

# Time consistency of dynamic risk measures

Alexander Shapiro\*

School of Industrial & Systems Engineering,  
Georgia Institute of Technology,  
765 Ferst Drive, Atlanta, GA 30332.

**Abstract.** In this paper we discuss time consistency of risk averse multistage stochastic programming problems. We show, in a framework of finite scenario trees, that composition of law invariant coherent risk measures can be law invariant only for the expectation or max-risk measures.

**Key Words.** Multistage stochastic programming, risk averse optimization, dynamic programming, time consistency, coherent risk measures

---

\*e-mail: ashapiro@isye.gatech.edu. This research was partly supported by the NSF award DMS-0914785 and ONR award N000140811104.

# 1 Introduction

Consider a decision process of the form

$$\begin{aligned} &\text{decision } (x_1) \rightsquigarrow \text{observation } (\mathcal{F}_2) \rightsquigarrow \text{decision } (x_2) \rightsquigarrow \\ &\dots \rightsquigarrow \text{observation } (\mathcal{F}_T) \rightsquigarrow \text{decision } (x_T). \end{aligned} \tag{1.1}$$

Here  $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_T$  is a sequence of sigma algebras (filtration) defined on a measurable space  $(\Omega, \mathcal{F})$ , with  $\mathcal{F}_1 = \{\emptyset, \Omega\}$  and  $\mathcal{F}_T = \mathcal{F}$ , representing flow of information and  $x_t \in \mathbb{R}^{n_t}$ ,  $t = 1, \dots, T$ , are decision vectors. The basic requirement of such decision process is the *nonanticipativity* constraint. That is, at time period  $t \in \{1, \dots, T\}$  our decision  $x_t$  should only depend on information available at time  $t$ , and should not depend on future observations. This can be formulated by considering  $x_t : \Omega \rightarrow \mathbb{R}^{n_t}$  as a mapping which is required to be  $\mathcal{F}_t$  measurable. Since  $\mathcal{F}_1 = \{\emptyset, \Omega\}$  is trivial, this implies that the first stage decision  $x_1$  is deterministic. If the set  $\Omega$  is finite this process can be represented as a scenario tree with every node of the tree representing the corresponding state of the process with its children nodes representing possible states at the next period of time (see, e.g., [7, section 6.7.1] for a detail description of this construction).

So far we didn't introduce any optimality criteria for our decisions. One standard approach is to minimize the expected value of the total cost:

$$\text{Min } \mathbb{E}[J(\omega)] \quad \text{subject to } (x_1, x_2(\cdot), \dots, x_T(\cdot)) \in \mathfrak{X}. \tag{1.2}$$

Here  $f_t : \mathbb{R}^{n_t} \times \Omega \rightarrow \mathbb{R}$  are cost functions,  $J(\omega) := f_1(x_1) + f_2(x_2(\omega), \omega) + \dots + f_T(x_T(\omega), \omega)$  and

$$\mathfrak{X} := \{(x_1, x_2(\cdot), \dots, x_T(\cdot)) : x_1 \in \mathcal{X}, x_2(\cdot) \in \mathcal{X}_2(x_1, \cdot), \dots, x_T(\cdot) \in \mathcal{X}_T(x_{T-1}(\cdot), \cdot)\}$$

is the set of feasibility constraints. The functions  $f_t(x_t, \cdot)$  and multifunctions  $\mathcal{X}_t : \mathbb{R}^{n_{t-1}} \times \Omega \rightrightarrows \mathbb{R}^{n_t}$  are  $\mathcal{F}_t$ -measurable, the expectation is computed according to a specified probability measure  $P$  on the measurable space  $(\Omega, \mathcal{F})$ , and the optimization (minimization) in (1.2) is performed over mappings  $x_t(\cdot)$ ,  $t = 1, \dots, T$ , satisfying the nonanticipativity constraint. We refer to problem (1.2), where the cost minimization is performed "on average", as the *risk neutral* formulation.

In recent years considerable attention was attracted to *risk averse* formulations of multistage stochastic programming problems (with  $T \geq 3$ ). That is, for a specified risk measure  $\rho : \mathcal{Z} \rightarrow \mathbb{R}$ , the following problem is considered

$$\text{Min } \rho[J(\omega)] \quad \text{subject to } (x_1, x_2(\cdot), \dots, x_T(\cdot)) \in \mathfrak{X}. \tag{1.3}$$

Here  $\rho$  is a functional assigning to a random variable  $Z \in \mathcal{Z}$  numerical value  $\rho(Z)$ , with  $\mathcal{Z}$  being a linear space of allowable random variables. In various applications it makes sense to use the space  $\mathcal{Z} := L_p(\Omega, \mathcal{F}, P)$ ,  $p \in [1, \infty)$ , of random variables  $Z : \Omega \rightarrow \mathbb{R}$  having finite  $p$ -th order moments. In the terminology introduced by Artzner et al [1], risk measure  $\rho : \mathcal{Z} \rightarrow \mathbb{R}$  is said to be *coherent* if it satisfies the following conditions (axioms):

(A1) *Monotonicity*: If  $Z, Z' \in \mathcal{Z}$  and  $Z \succeq Z'$ , then  $\rho(Z) \geq \rho(Z')$ .

(A2) *Convexity*:

$$\rho(tZ + (1-t)Z') \leq t\rho(Z) + (1-t)\rho(Z')$$

for all  $Z, Z' \in \mathcal{Z}$  and all  $t \in [0, 1]$ .

(A3) *Translation Equivariance*: If  $a \in \mathbb{R}$  and  $Z \in \mathcal{Z}$ , then  $\rho(Z + a) = \rho(Z) + a$ .

(A4) *Positive Homogeneity*: If  $t \geq 0$  and  $Z \in \mathcal{Z}$ , then  $\rho(tZ) = t\rho(Z)$ .

The notation  $Z \succeq Z'$  means that  $Z(\omega) \geq Z'(\omega)$  for a.e.  $\omega \in \Omega$ .

It is said that risk measure  $\rho$  is *law invariant* if  $\rho(Z) = \rho(Z')$  for any two random variables  $Z, Z' \in L_p(\Omega, \mathcal{F}, P)$  having the same distribution. An important example of a (nonlinear) law invariant coherent risk measure is the so-called Average Value-at-Risk, which can be defined as

$$\text{AVaR}_\alpha(Z) := \inf_{t \in \mathbb{R}} \{t + (1 - \alpha)^{-1} \mathbb{E}[Z - t]_+\}, \quad (1.4)$$

with  $\mathcal{Z} := L_1(\Omega, \mathcal{F}, P)$  and  $\alpha \in [0, 1)$ . It can be noted that for  $\alpha = 0$ , the  $\text{AVaR}_0(\cdot)$  coincides with the expectation operator  $\mathbb{E}(\cdot)$ . For another extreme case of  $\alpha$  tending to 1,

$$\lim_{\alpha \uparrow 1} \text{AVaR}_\alpha(Z) = \text{ess sup } Z.$$

That is,  $\text{AVaR}_1(\cdot)$  can be defined as the essential supremum operator defined on the space  $\mathcal{Z} := L_\infty(\Omega, \mathcal{F}, P)$ .

From the point of view of the dynamic decision process (1.1) it is natural to consider the following conceptual requirement of *time consistency* (cf., [8]).

- Optimality of the decision at a state of the process at time  $t \in \{1, \dots, T-1\}$  should not involve states which do not follow that state, i.e., cannot happen in the future.

If the decision process is represented by the corresponding scenario tree, this means that if at a time  $t$  we are at a certain node of the tree, then optimality of our future decisions should not depend on scenarios which do not pass through this node.

In order to formalize this concept of time consistency we need to define what do we optimize (say minimize) at every state of the process looking into the future. That is, in the scenario tree formulation, at every node at time  $t \in \{1, \dots, T-1\}$  we should define a risk measure on the space of children nodes of that node, and hence construct the corresponding conditional risk mapping  $\rho_{t+1|\mathcal{F}_t}$  (see [7, section 6.7.1] for a detail description of such construction). Recall that the sigma algebra  $\mathcal{F}_1$  is trivial and hence  $\rho_{2|\mathcal{F}_1}$  is a real valued risk measure.

This will define a nested formulation of the associated optimization problem (see also [6]). This nested formulation can be written in the equivalent form

$$\text{Min } \varrho[J(\omega)] \quad \text{subject to } (x_1, x_2(\cdot), \dots, x_T(\cdot)) \in \mathfrak{X}, \quad (1.5)$$

where  $\varrho : \mathcal{Z} \rightarrow \mathbb{R}$  is the risk measure defined as the composition

$$\varrho(Z) = \rho_{2|\mathcal{F}_1}(\dots \rho_{T|\mathcal{F}_{T-1}}(Z)) \quad (1.6)$$

of the constructed conditional risk mappings. Assuming that each conditional risk measure is coherent, we have that the composite risk measure  $\varrho$  is also coherent.

The composite risk measure  $\varrho$  assigns to each state of the process (node of the scenario tree) the corresponding conditional risk measure. Consequently the formulation (1.5) is time consistent with respect to optimization at every state performed by minimizing the corresponding conditional risk measure (cf., [5]). This raises a natural question whether for a given risk measure  $\rho : \mathcal{Z} \rightarrow \mathbb{R}$ , the optimization problem (1.3) is equivalent to the time consistent problem (1.5) for some choice of the conditional risk measures  $\rho_{t|\mathcal{F}_{t-1}}$ . Of course, a sufficient condition for such time consistency of problem (1.3) is that risk measure  $\rho$  can be represented as the composite risk measure  $\varrho$ , i.e.,  $\rho(\cdot) = \varrho(\cdot)$  for some choice of the conditional risk measures. We say that risk measure  $\rho : \mathcal{Z} \rightarrow \mathbb{R}$  is *decomposable* into a sequence  $\rho_{t+1|\mathcal{F}_t}$ ,  $t = 2, \dots, T$ , of conditional risk mappings if

$$\rho(\cdot) = \rho_{2|\mathcal{F}_1}(\dots \rho_{T|\mathcal{F}_{T-1}}(\cdot)). \quad (1.7)$$

Note that with a law invariant risk measure is associated its conditional analogue. If the decomposition (1.7) holds for a law invariant risk measure  $\rho$  with the conditional risk mappings  $\rho_{t+1|\mathcal{F}_t}$  being the conditional analogues of  $\rho$ , then  $\rho$  is called time consistent in the mathematical finance literature (e.g., [3]). An example of time consistent risk measure is the expectation operator, i.e.,  $\rho(\cdot) := \mathbb{E}(\cdot)$  and  $\rho_{t|\mathcal{F}_{t-1}}(\cdot) = \mathbb{E}[\cdot|\mathcal{F}_{t-1}]$ . Another example is the essential supremum operator. It is shown in Kupper and Schachermayer [4] that (under certain regularity conditions) for the class of law invariant coherent risk measures these two examples are the only possible time consistent risk measures. In fact a more general result is proved in [4]; it is shown that in the class of law invariant convex risk measures, time consistent risk measures can be only entropic risk measures.

We approach the question of decomposability of a coherent risk measure from a somewhat different point of view. In the next section we show that if each conditional risk mapping in (1.6) is law invariant and coherent, then the corresponding composite risk measure can be law invariant only in the two cases of expectation and essential supremum operators. This implies, for example, that  $\rho(\cdot) := \text{AVaR}_\alpha(\cdot)$ ,  $\alpha \in (0, 1)$ , risk measure is not decomposable into a sequence of law invariant coherent risk mappings. We will use a framework of finite scenario trees and the derivations will be rather elementary.

## 2 Main result

Let us consider the following construction. Let  $\Omega = \{\omega_1, \dots, \omega_n\}$ ,  $n \geq 3$ , be a finite sample space equipped with equal probabilities  $p_i = 1/n$ ,  $i = 1, \dots, n$ . Consider the sample space  $\Omega \times \Omega$  equipped with equal probabilities  $1/n^2$  and sigma algebra  $\mathcal{F}$  of all its subsets. Denote  $\mathcal{Z}$  the linear space of functions (random variables)  $Z : \Omega \times \Omega \rightarrow \mathbb{R}$ . We can identify random variable  $Z \in \mathcal{Z}$  with the corresponding  $n \times n$  matrix with entries  $Z_{ij} := Z(\omega_i, \omega_j)$ ,  $i, j = 1, \dots, n$ . Denote by  $Z_{i*} := (Z_{i1}, \dots, Z_{in})$  and  $Z_{*i} := (Z_{1i}, \dots, Z_{ni})$  the respective rows and columns of matrix  $Z$ . Let  $\mathcal{G}$  be the subalgebra of  $\mathcal{F}$  generated by events  $\{\omega_i\} \times \Omega$ ,  $i = 1, \dots, n$ , and  $\mathcal{Z}_{\mathcal{G}}$  be the linear subspace of  $\mathcal{Z}$  of  $\mathcal{G}$ -measurable functions, i.e.,  $Z \in \mathcal{Z}_{\mathcal{G}}$  if  $Z_{i1} = \dots = Z_{in}$  for every  $i \in \{1, \dots, n\}$ .

Let  $\rho_0 : \mathcal{Z}_{\mathcal{G}} \rightarrow \mathbb{R}$  be a real valued law invariant coherent risk measure, and  $\rho_i$ ,  $i = 1, \dots, n$ , be real valued law invariant coherent risk measures defined on the space of random variables  $Y : \Omega \rightarrow \mathbb{R}$ . We can view random variable  $Y : \Omega \rightarrow \mathbb{R}$  as  $n$ -dimensional vector with entries  $Y_j := Y(\omega_j)$ ,  $j = 1, \dots, n$ , and hence to consider values  $\rho_i(Z_{i*})$  for  $i \in \{1, \dots, n\}$ . Also we can view risk measures  $\rho_i$ ,  $i = 0, 1, \dots, n$ , as functions  $\rho_i : \mathbb{R}^n \rightarrow \mathbb{R}$ . Consider the risk mapping  $\rho_{|\mathcal{G}} : \mathcal{Z} \rightarrow \mathcal{Z}_{\mathcal{G}}$  defined as

$$[\rho_{|\mathcal{G}}(Z)](\{\omega_i\} \times \Omega) = \rho_i(Z_{i*}), \quad i = 1, \dots, n, \quad (2.1)$$

and the composite risk measure  $\varrho : \mathcal{Z} \rightarrow \mathbb{R}$  defined as

$$\varrho(Z) := \rho_0(\rho_{|\mathcal{G}}(Z)). \quad (2.2)$$

Note that law invariance of risk measures  $\rho_i$ ,  $i = 0, 1, \dots, n$ , means that if  $Y, Y' \in \mathbb{R}^n$  are such that components of  $Y'$  are obtained by a permutation of components of  $Y$ , then  $\rho_i(Y) = \rho_i(Y')$ . Similarly,  $\varrho$  is law invariant if  $Z \in \mathcal{Z}$  and  $Z'$  is obtained from  $Z$  by a permutation of components of  $Z$ , then  $\varrho(Z) = \varrho(Z')$ .

**Lemma 2.1** *Let  $\rho_i$ ,  $i = 0, 1, \dots, n$ , be law invariant coherent risk measures and  $\varrho$  be the corresponding composite risk measure defined in (2.2). Suppose that  $\varrho$  is law invariant. Then  $\rho_0 = \rho_1 = \dots = \rho_n$ .*

**Proof.** First, let us show that  $\rho_1 = \dots = \rho_n$ . Indeed, suppose for example that  $\rho_1 \neq \rho_2$ . This means that there exists  $Y \in \mathbb{R}^n$  such that  $\rho_1(Y) \neq \rho_2(Y)$ . We can assume that  $\rho_1(Y) > \rho_2(Y) > 0$ .

Let  $Z \in \mathcal{Z}$  be such that  $Z_{1*} = Y$  and all other components of  $Z$  are zeros, and  $Z' \in \mathcal{Z}$  be such that  $Z'_{2*} = Y$  and all other components of  $Z'$  are zeros. Note that  $Z'$  can be obtained from  $Z$  by permutation of its components and hence by law invariance of  $\varrho$  it follows that  $\varrho(Z) = \varrho(Z')$ . We have that  $\rho_{|\mathcal{G}}(Z)$  and  $\rho_{|\mathcal{G}}(Z')$  are given by vectors  $X := (\rho_1(Y), 0, \dots, 0) \in \mathbb{R}^n$  and  $X' := (0, \rho_2(Y), \dots, 0)$ , respectively. Since  $\rho_0$  is law invariant we have that  $\rho_0(X) = \rho_0(X')$ . Also by positive homogeneity of  $\rho_0$  we have that  $\rho_0(X) = \rho_1(Y)\rho_0(e_1)$  and  $\rho_0(X') = \rho_2(Y)\rho_0(e_1)$ , where  $e_i \in \mathbb{R}^n$  denotes the  $i$ -th coordinate vector. Also

$$1 = \rho_0(e_1 + \dots + e_n) \leq \rho_0(e_1) + \dots + \rho_0(e_n) = n\rho_0(e_1),$$

and hence  $\rho_0(e_1) \geq 1/n > 0$ . Thus  $\varrho(Z) > \varrho(Z')$ , a contradiction.

Let  $\rho$  denote the common risk measure  $\rho_i$ ,  $i = 1, \dots, n$ . Next let us show that  $\rho_0 = \rho$ . Consider  $Y \in \mathbb{R}^n$  and let  $Z \in \mathcal{Z}$  be such that  $Z_{i*} = Y$ ,  $i = 1, \dots, n$ . Then  $\rho(Z_{i*}) = \rho(Y)$ ,  $i = 1, \dots, n$ , and hence  $\rho_{\mathcal{G}}(Z) = (\rho(Y), \dots, \rho(Y))$  and thus  $\varrho(Z) = \rho(Y)$ . On the other hand, let  $Z' \in \mathcal{Z}$  be such that  $Z'_{*i} = Y$ ,  $i = 1, \dots, n$ , i.e.,  $Z' = Z^{\top}$ . Note that  $Z'$  can be obtained from  $Z$  by permutation of its components and hence  $\varrho(Z') = \varrho(Z)$ . Moreover,  $Z_{i*} = (Y_i, \dots, Y_i)$  and hence  $\rho(Z_{i*}) = Y_i$ . Thus  $\rho_{\mathcal{G}}(Z') = Y$  and  $\varrho(Z') = \rho_0(Y)$ . It follows that  $\rho_0(Y) = \rho(Y)$ . ■

**Remark 1** It could be noted that positive homogeneity of the considered risk measures was used in the above proof only to ensure that if  $c_1 > c_2 > 0$ , then  $\rho_0(c_1 e_1) > \rho_0(c_2 e_1)$ .

For a finite set  $\Omega$  and  $Y : \Omega \rightarrow \mathbb{R}$  denote

$$\rho^{\max}(Y) := \max_{\omega \in \Omega} Y(\omega) \tag{2.3}$$

the max-measure. Note that if  $\rho_i := \rho^{\max}$ ,  $i = 0, 1, \dots, n$ , then the composite risk measure  $\varrho$  is also the max-measure.

**Theorem 2.1** *Let  $\rho_i$ ,  $i = 0, 1, \dots, n$ , be law invariant coherent risk measures and  $\varrho$  be the corresponding composite risk measure defined in (2.2). Suppose that  $\varrho$  is law invariant. Then  $\rho_i = \rho_0$ ,  $i = 1, \dots, n$ , and  $\rho_0$  is either the expectation  $\rho_0(\cdot) = \mathbb{E}[\cdot]$  or the max-measure  $\rho_0(\cdot) = \rho^{\max}(\cdot)$ .*

**Proof.** By lemma 2.1 we have that risk measures  $\rho_i$ ,  $i = 0, 1, \dots, n$ , are equal to each other. Denote by  $\rho$  the common risk measure  $\rho_i$ . Since  $\rho$  is coherent it has the following dual representation (cf., [1])

$$\rho(Y) = \sup_{a \in \mathfrak{A}} a^{\top} Y, \tag{2.4}$$

where  $\mathfrak{A}$  is a convex compact subset of  $\Delta_n = \{a \in \mathbb{R}_+^n : \sum_{i=1}^n a_i = 1\}$ . It follows that for any  $Z \in \mathcal{Z}$ ,

$$\rho_{\mathcal{G}}(Z) = \left( \sup_{b_1 \in \mathfrak{A}} b_1^{\top} Z_{1*}, \dots, \sup_{b_n \in \mathfrak{A}} b_n^{\top} Z_{n*} \right), \tag{2.5}$$

and hence for any  $Z \in \mathcal{Z}$  with nonnegative components  $Z_{ij} \geq 0$ ,

$$\varrho(Z) = \sup_{a \in \mathfrak{A}} a^{\top} \rho_{\mathcal{G}}(Z) \geq \sup_{a \in \mathfrak{A}} \sum_{i,j=1}^n a_i a_j Z_{ij}. \tag{2.6}$$

Now consider  $Y \in \mathbb{R}^n$  with positive components arranged in the decreasing order, i.e.,  $Y_1 > \dots > Y_n > 0$ , and  $Z \in \mathcal{Z}$  such that  $Z_{i*} = Y$ ,  $i = 1, \dots, n$ . Since the maximum in the right hand side of (2.4) is attained, there exists  $\bar{a} \in \mathfrak{A}$  (depending on  $Y$ ) such that  $\rho(Y) = \bar{a}^{\top} Y$ . Recall that  $\bar{a} \geq 0$  and  $\bar{a}_1 + \dots + \bar{a}_n = 1$ , and note that since components of  $Y$  are decreasing and  $\rho$  is law invariant, it

follows that  $\bar{a}_1 \geq \dots \geq \bar{a}_n$ . We have that  $\rho_{\mathcal{G}}(Z) = (\rho(Y), \dots, \rho(Y))$  and hence  $\bar{a}^\top \rho_{\mathcal{G}}(Z) = \rho(Y)$ , and thus

$$\varrho(Z) = \rho(Y) = \bar{a}^\top \rho_{\mathcal{G}}(Z) = \sum_{i=1}^n \bar{a}_i \sum_{j=1}^n \bar{a}_j Y_j = \sum_{i,j=1}^n \bar{a}_i \bar{a}_j Y_j. \quad (2.7)$$

Since  $\varrho$  is law invariant, we have that if  $Z' \in \mathcal{Z}$  is obtained from the matrix  $Z$  by a permutation of its elements, then  $\varrho(Z) = \varrho(Z')$  and hence by (2.6) and (2.7) it follows that

$$\sum_{i,j=1}^n \bar{a}_i \bar{a}_j Y_j \geq \sum_{i,j=1}^n \bar{a}_i \bar{a}_j Z'_{ij}. \quad (2.8)$$

In particular, consider  $Z' \in \mathcal{Z}$  obtained by interchanging elements  $Z_{12} = Y_2$  and  $Z_{i1} = Y_1$ ,  $i \geq 3$ , of  $Z$ . By (2.8) we obtain that

$$\bar{a}_1 \bar{a}_2 Y_2 + \bar{a}_i \bar{a}_1 Y_1 \geq \bar{a}_1 \bar{a}_2 Y_1 + \bar{a}_i \bar{a}_1 Y_2.$$

It follows that  $\bar{a}_1 \bar{a}_i (Y_1 - Y_2) \geq \bar{a}_1 \bar{a}_2 (Y_1 - Y_2)$ , which implies that  $\bar{a}_i \geq \bar{a}_2$  and hence  $\bar{a}_i = \bar{a}_2$ ,  $i = 3, \dots, n$ . That is, we obtain that  $\bar{a}_2 = \dots = \bar{a}_n$ .

Now let  $Z' \in \mathcal{Z}$  be obtained from  $Z$  by interchanging elements  $Z_{22} = Y_2$  and  $Z_{13} = Y_3$ . Again by (2.8) we have that

$$\bar{a}_2^2 Y_2 + \bar{a}_1 \bar{a}_3 Y_3 \geq \bar{a}_2^2 Y_3 + \bar{a}_1 \bar{a}_3 Y_2,$$

and since  $\bar{a}_2 = \bar{a}_3$  it follows that  $\bar{a}_2^2 \geq \bar{a}_1 \bar{a}_2$ . Consequently either  $\bar{a}_2 = 0$  or  $\bar{a}_1 = \bar{a}_2$ . That is, either  $\bar{a} = e_1$  or  $\bar{a} = (e_1 + \dots + e_n)/n$ , i.e., either  $\rho(Y) = \max_{1 \leq i \leq n} Y_i$  or  $\rho(Y) = n^{-1}(Y_1 + \dots + Y_n) = \mathbb{E}[Y]$ . By continuity of  $\rho(\cdot)$  this shows that either  $\rho(\cdot) = \mathbb{E}[\cdot]$  or  $\rho$  is the max-measure on  $\mathbb{R}_+^n$ . Since  $\rho(Y + c) = \rho(Y) + c$  for any  $Y \in \mathbb{R}^n$  and  $c \in \mathbb{R}$ , this holds on  $\mathbb{R}^n$  as well. ■

**Remark 2** It is essential in the above result that all probabilities  $p_i$  are the same. Otherwise suppose that the probabilities  $p_i$  are such that  $\sum_{i \in I} p_i = \sum_{j \in J} p_j$ , for  $I, J \subset \{1, \dots, n\}$ , only if  $I = J$ , and the same holds for probabilities  $p_{ij} = p_i p_j$  on the space  $\Omega \times \Omega$ . Then any risk measures defined on the spaces of random variables  $Y : \Omega \rightarrow \mathbb{R}$  and  $Z : \Omega \times \Omega \rightarrow \mathbb{R}$  are law invariant.

### 3 Concluding remarks

The result of Theorem 2.1 has the following interpretation. Consider the following 3-stage scenario tree. Root node at the first stage has  $n \geq 3$  children nodes and every node at the second stage has  $n$  children nodes. This results in  $n^2$  scenarios, to each scenario we assign equal probability  $1/n^2$ , so that the probability of moving from one node to a node at the next stage is  $1/n$ . Suppose that to the root node is assigned a law invariant coherent risk measure defined on the space of its children nodes, and to every node at the second stage is assigned a law invariant coherent risk measure on its children nodes (these risk measures can be different from each other). This defines a composite risk measure  $\varrho$  on the considered scenario tree (cf., [7, section 6.7.1]). By Theorem 2.1 we have that if  $\varrho$  is law invariant, then  $\varrho$  is either the expectation or the max-risk measure. For example the  $\text{AVaR}_\alpha$ ,  $\alpha \in (0, 1)$ , cannot be represented as such composite risk measure.

In a similar way this can be extended to the following continuous case. Let  $\Xi_1 \subset \mathbb{R}^{d_1}$  and  $\Xi_2 \subset \mathbb{R}^{d_2}$  be two closed sets equipped with respective probability distributions  $P_1$  and  $P_2$ . Suppose that the probability measures  $P_1$  and  $P_2$  are atomless and complete and let  $\Xi := \Xi_1 \times \Xi_2$  be equipped with the probability measure  $P := P_1 \times P_2$  and the sigma algebra  $\mathcal{F}$  of the corresponding Lebesgue sets (see, e.g., [2, p. 45]). Consider space  $\mathcal{Z} := L_p(\Xi, \mathcal{F}, P)$ , with  $p \in [1, \infty)$ , of random

variables  $Z : \Xi \rightarrow \mathbb{R}$  having finite  $p$ -th order moments. Conditional on  $\xi_1$  define a (real valued) law invariant coherent risk measure  $\rho_{|\xi_1}(\cdot)$  on the space of random variables  $Y : \Xi_2 \rightarrow \mathbb{R}$  having finite  $p$ -th order moments. For a (real valued) law invariant coherent risk measure  $\rho_0$ , defined on the space of random variables  $Y : \Xi_1 \rightarrow \mathbb{R}$  having finite  $p$ -th order moments, consider the corresponding composite risk measure  $\varrho(\cdot) := \rho_0(\rho_{|\xi_1}(\cdot)) : \mathcal{Z} \rightarrow \mathbb{R}$ . Then  $\varrho(\cdot)$  can be law invariant only if  $\varrho(\cdot) = \mathbb{E}[\cdot]$ . Recall that we consider real valued risk measures and note that the essential supremum risk measure  $\varrho(Z) = \text{ess sup}(Z)$  can have  $\infty$  values for some  $Z \in L_p(\Xi, \mathcal{F}, P)$ , with  $p \in [1, \infty)$ . If we consider the space  $\mathcal{Z} := L_\infty(\Xi, \mathcal{F}, P)$  of essentially bounded random variables, then the composite risk measure can be law invariant if either  $\varrho(\cdot) = \mathbb{E}[\cdot]$  or  $\varrho(\cdot) = \text{ess sup}(\cdot)$ .

## References

- [1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, Coherent measures of risk, *Mathematical Finance*, 9 (1999), 203–228.
- [2] P. Billingsley, *Probability and Measure*, Wiley, Third Edition, New York, 1995.
- [3] H. Föllmer and A. Schied, *Stochastic Finance: An Introduction in Discrete Time*, Walter de Gruyter, Berlin, 2002 (Second Edition 2004).
- [4] M. Kupper and W. Schachermayer, Representation results for law invariant time consistent functions, *Mathematics and Financial Economics*, 2 (2009), 189–210.
- [5] A. Ruszczyński and A. Shapiro, Conditional risk mappings, *Mathematics of Operations Research*, 31 (2006), 544–561.
- [6] A. Ruszczyński, Risk-averse dynamic programming for Markov decision processes, *Math. Program., Ser. B*, 125 (2010), 235–261.
- [7] A. Shapiro, D. Dentcheva and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia, 2009.
- [8] A. Shapiro, On a time consistency concept in risk averse multi-stage stochastic programming, *Operations Research Letters*, 37 (2009), 143–147.