A Risk Measure For Income Streams

Georg Ch. Pflug¹ and Andrzej Ruszczyński²

Abstract

A new measure of risk is introduced for a sequence of random incomes adapted to some filtration. This measure is formulated as the optimal net present value of a stream of adaptively planned commitments for consumption.

The new measure is calculated by solving a stochastic dynamic linear optimization problem which, for finite filtrations, reduces to a deterministic linear program.

We analyze properties of the new measure by exploiting the convexity and duality structure of the stochastic dynamic linear problem. The measure depends on the full distribution of the income process (not only on its marginal distributions) as well as on the filtration, which is interpreted as the available information about the future. The features of the new approach are illustrated by numerical examples.

¹Department of Statistics and Decision Support Systems, Universitätsstraße 5, University of Vienna, 1090 Wien-Vienna, Austria; e-mail: georg.pflug@univie.ac.at
²Department of Management Science and Information Systems and RUTCOR, Rutgers University, 94 Rockefeller Rd, Piscataway, NJ 08854, U.S.A.; e-mail rusz@rutcor.rutgers.edu
1 Motivation

Since the seminal work of Markowitz it is well understood that consequences of economic activity with uncertain success must be judged in two different and well distinguished dimensions. The mean refers to the average result among a set of possible scenarios, while the risk dimension describes the possible variation of the results under varying scenarios. In the Markowitz model the risk is measured by the variance of the outcome (cf. [8, 9]). In the mean–risk setting the decision maker is faced with a two-objective situation: he/she wants to maximize the mean return and to minimize the risk at the same time. As for all multi-objective situations, there is in general no uniquely defined best decision, which is optimal in both dimensions and one has to seek for compromises. The set of solutions which are Pareto-efficient in the sense of these two objectives is called the mean–risk efficient frontier.

In some models for optimal decision making the two dimensions are often mixed by introducing a nondecreasing concave utility function. Risk aversion, i.e. the degree of taking the risk dimension into account, can be modeled by the negative curvature of the utility function.

However, it is highly desirable to clearly separate the two dimensions and to make the compromising strategy as transparent as possible, and the efficient frontier approach provides such a transparency. In the first step, the efficient frontier is calculated for a given decision problem and the non-dominated decisions are identified. In the second step, the compromise decision may be chosen among the efficient candidates.

There is a vast literature on one-period decision models using several notions of measuring risk (see, e.g., [1, 7, 13, 14, 19]). In the multiperiod situation, however, most proposals focus on the risk contained in the terminal wealth (see [3, 11, 12] and the references therein). The purpose of this paper is to propose a risk measure for multiperiod models which incorporates the risk contained in intermediate incomes.

Suppose that \( I_1, \ldots, I_T \) is a stream of random incomes. A simple but inappropriate way of defining the multiperiod risk would be to look at the marginal variables separately and fabricate a combined risk measure as a combination of the univariate risk measures.

The distinction can be made clear by advocating an example which goes back to Philippe Artzner. Suppose that a coin is thrown three times. In situation 1, a reward of 1 is paid if the coin shows more heads than tails. In situation 2, the same reward is paid if the last throw shows head. Do the two situations reflect the same risk for the decision maker? If the whole experiment is done in a few seconds, one is inclined to say ‘yes’. But suppose that the throw of the coin happens just once a year. Then in situation 1 the decision maker knows her income one year ahead, which is a clear advantage over situation 2. Thus situation 1 should turn out to be less risky than 2, although their income variables have identical marginals. We shall return to this example in section 7.

To valuate the entire income stream and not just the terminal wealth appears to be appropriate in many models. For instance, pension funds promise a income streams to their clients. Since the rights emerging from a pension fund membership are not bequeathable, clients are not interested in the terminal wealth at some future moments of time, at which they may not be able to consume it. At least in Europe, pension funds are
only administrators and not owners of the funds. They are primarily interested in high management fees, which come from a large number of customers. Customers can only be attracted if a good income stream can be guaranteed. Thus it is in the own interest of a pension fund to keep an eye on the customer’s income stream process (see, e.g., [10, 17]).

Besides that, the income stream risk must also be considered in other cases, where the primal investment is just made for the purpose of getting the income at later periods. Real options, or loans to companies are good examples here.

The paper is organized as follows. After an introductory section about one-period measures, we introduce our concept of a multiperiod measure in section 3. Its properties are analyzed in section 4, and section 5 contains explicit linear programming models for the case of finitely many scenarios. In section 6 we consider mean–risk models for our measures and compare them to models based on the terminal wealth distribution. Illustrative examples are contained in section 7. All calculations can be done by standard linear programming packages.

2 The one-period case

Let $I$ be a random income variable defined on some probability space $(\Omega, \mathcal{F}, P)$. The risk contained in $I$ is caused by the lack of information about its exact value. A variable, but predictable value of $I$ is riskless. If a natural catastrophe, e.g. a flood, were completely predictable, there would be no risk and no company would insure against it.

If a decision maker were clairvoyant, he/she would face no risk since he/she would see the future in a deterministic way and would be able to adapt to it. For us, normal humans, some but not all information about the future may be available.

The amount of information available may be expressed in terms of some $\sigma$-algebra $\mathcal{F} \subseteq \mathcal{F}$. The extreme cases are the clairvoyant ($\mathcal{F} = \mathcal{F}$) and the totally uninformed ($\mathcal{F} = \mathcal{F}_0 = \{\Omega, \emptyset\}$).

The ultimate goal of engaging in risky enterprises with uncertain income opportunities is consumption. Consumption, however, can only be realized after deciding about the amount one wants to commit for this purpose (to buy a house, a car etc.).

Suppose that the decision maker decides to commit an amount $a$. In this case, he/she risks not achieving this decided target, since $I$ may be less than $a$. However, he/she may insure against the shortfall event, i.e. the event that $I < a$. Insurance comes at the price of $\mathbb{E}(q|I - a|^-)$, for $q > 1$. The costs for insurance decrease the possible consumption.

If, on the other hand, some surplus is left after consumption, this surplus is discounted by a factor $d < 1$, since saving does not provide the same satisfaction as the consumption committed for.

The Expected Net Present Value (ENPV) of the consumption and savings is therefore

$$\mathbb{E}(a + d[I - a]^+ - q[I - a]^-).$$

A rational decision maker maximizes the ENPV with respect to the available information

\footnote{We use the notation $[x]^+ = \max(x, 0)$ and $[x]^− = \max(−x, 0)$.}
\( \mathcal{F} \); i.e., his/her utility functional is

\[
U_F(I) = \max \{ \mathbb{E}(a + d[I - a]^+ - q[I - a]^-) : a \text{ is } \mathcal{F} \text{ measurable} \}. \tag{1}
\]

It is evident that \( \mathcal{F}_1 \subseteq \mathcal{F}_2 \) implies that \( U_{\mathcal{F}_1}(I) \leq U_{\mathcal{F}_2}(I) \), i.e., more information provides more utility.

Since \( I \) is \( \tilde{\mathcal{F}} \)-measurable and

\[
a + d[v - a]^+ - q[v - a]^- = v - (1 - d)[v - a]^+ - (q - 1)[v - a]^-
\]

is concave in \( a \), one sees that \( \mathcal{U}_\tilde{F}(I) = \mathbb{E}(I) \)

and

\[
\mathcal{U}_F(I) \leq \mathbb{E}(I)
\]

for any other \( \sigma \)-subalgebra \( \mathcal{F} \) of \( \tilde{\mathcal{F}} \).

The risk \( R \) contained in the random variable \( I \) and the information \( \mathcal{F} \) is defined as the difference between the maximal utility (the utility of the clairvoyant) and the actual utility.

\[
R_\mathcal{F}(I) = \mathbb{E}(I) - U_\mathcal{F}(I). \tag{2}
\]

The just defined utility and risk measures are closely related to the notion of the conditional-value-at-risk (CV@R). Recall that the CV@R is defined as

\[
\mathbb{C}V@R_\beta(I) = \max \{ a - \frac{1}{\beta} \mathbb{E}([I - a]^-) : a \in \mathbb{R} \}
\]

(see Rockefellar and Uryasev [18]). We have the following representation (see [15]):

\[
\mathbb{C}V@R_\beta(I) = \frac{1}{\beta} \int_0^\beta G^{-1}(p) \, dp = \mathbb{E}(I | I \leq G^{-1}(\beta)) - \left( \frac{G^{-1}(\beta)}{\beta} \right) G^{-1}(\beta),
\]

where \( G(u) = \mathbb{P}\{ I \leq u \} \) and \( G^{-1}(p) = \inf\{ u : G(u) \geq p \} \). Clearly, \( \mathbb{C}V@R_\beta(I) \leq \mathbb{C}V@R_1(I) = \mathbb{E}(I) \). If \( I \) is a constant, then

\[
\mathbb{C}V@R_\beta(I) = \mathbb{E}(I) = I. \tag{3}
\]

**Lemma 1** Let \( I|\mathcal{F} \) be the conditional distribution of \( I \) given the \( \sigma \)-algebra \( \mathcal{F} \). Then

\[
U_\mathcal{F}(I) = d\mathbb{E}(I) + (1 - d)\mathbb{E}[\mathbb{C}V@R_\beta(I|\mathcal{F})]
\]

and

\[
R_\mathcal{F}(I) = (1 - d)[\mathbb{E}(I) - \mathbb{E}[\mathbb{C}V@R_\beta(I|\mathcal{F})]],
\]

where \( \beta = (1 - d)/(q - d) \).
Recall the following ordering relations for random variables.

If the correlation $\rho \neq 0$, then conditionally on $F$ and taking the expectation afterwards, one gets the general formulas

$$ U_{\beta}(I) = dE(I) + (1 - d)C\nu @R_\beta(I), $$

with $\beta = (1 - d)/(q-d)$. Repeating now the same argument for the conditional distribution of $I$ given $F$ and taking the expectation afterwards, one gets the general formulas

$$ U_{F}(I) = E[dE(I|F) + (1 - d)C\nu @R_\beta(I|F)] = dE(I) + (1 - d)E[C\nu @R_\beta(I|F)] $$

and

$$ R_{F}(I) = (1 - d)[E(I) - E[C\nu @R_\beta(I|F)]. $$

If $I$ is $\mathcal{F}$ measurable, then conditionally on $\mathcal{F}$, $I$ is a constant and therefore by (3), $U_{F}(I) = E(I)$ and $R_{F}(I) = 0$.

**Example.** Suppose that $I$ follows a lognormal distribution and that the available information is contained in a random variable $J$, which is correlated with $I$. To be more precise, assume that $I = \exp(X_1)$, $J = \exp(X_2)$, where

$$ \left( \begin{array}{c} X_1 \\ X_2 \end{array} \right) \sim N \left( \left( \begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right), \left( \begin{array}{cc} \sigma_1^2 & \sigma_1 \sigma_2 \rho \\ \sigma_1 \sigma_2 \rho & \sigma_2^2 \end{array} \right) \right) $$

The $C\nu @R_\beta(Y)$ of a Lognormal$(\mu, \sigma^2)$ distributed $Y$ is

$$ C\nu @R_\beta(Y) = \frac{1}{\beta} e^{\mu + \sigma^2/2} \Phi(\Phi^{-1}(\beta) - \sigma), $$

as can be seen by calculation. Let $\mathcal{F}_J$ be the $\sigma$-algebra generated by $J$. Using the conditional distribution of $I$ given $J$ and the formula $E(I) = e^{\mu_1 + \sigma_1^2/2}$ one gets that

$$ U_{\mathcal{F}_J}(I) = de^{\mu_1 + \sigma_1^2/2} + (q - d)e^{\mu_1 + \sigma_1^2/2} \Phi(\Phi^{-1}(\beta) - \sigma_1 \sqrt{1 - \rho^2}) $$

and

$$ R_{\mathcal{F}_J}(I) = (1 - d)e^{\mu_1 + \sigma_1^2/2} - (q - d)e^{\mu_1 + \sigma_1^2/2} \Phi(\Phi^{-1}(\beta) - \sigma_1 \sqrt{1 - \rho^2}). $$

If the correlation $\rho = 1$, then $R_{\mathcal{F}_J} = 0$ and if $\rho = 0$, then $R_{\mathcal{F}_J} = R_{\mathcal{F}_0}$. Notice that $v \mapsto U_a(v)$ is a concave, nondecreasing utility function for every fixed $a$. Recall the following ordering relations for random variables.
Definition 1 Let \( I^{(1)} \) and \( I^{(2)} \) two random income variables.

- We say that **first order stochastic dominance** \((I^{(1)} \prec_{\text{FSD}} I^{(2)})\) holds, if \( \mathbb{E}[U(I^{(1)})] \leq \mathbb{E}[U(I^{(2)})] \) for all nondecreasing functions \( U \) for which these expected values are finite.

- We say that **second order stochastic dominance** \((I^{(1)} \prec_{\text{SSD}} I^{(2)})\) holds, if \( \mathbb{E}[U(I^{(1)})] \leq \mathbb{E}[U(I^{(2)})] \) for all nondecreasing and concave, integrable functions \( U \) for which these expectations are finite.

- We say that **concave dominance** \((I^{(1)} \prec_{\text{CC}} I^{(2)})\) holds, if \( \mathbb{E}[U(I^{(1)})] \leq \mathbb{E}[U(I^{(2)})] \) for all concave functions \( U \) for which these expectations are finite.

Since all \( U_a \) are nondecreasing and concave by (4), it follows that

\[
\mathcal{U}_F(I) = \max\{\mathbb{E}(U_a(I)) : a \in \mathbb{R}\}
\]

is monotonic w.r.t. second order stochastic dominance \( \prec_{\text{SSD}} \) and a fortiori with first order stochastic dominance \( \prec_{\text{FSD}} \) and concave dominance \( \prec_{\text{CC}} \). By a similar argument, \( \mathcal{R}_F \) is antitonic w.r.t. \( \prec_{\text{CC}} \).

More generally, if \( I^{(1)} \) and \( I^{(2)} \) are defined on the same probability space, and all the conditional distributions satisfy \((I^{(1)}|\mathcal{F}) \prec_{\text{SSD}} (I^{(2)}|\mathcal{F})\), then \( \mathcal{U}_F(I^{(1)}) \leq \mathcal{U}_F(I^{(2)}) \). Similarly, if \((I^{(1)}|\mathcal{F}) \prec_{\text{CC}} (I^{(2)}|\mathcal{F})\), then \( \mathcal{R}_F(I^{(1)}) \geq \mathcal{R}_F(I^{(2)}) \).

It is necessary to require the ordering of all conditional distributions.

**Example 1** Let the probability space have three points, \( \omega_1, \omega_2, \omega_3 \), each having probability \( 1/3 \). Let \( I^{(1)}(\omega_1) = 1.01, I^{(1)}(\omega_2) = 1.015, I^{(1)}(\omega_3) = 1.03; I^{(2)}(\omega_1) = 1.01501, I^{(2)}(\omega_2) = 1.0301, I^{(2)}(\omega_3) = 1.0101 \). Choose \( q = 1.2, d = 0.93 \) and \( \mathcal{F} = \{\{\omega_1, \omega_2\}, \omega_3\} \). Then

\[
\mathcal{U}_F(I^{(1)}) = 1.0175 > \mathcal{U}_F(I^{(2)}) = 1.0174,
\]

but \( I^{(1)} \prec_{\text{SSD}} I^{(2)} \).

Notice that \( \mathcal{U}_F \) is translation-equivariant, i.e. for all constant \( b \)

\[
\mathcal{U}_F(I + b) = \mathcal{U}_F(I) + b.
\]

This follows directly from the definition.

In contrast, \( \mathcal{R}_F(I) \) is translation-invariant, i.e. for all constant \( b \)

\[
\mathcal{R}_F(I + b) = \mathcal{R}_F(I).
\]

Since \( U_{\lambda a}(\lambda v) = \lambda U_a(v), \mathcal{U}_F \) and \( \mathcal{R}_F \) are (positively) homogeneous, i.e.

\[
\mathcal{U}_F(\lambda I) = \lambda \mathcal{U}_F(I)
\]

\[
\mathcal{R}_F(\lambda I) = \lambda \mathcal{R}_F(I).
\]
\( \mathcal{U} \) is concave and \( \mathcal{R} \) is convex in the following sense: If \( I_1 \) and \( I_2 \) are two income variables (possibly dependent), then

\[
\mathcal{U}_F(pI_1 + (1-p)I_2) \geq p\mathcal{U}_F(I_1) + (1-p)\mathcal{U}_F(I_2) \tag{7}
\]

and

\[
\mathcal{R}_F(pI_1 + (1-p)I_2) \leq p\mathcal{R}_F(I_1) + (1-p)\mathcal{R}_F(I_2). \tag{8}
\]

To prove (7) suppose that \( \mathcal{U}_F(I_1) = \mathbb{E}(u_a(I_1)) \) and \( \mathcal{U}_F(I_2) = \mathbb{E}(u_a(I_2)) \). Then, using (4), we obtain

\[
\mathcal{U}_F(pI_1 + (1-p)I_2) = \max_a \mathbb{E}(u_a(pI_1 + (1-p)I_2)) \geq \mathbb{E}(p u_{a_1}(I_1) + (1-p)u_{a_2}(I_2)) = p\mathcal{U}_F(I_1) + (1-p)\mathcal{U}_F(I_2).
\]

Inequality (8) is easily deduced from that.

If we compound \( I_1 \) and \( I_2 \) with probability \( p \), i.e.,

\[
I = \begin{cases} I_1 & \text{with probability } p, \\ I_2 & \text{with probability } 1 - p, \end{cases}
\]

then \( \mathbb{E}(u_a(I)) = p\mathbb{E}(u_a(I_1)) + (1-p)\mathbb{E}(u_a(I_2)) \) and thus

\[
\mathcal{U}_F(I) \leq p\mathcal{U}_F(I_1) + (1-p)\mathcal{U}_F(I_2).
\]

Artzner, Delbaen, Eber and Heath [2] have introduced the notion of a coherent risk measure as a measure being translation-equivariant (they call it translation-invariant), positive homogenous, convex in the sense of (8) and monotonic w.r.t. pointwise ordering (see also [5]). Thus \( -\mathcal{U}_F \) is a coherent risk measure in the sense of [2], but \( \mathcal{R}_F \) is not since it is translation invariant in the sense of (6).

3 Risk of multiperiod income streams

Suppose now that \( I_1, I_2, \ldots, I_T \) is a stream of random incomes which arrive at times \( 1, 2, \ldots, T \). We denote by \( (\Omega, \mathcal{F}, \mathbb{P}) \) the probability space on which these random variables are defined. Together with that, a filtration \( \{\mathcal{F}_t\}, t = 1, \ldots, T \), is defined, so that \( I_t \) is \( \mathcal{F}_t \)-measurable for each \( t = 1, \ldots, T \). The \( \sigma \)-subfield \( \mathcal{F}_t \) represents the information available at time \( t \). We take the convention that \( \mathcal{F}_0 = \{\emptyset, \Omega\} \).

Analogously to the static case, let \( a_t \) be the amount to be consumed at time \( t \). The decision about \( a_t \) must be made at time \( t - 1 \), so \( a_t \) must be \( \mathcal{F}_{t-1} \)-measurable. The consumption of one unit at time \( t \) gives a Net Present Value (NPV) of \( c_t \geq 0 \). The shortfall costs are \( q_t \geq 0 \).

The expected shortfall costs are immediately subtracted from the consumption before period \( t \) (this can be interpreted as an insurance cost). Any surplus occurring in period \( t \)
increases the income of the next period. The final surplus is discounted by a factor \( d \geq 0 \).

We make the following assumptions about the sequences \( \{c_t\}, \{q_t\} \) and the constant \( d \):

\[
c_t \leq q_t, \quad t = 1, \ldots, T, \\
c_{t+1} \leq c_t, \quad t = 1, \ldots, T - 1, \\
d \leq c_T.
\]

Let \( K_t \) be the (random) surplus carried from period \( t \) to period \( t + 1 \). We have \( K_0 = 0 \) and

\[
K_t = [K_{t-1} + I_t - a_t]^+, \quad t = 1, \ldots, T.
\]

The shortfall \( M_t \) at period \( t \) is given by

\[
M_t = [K_{t-1} + I_t - a_t]^-. \tag{11}
\]

Our objective is to maximize the expected consumption minus the expected shortfall costs. This can be written as the following optimization problem:

\[
U(I_1, I_2, \ldots, I_T) = \max E \left[ \sum_{t=1}^{T} (c_t a_t - q_t M_t) + dK_T \right] \tag{12}
\]

\[\text{s.t. } a_t \text{ is } F_{t-1}\text{-measurable for } t = 1, \ldots, T. \tag{13}\]

Similarly to the static case, we introduce the dynamic risk measure of the sequence \( \{I_t\} \) as

\[
\mathcal{R}(I_1, \ldots, I_T) = U(EI_1, \ldots, EI_T) - U(I_1, \ldots, I_T). \tag{14}
\]

We shall prove in the next section that it is always non-negative, and that it possesses most of the properties of the risk measure in the static case.

In order to analyze problem (12)–(13) we shall formalize it as a stochastic control problem. We denote by \( \mathcal{X}_t \) the space of \( F_t\)-measurable random variables having a finite expected value: \( \mathcal{X}_t = L^1(\Omega, \mathcal{F}_t, \mathbb{P}) \). We also use the notation \( E_t(\cdot) \) for \( E(\cdot|\mathcal{F}_t) \).

Problem (12)–(13) can be now stated as follows. Find random variables \( a_t \in \mathcal{X}_{t-1} \), \( M_t \in \mathcal{X}_t \), and \( K_t \in \mathcal{X}_t \), \( t = 1, \ldots, T \), that constitute the solution of the problem

\[
\max E \left[ \sum_{t=1}^{T} (c_t a_t - q_t M_t) + dK_T \right] \tag{15}
\]

\[\text{s.t. } K_t = K_{t-1} + I_t - a_t + M_t, \quad t = 1, \ldots, T, \tag{16}\]

\[
K_t \geq 0, \quad M_t \geq 0, \quad t = 1, \ldots, T, \tag{17}\]

where \( K_0 = 0 \) and the constraints (16)–(17) are understood in the ‘almost sure’ sense.

We can view (15)–(17) as a linear programming problem in abstract spaces. Let us introduce Lagrange multipliers \( \lambda_t \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P}) \) associated with the constraints (16),
The lagrangian takes on the form
\[
L(a, M, K, \lambda) = \mathbb{E} \sum_{t=1}^{T} (c_t a_t - q_t M_t) + d \mathbb{E} K_T - \mathbb{E} \sum_{t=1}^{T} \lambda_t (K_t - K_{t-1} - I_t + a_t - M_t). \tag{18}
\]

The dual functional is defined as
\[
D(\lambda) = \sup_{(a, M, K) \in X_0} L(K, a, M, \lambda),
\]
where
\[
X_0 = \{(a, M, K) : a_t \in \mathcal{X}_{t-1}, M_t \in \mathcal{X}_t, M_t \geq 0, K_t \in \mathcal{X}_t, K_t \geq 0, t = 1, \ldots, T\}. \tag{20}
\]
We have
\[
L(a, M, K, \lambda) = \mathbb{E} \sum_{t=1}^{T} (c_t - \lambda_t) a_t + \mathbb{E} \sum_{t=1}^{T} (\lambda_t - q_t) M_t + \mathbb{E} (d - \lambda_T) K_T
\]
\[+ \mathbb{E} \sum_{t=1}^{T-1} (\lambda_{t+1} - \lambda_t) K_t + \mathbb{E} \sum_{t=1}^{T} \lambda_t I_t
\]
\[= \mathbb{E} \sum_{t=1}^{T} (c_t - \mathbb{E}_{t-1} \lambda_t) a_t + \mathbb{E} \sum_{t=1}^{T} (\lambda_t - q_t) M_t + \mathbb{E} (d - \lambda_T) K_T
\]
\[+ \mathbb{E} \sum_{t=1}^{T-1} (\mathbb{E}_{t} \lambda_{t+1} - \lambda_t) K_t + \mathbb{E} \sum_{t=1}^{T} \lambda_t I_t,
\]
where we have manipulated (by conditioning) the coefficients in front of \(a_t, M_t\) and \(K_t\) to obtain elements of the corresponding dual spaces \(L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})\), \(L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})\), and \(L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})\). It follows that \(D(\lambda) < +\infty\) if and only if the following conditions are satisfied:
\[
\mathbb{E}_{t-1} \lambda_t = c_t, \quad t = 1, \ldots, T, \tag{21}
\]
\[
\lambda_t \leq q_t, \quad t = 1, \ldots, T, \tag{22}
\]
\[
\lambda_T \geq d, \tag{23}
\]
\[
\lambda_t \geq \mathbb{E}_t \lambda_{t+1}, \quad t = 1, \ldots, T - 1, \tag{24}
\]
and the dual problem is to find
\[
\min \mathbb{E} \sum_{t=1}^{T} \lambda_t I_t \tag{25}
\]
subject to (21)–(24). It is worth noting that the multiplier process, \( \{ \lambda_t \} \), is a supermartingale.

Kuhn–Tucker optimality conditions and duality relations hold for our model (15)–(17), similarly to the finite-dimensional case.

**Theorem 1** The processes \( \hat{a}_t, \hat{M}_t, \) and \( \hat{K}_t, t = 1, \ldots, T \), constitute an optimal solution of (15)–(17) if and only if there exists multipliers \( \hat{\lambda}_t \in L_\infty(\Omega, \mathcal{F}_t, \mathbb{P}) \), \( t = 1, \ldots, T \), such that conditions (21)–(24) are satisfied together with the complementary slackness conditions (understood in the ‘almost sure’ sense):

\[
\begin{align*}
\hat{M}_t(q_t - \hat{\lambda}_t) &= 0, \quad t = 1, \ldots, T, \quad (26) \\
\hat{K}_T(\hat{\lambda}_T - d) &= 0, \quad (27) \\
\hat{K}_t(\hat{\lambda}_t - E_t \hat{\lambda}_{t+1}) &= 0, \quad t = 1, \ldots, T - 1. \quad (28)
\end{align*}
\]

**Proof.** Consider the affine operator \( G = (G_1, \ldots, G_T) \) involved in (16):

\[
G_t(a, M, K) = K_t - K_{t-1} - I_t + a_t - M_t, \quad t = 1, \ldots, T.
\]

We treat it as an operator from the space on which \( (a, M, K) \) are defined (the product of the corresponding \( L_1 \) spaces) to \( \mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_T \). Since the image of the set (20) under \( G \) contains a neighborhood of 0 in \( \mathcal{X} \), our result follows from [4, Thm. 4, §1.1].

**Theorem 2** Suppose that conditions (9) hold. Then for every sequence \( I_1, \ldots, I_T \) such that \( E |I_t| < +\infty, t = 1, \ldots, T \), the optimal values of problems (15)–(17) and (21)–(25) are finite and equal.

**Proof.** A feasible solution to the primal problem (15)–(17) is given by \( a_t = E(I_t) \), with the other variables determined by (10)–(11). The objective value at this point provides a lower bound for the optimal value of the dual problem. The feasible set of the dual problem, given by (21)–(24), is convex, closed and bounded in \( L_\infty(\Omega, \mathcal{F}_1, \mathbb{P}) \times \cdots \times L_\infty(\Omega, \mathcal{F}_T, \mathbb{P}) \). Hence, it is weakly* compact (Alaoglu theorem, see [6, Thm. 6, p. 179]). Therefore the dual problem has an optimal solution, \( \hat{\lambda} \). Then every solution \( (a, M, K) \) of the conditions (26)–(28) which satisfies equation (16) is, by Theorem 1, an optimal solution of the primal problem. Such a solution exists, because we can determine \( K \) and \( M \) from (26)–(28), and then choose \( a \) (which is not constrained) to ensure (16).

It is clear that the optimal Lagrange multipliers \( \hat{\lambda}_t(\omega) \) can be interpreted as the (random) costs of a unit of a credit at time \( t \) and scenario \( \omega \). With such costs neither is it profitable to borrow nor to lend at each time \( t \).

4 Properties of the dynamic risk measure

Directly from the definition, the functionals \( U \) and \( R \) are homogeneous. \( U \) is monotonic in the following sense: If two income processes \( \{ I_t^{(1)} \} \) and \( \{ I_t^{(2)} \} \) are defined on the same
probability space \((\Omega, (\mathcal{F}_t), \mathbb{P})\) with the same filtration \(\mathcal{F}_t\) and if \(I_t^{(1)} \leq I_t^{(2)}\) a.s. for all \(t\), then \(U(I_t^{(1)}, \ldots, I_T^{(1)}) \leq U(I_t^{(2)}, \ldots, I_T^{(2)})\). More generally, if all conditional distributions satisfy \(I_t^{(1)}|\mathcal{F}_{t-1} \prec_{SSD} I_t^{(2)}|\mathcal{F}_{t-1}\) for \(t = 1, \ldots, T\), then \(U(I_1^{(1)}, \ldots, I_T^{(1)}) \leq U(I_1^{(2)}, \ldots, I_T^{(2)})\).

Finally, \(U\) is translation equivariant in the following sense:

\[
U(I_1 + b_1, \ldots, I_T + b_T) = U(I_1, \ldots, I_T) + c_1b_1 + c_2b_2 + \ldots + c_Tb_T,
\]

where \(b_1, \ldots, b_T\) are constants. We shall also show in this section that \(U\) is concave, so it makes sense to call \(-U\) coherent in the sense of [2].

Let us start from the following observation.

**Lemma 2** Suppose that conditions (9) hold and that each \(I_t\) is \(\mathcal{F}_{t-1}\)-measurable and integrable, \(t = 1, \ldots, T\). Then

\[
U(I_1, \ldots, I_T) = \sum_{t=1}^{T} c_t \mathbb{E}(I_t).
\]

**Proof.** The solution

\[
a_t = I_t, \quad M_t = 0, \quad K_t = 0, \quad t = 1, \ldots, T,
\]

is feasible for the primal problem (15)–(17), while the solution

\[
\lambda_t = c_t, \quad t = 1, \ldots, T,
\]

is feasible for the dual problem (21)–(25). Since both have the same objective values, \(\sum_{t=1}^{T} c_t \mathbb{E}(I_t)\), by virtue of Theorem 2 they are optimal for their problems.

As a conclusion from this result we obtain a basic property of our risk measure.

**Theorem 3** Suppose that conditions (9) hold. Then for every sequence \(I_1, \ldots, I_T\) such that \(\mathbb{E}|I_t| < +\infty, t = 1, \ldots, T,\) the risk measure (14) is finite and non-negative.

**Proof.** Under conditions (9) the deterministic solution (29) is feasible for (21)–(25). Since a feasible solution for a dual problem always provides an upper bound for the primal problem, for every sequence \(I_1, \ldots, I_T\) such that \(\mathbb{E}|I_t| < +\infty, t = 1, \ldots, T,\) we have

\[
U(I) \leq D(c) = \sum_{t=1}^{T} c_t \mathbb{E}(I_t) = U(\mathbb{E}(I)),
\]

where the last equality follows from Lemma 2.

**Theorem 4** Let \(\mathcal{B}_t, t = 1, \ldots, T,\) be \(\sigma\)-subalgebras such that \(\mathcal{B}_{t-1} \subseteq \mathcal{B}_t \subseteq \mathcal{F}_t\) for \(t = 1, \ldots, T\). Then for every sequence \(I_1, \ldots, I_T\) with \(\mathbb{E}|I_t| < \infty\) we have

\[
\mathcal{R}(\mathbb{E}(I_1|\mathcal{B}_1), \ldots, \mathbb{E}(I_T|\mathcal{B}_T)) \leq \mathcal{R}(I_1, \ldots, I_t).
\]
Proof. By Theorem 2 both $U(I_1, \ldots, I_t)$ and $U(E(I_1|B_1), \ldots, E(I_T|B_T))$ are finite. Let $\hat{\mu}_t, t = 1, \ldots, T$, be the optimal solution of the dual problem (21)–(25) with the income stream $E(I_t|B_t), t = 1, \ldots, T$. Then the multipliers $\bar{\mu}_t = E(\mu_t|B_t), t = 1, \ldots, T$, are also optimal solutions of this problem. Indeed, the feasibility follows from

$$ E_t \bar{\mu}_t = E_t E(\mu_t|B_t) = E_t \mu_t = c_t, \quad t = 1, \ldots, T, $$

and the optimality is guaranteed by

$$ E \sum_{t=1}^T \mu_t E\{I_t|B_t\} = E \sum_{t=1}^T \bar{\mu}_t E\{I_t|B_t\}. $$

The multipliers $\bar{\mu}_t$ are also feasible for (21)–(25) with the income stream $I_t, t = 1, \ldots, T$. Therefore,

$$ U(I_1, \ldots, I_t) \leq E \sum_{t=1}^T \bar{\mu}_t I_t = E \sum_{t=1}^T \bar{\mu}_t E(I_t|B_t). $$

Combining the last two relations and using (14) we obtain the required result.

A simple interpretation of Theorem 4 is that the additional information, represented by $B_t$, reduces risk. In particular, if each $I_t$ becomes known at the preceding period, there is no risk at all, as we have shown it in Lemma 2.

Also, combining two income streams cannot increase risk.

**Theorem 5** Let $I = (I_1, \ldots, I_T)$ and $J = (J_1, \ldots, J_T)$ be two streams of integrable incomes. Then for every $p \in (0, 1)$

$$ R(pI + (1 - p)J) \leq pR(I) + (1 - p)R(J), $$

that is, the functional $R(\cdot)$ is convex.

**Proof.** The result follows from Theorem 2. Let us denote by $\Lambda$ the set of multipliers defined by (21)–(24). We have

$$ U(pI + (1 - p)J) = \min_{\lambda \in \Lambda} \left[ pE \sum_{t=1}^T \lambda_t I_t + (1 - p)E \sum_{t=1}^T \lambda_t J_t \right] $$

$$ \geq p \min_{\lambda \in \Lambda} E \sum_{t=1}^T \lambda_t I_t + (1 - p) \min_{\lambda \in \Lambda} E \sum_{t=1}^T \lambda_t J_t $$

$$ = pU(I) + (1 - p)U(J). $$

Since Lemma 2 implies that

$$ U(pE(I) + (1 - p)E(J)) = pU(E(I)) + (1 - p)U(E(J), $$

our result follows. 

2

12
5 Finite filtrations

Let us consider in more detail the case when the filtration $\mathcal{F} = (\mathcal{F}_1, \ldots, \mathcal{F}_T)$ is finite. This filtration generates partitions of the probability space $\Omega$, which may be represented by a rooted tree of height $T$. Each node of the tree at layer $t$ represents an atom of the $\sigma$-algebra $\mathcal{F}_t$. Subtrees represent subpartitions.

Suppose that the nodes of this tree are numbered $\{0, 1, 2, \ldots, N\}$, with 0 being the root. Let $\mathcal{N} = \{1, 2, \ldots, N\}$ be the node set (not including the root). We assume that there are $N_0 − 1$ nonterminal nodes in $\mathcal{N}$ and that $\mathcal{T} = \{N_0, \ldots, N\}$ is the set of terminal nodes (leaves). If $n$ is a node in $\mathcal{N}$, then $n −$ denotes its predecessor and $t(n)$ denotes its time stage (its distance from the root).

The nodes of the tree are marked by the probabilities of the corresponding elements of the partitions. Evidently, such a tree represents the filtered probability space $(\Omega, (\mathcal{F}_t)_{t=1,\ldots,T}, P)$.

An income stream $I = \{I_t\}$, which is adapted to the filtration $\mathcal{F}_t$, assigns values $I_n$ to each node $n \in \mathcal{N}$.

We call such a valued tree an income stream tree.

The commitment decisions are made at the nonterminal nodes (including the root), i.e. $a$ is a vector of length $N_0$ with components $a_0, \ldots, a_{N_0−1}$.

The calculation of the dynamic utility functional $U$ turns out to be a linear program defined on income stream trees. It reads

$$
\max_{a,M,K} \sum_{n \in \mathcal{N}} p_n c_{t(n)} a_{n−} - \sum_{n \in \mathcal{N}} p_n q_{t(n)} M_n + \sum_{n \in \mathcal{T}} p_n d K_n \\
\text{s.t. } K_n + a_{n−} - M_n = I_n, \quad t(n) = 1, \quad n \in \mathcal{N}, \\
K_n - K_{n−} + a_{n−} - M_n = I_n, \quad t(n) > 1, \quad n \in \mathcal{N}, \\
M_n \geq 0, \quad K_n \geq 0 \quad n \in \mathcal{N}.
$$

(31)

This linear program has $N_0 + 2N$ variables and $N$ equality constraints. Its optimal value is $U(I)$. The risk is defined as

$$
\mathcal{R}(I) = \sum_{n \in \mathcal{N}} c_{t(n)} p_n I_n - U(I).
$$

Let $(z_n)$ be the vector of dual variables of (31). We introduce the notation $n+$ for the set of all successors of the node $n \in \mathcal{N} \setminus \mathcal{T}$. Setting $z_n = p_n \lambda_n$, we obtain the following
form of the dual problem:

\[
\min_{\lambda} \sum_{n \in \mathcal{N}} p_n \lambda I_n \\
\text{s.t. } \lambda_n \geq \frac{1}{p_n} \sum_{m \in n^+} p_m \lambda_m, \quad n \in \mathcal{N} \setminus \mathcal{T}, \\
c_n = \frac{1}{p_n} \sum_{m \in n^+} p_m \lambda_m, \quad n \in \mathcal{N} \setminus \mathcal{T}, \\
\lambda_n \leq q_n, \quad n \in \mathcal{N}, \\
\lambda_n \geq d, \quad n \in \mathcal{T}.
\]

(32)

It is the discrete counterpart of the general dual problem (21)–(25). As before, the dual process \( \{\lambda_n\} \) is a submartingale.

6 Mean–risk models

To illustrate the way in which our dynamic risk measure can be employed, let us now consider the following situation. We have \( n \) random income streams \( I^j_t, \ t = 1, \ldots, T, \ j = 1, \ldots, n \). Our objective is to create a portfolio,

\[
I(x) = \sum_{j=1}^{n} x_j I^j
\]

with \( x_j \geq 0, \sum_{j=1}^{n} x_j \leq C \). With no loss of generality we may assume that \( C = 1 \). The portfolio income stream can be characterized by two measures: the mean represented by the expected net present value,

\[
\mu(x) = \sum_{t=1}^{T} c_t \mathbb{E}(I_t(x)) = \sum_{j=1}^{n} x_j \sum_{t=1}^{T} c_t \mathbb{E}(I^j_t),
\]

and our risk measure

\[
r_d(x) = \mathcal{R}(I_1(x), \ldots, I_t(x)).
\]

The mean–risk optimization model can be formulated as follows

\[
\begin{align*}
\min & \quad r_d(x) \\
\text{s.t.} & \quad \mu(x) = m, \\
& \quad \sum_{j=1}^{n} x_j = 1, \\
& \quad x \geq 0.
\end{align*}
\]

(33)
In the above problem the parameter $m$ is varied over all values for which a solution exists. Using Lemma 2 we can write this problem in a more explicit form:

$$\min \mathbb{E} \left[ \sum_{t=1}^{T} \left( c_t \left( \sum_{j=1}^{n} x_j I_t^j - a_t \right) + q_t M_t \right) - dK_T \right]$$

s.t. $K_t = K_{t-1} + \sum_{j=1}^{n} x_j I_t^j - a_t + M_t, \quad t = 1, \ldots, T,$

$$\sum_{j=1}^{n} x_j \sum_{t=1}^{T} c_t \mathbb{E}(I_t^j) = m,$$

$$\sum_{j=1}^{n} x_j \leq 1,$$

$$x_j \geq 0, \quad j = 1, \ldots, n, \quad K_t \geq 0, \quad M_t \geq 0, \quad t = 1, \ldots, T.$$  \hspace{1cm} (34)

By varying $m$ over possible values of the mean we can reconstruct the mean–risk efficient frontier of our model.

Another possibility is to base the mean–risk model on the distribution of the net present value of $I(x)$, given by

$$N(x) = \sum_{j=1}^{n} x_j \sum_{t=1}^{T} c_t I_t^j.$$  

The mean is the same as before, while the risk can be measured by the static measure based on (1):

$$r_s(x) = \mathbb{E}(N(x)) - \max_a \{ \mathbb{E}(a + d[N(x) - a]^+ - q[N(x) - a]^-) \}.$$  

The explicit linear programming formulation of model (33) (with the rik measure $r_d$ replaced by $r_s$) reads:

$$\min \mathbb{E} \left[ \sum_{j=1}^{n} x_j \sum_{t=1}^{T} c_t I_t^j - a + qM - dK \right]$$

s.t. $\sum_{j=1}^{n} x_j \sum_{t=1}^{T} c_t I_t^j - a + M - K = 0,$

$$\sum_{j=1}^{n} x_j \sum_{t=1}^{T} c_t \mathbb{E}(I_t^j) = m,$$

$$\sum_{j=1}^{n} x_j \leq 1,$$

$$x_j \geq 0, \quad j = 1, \ldots, n, \quad K \geq 0, \quad M \geq 0.$$  \hspace{1cm} (35)

The disadvantage of the latter approach is that it concentrates exclusively on the net present value, ignoring the temporal variations.
7 Examples

**Example 2** This example has already been mentioned in the Introduction. Suppose a fair coin is thrown three times. Consider two situations:

**Situation 1:** The income is 1 at the final stage, if more heads than tails were counted.

**Situation 2:** The income is 1 at the final stage, if the last throw shows heads.

The corresponding income stream trees are shown in Figure 1, where an upmove means heads and a downmove represents tails.

Evidently, the two cases leads to exactly the same marginal income distributions at each stage. On the other hand, Situation 1 is more predictable and should lead to a smaller risk.

We have solved the linear program (31) with the specification
\[ c = [1, 0.95, (0.95)^2], \quad q = [1.2, 1.2 \cdot 0.95, 1.2 \cdot (0.95)^2], \quad d = (0.93)^2, \]
and we have obtained the following results:
\[ U(I^{(1)}) = 0.4419, \quad U(I^{(2)}) = 0.4325, \]
\[ U(\mathbb{E}(I^{(1)})) = U(\mathbb{E}(I^{(2)})) = 0.5 \cdot (0.95)^2 = 0.4512 \]
and therefore
\[ R(I^{(1)}) = 0.0093 < R(I^{(2)}) = 0.0188. \]

This analysis shows that Situation 2 is riskier than Situation 1, indeed.

It is also interesting to look at the dual variables \( \lambda^{(1)} \) and \( \lambda^{(2)} \) given by (32). They generate dual submartingale processes which live on the same tree as the income processes. This is illustrated in Figure 2.
Example 3 We modify Example 2 in such a way that a positive income may also occur at stages 1 and 2, but the marginals remain identical. Consider the income trees shown in Figure 3.

Assuming that all arc probabilities are 0.5 one gets the result

\[
U(I^{(3)}) = 1.3919, \quad R(I^{(3)}) = 0.0344
\]

\[
U(I^{(4)}) = 1.3775, \quad R(I^{(4)}) = 0.0487
\]

Since the predictability occurs earlier in Tree 3, its risk is smaller, as expected. We may illustrate here that hiding some information leads to larger risk. Suppose that the outcome of throw 2 is not revealed. In this case Tree 3 changes to Tree 3a shown in Figure 4.

The utility and risk for Tree 3a are

\[
U(I^{(3a)}) = 1.3825, \quad R(I^{(3a)}) = 0.0438.
\]
As expected, the risk of Tree 3a is larger than the risk of Tree 3.

**Example 4** Our next example concerns portfolio models for income streams. Let us consider two streams as shown on Figure 5.

The third income stream is a shifted random walk: each upmove corresponds to the income of 3 and each downmove to the income of 0. The mean and the dynamic risk associated with these three income streams are as follows

\[
\mu(I^1) = 2.13, \quad \mathcal{R}(I^1) = 0.034 \\
\mu(I^2) = 2.46, \quad \mathcal{R}(I^2) = 0.065 \\
\mu(I^3) = 4.27, \quad \mathcal{R}(I^3) = 0.101.
\]

For these three assets we have formulated and solved the two mean–risk portfolio problems: the model with the dynamic risk measure (34) and the model based on the net present value (35).
We assumed the mean value $m = 3$. In the dynamic model we have obtained the following portfolio

$$x^d = (0.6, 0, 0.4),$$

while the static model (based on the net present value) yields

$$x^s = (0, 0.69, 0.23).$$

As expected, the static model uses inflow 2, which has a better mean and a lower risk in terms of the total cumulative flow, while the dynamic model uses inflow 1, which is more predictable.

The corresponding cumulative flow trajectories for all eight scenarios are illustrated in Figures 7 and 7. The cumulative flow for income $\{I_t\}$ is defined recursively by

$$W(0) = 0, \quad W_t = W_{t-1} + c_t I_t, \quad t = 1, 2, 3.$$  

Our results confirm our expectations: the dynamic risk measure, when used in a mean–risk optimization model, prefers solutions with more predictable income streams.

The images of the dynamic mean–risk frontier and the static frontier in the space corresponding to our dynamic measure is presented in Figure 8. We see that for low values of risk, there is a significant difference between the two approaches.
Figure 7: Performance of the portfolio generated by the static mean–risk model.

Figure 8: The dynamic mean–risk frontier and the image of the static mean–risk frontier.
References


