}(f) = min\{Q(f) : Q \in \mathcal{L}(\underline{P})\}$ for all $f \in D(\underline{P})$.
(3) If $\underline{P}$ avoids sure loss, then its natural extension is the lower envelope of the non-empty set $\mathcal{L}(\underline{P})$, i.e., $\underline{E}_{\underline{P}}(f) = min\{Q(f) : Q \in \mathcal{L}(\underline{P})\}$ for all $f \in \mathcal{G}$.

$\Diamond$ (3). Suppose that $\underline{P}$ avoids sure loss. For a given $f \in \mathcal{G}$, define $\rho(g) := \underline{E}_{\underline{P}}(g + f) - \underline{E}_{\underline{P}}(f)$, $g \in \mathcal{G}$. The functional ρ can be shown to be concave since $\underline{E}_{\underline{P}}$ is concave. Consider the linear space $\mathcal{G}$ and the subspace $\{0\}$. By the Hahn-Banach Theorem 3.6, there exists an extension to a real linear functional φ on $\mathcal{G}$ that dominates ρ. Then, by using the fact that $\underline{E}_{\underline{P}}$ is super-additive, we get, $\varphi(g) \geq \rho(g) = \underline{E}_{\underline{P}}(g + f) - \underline{E}_{\underline{P}}(f) \geq \underline{E}_{\underline{P}}(g)$. We see that $\varphi(g) \geq \underline{E}_{\underline{P}}(g)$, so by Proposition 4.30, φ is a linear prevision, and by Definition 4.27, φ is in the dual model $\mathcal{L}(\underline{E}_{\underline{P}})$. By Proposition 4.29, $\mathcal{L}(\underline{E}_{\underline{P}}) = \mathcal{L}(\underline{P})$. It follows from Definition 4.27 that $\underline{E}_{\underline{P}}(g) \leq \inf\{Q(g) : Q \in \mathcal{L}(P)\}$.

By Theorem 4.20 (2), $\underline{E}_{\underline{P}}(0) = 0$, and given that φ is self-conjugate, we get $-\varphi(f) = \varphi(-f) \geq \rho(-f) = \underline{E}_{\underline{P}}(-f + f) - \underline{E}_{\underline{P}}(f) = -\underline{E}_{\underline{P}}(f)$, i.e., $\underline{E}_{\underline{P}}(f) \geq \varphi(f)$.

Combining the two above, we get $\underline{E}_{\underline{P}}(f) = \varphi(f)$.

(2). From (3) above, we see that $\underline{P}$ avoids sure loss and is a restriction of its natural extension. Therefore from Definition 4.16, $\underline{P}$ is coherent.

(1). From (3), we know that φ exists, therefore $\mathcal{L}(\underline{E}_P)$ is non-empty. But $\mathcal{L}(\underline{E}_P) = \mathcal{L}(\underline{P})$. So $\mathcal{L}(\underline{P})$ is non-empty. Conversely, suppose that $\mathcal{L}(\underline{P})$ is non-empty. Then by Definition 4.27, there exists a linear prevision P that dominates $\underline{P}$. By Theorem 4.18 (2), P avoids sure loss (as a lower prevision). Therefore by Proposition 4.24 (1), $\underline{P}$ avoids sure loss. $\Diamond$

A key result that we are establishing here is as follows. First, a non-empty credal set associated with a lower prevision implies a lower prevision that avoids sure loss, and vice-versa. Second, the minimum of a credal set associated with a lower prevision is a coherent lower prevision given that the lower prevision avoids sure loss, and vice-versa. Essentially, we are establishing a one-to-one correspondence (an equivalence) between coherent lower previsions and sets of linear previsions. In fact, the linear prevision for each gamble can be written as a linear inequality that is bounded from below. The lower bound is given by the coherent lower prevision which taken together for all the gambles considered becomes the lower envelope for the linear previsions.

4.9 Summary and Comments

In this chapter we see how the theory of lower previsions is developed, as in the works of Walley [78] and Troffaes & de Cooman [67]. We would now use the theory of lower previsions to apply to a practical situation, i.e., to a question that we would ask concerning financial risk. It is a computational question. How do we calculate financial risk measures from data? How do we see risk measures as imprecise probabilities and then determine whether they satisfy the criteria of rationality, viz., consistency and coherence when seen as lower previsions. We discuss these aspects from Chapter 5 onward and then conclude our study.

From the discussions so far, we understand that beliefs can be expressed by alternative models of imprecise probability, and that they are subjective probability models. The alternative models we discussed are a coherent set of desirable gambles, coherent lower previsions, and a set of linear previsions. We also now know that lower and upper previsions, as super- and sub-linear previsions, can be non-additive probabilities. Lower and upper previsions are interpreted as the supremum buying and infimum selling prices of gambles and perhaps they model best our beliefs.

An individual would not use a set of desirable gambles, but would use lower and upper previsions he has assessed for the gambles to make his decisions for them, against them, or take no action on them. Actually coherent lower prevision is quite very general as a model of uncertainty in imprecise probability, therefore sometimes additional restrictions are placed in order to use the models for statistical inference. When restrictions are placed on the general model we obtain special cases of imprecise probability models. The special cases are discussed further in the works that we have cited earlier, but are not considered here.

Therefore from the next chapter onward, our study moves from theory of coherent lower previsions to its applications. Here risk measures in finance are seen as lower previsions, and we ask whether the risk measures in finance are consistent and coherent. That lead us first to the task of obtaining risk measures from actual data, interpreting them as upper previsions, and then see them as lower previsions using the notion of conjugacy. We would then computationally determine whether the risk measures so obtained are coherent from the imprecise probability point of view, and if not, then find natural extensions that are coherent. We now move to this application.

Chapter 5

Determining Consistency and Coherence of Lower Prevision from Data

In this section, we discuss how we may computationally decide whether the assessments of risk (which are interpreted as lower previsions) are mutually consistent and coherent. The central idea on how to determine this computationally comes to us from the article by Walley, Pelessoni and Vicig [76]. Using the method of linear programming, the article presents algorithms for checking consistency and coherence of given assessments. The main ideas from this article needed for our study are summarized and discussed below. We also present here examples to illustrate and clarify the method.

We first define a linear program. A linear program is written in the following form.

Definition 5.1.

Maximize $c^T x$

Subject to,

(1) $Ax \leq b$, *and (2)* $x \geq 0$, *where* $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$.

The primal variables are x *which is an* $(n \times 1)$ *vector of real numbers.*

The dual form of this linear program is written as:

Minimize $b^T y$

Subject to,

(1) $A^T y \geq c$, *and (2)* $y \geq 0$.

The dual variables are y *which is an* $(m \times 1)$ *vector of real numbers.*

We may use either the primal or the dual form to solve our optimization problem. The following results are from Walley et al. [76].

Consistency of a set of assessments is characterized by the notion of avoiding uniform loss. A set of assessments avoid uniform loss when the system of linear inequalities $\Sigma_{i=1}^{k} \lambda_i (G_i + \epsilon B_i) \leq 0$, and $\lambda > 0$ has no solution (λ, ϵ), where $\epsilon > 0$, λ_i is the stake on gamble i, $G_i = B_i(A_i - c_i)$ is the gain from gamble i (for all gambles $i = 1, ..., k$), and $\underline{P}(A_i|B_i) = c_i$ is the assessment (probability) of event A on gamble i, given that the conditioning event B occurs (equal to one) for gamble i. This means that to avoid uniform loss, the net gain from a linear combination of gambles cannot be negative or zero.

They give the following lemma which states that avoiding uniform loss is equivalent to the condition that the supremum of the net gain from a linear combination of gambles must at least be non-negative.

Lemma (Walley et al. [76]).

The assessments avoid uniform loss if and only if $\sup(\Sigma_{i=1}^{k}\lambda_i G_i|S(\lambda)) \geq 0$, *whenever* $\lambda > 0$, *and* $S(\lambda) = \cup_{i\in I(\lambda)} B_i$ *is the union of the conditioning events* B_i *for which* λ_i *is non-zero, i.e.,* $I(\lambda) = \{i : \lambda_i \neq 0, i = 1, ..., k\}$.

This means that incurring uniform loss for the assessments is equivalent to having a non-empty set $I \subseteq 1, ..., k$ such that $\Sigma_{i\in I}\lambda_i G_i + \Sigma_{i\in I} B_i \leq 0$, for some $\lambda \geq 0$.
Walley et al. [76] then give the algorithm for checking consistency (which is Algorithm 2 of their paper) and is as follows.

Algorithm (Walley et al. [76]).

(a) Set $I = 1, ..., k$.

(b) Maximize $\Sigma_{i\in I}\tau_i$

Subject to,

(1) $\lambda_i \geq 0$, *(2)* $0 \leq \tau_i \leq 1$ *for all* $i \in I$, *and (3)* $\Sigma_{i\in I}\lambda_i G_i + \Sigma_{i\in I}\tau_i B_i \leq 0$

(c) If $\tau_i = 1$ *for all* $i \in I$ *then the assessments incur uniform loss. Otherwise replace* I *by the subset* $i \in I : \tau_i = 1$. *If* I *is empty then the assessments avoid uniform loss. Otherwise, return to (b).*

In this algorithm, we are trying to find the largest set such that the inequality is satisfied. If this set is non-empty, then the assessments do not avoid uniform loss. On the other hand, if the set is empty then there are no assessments that incur uniform loss, consequently the assessments are mutually consistent.

In our study, all events in the possibility space are considered to be likely to happen. Also, each event occurring is not conditional on some event occurring. Therefore we use unconditional imprecise probability, and omit the conditional probability statement. Therefore, to check whether assessments avoid uniform loss, we solve the following maximization problem.

Method for Checking Consistency (Walley et al. [76]).

Maximize $\Sigma_{i\in I}\tau_i$

Subject to,

(1) $\lambda_i \geq 0$, *(2)* $0 \leq \tau_i \leq 1$ *for all* $i \in I$, *and (3)* $\Sigma_{i\in I}\lambda_i(f_i - \underline{P}(f_i)) + \Sigma_{i\in I}\tau_i \leq 0$

If $\tau_i = 1$ *then the assessments incur sure loss. If* $\tau_i = 0$ *then the assessments avoid uniform loss.*

Thus by the above method given by the algorithm, when we find that the assessments avoid sure loss we go to the next step and check for coherence.

Determining coherence of assessments is seen as the inference problem. Given a set of assessments for gambles, the inference actually may be for an existing gamble in the set, or for a new gamble. In fact, the solution to the inference problem gives us the natural extension for the assessment of an existing gamble which will now be coherent if not already coherent. Here we look for inference for the existing gamble in our set of gambles.

For the inference problem, Algorithm 4 of Walley at al. [76], is as follows.

Algorithm (Walley et al. [76]).

(a) Set $I = 1, 2, ..., k$.

(b) Maximize $\Sigma_{i \in I} \tau_i$

Subject to,

(1) $\lambda_i \geq 0$, *(2)* $0 \leq \tau_i \leq 1$ *for all* $i \in I$, *and (3)* $sup(\Sigma_{i \in I} \lambda_i G_i + \Sigma_{i \in I} \tau_i B_i | B^c) \leq 0$

(c) If $\tau_i = 1$ *for all* $i \in I$ *then go to (d). Otherwise replace* I *by the subset* $i \in I : \tau_i = 1$. *If* I *is non-empty then return to (b).*

(d) Solve the linear program LP (see below). The maximized value of μ *is the exact value of its natural extension.*

The first part of the two part algorithm essentially checks for consistency of assessments, which is similar to Algorithm 2 given earlier. Once consistency is checked, then we move to the second part where the natural extension is calculated from the linear program. As one can see from the maximization, the natural extension will dominate the assessment. This calculation of the natural extension is repeated for each gamble in the set until all corrected assessments are now coherent. The linear program to be solved is as follows.

Linear Program (Walley et al. [76]).

Maximize μ

Subject to,

(1) $\lambda_i \geq 0$ *for all* $i \in I$, *and (2)* $\Sigma_{i \in I} \lambda_i G_i + \mu B \leq AB$

If an assessment avoids sure loss but is not coherent, then the above linear program is run and the maximized value is the coherent lower prevision. This linear program is similarly repeated for each assessment to calculate its natural extension. As earlier, we omit conditional probability and write the maximization problem as follows.

Method for Finding Natural Extension (Walley et al., [76]).

Maximize μ

Subject to,

(1) $\lambda_i \geq 0$ *for all* $i \in I$, *and (2)* $\Sigma_{i \in I} \lambda_i (f_i - \underline{P}(f_i)) \leq (f_0 - \mu)$

We now discuss how the calculations are carried out. To determine whether the assessments are consistent and coherent by the above methods, we use the Python library *improb* written by Troffaes [66]. Library *improb* contains an implementation of Algorithm 2 and 4, and the above LP problem (as Lemma 5) by Walley et al. [76] as mentioned above. To do the above calculations in *improb*, Troffaes uses *pycddlib*

(Troffaes [68]) which is a Python wrapper he has written for the library *cddlib*. Library *cddlib* was developed by Fukuda [17] and it contains a class that performs polyhedral computations for solving linear programming problems, it is an implementation of the double description method in C. A double description pair is defined as follows (Fukuda et al. [18]). A pair $(\tilde{A}, R)$ of real matrices is said to be a double description pair if it is true that $\tilde{A}x \geq 0$ if and only if $x = R\,\lambda$ for some $\lambda \geq 0$. This means that $\{\tilde{A}x \geq 0\}$ and $\{x = R\lambda; \lambda \geq 0\}$ are simultaneous representations.

A polyhedron P can be given by the H-representation (see Quaeghebeur [49]). H-representation is a set of linear constraints, which are half-spaces $P = \{Ax \leq b | x \in \mathbb{R}^n\}$, with short description $\tilde{A} = [b, -A]$, where $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $\tilde{A} \in \mathbb{R}^{m\times(n+1)}$, and corresponding to the first one of the double description pair. Maximization or minimization is done over this polyhedron. We shall see the short description for H-representation in the examples shown below

We shall now discuss some examples. They illustrate the method for finding whether assessments avoid sure loss and if so, to correct for coherence when assessments are not coherent. In the examples given below gambles are defined on a possibility space with three elements. So the possibility space is $X = \{a, b, c\}$. Two gambles f_1 and f_2 are specified on this possibility space, along-with their corresponding rewards. The lower previsions assessed for each gamble, i.e., $\underline{P}(f_1)$ and $\underline{P}(f_2)$ are also given.

The first example illustrates the case where the assessments avoid sure loss (ASL).

Example 1: Checking whether Assessments Avoid Sure Loss

Gamble	a	b	c	$\underline{P}(f_i)$
f_1	1	3	5	4.0
f_2	2	4	6	5.0

Notice that by inspection, in the example above, the assessments are less than or equal to the supremum of the gambles. So, we know that they avoid sure loss and therefore assessments are consistent. We check this by formulating the consistency problem as a linear programming problem and then solve it, as shown below.

To solve the consistency problem that we have here, we first refer back to the LP formulation we have given in “Method for Checking Consistency” (see above). A brief explanation of the objective and the constraints needed here is as follows. The objective function τ that is to be maximized, when it attains a value of 1, means that we are incurring sure loss. When the objective attains a value zero, then this means that the assessments avoid sure loss. The constraints written below are as follows. The first two constraints give us $0 \leq \tau \leq 1$. The third and fourth constraints give us $\lambda_i \geq 0$ for $i = 1, 2$ corresponding to the two gambles. The remaining three constraints give us $\Sigma_{i\in I}\lambda_i(f_i - \underline{P}(f_i)) + \tau \leq 0$, where each constraint corresponds to each element in the possibility space a, b, c in that order, and where the gain for each gamble is the difference between the reward and the assessed lower prevision for that gamble, but now written as a negative since we are using the H-representation. Thus the consistency problem is written as follows.

Maximize τ subject to the following constraints (written in $\tilde{A} = [b, -A]$, the H-representation):

$$
\begin{aligned}
0 + 1.\tau + 0.\lambda_1 + 0.\lambda_2 &\geq 0 \\
1 - 1.\tau + 0.\lambda_1 + 0.\lambda_2 &\geq 0 \\
0 + 0.\tau + 1.\lambda_1 + 0.\lambda_2 &\geq 0 \\
0 + 0.\tau + 0.\lambda_1 + 1.\lambda_2 &\geq 0 \\
0 - 1.\tau + (4-1).\lambda_1 + (5-2).\lambda_2 &\geq 0 \\
0 - 1.\tau + (4-3).\lambda_1 + (5-4).\lambda_2 &\geq 0 \\
0 - 1.\tau + (4-5).\lambda_1 + (5-6).\lambda_2 &\geq 0
\end{aligned}
$$

Solution is $(\tau, \lambda_1, \lambda_2) = (0, 0, 0)$. This implies that the assessments avoid sure loss.

Given below is the H-representation of the formulation and the solution from *improb* output.

```
Example 1: Assessments ASL
(H-representation and Solution)

 a   b   c
1.0 3.0 5.0 | a b c : [4.0,    ]
2.0 4.0 6.0 | a b c : [5.0,    ]
H-representation
begin
 7 4 real
  0  1  0  0
  1 -1  0  0
  0  0  1  0
  0  0  0  1
  0 -1  3  3
  0 -1  1  1
  0 -1 -1 -1
end
maximize
  0  1  0  0
(0.0, 0.0, 0.0)
True
```

The second example illustrates the case when assessments do not avoid sure loss.

Example 2: Checking whether Assessments do not ASL

Gamble	a	b	c	$\underline{P}(f_i)$
f_1	1	3	5	4.0
f_2	2	4	6	8.0

Notice that in this example, the assessment on the second gamble is greater than its supremum value, which is making the net gain negative. There is at least a sure loss of 1 unit, so the assessments are not avoiding sure loss.

Maximize τ subject to the following constraints:

$$\begin{aligned}
0 + 1.\tau + 0.\lambda_1 + 0.\lambda_2 &\geq 0 \\
1 - 1.\tau + 0.\lambda_1 + 0.\lambda_2 &\geq 0 \\
0 + 0.\tau + 1.\lambda_1 + 0.\lambda_2 &\geq 0 \\
0 + 0.\tau + 0.\lambda_1 + 1.\lambda_2 &\geq 0 \\
0 - 1.\tau + (4-1).\lambda_1 + (8-2).\lambda_2 &\geq 0 \\
0 - 1.\tau + (4-3).\lambda_1 + (8-4).\lambda_2 &\geq 0 \\
0 - 1.\tau + (4-5).\lambda_1 + (8-6).\lambda_2 &\geq 0
\end{aligned}$$

Solution is $(\tau, \lambda_1, \lambda_2) = (1, 0, 0.5)$. Since $\tau \neq 0$, this implies that the assessments do not avoid sure loss.

The H-representation of the formulation and solution from *improb* output is given below.

```
Example 2: Assessments do not ASL
(H–representation and Solution)

 a   b   c
1.0 3.0 5.0 | a b c : [4.0,    ]
2.0 4.0 6.0 | a b c : [8.0,    ]
H–representation
begin
 7 4 real
  0  1  0  0
  1 −1  0  0
  0  0  1  0
  0  0  0  1
  0 −1  3  6
  0 −1  1  4
  0 −1 −1  2
end
maximize
  0  1  0  0
(1.0, 0.0, 0.5)
False
```

In the first example we saw that the assessments are consistent. There, given two gambles f_1 and f_2 along with their rewards, where the rewards are known to the individual who is placing the bets (but actual outcomes are unknown to him), he has made an assessment of lower prevision for each gamble. So the first example illustrates the case where we first check whether assessments he had made are consistent. The third example given below illustrates the case where those assessments are now checked for coherence, i.e., once the assessments are found to be consistent, we proceed to check whether his assessments are also coherent. The purpose of the third example here is to illustrate this next step. We note that consistency and coherence are different conditions placed on assessments therefore the checks are not the same for both.

Example 3: Assessments are Coherent

Gamble	a	b	c	$\underline{P}(f_i)$
f_1	1	3	5	4.0
f_2	2	4	6	5.0

The first part given below shows the primal formulation for checking coherence and its solution. The second part that follows illustrates the case where the problem has been formulated in terms of its dual, where the dual formulation is used for checking coherence. The dual formulation is shown in H-representations form, along with the solution.

Notice that by inspection of the above example, the assessments are first of all consistent. Also there is a certain symmetry in the values of the assessments, in the sense that they are close to the supremum of the gambles, and also mid-way between outcomes b and c. To verify that this assessment is coherent we solving the following problem.

Here we refer to the LP formulation given in "Method for Finding Natural Extension" (see above). A brief explanation of the objective and the constraints written below is as follows. The objective function μ that is to be maximized is the lower prevision for gamble f_0, given the assessment for the other gamble, and the rewards for all gambles. When the objective attains its maximum (supremum) value then this is the natural extension which is coherent. The first three constraints are written where the value μ for the second gamble is found such that the gain dominates that of the first gamble. The three constraints refer to a, b, c respectively. The fourth constraint ensures that the stake on the gamble is non-negative.

Maximize μ subject to the following constraints:

$$\begin{aligned} \lambda(1-4) &\leq 2-\mu \\ \lambda(3-4) &\leq 4-\mu \\ \lambda(5-4) &\leq 6-\mu \\ \lambda &\geq 0 \end{aligned}$$

Solution is $\lambda = 1$, and objective $\mu = 5$. This is the same as the assessment. So the assessment $\underline{P}(f_2) = 5$ is coherent. We next solve using the dual formulation and also show the H-representation of the formulation.

In this formulation, the dual variables are x_1, x_2 and x_3 corresponding to a, b, c in the possibility space. The objective that is to be minimized is the same as that maximized in the primal (μ) but now written in terms of the dual variables. The first two constraints written below give us $\sum_{i=1}^{3} x_i = 1$ written as two inequalities. The next three constraints give us $x_i \geq 0$ for $i = 1, 2, 3$. The last two constraints correspond to each gamble and the net gain.

Minimize $\mu = 0 + 2x_1 + 4x_2 + 6x_3$ subject to the following constraints:

$$\begin{aligned}
-1 + 1.x_1 + 1.x_2 + 1.x_3 &\geq 0 \\
1 - 1.x_1 - 1.x_2 - 1.x_3 &\geq 0 \\
0 + 1.x_1 + 0.x_2 + 0.x_3 &\geq 0 \\
0 + 0.x_1 + 1.x_2 + 0.x_3 &\geq 0 \\
0 + 0.x_1 + 0.x_2 + 1.x_3 &\geq 0 \\
0 + (1-4)x_1 + (3-4)x_2 + (5-4)x_3 &\geq 0 \\
0 + (2-5)x_1 + (4-5)x_2 + (6-5)x_3 &\geq 0
\end{aligned}$$

Solution is, $(x_1, x_2, x_3) = (0, 0.5, 0.5)$ which from the objective function above gives us the optimum value as $\mu = 5$. This is the same as our initial assessment for this gamble. Therefore the assessment $\underline{P}(f_2) = 5$ is coherent. Here is the H-representation and solution from *improb* output.

```
H-representation
linearity 1  1
begin
 6 4 real
  1 -1 -1 -1
  0  1  0  0
  0  0  1  0
  0  0  0  1
  0 -3 -1  1
  0 -3 -1  1
end
minimize
  0  2  4  6
* Algorithm: dual simplex algorithm
* minimization is chosen
* Objective function is
  0 +  2 X[  1] +  4 X[  2] +  6 X[  3]
* LP status: a dual pair (x,y) of optimal solutions found.
begin
  primal_solution
    1 :   0
    2 :   5.000000000E-01
    3 :   5.000000000E-01
  dual_solution
    2 :   0
    7 :  -5
    5 :  -1
  optimal_value :   5
end
True
```

The examples discussed here give us an idea about the calculation needed to assess consistency and coherence given data . In the actual problem considered in the next part of this study, six gambles are used in each assessment set, the reward from the gamble is its rate of return calculated from closing prices of stocks traded, the possibility space for each gamble has seven elements, and the risk measure computed from data is taken as the assessed lower prevision of the gamble.

Following this, in the next chapter we define a risk measure and see how it is interpreted as a lower prevision. Then methods of measuring risk from data are discussed, and having found risk measures from data, we computationally determine as explained in this chapter whether assessments are coherent.

Chapter 6

Application of the Theory of Lower Previsions to Risk Measures

6.1 Financial Risk and Lower Previsions

Models of economic and financial behaviour are in some sense less much less "precise" than "exact" models that describe physical laws of motion or laws of thermodynamics. Besides this, any decision-making model that incorporates a measure of uncertainty in terms of a precise probability can be generalized to a model of imprecise probability. In this section we discuss financial risk where risk measures are seen as upper previsions. One is able to reason and develop a more elaborate set of procedures with previsions for assessing and measuring risk than what classical probability methods can offer.

Within imprecise probability, risk measurement in finance has been well and fairly extensively discussed by Pelessoni and Vicig ([45], [46]), and Vicig ([70] and Chapter 12 of [4]). This section given below draws from these works.

Let $\mathcal{G}$ be the set of gambles as defined earlier. For any gamble $f \in \mathcal{G}$ a risk measure of f is a real number $\rho(f)$. Risk here is about the future prospect of achieving a negative outcome (reward) from the gamble. If the reward from the gamble f is expressed in some monetary unit, then the risk measure $\rho(f)$ also is expressed in the same monetary unit. It would be the amount that can be held as reserve to hedge in case gamble f makes a loss. It is assumed that the return from each gamble $f \in \mathcal{G}$ has a probability distribution given by P_f. Later in our study, we are going to say that the return from the gamble is its reward, and therefore this distribution becomes the probability distribution of returns from the gamble f. We can calculate this from data. The quantified value of the risk measure $\rho(f)$ would depend on this probability distribution, as we now see below.

One monetary quantification of risk is the value-at-risk denoted by VaR, and is specified for any given level α in P_f. It is defined as follows. Quoting from Artzner et al. [3], Vicig defines VaR as follows (Ch 12 of [4]). We omit the subscript f.

Definition 6.1. *Let $\alpha \in (0, 1)$. If $P(f < q) \leq \alpha \leq P(f \leq q)$ then q is an α-quantile for f. Then* $q_\alpha^+ \coloneqq \inf\{x\colon P(f \leq x) > \alpha\}$, *and* $VaR_\alpha(f) \coloneqq -q_\alpha^+(f)$.

Thus VaR is a quantile-based measure of risk. It is the negative of the lower threshold, which for the given level α, will not be exceeded. The first part of the definition gives the greatest lower bound for the quantile value such that the probability of obtaining a value less than or equal to this quantile value is at least α. If we are given the probability distribution of profits that, say, an organization makes, then values in the lower tail of this distribution would indicate the organization's losses, and therefore α-quantile values will have negative signs. Since VaR is a loss measure, we remove the negative sign and indicate the loss as a positive value. This is achieved by the second part of the definition.

In the finance literature, some authors refer to α as confidence (for example, Tasche [65]). In our study we shall in general refer to this as "level α". However, we shall take confidence level to be $1 - \alpha$. Using the definition above, suppose we find that for given level $\alpha = 0.05$, there is a prospect of making a loss of at least 10,000 for a gamble g. Then the value-at-risk is $VaR_{0.05}(g) = 10,000$. Here, we are saying that given the observed past pattern of losses and gains from gamble g, there is a 5% chance of making a loss of at least $10,000$. Therefore for the given 5% we would create a hedge for an amount of at least $10,000$ for g. The measure VaR is also for a specified time duration, losses say in a day, month or year.

It is desirable if risk can be reduced. Therefore when we make investment decisions or allocation decisions to cover risk, we would like the criteria that we use to make these decisions to have certain desirable properties so as to minimize unfavourable outcomes.

Artzner, Delbaen, Eber, and Heath [3] state certain desirable properties for a risk measure to possess. These properties are specified by a set of axioms on risk measures, and a risk measure is considered coherent if it satisfies these properties. In order to distinguish coherence here from the notion of coherence that is developed earlier for imprecise probability, Vicig (Chapter 12 of [4]) calls this ADEH-coherence (named after the authors above). The axioms are as follows.

Axiom 6.2. *Let $\mathcal{G}$ be a linear space of gambles which contains real constants. A mapping $\rho\colon \mathcal{G} \to \mathbb{R}$ is an ADEH-coherent risk measure if and only if it satisfies the following axioms.*

(1) Translation invariance

$$\rho(f + \delta) = \rho(f) - \delta \text{ for all } f \in \mathcal{G},\ \delta \in \mathbb{R}$$

(2) Positive homogeneity

$$\rho(\lambda f) = \lambda \rho(f) \text{ for all } f \in \mathcal{G},\ \lambda \geq 0$$

(3) Monotonicity

$$f \leq g \text{ implies that } \rho(f) \leq \rho(g) \text{ for all } f, g \in \mathcal{G}$$

(4) Subadditivity

$$\rho(f + g) \leq \rho(f) + \rho(g) \text{ for all } f, g \in \mathcal{G}$$

Suppose that a gamble reward is expressed in monetary units, say currency units. Then translation invariance means that addition of cash (a risk-less capital when held) reduces the risk by the same amount.

If $\rho(f) - \delta = 0$ then δ is the amount of cash that is added to the gamble f in order to reduce its risk to zero units. This would be the risk neutral position. Positive homogeneity suggests that larger a position (a stake) in an asset, the risk is proportionately larger. So when we double our position, the risk is two times larger. This however may not always be true. Monotonicity implies that gambles with a better valuation (in terms of better future returns) would entail lesser risk. Sub-additivity is the principle behind diversification. Dividing the wealth into different asset classes is likely to lower risk when compared to investing in only one asset class. This means that a merger does not create extra risk. Here, the gamble $f + g$, which is a sum or a combination of gambles f and g, can be thought of as a merger. Essentially, the axiom states that the risk of a sum of gambles cannot exceed the sum of the risks of individual gambles. Thus the property of sub-additivity helps reduce risk through diversification, where wealth is staked in several gambles rather than in just one gamble.

Artzner et al. [3] state that risk measure VaR may not have the sub-additivity property. If this is true then use of this risk measure may induce a firm to take a wrong decision regarding diversification. This motivated the search for a family of risk measures with the above desired properties.

The risk measure $\rho(f)$ has an interpretation as an imprecise prevision. The measure $\rho(f)$ can be seen as the smallest amount of money that is needed to hedge against the risk of holding (betting for) gamble f. Buying f is the same as selling $-f$. So $\rho(f)$ is seen as the infimum selling price of $-f$, which is actually the upper prevision $\overline{P}$ of $-f$. Then along with the conjugacy property $\overline{P}(-f) = -\underline{P}(f)$, we get

$$\rho(f) = \overline{P}(-f) = -\underline{P}(f).$$

Therefore, $\rho(f)$ is either an upper prevision for $-f$, or negative of a lower prevision for f.

A coherent risk measure on some set of gambles $\mathcal{K}$ is defined as follows (Vicig in Chapter 12 of [4]).

Definition 6.3.

1. A mapping $\rho\colon \mathcal{K} \to \mathbb{R}$ is a coherent risk measure on $\mathcal{K}$ if and only if there exists a coherent upper prevision $\overline{P}$ on $-\mathcal{K} = \{-f \colon f \in \mathcal{K}\}$ such that $\rho(f) = \overline{P}(-f)$, for all $f \in \mathcal{K}$.

This is restated as:

1'. The mapping $\rho\colon \mathcal{K} \to \mathbb{R}$ is a coherent risk measure on $\mathcal{K}$ if and only if for all $n \geq 0$, $f_0, f_1, \dots f_n \in \mathcal{K}$, and $\lambda_0, \lambda_1, \dots, \lambda_n \in \mathbb{R}^+$, it is true that $\sup \overline{G} \geq 0$, where $\overline{G} = \sum_{i=1}^{n} \lambda_i(\rho(f_i) + f_i) - \lambda_0(\rho(f_0) + f_0)$.

ADEH-coherent risk measures are a particular case of coherent risk measures. Vicig (in [4]) shows that ρ is ADEH-coherent on $\mathcal{K}$ if and only if it is a coherent risk measure.

We next look at another risk measure that is called conditional value-at-risk ($CVaR$), which is known to be ADEH-coherent (Rockafeller and Uryasev [52], Follmer and Schied [16]). It is not yet known whether $CVaR$ is coherent as an imprecise prevision.

Let Y be a random cost variable and let F be its (loss) distribution function. Then VaR is,

$$VaR_\alpha(Y) = F^{-1}(\alpha) = \inf\{u : F(u) > \alpha\}$$

$CVaR$ is defined as the mean of the α tail of the distribution F. It is the average loss for the $100\alpha\%$ of the worst case scenarios. So it is also called as mean excess loss, mean shortfall, or tail VaR.

Rockafellar and Uryasev [51] define $CVaR$ as a conditional expectation of Z, given that $Z \geq VaR_\alpha(Y)$.

$$CVaR_\alpha(Y) \coloneqq \mathbb{E}(Z|Z \geq VaR_\alpha(Y))$$

They also state that $CVaR$ is continuous with respect to α unlike some other risk measures. Therefore $CVaR$ risk measure can be evaluated also for random variables (gambles) with discontinuous distributions. In our study the returns distribution we obtain from historical data is assumed to be continuous. We use this data to calculate VaR and $CVaR$ which are our imprecise probability measures. Another property of VaR and $CVaR$ is that they are monotonic in α. A smaller α level implies greater risk.

Acerbi and Tasche [1] define $CVaR$ as expected shortfall ES,

$$ES_\alpha(Y) \coloneqq 1/\alpha \int_0^\alpha VaR_\beta(Y)d\beta$$

They show that ES is ADEH-coherent if it has the following properties. Similarly Phlug [48] shows that $CVaR$ is coherent if it has the following four properties.

Axiom 6.4.

(1) Monotonicity

$$f \in \mathcal{G}, \textit{ and } f \geq 0 \textit{ implies that } \rho(f) \leq 0$$

(2) Subadditivity

$$f, g, f+g \in \mathcal{G} \textit{ implies that } \rho(f+g) \leq \rho(f) + \rho(g)$$

(3) Positive homogeneity

$$f \in \mathcal{G}, \lambda > 0, \lambda f \in \mathcal{G} \textit{ implies that } \rho(\lambda f) = \lambda \rho(f)$$

(4) Translation invariance

$$f \in \mathcal{G}, \delta \in \mathbb{R} \textit{ implies that } \rho(f+\delta) = \rho(f) - \delta$$

The properties given here are identical to the properties for ADEH-coherence that we discussed above. We study coherence in imprecise probability and finance because coherence implies the absence of possibility of incurring sure loss. More importantly, coherence implies the absence of opportunity for arbitrage. This makes coherence a desirable property to possess.

Having defined risk and its properties, and interpreting it as an assessment of a lower prevision, we now turn to aspects concerning obtaining a measure of the assessment from actual data, and then examine ways to determine whether the assessments are consistent and coherent as an imprecise probability.

6.2 Data Used in the Study

Let us consider an equity-based mutual fund. Equity-based means that the mutual fund is composed of a collection of equities. An equity is a non fixed-income asset or a non fixed-income security. An example of a fixed-income security is a fixed-deposit held in a bank which gives fixed returns over its tenure. We can think of a fixed deposit as a constant gamble. An equity is a share or stock of a company that gives uncertain returns to the owner who holds it. An equity can be thought of as a gamble. It is a risky asset to hold. Therefore an equity-based mutual fund is a collection of gambles or risky assets. A collection of assets, we may also call it a portfolio of assets. We can think of a fund as a portfolio of individual assets. A well diversified equity-based mutual fund may contain more than 200 stocks of companies. A sectoral equity-based mutual fund may contain as few as 15 stocks of companies in that sector. An example of a sector is the banking sector, and a banking sector equity-based mutual fund would be composed of stocks of banks that the fund manager chooses to include in the fund the composition of which over time he may be able to change based on its performance.

We select some mutual funds to study coherence of risk measures. Although the riskiness of a fund would be of interest to the fund manager who could be a financial institution or the regulator of a financial institution, this study is not developing an overall measure of risk for assessing or managing fund riskiness and examine coherence. Rather, this study looks at the risk measure of a particular share in that fund, calculates the risk measure from its historical price data, and asks the question whether that risk measure when seen as a lower prevision is coherent.

The raw data that we use in the study is the price data of a stock (share) that is part of a fund. The price that we refer to here is the closing price of that stock at the end of a trading day. We call it the adjusted daily price basing it on dividends declared by the company on that date. This is a time-series data, and from this we can calculate daily return using Definition 6.5.

Three separate sets of data were used in the study. A description of the data sets used is as follows. FSPCX is an US fund, RBF is an India fund, and EDHEC is an International fund. RBF shows the greatest volatility and FSPCX the least, a characteristic possibly of their respective financial markets. We will see later, that the data sets have characteristics that are both common and different. Three different data sets were used with the objective that our analysis of data and results are made, as much as possible, independent of the characteristics of data or of financial markets.

The first data set is from the New York Stock Exchange (NYSE). This data set is that of the mutual fund Fidelity Select Insurance Portfolio (FSPCX) from the Financial Services Sector. The fund consists of sixty-one holdings, primarily investing in common stocks, that are traded in NYSE. Its risk category is below average, and return category is above average. Six stocks, comprising of about 58% of the portfolio, is selected for the study. The benchmark index for this sector is the NYSE Financial Services Index (NYKID), the index is made up of 398 components. Total number of data points is 504, which consists of time series data on adjusted daily prices for all trading days in USA from Jan 1, 2013 to Dec 31, 2014. The code names of the 8 assets are: ACE, AIG, AFL, ALL, CB, MET.

The second data set is from the Bombay Stock Exchange (BSE). This data set is that of the mutual fund Reliance Banking Fund (RBF) from the Banking Sector. The fund consists of twenty-eight holdings, investing in equity, that are traded in BSE. Its risk category is below average, and return category is above average. Six stocks, comprising of about 66% of the total portfolio, is selected for the study. The bench mark index for this sector is the BSE BANKEX, the index is made up of twelve components. Total number of data points is 494, which consists of time series data on adjusted daily prices for all trading days in India from Jan 1, 2013 to Dec 31, 2014. The code names of the six assets are: AXIS, HDFC, ICICI, IDFC, SBI, and YES.

The third data set is *edhec*, a data series used in the R-package *PerformanceAnalytics* that is maintained by Peterson [47]. The data source is the EDHEC Risk and Asset Management Research Center. The R-package gives the following description. EDHEC data series is a composite hedge fund style index returns. The dataset contains thirteen indices of global funds. Total number of data points is 152, which consists of monthly returns from Jan 31, 1997 to Aug 31, 2009. Six indices were used for the study. Further description about the data set can be found in the R-package reference manual.

6.3 Calculating Gamble Values and Assessments

From the adjusted daily prices of shares available in the first two data sets, the daily (financial) returns data is computed for each share. This is calculated by using the returns function of the *timeSeries* R-package. The third data set is already in the form of monthly financial returns that had been computed from indices and available to us as an R-package. Therefore there was no need to calculate returns data for the shares in this fund.

We first define return from a share (which is our proxy for reward from a gamble) as follows.

Definition 6.5. *Suppose that P_t is the price of the asset at end of period t. Then the simple net return on the asset during $t-1$ and t is given by $R_t = (P_t - P_{t-1})/P_{t-1}$. The simple gross return is $1 + R_t$.*

Using this definition, we calculate returns data of each share from each fund. This is a time series data using this we can plot a returns distribution for each share. These are given in Appendix C (Figures C.1–C.4). The risk measures VaR and $CVaR$ for each share are calculated from the above returns data using functions from the R-package *PerformanceAnalytics* (Peterson [47]). It is possible to calculate the risk measures using either from historical data or by assuming that the returns distribution is normal. This option is available in the R-package. The historical method is used in our analysis because we want to know the consequences of data being approximately normal or not. This is discussed later in this section.

From the returns data obtained from daily closing prices, we are calculating VaR and $CVaR$. The measures VaR and $CVaR$ are calculated for specified levels of α (in fact $\alpha = 0.01$ and 0.25). We used these levels because a level of 0.05 did not produce different results from that of 0.01, but we wanted to examine whether larger levels like 0.25 would give different results. This is discussed later.

Note that the $CVaR$ is actually the average loss for the $100\alpha\%$ of the worst case scenarios. Also note that risk measures VaR and $CVaR$ are calculated directly from the returns data using R-packages. So essentially we are using continuous data to find the risk measures, i.e., a continuous returns distribution.

It is clarified here that VaR and $CVaR$ are not calculated from discretized returns data. Therefore, the way the discretization of returns is done does not affect calculation of risk measures, the calculated risk measure would remain unchanged.

The purpose of discretizing the returns distribution is to obtain a finite number of rewards for each gamble to match the finite number of outcomes in the possibility space. So in our case, the reward for a gamble (stock) when a particular outcome occurs is the discretized value of the return for that stock. If we have seven elementary outcomes in the possibility space, then for each gamble there will be seven rewards or returns. We now look at how discrete values of returns are obtained.

The discrete values are obtained from our returns data set in the following manner. Suppose we want five discrete values of returns from the returns distribution. If the ordered data set (sorted in ascending order) contains 500 points, then the average of the first 100 points will be the first discrete value, the average of the next 100 points will be the second discrete value, and so on. So now we have five discrete return values from our original returns data.

To the above five values, we add the minimum value and the maximum value of the returns. This minimum and maximum will correspond to the infimum and the supremum of the gamble rewards. We need these two values to check for avoiding sure loss and for coherence. So essentially we now have seven discrete values. This will correspond to a possibility space that has seven elementary outcomes.

Above, if instead of 100 data points we take 50 data points each time and average them, then we would be able to get ten discrete return values. We can then add the infimum and the supremum to these to get twelve discrete returns which is a finer discretization. The possibility space here will have twelve elementary outcomes.

There is a question here as to whether we would want to discretize more or discretize less. Therefore a check on coherence was done for the EDHEC fund using a discretization of seven values for one case and twelve values for another case. It was found that in both cases the risk measures had to be corrected for coherence. Therefore it was decided to continue all experiments with seven discrete values. Results of the two methods can be compared in Appendix B (Tables B.3 and B.8)

Now we have the data processed for calculations. The data now give us for each share (gamble) its seven returns values (its seven rewards), and an assessment of its lower prevision (in terms of either VaR or $CVaR$). Therefore the input data for checking coherence are the rewards gamble-wise, and the assessed lower prevision for the gamble. Using the rewards one is able to check coherence of lower prevision for each gamble.

These are given in Appendix A (Tables A.1–A.4) fund-wise and for each of the six share in each fund . A data precision of upto 12 decimal places was used to avoid rounding off errors and to be able to detect small corrections while checking coherence.

The rewards and the risk measure data are input to python module *improb* case-wise (either for VaR or for $CVaR$). In Chapter 5 we have discussed the method for determining consistency and coherence of each risk measure. Appendix F (Tables F.1–F.3) gives the Python code and R-script used for performing the calculations. We discuss results from calculations in the next section.

6.4 Discussion of Results and Coherence

If the assessment of lower prevision is less than the infimum of the gamble, then the assessment avoids sure loss but is not coherent. On the other hand, if the assessment is greater than the supremum of the gamble then we are not avoiding sure loss. One of the properties of consistency and coherence is that the assessment in within the infimum and the supremum of the gamble. Therefore if the assessment made is at least the minimum and at most the maximum value of the gamble return, then the assessment would be consistent. This is automatically ensured by the method we have used to calculate VaR and $CVaR$ measures for each gamble. The method that we have used to construct the risk measure ensures that the α-quantile will be greater than the minimum of the returns data, unless $\alpha = 0$ in which case the risk measure (lower prevision) will be equal to the minimum return but is still within bounds. What we are stating here is that if the risk measure is within bounds as described above then they are consistent, and only coherence of the risk measure is to be ascertained. We now shall discuss results.

For all the three funds, calculations show that the assessments are not coherent. This means that for each gamble, the natural extension involves a correction, after which the corrected lower prevision is coherent. This is true for all the assets of all the funds, for both the $CVaR$ and VaR risk measures. This is the first result, i.e, that the assessments are not coherent. This was tried for two levels of α as mentioned earlier, viz., $\alpha = 0.01$ and $\alpha = 0.25$. What was observed was that the correction required is more for level 0.01 than for level 0.25. So incoherence could be more of an issue when we are more conservative in our assessment. The results are given in Appendix B (Tables B.1–B.3) for each fund.

This led to the conjecture that it is the left-tail characteristic of the returns distribution of the gamble that is causing more correction when lower and more conservative α levels are specified for the risk measure. Gambles that had long left tails possibly needed more correction.

The three funds FSPCX, EDHEC, and RBF were checked for volatility by calculating their standard deviations of daily return. The EDHEC data is in monthly returns, therefore this was converted to daily return by assuming that there are 21 trading days in one month. This is presented in Appendix G (Figure G.1). It is to be noted that the three funds are from different financial markets. Further, FSPCX and RBF data is for period 2013-2014, whereas EDHEC is for period 1997-2009. So we are looking at two different periods, and volatility may not be the same. Also, EDHEC is an index type of fund (which is a fund of funds), and therefore this fund would have lower volatility. However, a comparison of the standard deviations between the three shows that RBF fund has a standard deviation approximately three times that of FSPCX, and standard deviation of fund FSPCX is approximately twice that of EDHEC. What was observed was that fund RBF which had the relatively larger standard deviation needed more correction.

This led to the second experiment and the second result. Therefore at this point we go back, and test all gambles used in the study for normality. The results are shown in the appendices. Appendix C (Figures C.1–C.3) gives the density plots against normal density, Appendix D (Figures D.1–D.3) gives the $Q-Q$ plot, and Appendix E (Figures E.1–E.3) gives the tests for normality along with the p-values. The null hypothesis of normality were strongly rejected for all gambles except in one case in the EDHEC fund. So it is observed from the data analysis that the results are consistently the same for all the three separate sets of data. The assessments are all coming as avoiding sure loss, but need correction to be coherent. The conjecture leads to the second experiment.

The second experiment now assumed that the returns were generated as though coming from a set of normal distributions. This was performed through a simulation of data. This was conducted only for the US fund FSPCX using the six gambles. Using FSPCX data, returns were generated for each gamble as though they came from normal distributions. The simulated values were again tested for normality and plotted. The results are shown in Appendices C (Figure C.4), D (Figure D.4), E (Figure E.4). For this set of data, VaR and $CVaR$ measures (assessments) are once more calculated for α levels of 0.01 and 0.25. This is shown in Appendix A (Table A.4). These assessments are now going to be checked for coherence.

It is observed that for all the six gambles in the fund, almost no correction was required as can been in Appendix B (Table B.4). In fact, for level 0.25 no correction was required at all, i.e., all the six assessments made were coherent. This leads to the second result that if the returns are normally distributed, then the risk measures (assessments) computed from them are coherent. Therefore, it is observed that the left tail characteristic of returns is important for deciding coherence of assessments, and that if the returns were generated from a common set of distributions it is possible that one may observe the assessments to be coherent.

The third experiment was conducted to overcome a possible weakness in the modelling of returns of gambles. In the market place it may not be reasonable to expect that all assets will move in the same direction in terms of prices of stocks or their returns. At a point in time, while some stocks may generate positive returns, others may move in the negative direction or not at all. This aspect needs to be taken into account. This was achieved as follows. In the same input data that was used for checking coherence, a permutation of those discretized values is now taken row-wise (i.e., gamble-wise). This randomized the returns. So now we have, for any particular outcome in the possibility space (which for our experimental setup has seven elementary outcomes), some gambles are giving positive returns, while others are giving negative or no returns. For example in Appendix B, the discrete return values in Table B.1 are ascending starting from negative values and going to positive values, whereas in Table B.5 the return values are not placed in any particular order. This is because a permutation of the row was taken and used for each of the six rows. This means that when an elementary outcome say $X = a$ occurs, in Table B.5, the six gambles will not all give negative returns, now some will give negative and some will give positive returns.

This data along-with its earlier assessments we have for VaR and $CVaR$ (which we already have for each of the two α levels and used in the first experiment), was checked for coherence. (Notice that assessments are identical in Table B.1 and B.5). This was then repeated for the other two funds.

It is observed that the corrections are needed at the 0.01 level only for some gambles in the fund. It appears that no corrections are needed for level 0.25, for both VaR and $CVaR$. The randomized returns data and the results of correction are shown in Appendix B (Tables B.5 to B.7) for the three funds.

We may be able to explain this result as follows. From Appendix B Table B.1, notice that the rewards are high across the six gambles when say outcome $X = f$ or $X = s$ occurs in the possibility space. When we look at Table B.6 we see that none of the outcomes give a pattern of high rewards. Therefore, for a given set of lower prevision values for each gamble, the net gain in the first case will be high, whereas the net gain will be low in the second case. The definition of coherence then implies that a high supremum buying rate will be selected in the first case, whereas a low supremum buying rate will be selected in the second case. This possibly explains why a correction to a higher value was forthcoming in the first case, whereas no correction was forthcoming in the second (the randomized) case.

Additionally, we may also need to revisit our initial motivation for this experiment, i.e., the need for randomization. FSPCX consists of stocks in the financial sector, RBF consists of stocks in the banking sector, EDHEC consists of funds in the global sector. Each sector will exhibit its own systematic influences which will affect all the stocks in that sector in the same manner. So while responding to these influences, stocks in that sector could move in the same direction. If this is true, then we may not that motivated and may hesitate to randomize. Therefore it is suggested that we may use the results of the third experiment with caution.

In the next chapter we summarize the work done so far, and explore avenues where we may expand this inquiry into imprecise probability.

Chapter 7

Summary, Conclusion, and Further Work

In this thesis, we have briefly traced the literature on imprecise probability, and discussed the theory of lower prevision as given in Troffaes and Cooman [67]. Financial risk measures viz., VaR and $CVaR$ can be seen as lower previsions, as interpreted by Vicig [70]. This study addresses the problem of computationally determining the consistency and coherence of assessments of risk from the imprecise probability point of view. The algorithm written by Walley, Pelessoni and Vicig [76], and implemented in Python by Troffaes [66] was used to computationally determine consistency and coherence.

Using data from the returns distribution of mutual funds, the following results were obtained. Risk measures are consistent (avoid sure loss) when we ensure that their values are at most the maximum value of the gamble. For coherence, we impose the minimum condition that each assessment is at least the minimum for that gamble. Then, when we expect high levels of confidence, the risk measures are not coherent, whereas when we lower the confidence level, the assessments are near-coherent. The reason appears to be attributed to the long left-tail characteristic of the returns distributions. When we assume that the returns distributions are generated from normal distributions, then the assessments are near-coherent with little or no correction required, irrespective of the level of α used, high or low. Therefore for smaller α levels, we may expect coherence of assessments to be not coherent. We attribute this to the fact that actual data is not normally distributed.

In secondary markets, we observe that assets at a particular point in time move in different directions. This reality was taken into account by randomizing the returns, so that when some assets may generate positive returns, others may fetch negative or no returns. Coherence of assessments was again examined in such a situation. It is observed that overall, correction was required only in a few instances. However, stocks in the same sector could move in the same direction due to some influences specific to the sector affecting all stocks in the same manner. Therefore the results of randomization may be viewed with some caution.

The first situation mentioned above also models an adverse event when all returns are moving in the negative direction. A stock market crash is such an event. The above result suggests that when such an event occurs (when accurate risk measurement is actually needed), we need not expect assessments of risk with high levels of confidence to be coherent. Whereas in the second, in a situation when assets in the market move in different directions, assessments of risk are more accurate in terms of coherence. However coherence of lower prevision is desirable since this will avoid the opportunity for arbitrage.

It may not be possible to model all diverse sources of uncertainty and risk only by a set of normal distributions. In fact, most returns data are non-normal. Therefore, the observation is that it may not be reasonable to expect coherence to be achieved in practice computationally, if we use historical data. Therefore in this aspect, the concern pertaining to the use of VaR appear to continue to be and remain the same also for $CVaR$.

Further work may be envisaged in three directions in particular, viz., enquiry into the theoretical side, enquiry into the computational (practice) side, and enquiry into the connection between theory and practice.

What is the possible connection between consequences of rationality axioms of imprecise probability (let us call this IP-coherence) and the rationality axioms that give ADEH-coherence? Is one a subset of the other? We may conjecture or expect that ADEH-coherence is more restrictive than IP-coherence, therefore IP-coherence may not imply ADEH-coherence. In that case, in what way can we relax the axioms of ADEH-coherence such that the two notions of coherence can become equivalent to one other?

We observe from this study that in order to establish IP-coherence, the returns distributions may need to come from a set of known distributions. Suppose that the returns came from say a set of gamma distributions. In that case, would it still be possible to establish coherence? One may conjecture here that it is possible to do so. What is the possible connection between the consequences of rational betting behaviour (viz., coherence) and the distributional characteristic of the returns distributions, such that coherence can be established only under certain known conditions? In that case, can we ask the larger question, that if the returns distributions were modelled as a consequence of rational market behaviour, then is it possible that one may be able to establish IP-coherence here? One may observe from the data analysis that it is not quite possible to model returns distributions from any given set of analytic distributions. There is varied volatility. So, what if it is true that financial markets are essentially irrational?

On the applications side, one may want to proceed from inference to decision-making. This can be say in deciding an optimal portfolio for an individual or for a firm investing. The mean-VaR or mean-$CVaR$ frontier could be given not as a locus of points, but as interval-valued, with an upper envelope and a lower-envelope, and used for making optimal portfolio decisions.

We admit that these exercises are difficult.

References

[1] C. Acerbi and D. Tasche, “On the coherence of expected shortfall”, *Journal of Banking and Finance*, **26:7**, 2002, 1487–1503. 52

[2] M. Allais, “Le comportement de l’homme rationnel devant le risque: Critique des postulats et axiomes de l’ecole Americaine”, *Econometrica*, **21**, 503–546. 10

[3] P. Artzner, F. Delbaen, J.M. Eber, D. Heath, “Coherent measures of risk”, *Mathematical Finance*, **9:3**, 1999. 49, 50, 51

[4] T. Augustin, F. Coolen, G. de Cooman, and M. Troffaes, Editors, *Introduction to Imprecise Probabilities*, 2014, Wiley. 10, 17, 18, 23, 30, 49, 50, 51

[5] R.J. Aumann, “Utility theory without the completeness axiom”, *Econometrica*, 1962, 445–462. 9

[6] J-M Bernard, “Imprecise Dirichlet model for multinomial data”, *Invited Tutorial at 3rd ISIPTA*, Lugano, 2003. 11

[7] G. Boole, *The Laws of Thought*, 1958 (1854), Dover. 8

[8] M.E. Brady and R. Arthmar, “Keynes, Boole and the interval approach to probability”, *History of Economic Ideas*, **20:3**, 2012, 65–84. 9

[9] A.P. Dempster, “Upper and lower probabilities induced by a multi-valued mapping”, *The Annals of Mathematical Statistics*, **48:2**, 1967, 325–339. 10

[10] B. de Finetti, *Probability Induction and Statistics*, 1972, John Wiley & Sons. 27

[11] B. de Finetti, *Theory of Probability*, Vol **I**, 1974, John Wiley & Sons. 5, 8, 17, 27, 28, 31

[12] D. Ellsberg, “Risk ambiguity and the Savage axioms”, *Quarterly Journal of Economics*, **75:4**, 1961, 643–669. 10

[13] T.L. Fine, *Theories of Probability*, 1973, Academic Press. 4, 5, 6, 7

[14] T.L. Fine, “Lower probability models for uncertainty and non-deterministic processes”, *Journal of Statistical Planning and Inference*, **20**, 1988, 389–411. 10

[15] P.C. Fishburn, “The axioms of subjective probability”, *Statistical Science*, **1:3**, 1986, 335–345. 7, 8

[16] H. Follmer and A. Schied, “Coherent and convex risk measures”, in R. Cont, Editor, *Encyclopedia of Quantitative Finance*, 2010, John Wiley & Sons, 355–363. 51

[17] K. Fukuda, *cddlib Reference Manual*, Swiss Federal Institute of Technology, Zurich, (available from https://www.inf.ethz.ch/personal/fukudak/cdd_home/index.html. Last accessed on April 4, 2017), 2016. 40

[18] K. Fukuda and A. Prodon, “Double description method revisited”, in M. Deza, R. Euler and I. Manoussakis, Editors, *Combinatorics and Computer Science: Lecture Notes in Computer Science*, **1120**, 1996, Springer, 91–111. 40

[19] I.J. Good, *Probability and the Weighing of Evidence*, 1950, Charles Griffin & Co. Ltd. 6, 7

[20] I.J. Good, “Subjective probability as the measure of a non-measurable set”, in E. Nagel, P. Suppes and A. Tarski, Editors, *Logic Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*, 1962, Stanford University Press. 9

[21] N.T. Gridgeman, “Reviewed work: Boole’s logic and probability. A critical exposition from the standpoint of contemporary algebra, logic and probability theory by Theodore Hailperin”, *The Journal of Symbolic Logic*, **53:4**, 1988, 1253–1254. 8

[22] I. Hacking, *The Emergence of Probability,* 1975, Cambridge University Press. 9

[23] I. Hacking, *An Introduction to Probability and Inductive Logic*, 2001, Cambridge University Press. 4, 7

[24] T. Hailperin, *Booles Logic and Probability: A Critical Exposition from the Standpoint of Contemporary Algebra, Logic, and Probability Theory*, 2nd Revised Edition, 1986 (1976), North-Holland. 8

[25] F. Hampel, "Non-additive probabilities in statistics", *Journal of Statistical Theory and Practice*, **3:1**, 2009, 11–23. 9

[26] J.A. Hartigan, *Bayes Theory*, 1983, Springer-Verlag. 6

[27] C. Howson, "Lecture" delivered on Feb 11, 2011 at the University of Western Ontario, *Rotman Instiute of Philosophy Speaker Series*, **8**, (video presentation available from http://ir.lib.uwo.ca/rotmanseries/8/. Last accessed on April 5, 2017), 2011. 9

[28] E. Jaynes, in L.G. Bretthorst, Editor, *Probability Theory: The Logic of Science*, 2003, Cambridge University Press. 6

[29] H. Jeffreys, *Theory of Probability*, 1961, Oxford University Press. 6

[30] R. Jeffrey, "Bayesianism with a human face", in J. Earman, Editor, *Testing Scientific Theories: Minnesota Studies in the Philosophy of Science*, **10**, 1983, University of Minnesota Press, 133–156. 9

[31] J.M. Keynes, *A Treatise on Probability*, 1921, Macmillan. 9

[32] A.N. Kolmogorov, *Foundations of the Theory of Probability*, 2nd English Edition, 1956, Chesea. 7

[33] B.O. Koopman, "The axioms and algebra of intuitive probability", *Annals of Mathematics*, **41:2**, 1940. 9

[34] I. Kozine, "A commentary on the book by Kuznetsov titled *Interval Statistical Models* (1991)", *SIPTA*, (available from http://www.sipta.org/index.php?id=kuznetsov. Last accessed on April 3, 2017), 2014. 11

[35] H.E. Kyburg, "Subjective probability: Criticisms, reflections, and problems", *Journal of Philosophical Logic*, **7**, 1978, 157–180. 4

[36] H.E. Kyburg, "Self profile", in R.J. Bogdan, Editor, *Henry E. Kyburg Jr. & Isaac Levi*, *Profiles: An International Series on Contemporary Philosophers and Logicians*, **3**, 1982, D. Reidel, 5–53. 10

[37] H.E. Kyburg, "Interval-valued probabilities", a contribution to the *Documentation Section* of the *Society for Imprecise Probability Theory and Applications (SIPTA)*, (available from http://www.sipta.org/documentation/interval_prob/kyburg.pdf. Last accessed on April 6, 2017), 1999. 9

[38] I. Levi, *The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance*, 1980, MIT Press. 10

[39] I. Levi, "Probability logic and logic of probability", *4th Annual Formal Epistemology Workshop*, Carnegie Mellon University, Pittsburgh, 2007. 4

[40] E. Miranda, "A survey of the theory of coherent lower previsions", *IJAR*, **48**, 2008, 628–658. 10

[41] E. Miranda, A presentation titled "General introduction to imprecise probability", *4th SIPTA Summer School on Imprecise Probabilities*, Durham, (available from https://www.sipta.org/ssipta10/material/coherent_lower_previsions/clp-miranda.pdf. Last accessed on April 6, 2017), 2010. 17

[42] A. de Moivre, *The Doctrine of Chances; or a Method of Calculating the Probability of Events in Play*, 3rd edition, (1st ed., 1718; 2nd ed., 1738), 1756, A. Millar. 4

[43] L. Narens, "On qualitative axiomatizations for probability theory", *Journal of Philosophical Logic*, **9**, 1980, 143–151. 7

[44] L. Narens, *Theories of Probability: An Examination of Logical and Qualitative Foundations*, 2007, World Scientific. 6

[45] R. Pelessoni and P. Vicig, "Coherent risk measures and upper previsions", *2nd ISIPTA*, Ithaca, 2001. 49

[46] R. Pelessoni and P. Vicig, "Imprecise previsions for risk measurement", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **11:4**, 2003, 393–412. 49

[47] B.G. Peterson, R-Package $PerformanceAnalytics$, (available from https://cran.r-project.org/package =PerformanceAnalytics. Last accessed on April 2, 2017), 2014. 54

[48] G. Ch. Pflug, "Some remarks on the value-at-risk and the conditional value-at-risk", in S. Uryasev, Editor, *Probabilistic Constrained Optimization: Methodology and Applications*, 2000, Kluwer Academic. 52

[49] E. Quaeghebeur, "Characterizing coherence, correcting incoherence", *8th ISIPTA*, Compiengne, 2013. 40

[50] F.P. Ramsey, "Truth and probability", in R.B. Braithwaite, Editor, *The Foundations of Mathematics and other Logical Essays*, **Ch. VII**, 1931 (1926), Harcourt Brace and Company, 156–198.

[51] R.T. Rockafeller and S. Uryasev, "Optimization of conditional value-at-risk", *Journal of Risk*, **2:3**, 2000, 21–41. 52

[52] R.T. Rockafeller and S. Uryasev, "Conditional value-at-risk for general loss distributions", *Journal of Banking and Finance*, **26:7**, 2002, 1443–1471. 51

[53] L.J. Savage, *Foundations of Statistics*, 1972, Dover. 6

[54] E. Schechter, *Handbook of Analysis and its Foundations*, 1997, Academic Press. 16

[55] G. Shafer, *A Mathematical Theory of Evidence*, 1976, Princeton University Press. 9

[56] G. Shafer, "Non-additive probabilities in the work of Bernoulli and Lambert, *Archive for History of Exact Sciences*, **19:4**, 1978, Springer, 309–370. 9

[57] G. Shafer, "Lambert, Johann Heinrich", in N.L. Johnson and S. Kotz, Editors, *Leading Personalities in Statistical Sciences: From the 17th Century to the Present*, 1997, John Wiley & Sons, 47–48. 9

[58] B. Skyrms, *Choice and Chance: An Introduction to Inductive Logic*, 2000, Wadsworth. 6

[59] C.A.B. Smith, "Consistency in statistical inference and decision", *Journal of the Royal Statistical Society, Series B*, **23**, 1961, 1–25. 10

[60] C.A.B. Smith, "Personal probability and statistical analysis", *Journal of the Royal Statistical Society, Series A*, **128**, 1965, 469–499. 10

[61] S. Spielman, "Kyburg's system of probability", in R.J. Bogdan, Editor, *Henry E. Kyburg Jr. & Isaac Levi*, 1981, D. Riedel, 57–96. 9

[62] *Stanford Encyclopedia of Philosophy (SEP)*, The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, (available from https://plato.stanford.edu/entries/prob ability-interpret. Last accessed on May 10, 2017), 2011. 3

[63] *Stanford Encyclopedia of Philosophy (SEP)*, The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, (available from https://plato.stanford.edu/entries/impr ecise-probabilities. Last accessed on May 10, 2017), 2014. 5, 8

[64] P. Suppes, "Qualitative theory of subjective probability", in G. Wright and P. Ayton, Editors, *Subjective Probability*, 1994, Wiley & Sons. 6

[65] D. Tasche, "Expected shortfall and beyond", *Journal of Banking and Finance*, **26:7**, 2002, 1519–1533. 50

[66] M. Troffaes, *improb* 0.1.1: A Python module for working with imprecise probabilities, Python Software Foundation, (available from https://pypi.python.org/pypi/improb. Last accessed on April 3, 2017), 2011. 2, 39, 59

[67] M. Troffaes and G. de Cooman, *Lower Previsions*, 2014, Wiley. 2, 12, 13, 16, 17, 18, 21, 22, 23, 24, 26, 29, 30, 35, 59

[68] M. Troffaes, *pycddlib* 1.0.5: A Python wrapper for Komei Fukuda's *cddlib*, Python Software Foundation, (available from https://pypi.python.org/pypi/pycddlib. Last accessed on April 3, 2017), 2015. 40

[69] P. Vicig, M. Zaffalon and F. Cozman, "Notes on "Notes on conditional previsions"", *IJAR*, **44**, 2007, 358–365. 10

[70] P. Vicig, "Financial risk measurement with imprecise probabilities", *IJAR*, **49**, 2008, 159–174. 2, 49, 59

[71] P. Vicig and T. Seidenfeld, "Bruno de Finetti and imprecision", *7th ISIPTA*, Innsbruk, 2011. 9

[72] C. Villegas, "On qualitative probability σ-algebras", *Annals of Mathematical Statistics*, **35:4**, 1964, 1787–1796. 7, 8

[73] R. von Mises, *Probability, Statistics, and Truth*, 1981 (1939), Dover. 4

[74] J. von Plato, *Creating Modern Probability*, 1994, Cambridge University Press. 6, 7, 8

[75] P. Walley, "Towards a unified theory of imprecise probability", *IJAR*, **24**, 2000, 125–148. 11

[76] P. Walley, R. Pelessoni, and P. Vicig, "Direct algorithms for checking consistency and making inferences from conditional probability assessments", *Journal of Statistical Planning and Inference*, **126**, 2004, 119–151. 2, 37, 38, 39, 59

[77] P. Walley and T.L. Fine, "Towards a frequentist theory of upper and lower probability", *Annals of Statistics*, **10:3**, 1982, 741–761. 10

[78] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, 1991, Chapman and Hall. 8, 10, 17, 30, 31, 33, 35

[79] P. Walley, "Inferences from multinomial data: Learning about a bag of marbles", *Journal of the Royal Statistical Society, Series B*, **58:1**, 1996, 3–57. 11

[80] L. Wasserman and J.B. Kadane, "Entropy when probabilities are imprecise", in D.A. Berry, K.M. Chaloner and J.K. Geweke, Editors, *Bayesian Analysis in Statistics and Econometrics*, 1996, Wiley. 10

[81] K. Weichselberger, "The theory of interval-probability as a unifying concept for uncertainty", *IJAR*, **24**, 2000, 149–170. 11

[82] P.M. Williams, "Notes on conditional previsions", *IJAR*, **44**, 2007, 366–383. 10, 17, 33

[83] S. Zabell, "Philosophy of inductive logic: The Bayesian perspective", in L. Haaparanta, Editor, *The Development of Modern Logic*, 2009, Oxford University Press. 6

Appendix A

Risk Measures and Gamble Values (Fund-wise) [6.3]

Each table given below corresponds to one fund and the assets in that fund. This Appendix presents the two risk measures (VaR and $CVaR$), and the five discretized values of the returns along-with the minimum and the maximum return (total seven values labeled [,1] to [,7]). The six rows correspond to the six stocks of the fund. VaR and $CVaR$ (which is ES, the expected shortfall) is calculated using historical data (given in the first column), and under the assumption that data is gaussian (given in the second column). Two levels of α are used ($p = 0.75$ corresponds to $\alpha = 0.25$, and $p = 0.99$ corresponds to $\alpha = 0.01$). The risk measures are used as lower previsions. The objective is to determine whether the risk measures are consistent and coherent. See Appendix F for more explanation about the format presented here.

Table A.1: Fund FSPCX

```
## Uses Fund FSCPX data
 # Calculating VaR and CVaR with "returns" data (risk measures for each gamble)

> var.h = VaR(S,  p = 0.75, method = "historical"); var.g = VaR(S, p = 0.75, method = "gaussian")
> var = cbind(t(var.h), t(var.g))
> var
                     VaR                  VaR
ACE -0.00416255398062501 -0.00485813818419109
AFL -0.00532209447579059 -0.00650122709958653
AIG -0.00662149332524494 -0.00789788948190040
ALL -0.00390077084348905 -0.00503549584341533
CB  -0.00412346102321256 -0.00472669057825151
MET -0.00694271863174766 -0.00863148123557469

> cvar.h = ES(S, p = 0.75, method = "historical"); cvar.g = ES(S, p = 0.75, method = "gaussian")
> cvar = cbind(t(cvar.h), t(cvar.g))
> cvar
                      ES                   ES
ACE -0.00979197426713049 -0.00988184543881942
AFL -0.01252857922008094 -0.01267747817417615
AIG -0.01494115078408693 -0.01577030958641306
ALL -0.00974518760013264 -0.01051767576000751
CB  -0.00961469655723048 -0.00952777994651768
MET -0.01658153456934341 -0.01717474257827291

> var.h = VaR(S,  p = 0.99, method = "historical"); var.g = VaR(S, p = 0.99, method = "gaussian")
> var = cbind(t(var.h), t(var.g))
> var
                    VaR                 VaR
ACE -0.0211325638447521 -0.0187673262078082
AFL -0.0277308187539020 -0.0236014746107027
AIG -0.0333801532228247 -0.0296943369826358
ALL -0.0209304318131117 -0.0202140616662799
CB  -0.0219412395766058 -0.0180195142554034
MET -0.0345183253970429 -0.0322852935155422

> cvar.h = ES(S, p = 0.99, method = "historical"); cvar.g = ES(S, p = 0.99, method = "gaussian")
> cvar = cbind(t(cvar.h), t(cvar.g))
> cvar
                     ES                  ES
ACE -0.0259386762439764 -0.0216206921583061
AFL -0.0344908645873863 -0.0271094625615778
AIG -0.0438486004524433 -0.0341657152568249
ALL -0.0247096666322696 -0.0233278309835696
CB  -0.0269576254314673 -0.0207464376676080
MET -0.0404791011206761 -0.0371376962902065

> # Calculating gamble (asset) returns discretized from a continuous possibility space --
> # discretized values of each gamble

> gamble = rbind(ace,afl,aig,all,cb,met)
> gamble
                   [,1]                [,2]                 [,3]                 [,4]
ace -0.0345690049510583 -0.0110457103299400 -0.00311056872490512 0.001103174758187344
afl -0.0431856429522155 -0.0140600313482144 -0.00409161821020374 0.000610452617346244
aig -0.0652468454267962 -0.0168151250779751 -0.00536133237037656 0.001215629688728013
all -0.0295249767342611 -0.0110576512039879 -0.00298562927585648 0.000672703278054171
cb  -0.0338507623929602 -0.0107621250372860 -0.00279896338200148 0.000912853838716419
met -0.0432892303022039 -0.0187375939382935 -0.00534912522483403 0.000781329951828878

                   [,5]               [,6]               [,7]
ace 0.00490758684743495 0.0122519528312329 0.0376086159280909
afl 0.00533382917747575 0.0146132549919804 0.0301556734502639
aig 0.00697809632431584 0.0189931118794075 0.0567291319558037
all 0.00528959401882367 0.0138922036684100 0.0567102422247075
cb  0.00489305742715011 0.0112604754417573 0.0321256106756230
met 0.00795664000824624 0.0204831839090866 0.0629905090929420
```

Table A.2: Fund RBF

```
## Uses Fund RBF data
  # Calculating VaR and CVaR with "returns" data (risk measures for each gamble)

> var.h = VaR(S,  p = 0.75, method = "historical"); var.g = VaR(S, p = 0.75, method = "gaussian")
> var = cbind(t(var.h), t(var.g))
> var
                       VaR                 VaR
axis  -0.00860986433427233 -0.02828357742149235
hdfc  -0.00685193209933265 -0.00862147433225266
icici -0.00975294552362036 -0.02800396736129441
idfc  -0.01397106953247553 -0.01528178426020931
sbi   -0.00996588640407770 -0.03088963270748658
yes   -0.01162224055898099 -0.01808374036241847

> cvar.h = ES(S, p = 0.75, method = "historical"); cvar.g = ES(S, p = 0.75, method = "gaussian")
> cvar = cbind(t(cvar.h), t(cvar.g))
> cvar
                       ES                  ES
axis  -0.0293429230222401 -0.0531226673729922
hdfc  -0.0153439741569217 -0.0169501360780711
icici -0.0272038502680895 -0.0521932964385219
idfc  -0.0267243270574585 -0.0287753372419719
sbi   -0.0266107733990566 -0.0571111077170961
yes   -0.0304362345820863 -0.0352576120462429

> var.h = VaR(S,  p = 0.99, method = "historical"); var.g = VaR(S, p = 0.99, method = "gaussian")
> var = cbind(t(var.h), t(var.g))
> var
                      VaR                 VaR
axis  -0.0626624545535133 -0.0970558117416051
hdfc  -0.0335162199694128 -0.0316811225832466
icici -0.0424683949967295 -0.0949772023144918
idfc  -0.0529525754611429 -0.0526415180219769
sbi   -0.0403885525681367 -0.1034892902271841
yes   -0.0825216039484376 -0.0656332089592825

> cvar.h = ES(S, p = 0.99, method = "historical"); cvar.g = ES(S, p = 0.99, method = "gaussian")
> cvar = cbind(t(cvar.h), t(cvar.g))
> cvar
                       ES                  ES
axis  -0.2155270142747848 -0.1111639215331671
hdfc  -0.0428800685731049 -0.0364116370488015
icici -0.1988633723182135 -0.1087162608336282
idfc  -0.0763941046987219 -0.0603055881355227
sbi   -0.2173712150267768 -0.1183825672777765
yes   -0.1054986046555254 -0.0753876270049792

> # Calculating gamble (asset) returns discretized from a continuous possibility space --
> # discretized values of each gamble

> gamble = rbind(axis.p,hdfc.p,icici.p,idfc.p,sbi.p,yes.p)
> gamble
                        [,1]                [,2]                 [,3]                  [,4]
axis.p  -0.7985692487550452 -0.0342002835314558 -0.00668691529049488  0.001229662099198810
hdfc.p  -0.0701485939173551 -0.0172331442936279 -0.00527086618467318  0.000632309921772782
icici.p -0.7968876037646193 -0.0312517458985536 -0.00755256157233249  0.000136673311017268
idfc.p  -0.1457027372138480 -0.0296328887733738 -0.01120362146869251 -0.000698787075985230
sbi.p   -0.8989883469005924 -0.0305851519574247 -0.00768965846908727 -0.000236017182371143
yes.p   -0.1303663628943856 -0.0347903336100709 -0.00885349929279001  0.001157533191786451

                        [,5]               [,6]               [,7]
axis.p  0.00942118334861200 0.0292244929024093 0.1247143073004269
hdfc.p  0.00586543474024665 0.0199776842880992 0.0641764820500203
icici.p 0.00829836400635116 0.0270826063155096 0.1048599180163077
idfc.p  0.01028456130694588 0.0311157910540921 0.0798703663523833
sbi.p   0.00799300736854518 0.0242897857743042 0.0779780520869169
yes.p   0.01160015808684604 0.0375448725919181 0.2222839247596091
```

Table A.3: Fund EDHEC

```
## Uses Fund from EDHEC data (first six assets are used)
 # Calculating VaR and CVaR with "returns" data (risk measures for each gamble)

> var.h = VaR(S,  p = 0.75, method = "historical")
VaR calculation produces unreliable result (inverse risk) for column: 1 : -0.000575
VaR calculation produces unreliable result (inverse risk) for column: 5 : -0.0024

> var.g = VaR(S, p = 0.75, method = "gaussian")
> var = cbind(t(var.h), t(var.g))
> var
                             VaR                  VaR
Convertible Arbitrage         NA -0.007226739302723630
CTA Global             -0.011700 -0.010744659699910581
Distressed Securities  -0.000725 -0.004543113317276576
Emerging Markets       -0.011800 -0.018210101218948921
Equity Market Neutral         NA -0.000158991660914751
Event Driven           -0.000275 -0.004869285231128635

> cvar.h = ES(S, p = 0.75, method = "historical")
> cvar.g = ES(S, p = 0.75, method = "gaussian")
> cvar = cbind(t(cvar.h), t(cvar.g))
> cvar
                                        ES                   ES
Convertible Arbitrage  -0.01597894736842105 -0.01922064414544062
CTA Global             -0.02398421052631579 -0.02565808234881528
Distressed Securities  -0.01426315789473684 -0.01551363557172443
Emerging Markets       -0.04154210526315789 -0.04093297257793453
Equity Market Neutral  -0.00408378378378378 -0.00550900733288450
Event Driven           -0.01524736842105263 -0.01583347288773414

> var.h = VaR(S,  p = 0.99, method = "historical")
> var.g = VaR(S, p = 0.99, method = "gaussian")
> var = cbind(t(var.h), t(var.g))
> var
                                      VaR                VaR
Convertible Arbitrage  -0.0680079999999999 -0.0404343824115108
CTA Global             -0.0503090000000000 -0.0520356006694859
Distressed Securities  -0.0649070000000000 -0.0349173069197818
Emerging Markets       -0.1159990000000000 -0.0811231401659164
Equity Market Neutral  -0.0211500000000000 -0.0149716330309245
Event Driven           -0.0626020000000000 -0.0352259401698589

> cvar.h = ES(S, p = 0.99, method = "historical")
> cvar.g = ES(S, p = 0.99, method = "gaussian")
> cvar = cbind(t(cvar.h), t(cvar.g))
> cvar
                            ES                  ES
Convertible Arbitrage  -0.11320 -0.0472466821648833
CTA Global             -0.05375 -0.0605061285555533
Distressed Securities  -0.08055 -0.0411483456872540
Emerging Markets       -0.16265 -0.0940292798247388
Equity Market Neutral  -0.04360 -0.0180103356820322
Event Driven           -0.07565 -0.0414533810115608

> gamble = rbind(s1.p,s2.p,s3.p,s4.p,s5.p,s6.p)
> gamble
        [,1]                [,2]                 [,3]                [,4]
s1.p -0.1237 -0.0201033333333333  0.00275000000000000 0.00895333333333333
s2.p -0.0543 -0.0269366666666667 -0.00786666666666667 0.00424000000000000
s3.p -0.0836 -0.0175633333333333  0.00140666666666667 0.00957666666666667
s4.p -0.1922 -0.0483066666666667 -0.00505333333333333 0.01313666666666667
s5.p -0.0587 -0.0055100000000000  0.00348000000000000 0.00637333333333333
s6.p -0.0886 -0.0188966666666667  0.00187666666666667 0.00976666666666667

                    [,5]                [,6]   [,7]
s1.p  0.0137333333333333 0.0263300000000000 0.0611
s2.p  0.0188100000000000 0.0423300000000000 0.0691
s3.p  0.0167233333333333 0.0291533333333333 0.0504
s4.p  0.0265766666666667 0.0510400000000000 0.1230
s5.p  0.0091600000000000 0.0159433333333333 0.0253
s6.p  0.0164966666666667 0.0283866666666667 0.0442
```

Table A.4: Simulating Fund FSPCX

```
## Uses Simulated Fund FSCPX data (6 assets)
 # Calculating VaR and CVaR with "returns" data (risk measures for each gamble)

# VaR (Historical and Gaussian) with 75% confidence
> var
                     VaR                  VaR
[1,] -0.00488268740482409 -0.00482731237639477
[2,] -0.00632629354482058 -0.00633223181457504
[3,] -0.00829040611998566 -0.00828110393755183
[4,] -0.00493489358458181 -0.00492943347400483
[5,] -0.00469619258653538 -0.00471664593181629
[6,] -0.00900363458594299 -0.00881243329550233

# CVaR (Historical and Gaussian) with 75% confidence
> cvar
                      ES                   ES
[1,] -0.00981500507338899 -0.00983007812376588
[2,] -0.01237678135982996 -0.01241225006154608
[3,] -0.01608983205668858 -0.01618192076686943
[4,] -0.01039694225411507 -0.01032860459389577
[5,] -0.00941727626300497 -0.00945039463811709
[6,] -0.01759486403566525 -0.01753347401345542

# VaR (Historical and Gaussian) with 99% confidence
> var
                    VaR                 VaR
[1,] -0.0191600082749048 -0.0186785194414936
[2,] -0.0237544005511929 -0.0231660385429091
[3,] -0.0309775139629290 -0.0301561737314559
[4,] -0.0201038888063480 -0.0198781720194403
[5,] -0.0183022952740826 -0.0178230228513375
[6,] -0.0316671095530977 -0.0329584650922147

# CVaR (Historical and Gaussian) with 99% confidence
> cvar
                     ES                  ES
[1,] -0.0214293532016757 -0.0215199910317648
[2,] -0.0269296395337209 -0.0266193681588675
[3,] -0.0348811678144752 -0.0346436807813792
[4,] -0.0228063394142126 -0.0229447939887469
[5,] -0.0203000738089036 -0.0205116981049220
[6,] -0.0362271294788530 -0.0379118430213500

> # Calculating gamble (asset) returns discretized from a continuous possibility space --
> # discretized values of each gamble

> gamble = rbind(ace,afl,aig,all,cb,met)
> gamble
                   [,1]                [,2]                 [,3]                 [,4]
ace -0.0271687940282466 -0.0108888691860873 -0.00370131167758230 0.000835931470848067
afl -0.0345174665231812 -0.0136928727524932 -0.00490123588210311 0.000489074632990284
aig -0.0424288153790054 -0.0178163576304495 -0.00643564920722349 0.000650485371440119
all -0.0277653639683724 -0.0115598332619903 -0.00361064309293271 0.001235289534078458
cb  -0.0248514239241166 -0.0104485479843596 -0.00357090123210605 0.000551178270285249
met -0.0439104132904057 -0.0194537415772671 -0.00689953989994164 0.000972377355401188
                   [,5]               [,6]               [,7]
ace 0.00533807775778085 0.0125583615974073 0.0268851234950690
afl 0.00604934199084488 0.0147625877094900 0.0315706544883724
aig 0.00758607803032093 0.0192702675992051 0.0442586797281074
all 0.00601137547082904 0.0137961261973543 0.0299729134356181
cb  0.00485310937587753 0.0117900316741958 0.0262736898242024
met 0.00902962221212172 0.0215858762123444 0.0473016630802702
```

Appendix B

Coherent Extensions (Fund-wise) [6.4]

This appendix gives the main results of this thesis. Given discrete values of return and an assessment, the coherent extension of the assessment is presented here. Notice that the assessment of each gamble is within the minimum and the maximum return for that gamble. This means that the assessments avoid sure loss, and that they also meet a necessary condition for coherence.

The corrected lower previsions of each of the six gambles of each fund are given. The gambles are given row-wise. For the possibility space $\mathcal{X} = \{i, a, b, c, d, e, s\}$, the columns correspond to the seven discretized values of the returns of each gamble. As discussed in Chapter 6, VaR and $CVaR$ are used as risk measures, and two levels of α (0.01 and 0.25) are used. Notice that correction is more for level 0.01 for both risk measures.

Table B.1: Fund FSPCX

Fund FSPCX (6 assets used)
Discretized values

i	a	b	c	d	e	s
-0.065246845	-0.016815125	-0.005361332	0.001215630	0.006978096	0.018993112	0.056729132
-0.043289230	-0.018737594	-0.005349125	0.000781330	0.007956640	0.020483184	0.062990509
-0.043185643	-0.014060031	-0.004091618	0.000610453	0.005333829	0.014613255	0.030155673
-0.034569005	-0.011045710	-0.003110569	0.001103175	0.004907587	0.012251953	0.037608616
-0.033850762	-0.010762125	-0.002798963	0.000912854	0.004893057	0.011260475	0.032125611
-0.029524977	-0.011057651	-0.002985629	0.000672703	0.005289594	0.013892204	0.056710242

CVaR
Using 99% confidence

Assessment	Coherent Extension
[-0.0438486004524]	[-0.0438486004524]
[-0.0404791011207]	[-0.0324417541925]
[-0.0344908645874]	[-0.0303193762594]
[-0.025938676244]	[-0.0241758738114]
[-0.0269576254315]	[-0.0236496728033]
[-0.0247096666323]	[-0.0213656887124]

CVaR
Using 75% confidence

Assessment	Coherent Extension
[-0.0149411507841]	[-0.0149411507841]
[-0.0165815345693]	[-0.0165815345693]
[-0.0125285792201]	[-0.0125285792201]
[-0.00979197426713]	[-0.00979197426713]
[-0.00961469655723]	[-0.00958277207039]
[-0.00974518760013]	[-0.00974518760013]

VaR
Using 99% confidence

Assessment	Coherent Extension
[-0.0333801532228]	[-0.0333801532228]
[-0.034518325397]	[-0.0271349530336]
[-0.0277308187539]	[-0.0240249425768]
[-0.0211325638448]	[-0.0190913472524]
[-0.0219412395766]	[-0.0186590967919]
[-0.0209304318131]	[-0.0173740026564]

VaR
Using 75% confidence

Assessment	Coherent Extension
[-0.00662149332524]	[-0.00662149332524]
[-0.00694271863175]	[-0.00694271863175]
[-0.00532209447579]	[-0.00532209447579]
[-0.00416255398063]	[-0.00406223125009]
[-0.00412346102321]	[-0.00409375319756]
[-0.00390077084349]	[-0.00390077084349]

Table B.2: Fund RBF

Fund from RBF (6 assets used)
Discretized values

i	a	b	c	d	e	s
-0.8989	-0.0305	-0.0076	-0.0002	0.0079	0.0242	0.0779
-0.7985	-0.0342	-0.0066	0.0012	0.0094	0.0292	0.1247
-0.7968	-0.0312	-0.0075	0.0001	0.0082	0.0270	0.1048
-0.1457	-0.0296	-0.0112	-0.0006	0.0102	0.0311	0.0798
-0.1303	-0.0347	-0.0088	0.0011	0.0116	0.0375	0.2222
-0.0701	-0.0172	-0.0052	0.0006	0.0058	0.0199	0.0641

CVaR
Using 99% confidence

Assessment	Coherent Extension
[-0.2173]	[-0.2173]
[-0.2155]	[-0.19860723169]
[-0.1988]	[-0.195886872409]
[-0.0763]	[-0.0545740672501]
[-0.1054]	[-0.0552643482266]
[-0.0428]	[-0.0285792261631]

CVaR
Using 75% confidence

Assessment	Coherent Extension
[-0.0266]	[-0.0266]
[-0.0293]	[-0.0293]
[-0.0272]	[-0.0270696484501]
[-0.0267]	[-0.0263333333333]
[-0.0304]	[-0.0301018115942]
[-0.0153]	[-0.0150695652174]

VaR
Using 99% confidence

Assessment	Coherent Extension
[-0.0403]	[-0.0403]
[-0.0626]	[-0.0428252187932]
[-0.0424]	[-0.0398398894519]
[-0.0529]	[-0.0309102026716]
[-0.0825]	[-0.0357788576693]
[-0.0335]	[-0.0177969829572]

VaR
Using 75% confidence

Assessment	Coherent Extension
[-0.0099]	[-0.0099]
[-0.0086]	[-0.0086]
[-0.0097]	[-0.00950803029345]
[-0.0139]	[-0.0125333333333]
[-0.0116]	[-0.0106768115942]
[-0.0068]	[-0.00606956521739]

Table B.3: Fund EDHEC

Fund from EDHEC (6 assets used)
Discretized values

i	a	b	c	d	e	s
-0.192200	-0.048306	-0.005053	0.013136	0.026576	0.051040	0.123000
-0.123700	-0.020103	0.002750	0.008953	0.013733	0.026330	0.061100
-0.088600	-0.018896	0.001876	0.009766	0.016496	0.028386	0.044200
-0.083600	-0.017563	0.001406	0.009576	0.016723	0.029153	0.050400
-0.058700	-0.005510	0.003480	0.006373	0.009160	0.015943	0.025300
-0.054300	-0.026936	-0.007866	0.004240	0.018810	0.042330	0.069100

CVaR
Using 99% confidence

Assessment	Coherent Extension
[-0.16265]	[-0.151350227486]
[-0.1132]	[-0.0942900601617]
[-0.07565]	[-0.0688118744125]
[-0.08055]	[-0.064852891521]
[-0.0436]	[-0.0436]
[-0.05375]	[-0.0465316901673]

CVaR
Using 75% confidence

Assessment	Coherent Extension
[-0.0415421052632]	[-0.0405005631438]
[-0.0159789473684]	[-0.0159789473684]
[-0.0152473684211]	[-0.0152473684211]
[-0.0142631578947]	[-0.0141398540512]
[-0.00408378378378	[-0.00408378378378]
[-0.0239842105263]	[-0.0234946275901]

VaR
Using 99% confidence

Assessment	Coherent Extension
[-0.115999]	[-0.0906166253055]
[-0.0680079]	[-0.0505646860312]
[-0.062602]	[-0.0393917804099]
[-0.064907]	[-0.036980534875]
[-0.02115]	[-0.02115]
[-0.050309]	[-0.0349821169393]

VaR *
Using 75% confidence

Assessment	Coherent Extension
[-0.0118]	[-0.00953197183709]
[-0.000275]	[-0.000275]
[-0.000725]	[-0.000614954794642]
[-0.0117]	[-0.00984075303293]

* 4 assets used. R-package could not give reliable VaR estimates for assets 1, 5

Table B.4: Simulated FSPCX

Fund FSPCX (6 assets) Simulated from Normal Distributions
Discretized values

i	a	b	c	d	e	s
-0.043910413	-0.019453742	-0.006899540	0.000972377	0.009029622	0.021585876	0.047301663
-0.042428815	-0.017816358	-0.006435649	0.000650485	0.007586078	0.019270268	0.044258680
-0.034517467	-0.013692873	-0.004901236	0.000489075	0.006049342	0.014762588	0.031570654
-0.027765364	-0.011559833	-0.003610643	0.001235290	0.006011375	0.013796126	0.029972913
-0.027168794	-0.010888869	-0.003701312	0.000835931	0.005338078	0.012558362	0.026885123
-0.024851424	-0.010448548	-0.003570901	0.000551178	0.004853109	0.011790032	0.026273690

CVaR
Using 99% confidence

CVaR
Using 75% confidence

Assessment	Coherent Extension	Assessment	Coherent Extension
[-0.0362271294789]	[-0.0349991713358]	[-0.0175948640357]	[-0.0175604824492]
[-0.0348811678145]	[-0.0334608098929]	[-0.0160898320567]	[-0.0160898320567]
[-0.0269296395337]	[-0.0269296395337]	[-0.0123767813598]	[-0.0123767813598]
[-0.0228063394142]	[-0.0218605782796]	[-0.0103969422541]	[-0.0103827241284]
[-0.0214293532017]	[-0.0212369014562]	[-0.00981500507339]	[-0.00980785732414]
[-0.0203000738089]	[-0.0196034689347]	[-0.009417276263]	[-0.009417276263]

VaR
Using 99% confidence

VaR
Using 75% confidence

Assessment	Coherent Extension	Assessment	Coherent Extension
[-0.0316671095531]	[-0.031270129691]	[-0.00900363458594]	[-0.00893448078334]
[-0.0309775139629]	[-0.0297080147047]	[-0.00829040611999]	[-0.00829040611999]
[-0.0237544005512]	[-0.0237544005512]	[-0.00632629354482]	[-0.00632629354482]
[-0.0201038888063]	[-0.0193896329445]	[-0.00493489358458]	[-0.00493489358458]
[-0.0191600082749]	[-0.018754612843]	[-0.00488268740482]	[-0.00486636024075]
[-0.0183022952741]	[-0.0174073841684]	[-0.00469619258654]	[-0.00469619258654]

Table B.5: Fund FSPCX Randomized Returns

Fund from FSPCX Randomized Returns
Discretized values

i	a	b	c	d	e	s
-0.005361332	-0.016815125	0.056729132	0.018993112	0.006978096	0.001215630	-0.065246845
-0.004091618	0.000610453	-0.014060031	-0.043185643	0.005333829	0.014613255	0.030155673
-0.002798963	-0.033850762	-0.010762125	0.000912854	0.004893057	0.011260475	0.032125611
0.004907587	-0.003110569	0.037608616	-0.034569005	-0.011045710	0.012251953	0.001103175
0.013892204	-0.029524977	0.056710242	-0.002985629	-0.011057651	0.000672703	0.005289594
0.020483184	0.007956640	-0.043289230	0.000781330	0.062990509	-0.005349125	-0.018737594

CVaR
Using 99% confidence

Assessment	Coherent Extension
[-0.0438486004524]	[-0.0438486004524]
[-0.0344908645874]	[-0.0344908645874]
[-0.0269576254315]	[-0.0269576254315]
[-0.025938676244]	[-0.025938676244]
[-0.0247096666323]	[-0.0247096666323]
[-0.0404791011207]	[-0.0404791011207]

CVaR
Using 75% confidence

Assessment	Coherent Extension
[-0.0149411507841]	[-0.0149411507841]
[-0.0125285792201]	[-0.0125285792201]
[-0.00961469655723]	[-0.00961469655723]
[-0.00979197426713]	[-0.00979197426713]
[-0.00974518760013]	[-0.00974518760013]
[-0.0165815345693]	[-0.0165815345693]

VaR
Using 99% confidence

Assessment	Coherent Extension
[-0.0333801532228]	[-0.0333801532228]
[-0.0277308187539]	[-0.0277308187539]
[-0.0219412395766]	[-0.0219412395766]
[-0.0211325638448]	[-0.0211325638448]
[-0.0209304318131]	[-0.0209304318131]
[-0.034518325397]	[-0.034518325397]

VaR
Using 75% confidence

Assessment	Coherent Extension
[-0.00662149332524]	[-0.00662149332524]
[-0.00532209447579]	[-0.00532209447579]
[-0.00412346102321]	[-0.00412346102321]
[-0.00416255398063]	[-0.00416255398063]
[-0.00390077084349]	[-0.00390077084349]
[-0.00694271863175]	[-0.00694271863175]

Table B.6: Fund RBF Randomized Returns

```
# Fund from RBF (6 assets used)

# Discretized Values
     i          a          b          c          d          e          s
-0.070148  -0.017233   0.019977   0.064176   0.005865  -0.005270   0.000632
-0.008853  -0.130366   0.011600   0.037544   0.001157   0.222283  -0.034790
 0.000136   0.104859  -0.007552  -0.796887   0.027082  -0.031251   0.008298
 0.007993  -0.898988  -0.030585  -0.007689   0.077978   0.024289  -0.000236
 0.029224   0.001229  -0.034200  -0.798569   0.124714   0.009421  -0.006686
 0.031115  -0.011203   0.010284  -0.000698  -0.145700  -0.029632   0.079870


# CVaR                                      # CVaR
# Using 99% confidence                      # Using 75% Confidence
Assessment          Coherent Extension      Assessment          Coherent Extension
[-0.0428]           [-0.0428]               [-0.0153]           [-0.0153]
[-0.1054]           [-0.0578735712189]      [-0.0304]           [-0.0304]
[-0.1988]           [-0.1988]               [-0.0272]           [-0.0272]
[-0.2173]           [-0.2173]               [-0.0266]           [-0.0266]
[-0.2155]           [-0.2155]               [-0.0293]           [-0.0293]
[-0.0763]           [-0.0763]               [-0.0267]           [-0.0267]


 # VaR                                      # VaR
# Using 99% confidence                      # Using 75% Confidence
Assessment          Coherent Extension      Assessment          Coherent Extension
[-0.0335]           [-0.0335]               [-0.0068]           [-0.0068]
[-0.0825]           [-0.0390508581933]      [-0.0116]           [-0.0116]
[-0.0424]           [-0.0424]               [-0.0097]           [-0.0097]
[-0.0403]           [-0.0403]               [-0.0099]           [-0.0099]
[-0.0626]           [-0.0626]               [-0.0086]           [-0.0086]
[-0.0529]           [-0.0529]               [-0.0139]           [-0.0139]
```

Table B.7: Fund EDHEC Randomized Returns

```
# Fund from EDHEC (6 assets used)

# Discretized Values
     i          a          b          c          d          e          s
-0.018896   0.001876   0.028386  -0.088600   0.016496   0.044200   0.009766
-0.005053   0.051040  -0.048306   0.026576  -0.192200   0.013136   0.123000
 0.0091600  0.015943   0.025300   0.006373   0.003480  -0.005510  -0.058700
 0.0095766  0.001406  -0.017563   0.029153  -0.083600   0.050400   0.016723
 0.0137333  0.002750  -0.123700   0.008953   0.061100   0.026330  -0.020103
 0.0188100  0.069100   0.004240  -0.026936  -0.054300  -0.007866   0.042330


# CVaR                                          # CVaR
# Using 99% confidence                          # Using 75% Confidence
Assessment          Coherent Extension          Assessment          Coherent Extension
[-0.07565]          [-0.07565]                  [-0.0152473684211]  [-0.0152473684211]
[-0.16265]          [-0.16265]                  [-0.0415421052632]  [-0.0415421052632]
[-0.0436]           [-0.0436]                   [-0.00408378378378] [-0.00408378378378]
[-0.08055]          [-0.0741946875]             [-0.0142631578947]  [-0.0142631578947]
[-0.1132 ]          [-0.1132]                   [-0.0159789473684]  [-0.0159789473684]
[-0.05375]          [-0.0506040551247]          [-0.0239842105263]  [-0.0239842105263]


# VaR                                           # VaR *
# Using 99% confidence                          # Using 75% Confidence
Assessment          Coherent Extension          Assessment          Coherent Extension
[-0.062602]         [-0.062602]                 [-0.000275]         [-0.000275]
[-0.115999]         [-0.115999]                 [-0.0118]           [-0.0118]
[-0.02115]          [-0.02115]                  [-0.000725]         [-0.000725]
[-0.064907]         [-0.0593463885681]          [-0.0117]           [-0.0117]
[-0.0680079]        [-0.0680079]
[-0.050309]         [-0.0447692150443]

* 4 assets used. R-package could not give reliable VaR estimates for assets 1, 5.
```

Table B.8: Fund EDHEC Returns Discretized Further

```
# Fund from EDHEC (6 assets used)

# Discretized Values
     i           a           b            c           d           e           f
-0.1922     -0.066813   -0.0298      -0.01230667  0.0022      0.0104467   0.0158267
-0.1237     -0.03284    -0.00736667   0.00046667  0.0050333   0.00792     0.0099867
-0.0886     -0.031093   -0.0067      -0.00059333  0.0043467   0.00814     0.0113933
-0.0836     -0.029333   -0.00579333  -0.00069333  0.0035067   0.0083933   0.01076
-0.0587     -0.011427    0.00040667   0.00248667  0.0044733   0.0058267   0.00692
-0.0543     -0.03706    -0.01681333  -0.01128    -0.0044533   2e-05       0.00846

                 g           h           j           k           s
            0.023767    0.029387    0.038493    0.063587    0.1230
            0.0126      0.014867    0.01852     0.03414     0.0611
            0.014207    0.018787    0.023773    0.033       0.0442
            0.01474     0.018707    0.02396     0.034347    0.0504
            0.00818     0.01014     0.01292     0.018967    0.0253
            0.0154      0.02222     0.032247    0.052413    0.0691


# CVaR                                      # CVaR
# Using 99% confidence                      # Using 75% Confidence
Assessment          Coherent Extension      Assessment          Coherent Extension
[-0.16265]          [-0.152148729719]       [-0.041542]         [-0.041542]
[-0.1132]           [-0.0946773866689]      [-0.015978]         [-0.015978]
[-0.07565]          [-0.0702310430901]      [-0.015247]         [-0.015247]
[-0.08055]          [-0.0662659678886]      [-0.014263]         [-0.014263]
[-0.0436]           [-0.0436]               [-0.004083]         [-0.004083]
[-0.05375]          [-0.048793177924]       [-0.023984]         [-0.0238072466178]


# VaR                                       # VaR *
# Using 99% confidence                      # Using 75% Confidence
Assessment          Coherent Extension      Assessment          Coherent Extension
[-0.11599]          [-0.0926023046982]      [-0.000725]         [-0.000725]
[-0.0680]           [-0.0515278721469]      [-0.0117]           [-0.00602119402985]
[-0.0626]           [-0.0429209051679]      [-0.000275]         [-0.000275]
[-0.0649]           [-0.0404945095509]      [-0.0118]           [-0.0058352801704]
[-0.02115]          [-0.02115]
[-0.0503]           [-0.0406058828507]

* 4 assets used. R-package could not give reliable VaR estimates for assets 1 and 5.
```

Appendix C

Density Plots of Data (Fund-wise) [6.3]

Each figure corresponds to one fund and the assets in that fund. This Appendix gives the probability density plot for each asset superimposed by the normal density. Notice that all assets (except the second asset in fund EDHEC) have long tails. Left extreme values are greatest for assets in fund RBF, implying significant downside risks for some of its assets. The simulated fund in Figure C.4 appears close to a normal distribution.

Figure C.1: Fund FSPCX

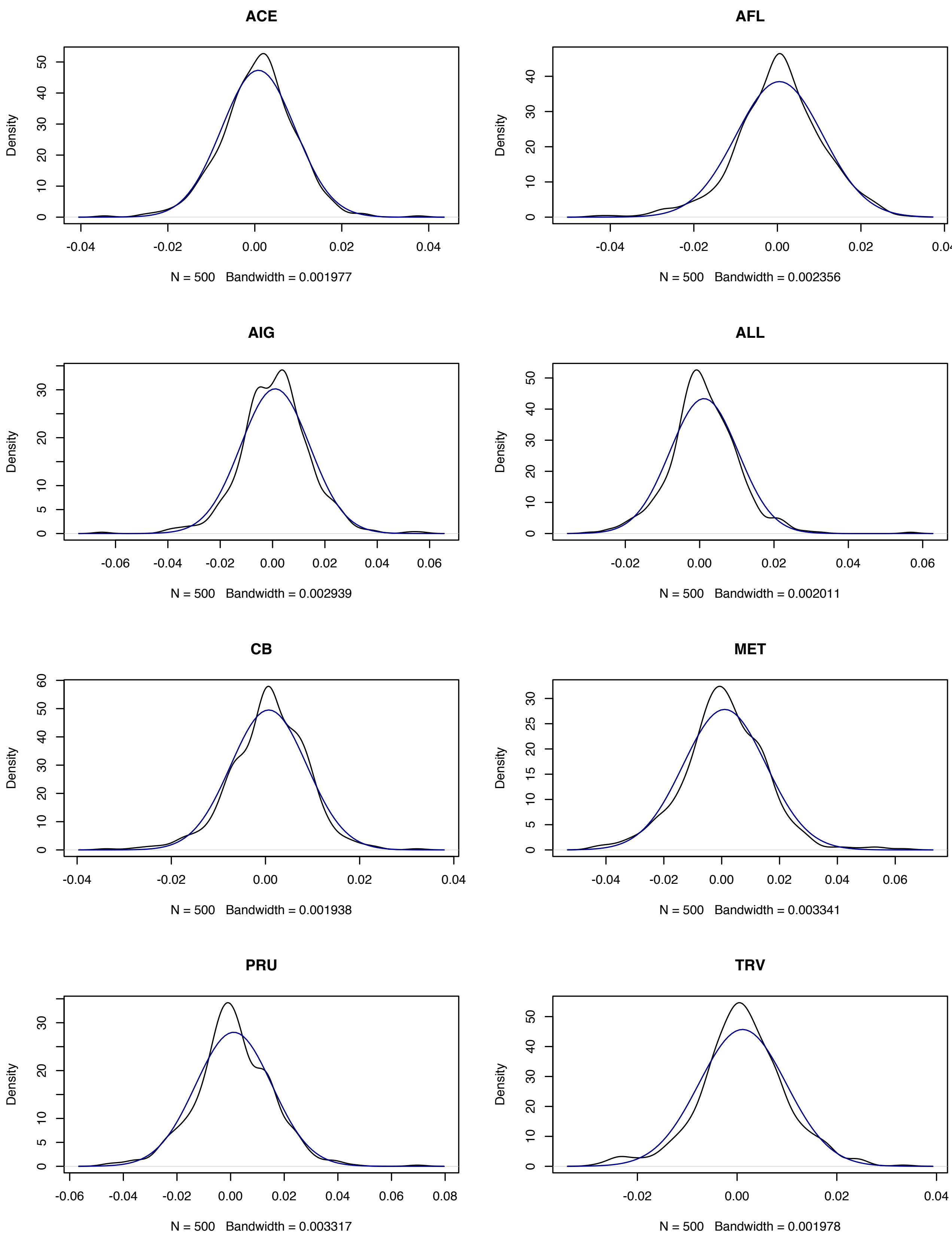


Figure C.2: Fund RBF

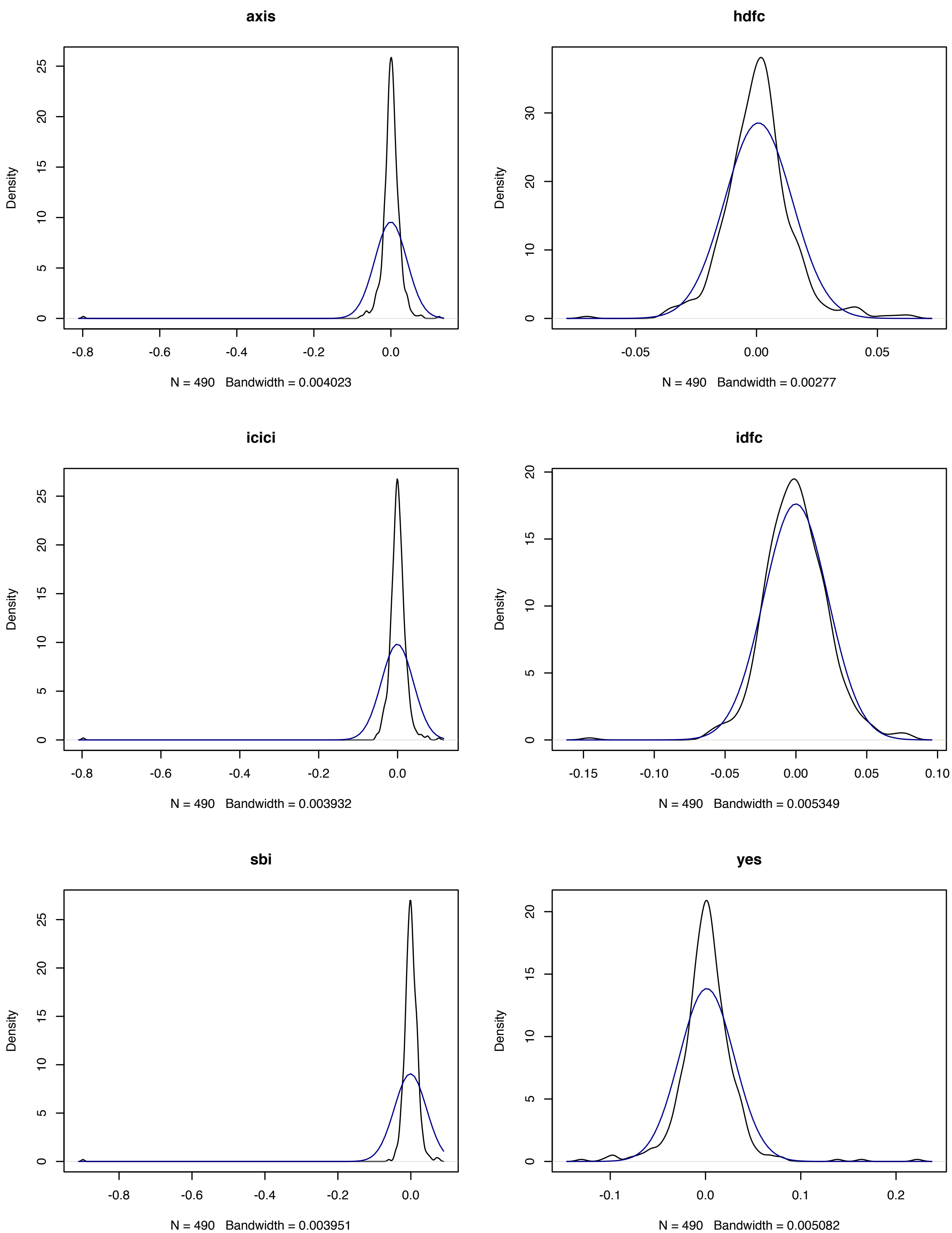


Figure C.3: Fund EDHEC

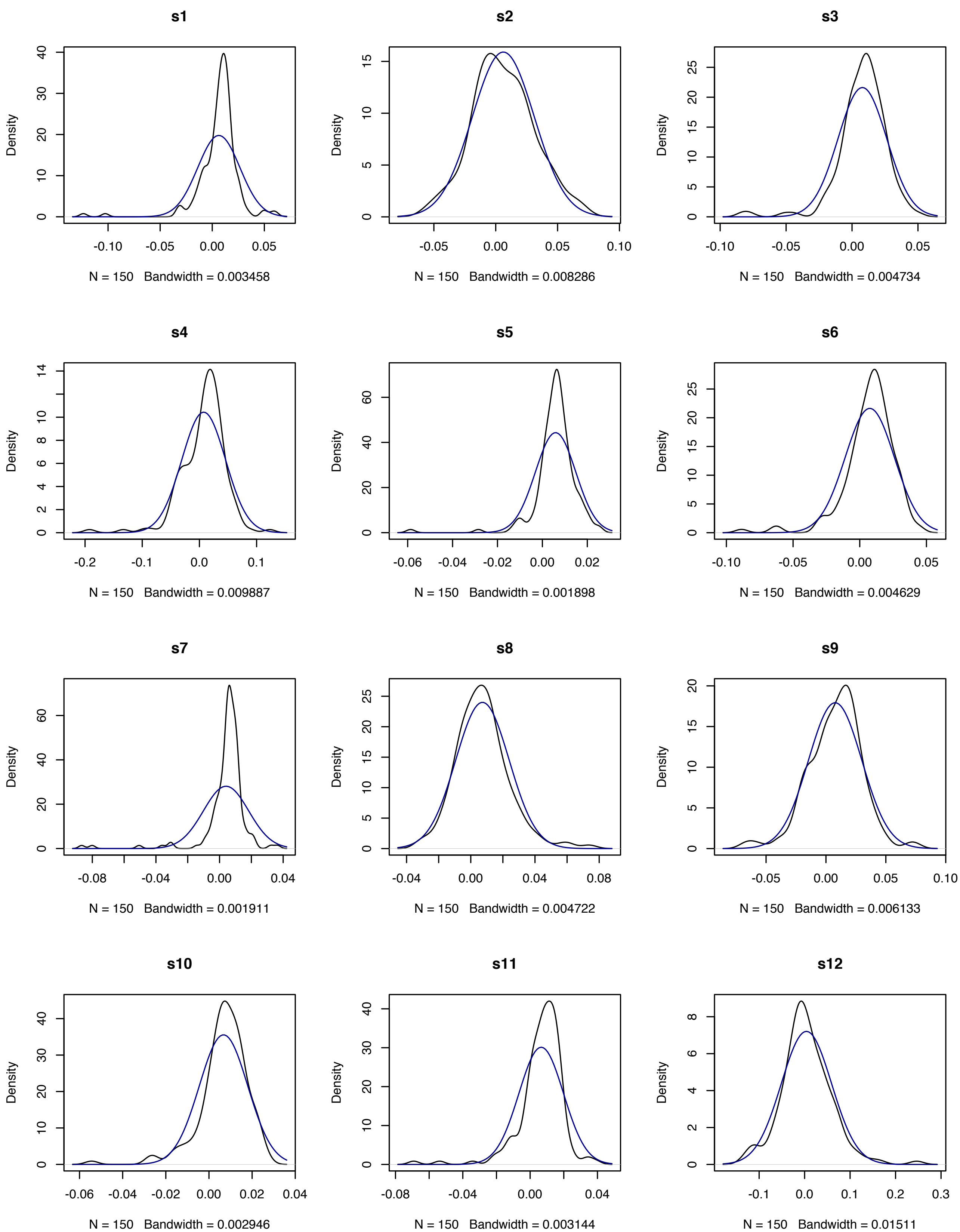
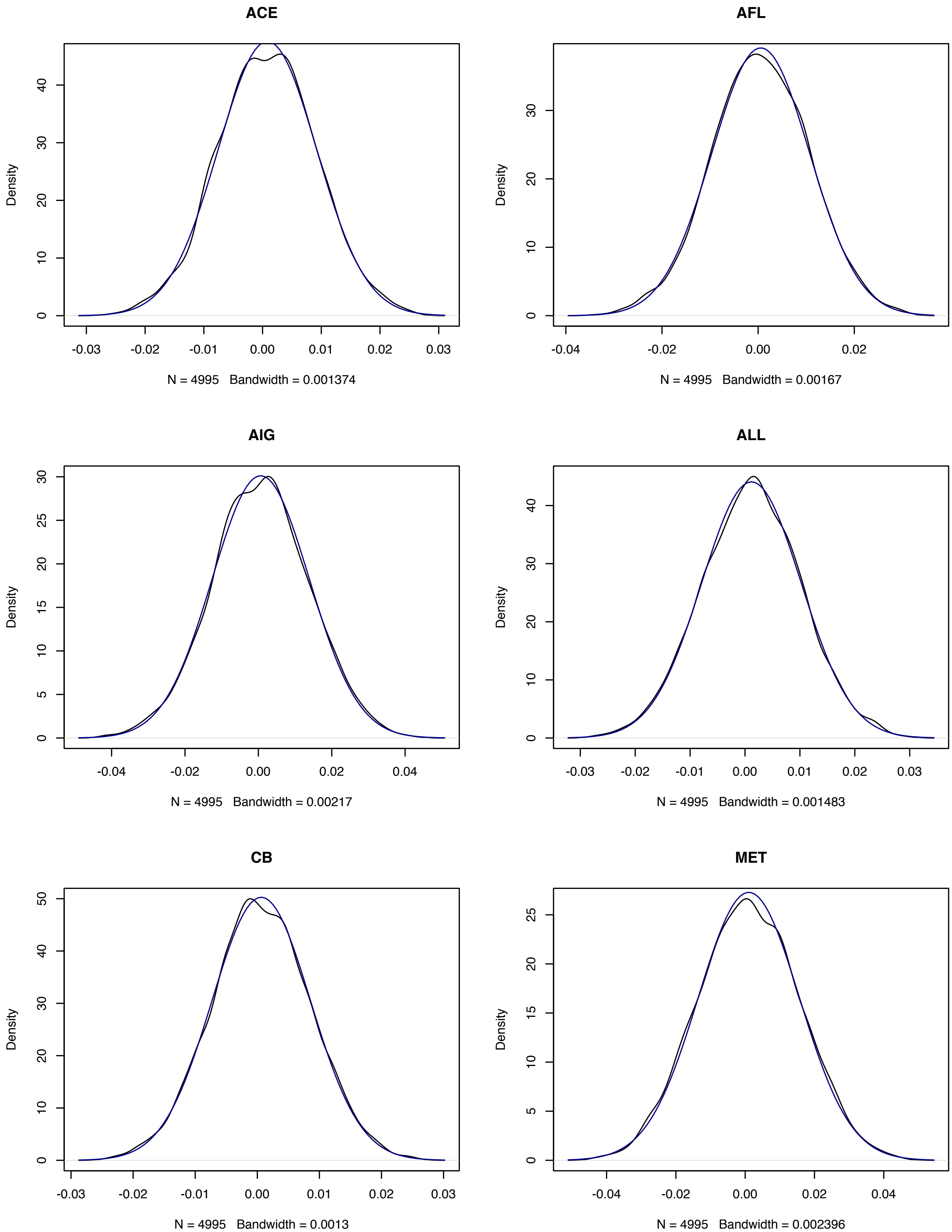


Figure C.4: Fund Simulated FSPCX



Appendix D

Q-Q Plots of Data (Fund-wise) [6.4]

The Q-Q plots indicate that most of the assets have heavy tails. Notice the single left extreme value for three assets in fund RBF. Left extreme values are also observed for assets in fund EDHEC. Asset returns in the simulated fund FSPCX appear normally distributed.

Figure D.1: Fund FSPCX

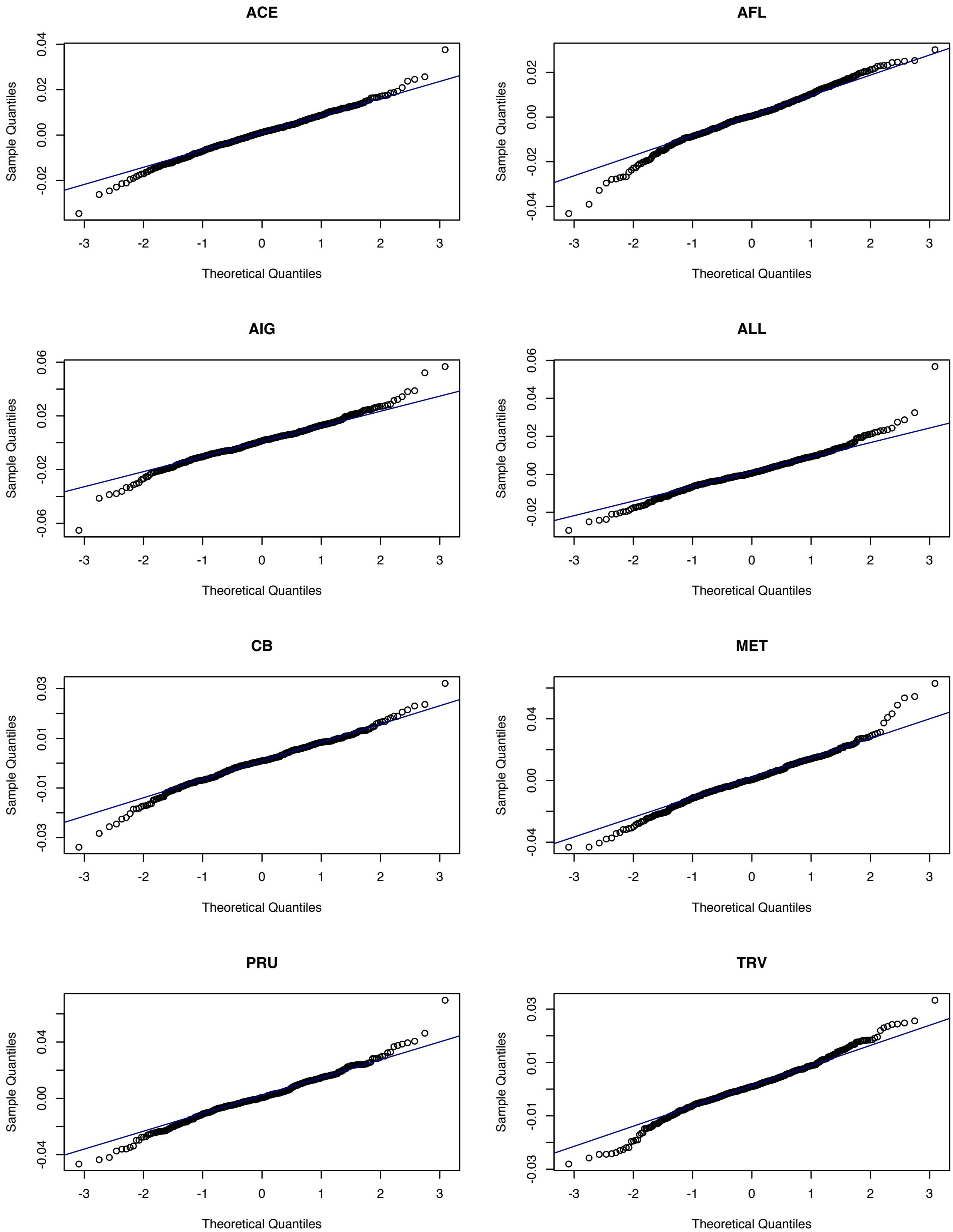


Figure D.2: Fund RBF

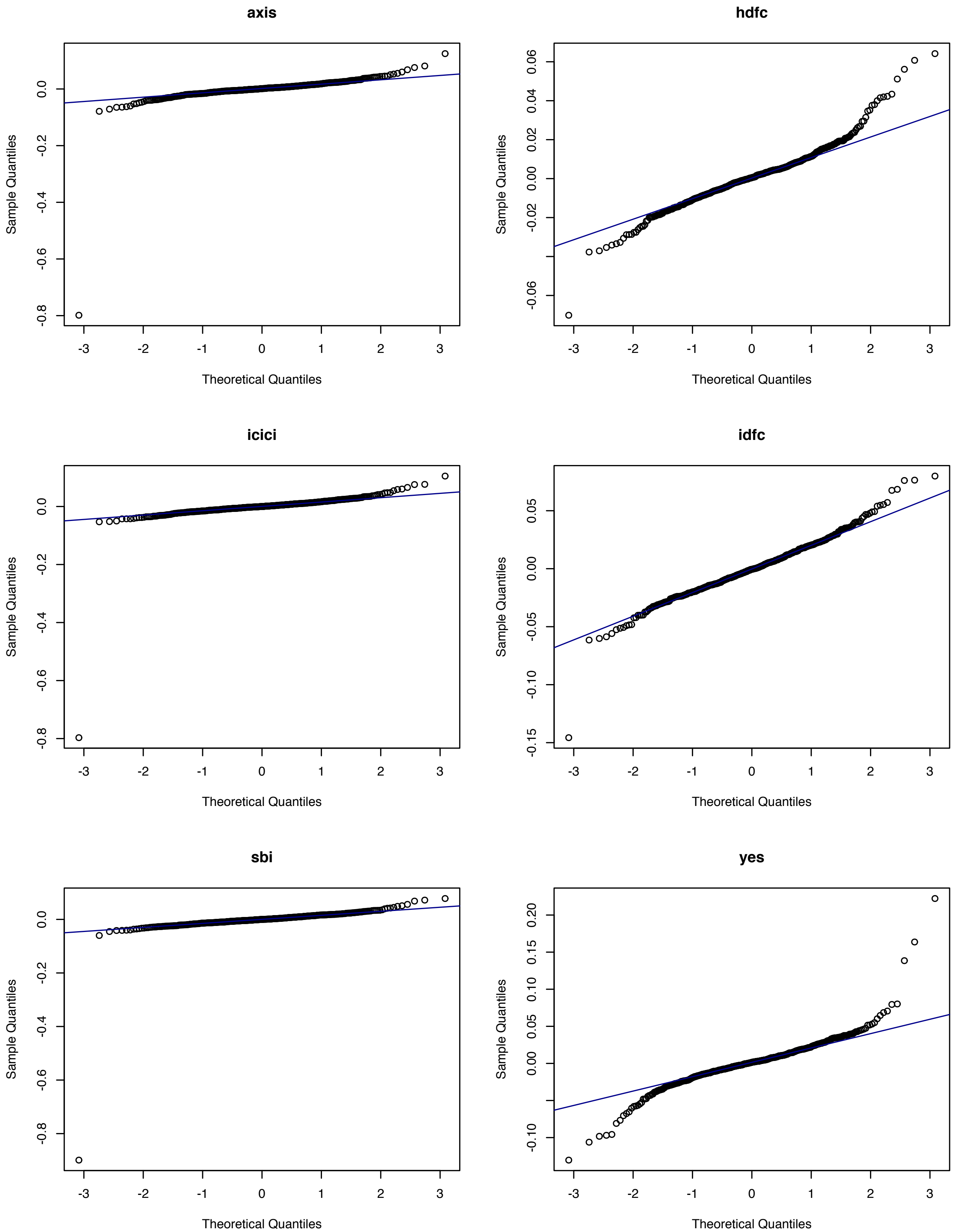


Figure D.3: Fund EDHEC

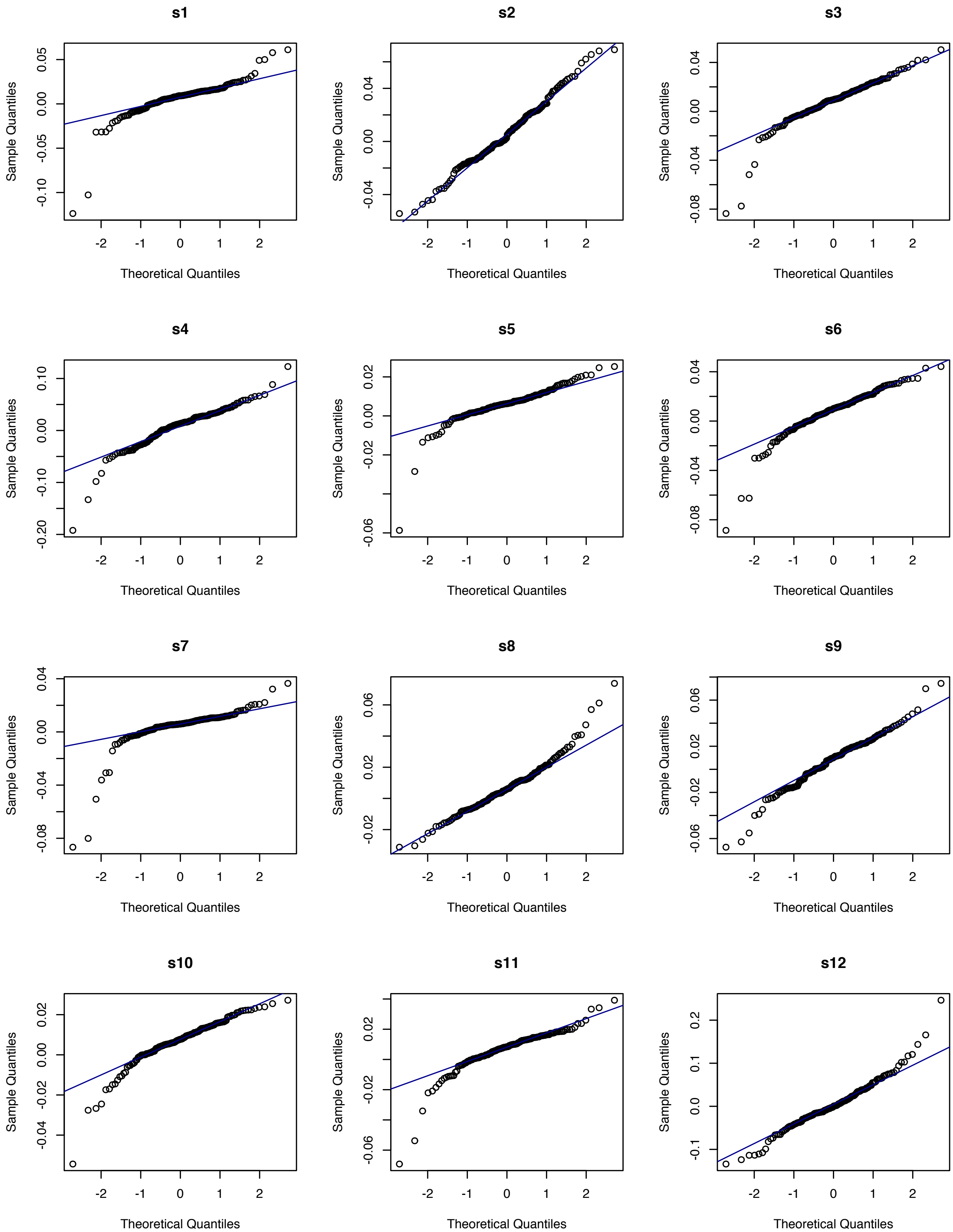
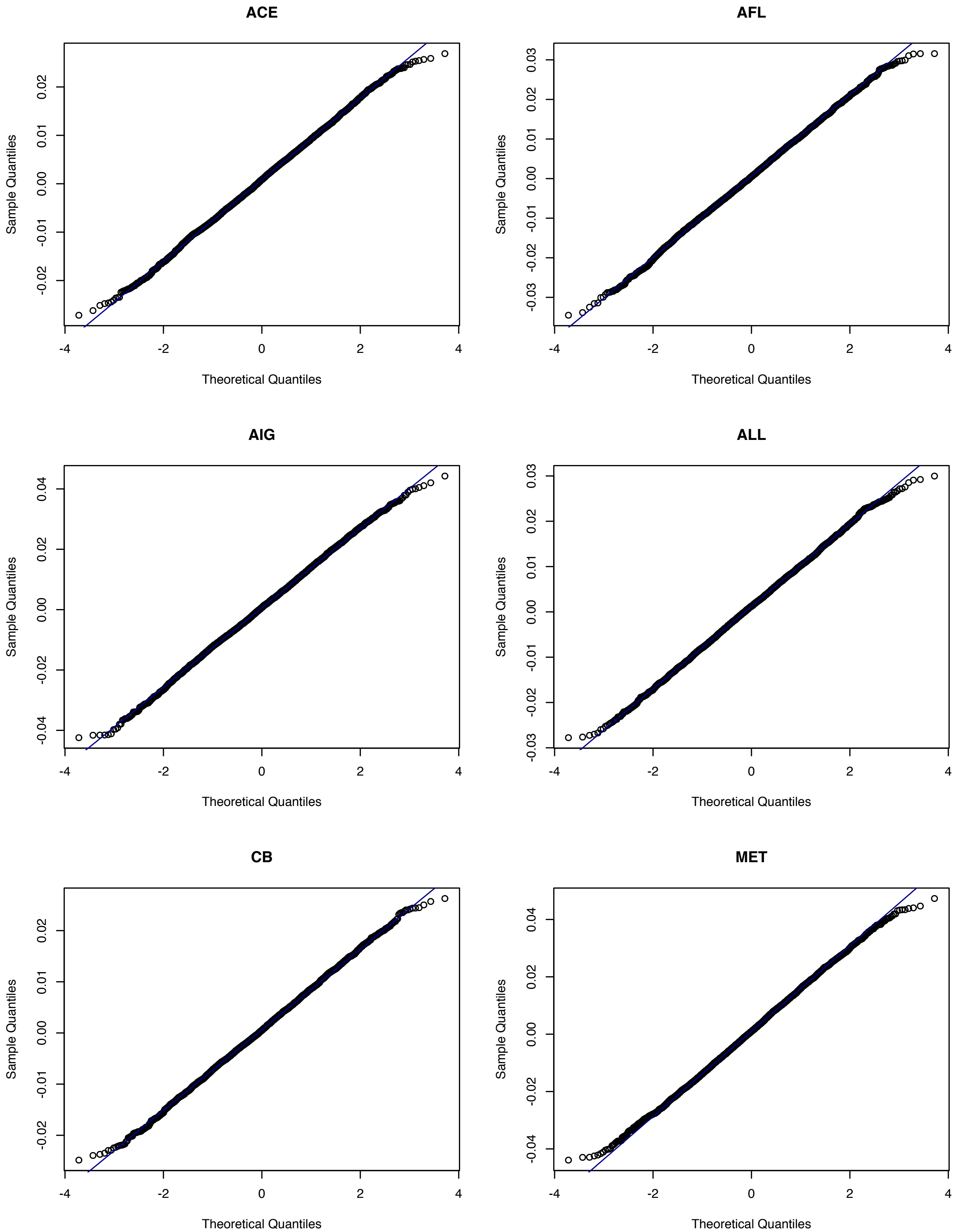


Figure D.4: Fund Simulated FSPCX



Appendix E

Tests for Normality: Fund-wise [6.4]

The kurtosis measures are large, in particular for assets in the RBF fund. The p-values for the tests of normality are very small indicating that the hypothesis of normality would be strongly rejected. The only exception is for the second asset in fund EDHEC. For the assets in the simulated fund, kurtosis values are close to zero and the hypothesis of normality is not rejected.

Figure E.1: Fund FSPCX

```
# Fund FSPCX –  Is data normal or long/short tailed?

> skewness(ACE)
[1] -0.117908914189545
> skewness(AFL)
[1] -0.363844881868591
> skewness(AIG)
[1] -0.105582673787017
> skewness(ALL)
[1] 0.449228774956291
> skewness(CB)
[1] -0.324257735474136
> skewness(MET)
[1] 0.164245293073127

> kurtosis(ACE)
[1] 1.21760709889597
> kurtosis(AFL)
[1] 1.06141518754044
> kurtosis(AIG)
[1] 2.16509902993908
> kurtosis(ALL)
[1] 2.91195135668584
> kurtosis(CB)
[1] 1.37447302317224
> kurtosis(MET)
[1] 1.47609484020773

# Anderson-Darling normality test
data:  ACE
A = 0.8067202230542, p-value = 0.0366378229463
data:  AFL
A = 1.515074497428, p-value = 0.000658143110713
data:  AIG
A = 1.673987400616, p-value = 0.000267809057916
data:  ALL
A = 2.121395308249, p-value = 2.14097816897e-05
data:  CB
A = 1.635911255269, p-value = 0.000332163656017
data:  MET

# Shapiro-Wilk normality test
data:  ACE
W = 0.9903828778493, p-value = 0.00238423848822
data:  AFL
W = 0.9855354445383, p-value = 7.13575693522e-05
data:  AIG
W = 0.9794896289028, p-value = 1.71770216346e-06
data:  ALL
W = 0.9733243058157, p-value = 6.64901878171e-08
data:  CB
W = 0.9841967161062, p-value = 2.95727146149e-05
data:  MET
W = 0.98279840341, p-value = 1.22149517526e-05
```

Figure E.2: Fund RBF

```
# Fund RBF – Is data normal or long/short tailed?

> skewness(axis)
[1] -14.2826002701802
> skewness(hdfc)
[1] 0.397651094555099
> skewness(icici)
[1] -15.3176468132711
> skewness(idfc)
[1] -0.222746608519844
> skewness(sbi)
[1] -17.2947525932027
> skewness(yes)
[1] 0.915678113553406

> kurtosis(axis)
[1] 272.27546268517
> kurtosis(hdfc)
[1] 3.51856537911607
> kurtosis(icici)
[1] 299.605698967051
> kurtosis(idfc)
[1] 3.78374953410891
> kurtosis(sbi)
[1] 350.939982629864
> kurtosis(yes)
[1] 10.9513725053887

# Anderson-Darling normality test
data:  axis
A = 51.04853788854, p-value < 2.220446049e-16
data:  icici
A = 54.93976074319, p-value < 2.220446049e-16
data:  hdfc
A = 4.808871060396, p-value = 6.4240775803e-12
data:  idfc
A = 1.674137748833, p-value = 0.000267503091013
data:  sbi
A = 67.98392546418, p-value < 2.220446049e-16
data:  yes
A = 9.05984915479, p-value < 2.220446049e-16

# Shapiro-Wilk normality test
data:  axis
W = 0.3683911776967, p-value < 2.220446049e-16
data:  hdfc
W = 0.9512424299299, p-value = 1.23915399443e-11
data:  icici
W = 0.3319717822556, p-value < 2.220446049e-16
data:  idfc
W = 0.9669193444041, p-value = 4.65026790486e-09
data:  sbi
W = 0.2597859761881, p-value < 2.220446049e-16
data:  yes
W = 0.8828068692471, p-value < 2.220446049e-16
```

Figure E.3: Fund EDHEC

```
# Fund EDHEC – Is data normal or long/short tailed?

> skewness(s1)
[1] -2.63271058726335
> skewness(s2)
[1] 0.159921475321309
> skewness(s3)
[1] -1.63851253897242
> skewness(s4)
[1] -1.29047653580327
> skewness(s5)
[1] -2.74737576659384
> skewness(s6)
[1] -1.68659446900765

> kurtosis(s1)
[1] 15.6750664163205
> kurtosis(s2)
[1] -0.106169825512085
> kurtosis(s3)
[1] 6.19913480167745
> kurtosis(s4)
[1] 5.21371321050983
> kurtosis(s5)
[1] 17.3177737079766
> kurtosis(s6)
[1] 5.9025821120063

# Anderson-Darling normality test
data:  ss1
A = 7.265134989838, p-value < 2.220446049e-16
data:  ss2
A = 0.4268835482754, p-value = 0.309654245102
data:  ss3
A = 2.584101499147, p-value = 1.50201833147e-06
data:  ss4
A = 2.44662205614, p-value = 3.2631305791e-06
data:  ss5
A = 4.652043665219, p-value = 1.39768694501e-11
data:  ss6
A = 2.734308184904, p-value = 6.43983059304e-07

# Shapiro-Wilk normality test
data:  ss1
W = 0.764409635175, p-value = 3.04357926527e-14
data:  ss2
W = 0.9907352666498, p-value = 0.432396492796
data:  ss3
W = 0.888269344426, p-value = 3.03237847088e-09
data:  ss4
W = 0.9144430181976, p-value = 9.44885503842e-08
data:  ss5
W = 0.8030300564608, p-value = 6.22468404336e-13
data:  ss6
W = 0.8888061116777, p-value = 3.23669825154e-09
```

Figure E.4: Fund Simulated FSPCX

```
> Simulating Fund FSPCX -- Is data normal or long/short tailed?
> # simulating FSPCX data from normal distributions

> skewness(ACE)
[1] -0.00675450815635633
> skewness(AFL)
[1] -0.0288391969809926
> skewness(AIG)
[1] -0.016561497499291
> skewness(ALL)
[1] -0.0154242609617082
> skewness(CB)
[1] 0.00568648301864284
> skewness(MET)
[1] 0.0108204947222428

> kurtosis(ACE)
[1] -0.0732340184830282
> kurtosis(AFL)
[1] -0.0285865619298291
> kurtosis(AIG)
[1] -0.0339827182677856
> kurtosis(ALL)
[1] -0.0707145969244327
> kurtosis(CB)
[1] -0.063905967475419
> kurtosis(MET)
[1] -0.209180579558843

# Anderson-Darling normality test
data:  ACE
A = 0.3723183438024, p-value = 0.420191825882
data:  AFL
A = 0.2875042335754, p-value = 0.620160046993
data:  AIG
A = 0.3638132188362, p-value = 0.439550607606
data:  ALL
A = 0.24349035874, p-value = 0.765585968586
data:  CB
A = 0.2786560222703, p-value = 0.648793271656
data:  MET
A = 0.5056459621637, p-value = 0.202052929833

# Shapiro-Wilk normality test
data:  ACE
W = 0.9994678821204, p-value = 0.169116932838
data:  AFL
W = 0.9994916797907, p-value = 0.201123906118
data:  AIG
W = 0.9995467426938, p-value = 0.295705534682
data:  ALL
W = 0.999553187232, p-value = 0.308812598368
data:  CB
W = 0.9995912911258, p-value = 0.395372321368
data:  MET
W = 0.9992527454127, p-value = 0.0322619164606
```

Appendix F

Python Code and R-script [6.3]

The input file for the Python code *improb* (Table F.1) contains the following data elements. The seven discretized values of the returns gamble-wise, and assessed lower prevision (which may be VaR or $CVaR$) again gamble-wise. The possibility space is $X = \{i, a, b, c, d, e, s\}$. The rows correspond to the gambles. The assessed lower prevision values for each gamble are entered in the last column appearing after the discretized values are entered. This is the same format as the examples given in Chapter 5.

The python code shown below first checks for consistency and returns a true or false. If true then coherence is checked and when corrected if necessary, it returns a true. The Python output file is printed in the same format, but will now replace the assessed lower prevision that was in the last column with the now corrected lower prevision.

In discretizing the returns data (Table F.2), five values are obtained to which we add the minimum and maximum returns. So there are seven values.

In simulating normally distributed data for the assets in a fund (Table F.3), mean and standard deviation of each asset is calculated from returns data, then 5000 random values are generated. The data is then used to find the risk measures as well as the discretized values, after which the risk measures are checked for coherence.

Table F.1: Input File for *improb* (in Python)

```
# Fund FSPCX
# Checks for almost sure loss and coherence of assessments of lower previsions (6 assets used)

from improb import PSpace, Event, Gamble
from improb.lowprev.lowpoly import LowPoly

# Using CVaR measures

# Uses 99% confidence
lpr = LowPoly(pspace= "iabcdes", number_type = "float")
lpr.set_lower({"i": -0.0345690049510583, "a": -0.0110457103299400, "b": -0.00311056872490512,
lpr.set_lower({"i": -0.0431856429522125, "a": -0.0140600313282144, "b": -0.00409161821020374,
lpr.set_lower({"i": -0.0652468454267962, "a": -0.0168151250779751, "b": -0.00536133237037656,
lpr.set_lower({"i": -0.0295249767342611, "a": -0.0110576512039879, "b": -0.00298562927585648,
lpr.set_lower({"i": -0.0338507623929602, "a": -0.0107621250372860, "b": -0.00279896338200148,
lpr.set_lower({"i": -0.0432892303022039, "a": -0.0187375939382935, "b": -0.00534912522483403,

"c": 0.00110317474758187344,
"c": 0.000610452617346244,
"c": 0.00121562968872801З,
"c": 0.000672703278054171,
"c": 0.000912853838716419,
"c": 0.000781329951828878,

"d": 0.00490758684743495, "e": 0.0122519528312329, "s": 0.0376086159280909}, "-0.0259386762439764")
"d": 0.00533382917747575, "e": 0.0146132549919804, "s": 0.0301556734502639}, "-0.0344908645873863")
"d": 0.00697809632431584, "e": 0.0189931118794075, "s": 0.0567291319558037}, "-0.0438486004524433")
"d": 0.00528959401882367, "e": 0.0138922036684100, "s": 0.0567102422247075}, "-0.0247096666322696")
"d": 0.00489305742715011, "e": 0.0112604754417573, "s": 0.0321256106756230}, "-0.0269576254314673")
"d": 0.00795664000824624, "e": 0.0204831839090866, "s": 0.0629905090929420}, "-0.0404791011206761")

# Print values of gambles and their assessments of lower previsions
print(lpr)

# Do the assessments avoid sure loss? (T/F)
print(lpr.is_avoiding_sure_loss())

# Are the assessments coherent? (T/F)
print(lpr.is_coherent())

# Get coherent extensions
lpr2 = lpr.get_coherent()

# Print results; are they coherent? (T/F)
print(lpr2)
print(lpr2.is_coherent())
```

Table F.2: Discretizing Returns (in R)

```
# Uses fund FSPCX data, checks normality density etc.,
# finds CVaR and VaR risk measures (lower previsions),
# and gamble rewards discretized into five finite states

library(timeSeries)
options(digits=15)

load(file = "R.RData") # File containing the returns data in time series
dim(R)

S = R[4:503,1:8]
dim(S)
attach(S)

# Calculating gamble (asset) returns discretized from a continuous possibility space

# ACE
a1 = sort(as.vector(S[,1]))
head(a1)
min(a1)
tail(a1)
max(a1)
length(a1)
n = length(a1)/5  # To divide the data into 5 parts, each with prob 1/5
n
cp1 = mean(a1[1:n])
cp2 = mean(a1[(n+1):(2*n)])
cp3 = mean(a1[((2*n)+1):(3*n)])
cp4 = mean(a1[((3*n)+1):(4*n)])
cp5 = mean(a1[((4*n)+1):(5*n)])

inf.ace = min(a1)
cp1.ace = cp1
cp2.ace = cp2
cp3.ace = cp3
cp4.ace = cp4
cp5.ace = cp5
sup.ace = max(a1)

# Discretized values for each gamble
ace = c(inf.ace, cp1.ace, cp2.ace, cp3.ace, cp4.ace, cp5.ace, sup.ace)

# Similarly other gamble returns are discretized

# gamble = rbind(ace,afl,aig,all,cb,met)
# gamble
```

Table F.3: Simulating Fund FSPCX (in R)

```
# Simulating FSPCX data from normal distributions
# uses fund FSPCX data, checks normality density etc. and
# finds lower previsions, and gamble rewards
# discretized into five finite states

library(timeSeries)


options(digits=15)

load(file = "R.RData")
dim(R)

S = R[4:503,1:8]
head(S)
tail(S)
dim(S)

# Simulate data
set.seed(100)

ACE1 = rnorm(5000, mean=mean(S[,1]), sd=sd(S[,1]))
AFL1 = rnorm(5000, mean=mean(S[,2]), sd=sd(S[,2]))
AIG1 = rnorm(5000, mean=mean(S[,3]), sd=sd(S[,3]))
ALL1 = rnorm(5000, mean=mean(S[,4]), sd=sd(S[,4]))
CB1 = rnorm(5000, mean=mean(S[,5]), sd=sd(S[,5]))
MET1 = rnorm(5000, mean=mean(S[,6]), sd=sd(S[,6]))

ACE1 = sort(ACE1)
AFL1 = sort(AFL1)
AIG1 = sort(AIG1)
ALL1 = sort(ALL1)
CB1 = sort(CB1)
MET1 = sort(MET1)

# Removing outliers

ACE = ACE1[3:4997]
AFL = AFL1[3:4997]
AIG = AIG1[3:4997]
ALL = ALL1[3:4997]
CB = CB1[3:4997]
MET = MET1[3:4997]

length(ACE)
```

Appendix G

Volatility (Fund-wise) [6.4]

The standard deviation of daily return of each stock and the corresponding time period is given here fund-wise. Fund RBF has the highest standard deviation and needs larger correction for coherence when compared to FSPCX. Fund EDHEC data shown here is after converting from monthly returns to daily returns. EDHEC is an index of funds (fund of funds) therefore its standard deviation is lower. FSPCX has a standard deviation approximately twice that EDHEC. RBF has a standard deviation approximately three times that of FSPCX.

Figure G.1: Volatility of Funds

```
# Standard Deviation of Daily Returns

1. Fund RBF (2013–2014)

        axis     hdfc    icici     idfc      sbi      yes
    0.041676 0.013974 0.040586 0.022640 0.043995 0.028815

Standard Deviation of the Portfolio = 0.033808


2. Fund FSPCX (2013–2014)

         ACE      AFL      AIG      ALL       CB      MET
    0.008429 0.010362 0.013208 0.009198 0.008055 0.014334

Standard Deviation of the Portfolio = 0.01086


3. Fund EDHEC (1997–2009)

    Convertible Arbitrage   CTA Global  Distressed Securities   Emerging Markets
                 0.004402     0.005473               0.004026           0.008339

    Equity Market Neutral   Event Driven
                 0.001963       0.004402

Standard Deviation of the Portfolio = 0.005085
```