

Coherent risk measures induced by partially specified probabilities

Ehud Lehrer*

December 6, 2007

A preliminary draft

ABSTRACT:

Partially specified probabilities induce coherent risk measures of a special kind. These measures are axiomatized using the four axioms that characterize coherent risk measures and an additional one, which requires that the risk measure be additive on the set of efficient portfolios.

Journal of Economic Literature classification numbers: C61, C72, D81, D82, D83

Keywords: partially specified probability, coherent risk measures, efficient portfolio

*School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel and INSEAD, Bd. de Constance, 77305 Fontainebleau Cedex, France . e-mail: lehrer@post.tau.ac.il. The author acknowledges the support of the Israel Science Foundation, Grant #762/045.

1 Introduction

Recently there has been considerable interest in how to measure the risk of portfolios, which was motivated in part by regulatory rules imposed on the financial sector. Classical risk measures, such as the variance and the Value at Risk, suffer from a few weaknesses. The variance, for instance, considers equally positive and negative deviations from the mean. Value at Risk on the other hand, might penalize diversification. That is, the risk might reduce due to a split of a portfolio into two parts.

Artzner et al. (1999) proposed an alternative approach to measuring risk by requiring that a risk measure have four plausible properties: monotonicity, homogeneity, sub-additivity and translation invariance. The latter means that if a sure asset is added to a portfolio, the risk of the combined portfolio is the risk of the original one minus the value of the sure asset. The risk measures that satisfy these axioms are called coherent.

Any coherent risk measure can be represented as the (negative value of) the maximum of the expectations over a convex set of probability distributions (see Delbaen (2002) and Gilboa and Schmeidler (1989) whose model of decisions under uncertainty is strongly related to coherent risk measures). This paper introduces a specific family of natural coherent risk measures, which is characterized with the help of an additional axiom.

Markowitz (1952) introduced the efficient frontier that consists of those portfolios that minimize the risk given a fixed return level, or, alternatively, of those that maximize the return given a fixed risk level. Here, however, there is no numerical index that measures the return. The partial order that plays a role in the monotonicity axiom replaces it. A portfolio X is said to dominate Y (denoted, $X \geq Y$) if at any state of the world X yields a payoff which is as high as that yielded by Y . Being dominated by X is analogous to having a lower growth rate than X .

Borrowing from Markowitz, we say that X is efficient if for any portfolio Y dominated by X , either the risk of X is strictly smaller than that of Y , or X is strictly greater than Y on an insignificant set of states (meaning, for instance, a set whose probability is zero).

The additional axiom, called additivity with respect to efficient portfolios, requires that the risk measure be additive whenever efficient portfolios are involved. That is, when two efficient portfolios are merged, the risk of the resulting portfolio is exactly the sum of the risks of the ingredients. The intuition of this axiom is that sub-additivity

of a risk measure is a consequence of a synergy between two portfolios: diversification reduces risk. However, when two portfolios are already efficient, there is no room for a further improvement in the risk. In other words, diversifying over efficient portfolios does not reduce risk.

It turns out that additivity with respect to efficient portfolios along with the other four axioms that correspond to coherent risk measures characterize a family of measures induced by partially specified probabilities.

Lehrer (2005b) introduced and axiomatized the notion of partially specified probability (PSP), which is a pair $(\mathbb{P}, \mathcal{Y})$, where \mathbb{P} is a probability measure and \mathcal{Y} is a set of random variables. The distribution \mathbb{P} is not fully specified. Rather, only the expectation of every variable in \mathcal{Y} with respect to \mathbb{P} is known. Typically, the algebra generated by \mathcal{Y} does not cover the entire space of random variables. This means that $(\mathbb{P}, \mathcal{Y})$ defines an affine space of probability distributions that are consistent with data available.

The risk measure induced by the partially specified probability $(\mathbb{P}, \mathcal{Y})$ is the (negative value of the) maximum of the expectations (w.r.t. \mathbb{P}) over all the variables in the algebra generated by \mathcal{Y} that are dominated by the portfolio under consideration. A possible interpretation of this measure is that there is a solid data about the risk of the portfolios in \mathcal{Y} . The risk measure coincides with the expectation w.r.t. \mathbb{P} on the portfolios in \mathcal{Y} . Using this information the risk measure is extended to all possible portfolios. The risk of an unknown portfolio, say X , is defined by approximating X with portfolios in the algebra generated by \mathcal{Y} that are dominated by it.

The paper is built as follows. Section 2 presents the model of coherent risk measures. For simplicity and brevity, I confine the discussion to finite state spaces. Section 3 explains what PSP is and what is the risk measure induced by it. Section 4 defines efficient portfolios and introduces the axiom of additivity w.r.t. efficient portfolios. The last section provides the representation theorem.

2 Coherent risk measures

Let S be a finite space and denote by \mathcal{X} the set of all bounded functions over S . Members of \mathcal{X} will often be referred to as portfolios.

Definition 1 (Artzner et al) A mapping $\varrho : \mathcal{X} \rightarrow \mathbb{R}$ is called a coherent risk measure if it is:

- *Monotone (M)*: if $X \leq Y$, then $\varrho(X) \geq \varrho(Y)$ for every $X, Y \in \mathcal{X}$;
- *Sub-additive (S-ADD)*: $\varrho(X + Y) \leq \varrho(X) + \varrho(Y)$ for every $X, Y \in \mathcal{X}$;
- *Homogeneous (H)*: $\varrho(cX) = c\varrho(X)$ for every $X \in \mathcal{X}$ and constant $c > 0$; and
- *Translation Invariant (TI)*: $\varrho(X + \mathbf{c}) = \varrho(X) - \varrho(\mathbf{c})$ for every constant $c \geq 0$ and every $X \in \mathcal{X}$, where \mathbf{c} is the sure asset (constant variable) corresponding to c .

These requirements rule out some of the classical measures of risk traditionally used in finance such as those that are based on second moments. The latter are ruled out due to the violation of the monotonicity requirement. Quantile-based measures like the value-at-risk are also ruled out due to lack of subadditivity.

3 Partially specified probabilities

3.1 Two variations on a theme by Ellsberg

Example 1 An urn contains 90 balls, 30 are Red, 30 are Black, and 30 are White. This information is equivalent to saying that the expectation of the random variables $[1(R), 0(B), 0(W)]$, $[0(R), 1(B), 0(W)]$ and $[0(R), 0(B), 1(W)]$ is $\frac{1}{3}$. Moreover, since these three random variable span all other variables defined over the state space $\{R, B, W\}$, one can calculate the expectation of every random variable.

Suppose now that it is known that 30 balls are Red and the others are either Black or White, but there is no indication as to how the Black and White balls are distributed; the distribution of colors is partially specified. The probability of only one non-trivial event is known: the probability of Red is $\frac{1}{3}$. Equivalently, the expectation of the random variable $[1(R), 0(B), 0(W)]$ is $\frac{1}{3}$, but the expectations of the two random variables $[0(R), 1(B), 0(W)]$ and $[0(R), 0(B), 1(W)]$ are unknown. Observe that the expectation of the random variable $[1(R), 1(B), 1(W)]$ is known as well — it is 1.

Suppose that, as before, there are 30 Red balls, 60 balls which are either Black or White, and an unknown number of Green balls. In addition, it is known that the number

of Green balls is the same as the number of White balls. The probability of Red is no longer $\frac{1}{3}$. Furthermore, the probability of no non-trivial event is known. Nevertheless, some information about the distribution of colors is available. It turns out that this information is given by the expectation of two random variables.

Denote by X the random variable $[0(R), 0(B), 1(W), -1(G)]$ that takes the value 0 on Red and Black, 1 on White, and -1 on Green. Since the probabilities of White and Green are equal, the expectation of X is 0. Furthermore, let Y be the random variable $[1(R), 0(B), 0(W), \frac{1}{3}(G)]$. It turns out¹ that the expectation of Y is $\frac{1}{3}$.

3.2 Partially specified probabilities

A *partially-specified probability* (PSP) (see Lehrer (2005) for an axiomatic approach to PSP) over S is a pair $(\mathbb{P}, \mathcal{Y})$, where \mathcal{Y} is a set of real functions over S , \mathcal{Y} contains $\mathbb{1}_S$, and \mathbb{P} is a probability over S .

Let Z be a bounded function defined over S and let $(\mathbb{P}, \mathcal{Y})$ be a partially-specified probability. Denote,

$$\int Zd(\mathbb{P}, \mathcal{Y}) = \max \left\{ \sum_{Y \in \mathcal{Y}} \lambda_Y \mathbb{E}_{\mathbb{P}}(Y); \sum_{Y \in \mathcal{Y}} \lambda_Y Y \leq Z \text{ and } \lambda_Y \in \mathbb{R} \text{ for every } Y \in \mathcal{Y} \right\}, \quad (1)$$

where $\mathbb{E}_{\mathbb{P}}(Y)$ is the expectation of Y w.r.t. \mathbb{P} . The decision maker knows $\mathbb{E}_{\mathbb{P}}(Y)$ for every $Y \in \mathcal{Y}$ and he can therefore calculate the expectation of any function of the form $\sum_{Y \in \mathcal{Y}} \lambda_Y Y$. The integral of Z is the maximal $\sum_{Y \in \mathcal{Y}} \lambda_Y \mathbb{E}_{\mathbb{P}}(Y)$ among all those functions of the form $\sum_{Y \in \mathcal{Y}} \lambda_Y Y$ that are below Z .

Denote by $\text{span}(\mathcal{Y})$ the algebra generated by \mathcal{Y} . Clearly, when the expectation w.r.t. \mathbb{P} of every portfolio in \mathcal{Y} is known, so is the expectation w.r.t. \mathbb{P} of every portfolio in $\text{span}(\mathcal{Y})$. Note that $\int \cdot d(\mathbb{P}, \mathcal{Y}) = \int \cdot d(\mathbb{P}, \text{span}(\mathcal{Y}))$.

3.3 Coherent risk measure induced by a partially specified probability

Let $(\mathbb{P}, \mathcal{Y})$ be a partially specified probability.

¹Denote by n_i the number of balls of color i . Then $\frac{n_r}{n_r+n_b+n_w} = \frac{1}{3} = \frac{\frac{1}{3}n_g}{n_g}$ and therefore, $\frac{n_r+\frac{1}{3}n_g}{n_r+n_b+n_w+n_g} = \frac{1}{3}$.

Definition 2 *The risk measure induced by $(\mathbb{P}, \mathcal{Y})$ is defined as*

$$\varrho_{(\mathbb{P}, \mathcal{Y})}(X) = - \int X d(\mathbb{P}, \mathcal{Y}) \quad (2)$$

for every $X \in \mathcal{X}$.

Lemma 1 *The mapping $\varrho_{(\mathbb{P}, \mathcal{Y})}(X)$ is a coherent risk measure.*

We say that X is ϱ -more risky (or simply more risky) than Y if $\varrho(X) > \varrho(Y)$ and they are ϱ -risk-equivalent if $\varrho(X') = \varrho(X)$. Define,

$$\mathcal{P}(p, \mathcal{Y}) = \{q; q \sim_{\mathcal{Y}} p\}. \quad (3)$$

Any coherent risk measure can be represented as the (negative value of) the maximum of the expectations over a convex set of probability distributions (see Delbaen (2002)). Here, $\varrho_{(\mathbb{P}, \mathcal{Y})}(X)$ can be represented as,

$$\varrho_{(\mathbb{P}, \mathcal{Y})}(X) = - \max_{\mathbb{Q} \in \mathcal{P}(p, \mathcal{Y})} \mathbb{E}_{\mathbb{Q}}(X). \quad (4)$$

That is, $\varrho_{(\mathbb{P}, \mathcal{Y})}$ is obtained by considering the set of all probability distributions that agree with \mathbb{P} on the random variables in \mathcal{Y} . This is an affine set of probability distributions (i.e. the intersection of an affine space with the set of all probability distributions). In other words, while a coherent risk measure is characterized by a (convex) set of probability distributions, when it is induced by PSP, this set has a more restrictive structure: it is an affine set.

The objective of this paper is to provide an additional property of coherent risk measures that characterizes those that are induced by partially specified probabilities.

4 Efficient portfolios

We say that $X \geq Y$ if $X(s) \geq Y(s)$ for every $s \in S$. Thus, there are two orders defined over portfolios: \geq induces a partial order and ‘being more risky w.r.t. ϱ ’ induces a complete order. For every portfolio X let² $\text{LE}(X)$ be the set of all portfolios X' that are both, risk-equivalent to X and $X \geq X'$.

²The symbol LE signifies that the set consists of those portfolios that are Lower than and Equivalent to X in the partial order and the complete order senses, respectively.

In what follows the notion of efficient portfolio is defined. When X is efficient and another portfolio Y such that $X \geq Y$ challenges its efficiency, either $\varrho(Y) > \varrho(X)$, or they share the same risk level. In the latter case, Y is smaller than X on an insignificant set of states (e.g., of probability zero). In this case it does not matter how far on this specific set a portfolio is below X , it is still risk-equivalent to X . That is, the portfolio $X + c(X - Y)$, which is far below X when c is very large is, like Y itself, risk-equivalent to X .

Definition 3 *A portfolio X is ϱ -efficient (or simply efficient), if for every portfolio Y , $X \geq Y$ either (i) $\varrho(Y) > \varrho(X)$ or (ii) $\varrho(X) = \varrho(Y)$ and $\varrho(X) = \varrho(X - c(X - Y))$ for every constant c .*

In other words, a portfolio X is efficient if, for every $Y \in \text{LE}(X)$ and for every positive constant c , the portfolio $X - c(X - Y)$ is also in $\text{LE}(X)$. This suggests that the set of states on which Y is smaller than X is insignificant and the gap between Y and X over this set can be expanded without increasing the risk measure of the resulting portfolio.

Example 2 Consider the urn that contains balls of four colors as in Example 1. There, \mathcal{Y} consists of the variables X , Y and $\mathbb{1}_S$. $\varrho_{(\mathbb{P}, \mathcal{Y})}(X) = 0$, $\varrho_{(\mathbb{P}, \mathcal{Y})}(Y) = -\frac{1}{3}$ and $\varrho_{(\mathbb{P}, \mathcal{Y})}(\mathbb{1}_S) = -1$. Consider $Z = [1(R), 0(B), 1(W), -\frac{2}{3}(G)]$. Since $Z = X + Y$, $\varrho_{(\mathbb{P}, \mathcal{Y})}(Z) = \varrho_{(\mathbb{P}, \mathcal{Y})}(X) + \varrho_{(\mathbb{P}, \mathcal{Y})}(Y) = -\frac{4}{3}$. Let $W = [1(R), 1(B), 1(W), -\frac{2}{3}(G)]$. The portfolio W is not efficient, because for calculating $\varrho_{(\mathbb{P}, \mathcal{Y})}(W)$ (see eq. (2)) the best approximation with variables generated by \mathcal{Y} and are dominated by W (see eq. (1)) is $X + Y$. Thus, $W \geq X + Y$ and $\varrho_{(\mathbb{P}, \mathcal{Y})}(W) = \varrho_{(\mathbb{P}, \mathcal{Y})}(X + Y)$. What makes W inefficient is the inequality $\varrho_{(\mathbb{P}, \mathcal{Y})}(W) < \varrho_{(\mathbb{P}, \mathcal{Y})}(W - 2(W - X - Y))$ (i.e., $c = 2$ violates Definition 3).

Lemma 2 *For every portfolio X there is an efficient portfolio Y in $\text{LE}(X)$.*

Proof Note that ϱ is continuous on \mathcal{X} . Fix a portfolio X , and denote $\mathcal{N} = \{R \subseteq S; X - c\mathbb{1}_R \in \text{LE} \text{ for every positive } c\}$. In the finite case there is a maximal set (w.r.t. inclusion) in \mathcal{N} , say T . Let M be a large constant.

Claim 1 There is a maximal (w.r.t \geq) portfolio Z such that $X - M\mathbb{1}_T - Z \in \text{LE}(X)$. Since T is maximal, the set of the portfolios Z whose support is contained in $S \setminus T$ and

$X - M\mathbb{1}_T - Z \in \text{LE}(X)$ is bounded. Thus, there is a maximum (w.r.t \geq), say Z . Due to continuity of ϱ , $X - M\mathbb{1}_T - Z \in \text{LE}(X)$.

Claim 2 For every $0 < M' < M$, $X - M'\mathbb{1}_T - Z \in \text{LE}(X)$.

Note that by (M), $X - Z \in \text{LE}(X)$. By convexity (i.e., (HO) and (S-ADD)) $\varrho(X - M'\mathbb{1}_T - Z) = \varrho(\frac{M'}{M}(X - M\mathbb{1}_T - Z) + \frac{M-M'}{M}(X - Z)) \leq \frac{M'}{M}\varrho(X - M\mathbb{1}_T - Z) + \frac{M-M'}{M}\varrho(X - Z) = \varrho(X)$. On the other hand, by (M), $\varrho(X - M'\mathbb{1}_T - Z) \geq \varrho(X - M\mathbb{1}_T - Z) = \varrho(X)$. Thus, $\varrho(X - M'\mathbb{1}_T - Z) = \varrho(X)$.

Claim 3 There is a maximal (w.r.t. \geq) portfolio Z such that $X - M\mathbb{1}_T - Z \in \text{LE}(X)$ for every $M > 0$.

This is an immediate conclusion of the previous two claims and (M). ■

4.1 Additivity w.r.t efficient portfolios

The key property that characterizes the risk measures induced by PSP is additivity with respect to efficient portfolios.

- *Additivity w.r.t. efficient portfolios (ADD-EFF): For every two efficient portfolios X and Y , $\varrho(X + Y) = \varrho(X) + \varrho(Y)$.*

The intuition behind this axiom is the following. The inequality $\varrho(X + Y) < \varrho(X) + \varrho(Y)$ means that the merger of X and Y strictly reduces risk. However, the axioms states that when X and Y are efficient, there is no reduction in the risk of the merged portfolio. Suppose, that two investments accounts are managed by two desks of the same firm. The axiom requires that if both accounts are managed efficiently (meaning that the portfolios produced are efficient), there would be no advantage in merging in the sense that the risk of the combined portfolio is no less than the sum of the risks. This property allows for a decentralized management, as long as the separate managing divisions produce efficient portfolios.

4.2 The representation theorem

Fix two vectors q and p in \mathbb{R}^S . We say that $q \sim_{\mathcal{Y}} p$ if $\langle Y, q \rangle = \langle Y, p \rangle$ for every $Y \in \text{span}(\mathcal{Y})$, where $\langle \cdot, \cdot \rangle$ is the inner product and Y is referred to as a vector in \mathbb{R}^S .

Theorem 1 *Let S be a finite space. A coherent risk measure ϱ satisfies additivity w.r.t efficient portfolios if and only if there is a partially specified probability $(\mathbb{P}, \mathcal{Y})$ such that $\varrho = \varrho_{(\mathbb{P}, \mathcal{Y})}$.*

Proof Denote by l the set of all efficient or constant portfolios. By (ADD-EFF) and (TI), ϱ is linear on $\text{span}(\mathcal{Y})$. Thus, there is a vector $p = (p(s))_{s \in S}$ in \mathbb{R}^S such that $\varrho(Y) = \langle Y, p \rangle$ for every $Y \in \text{span}(\mathcal{Y})$.

Due to (TI), $\sum_{s \in S} p(s) = 1$. Note that it is not yet proven that $p(s) \geq 0$ for every $s \in S$. Indeed, it might be that p is not a probability vector. This is why (p, \mathcal{Y}) cannot be referred to as a partially specified probability at this point. Define

$$\psi(X) = - \max \left\{ \sum_{Y \in \mathcal{Y}} \lambda_Y \langle Y, p \rangle; \sum_{Y \in \mathcal{Y}} \lambda_Y Y \leq Z \text{ and } \lambda_Y \in \mathbb{R} \text{ for every } Y \in \mathcal{Y} \right\}. \quad (5)$$

The function ψ is convex, monotonically decreasing and it coincides with ϱ on $\text{span}(\mathcal{Y})$. Moreover, Lemma 2 implies that for every X there is an efficient $Y(X) \in \text{LE}$. Thus, $\varrho(X) = \varrho(Y(X)) = \psi(Y(X))$. However, since ψ is monotonic, $\psi(Y) \geq \psi(X)$. Thus, $\varrho(X) \geq \psi(X)$ for every X . On the other hand, ψ is the greatest convex function over \mathcal{X} that coincides with ϱ on $\text{span}(\mathcal{Y})$. Therefore, ϱ and ψ coincide.

It remains to show that there exists $q \sim_{\mathcal{Y}} p$ which is a probability vector. Assume on the contrary that all $q \sim_{\mathcal{Y}} p$ are not probability vectors. It means that the set $A = \mathcal{P}(p, \mathcal{Y})$ does not intersect the unit cube, $B = \{(v_1, \dots, v_{|S|}) \in \mathbb{R}^S; 0 \leq v_i \leq 1 \text{ for every } i = 1, \dots, |S|\}$. Since both A and B are convex and closed and B is bounded, there is a non-zero $W \in \mathbb{R}^S$ such that

$$\langle a, W \rangle < 0 < \langle b, W \rangle \quad (6)$$

for every $a \in A$ and $b \in B$. The right inequality implies that W is a positive vector.

If $W \in \text{span}(\mathcal{Y})$, then $\varrho(W) = \psi(W) = -\langle a, W \rangle > 0$ for every $a \in A$. This contradicts monotonicity (because $0 \leq W$ and thus, $0 = \varrho(0) \geq \varrho(W)$). Therefore, $W \notin \text{span}(\mathcal{Y})$. But then, the linear system of equations (with u being the vector of unknowns), $\psi(Y) = -\langle u, Y \rangle$, $Y \in \mathcal{Y}$ and the additional equation $\langle u, W \rangle = 17$ (of course, 17 could be replaced by any negative number) has a solution, which contradicts the left inequality of eq. (6).

We conclude that the sets A and B are not disjoint, and an element in the intersection, say \mathbb{P} , is necessarily a probability vector. Thus, $\varrho = \psi = \varrho_{(\mathbb{P}, \mathcal{Y})}$ as desired. \blacksquare

5 The infinite case

This section is primarily devoted to show that Lemma 1 is no longer true when S is infinite. It suggests that additional axiom, such as point-wise convergence continuity or one that guarantees an efficient portfolio in every $\text{LE}(X)$, is needed.

Example 3 Suppose that S is the set of natural numbers, \mathbb{N} . Let $A_k = \{2k - 1, 2k\}$, $k \in \mathbb{N}$ and B be the set of all even numbers. Let $v(A_k) = 2^{-k-2}$; for any $T \subseteq \mathbb{N}$ such that $|B \setminus T| < \infty$ let $v(T) = \frac{1}{4}$. Finally, let $v(S) = 1$ and for any other $T \subseteq \mathbb{N}$, set $v(T) = 0$. Define,³

$$\varrho(X) = -\sup \left\{ \sum_{T \subseteq \mathbb{N}} \alpha_T v(T); \sum_{T \subseteq \mathbb{N}} \alpha_T \mathbb{1}_T \leq X \right\},$$

where α_T are real numbers. ϱ is a coherent risk measure. Now consider $X = \mathbb{1}_B$. $\varrho(X) = -\frac{1}{4}$.

Let Y_k be the portfolio $X - \mathbb{1}_{\{2\ell; \ell \leq k\}}$. Note that $Y_k \leq X$ and $\varrho(Y_k) = \varrho(X)$. However, Y_k is not efficient because $\varrho(Y_k) = \varrho(Y_{k+1})$ and $\varrho(Y_k - 2(Y_k - Y_{k+1})) = -(-2^{-(k+1)-2}) = 2^{-k-3} > \varrho(Y_k)$. The sequence Y_k is decreasing. Define $Y = \lim Y_k$. Since $Y = 0$, $\varrho(Y) = 0 > \varrho(X)$. Thus, there is no efficient portfolio in $\text{LE}(X)$.

6 References

- Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath, (1999) "Coherent measures of risk," *Mathematical Finance*, **9**, 203-228.
- Delbaen, F. (2002) "Coherent risk measures on general probability spaces," in Klaus Sandmann et al., editors, *Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann*, 1-37. Springer.
- Gilboa, I. and D. Schmeidler (1989) "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, **18**, 141-153.
- Lehrer, E. (2005a) "A new integral for capacities," to appear in *Economic Theory*.
- Lehrer, E. (2005b) "Partially specified probabilities: decisions and games," mimeo.
- Markowitz, H. (1952) "Portfolio selection," *Journal of Finance*, **7**, 77-91.

³The following is in the spirit of the concave integral for capacities (See Lehrer, 2005a).