Risk and Risk Aversion
when States of Nature Matter*

Jan Werner

Department of Economics,
University of Minnesota

December 2002,
revised May 2006

Abstract: When uncertainty is associated with some intrinsically relevant states of nature, there is no reason for an agent to base his or her preferences only on probability distribution of claims. We propose a new concept of risk for state-contingent claims that, unlike the standard concept of Rothschild-Stiglitz, does not identify state-contingent claims with their probability distribution. This concept is called mean-independent risk, and we provide a simple characterization in terms of marginal utilities of (non-expected) utility functions that exhibit aversion to mean-independent risk. We study implications of aversion to mean-independent risk on agents' choices under uncertainty.

---

1. Introduction

According to the classical theory of risk aversion, due to Pratt (1964) and Arrow (1965), an agent is risk averse if she or he is unwilling to take any actuarially fair (i.e., zero expectation) gamble when starting from a position of no risk. The Arrow-Pratt theory is rooted in the expected utility theory of decision making. A necessary and sufficient condition for an expected-utility maximizing agent to be risk averse is that the agent's von Neumann-Morgenstern utility function be concave. Concavity of expected utility implies that the agent is unwilling to take a mean-independent gamble when starting from a risky position. Furthermore, the greater the scale of a mean-independent gamble, the more unwilling is the agent to take the gamble. Mean-independent gamble has zero expectation conditional on each possible realization of the initial risk.

There are many important implications of risk aversion under expected utility. Examples include Consumption-Based Security Pricing that establishes a relationship between expected return of a risky security and covariance of the return with consumption growth, and comonotonicity of Pareto-optimal risk sharing (see LeRoy and Werner (2001) for a textbook treatment). Expected utility has for many years been loosing its appeal as the primary model of decision making under uncertainty. The list of problems with expected utility is long: the paradoxes of Allais and Ellsberg, the negative experimental evidence, the failure of expected-utility based asset pricing models, etc. Rabin (2000) has pointed out some limitations of risk aversion under expected utility.

The objective of this paper is to extend the theory of risk aversion beyond the expected utility theory of decision making. The original Arrow and Pratt definition is very weak in the absence of expected utility. In particular, it does not imply that the agent is unwilling to take a mean-independent gamble when starting from a risky position, or that the unwillingness to take a mean-independent gamble increases with the scale of the gamble. We take this latter condition as the definition of risk aversion.

We consider an agent facing a choice of a contingent claim under uncertainty. Contingent claims are real valued random variables on a finite set of states of nature

2

equipped with a probability measure. The agent's preferences are described by an arbitrary (continuous) utility function on contingent claims. We first introduce the concept of mean-independent gamble (or risk): contingent claim ϵ is a *mean-independent risk* at another contingent claim z if the conditional expectation of ϵ on z is zero. We say that one contingent claim differs from another by mean-independent risk if they differ by the scale of mean-independent risk at a common "initial" contingent claim. More precisely, x differs from y by mean-independent risk if $x = z + \epsilon$ and $y = z + \lambda\epsilon$ for some $0 \leq \lambda \leq 1$ and some ϵ and z such that ϵ is a mean-independent risk at z. To differ by mean-independent risk is an antisymmetric relation, but it is usually not transitive. We are led to consider its transitive closure which we call *mean-independent more risky.*

Utility function that are monotone decreasing with respect to the relation of mean-independent more risky are called *averse to mean-independent risk.* We prove in Theorem 4.1 that these utility functions are characterized by the property that marginal utilities of consumption in different states rescaled by probabilities of states are inversely related to consumption levels in the states. That is, utility function U exhibits aversion to mean-independent risk, if $\partial_s U(x)/\pi_s \leq \partial_t U(x)/\pi_t$, whenever $x_s \geq x_t$, where $\partial_s U$ denotes $\partial U/\partial x_s$. For an expected utility function, this condition means decreasing marginal utility and is equivalent to concavity of the von Neumann-Morgenstern utility function.

Our concept of risk is related to, but different from the standard Rothschild-Stiglitz (1970) (or second-order stochastic dominance) concept. We prove that if x is mean-independent more risky than y, then x is more risky than y in the Rothschild-Stiglitz sense. The converse does not hold. The main difference between the two concepts is that the Rothschild-Stiglitz relation is defined on (cumulative) distribution functions of random variables and hence it identifies random variables with their distributions, while our relation is defined directly on random variables and it does not identify random variables with distributions. Utility functions that are monotone decreasing with respect to the Rothschild-Stiglitz relation of more risky – these are called strongly risk averse – are averse to mean-independent risk and distribution invariant. In contrast, aversion to mean-independent risk

3

does not necessitate distribution invariance. Often, when states of nature are associated with some intrinsically relevant events, there is no reason for an agent to base her preferences over state-contingent claims only on distribution of claims. For instance, if the two possible states of nature are sunshine and rain, then even if they have equal probabilities, an agent may not be indifferent between, say, contingent claims $(100, 0)$ and $(0, 100)$. For an excellent discussion of this issue see Arrow (1971).

Utility functions that exhibit aversion to mean-independent risk have a simple characterization of Theorem 4.1. This is in contrast to strongly risk averse utility functions for which there exist no such characterization (with an exception of the case of uniform probabilities of states, see Section 4).

Our theory of risk aversion is aimed at applications to models of decision making other than the expected utility. The most important models in the setting of contingent claims on a set of states of nature are the multiple-prior (or maxmin) expected utility of Gilboa and Schmeidler (1989) and the Choquet expected utility with non-additive probabilities of Schmeidler (1979). Many (but not all) multiple-prior expected utility and Choquet expected utility functions are not distribution invariant under any probabilities of states.[1] In the multiple-prior expected utility model the agent has a set of probability measures on states instead of a single measure. In the companion paper (Werner (2005b)) we study the question of whether a multiple-prior expected utility is averse to mean-independent risk with respect to some probability measure in the set of multiple priors. We show that for many sets of priors there exists one such probability measure. It should be noted that some important sets of multiple priors come with a specific "reference" probability measure. Such are neighborhoods of a measure with respect to a metric such as, for instance, relative entropy (introduced in this context by Hansen and Sargent (2001)). These reference measures are of special relevance in the question of aversion to mean-independent risk.

Karni (1985) extended the Arrow-Pratt theory of risk aversion to state-dependent

---

[1]That is, they are not probabilistically sophisticated, using the terminology of Machina and Schmeidler (1992).

utility functions. He defined risk aversion as unwillingness to take actuarially fair gamble when starting from a position of risk-free marginal utility of wealth. This approach is designed specifically for utility functions that are state separable, and it should be considered complementary to our approach. Multiple-prior and Choquet expected utilities are non-separable. We show in Section 4 that a state-separable utility function exhibits aversion to mean-independent risk if and only if it is an expected utility.

The paper is organized as follows: In Section 2 we introduce the concept of mean-independent risk and the relation of mean-independent more risky. A comparison with the Rothschild-Stiglitz relation is in Section 3. A characterization of differentiable utility functions that exhibit aversion to mean-independent risky is proved in Section 4. In Section 5 we show that a property termed preference for conditional expectation is equivalent to aversion to mean-independent risk for quasi-concave utility functions. In Section 6 we extend our characterization result of Section 4 to non-differentiable utility functions using the concept of the superdifferential. This extension is important since most of multiple-prior expected utility functions are not differentiable. In Section 7 we study some implications of aversion to mean-independent risk on agents' choices under uncertainty. We extend Peleg and Yaari (1973) and Dybvig and Ross (1995) characterizations of efficient claims from risk-averse expected utility functions to utility functions that are averse to mean-independent risk.

2. Mean-Independent Risk

There is a finite set $S = \{1, \ldots, S\}$ of states of nature (with $S > 1$) and a probability measure π on the set $\mathcal{S}$ of all subsets of S. Any S-dimensional vector $x = (x_1, \ldots, x_S) \in \mathcal{R}^S$ can be regarded as random variable on $(S, \mathcal{S}, \pi)$ and will be termed a *contingent claim*. Probability measure π can be identified with a vector $(\pi_1, \ldots, \pi_S)$ of probabilities of states. We assume that $\pi_s > 0$ for every state s. We use $E(x)$ to denote the expected value of x under π, and $\mathbf{E}(x)$ to denote the contingent claim (or vector) that takes value $E(x)$ in every state.

We will frequently use conditional expectations. For partition F of states, $E(x|F)$ denotes conditional expectation of x on F. For contingent claim z, $E(x|z)$

denotes conditional expectation of x on z, that is, conditional expectation of x on the partition generated by z. All these conditional expectations are contingent claims (or vectors).

We say that $\epsilon \in \mathcal{R}^S$ is a *mean-independent risk* at $z \in \mathcal{R}^S$ if ϵ is mean-independent of z with zero mean, that is,

$$E(\epsilon|z) = 0. \tag{1}$$

Thus, a mean-independent risk at z is a contingent claim that has zero expectation conditional on each possible value that z may take.

Note that if $\epsilon \neq 0$ is a mean-independent risk at z, then z must be state independent for some subset of states, that is, $z_s = z_{s'}$ for at least two states $s \neq s'$. If there are only two states, then $\epsilon \neq 0$ is mean-independent risk at z if only if z is deterministic and $E(\epsilon) = 0$. In general, for any partition F of states, if $E(\epsilon|F) = 0$ and z is measurable with respect to F, then ϵ is a mean-independent risk at z. In Appendix A we provide characterizations of mean-independent risk.

Consider two contingent claims $x, y \in \mathcal{R}^S$ with the same expectation, $E(x) = E(y)$. We say that x *differs from* y *by mean-independent risk* if there exist $z, \epsilon \in \mathcal{R}^S$ and $0 \leq \lambda \leq 1$ such that ϵ is a mean-independent risk at z and

$$x = z + \epsilon \quad \text{and} \quad y = z + \lambda\epsilon. \tag{2}$$

Thus, x and y differ by the scale of mean-independent risk at a common "initial" contingent claim, with the former having a greater scale than the latter. If (2) holds with $\epsilon \neq 0$ and $\lambda < 1$ (equivalently, if $x \neq y$), then x *strictly* differs from x by mean-independent risk.

For any contingent claim x and a partition F of states, x differs from $E(x|F)$ by mean-independent risk. This is so because $x - E(x|F)$ is a mean-independent risk at $E(x|F)$. Further, x differs from $E(x|F) + \lambda[x - E(x|F)]$ by mean-independent risk for any $0 \leq \lambda \leq 1$ (see Appendix A).

Example 2.1: Let there be three equally probable states. Consider $y = (5, 1, 3)$ and $x = (5, 0, 4)$. Set $z = (5, 2, 2)$ and $\epsilon = (0, -2, 2)$. Clearly, $E(\epsilon|z) = 0$, hence ϵ is a mean-independent risk at z. Since $x = z + \epsilon$ and $y = z + \frac{1}{2}\epsilon$, x differs from y by

mean-independent risk. Note also that $z = E(x|F)$ for partition $F = \{\{1\}, \{2, 3\}\}$.
□

If x differs from y by mean-independent risk and $x \neq y$, then z, ϵ and λ satisfying (2) are unique. First, we note that if $x \neq y$, then $\lambda < 1$ and

$$\epsilon = \frac{1}{1 - \lambda}(x - y). \tag{3}$$

If there exist another triple z', ϵ', λ' satisfying (2), then, because of (3) ϵ and ϵ' are related via $\epsilon' = \frac{1-\lambda}{1-\lambda'}\epsilon$. This implies that $0 = (z' - z) + \frac{\lambda'-\lambda}{1-\lambda'}\epsilon$, and that $E(\epsilon|z') = 0$. Since also $E(\epsilon|z) = 0$, it follows that ϵ is uncorrelated with z and z', and hence also with $z' - z$. Therefore $z = z'$ and $\lambda = \lambda'$, which implies $\epsilon = \epsilon'$.

To differ by mean-independent risk (DMIR, for short) is a binary relation on the set of contingent claims. Of course, it is reflexive. It is also antisymmetric as established by the following

Proposition 2.2 *If x DMIR y and y DMIR x, then $x = y$*

PROOF: We have (2) and

$$y = z' + \epsilon' \quad \text{and} \quad x = z' + \lambda'\epsilon' \tag{4}$$

for some ϵ' and z' such that $E(\epsilon'|z') = 0$. Suppose that $x \neq y$. Then $\lambda < 1$ and $\lambda' < 1$. Using (3) for ϵ and for ϵ', we obtain $\epsilon' = -\frac{1-\lambda}{1-\lambda'}\epsilon$. It follows now from (2) and (4) that $0 = (z' - z) + (\frac{\lambda\lambda'-1}{1-\lambda'})\epsilon$. Since ϵ is uncorrelated with $z - z'$, this implies that $z = z'$ and $\lambda\lambda' = 1$. The latter contradicts $\lambda < 1$ and $\lambda' < 1$. □

It is clearly desirable that a relation of more risky be transitive. If there are only two states, then relation DMIR is transitive (see Appendix A). However, if $S \geq 3$, then DMIR is not transitive.

Example 2.4: As in Example 2.1, there are three equally probable states. We have seen that x DMIR y for $x = (5, 0, 4)$ and $y = (5, 1, 3)$. It is also true that y DMIR w for $w = (4, 2, 3)$ (take $\epsilon = (2, -2, 0)$ and $z = (3, 3, 3)$). It is however not true that x DMIR w. □

We are led to consider the transitive closure of DMIR. The transitive closure of DMIR – called *mean-independent more risky* – is defined as follows: for two

contingent claims x and y, x is mean-independent more risky than y, if there exists a finite collection $x_1, \ldots, x_N$ of contingent claims such that x DMIR x_1, x_n DMIR x_{n+1} for each $1 \leq n \leq N-1$, and x_N DMIR y. If, in addition, $x \neq y$, then x is *strictly mean-independent more risky* than y. Mean-independent more risky is a reflexive and transitive relation. It is also antisymmetric as will be seen in Section 3.

3. Mean-Independent More Risky and Rothschild-Stiglitz More Risky

In this section we compare the relation of mean-independent more risky with the Rothschild-Stiglitz (1970) (or second-order stochastic dominance) relation of more risky. We begin by showing that every concave expected utility is decreasing with respect to the relation of mean-independent more risky. One consequence of this is that mean-independent more risky is an antisymmetric relation.

Proposition 3.1 *If x is mean-independent more risky than y, then $E[v(y)] \geq E[v(x)]$ for every concave function $v : \mathcal{R} \to \mathcal{R}$. Further, if x is strictly mean-independent more risky than y, then $E[v(y)] > E[v(x)]$ for every strictly concave v.*

PROOF: It suffices to show that the claimed inequalities hold if x DMIR y, or x strictly DMIR y, respectively. Suppose that x DMIR y, so that (2) holds. It follows that $y = \lambda x + (1 - \lambda)z$ for $\lambda \in [0, 1]$. Concavity of v implies that

$$v(y) \geq \lambda v(x) + (1 - \lambda)v(z). \tag{5}$$

Taking expectations in (5) we obtain

$$E[v(y)] \geq \lambda E[v(x)] + (1 - \lambda)E[v(z)]. \tag{6}$$

Since $x = z + \epsilon$ and $E(\epsilon|z) = 0$, conditional Jensen's inequality implies that $E[v(z)] \geq E[v(x)]$. Now, (6) implies that

$$E[v(y)] \geq E[v(x)]. \tag{7}$$

The strict version follows in the same way since Jensen's inequality is strict for $\epsilon \neq 0$ and strictly concave v. □

8

Contingent claim x is more risky in the Rothschild-Stiglitz sense than y (*R-S more risky*) if

$$x =^d y + \epsilon \tag{8}$$

for some random variable ϵ such that $E(\epsilon|y) = 0$. Relation $=^d$ indicates that the two random variables have the same distribution. Note that ϵ need not be a random variable on S. If x is R-S more risky than y and y is R-S more risky than x, i.e. if x and y are R-S equally risky, then x and y have the same distribution. A well-known necessary and sufficient condition for x to be R-S more risky than y is that $E[v(y)] \geq E[v(x)]$ holds for every strictly increasing and concave utility function v.

Proposition 3.1 implies the following

Theorem 3.2: *If x is mean-independent more risky than y, then x is R-S more risky than y.*

A converse to Theorem 3.2 does not hold.

Example 3.3: Let there be two equally probable states. Consider $x = (5, 2)$ and $y = (3, 4)$. Clearly, x is R-S more risky than y. However, it is not true that x DMIR y. Since DMIR coincides with mean-independent more risky in the case of two states, x is not mean-independent more risky than y. □

4. Aversion to Mean-Independent Risk

An agent's preferences over contingent claims on $(S, \mathcal{S}, \pi)$ are described by a strictly increasing and continuous utility function $U : \mathcal{R}^S \to \mathcal{R}$. We say that U is *averse to mean-independent risk* under π if $U(y) \geq U(x)$ whenever x is mean-independent more risky than y. Most of the time we drop the phrase "under π," except when a reference to a specific probability measure is required.

By Proposition 3.1, every utility function that has concave expected utility representation is averse to mean-independent risk. Concave expected utilities are in fact the only state-separable utility functions that are averse to mean-independent risk. It follows from Werner (2005a) that any utility function that satisfies the sure-thing principle (and hence is state separable) and is averse to mean-independent risk must have concave expected utility representation. By Theorem 3.2, every

utility function that is decreasing with respect to the relation of R-S more risky is averse to mean-independent risk. Utility functions that are decreasing with respect to the relation of R-S more risky are called *strongly risk averse* (see Cohen (1995)), or *Schur concave*. Examples of non-expected utility functions that are strongly risk averse include mean-variance utility functions that are variance averse, that is, utility functions that depend only on the expectation and the variance of a contingent claim and are decreasing in variance.

Strongly risk averse utility functions are distribution invariant. Utility function U is *distribution invariant* under π if $U(x) = U(y)$ whenever $x =^d y$. Utility functions that are averse to mean-independent risk need not be distribution invariant. For example, the multiple-prior expected utility on two-state contingent claims

$$\min_{0.5-\delta_l \leq \mu \leq 0.5+\delta_u} \{\mu v(x_1) + (1-\mu)v(x_2)\} \tag{9}$$

with $0 < \delta_u, \delta_l < 0.5$, $\delta_u \neq \delta_u$ and concave v, is averse to mean-independent risk under equal probabilities of states, but it is not distribution invariant under such probabilities. For instance, claims $(1, 2)$ and $(2, 1)$ have the same distribution, but different utilities. It is not hard to verify that multiple-prior expected utility (9) is averse to mean-independent risk using the definition, but we defer the proof until Section 5.

We provide now a characterization of differentiable utility functions that are averse to mean-independent risk. For a differentiable function U, we use $DU(x) = (\partial_1 U(x), \ldots, \partial_S U(x))^2$ to denote the gradient of U at x, and $DU(x)/\pi$ to denote the gradient rescaled by probabilities π, that is, the vector with $\partial_s U(x)/\pi_s$ as its s-th coordinate. Further, two vectors $\phi, x \in \mathcal{R}^S$ are said to be *negatively comonotone* if $(\phi_s - \phi_t)(x_s - x_t) \leq 0$ for every s, t.

Theorem 4.1: *Differentiable utility function U is averse to mean-independent risk if and only if $DU(x)/\pi$ is negatively comonotone with x, for every x.*

PROOF: Utility function U is averse to mean-independent risk if and only if $U(z + \lambda\epsilon)$ is a non-increasing function of λ for $\lambda \geq 0$, for every $z, \epsilon \in \mathcal{R}^S$ such that ϵ is

---

[2]$\partial_s U(x)$ denotes $(\partial U/\partial x_s)(x)$.

10

a mean-independent risk at z. Since U is differentiable, a necessary and sufficient condition for $U(z + \lambda\epsilon)$ non-increasing in λ is that $DU(z + \lambda\epsilon)\epsilon \leq 0$ for every $\lambda \geq 0$. Since ϵ can be rescaled, this is equivalent to $DU(z + \epsilon)\epsilon \leq 0$ for every ϵ and z. Using Proposition A.1, Appendix A, we have $DU(x)[x - E(x|F)] \leq 0$ for every x and every partition F. Applying Lemma B.1 from Appendix B concludes the proof. □

For an expected utility function $U(x) = E[v(x)]$ with differentiable von Neumann-Morgenstern utility v, it holds $\partial_s U(x)/\pi_s = v'(x_s)$. Thus, negative comonotonicity of $DU(x)/\pi$ with x means that marginal utility v' is decreasing, or equivalently, that v is concave.

Since strongly risk averse utility functions are averse to mean-independent risk, Theorem 4.1 implies that negative comonotonicity of the rescaled gradient with consumption holds for strongly risk averse utility functions (see Chew and Mao (1995) for a related result.) A fundamental result of Schur and Ostrowski (see Marshall and Olkin (1979)) says that when all states are equally probable, then utility function U is strongly risk averse if and only if it is (1) symmetric and (2) $DU(x)$ is negatively comonotone with x for every x. Utility U is symmetric (or rearrangement invariant) if it is invariant to permutations of coordinates. This is a necessary and sufficient condition for U to be distribution invariant when probabilities of states are equal. No characterization of strongly risk averse utility functions is known for the case of unequal probabilities. It follows from Theorem 4.1 that if probabilities are equal, then symmetric utility function is strongly risk averse if and only if it is averse to mean-independent risk.

5. Preference for Conditional Expectations

We say that utility function U exhibits *preference for conditional expectations* under π if $U(E(x|F)) \geq U(x)$ for every $x \in \mathcal{R}^S$ and every partition F. Equivalently, U exhibits preference for conditional expectations if the agent always rejects a mean-independent risk, that is, $U(z) \geq U(z + \epsilon)$ for every $\epsilon, z \in \mathcal{R}^S$ such that ϵ is mean-independent risk at z (see Proposition A.1, Appendix A). If U is averse to mean-independent risk, then it exhibits preference for conditional expectations.

11

For quasi-concave utility function the converse is also true.

Theorem 5.1: *Suppose that U is quasi-concave. Then U is averse to mean-independent risk if and only if exhibits preference for conditional expectations.*

PROOF: We already proved that aversion to mean-independent risk implies preference for conditional expectations. For the converse implication suppose that U exhibits preference for conditional expectations and let x be mean-independent more risky than y. By Proposition A.2 there exists a partition F such that $y = (1 - \lambda)E(x|F) + \lambda x$. Since $U(E(x|F)) \geq U(x)$, quasi-concavity implies that $U(y) \geq U(x)$, and hence that U is averse to mean-independent risk. $\square$

Preference for conditional expectations implies preference for unconditional expectation, $U(x) \leq U(E(x))$ for all x. For expected utility, the two are equivalent and also equivalent to concavity of the von Neumann-Morgenstern utility function. Also, if there are only two states, preference for conditional expectations is equivalent to preference for unconditional expectation. This makes it is easy to see that multiple-prior expected utility function (9) exhibits preference for conditional expectations. Since function (9) is concave, Theorem 5.1 implies that it is averse to mean-independent risk.

A characterization of preference for conditional expectations for differentiable and quasi-concave utility functions is given in the following

Theorem 5.2: *Suppose that U is quasi-concave and differentiable. U exhibits preference for conditional expectations if and only if, for every x and every s, s',*

$$\text{if } x_s = x_{s'}, \quad \text{then} \quad \frac{\partial_s U(x)}{\pi_s} = \frac{\partial_{s'} U(x)}{\pi_{s'}}. \tag{10}$$

PROOF: Consider arbitrary contingent claim x, and let F_x be the partition of states generated by x. Let $A_1, \ldots A_k$ be the elements of partition F_x. Of course, $x_s = x_{s'}$ for every $s, s' \in A_l$, for every l. If U exhibits preference for conditional expectations, then x solves the following maximization problem

$$\max_{y} \ U(y) \tag{11}$$

$$\text{subject to} \quad E(y|F_x) = x.$$

12

The constraint $E(y|F_x) = x$ can be equivalently written as k constraints $E(y|A_l) = x_l$, for $l = 1, \ldots, k$, where x_l denotes the common value of x_s on the set A_l. The first-order Lagrange conditions for the solution x to this maximization problem can be written as $\frac{\partial_s U(x)}{\pi_s} = \frac{\partial_{s'} U(x)}{\pi_{s'}}$ for every $s, s' \in A_l$ and every l. This implies that (10) holds. Conversely, if (10) holds, then x satisfies the first-order conditions for the maximization problem (11). Since U is quasi-concave, the first-order conditions are sufficient so that x is a solution. This implies that U exhibits preference for conditional expectations. $\square$

Of course, preference for conditional expectations in Theorem 5.2 can be substituted by aversion to mean-independent risk, because of Theorem 5.1. Condition (10) can be equivalently stated as that the rescaled gradient $DU(x)/\pi$ is measurable with respect to the partition generated by x.

When all states are equally probable, every symmetric differentiable utility function satisfies condition (10). If, in addition, U is quasi-concave, then it follows that it is averse to mean-independent risk under equal probabilities. Dasgupta, Sen and Starrett (1973) (see also Rothschild and Stiglitz (1973)) proved that every symmetric and quasi-concave utility function is strongly risk averse when states are equally probable.

6. Mean-Independent Risk Aversion without Differentiability

The drawback of our characterization results in sections 4 and 5 is that they require differentiability of the utility function. Many interesting non-expected utility functions are not differentiable. Multiple-prior expected utility functions are non-differentiable, although usually on small sets of contingent claims. Multiprior expected utility (9) is non-differentiable at every deterministic contingent calim.

The results of Sections 4 and 5 can be extended to non-differentiable utility functions by the use of superdifferentials. Let utility function U be concave. The *superdifferential* of U at x is defined as the set of all vectors $\phi \in \mathcal{R}^S$ that satisfy

$$U(y) \leq U(x) + \phi(y - x), \tag{12}$$

for every y. The superdifferential of U at x is denoted by $\partial U(x)$. It is a nonempty

13

set that reduces to single vector $DU(x)$ if U is differentiable at x.

An extension of Theorem 5.2 to non-differentiable, concave utility functions is the following

Theorem 6.1: *Suppose that U is concave. U is averse to mean-independent risk if and only if for every x there exists $\hat{\phi} \in \partial U(x)$ such that*

$$\text{if } x_s = x_{s'}, \quad \text{then} \quad \frac{\hat{\phi}_s}{\pi_s} = \frac{\hat{\phi}_{s'}}{\pi_{s'}}. \tag{13}$$

PROOF: By Theorem 5.1, U is averse to mean-independent risk if and only if it exhibits preference for conditional expectations. This is in turn equivalent (see the proof of Theorem 5.2) to x solving the maximization problem (11), for every x. The maximization problem (11) can be rewritten as

$$\max_{\epsilon} \ U(x + \epsilon) \tag{14}$$

$$\text{subject to} \quad E(\epsilon|F_x) = 0.$$

Clearly, x is a solution to (11) if and only if 0 is a solution to (14). The constraint set in (14) is the linear subspace of all ϵ that are mean-independent risk at x. In Appendix A we discuss geometrical structure of this subspace.

A necessary and sufficient condition for 0 to be a solution to (14) for concave function U is that there exists $\hat{\phi} \in \partial U(x)$ such that $\hat{\phi}$ is orthogonal to the constraint set at x (see Rockafellar [12], page 54). It follows from Proposition A.3 in Appendix A that $\hat{\phi}$ is orthogonal to the subspace of all ϵ that are mean-independent risk at x if and only if $\hat{\phi}/\pi$ is measurable with respect to F_x. This is equivalent to condition (13). $\square$

A partial extension of Theorem 4.1 to non-differentiable, concave functions is the following

Proposition 6.1: *Suppose that U is concave. If, for every $\phi \in \partial U(x)$ and every x, ϕ/π is negatively comonotone with x, then U is averse to mean-independent risk.*

PROOF: We first observe, as in the proof of Theorem 4.1, that U is averse to mean-independent risk if and only if $U(z + \lambda\epsilon)$ is a non-increasing function of λ

14

for $\lambda \geq 0$, for every $z, \epsilon \in \mathcal{R}^S$ such that ϵ is a mean-independent risk at z. Let $g(\lambda) = U(z + \lambda\epsilon)$. Single-variable function g is concave. For $x = z + \lambda\epsilon$ it holds $\partial g(\lambda) = \{\epsilon\phi : \phi \in \partial U(x)\}$, where $\partial g(\lambda)$ is the superdifferential of g (see Aubin (1998)). By Proposition A.1, Appendix A, ϵ can be written as $x - E(x|F_z)$ where F_z is the partition generated by z. Using Lemma B.1 we obtain that $\partial g(x) \leq 0$. This implies (Aubin (1998)) that g is a non-increasing function of λ. $\square$

7. Efficient Claims Under Aversion to Mean-Independent Risk

In this section we study some implications of aversion to mean-independent risk on agents' choices under uncertainty. We extend characterizations of efficient claims of Peleg and Yaari (1973) and Dybvig and Ross (1995) from risk-averse expected utility functions to utility functions that are averse to mean-independent risk. Chew and Zilcha (1990) demonstrated that the Peleg-Yaari characterization remains unchanged for strongly risk averse utility functions. We show that the same characterizations hold for utility functions that are averse to mean-independent risk. This implies that such features of expected utility as state separability and distribution invariance have no implications on efficient claims. All the implications come from risk aversion.

Let $X \subset \mathcal{R}^S$ be a closed and convex set of contingent claims on $(S, \mathcal{S}, \pi)$ that are available to a decision maker. A claim $x^* \in X$ is *supported by a price system* $p \in \mathcal{R}^S$, $p \neq 0$ if

$$px^* \geq px \qquad \forall x \in X. \tag{15}$$

We say that $x^* \in X$ is *EUC-efficient* if there exists a strictly increasing and concave utility function $v : \mathcal{R} \to \mathcal{R}$ such that

$$E[v(x^*)] \geq E[v(x)] \qquad \forall x \in X. \tag{16}$$

Similarly, we say that $x^* \in X$ is *MIC-efficient* if there exists a strictly increasing and quasi-concave utility function $U : \mathcal{R}^S \to \mathcal{R}$ exhibiting aversion to mean-independent risk such that

$$U(x^*) \geq U(x) \qquad \forall x \in X. \tag{17}$$

15

Theorem 7.1: *The following conditions are equivalent*

  (i) x^* is EUC-efficient,

  (ii) x^* is MIC-efficient,

  (iii) x^* is supported by some price system $p \in \mathcal{R}_{++}^S$ such that x^* and p/π are *negatively comonotone.*

PROOF: Obviously, (i) implies (ii). Suppose that (ii) holds, that is, that x^* is utility maximizing for U. Let P denote the preferred-to-x^* set for U, and P^o the strictly preferred-to-x^* set. Clearly $P = \mathrm{cl}P^o$, and P is closed and convex. Further, $E(x^*|F) \in P$ for every partition F. Let p be a price vector that separates X from P^o. Then p supports x^*, and $py \geq px^*$ for all $y \in P$. In particular, $pE(x^*|F) \geq px^*$. Using Lemma B.1 we obtain (iii).

A proof that (iii) implies (i) can be found in Peleg and Yaari (1973). □

When the choice set X is given as budget feasible payoffs of marketed assets, the decision maker's problem is termed the *canonical portfolio problem* (Dybvig and Ross (1995)). Formally, this choice set is

$$X_A = \{x \in \mathcal{R}^S : x = \sum_{j=1}^{k} h_j a_j \text{ for some } h \in \mathcal{R}^k \text{ satisfying } \sum_{j=1}^{k} h_j q_j = w\}, \quad (18)$$

where $a_j \in \mathcal{R}_+^S$ and $q_j \in \mathcal{R}_+$ are, respectively, the payoff and the price of asset j, k is the number of assets, and $w \in \mathcal{R}_+$ is the decision maker's wealth. Since the choice set X_A is affine, it follows that p supports $x^* \in X_A$ if and only if $q_j = pa_j$ for every j. If p is strictly positive, then this is, of course, the no-arbitrage condition on asset price vector q. We have the following corollary to Theorem 7.1

Corollary 7.2: *The following conditions are equivalent for the canonical portfolio problem*

  (i) x^* is EUC-efficient,

  (ii) x^* is MIC-efficient,

  (iii) x^* and p/π are negatively comonotone for some $p \in \mathcal{R}_{++}^S$ such that $q_j = pa_j$ *for all j.*

Appendix

A. The Structure of Mean-Independent Risk.

Characterizations of mean independent risk and the relation DMIR can be given using conditional expectations.

Proposition A.1 ϵ *is a mean-independent risk at* z *if and only if there exist* x *and a partition* F *such that*

$$z = E(x|F), \quad \text{and} \quad \epsilon = x - E(x|F.) \tag{19}$$

PROOF: Suppose that ϵ is mean-independent risk at z. Let partition F be generated by z, and let $x = z + \epsilon$. Then $E(x|F) = E(x|z) = z$ and $\epsilon = x - E(x|F)$. Conversely, if (19) holds for some x and F, then $E(\epsilon|z) = E(\epsilon|F) = E([x - E(x|F)]|F) = 0$, so that ϵ is mean-independent risk at z. $\square$

An immediate consequence of Proposition A.1 is the following

Proposition A.2 x *DMIR* y *if and only if there exists partition* F *such that*

$$y = E(x|F) + \lambda[x - E(x|F)] \tag{20}$$

for some $0 \leq \lambda \leq 1$.

Note that (20) implies that $E(x|F) = E(y|F)$. Taking $\lambda = 0$ in Proposition A.2, we obtain that x DMIR $E(x|F)$ for ever partition F and every x. In particular, x DMIR $\mathbf{E}(x)$.

With two states, $\epsilon \neq 0$ is mean-independent risk at z only if z is deterministic, and consequently x DMIR y if and only if $y = \mathbf{E}(x) + \lambda[x - \mathbf{E}(x)]$ for some $0 \leq \lambda \leq 1$. It is easy to show that this a transitive relation on $\mathcal{R}^2$.

For a contingent claim z, let F_z denote the partition generated by z and X_z the set of all contingent claims that are measurable with respect F_z. If ϵ is mean-independent risk at z, then it is also mean-independent risk at any $z' \in X_z$. Indeed, since z' is F_z-measurable, it follows that $E(\epsilon|z') = E[E(\epsilon|z)|z'] = 0$.

The set of ϵ that are mean-independent risks at z is a linear subspace of $\mathcal{R}^S$. Of course, X_z is a linear subspace, too. There is a simple geometry of these

subspaces. To explain this, we introduce the *expectations inner product* in $\mathcal{R}^S$ defined by the expectation $E(xy)$ with respect to π of the (state-by-state) product any two contingent claims x and y. Whenever two claims or two subspaces of claims are orthogonal to each other under the expectation inner product, we refer to them as π-*orthogonal*. Note that x is π-orthogonal to y if and only if x/π is orthogonal to y in the usual Euclidean inner product. We have the following

Proposition A.3 ϵ *is mean-independent risk at z if and only if ϵ is π-orthogonal to X_z.*

PROOF: If ϵ is mean-independent risk at z, then, as we already proved, $E(\epsilon|z') = 0$ for every $z' \in X_z$. This implies $E(\epsilon z') = 0$ for every $z' \in X_z$, hence ϵ is π-orthogonal to X_z. For the converse implication, suppose that ϵ is π-orthogonal to X_z. Let $A_1, \ldots A_k$ be the elements of partition F_z. Since the indicator function 1_{A_j} lies in X_z, for every j, we have $E(\epsilon 1_{A_j}) = 0$. This implies $E(\epsilon|A_j) = 0$ for every j, and hence $E(\epsilon|F_z) = 0$, or equivalently $E(\epsilon|z) = 0$. This concludes the proof. $\square$

B. Negative Comonotonicity.

Lemma B.1: *For $x, \phi \in \mathcal{R}^S$, ϕ/π and x are negatively comonotone if and only if*

$$\phi x \le \phi E[x|F] \tag{21}$$

for every partition F of S.

PROOF: Let F be an arbitrary partition of S consisting of disjoint subsets $S_1, \ldots, S_k$. For any $\phi, x \in \mathcal{R}^S$ we have

$$\phi[x - E(x|F)] = \sum_j \sum_{s \in S_j} \phi_s[x_s - E(x|S_j)] = \tag{22}$$

$$= \sum_j \left(\frac{1}{\pi(S_j)} \sum_{s \in S_j} \phi_s[\pi(S_j)x_s - \sum_{t \in S_j} \pi_t x_t] \right). \tag{23}$$

Using $\pi(S_j) = \sum_{t \in S_j} \pi_t$, we can rewrite each summand in (23) as

$$\frac{1}{\pi(S_j)} \sum_{s,t \in S_j} \phi_s \pi_t[x_s - x_t] = \frac{1}{\pi(S_j)} \sum_{s,t \in S_j, s < t} (\phi_s \pi_t[x_s - x_t] + \phi_t \pi_s[x_t - x_s]). \tag{24}$$

18

Finally, we obtain that

$$\phi[x - E(x|F)] = \sum_j \left(\frac{1}{\pi(S_j)} \sum_{s,t \in S_j, s<t} \pi_s \pi_t \left[\frac{\phi_s}{\pi_s} - \frac{\phi_t}{\pi_t} \right] [x_s - x_t] \right). \quad (25)$$

If ϕ/π and x are negatively comonotone, then it follows from (25) that $\phi[x - E(x|F)] \leq 0$ for every partition F.

Conversely, suppose that (21) holds for every partition F. For any $s, t \in S$, consider partition F_{st} of S that consists of the set $\{s, t\}$ and singleton sets for all states other than s or t. Applying (25) to F_{st}, we have

$$0 \geq \phi[x - E(x|F_{st})] = \frac{\pi_s \pi_t}{\pi_s + \pi_t} \left[\frac{\phi_s}{\pi_s} - \frac{\phi_t}{\pi_t} \right] [x_s - x_t]. \quad (26)$$

Since (26) holds for every $s, t \in S$, it follows that ϕ/π and x are negatively comonotone. $\square$

References

Arrow, K. *Aspect of the Theory of Risk Bearing,* Yrjö Jahnsson Foundation, Helsinki, 1965.

Arrow, K. Exposition of the Theory of Choice under Uncertainty, in C. B. McGuire and R. Radner eds., *Decision and Organization*, North-Holland, 1971.

Aubin, J-P. *Optima and Equilibria, An Introduction to Nonlinear Analysis*, Springer. 1998.

Chew, S.H. and M.H. Mao, "A Schur Concave Characterization of Risk Aversion for Non-expected Utility Preferences," *Journal of Economic Theory*, 67, 402-435, (1995).

Chew, S.H. and I. Zilcha, "Invariance of the Efficient Sets when the Expected Utility Hypothesis is Relaxed," *Journal of Economic Behavior and Organization*, 13, 125-131, (1990).

Cohen, M. D. "Risk-Aversion Concepts in Expected- and Non-Expected-Utility Models," *The Geneva Papers on Risk and Insurance Theory*, 20, 73-91, (1995).

Dybvig P. and S. Ross, "Portfolio Efficient Sets," *Econometrica*, 50, 1523-1546, (1995).

Dasgupta, P., A. Sen, and D. Starrett, "Notes on the Measurement of Inequality," *Journal of Economic Theory*, 6, 180-187, (1973).

Gilboa, I. and D. Schmeidler, "Maxmin expected utility with non unique prior," *Journal of Mathematical Economics*, 18, (1989), 141-153.

Hansen, L. and Th. Sargent, "Robust Control and Model Uncertainty," *American Economic Review*, **91**, pp 60-66, (2001).

Karni, E. *Decision Making Under Uncertainty*, Harvard University Press, 1985.

LeRoy, S. and J. Werner, *Principles of Financial Economics*, Cambridge University Press 2001.

Machina, M. and D. Schmeidler, "A More Robust Definition of Subjective Probability", *Econometrica*, **60**, No. 4, 745-780, (1992)

Marshall A. W. and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, 1979.

Peleg, B. and M. Yaari, "A Price Characterization of Efficient Random Variables,"

Econometrica, 43, (1973), 283-292.

Pratt, J. "Risk Aversion in the Small and in the Large" *Econometrica*, 32: 132-136, (1964).

Rabin, M. "Risk Aversion and Expected Utility Theory: A Calibration Theorem," *Econometrica*, 68, 1281-1292, (2000).

Rothschild, M. and J.E. Stiglitz, "Increasing Risk. I: A Definition," *Journal of Economic Theory*, 2, (1970), 225-243.

Rothschild, M. and J.E. Stiglitz, "Some Further Results on the Measurement of Inequality," *Journal of Economic Theory*, 6, (1973), 188-204.

Rockafellar, T. *The Theory of Subgradients and Its Applications to Problems of Optimization.*, Helderman, 1981.

Schmeidler, D. "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57, (1989), 571-587.

Werner, J. "A Simple Axiomatization of Risk-Averse Expected Utility," *Economics Letters,* 88: 73-77, (2005a).

Werner, J. ' 'Risk Aversion for Multiple-Prior Expected Utility," mimeographed, (2005b).