INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH
ICER

WORKING PAPER SERIES

Fabio Maccheroni, Massimo Marinacci, Aldo Rustichini and Marco Taboga

PORTFOLIO SELECTION WITH
MONOTONE MEAN-VARIANCE PREFERENCES

Working Paper no. 27/2004
December 2004

APPLIED MATHEMATICS
WORKING PAPER SERIES

Portfolio Selection with Monotone Mean-Variance Preferences

Fabio Maccheroni
Istituto di Metodi Quantitativi and IGIER, Università Bocconi

Massimo Marinacci
Dipartimento di Statistica e Matematica Applicata and ICER, Università di Torino

Aldo Rustichini
Department of Economics, University of Minnesota

Marco Taboga*
Research Department, Banca d'Italia and Università di Torino

First draft: Spring 2004
This draft: December 2004

Abstract

We propose a portfolio selection model based on a class of preferences that coincide with mean-variance preferences on their domain of monotonicity, but differ where mean-variance preferences fail to be monotone.

1 Introduction

Since Markowitz's ([M]) seminal paper on portfolio selection, mean-variance preferences have been extensively used to model the behavior of economic

*The views expressed in the article are those of the author and do not involve the responsibility of the bank.

agents choosing among uncertain prospects. These preferences, denoted by $\succeq_{mv}$ assign to an uncertain prospect f the following utility score:

$$U_\theta(f) = \mathrm{E}^P[f] - \frac{\theta}{2}\mathrm{Var}^P[f],$$

where P is a given probability measure and θ is an index of the agent's uncertainty aversion.

The success of such a specification of preferences is due to its analytical tractability and clear intuitive meaning. Mean-variance preferences have, however, a major theoretical drawback: they may fail to be monotone. It may happen that an agent with mean-variance preferences strictly prefers less to more, thus violating one of the most compelling principles of economic rationality.

This is a well known problem, which can be illustrated with a simple example. Consider a mean-variance agent with $\theta = 2$. Suppose she has to choose among the two following prospects f and g:

States of Nature	s_1	s_2	s_3	s_4
Probabilities	0.25	0.25	0.25	0.25
Payoff to act f	1	2	3	4
Payoff to act g	1	2	3	5

Prospect g yields a strictly higher payoff than f in every state. Any rational agent should prefer g to f. However, it turns out that our mean-variance agent strictly prefers f to g. In fact:

$$U_2(f) = 1.25 > 0.5625 = U_2(g).$$

The reason why monotonicity fails here is fairly intuitive. By choosing g rather than f, the payoff in state s_4 increases by one unit. This additional unit increases the mean payoff, but it also makes the distribution of payoffs more spread out, thus increasing the variance. The increase in the mean is more than compensated by the increase in the variance, and this makes our mean-variance agent worse off.

In this paper we propose an adjusted version of mean-variance preferences that satisfies monotonicity, based on Maccheroni, Marinacci and Rustichini [MMR-1] and [MMR-2]. As it will be detailed in the next section, they axiomatize a class of preferences, called variational preferences, that includes as a special case the preferences $\succeq^{mmv}$ represented by the choice functional

$$V_\theta(f) = \min_Q \left\{ \mathrm{E}^Q[f] + \frac{1}{2\theta} C(Q||P) \right\}, \qquad \forall f \in \mathcal{L}^2(P),$$

where Q ranges over all probability measures with square-integrable density with respect to P, and $C(Q||P)$ is the relative Gini concentration index, a concentration index that enjoys properties similar to those of the relative entropy (see the next section for details). Unlike mean-variance preferences, the preferences $\succeq^{mmv}$ are monotone.

Consider the domain of monotonicity $\mathcal{G}_\theta$ of mean-variance preferences, that is, the set of prospects on which they are monotone. The set $\mathcal{G}_\theta$ is where mean-variance preferences are economically meaningful. [MMR-1] and [MMR-2] show that the two choice functionals U_θ and V_θ coincide on $\mathcal{G}_\theta$, that is,

$$\mathrm{E}^P[f] - \frac{\theta}{2}\mathrm{Var}^P[f] = \min_Q \left\{ \mathrm{E}^Q[f] + \frac{1}{2\theta} C(Q||P) \right\}, \qquad \forall f \in \mathcal{G}_\theta.$$

Moreover, they show that V_θ is the minimal monotone choice functional that extends the mean-variance functional U_θ outside the domain of monotonicity $\mathcal{G}$, and that V_θ pointwise dominates U_θ.

The preferences $\succeq_{mmv}$ have, therefore, the following key properties:

- They agree with mean-variance preferences where mean-variance preferences are economically meaningful.
- Their choice functional V_θ is the minimal, and so the most cautious, monotone functional that extends the mean-variance functional outside its domain of monotonicity.
- The functional V_θ is also the best best possible monotone approximation of U_θ: if V'_θ is any other monotone extension of U_θ outside $\mathcal{G}_\theta$, then

  $$|V_\theta(f) - U_\theta(f)| \leq |V'_\theta(f) - U_\theta(f)|$$

  for each prospect f.

These properties make the preferences $\succeq_{mmv}$ a natural adjusted version of mean-variance preferences that satisfies monotonicity. For this reason we call them monotone mean-variance preferences.

In view of all this, it is natural to wonder what happens in a portfolio problem à la Markowitz if we use monotone mean-variance preferences in place of standard mean-variance preferences. This is the main subject matter of this paper. Markowitz's well-known optimal allocation rule under mean-variance preferences is:

$$\alpha_{mv}^{*} = \frac{1}{\theta}\mathrm{Var}^{P}\left[X\right]^{-1}\mathrm{E}^{P}\left[X - \vec{1}R\right],$$

where α_{mv}^{*} is the optimal portfolio of risky assets, X is the vector of gross returns on the risky assets, R is the gross return on the risk-free asset, and $\vec{1}$ is a vector of 1s. Our main result, Theorem 3, shows that with monotone mean-variance preferences the optimal allocation rule becomes:

$$\alpha_{mmv}^{*} = \frac{1}{\theta P\left(W \leq \kappa\right)}\mathrm{Var}^{P}\left[X \left| W \leq \kappa\right.\right]^{-1}\mathrm{E}^{P}\left[X - \vec{1}R \left| W \leq \kappa\right.\right],$$

where W is future wealth and κ is a constant determined along with α_{mmv}^{*} by solving a suitable system of equations.

Except for a scaling factor, the difference between Markowitz's optimal portfolio α_{mv}^{*} and the above portfolio α_{mmv}^{*} is that in the latter conditional moments of asset returns $\mathrm{E}^{Q}\left[\cdot \left| W \leq \kappa\right.\right]$ and $\mathrm{Var}^{Q}\left[\cdot \left| W \leq \kappa\right.\right]$ are used instead of unconditional moments, so that the allocation α_{mmv}^{*} ignores the part of the distribution of X where wealth is higher than κ. As a result, a monotone mean-variance agent does not take into account those high payoff states which contribute to increase the mean return, but give an even greater contribution to increase the variance. By doing so, this agent does not incur in violations of monotonicity caused for mean-variance preferences by an exaggerate penalization of "positive deviations from the mean".

This is a key feature of monotone mean-variance preferences, which we illustrate in Section 5 by showing how it avoids some pathological situations in which the more the payoff to an asset is increased in some states, the more a mean-variance agent reduces the quantity of it in her portfolio, until in the limit she ends up holding none.

The paper is organized as follows. Section 2 illustrates in detail monotone mean-variance preferences. Sections 3 and 4 derive the optimal allocation rule under the proposed specification of preferences. Section 5 presents some examples to illustrate the difference between the optimal allocation rule proposed here and Markowitz's. All proofs are collected in the Appendix.

2 Monotone Mean-Variance Preferences

We consider a measurable space (S, Σ) of states of nature. An uncertain prospect is a Σ-measurable real valued function $f : S \to \mathbb{R}$, that is, a stochastic monetary payoff.

The agent's preferences are described by a binary relation $\succeq$ on a set of uncertain prospects. [MMR-1] provides a set of simple behavioral conditions that guarantee the existence of a utility function $u : \mathbb{R} \to \mathbb{R}$ and an uncertainty index $c : \Delta \to [0, \infty]$ on the set Δ of all probability measures, such that

$$f \succeq g \Leftrightarrow \inf_{Q \in \Delta} \left\{ \mathrm{E}^Q \left[u \left(f \right) \right] + c \left(Q \right) \right\} \geq \inf_{Q \in \Delta} \left\{ \mathrm{E}^Q \left[u \left(g \right) \right] + c \left(Q \right) \right\} \tag{1}$$

for all (simple) prospects f, g.

Preferences having such a representation are called *variational*, and two important special cases of variational preferences are the multiple priors preferences of Gilboa and Schmeidler [GS], obtained when c only takes on values 0 and ∞, and the multiplier preferences of Hansen and Sargent [HS], obtained when $c(Q)$ is proportional to the relative entropy of Q with respect to a fixed probability measure P.[1] Variational preferences satisfy the basic tenets of economic rationality. In particular, they are monotone, that is, given any two prospects f and g, we have $f \succeq g$ whenever $f(s) \succeq g(s)$ for each $s \in S$.

For concreteness, given a probability measure P on (S, Σ), we consider the set $\mathcal{L}^2(P)$ of all square integrable uncertain prospects. A mean-variance preference relation $\succeq_{mv}$ on $\mathcal{L}^2(P)$ is represented by the choice functional $U_\theta : \mathcal{L}^2(P) \to \mathbb{R}$ given by

$$U_\theta \left(f \right) = \mathrm{E}^P \left[f \right] - \frac{\theta}{2} \mathrm{Var}^P \left[f \right],$$

[1] The relative entropy of Q given P is $\mathrm{E}^P \left[\frac{dQ}{dP} \ln \frac{dQ}{dP} \right]$ if $Q \ll P$ and ∞ otherwise.

with $\theta > 0$. Let $\mathcal{G}_\theta$ be the domain of monotonicity of the preference $\succeq_{mv}$, that is, the subset of $\mathcal{L}^2(P)$ where $\succeq_{mv}$ is monotone. Formally, $\mathcal{G}_\theta$ is the convex set where the Gateaux differential of U_θ is positive (as a linear functional). Some algebra shows that

$$\mathcal{G}_\theta = \left\{ f \in \mathcal{L}^2(P) : f - \mathrm{E}^P[f] \leq \frac{1}{\theta} \ \ P\text{-a.s.} \right\}$$

As we discussed in the Introduction, the domain of monotonicity $\mathcal{G}_\theta$ is where the mean-variance preference $\succeq_{mv}$ is economically meaningful. [MMR-1] and [MMR-2] show that the restriction of $\succeq_{mv}$ to $\mathcal{G}_\theta$ is a variational preference, and that

$$U_\theta(f) = \min_{Q \in \Delta^2(P)} \left(\mathrm{E}^Q[f] + \frac{1}{2\theta} C(Q||P) \right), \quad \forall f \in \mathcal{G}_\theta,$$

where $\Delta^2(P)$ is the set of all probability measures with square-integrable density with respect to P, and $C(Q||P)$ is the relative Gini concentration index given by

$$C(Q||P) = \begin{cases} \mathrm{E}^P\left[\left(\frac{dQ}{dP}\right)^2\right] - 1 & \text{if } Q \ll P, \\ \infty & \text{otherwise.} \end{cases}$$

Along with the Shannon entropy, the Gini index is the most classic concentration index. For discrete distributions it is given by $\sum_{i=1}^n Q_i^2 - 1$, and $C(Q||P)$ is its continuous and relative version. [MMR-2] studies in detail the properties of $C(Q||P)$, which turn out to be similar to those of the relative entropy.

Now, say that a preference is a *monotone mean-variance preference*, written $\succeq_{mmv}$, if it is represented by the choice functional $V_\theta : \mathcal{L}^2(P) \to \mathbb{R}$ given by

$$V_\theta(f) = \min_{Q \in \Delta^2(P)} \left\{ \mathrm{E}^Q[f] + \frac{1}{2\theta} C(Q||P) \right\}, \quad \forall f \in \mathcal{L}^2(P), \tag{2}$$

where $\theta > 0$. [MMR-2] proves the following result:

Theorem 1 *The functional $V_\theta : \mathcal{L}^2(P) \to \mathbb{R}$ given by (2) is the minimal monotone functional on $\mathcal{L}^2(P)$ such that $V_\theta(f) = U_\theta(f)$ for all $f \in \mathcal{G}_\theta$; that is,*

$$V_\theta(f) = \sup\{U_\theta(g) : g \in \mathcal{G}_\theta \text{ and } g \leqslant f\}, \quad \forall f \in \mathcal{L}^2(P). \tag{3}$$

Moreover, $V_\theta(f) \geq U_\theta(f)$ for each $f \in \mathcal{L}^2(P)$.

The functional V_θ is concave, continuous, and in view of this theorem it has the following fundamental properties:

(i) V_θ coincides with the mean-variance choice functional U_θ on its domain of monotonicity $\mathcal{G}_\theta$;

(ii) V_θ is the minimal monotone extension of U_θ outside the domain of monotonicity $\mathcal{G}_\theta$, and so it is the most cautious monotone adjustment of the mean-variance choice functional.

(iii) V_θ is the best possible monotone approximation of U_θ: if V'_θ is any other monotone extension of U_θ outside the domain of monotonicity $\mathcal{G}_\theta$, then

$$|V_\theta(f) - U_\theta(f)| \leq |V'_\theta(f) - U_\theta(f)|, \quad \forall f \in \mathcal{L}^2(P).$$

In view of (i)-(iii), the monotone choice functional V_θ provides a natural adjustment of the mean-variance choice functional. It also has the remarkable feature of involving, like multiplier preferences ([HS]), a classic concentration index. This ensures to V_θ a good analytical tractability, as it will be seen in the next section.

The next theorem further illustrates the nature of V_θ.

Theorem 2 *Let $f \in \mathcal{L}^2(P)$. Then:*

$$V_\theta(f) = \begin{cases} U_\theta(f) & \text{if } f \in \mathcal{G}_\theta, \\ U_\theta(f \wedge \kappa_f) & \text{else,} \end{cases}$$

where

$$\kappa_f = \max\{t \in \mathbb{R} : f \wedge t \in \mathcal{G}_\theta\}. \tag{4}$$

A monotone mean-variance agent can thus be regarded as still using the mean-variance functional U_θ even in evaluating prospects outside the domain of monotonicity $\mathcal{G}_\theta$. In this case, however, the agent no longer considers the original prospects, but rather their truncations at κ_f, the largest constant t such that $f \wedge t$ belongs to $\mathcal{G}_\theta$.

Observe that, besides depending on the given act f, the constant κ_f also depends on the parameter θ. Corollary 9 in Appendix provides an explicit formula for κ_f, which shows that κ_f becomes lower when θ increases.

3 The Portfolio Selection Problem

We now present the static portfolio choice problem we are going to study. We consider the one-period allocation problem of an agent who has to decide how to invest a unit of wealth at time 0, dividing it among $n+1$ assets. The first n assets are risky, while the $(n+1)$-th is risk-free. The gross return on the i-th asset after one period is denoted by X_i. The $(n \times 1)$ vector of the returns on the first n assets is denoted by X and the $(n \times 1)$ vector of portfolio weights, indicating the fraction of wealth invested in each of the first n assets, is denoted by α. The return on the $(n+1)$-th asset is risk-free, and so equal to a constant R.

The end-of-period wealth W_α induced by a choice of α is given by:

$$W_\alpha = R + \alpha \left(X - \vec{1}R \right).$$

We assume that there are no frictions of any kind: securities are perfectly divisible; there are no transaction costs or taxes; agents are price-takers, in that they believe that their choices do not affect the distribution of asset returns; there are no institutional restrictions, so that agents are allowed to buy, sell or short sell any desired amount of any security.[2] As a result, α can be chosen in $\mathbb{R}^n$.

We adopt $\succeq_{mmv}$ as a specification of preferences, and so portfolios α are ranked according to the preference functional:

$$V_\theta \left(W_\alpha \right) = \min_{Q \in \Delta^2(P)} \left(\mathrm{E}^Q \left[W_\alpha \right] + \frac{1}{2\theta} C \left(Q || P \right) \right),$$

where P is the reference probability measure. Hence, the portfolio problem can be written as:

$$\max_{\alpha \in \mathbb{R}^n} \min_{Q \in \Delta^2(P)} \left(\mathrm{E}^Q \left[W_\alpha \right] + \frac{1}{2\theta} C \left(Q || P \right) \right).$$

4 The Optimal Portfolio

In this section we give a solution to the portfolio selection problem outlined in the previous section. The characterization of the optimal portfolio is given

[2] This assumption can be weakened, by simply requiring that at an optimum institutional restrictions are not binding.

by the following theorem.[3]

Theorem 3 *The vector* α^* *is a solution of the portfolio selection problem if and only if there exists* $\kappa^* \in \mathbb{R}$ *such that* (α^*, κ^*) *satisfies the system of equations:*

$$\begin{cases} \theta P\left(W_\alpha \le \kappa\right) \mathrm{Var}^P\left[X\left|W_\alpha \le \kappa\right.\right]\alpha = \mathrm{E}^P\left[X - \vec{1}R\left|W_\alpha \le \kappa\right.\right], \\ P\left(W_\alpha \le \kappa\right)\left(\kappa - \mathrm{E}^P\left[W_\alpha\left|W_\alpha \le \kappa\right.\right]\right) = \frac{1}{\theta}. \end{cases}$$

The optimal portfolio α^* is thus determined along with the threshold κ^* by solving a system of $n+1$ equations in $n+1$ unknowns. Although it is not generally possible to find explicitly a solution of the above system of equations, numerical calculation with a standard equation solver is straightforward: given an initial guess (α, κ), one is able to calculate the first two moments of the conditional distribution of wealth; if the moments thus calculated, together with the initial guess (α, κ), satisfy the system of equations, then $(\alpha, \kappa) = (\alpha^*, \kappa^*)$ and numerical search stops; otherwise, the search procedure continues with a new initial guess for the parameters.[4] In the next Section we will solve in this way few simple portfolio problems in order to illustrate some features of the model.

The optimal allocation rule of Theorem 3 is easily compared to the rule that would obtain in a classic Markowitz's setting. In the traditional mean-variance model we would have:

$$\alpha^* = \frac{1}{\theta}\left[\mathrm{Var}^P\left[X\right]\right]^{-1}\mathrm{E}^P\left[\left(X - \vec{1}R\right)\right].$$

Provided $\mathrm{Var}^P\left[X\left|W_\alpha \le \kappa\right.\right]$ is invertible, the first n equations can be written as:

$$\alpha^* = \frac{1}{\theta P\left(W_{\alpha^*} \le \kappa^*\right)}\mathrm{Var}^P\left[X\left|W_{\alpha^*} \le \kappa^*\right.\right]^{-1}\mathrm{E}^P\left[X - \vec{1}R\left|W_{\alpha^*} \le \kappa^*\right.\right],$$

which is easily compared to Markowitz's optimal allocation. The unconditional mean and variance of the vector of returns X are replaced by a conditional mean and a conditional variance, both calculated by conditioning on

[3] $E^P\left[\cdot\left|W_\alpha \le \kappa\right.\right]$ and $\mathrm{Var}^P\left[\cdot\left|W_\alpha \le \kappa\right.\right]$ are the expectation and variance conditional on the event $\{W_\alpha \le \kappa\}$. Note that $\mathrm{Var}^P\left[\cdot\left|W_\alpha \le \kappa\right.\right]$ is an $(n \times n)$ matrix.

[4] A R (S-Plus) routine to calculate the optimal portfolio in an economy with finitely many states of nature and assets is available upon request from the authors.

the event $\{W_{\alpha^*} \leq \kappa^*\}$. Furthermore a scaling factor is introduced, which is inversely proportional to the probability of exceeding the threshold κ^*. To close the system, an extra equation is added to determine the threshold κ^*.

Roughly speaking, when computing the optimal portfolio we ignore that part of the distribution of X where wealth is higher than κ^*. To see why it is optimal to ignore the part of the distribution of X where one obtains the highest returns, recall the example of non-monotonicity of mean-variance illustrated in the Introduction. In that example, high payoffs were increasing the variance more than the mean, thus leading the mean-variance agent to prefer a strictly smaller prospect, contrary to what would commonly be reputed economically reasonable. With monotone mean-variance preferences, this kind of behavior is avoided by artificially setting the probability of some high payoff states equal to zero. In our portfolio selection problem we set the probability of the event $\{W_{\alpha^*} > \kappa^*\}$ equal to zero.

When there is only one risky asset, the optimal quantity of it to be held according to our model and according to the mean-variance model can be compared by means of the following result. Here α^*_{mmv} denotes the optimal portfolio according to the monotone mean-variance model and α^*_{mv} the optimal portfolio according to the mean-variance model.

Proposition 4 *Suppose that S is finite, with $P(s) > 0$ for all $s \in S$, and that there is only one risky asset. Then, either*

$$\alpha^*_{mmv} \geq \alpha^*_{mv} \geq 0$$

or

$$\alpha^*_{mmv} \leq \alpha^*_{mv} \leq 0.$$

If, in addition, $P\left(W_{\alpha^*_{mmv}} > \kappa^*\right) > 0$, then:

$$\begin{array}{lcll} \alpha^*_{mmv} & > & \alpha^*_{mv} & \text{if } \alpha^*_{mmv} > 0, \\ \alpha^*_{mmv} & < & \alpha^*_{mv} & \text{if } \alpha^*_{mmv} < 0. \end{array}$$

In the presence of only one risky asset, an investor with monotone mean-variance preferences always holds a portfolio which is more aggressive than the portfolio held by a mean-variance investor. If she buys a positive quantity of the risky asset, this is greater than or equal to the quantity that would be bought by a mean-variance investor; on the contrary, if she sells the risky asset short, she sells more or as much as a mean-variance investor would do. This kind of behavior will be thoroughly illustrated by the examples in the next section: the intuition behind it is that in some cases a favorable investment opportunity is discarded by a mean-variance investor because of non-monotonicity of her preferences, while a monotone mean-variance investor exploits the opportunity, thus taking a more leveraged position.

5 Some Examples

In this section we present three simple examples to illustrate the optimal portfolio rule we derived above. In every example there are five possible states of Nature. Each of them obtains with a probability $P(s_i)$ that remains fixed throughout the examples. In all examples we also set $\theta = 10$.

Example 1 is a case in which our model and the traditional mean-variance model deliver the same optimal composition of the portfolio. This is not interesting *per se*, but it serves as a benchmark and it helps to introduce Example 2, where the two optimal portfolios differ. In Example 1 there is only one risky asset, whose gross return is denoted by X_1 and is reported in the next table, and a risk-free asset, whose gross return R is equal to 1 across all states. In this example, the optimal portfolio α^*_{mmv} calculated according to to our rule is equal to the mean-variance optimal portfolio α^*_{mv}. W_{mv} and W_{mmv} represent the overall return to the two optimal portfolios for each state of the world. The table also displays the value of the constant κ^* at which it is optimal to truncate the distribution of the return to the portfolio of risky assets.

	$P(s_i)$	$P(s_i \mid W_{mmv} \leq \kappa^*)$	R	X_1	W_{mv}	W_{mmv}
s_1	0.1	0.1	1	0.97	0.9375	0.9375
s_2	0.2	0.2	1	0.99	0.9791	0.9791
s_3	0.4	0.4	1	1.01	1.0208	1.0208
s_4	0.2	0.2	1	1.03	1.0620	1.0620
s_5	0.1	0.1	1	1.05	1.1041	1.1041
$\alpha^*_{mv} = 2.083$						
$\alpha^*_{mmv} = 2.083$			$\kappa^* = 1.1211$			

Example 2 is a slight modification of Example 1. We increase the payoff to the risky asset in state s_5 from 1.05 to 1.10, leaving everything else unchanged. The effect of this change is an increase in both the mean and the variance of X_1, the payoff to the risky asset. The optimal behavior according to the mean-variance model is to reduce the fraction of wealth invested in the risky asset from 2.083 to 1.3574. In contrast, according to our model it is also optimal to decrease the position in the risky asset, but less, from 2.083 to 1.8382.

	$P(s_i)$	$P(s_i \mid W_{mmv} \leq \kappa^*)$	R	X_1	W_{mv}	W_{mmv}
s_1	0.1	0.1111	1	0.97	0.9592	0.9448
s_2	0.2	0.2222	1	0.99	0.9864	0.9816
s_3	0.4	0.4444	1	1.01	1.0135	1.0183
s_4	0.2	0.2222	1	1.03	1.0407	1.0551
s_5	0.1	0	1	1.10	1.1357	1.1838
$\alpha^*_{mv} = 1.3574$						
$\alpha^*_{mmv} = 1.8382$			$\kappa^* = 1.1213$			

In both cases the optimal behavior might seem puzzling at a first sight: when the payoff of an asset increases in one state, it is optimal to hold less of that asset. This behavior can be understood by looking at the distributions of the overall return in the two tables. By reducing the fraction of wealth invested in the risky asset, the overall return increases in the states where the risky asset pays less than the risk-free asset. On the contrary, the overall return decreases in the states where the risky asset pays more than the risk-free asset. In state s_5, however, the effect of this decrease is compensated by the fact that we have raised the payoff to the risky asset from 1.05 to 1.10. Hence, by reducing the amount of wealth invested in the risky asset, the investor gives up some of the extra payoff received in state s_5 in order to guarantee himself a higher overall return in the states where the risky asset has a low payoff. The problem with this kind of behavior is that it can become pathological with mean-variance preferences. The next table shows what happens if we further increase the payoff in state s_5.

$X_1(s_5)$	1.05	1.10	1.15	1.20	1.50	2	3
α^*_{mv}	2.0830	1.3574	0.9174	0.6747	0.2465	0.1175	0.0572
α^*_{mmv}	2.0830	1.8382	1.8382	1.8382	1.8382	1.8382	1.8382

The more we increase the payoff in state s_5, the more the mean-variance optimal fraction α^*_{mv} of wealth invested in the risky asset decreases, until it goes to zero when the payoff in state s_5 becomes very large. In our model this does not happen. At first α^*_{mmv} decreases, but it then stops to decrease and it remains fixed at the same value, though the payoff in state s_5 is further increased. The reason why this happens is that, once probabilities have been optimally reassigned to states and a zero probability has been assigned to state s_5, any further increases of the payoff in s_5 are disregarded and have no influence on the formation of the optimal portfolio.

Example 3 is slightly more complicated. Everything is as in Example 2, but a second risky asset is added. The payoff to this new asset, denoted by X_2, is high in the states where X_1 is low and low where X_1 is high.

	$P(s_i)$	$P(s_i \vert W_{mmv} \leq \kappa^*)$	R	X_1	X_2	W_{mv}	W_{mmv}
s_1	0.1	0.1111	1	0.97	1.05	1.002	1.0231
s_2	0.2	0.2222	1	0.99	1.00	0.9833	0.9570
s_3	0.4	0.4444	1	1.01	0.99	1.0061	1.0125
s_4	0.2	0.2222	1	1.03	0.99	1.0393	1.0985
s_5	0.1	0	1	1.10	0.99	1.1556	1.3994
$\alpha^*_{mv} = (1.6613, 1.0495)$							
$\alpha^*_{mmv} = (4.2989, 3.0423)$			$\kappa^* = 1.1316$				

Also in this case the optimal portfolios suggested by our model and by the traditional model are different. To get an intuitive idea of what is happening, note that, although the market is still arbitrage-free, asset 2 allows to hedge away almost completely the risks taken by investing in asset 1. Consider for example a portfolio formed by 0.5 units of asset 1 and 0.5 units of asset 2. Its payoffs in the five states are collected in the following vector:

$$(1.01, 0.995, 1, 1.01, 1.045)$$

A qualitative inspection of this payoff vector reveals that in state s_2 this portfolio pays off slightly less than the risk-free asset, while in all other states it pays off more and in some states considerably more. Roughly speaking, if it was not for the slightly low payoff in state s_2, there would be an arbitrage opportunity because the portfolio would pay off more than the risk-free asset in every state. As a consequence, we would expect an optimal portfolio rule to exploit this favorable configuration of payoffs by prescribing to take a highly levered position. As reported in the last table, according to our model it is optimal to take a highly levered position in the risky assets in order to exploit this opportunity, at the cost of facing a low payoff in state s_2. In contrast, with the mean-variance model the optimal portfolio is much less aggressive, and the investor is overly concerned with the unique state in which the payoff is lower than the payoff to the risk-free asset.

A Appendix: Proofs

For all $f \in \mathcal{L}(P)$ define $g_f : \mathbb{R} \to \mathbb{R}$ by

$$g_f(z) = zP(f \le z) - \int f 1_{\{f \le z\}} dP \quad \forall z \in \mathbb{R}.$$

Notice that

$$\begin{aligned} g_f(z) &= zP(f \le z) - \int_{\{f \le z\}} f dP \\ &= zP(f < z) + zP(f = z) - \int_{\{f<z\}} f dP - \int_{\{f=z\}} f dP \\ &= zP(f < z) + zP(f = z) - \int_{\{f<z\}} f dP - zP(f = z) \\ &= zP(f < z) - \int_{\{f<z\}} f dP \end{aligned}$$

for all $z \in \mathbb{R}$.

Lemma 5 *For all $z \in \mathbb{R}$, $g_f(z) = \int_{-\infty}^{z} F_f(t)\, dt$.*[5]

Proof. Choose $z \in \mathbb{R}$. Since $f 1_{\{f \le z\}} = (f \wedge z) - z 1_{\{f>z\}}$, then

$$\begin{aligned} g_f(z) &= zP(f \le z) - \int f 1_{\{f \le z\}} dP = zP(f \le z) - \int \left(f \wedge z - z 1_{\{f>z\}}\right) dP \\ &= zP(f \le z) - \int (f \wedge z)\, dP + z \int 1_{\{f>z\}} dP \\ &= zP(f \le z) + zP(f > z) - \int (f \wedge z)\, dP = \int z - (f \wedge z)\, dP, \end{aligned}$$

that is

$$g_f(z) = \int z - (f \wedge z)\, dP. \tag{5}$$

Observe that $z - (f \wedge z) \ge 0$, and so

$$\int z - (f \wedge z)\, dP = \int_0^{\infty} P(z - (f \wedge z) \ge u)\, du.$$

[5] $F_f : \mathbb{R} \to [0, 1]$ is the cumulative distribution function given by $F_f(t) = P(f \le t)$ for each $t \in \mathbb{R}$.

On the other hand, $\{z - (f \wedge z) \geq u\} = \{f \leq z - u\}$ for all $u > 0$. In fact,

$$z - (f \wedge z) \geq u \Rightarrow (f \wedge z) \leq z - u < z \Rightarrow (f \wedge z) = f \Rightarrow f \leq z - u$$

and

$$f \leq z - u \Rightarrow f \wedge z \leq (z - u) \wedge z \Rightarrow f \wedge z \leq z - u \Rightarrow z - (f \wedge z) \geq u.$$

Hence,

$$\begin{aligned} g_f(z) &= \int_0^\infty P(z - (f \wedge z) \geq u)\, du \\ &= \int_0^\infty P(f \leq z - u)\, du = \int_{-\infty}^z P(f \leq t)\, dt, \end{aligned}$$

as desired. ■

Corollary 6 *For all $z \in \mathbb{R}$,*

$$g_f(z) = \int_{-\infty}^z P(f < t)\, dt, \quad \forall z \in \mathbb{R}.$$

Proof. Notice that $P(f < t) = \lim_{u \to t^-} P(f \leq u) \neq P(f \leq t)$ for at most a countably many ts. ■

Lemma 7 *The function g_f is continuous on $\mathbb{R}$, and*

$$F_f(z) = \lim_{\varepsilon \to 0^+} \left[\frac{g_f(z + \varepsilon) - g_f(z)}{\varepsilon} \right], \quad \forall z \in \mathbb{R}.$$

That is, F_f is the right derivative of g_f, and $F_f(z)$ is the derivative of g_f at every point z at which F_f is continuous. Moreover, setting $\zeta = \operatorname{essinf}(f)$, g_f is strictly increasing on (ζ, ∞), $g_f \equiv 0$ on $(-\infty, \zeta]$,[6] *$\lim_{z \to \zeta^+} g_f(z) = 0^+$, and $\lim_{z \to \infty} g_f(z) = \infty$.*

Proof. The Fundamental Theorem of Calculus guarantees the continuity and derivability properties of g_f. Recall that

$$\operatorname{essinf}(f) = \sup\{\alpha \in \mathbb{R} : P(f < \alpha) = 0\}.$$

[6] With the convention $(-\infty, -\infty] = \varnothing$.

If $z \in \mathbb{R}$ and $z \leq \operatorname{essinf}(f)$, for all $t < z$ there exists $\alpha > t$ such that $P(f < \alpha) = 0$. Then,

$$0 \leq P(f < t) \leq P(f < \alpha) = 0.$$

This implies $g_f(z) = 0$ for all $z \in (-\infty, \zeta]$.

On the other hand, if $\zeta < z < z'$, then

$$g_f(z') - g_f(z) = \int_z^{z'} P(f < t)\, dt \geq P(f < z)(z' - z).$$

But $P(f < z) = 0$ would imply $z \leq \zeta$, a contradiction. Therefore, $g_f(z') - g_f(z) > 0$. That is, g_f is strictly increasing on (ζ, ∞).

Notice that $\lim_{t\to\infty} F_f(t) = 1$. Then, for all $n > 1$ there exists $k \geq 1$ $(n, k \in \mathbb{N})$ such that $F_f(t) > 1 - \frac{1}{n}$ for all $t \geq k$. Therefore,

$$\begin{aligned} g_f(k+n) &= \int_{-\infty}^{k+n} F_f(t)\, dt \geq \int_k^{k+n} F_f(t)\, dt \\ &\geq n\left(1 - \frac{1}{n}\right) = n - 1. \end{aligned}$$

Since g_f is increasing on $\mathbb{R}$, then $\lim_{z\to\infty} g_f(z) = \infty$.

If $\zeta > -\infty$, $g_f(\zeta) = 0$, continuity and nonnegativity imply $\lim_{z\to\zeta^+} g_f(z) = 0^+$. Let $\zeta = -\infty$. By (5),

$$\begin{aligned} g_f(-n) &= \int -n - (f \wedge (-n))\, dP = \int ((-f) \vee n) - n dP \\ &= \int (-f) - ((-f) \wedge n)\, dP. \end{aligned}$$

The Monotone Convergence Theorem guarantees that $\lim_{n\to\infty} g_f(-n) = 0$. Monotonicity and nonnegativity imply $\lim_{z\to\zeta^+} g_T(z) = 0^+$. ■

For the rest of the Appendix we will equivalently write E^P or just E.

Lemma 8 *Let $f \in \mathcal{L}^2(P) - \mathcal{G}_\theta$ and $t \in \mathbb{R}$. Then*

$$f \wedge t \in \mathcal{G}_\theta \Leftrightarrow g_f(t) \leq \frac{1}{\theta}. \tag{6}$$

Proof. Notice that

$$\begin{aligned} f\wedge t-\mathrm{E}[f\wedge t] &= f1_{\{f\leq t\}}+t1_{\{f>t\}}-tP(f>t)-\mathrm{E}\left[f1_{\{f\leq t\}}\right]= \\ &= f1_{\{f\leq t\}}+t1_{\{f>t\}}-t+tP(f\leq t)-\mathrm{E}\left[f1_{\{f\leq t\}}\right]= \\ &= (f-t)\,1_{\{f\leq t\}}+g_f(t)\,. \end{aligned}$$

Since $(f-t)\,1_{\{f\leq t\}}\leq 0$,

$$g_f(t)\leq\frac{1}{\theta}\Rightarrow f\wedge t-\mathrm{E}[f\wedge t]\leq\frac{1}{\theta},$$

i.e., $g_f(t)\leq\frac{1}{\theta}\Rightarrow f\wedge t\in\mathcal{G}_\theta$.

For the converse implication, note that we are assuming $f\notin\mathcal{G}_\theta$. Then, being $f\wedge t\in\mathcal{G}_\theta$, it cannot be $f\wedge t=f$ a.s. Hence, $\mathrm{essup}(f\wedge t)=t$. It follows that:

$$\begin{aligned} f\wedge t-\mathrm{E}[f\wedge t]\leq\frac{1}{\theta}\text{ a.s.} &\Rightarrow \mathrm{essup}\,(f\wedge t)-tP(f>t)-\mathrm{E}\left[f1_{\{f\leq t\}}\right]\leq\frac{1}{\theta} \\ &\Rightarrow t-tP(f>t)-\mathrm{E}\left[f1_{\{f\leq t\}}\right]\leq\frac{1}{\theta} \\ &\Rightarrow tP(f\leq t)-\mathrm{E}\left[f1_{\{f\leq t\}}\right]\leq\frac{1}{\theta} \\ &\Rightarrow g_f(t)\leq\frac{1}{\theta}, \end{aligned}$$

i.e., $f\wedge t\in\mathcal{G}_\theta\Rightarrow g_f(t)\leq\frac{1}{\theta}$. ■

Lemmas 7 and 8 immediately yield the following:

Corollary 9 *Let $f\in\mathcal{L}^2(P)-\mathcal{G}_\theta$ and $t\in\mathbb{R}$. Then*

$$g_f^{-1}\left(\frac{1}{\theta}\right)=\max\{t\in\mathbb{R}:f\wedge t\in\mathcal{G}_\theta\}\,.$$

Theorem 10 *Let $f\in\mathcal{L}^2(P)$. Then*

$$V_\theta(f)=\begin{cases}\mathrm{E}[f]-\frac{\theta}{2}\mathrm{Var}[f] & \text{if } f\in\mathcal{G}_\theta,\\ \mathrm{E}[f\wedge\kappa]-\frac{\theta}{2}\mathrm{Var}[f\wedge\kappa] & \text{else,}\end{cases}$$

where $\kappa=g_f^{-1}\left(\frac{1}{\theta}\right)$. Moreover, the Gateaux differential of V_θ at f is

$$\nabla V_\theta(f)=\theta(\kappa-f)\,1_{\{f\leq\kappa\}}.$$

Proof. Let $f \in \mathcal{L}^2(P)$. [MMR-2] shows that

$$V_\theta(f) = \min_{Q \in \Delta^2(P)} \left\{ \mathrm{E}^Q(f) + \frac{1}{2\theta} C(Q \| P) \right\}.$$

That is, $V_\theta(f)$ is the value of the problem:

$$\begin{cases} \min \left\{ \mathrm{E}(fY) + \frac{1}{2\theta}\mathrm{E}(Y^2) - \frac{1}{2\theta} \right\} \\ Y \geq 0 \\ E(Y) = 1 \end{cases} . \tag{7}$$

They also observe that the solution of such problem exists, is unique, and it coincides with the Gateaux derivative of V_θ at f. Notice that Y is a solution of problem (7) if and only if it is a solution of the constrained optimization problem:

$$\begin{cases} \min \left\{ \mathrm{E}(fY) + \frac{1}{2\theta}\mathrm{E}(Y^2) \right\} \\ Y \geq 0 \\ \mathrm{E}(Y) = 1 \end{cases}$$

The Lagrangian is

$$L(Y, \mu, \lambda) = \mathrm{E}(fY) + \frac{1}{2\theta}\mathrm{E}\left(Y^2\right) - \mathrm{E}(\mu Y) - \lambda(\mathrm{E}(Y) - 1),$$

with $\mu \in \mathcal{L}^2_+(P)$, $\lambda \in \mathbb{R}$. The Karush-Kuhn-Tucker optimality conditions are:

$$\begin{aligned} & f + \frac{1}{\theta}Y - \mu - \lambda = 0 \quad P\text{-a.s.} \\ & \mathrm{E}(\mu Y) = 0 \\ & Y \geq 0, \mu \geq 0 \quad P\text{-a.s.} \\ & \mathrm{E}(Y) = 1 \end{aligned}$$

Since $\mu, Y \geq 0$, they are equivalent to:

$$\begin{aligned} & f + \frac{1}{\theta}Y - \mu - \lambda = 0 \quad P\text{-a.s.} \\ & \mu Y = 0 \quad P\text{-a.s.} \\ & Y \geq 0, \mu \geq 0 \quad P\text{-a.s.} \\ & \mathrm{E}(Y) = 1 \end{aligned}$$

that is,

$$f+\frac{1}{\theta}Y-\lambda\geq 0\quad P\text{-a.s.}\tag{8}$$

$$\left[f+\frac{1}{\theta}Y-\lambda\right]Y=0\quad P\text{-a.s.}\tag{9}$$

$$Y\geq 0\quad P\text{-a.s.}\tag{10}$$

$$\mathrm{E}\left[Y\right]=1\tag{11}$$

Assume (Y^*,λ^*) satisfy (8) - (11). W.l.o.g. we can assume that (8) - (10) are satisfied everywhere (not only P-a.s.).

If $s\in\{Y^*>0\}$, then by (9) $f(s)+\frac{1}{\theta}Y^*(s)-\lambda^*=0$ and

$$Y^*(s)=\theta\left(\lambda^*-f(s)\right).\tag{12}$$

In particular, $\lambda^*-f(s)>0$, and $s\in\{f<\lambda^*\}$. Conversely, if $s\in\{f<\lambda^*\}$, then by (8) $Y^*(s)\geq\theta\left(\lambda^*-f(s)\right)>0$ and $s\in\{Y^*>0\}$. In sum,

$$\begin{array}{l}\{Y^*>0\}=\{f<\lambda^*\}\ \text{ and}\\ Y^*=\theta\left(\lambda^*-f\right)1_{\{f<\lambda^*\}}\,.\end{array}$$

By (11),

$$\begin{aligned}1&=\mathrm{E}\left[Y^*\right]=\mathrm{E}\left[\theta\left(\lambda^*-f\right)1_{\{f<\lambda^*\}}\right]\\&=\theta\left(\lambda^*P\left(f<\lambda^*\right)-\mathrm{E}\left[f1_{\{f<\lambda^*\}}\right]\right),\end{aligned}$$

that is,

$$g_f\left(\lambda^*\right)=\lambda^*P\left(f<\lambda^*\right)-\mathrm{E}\left[f1_{\{f<\lambda^*\}}\right]=\frac{1}{\theta}.$$

In other words

$$\lambda^*=g_f^{-1}\left(\frac{1}{\theta}\right)\equiv\kappa,\tag{13}$$

and λ^* is unique. A fortiori, Y^* is unique and

$$Y^*=\theta\left(\kappa-f\right)1_{\{f<\kappa\}}.\tag{14}$$

By construction, the pair (Y^*,λ^*) defined by (13) and (14) is a solution of (8) - (11). Since the solution of (7) exists and it is unique, we conclude that Y^* defined as in Eq. (14) is the unique solution of (7).

Notice that $Y^* = \theta(\kappa - f) 1_{\{f<\kappa\}} + \theta(\kappa - f) 1_{\{f=\kappa\}} = \theta(\kappa - f) 1_{\{f\le\kappa\}}$,

$$(Y^*)^2 = \theta^2 \left(f^2 1_{\{f\le\kappa\}} + \kappa^2 1_{\{f\le\kappa\}} - 2\kappa f 1_{\{f\le\kappa\}}\right)$$

and

$$\mathrm{E}\left[(Y^*)^2\right] = \theta^2 \left(\int f^2 1_{\{f\le\kappa\}} dP + \kappa^2 P(f\le\kappa) - 2\kappa \int f 1_{\{f\le\kappa\}} dP\right).$$

Moreover,

$$\begin{aligned}\mathrm{E}[fY^*] &= \mathrm{E}\left[f\theta(\kappa - f) 1_{\{f\le\kappa\}}\right] = \mathrm{E}\left[\theta\kappa f 1_{\{f\le\kappa\}} - \theta f^2 1_{\{f\le\kappa\}}\right] \\ &= \theta\kappa \int f 1_{\{f\le\kappa\}} dP - \theta \int f^2 1_{\{f\le\kappa\}} dP.\end{aligned}$$

Therefore,

$$\begin{aligned}V_\theta(f) &= \mathrm{E}[fY^*] + \frac{1}{2\theta}\mathrm{E}\left[(Y^*)^2\right] - \frac{1}{2\theta} \\ &= \theta\kappa \int f 1_{\{f\le\kappa\}} dP - \theta \int f^2 1_{\{f\le\kappa\}} dP + \\ &\quad + \frac{\theta}{2}\left(\int f^2 1_{\{f\le\kappa\}} dP + \kappa^2 P(f\le k) - 2\kappa \int f 1_{\{f\le\kappa\}} dP\right) - \frac{1}{2\theta} \\ &= -\frac{\theta}{2}\int f^2 1_{\{f\le\kappa\}} dP + \frac{\theta}{2}\kappa^2 P(f\le k) - \frac{1}{2\theta}\end{aligned}$$

Also observe that $f 1_{\{f\le\kappa\}} + \kappa 1_{\{f>\kappa\}} = f \wedge \kappa$, whence

$$\begin{aligned}\mathrm{E}[f\wedge\kappa] &= \mathrm{E}\left[f 1_{\{f\le\kappa\}}\right] + \kappa P(f>\kappa) = \mathrm{E}\left[f 1_{\{f\le\kappa\}}\right] - \kappa P(f\le\kappa) + \kappa \\ &= -g_f(\kappa) + \kappa = \kappa - \frac{1}{\theta},\end{aligned}$$

and

$$\begin{aligned}\mathrm{Var}[f\wedge\kappa] &= \mathrm{E}\left[\left(f 1_{\{f\le\kappa\}} + \kappa 1_{\{f>\kappa\}}\right)^2\right] - \left(\kappa - \frac{1}{\theta}\right)^2 = \\ &= \int f^2 1_{\{f\le\kappa\}} dP + \kappa^2 P(f>k) - \kappa^2 - \frac{1}{\theta^2} + 2\frac{\kappa}{\theta} \\ &= \int f^2 1_{\{f\le\kappa\}} dP - \kappa^2 P(f\le k) - \frac{1}{\theta^2} + 2\frac{\kappa}{\theta}.\end{aligned}$$

Finally

$$
\begin{aligned}
\mathrm{E}[f\wedge\kappa]-\frac{\theta}{2}\mathrm{Var}[f\wedge\kappa] &= \kappa-\frac{1}{\theta}-\frac{\theta}{2}\left(\int_S f^2 1_{\{f\leq\kappa\}}dP-\kappa^2 P(f\leq k)-\frac{1}{\theta^2}+2\frac{\kappa}{\theta}\right)\\
&= \kappa-\frac{1}{\theta}-\left(\frac{\theta}{2}\int_S f^2 1_{\{f\leq\kappa\}}dP-\frac{\theta}{2}\kappa^2 P(f\leq k)-\frac{1}{2\theta}+\kappa\right)\\
&= -\frac{\theta}{2}\int_S f^2 1_{\{f\leq\kappa\}}dP+\frac{\theta}{2}\kappa^2 P(f\leq k)-\frac{1}{2\theta}\\
&= V_\theta(f).
\end{aligned}
$$

■

Proof of Theorem 2. It is now enough to combine Corollary 9 and Theorem 10. ■

Remark 1 Notice that:

- Inspection of the proof of Theorem 10 shows that for all $f\in\mathcal{L}^2(P)$ (not only for $f\in\mathcal{L}^2(P)-\mathcal{G}_\theta$), setting $\kappa=g_f^{-1}\left(\frac{1}{\theta}\right)$ we have

$$
\begin{aligned}
V_\theta(f) &= \mathrm{E}[f\wedge\kappa]-\frac{\theta}{2}\mathrm{Var}[f\wedge\kappa],\\
\nabla V_\theta(f) &= \theta(\kappa-f)1_{\{f\leq\kappa\}},\\
\{\nabla V_\theta(f)\} &= \arg\min_{Q\in\Delta^2(P)}\left\{\mathrm{E}^Q(f)+\frac{1}{2\theta}C(Q\|P)\right\}.
\end{aligned}
$$

- The properties of g_f guarantee that κ exists is unique and $\frac{1}{\theta}=\kappa P(f\leq\kappa)-\int f1_{\{f\leq\kappa\}}dP$; therefore, $P(f\leq\kappa)>0$.
- Moreover, $\frac{1}{\theta}=\kappa P(f\leq\kappa)-\int f1_{\{f\leq\kappa\}}dP$ implies $\frac{1}{\theta P(f\leq\kappa)}+\int f dP_{\{f\leq\kappa\}}=\kappa$, and so

$$
\begin{aligned}
\nabla V_\theta(f) &= \theta(\kappa-f)1_{\{f\leq\kappa\}}=\left(\frac{1}{P(f\leq\kappa)}+\theta\int_S f dP_{\{f\leq\kappa\}}-\theta f\right)1_{\{f\leq\kappa\}}\\
&= \left(\frac{1}{P(f\leq\kappa)}-\theta\left(f-\mathrm{E}[f|f\leq\kappa]\right)\right)1_{\{f\leq\kappa\}}.
\end{aligned}
$$

▲

Proof of Theorem 3. The maximization problem is

$$\sup_{\alpha\in\mathbb{R}} V_\theta\left(W_\alpha\right)$$

where $W_\alpha = R+\alpha\left(X-R\right)$ (remember that $\alpha\in\mathbb{R}^n$, $X\in\mathcal{L}^2\left(P\right)^n$, and $\vec{1}$ is a vector of 1s). From theorem (10) we know that V_θ is Gateaux differentiable and

$$\nabla V_\theta\left(W_\alpha\right) = \left(\frac{1}{P\left(W_\alpha\le\kappa_\alpha\right)} - \theta\left(W_\alpha - \mathrm{E}\left[W_\alpha|W_\alpha\le\kappa_\alpha\right]\right)\right)1_{\{W_\alpha\le\kappa_\alpha\}},$$

where κ_α solves:

$$P\left(W_\alpha\le\kappa_\alpha\right)\left(\kappa_a - \mathrm{E}^P\left[W_\alpha\left|W_\alpha\le\kappa_\alpha\right.\right]\right) = \frac{1}{\theta}.$$

Since for all $i=1,\ldots,n$

$$\frac{\partial V_\theta\left(W_\alpha\right)}{\partial\alpha_i} = \mathrm{E}\left[\nabla V_\theta\left(W_\alpha\right)\left(X_i-R\right)\right],$$

the first order conditions for an optimum are:

$$\mathrm{E}\left[X\nabla V_\theta\left(W_\alpha\right)\right] = \vec{1}R.$$

Substituting $\nabla V_\theta\left(W_\alpha\right)$:

$$\mathrm{E}\left[\frac{1_{\{W_\alpha\le\kappa_\alpha\}}}{P\left(W_\alpha\le\kappa_\alpha\right)}X - \theta\left(\left(\alpha X\right)1_{\{W_\alpha\le\kappa_\alpha\}}X - \mathrm{E}\left[\alpha X\left|W_\alpha\le\kappa_\alpha\right.\right]1_{\{W_\alpha\le\kappa_\alpha\}}X\right)\right] = \vec{1}R$$

set $A=\{W_\alpha\le\kappa_\alpha\}$ to obtain

$$\mathrm{E}\left[X\left|A\right.\right] - \theta P\left(A\right)\left(\mathrm{E}\left[\left(\alpha X\right)X\left|A\right.\right] - \mathrm{E}\left[X\left|A\right.\right]\mathrm{E}\left[\alpha X\left|A\right.\right]\right) = \vec{1}R$$

the observation that $\mathrm{E}\left[\left(\alpha X\right)X\left|A\right.\right] - \mathrm{E}\left[X\left|A\right.\right]\mathrm{E}\left[\alpha X\left|A\right.\right] = \mathrm{Var}\left[X\left|A\right.\right]\alpha$ yields:

$$\mathrm{E}\left[X-\vec{1}R\left|W_\alpha\le\kappa_\alpha\right.\right] = \theta P\left(W_\alpha\le\kappa_\alpha\right)\mathrm{Var}\left[X\left|W_\alpha\le\kappa_\alpha\right.\right]\alpha.$$

These are the first n equations. The $(n+1)$-th is the equation which determines κ_α:

$$P\left(W_\alpha\le\kappa\right)\left(\kappa - \mathrm{E}\left[W_\alpha\left|W_\alpha\le\kappa\right.\right]\right) = \frac{1}{\theta}.$$

Concavity of V_θ guarantees the sufficiency of the first order condition. ■

Proof of Proposition 4. Set $\alpha^* = \alpha^*_{mmv}$. The maximization problem it solves is

$$\max_{\alpha\in\mathbb{R}} \min_{Y\in\mathbb{Y}} \left(\mathrm{E}\left[\left(R+\alpha\left(X-R\right)\right)Y\right] + \frac{1}{2\theta}\mathrm{E}\left[Y^2\right] - \frac{1}{2\theta} \right) \tag{15}$$

where $\mathbb{Y} = \left\{Y \in \mathbb{R}^S_+ : \mathrm{E}\left[Y\right] = 1\right\}$. Clearly, $\mathbb{Y}$ is convex and compact, and (α^*, Y^*) is a solution of (15) if and only if it is a solution of

$$\max_{\alpha\in\mathbb{R}} \min_{Y\in\mathbb{Y}} G\left(\alpha, Y\right)$$

where $G\left(\alpha, Y\right) = \mathrm{E}\left[\left(R+\alpha\left(X-R\right)\right)Y\right] + \frac{1}{2\theta}\mathrm{E}\left[Y^2\right]$. Moreover, notice that $G : \mathbb{R}\times\mathbb{Y} \to \mathbb{R}$ is continuous, it is affine in α (for each fixed Y) and strictly convex in Y (for each fixed α). Set $v = \max_{\alpha\in\mathbb{R}} \min_{Y\in\mathbb{Y}} G\left(\alpha, Y\right)$, by (a version of) the Min-Max Theorem (e.g. [A, p. 134]) there exists $\bar{Y} \in \mathbb{Y}$ such that

$$v = \sup_{\alpha\in\mathbb{R}} G\left(\alpha, \bar{Y}\right).$$

Moreover,

$$G\left(\alpha^*, \bar{Y}\right) \geq \min_{Y\in\mathbb{Y}} G\left(\alpha^*, Y\right) = G\left(\alpha^*, Y^*\right) = v = \sup_{\alpha\in\mathbb{R}} G\left(\alpha, \bar{Y}\right) \geq G\left(\alpha^*, \bar{Y}\right),$$

therefore, $G\left(\alpha^*, \bar{Y}\right) = \min_{Y\in\mathbb{Y}} G\left(\alpha^*, Y\right)$, strict convexity implies $\bar{Y} = Y^*$. In turn, this yields $\sup_{\alpha\in\mathbb{R}} G\left(\alpha, Y^*\right) = v \neq \infty$ and it cannot be

$$\sup_{\alpha\in\mathbb{R}} \left(R + \alpha\mathrm{E}\left[\left(X-R\right)Y^*\right] + \frac{1}{2\theta}\mathrm{E}\left[\left(Y^*\right)^2\right] \right) = \infty,$$

therefore

$$\mathrm{E}\left[Y^*\left(X-R\right)\right] = 0. \tag{16}$$

Y^* is the solution of problem (7) in the proof of Theorem 10 with $f = R + \alpha^*\left(X-R\right) = W_{\alpha^*}$. Therefore, there exist $\lambda^* \in \mathbb{R}$ and $\mu \in \mathcal{L}^2\left(P\right)$ such that Y^* satisfies the following conditions:

$$R + \alpha^*\left(X-R\right) + \frac{1}{\theta}Y^* - \mu - \lambda^* = 0, \tag{17}$$

$$\mathrm{E}\left[Y^*\right] = 1, \tag{18}$$

$$Y^* \geq 0, \mu \geq 0, \mu Y^* = 0. \tag{19}$$

Taking the expectation of both sides of (17) we obtain:

$$(1-\alpha)R + \alpha^* \mathrm{E}[X] + \frac{1}{\theta}\mathrm{E}[Y^*] - \mathrm{E}[\mu] - \lambda^* = 0 \tag{20}$$

and, subtracting (20) from (17):

$$\alpha^* (X - \mathrm{E}[X]) + \frac{1}{\theta}(Y^* - \mathrm{E}[Y^*]) - (\mu - \mathrm{E}[\mu]) = 0.$$

Rearranging and using (18):

$$Y^* = 1 - \theta\alpha^* (X - \mathrm{E}[X]) + \theta(\mu - \mathrm{E}[\mu]). \tag{21}$$

Multiply both sides by μ, take expectations and use (19) to get:

$$\mathrm{E}[\mu] - \theta\alpha^* \mathrm{Cov}[\mu, X] + \theta \mathrm{Var}[\mu] = 0$$

and, rearranging terms:

$$\theta\alpha^* \mathrm{Cov}[\mu, X] = \mathrm{E}[\mu] + \theta \mathrm{Var}[\mu]. \tag{22}$$

Since $\mu \geq 0$, then $\mathrm{E}[\mu] \geq 0$ thus:

$$\alpha^* = 0 \Rightarrow \mu = 0 \Rightarrow \mathrm{Cov}[\mu, X] = 0 \tag{23}$$
$$\alpha^* > 0 \Rightarrow \mathrm{Cov}[\mu, X] \geq 0 \tag{24}$$
$$\alpha^* < 0 \Rightarrow \mathrm{Cov}[\mu, X] \leq 0 \tag{25}$$

Now, plugging (21) into (16) we obtain:

$$\mathrm{E}[(1 - \theta\alpha^* (X - \mathrm{E}[X]) + \theta(\mu - \mathrm{E}[\mu]))X] = R$$

or:

$$\mathrm{E}[X] - \theta\alpha^* \mathrm{Var}[X] + \theta \mathrm{Cov}[\mu, X] = R$$

which becomes:

$$\alpha^* = \frac{1}{\theta}\frac{\mathrm{E}[X-R]}{\mathrm{Var}[X]} + \frac{\mathrm{Cov}[\mu, X]}{\mathrm{Var}[X]}$$

Recalling that:

$$\alpha^*_{mv} = \frac{1}{\theta}\frac{\mathrm{E}[X-R]}{\mathrm{Var}[X]}$$

we obtain:

$$\alpha^* = \alpha^*_{mv} + \frac{\mathrm{Cov}\left[\mu, X\right]}{\mathrm{Var}\left[X\right]} \tag{26}$$

Using (23) - (25), it is now obvious that:

$$\alpha^* = 0 \Rightarrow \alpha^* = \alpha^*_{mv} = 0 \tag{27}$$

$$\alpha^* > 0 \Rightarrow \alpha^* \geq \alpha^*_{mv} \tag{28}$$

$$\alpha^* < 0 \Rightarrow \alpha^* \leq \alpha^*_{mv} \tag{29}$$

From the proof of Theorem 10 – Eq. (13) – we know that $\lambda^* = g^{-1}_{W_{\alpha^*}}\left(\frac{1}{\theta}\right) = \kappa^*$. Furthermore, if $P\left(W_{\alpha^*} > \kappa^*\right) > 0$, since S is finite, there exists s such that

$$R + \alpha^*\left(X\left(s\right) - R\right) = W_{\alpha^*}\left(s\right) > \kappa^* = \lambda^*,$$

that is $R + \alpha^*\left(X\left(s\right) - R\right) - \lambda^* > 0$. Since $Y^*\left(s\right) \geq 0$, by (17), we have

$$\mu\left(s\right) = R + \alpha^*\left(X\left(s\right) - R\right) - \lambda^* + \frac{1}{\theta}Y^* > 0,$$

and $\mathrm{E}\left[\mu\right] > 0$. Thus, in this case (22) implies $\alpha^*\mathrm{Cov}\left[\mu, X\right] > 0$ and the inequalities in (28) and (29) become strict.

Finally, we want to show that $\alpha^*\alpha^*_{mv} \geq 0$. By contradiction, suppose $\alpha^*\alpha^*_{mv} < 0$. Then, either $\alpha^* > 0$ and $\alpha^*_{mv} < 0$ or $\alpha^* < 0$ and $\alpha^*_{mv} > 0$. Suppose $\alpha^* > 0$ and $\alpha^*_{mv} < 0$, since

$$\alpha^*_{mv} = \frac{1}{\theta}\frac{\mathrm{E}\left[X - R\right]}{\mathrm{Var}\left[X\right]},$$

it must be $\mathrm{E}\left[X - R\right] < 0$ and

$$\alpha^*\mathrm{E}\left[X - R\right] < 0. \tag{30}$$

Clearly, if $\alpha^* < 0$ and $\alpha^*_{mv} > 0$, (30) still holds. Remember that $\left(\alpha^*, Y^*\right)$ is a saddle point for

$$G\left(\alpha, Y\right) = \mathrm{E}\left[\left(R + \alpha\left(X - R\right)\right)Y\right] + \frac{1}{2\theta}\mathrm{E}\left[Y^2\right]$$

and so:

$$G\left(\alpha^*, Y^*\right) \leq G\left(\alpha^*, 1_S\right) = R + \alpha^*\mathrm{E}\left[X - R\right] < R$$

where the last inequality follows from (30). But,

$$\min_{Y \in \mathbb{Y}} G\left(0, Y\right) = G\left(0, 1_S\right) = R > G\left(\alpha^*, Y^*\right) = \max_{\alpha \in \mathbb{R}}\min_{Y \in \mathbb{Y}} G\left(\alpha, Y\right),$$

which is impossible. ■

References

[A] Aubin, J-P., *Optima and equilibria*, Springer-Verlag, Berlin, 1998.

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

[HS] Hansen, L. and T. Sargent, Robust control and model uncertainty, *American Economic Review*, 91, 60-66, 2001.

[MMR-1] Maccheroni, F., M. Marinacci and A. Rustichini, Variational representation of preferences under ambiguity, ICER WP 5/04, 2004

[MMR-2] Maccheroni, F., M. Marinacci and A. Rustichini, 2004, A variational formula for the relative Gini concentration index", mimeo, 2004.

[M] Markowitz, H. M., Portfolio selection, *Journal of Finance*, 7, 77-91, 1952.

INTERNATIONAL CENTRE FOR ECONOMIC RESEARCH

APPLIED MATHEMATICS WORKING PAPER SERIES

1. Luigi Montrucchio and Fabio Privileggi, “On Fragility of Bubbles in Equilibrium Asset Pricing Models of Lucas-Type,” *Journal of Economic Theory* 101, 158-188, 2001 (ICER WP 2001/5).
2. Massimo Marinacci, “Probabilistic Sophistication and Multiple Priors,” *Econometrica* 70, 755-764, 2002 (ICER WP 2001/8).
3. Massimo Marinacci and Luigi Montrucchio, “Subcalculus for Set Functions and Cores of TU Games,” *Journal of Mathematical Economics* 39, 1-25, 2003 (ICER WP 2001/9).
4. Juan Dubra, Fabio Maccheroni, and Efe Ok, “Expected Utility Theory without the Completeness Axiom,” *Journal of Economic Theory*, forthcoming (ICER WP 2001/11).
5. Adriana Castaldo and Massimo Marinacci, “Random Correspondences as Bundles of Random Variables,” April 2001 (ICER WP 2001/12).
6. Paolo Ghirardato, Fabio Maccheroni, Massimo Marinacci, and Marciano Siniscalchi, “A Subjective Spin on Roulette Wheels,” *Econometrica*, forthcoming (ICER WP 2001/17).
7. Domenico Menicucci, “Optimal Two-Object Auctions with Synergies,” July 2001 (ICER WP 2001/18).
8. Paolo Ghirardato and Massimo Marinacci, “Risk, Ambiguity, and the Separation of Tastes and Beliefs,” *Mathematics of Operations Research* 26, 864-890, 2001 (ICER WP 2001/21).
9. Andrea Roncoroni, “Change of Numeraire for Affine Arbitrage Pricing Models Driven By Multifactor Market Point Processes,” September 2001 (ICER WP 2001/22).
10. Maitreesh Ghatak, Massimo Morelli, and Tomas Sjoström, “Credit Rationing, Wealth Inequality, and Allocation of Talent”, September 2001 (ICER WP 2001/23).
11. Fabio Maccheroni and William H. Ruckle, “BV as a Dual Space,” *Rendiconti del Seminario Matematico dell'Università di Padova*, 107, 101-109, 2002 (ICER WP 2001/29).
12. Fabio Maccheroni, “Yaari Dual Theory without the Completeness Axiom,” *Economic Theory*, forthcoming (ICER WP 2001/30).
13. Umberto Cherubini and Elisa Luciano, “Multivariate Option Pricing with Copulas,” January 2002 (ICER WP 2002/5).
14. Umberto Cherubini and Elisa Luciano, “Pricing Vulnerable Options with Copulas,” January 2002 (ICER WP 2002/6).
15. Steven Haberman and Elena Vigna, “Optimal Investment Strategies and Risk Measures in Defined Contribution Pension Schemes,” *Insurance: Mathematics and Economics* 31, 35-69, 2002 (ICER WP 2002/10).
16. Enrico Diecidue and Fabio Maccheroni, “Coherence without Additivity,” *Journal of Mathematical Psychology*, forthcoming (ICER WP 2002/11).
17. Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci, “Ambiguity from the Differential Viewpoint,” April 2002 (ICER WP 2002/17).

18. Massimo Marinacci and Luigi Montrucchio, "A Characterization of the Core of Convex Games through Gateaux Derivatives," *Journal of Economic Theory*, forthcoming (ICER WP 2002/18).
19. Fabio Maccheroni and Massimo Marinacci, "How to Cut a Pizza Fairly: Fair Division with Decreasing Marginal Evaluations," *Social Choice and Welfare*, 20, 457-465, 2003 (ICER WP 2002/23).
20. Erio Castagnoli, Fabio Maccheroni and Massimo Marinacci, "Insurance Premia Consistent with the Market," *Insurance: Mathematics and Economics* 31, 267-284, 2002 (ICER WP 2002/24).
21. Fabio Privileggi and Guido Cozzi, "Wealth Polarization and Pulverization in Fractal Societies," September 2002 (ICER WP 2002/39).
22. Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci, "Certainty Independence and the Separation of Utility and Beliefs," December 2002 (ICER WP 2002/40).
23. Salvatore Modica and Marco Scarsini, "The Convexity-Cone Approach to Comparative Risk and Downside Risk", January 2003 (ICER WP 2003/1).
24. Claudio Mattalia, "Existence of Solutions and Asset Pricing Bubbles in General Equilibrium Models", January 2003 (ICER WP 2003/2).
25. Massimo Marinacci and Luigi Montrucchio, "Cores and Stable Sets of Finite Dimensional Games", March 2003 (ICER WP 2003/7).
26. Jerome Renault, Sergio Scarlatti, and Marco Scarsini, "A Folk Theorem for Minority Games", April 2003 (ICER WP 2003/10).
27. Peter Klibanoff, Massimo Marinacci, and Sujoy Mukerji, "A Smooth Model of Decision Making under Ambiguity", April 2003 (ICER WP 2003/11).
28. Massimo Marinacci and Luigi Montrucchio, "Ultramodular Functions", June 2003 (ICER WP 2003/13).
29. Erio Castagnoli, Fabio Maccheroni, and Massimo Marinacci, "Choquet Insurance Pricing: a Caveat", June 2003 (ICER WP 2003/14).
30. Thibault Gajdos and Eric Maurin, "Unequal Uncertainties and Uncertain Inequalities: an Axiomatic Approach, June 2003 (ICER WP 2003/15).
31. Thibault Gajdos and John A. Weymark, "Multidimensional Generalized Gini Indices", June 2003 (ICER WP 2003/16).
32. Thibault Gajdos, Jean-Marc Tallon, and Jean-Christophe Vergnaud, "Decision Making with Imprecise Probabilistic Information", June 2003 (ICER WP 2003/18).
33. Alfred Müller and Marco Scarsini, "Archimedean Copulae and Positive Dependence", July 2003 (ICER WP 2003/25).
34. Bruno Bassan, Olivier Gossner, Marco Scarsini, and Shmuel Zamir, "Positive Value of Information in Games", *International Journal of Game Theory*, forthcoming (ICER WP 2003/26).
35. Marco Dall'Aglio and Marco Scarsini, "Zonoids, Linear Dependence, and Size-Biased Distributions on the Simplex", *Advances in Applied Probability* forthcoming (ICER WP 2003/27).
36. Taizhong Hu, Alfred Muller, and Marco Scarsini, "Some Counterexamples in Positive Dependence", *Journal of Statistical Planning and Inference*, forthcoming (ICER WP 2003/28).
37. Massimo Marinacci, Fabio Maccheroni, Alain Chateauneuf, and Jean-Marc Tallon, "Monotone Continuous Multiple Priors" August 2003 (ICER WP 2003/30).

38. Massimo Marinacci and Luigi Montrucchio, “Introduction to the Mathematics of Ambiguity”, October 2003 (ICER WP 2003/34).
39. Fabio Maccheroni, Massimo Marinacci and Aldo Rustichini, “Variational representation of preferences under ambiguity”, June 2004 (ICER WP 2004/5).
40. Antonio Lijoi, Ramsés H. Mena, and Igor Prünster, “Hierarchical mixture modelling with normalized Gaussian priors”, November 2004 (ICER WP 2004/12).
41. Stephen G. Walker, Antonio Lijoi, and Igor Prünster, “Contributions to the understanding of Bayesian consistency”, November 2004 (ICER WP 2004/13).
42. Antonio Lijoi, Igor Prünster, and Stephen G. Walker, “On consistency of nonparametric normal mixtures”, October 2004 (ICER WP 2004/23).
43. Stephen G. Walker, Antonio Lijoi, and Igor Prünster, “On rates of convergence for posterior distributions”, October 2004 (ICER WP 2004/24).
44. Fabio Maccheroni, Massimo Marinacci, Aldo Rustichini and Marco Taboga, “Portfolio selection with monotone mean-variance preferences”, December 2004 (ICER WP 2004/27).
45. Fabio Maccheroni and Massimo Marinacci, “A Strong Law of Large Numbers for Capacities”, December 2004 (ICER WP 2004/28).