



Multistage stochastic dominance: an application to pension fund management

Miloš Kopa¹ · Vittorio Moriggia² · Sebastiano Vitali^{1,2}

Received: 24 December 2022 / Accepted: 8 October 2023
© The Author(s) 2023

Abstract

A pension fund manager typically decides the allocation of the pension fund assets taking into account a long-term sustainability goal. Many asset and liability management models, in the form of multistage stochastic programming problem, have been proposed to help the pension fund manager to define the optimal allocation given a multi-objective function. The recent literature proposes univariate stochastic dominance constraints to guarantee that the optimal strategy is able to stochastically dominate a benchmark portfolio. In this work we extend previous results (i) considering alternative types of multivariate stochastic dominance that appear more suitable in a multistage framework, (ii) proposing a way to measure the economic cost of introducing stochastic dominance constraints, (iii) proposing a sort of augmented stochastic dominance through a safety margin. Numerical results show the difference between the alternative ways to interpret and apply the multivariate stochastic dominance. These results are evaluated thanks to the proposed economic cost of the stochastic dominance constraints and either in presence or not of a safety margin.

Keywords Stochastic programming · Portfolio selection · Sensitivity analysis · Asset and liability management · Pension fund · Stochastic dominance

1 Introduction

Asset and Liability Management (ALM) models are efficient and widely used tools to manage pension funds. The study of this class of problems moved from the work of Bradley and Crane

Miloš Kopa, Vittorio Moriggia, Sebastiano Vitali have contributed equally to this work.

✉ Vittorio Moriggia
vittorio.moriggia@unibg.it

Miloš Kopa
kopa@karlin.mff.cuni.cz

Sebastiano Vitali
sebastiano.vitali@unibg.it

¹ Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovska 83, 186 75 Prague, Czech Republic

² Department of Economics, University of Bergamo, Via dei Caniana 2, 24127 Bergamo, BG, Italy

(1972, 1980) where the authors proposed one of the first ALM model called BONDS model to help a portfolio manager to select bonds under uncertainty. Nevertheless, in that model the liability side was not analysed in deep. After this seminal work, other models were studied, see Kusy and Ziemba (1986), until the milestone model proposed in Dempster and Ireland (1988, 1989, 1991) and called MIDAS, focusing on the immunization of the liability side. The multistage stochastic optimization turned out to be a natural way to face ALM models. The Russell–Yasuda Kasai Model described in Cariño et al. (1994); Cariño and Ziemba (1998a); Cariño et al. (1998b) has been the first work to formulate an ALM problem as a multistage stochastic problem. In the same period, Mulvey and Zenios performed a complete analysis of the multistage stochastic problem applied to ALM problems considering the fixed income investment, see e.g. Mulvey (1994a, b), Nielsen and Zenios (1996) and Zenios (1995). Finally, their research produced the Tower Perrin scenario generation system and the well-known Towers Perrin-Tillinghast ALM model, see Mulvey et al. (2000).

All these ALM models did not specifically tackle a pension fund management problem which typically requires to consider additional features. For example, Pflug and Świetanowski (1999) introduced an ALM model for pension funds considering the specificities of both the asset and the liability side, while Consigli and Dempster (1998a, b) and Dempster et al. (2003) proposed the CALM model that considers a long-term target and jointly allows to have different types of pension contracts on the liability side. In general, there are two main characteristics that define the type of pension fund and, thus, the suitable type of ALM model. The former distinguishes the pension funds that follow a defined contribution from those that follow a defined benefit schema. In particular, in the defined contribution pension fund the final pension benefit is unknown and will be the result of the contribution investment, i.e. the pensioner bears the risk; in a defined benefit pension fund the final pension benefit is predetermined in the pension contract, i.e. the pension fund sponsor bears the risk. A specific focus for the defined benefit pension fund can be found in Dert (1998). The second characteristic relies on the pension pillar in which the pension fund is classified. Most of the countries distinguish three pension pillars: the first is the state pension system, the second is based on the worker category and/or on the employee's employer, the third is composed of private insurance contracts. For example, the InnoALM model developed in Ziemba (2007) and Geyer and Ziemba (2008) implicates a second pillar pension fund since it considers the pension fund of the employees of an electricity company. For a review, we suggest Zenios and Ziemba (2006, 2007) and Ziemba and Mulvey (1998). More innovative and comprehensive formulations of ALM models for second pillar pension funds having a defined benefit schema can be found in Consigli and Moriggia (2014), Consigli et al. (2017), and Moriggia et al. (2019).

In particular, our work moves from the model developed in Consigli et al. (2017); Moriggia et al. (2019); Vitali and Moriggia (2021) that will be discussed in the next sections.

More recent approaches extended the pension fund problem to consider not only the pension fund manager point of view, but also the pension fund sponsor (the issuer of the pension fund), see e.g. Vitali et al. (2017), and the pension fund investor, see e.g. Consigli et al. (2012), Kopa et al. (2018) and Consigli et al. (2019).

The main challenge of an ALM model is to improve the adaptability of the mathematical formulation to the real characteristics of the pension fund implementing the state-of-the-art in terms of risk control and portfolio selection. In particular, one of the most appreciated tools to improve the quality of a portfolio is the stochastic dominance which allows to compare different portfolios and, simultaneously, to prove the preference of one portfolio with respect to a benchmark portfolio for large classes of utility functions and, therefore, of economic agents.

The notion of stochastic dominance was introduced in statistics more than 50 years ago and it was firstly applied to economics and finance in Quirk and Saposnik (1962), Hadar and Russell (1969) and Hanoch and Levy (1969). Later on, the second-order stochastic dominance constraints were applied to static stochastic programs in Dentcheva and Ruszczyński (2003) and Luedtke (2008) and to portfolio efficiency analysis, see e.g. Post (2003), Kuosmanen (2004), Dupačová and Kopa (2012) and, more recently, Kopa and Post (2015). Similarly, the first-order stochastic dominance constraints were used in Kuosmanen (2004), Dentcheva and Ruszczyński (2004) or Dupačová and Kopa (2014). In multistage stochastic programming, the second-order stochastic dominance constraints were applied to asset-liability modeling in Yang et al. (2010) or Moriggia et al. (2019) and in an individual pension allocation problem in Kopa et al. (2018). In all the mentioned papers, the stochastic dominance constraints have been applied either on a single stage or on multiple stages separately. The latter approach is named multistage stochastic dominance. In particular, such approach requires the dominance on several stage, but the different relations are independent from each other, meaning that the way in which the dominance is fulfilled in one stage has no relation with the way it is fulfilled in another stage and, thus, the dominance relates a node with another node. In a recent contribution, Armbruster and Luedtke (2015) formulate a condition stronger than the multistage stochastic dominance, namely the multi-dimension (multivariate) stochastic dominance. Such approach requires that the dominance relation on a specific stage is the same of the dominance relation on another stage, i.e. the dominance is between scenarios instead of single nodes. This relation implies a time persistence of the stochastic dominance changing the interpretation of the dominance itself that becomes a scenario dominance rather than a nodal dominance. Further than this multi-dimension multivariate second-order stochastic dominance, we also analyze and discuss componentwise (Muller & Stoyan, 2002; Armbruster & Luedtke, 2015), linear (Dentcheva & Ruszczyński, 2009) and weak types of multivariate second degree stochastic orders.

Moving from the ALM models proposed in Consigli et al. (2017) and Moriggia et al. (2019), that already incorporate univariate stochastic dominance, the main contributions of this paper are the following:

- the comparison of the most recent alternatives proposed to formulate multivariate stochastic dominance orderings with a specific focus on the hierarchies among them;
- the implementation of each type of multivariate stochastic dominance in an ALM model in order to compare their different impact on the ALM objective variables;
- a detailed analysis of a decision variable of the model, the unexpected sponsor contribution, highlighting its relevance in the model and showing how this variable can be used to measure the cost of imposing stochastic dominance constraints;
- the proposal of an innovative way of stress testing of portfolios when stochastic dominance constraints are applied in pension fund models through a safety margin.

Thanks to these improvements, we extend the understanding of the ALM model and the comparison of its solution to a given benchmark. We introduce the multivariate stochastic dominance into the model concluding that it is more suitable than the univariate stochastic dominance in a multistage framework.

The paper is structured as follows. Section 2 recalls the formulation of the ALM model. Section 3 analyses the different types of multivariate stochastic dominance. Section 4 presents the setting of the model and the definition of the benchmark. Section 5 shows the numerical results and Sect. 6 concludes the paper.

2 Model description

As already mentioned, we propose an extension of the ALM model for pension fund proposed in Consigli et al. (2017); Moriggia et al. (2019). Therefore, following Consigli et al. (2017), the stochastic tree is differentiated between decisional nodes and intermediate nodes: in the decisional nodes the stochastic tree branches and the portfolio can be rebalanced, while in the intermediate nodes the stochastic tree does not branch and the model accounts the financial evolution of the variables (income cashflows, pension payments, etc.) but the portfolio is not rebalanced. The intermediate stages are every year. The set of decisional stages is denoted $\mathcal{T} = \{t_h\}_{h=0, \dots, H}$, where $t_0 = 0$ represents the here-and-now stage, then we take into account five further decisional stages at time 1, 2, 3, 5 and 10, and the final horizon at $t_H = 20$ years. The scenario tree is represented by the nodal notation and contains the asset returns: both the price returns and the income returns. For each node n we define $t(n)$ as the corresponding stage time and with $C(n)$ we denote the set of the children and nephews of n in all subsequent stages. For more detail on the nodal notation for a scenario tree, cf. Consigli et al. (2017).

Following Moriggia et al. (2019), only the already pensioned people generate liabilities. Therefore, we assume that the pension fund manager wants to ensure its sustainability in case there will not be new contracts in the next years and, then, we deal with the so called runoff case. In such framework, as time goes by, the number of pensioners decreases and consequently, in each node n also the net liability flows Λ_n^{NET} decrease. The pension fund has still to pay the net liabilities with the asset financial incomes and/or by selling them. Therefore, the liquidity gap and the ALM risk proposed in Consigli et al. (2017) have been adjusted considering the new liability flows and their duration. Moreover, contrary to the original model that adopts a portfolio replication approach, we define in each node n the Defined Benefit Obligation (DBO) D_n as the present value of the future payments in all children nodes weighted by the probability of each node.

$$D_n = \sum_{m \in C(n)} \left[e^{-r_{n,m}(t(m)-t(n))} \cdot \mathbb{E} \left[\Lambda_m^{NET} | t(m) = t_h \right], \forall t_h > t(n) \right] \quad (1)$$

where $r_{n,m}$ is the forward interest rate from node n till node m . The maximum surviving horizon for pensioners is 50 years. Then, we take into account the final conditions in 20-year final horizon because the fund continues for other 30 years.

For all other constraints of the ALM model, refer to Appendix A and to the cited references.

In the current work, it is of particular interest the role played by a variable that represents the unexpected sponsor contribution Φ_n . Indeed, for the purposes of our analysis, we assume that the pension fund is remarkably underfunded, i.e. the value of D_n is higher than the value of the portfolio. Nevertheless, the sponsor of the pension fund - who is the payer of last resort of the pension fund obligations - is willing to pay in case in some future stage the current value of the portfolio is not sufficient to pay the liabilities. This unexpected sponsor contribution is quantified by Φ_n and represents a direct injection of cash. In the optimal model Φ_n is a decision variable that appears also in the objective function where it is significantly penalized. However, its presence is fundamental for the definition of the benchmark used for the SD constraints and for the quantification of the economic costs generated by the introduction of the various types of SD constraints, as explained in Sect. 3.

2.1 Objective formulation

The objective function is a representation of a multicriteria approach that synthesizes a short-term risk control, a medium-term profitability and a long-term sustainability. The objective function aims to minimize the expected shortfall of a set K of variables Y_k with respect to a specific threshold \bar{Y}_k and, jointly, to maximize the expected values of the variables Y_k . The k -th variable is defined only in a specific stage t^k . The expected value of the different variables, as well as the expected shortfalls, are summed up considering a weight $\lambda_k > 0$ for each objective variable, with $\sum_{k=1}^K \lambda_k = 1$. The total expected value and the total expected shortfall are then combined through a risk-aversion coefficient α . Therefore, the objective function is:

$$\max \left\{ (1 - \alpha) \cdot \sum_{k=1}^K \lambda_k \cdot \mathbb{E}[Y_{k,n}] - \alpha \cdot \sum_{k=1}^K \lambda_k \cdot \mathbb{E}[\bar{Y}_k - Y_{k,n} | Y_{k,n} < \bar{Y}_k], \quad \forall n | t(n) = t^k \right\} \tag{2}$$

In the objective function, we consider the same four target variables $Y_{k,n}$ adopted in Moriggia et al. (2019) (for more details refer to Appendix A):

- $Y_{1,n}$: a joint measure of the ALM risk and of the liquidity gap,
- $Y_{2,n}$: a measure of the return adjusted by the risk,
- $Y_{3,n}$: the cumulative of the sponsor unexpected contribution Φ_n ,
- $Y_{4,n}$: the difference between the DBO and the portfolio value.

at stages $t^k = 1, 3, 10, 20$ and weights $\lambda_1 = 10\%$, $\lambda_2 = 30\%$, $\lambda_3 = 40\%$ and $\lambda_4 = 20\%$, respectively, while the risk-aversion coefficient α is set to 50%.

3 Multivariate stochastic dominance

Before discussing the multivariate second degree stochastic orders, we present a basic definition of the second-order stochastic dominance (SSD) in the univariate case:

A random variable A SSD dominates a random variable B ($A \geq B$) if the integrated cumulative distribution function of A is below the integrated cumulative distribution function of B , that is:

$$\int_{-\infty}^x F_A(y)dy \leq \int_{-\infty}^x F_B(y)dy \quad \forall x \in \mathbb{R}.$$

Equivalently, the second-order stochastic dominance holds if and only if no risk averse decision maker prefers B to A , that is:

$$Eu(A) \geq Eu(B)$$

for all concave non-decreasing functions u . Moving to the multivariate case, we consider the following four types of multivariate second-order stochastic dominance (MSSD) relations:

1. *Multidimension MSSD (MD-MSSD)*: A random vector $\mathbf{A} = (A_1, \dots, A_n)$ dominates a random vector $\mathbf{B} = (B_1, \dots, B_n)$ if

$$Eu(\mathbf{A}) \geq Eu(\mathbf{B})$$

for all concave non-decreasing (in each component) functions $u : \mathbb{R}^n \rightarrow \mathbb{R}$.

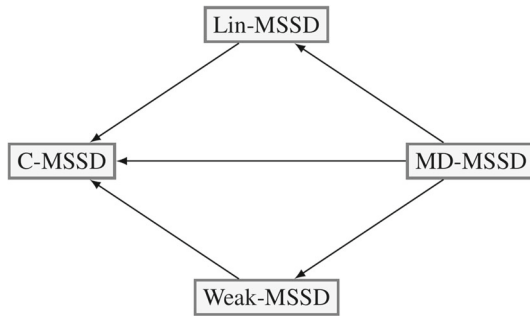


Fig. 1 Multivariate SSD relations

2. *Linear MSSD (Lin-MSSD)*: A random vector $\mathbf{A} = (A_1, \dots, A_n)$ dominates a random vector $\mathbf{B} = (B_1, \dots, B_n)$ if

$$\sum_{i=1}^n c_i A_i \geq \sum_{i=1}^n c_i B_i$$

for all $c_i \geq 0$ such that $\sum_{i=1}^n c_i = 1$.

3. *Componentwise MSSD (C-MSSD)*: A random vector $\mathbf{A} = (A_1, \dots, A_n)$ dominates a random vector $\mathbf{B} = (B_1, \dots, B_n)$ if $A_i \geq B_i$ for all $i = 1, 2, \dots, n$.
4. *Weak MSSD (Weak-MSSD)*: A random vector $\mathbf{A} = (A_1, \dots, A_n)$ dominates a random vector $\mathbf{B} = (B_1, \dots, B_n)$ if

$$\int_M F_{\mathbf{A}}(\mathbf{y}) d\mathbf{y} \leq \int_M F_{\mathbf{B}}(\mathbf{y}) d\mathbf{y}$$

for all n -dimensional intervals $M = (-\infty, m_1) \times (-\infty, m_2) \times \dots \times (-\infty, m_n)$.

For more details about MD-MSSD and C-MSSD we refer to Muller and Stoyan (2002) and Armbruster and Luedtke (2015). Lin-MSSD relation was extensively discussed, among others, in Dentcheva and Ruszczyński (2009), Dentcheva and Wolfhagen (2015), Dentcheva and Wolfhagen (2016) and Dentcheva et al. (2016). The last type of multivariate second-order stochastic dominance order (Weak-MSSD) can be seen as a modification of orthant dominance known as the multivariate extension of the first-order stochastic dominance, see Muller and Stoyan (2002) for more details.

Figure 1 summarizes implications among these four types of MSSD. In particular:

- when only positively affine (linear and non-decreasing in each variable) multivariate utility functions are considered in the definition of MD-MSSD, one gets directly a condition equivalent to Lin-MSSD, hence MD-MSSD implies Lin-MSSD;
- C-MSSD can be seen as a special case of Lin-MSSD when only binary values of c_i are considered. Therefore, Lin-MSSD implies C-MSSD;
- similarly to the univariate case, using integration by parts and properties of multivariate concave non-decreasing functions, one can derive that MD-MSSD implies Weak-MSSD, c.f. Levy (2006);
- similarly to the first-order multivariate stochastic dominance where the orthant dominance implies componentwise MFSD, one can show that Weak-MSSD implies C-MSSD, c.f. Muller and Stoyan (2002).

To show that the opposite implications do not hold we consider the following examples.

Example 1

Let $\mathbf{A} = (A_1, A_2)$ take values (1,0) and (0,1) with the same probabilities and let $\mathbf{B} = (B_1, B_2)$ take values (0,0) and (1,1) again with probabilities 0.5. Since the marginal distributions of \mathbf{A} are exactly the same as of \mathbf{B} , C-MSSD between \mathbf{A} and \mathbf{B} is trivially fulfilled. Moreover, $\sum_{i=1}^2 c_i A_i$ takes values c_1 and $c_2 = 1 - c_1$ with the same probabilities while $\sum_{i=1}^2 c_i B_i$ has alternative distribution with parameter 0.5. Hence $\sum_{i=1}^2 c_i A_i$ dominates $\sum_{i=1}^2 c_i B_i$ with respect to SSD, that is, Lin-MSSD holds true. Finally, it is easy to see that $F_{\mathbf{A}}(\mathbf{y}) \leq F_{\mathbf{B}}(\mathbf{y})$ for all $\mathbf{y} \in \mathbb{R}^2$, therefore \mathbf{A} dominates \mathbf{B} with respect to Weak-MSSD, too. Summarizing, we proved C-MSSD, Lin-MSSD and Weak-MSSD between \mathbf{A} and \mathbf{B} .

Let $u(x_1, x_2) = \min(x_1, x_2)$ that is a non-decreasing and concave function. Then $Eu(\mathbf{A}) = 0$ while $Eu(\mathbf{B}) = 0.5$ Hence, \mathbf{A} does not dominate \mathbf{B} with respect to MD-MSSD.

Example 2

Let \mathbf{A} and \mathbf{B} have a bivariate normal distribution with zero means and variance-covariance matrices:

$$V_{\mathbf{A}} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}, \quad V_{\mathbf{B}} = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 4 \end{pmatrix}$$

Since the marginal distributions of \mathbf{A} are the same as of \mathbf{B} , \mathbf{A} dominates \mathbf{B} with respect to C-MSSD.

The distributions of $\sum_{i=1}^2 c_i A_i$ and $\sum_{i=1}^2 c_i B_i$ are normal with zero mean and $\sigma_{\sum_{i=1}^2 c_i A_i}^2 = c_1^2 + 4c_2^2$, $\sigma_{\sum_{i=1}^2 c_i B_i}^2 = c_1^2 + 4c_2^2 - c_1c_2$, respectively. Since the variance of $\sum_{i=1}^2 c_i A_i$ is greater than the variance of $\sum_{i=1}^2 c_i B_i$ for $c_1 > 0$ and $c_2 = 1 - c_1 > 0$, $\sum_{i=1}^2 c_i A_i$ does not dominate $\sum_{i=1}^2 c_i B_i$ with respect to SSD for strictly positive c_1, c_2 and, hence, \mathbf{A} does not dominate \mathbf{B} with respect to Lin-MSSD.

Finally, one can easily show that $F_{\mathbf{A}}(\mathbf{y}) \geq F_{\mathbf{B}}(\mathbf{y})$ for all $\mathbf{y} \in \mathbb{R}^2$. Therefore, \mathbf{A} does not dominate \mathbf{B} with respect to Weak-MSSD.

Since the random variables in our model are discrete with equiprobable realizations it is useful to formulate the SSD conditions using a double stochastic matrix as proposed in Kuosmanen (2004), Luedtke (2008) and Armbruster and Luedtke (2015). In particular, if we define \mathbf{w}_{t_h} the vector of the optimal portfolio wealth realizations occurring in all nodes at stage t_h with the same probability and, similarly, we define $\mathbf{w}_{t_h}^B$ the vector of a benchmark portfolio wealth realizations occurring in all nodes at stage t_h with the same probabilities, we can assert that the optimal portfolio SSD dominates the benchmark portfolio at stage t_h if and only if

$$\mathbf{w}_{t_h} \geq \mathbf{Q}^{t_h} \cdot \mathbf{w}_{t_h}^B \tag{3}$$

for some matrix \mathbf{Q}^{t_h} which is double stochastic, i.e. satisfies the following conditions:

$$\sum_i Q_{i,j}^{t_h} = 1 \tag{4}$$

$$\sum_j Q_{i,j}^{t_h} = 1. \tag{5}$$

and elements of \mathbf{Q}^{t_h} have to belong to the interval $[0, 1]$, so each row and each column represents a convex combination.

Finally, the C-MSSD is obtained just by selecting jointly more than one stage, i.e. defining a subset $\mathcal{T}^{SSD} \subseteq \mathcal{T}$, and then defining the constraint:

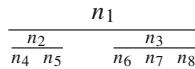
$$\mathbf{w}_{t_h} \geq \mathbf{Q}^{t_h} \cdot \mathbf{w}_{t_h}^B, \quad \forall t_h \in \mathcal{T}^{SSD} \tag{6}$$

Notice that the matrices \mathbf{Q}^{t_h} can differ from stage to stage.

The corresponding constraint becomes simply

$$\mathbf{w}_{t_h} \geq \mathbf{Q} \cdot \mathbf{w}_{t_h}^B, \quad \forall t_h \in \mathcal{T}^{SSD} \tag{7}$$

where, clearly, the matrix \mathbf{Q} is the same for all stages $t_h \in \mathcal{T}^{SSD}$ and then the vectors \mathbf{w}_{t_h} and $\mathbf{w}_{t_h}^B$ must be expanded to meet the dimension of \mathbf{Q} . For instance, assume we have a 5 scenarios tree having the following structure:



Then, the SD constraint on the last stage becomes:

$$\begin{bmatrix} w(n_4) \\ w(n_5) \\ w(n_6) \\ w(n_7) \\ w(n_8) \end{bmatrix} \geq \begin{bmatrix} Q(1, 1) & Q(1, 2) & Q(1, 3) & Q(1, 4) & Q(1, 5) \\ Q(2, 1) & Q(2, 2) & Q(2, 3) & Q(2, 4) & Q(2, 5) \\ Q(3, 1) & Q(3, 2) & Q(3, 3) & Q(3, 4) & Q(3, 5) \\ Q(4, 1) & Q(4, 2) & Q(4, 3) & Q(4, 4) & Q(4, 5) \\ Q(5, 1) & Q(5, 2) & Q(5, 3) & Q(5, 4) & Q(5, 5) \end{bmatrix} \cdot \begin{bmatrix} w^B(n_4) \\ w^B(n_5) \\ w^B(n_6) \\ w^B(n_7) \\ w^B(n_8) \end{bmatrix} \tag{8}$$

And, in the MD-MSSD case, the SD constraint on the previous stage becomes:

$$\begin{bmatrix} w(n_2) \\ w(n_2) \\ w(n_3) \\ w(n_3) \\ w(n_3) \end{bmatrix} \geq \begin{bmatrix} Q(1, 1) & Q(1, 2) & Q(1, 3) & Q(1, 4) & Q(1, 5) \\ Q(2, 1) & Q(2, 2) & Q(2, 3) & Q(2, 4) & Q(2, 5) \\ Q(3, 1) & Q(3, 2) & Q(3, 3) & Q(3, 4) & Q(3, 5) \\ Q(4, 1) & Q(4, 2) & Q(4, 3) & Q(4, 4) & Q(4, 5) \\ Q(5, 1) & Q(5, 2) & Q(5, 3) & Q(5, 4) & Q(5, 5) \end{bmatrix} \cdot \begin{bmatrix} w^B(n_2) \\ w^B(n_2) \\ w^B(n_3) \\ w^B(n_3) \\ w^B(n_3) \end{bmatrix} \tag{9}$$

where \mathbf{Q} of (9) is the same of (8). The SSD constraints (6) and (7) can be imposed in any stage $t_h, h > 0$.

Finally, we present a necessary and sufficient condition for Weak-MSSD between random vectors with equiprobable realizations $\mathbf{w}_{t_h} = (w_{t_h,1}, \dots, w_{t_h,K})'$ and $\mathbf{w}_{t_h}^B = (w_{t_h,1}^B, \dots, w_{t_h,K}^B)'$.

Theorem 1 *A random vector $(\tilde{\mathbf{w}}_{t_1}, \dots, \tilde{\mathbf{w}}_{t_H})'$ with equiprobable realizations $\mathbf{w}_{t_h} = (w_{t_h,1}, \dots, w_{t_h,K})', h = 1, \dots, H$ dominates random vector $(\tilde{\mathbf{w}}_{t_1}^B, \dots, \tilde{\mathbf{w}}_{t_H}^B)$ with equiprobable realizations $\mathbf{w}_{t_h}^B = (w_{t_h,1}^B, \dots, w_{t_h,K}^B)'$, $h = 1, \dots, H$ with respect to Weak-MSSD if and only if*

$$\sum_{k=1}^K \prod_{h=1}^H (z_h - w_{t_h,k})^+ \leq \sum_{k=1}^K \prod_{h=1}^H (z_h - w_{t_h,k}^B)^+ \quad \forall (z_1, \dots, z_H) \in \mathbb{R}^H, \tag{10}$$

where $(y)^+ = \max(0, y)$.

Proof: Since the random vector $(\tilde{\mathbf{w}}_{t_1}, \dots, \tilde{\mathbf{w}}_{t_H})'$ takes realizations with equal probabilities $1/K$ its cumulative distribution function could be formulated as follows:

$$\begin{aligned} F(y_1, \dots, y_H) &= P(\tilde{\mathbf{w}}_{t_1} \leq y_1, \dots, \tilde{\mathbf{w}}_{t_H} \leq y_H) = \frac{1}{K} \sum_{k=1}^K I(w_{t_1,k} \leq y_1, \dots, w_{t_H,k} \leq y_H) \\ &= \frac{1}{K} \sum_{k=1}^K \prod_{h=1}^H I(w_{t_h,k} \leq y_h) \end{aligned}$$

where $I(\dots)$ is the indicator function. Hence

$$\begin{aligned} \int_{-\infty}^{z_1} \dots \int_{-\infty}^{z_H} F(y_1, \dots, y_H) dy_H \dots dy_1 &= \int_{-\infty}^{z_1} \dots \int_{-\infty}^{z_H} \frac{1}{K} \sum_{k=1}^K \prod_{h=1}^H I(w_{t_h,k} \leq y_h) dy_H \dots dy_1 \\ &= \frac{1}{K} \sum_{k=1}^K \int_{-\infty}^{z_1} \dots \int_{-\infty}^{z_H} \prod_{h=1}^H I(w_{t_h,k} \leq y_h) dy_H \dots dy_1. \end{aligned}$$

Since

$$\begin{aligned} &\int_{-\infty}^{z_1} \dots \int_{-\infty}^{z_H} \prod_{h=1}^H I(w_{t_h,k} \leq y_h) dy_H \dots dy_1 \\ &= \int_{-\infty}^{z_1} \dots \int_{-\infty}^{z_{H-1}} \prod_{h=1}^{H-1} I(w_{t_h,k} \leq y_h) (z_H - w_{t_H,k})^+ dy_{H-1} \dots dy_1 \\ &= \int_{-\infty}^{z_1} \dots \int_{-\infty}^{z_{H-2}} \prod_{h=1}^{H-2} I(w_{t_h,k} \leq y_h) (z_H - w_{t_H,k})^+ (z_{H-1} - w_{t_{H-1},k})^+ dy_{H-2} \dots dy_1 \\ &\dots \\ &= \prod_{h=1}^H (z_h - w_{t_h,k})^+ \end{aligned}$$

for all $k = 1, 2, \dots, K$, we conclude that

$$\int_{-\infty}^{z_1} \dots \int_{-\infty}^{z_H} F(y_1, \dots, y_H) dy_H \dots dy_1 = \frac{1}{K} \sum_{k=1}^K \prod_{h=1}^H (z_h - w_{t_h,k})^+.$$

The same derivation could be done for cumulative distribution function of $\tilde{\mathbf{w}}_{t_h}^B$ what completes the proof.

To verify (10), one can solve the following optimization problem:

$$\max_{(z_1, \dots, z_H) \in \mathbb{R}^H} \sum_{k=1}^K \prod_{h=1}^H (z_h - w_{t_h,k})^+ - \sum_{k=1}^K \prod_{h=1}^H (z_h - w_{t_h,k}^B)^+. \tag{11}$$

Then (10) holds if and only if optimal objective value of (11) equals to zero, because: (i) positive optimal objective value implies violation of (10) and (ii) optimal objective value can not be negative (sufficiently small z_1, \dots, z_H give zero objective value).

Table 1 Asset universe

i	Asset	Asset type	Index/Underlying	Upper bound (%)	Initial portfolio
1	Bank account	Cash	EURIBOR 3-month	30	EUR 10,000
2	1–3 years	Treasury bond	Euro-Aggregate 1–3year	100	EUR 10,000
3	3–5 years	Treasury bond	Euro-Aggregate 3–5year		EUR 10,000
4	5–7 years	Treasury bond	Euro-Aggregate 5–7year		EUR 10,000
5	7–10 years	Treasury bond	Euro-Aggregate 7–10year		EUR 40,000
6	10+ years	Treasury bond	Euro-Aggregate 10+year		EUR 60,000
7	Securitized	Securitized bond	Euro-Aggregate: Securitized		EUR 20,000
8	Investment grade	Corporate bond	Euro Corporate ex Subordinated 1 Cap		EUR 20,000
9	High yield	Corporate bond	Euro HY B and above		EUR 20,000
10	Real estate	Real estate	GPR General Europe	20	EUR 150,000
11	MSCI	Public equity	MSCI Europe	50	EUR 150,000

4 Model setting

We consider an asset universe composed by 11 assets split in 4 classes: Cash, $i = 1$; Bonds, $i = 2, \dots, 9$; Real Estate, $i = 10$; Public Equity, $i = 11$. For each asset class we define an allocation upper bound according to the pension fund policy and regulatory constraints, see Table 1. Such limits are relatively loose and refer to a real case of a pension fund managed by an European insurance company. However, the main findings on the multivariate stochastic dominance have been confirmed also considering other asset settings.

The Treasury bond asset type is represented by five different maturity buckets. The Corporate bonds are subdivided into investment grade (rating higher than Baa3 Moody's or BBB-Standard&Poor's) and high yield (rating in the interval (Ba1, B3) Moody's or (BB+, B-) Standard&Poor's), see (Bertocchi et al, 2013, Chapter 5).

As shown in Table 1, each asset refers to a security that replicates a given index. For each index we have historical series of 17 years of quarterly returns, from the beginning of 1999 till the end of 2015. All series have been downloaded by DataStream. The scenario generation approach assumes that the process of the returns of each asset can be described as a linear regression of all assets, of the GDP and of the CPI. For more details on the scenario generation approach, please refer to Appendix B and to the cited references.

Moreover, we generate in each node the yield-curve for nominal interest rate using the Svensson model adopted by the European Central Bank. Consequently, using the simulated CPI, we generate also the real interest rate curves for each node.

The branching structure of the stochastic tree is 8-4-2-2-2-2, i.e. the root node has 8 children, each of them has 4 children, etc. Then, the tree is composed of 512 scenarios which grow over the 7 stages. Such choice is consistent with the number of scenarios in Consigli et al. (2017) that is adopted in practice by an insurance company that manages a pension fund and, thus, it is considered as representative of the degree of uncertainty faced by the portfolio manager. As in Moriggia et al. (2019), we slightly reduced the branching in the first stage to compensate the complexity induced by the stochastic dominance constraints.

4.1 Benchmark definition

The aim of this work is to analyse how the multivariate stochastic dominance impacts on the dynamic strategy of an underfunded pension fund, meaning that the current value of the asset is lower than the discounted value of the future pension payments. Indeed, we are tackling the case of a bank or an insurance (the sponsor of the pension fund) that assigns to a financial company (internal in the bank or in the insurance, or external, the pension fund manager) the task to manage the asset of the pension fund in order to pay the pension to pensioners. If, for any reason, the pension fund assets are not able to pay these liabilities, then the sponsor guarantees the pensions using the so-called unexpected sponsor contribution. In this perspective, our analysis considers the unexpected sponsor contribution needed to make the strategy feasible and the pension fund able to pay the future liabilities. Under this perspective, the benchmark against which we compare through the SD constraints is the wealth achieved by the pension fund keeping constant the allocation of the initial portfolio. This means that we assume that the most significant competitor for the optimal portfolio is the allocation actually implemented, i.e. the portfolio that is considered the best by the pension fund manager. Moreover, given the current situation of underfunding of the pension fund, we assume that the sponsor is already willing to cover the 50% of the future liabilities and, therefore, the expected sponsor contribution will correspond to the 50% of the pension payments in each node. Therefore, the benchmark wealth w_t^B is computed as follows:

1. We consider the initial portfolio wealth of EUR 500,000
2. We assume it is currently allocated as shown in Table 1
3. We assume it evolves over the same scenario tree used for the optimization (the same asset returns)
4. At each decisional node, we assume that the portfolio is rebalanced to return to the initial allocation
5. We assume that at each node (decisional and intermediate) the sponsor pays the 50% of the liabilities

The amount of sponsor's payments in this case, is clearly not optimal and not unexpected. Indeed, we call it *expected* sponsor contribution. Our aim is to run the optimal model without this expected sponsor contribution and leaving the possibility to the solver to ask for some *unexpected* sponsor contribution only when it is strictly needed to guarantee the sustainability of the pension fund. Then, comparing the expected sponsor contribution considered in the benchmark construction with the unexpected sponsor contribution required to obtain the optimal solution, we will be able to measure, on the one hand, the savings in terms of sponsor contribution by implementing an optimal investment strategy and, on the other hand, the economic costs in terms of sponsor contribution increment due to the imposition of the SD constraints. The results obtained using this benchmark are reported in Sect. 5.1. As it is reported in the next section, even considering an expected sponsor contribution of 50% of the pension payments, the benchmark portfolio allocation is so non-optimal that the benchmark turns out to be relatively weak and, therefore, not useful to investigate the effect of the different SD constraints. Thus, we assume that the pension fund sponsor wants to observe the optimal solution under stressed conditions represented by a so-called safety margin, i.e. the sponsor requires that the optimal portfolio is capable to generate a relevant extra amount of wealth with respect to the benchmark in the stages where the SD constraints are applied. In particular, when the SD is applied to the fourth stage the sponsor requires an extra wealth of EUR 300,000, when the SD is applied to the fifth stage the sponsor requires an extra wealth of EUR 300,000, and when the SD is applied to the sixth stage the sponsor requires an extra

Table 2 Size of the models according to the number of equations, variables and non zero elements

	Equations	Variables	Non zero elements
no SSD	845,656	921,655	4,201,757
SSD 4	845,848	925,751	4,214,109
SSD 5	846,040	938,039	4,251,037
SSD 6	846,424	987,191	4,398,621
SSD 7	847,192	1,183,799	4,988,701
C-MSSD 5-6-7	848,344	1,265,719	5,234,845
C-MSSD 4-5-6-7	848,536	1,269,815	5,247,197
MD-MSSD 5-6-7	848,984	1,265,719	5,677,853
MD-MSSD 4-5-6-7	849,624	1,269,815	5,948,701

wealth of EUR 250,000. The results obtained using this stronger benchmark that includes the safety margin are reported in Sect. 5.2.

5 Results

In this section, we produce the following results:

- The case without applying SSD constraints, i.e. guaranteeing the sustainability of the pension fund but without comparing the obtained wealth with any benchmark;
- The case with univariate SSD constraints on stages 4, 5, 6 and 7 (disjointly);
- The case with C-MSSD either on the set of stages 5-6-7 or on the set of stages 4-5-6-7;
- The case with MD-MSSD either on the set of stages 5-6-7 or on the set of stages 4-5-6-7.

When the SSD constraints are used, the benchmark is considered without safety margin in Sect. 5.1, and with safety margin in Sect. 5.2.

All the results are given by the implementation of a linear programming model solved by CPLEX 12.1.0 in GAMS, with an Intel(R) Core(TM) i7-8650U CPU 1.90GHz with 16.00GB RAM running Windows 10. Input management, parameter and coefficient computations, and output analysis are performed in MATLAB R2019a. Problem dimensions for each model are reported in Table 2.

For the larger size case, the solvers take less than one hour to find the optimal solution which is compatible with the pension fund manager requirements. Problem instances with larger dimension gave similar results but inducing memory issues.

5.1 Results considering the benchmark without safety margin

The results are reported highlighting the optimal here-and-now solution and showing the objective value of the multistage stochastic model. Moreover, we report the statistics of the distribution of the wealth at the final horizon and the amount of unexpected sponsor contribution required to make the pension fund able to pay the liabilities.

We begin analyzing the solution of the model applied to the non-stressed benchmark. Figure 2 presents the optimal results obtained solving the multistage stochastic model either without SD constraints, or with univariate SSD constraints, or with multivariate SSD constraints.

Let's first compare the no SSD optimal solution with the benchmark strategy. As we explained previously, implementing the benchmark strategy, i.e. a constant rebalancing to keep the initial proportion, the sponsor is willing to pay the 50% of the liabilities. This amount cumulated over the stages is, on average, EUR 290,870. Since the V@R and the AV@R of the wealth at the final horizon are below the average of the unexpected sponsor contribution, it is clear that without this injection of money the pension fund would not be able to pay the liability, at least on some scenarios. The optimal solution remarkably changes the here-and-now portfolio composition, reducing the allocation in the riskier assets, i.e. Real Estate and Public Equity, and increasing the allocation in Cash. Doing this, the portfolio is able to guarantee a sustainable solution requiring only EUR 12,589 from the sponsor, on average. However, the other statistics of the final wealth are somehow worse than the benchmark since the optimal solution does not aim at maximizing the final wealth, but it only aims at having a sustainable portfolio reaching the targets.

Since we also want to compare the wealth produced by the optimal solution to the wealth produced by the benchmark, we run the model with the SSD constraints. In the case of SSD on the fourth stage (SSD 4) we notice that the solution is very similar to the no SSD because the distribution of the wealth of the benchmark at the fourth stage is not particularly challenging so the achievement of the SD relation is relatively easy. The objective value slightly worsens and the unexpected sponsor contribution increases to EUR 28,281 while the statistics of the final wealth improves. We have the same results in the SSD 5 case, meaning that dominating the benchmark at stage 4 implies the dominance also at stage 5 and vice-versa. Moving to SSD 6 and SSD 7, the benchmark becomes more and more challenging. Indeed, the objective value reduces significantly while the unexpected sponsor contribution increases as well as the final wealth statistics. Especially in the hardest case SSD 7 we can notice that the optimal strategy achieves a final wealth improving all the statistics with respect to the benchmark. Moreover, the unexpected sponsor contribution reduces to EUR 180,721 compared to EUR 290,870 required in the benchmark case.

Considering the multivariate SSD cases, it is clear that the dominance at stage 5 implies the dominance at stage 4 as highlighted in the univariate case. Moreover, since the SSD 7 solution coincides with the C-MSSD 5-6-7, it is obvious that the dominance at stage 7 implies all the others. The difference between the MD-MSSD and the C-MSSD is insignificant both in terms of unexpected sponsor contribution and in terms of objective value, while it is null in terms of here-and-now solution.

Indeed, it appears that imposing univariate SSD constraints or C-MSSD or MD-MSSD does not have a clear effect on the optimal solution. Such behaviour should be driven by the fact that the benchmark is relatively weak in the inner stages and the really binding constraints are only at stage 7. Such conclusions are strengthened by the evidences in Table 3 where we show the distributions of the wealth achieved by the benchmark and by the optimal solution at the stages 4, 5, 6, 7. In particular, in the first row, we can see the results obtained applying the univariate SSD in each single stage 4, or 5, or 6, or 7. In the second row, there are the results obtained by applying the C-MSSD on stages 5-6-7, and, in the third row, by applying the C-MSSD to stages 4-5-6-7. Similarly, in the fourth row, we can compare the results obtained by applying the MD-MSSD to stages 5-6-7, and, in the fifth row, by applying the MD-MSSD on stages 4-5-6-7. As already explained above, when the solution dominates the distribution of the wealth of the benchmark at the stage 7, the dominance holds also in the previous stages. Therefore, we can conclude that when the benchmark is relatively weak, C-MSSD and MD-MSSD almost coincides. For this reason, in Sect. 5.2, we apply a stressed benchmark making the SD constraints active in all stages when applying the multivariate SD.

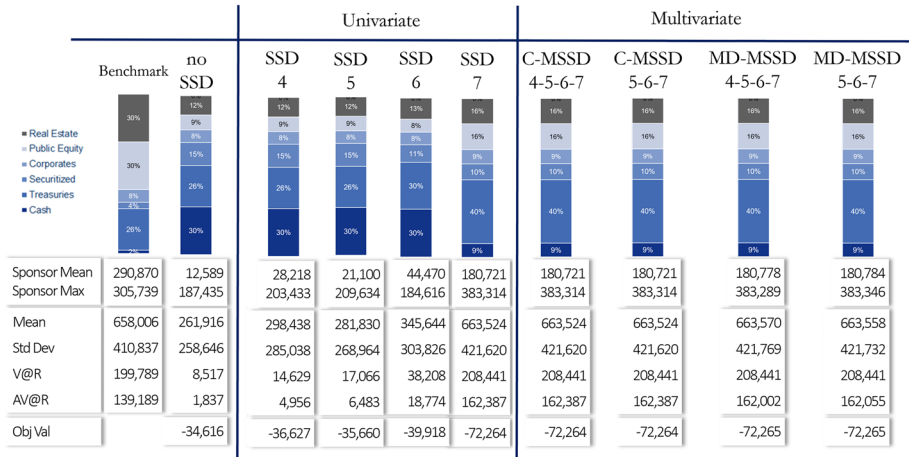
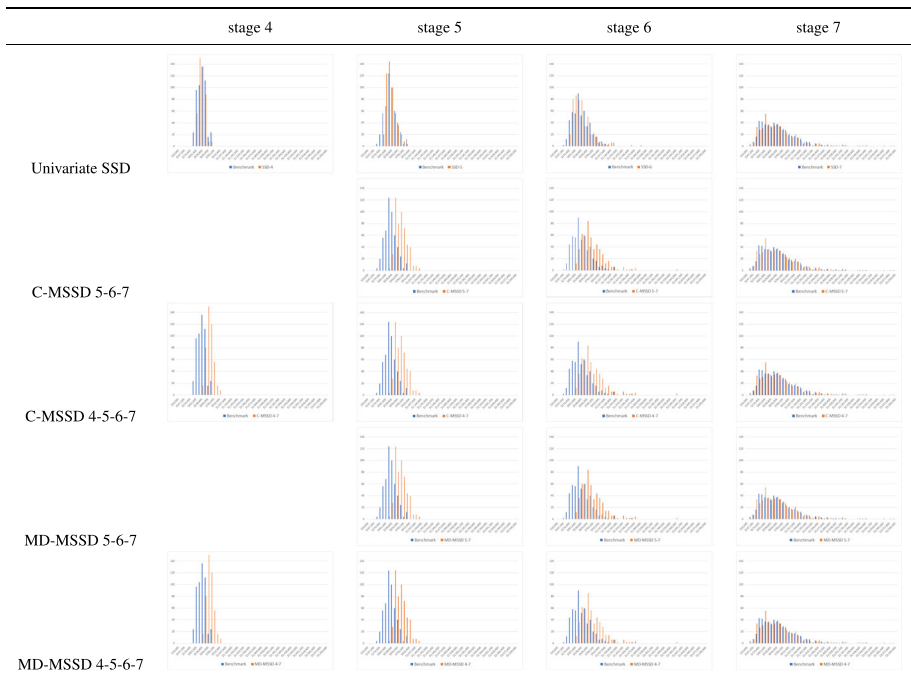


Fig. 2 Benchmark allocation and optimal here-and-now allocations obtained with the proposed model using the non-stressed benchmark. For each solution, the unexpected sponsor contribution, the statistics of the wealth of the pension fund at the final horizon, and the optimal objective value are reported

Table 3 Distributions of the wealth variable of the optimal portfolio (orange) and of the benchmark (blue) without safety margin



5.2 Results considering the benchmark with safety margin

Figure 3 reports the optimal solution of the model without SD constraints for sake of comparison, and the optimal solution of the model with univariate SSD constraints and with multivariate SSD constraints applied with respect to the stressed benchmark. As mentioned before, the stressed benchmark represents a situation in which the pension fund is more robust in terms of the dominance and, thus, requires a safety margin of EUR 300,000, in the fourth and fifth stage and of EUR 250,000 in the sixth stage. These margins produce the effect to activate the multivariate SD constraints also in the inner stages. Indeed, we observe that the univariate SSD indicates more aggressive here-and-now allocation that makes possible to beat the stressed benchmark still requiring less unexpected contribution to the sponsor.

Of particular interest is the behaviour of the multivariate SSD. Here, the difference between the C-MSSD 4-5-6-7 and the C-MSSD 5-6-7 is more sensible showing that dominating the stages 5-6-7 does not imply to dominate the benchmark also in the fourth stage. Moreover, the here-and-now portfolios are different and the C-MSSD 5-6-7 allocation is slightly less risky. It is also tangible the difference between the C-MSSD and the MD-MSSD proving that the different mathematical formulation reflects in an empirical and real-life application when the constraints are active on all the considered stages. Indeed, C-MSSD 4-5-6-7 and MD-MSSD 4-5-6-7 solutions show that the MD-MSSD is stronger because it requires more sponsor contribution but also provides better final wealth statistics. Similarly, MD-MSSD 5-6-7 solution is more expensive for the sponsor than C-MSSD 5-6-7 solution but, again, better in terms of final wealth statistics. However, the stronger MD-MSSD does not cost too much neither in terms of unexpected sponsor contribution (EUR 269,461 vs EUR 268,824 and EUR 226,591 vs EUR 221,591) nor in terms of objective value, i.e. of target achievement (-92,817 vs -92,029 and -84,203 vs -83,322). It is interesting to notice that now the here-and-now portfolio allocation of the MD-MSSD 4-5-6-7 is less risky than C-MSSD 4-5-6-7 which is instead equal to SSD 4. The same effect (even if less relevant) appears comparing the here-and-now portfolio of MD-MSSD 5-6-7, C-MSSD 5-6-7 and SSD 5. Such evidence suggests that to dominate the benchmark in stage 4 (or in stage 5) almost implies to dominate the benchmark in the subsequent stages. In our opinion, this is one of the effect of the multistage stochastic dominance when it is applied to a variable that evolves and accumulates over the stages, like the wealth variable that we use in our model. To prove this observation, we ran again all the cases for several scenario trees (both with safety margin and without) and we observed that these similarities persist.

In Table, we show the distributions of the wealth achieved by the stressed benchmark and by the optimal solution at the stages 4, 5, 6, 7. In particular, in the first row, we show the results applying the univariate SSD either on the stage 4, or 5, or 6, or 7. In the second row, we show the results obtained by applying the C-MSSD on stages 5-6-7, and, in the third row, by applying the C-MSSD on stages 4-5-6-7. Similarly, in the fourth row, we show the results obtained by applying the MD-MSSD on stages 5-6-7, and, in the fifth row, by applying the MD-MSSD on stages 4-5-6-7. We can notice that in the multivariate SSD cases the distribution of the wealth of the optimal solution and of the benchmark are very close to each other showing that the SD constraints are active. Still, a careful look indicates how the distributions obtained under the C-MSSD and under the corresponding MD-SSD differs from each other.

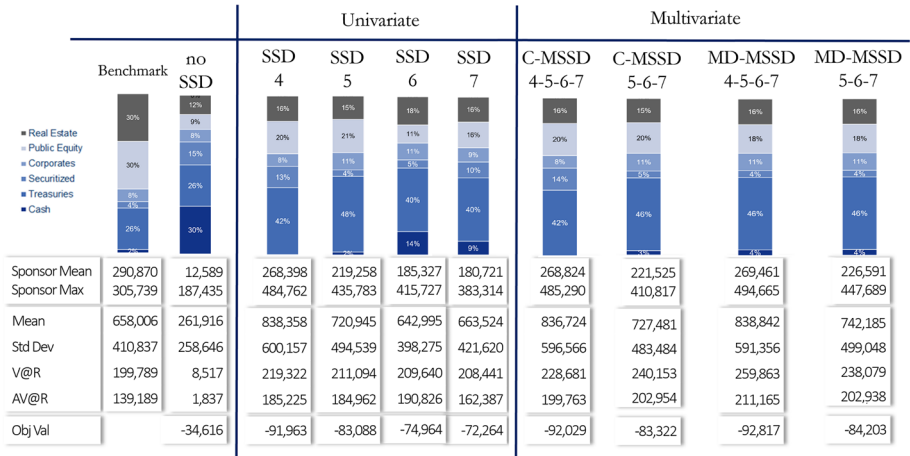
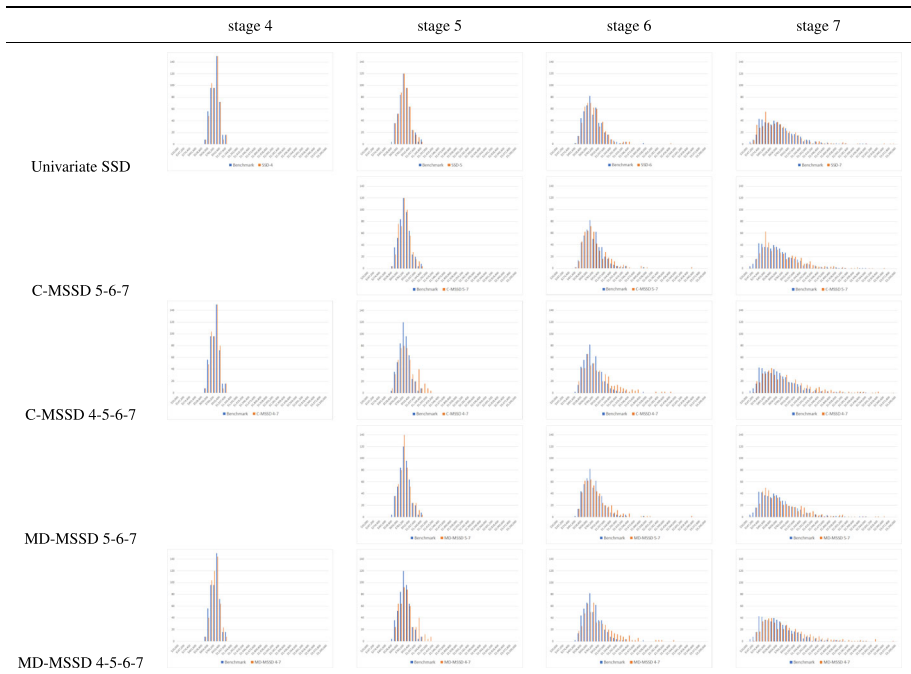


Fig. 3 Benchmark allocation and optimal here-and-now allocations obtained with the proposed model using the stressed benchmark. For each solution, we indicate the unexpected sponsor contribution, the statistics of the wealth of the pension fund at the final horizon, and the optimal objective value

Table 4 Distributions of the wealth variable of the optimal portfolio (orange) and of the benchmark (blue) with safety margin



6 Conclusion

In this work, we propose three main innovations with respect to the state-of-the-art of ALM models. The MD-MSSD shows to be an effective tool for a pension fund manager and, more generally, for a decision maker, to compare the optimal solution with a given benchmark. Indeed, the results of the MD-MSSD case are consistent with the results obtained using other types of SSD constraints. In our opinion, this happens when the stochastic dominance is imposed considering a variable that accumulates over the stages. However, in a multistage problem, the MD-MSSD is more suitable than other types of SSD constraints since it provides a dominance relation between scenarios (and not only between nodes). Moreover, when the benchmark is particularly hard to dominate, the differences between the different types of SSD constraints emerge in the sense that MD-MSSD suggests a more conservative portfolio than C-MSSD; it is a stronger and more demanding condition and, thus, it requires more sponsor contribution; and it generates a wealth process with better statistics on the final wealth. In a similar way the C-MSSD relates to the univariate SSD. The computing effort to solve a model with MD-MSSD increases, but the problem is still tractable.

The other contributions regard the economic cost of the stochastic dominance constraints and the idea of including a safety margin on the benchmark. In our contest, we adopt the sponsor contribution variable as measure of the cost of including SSD constraints. It is easy to interpret and it gives an effective picture of the impact of the dominance constraints on the solution. Such measure could be easily extended to other models in which it is possible to include a slack variable to measure the extra cost that the decision maker should pay to dominate the benchmark. On the other hand, if the benchmark appears unrealistically too much easy to dominate, our results show that the decision maker can introduce a safety margin that makes the solution more robust, especially when the problem requires some intermediate targets.

Acknowledgements The research of Miloš Kopa was partially supported by the Czech Science Foundation (Grant No. 19-28231X). The research of Vittorio Moriggia was supported by MIUR-ex60% 2021 sci.resp. Vittorio Moriggia. The research of Sebastiano Vitali was supported by MIUR-ex60% 2021 and 2022 sci.resp. Sebastiano Vitali.

Funding Open access funding provided by Università degli studi di Bergamo within the CRUI-CARE Agreement.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Consent for publication The authors have consented to the submission of the work to the journal.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

A The ALM model

The ALM problem is formulated as a linear programming problem. Here we describe the equations of the constraints used in the version of the model adopted in this paper to produce the results shown within the empirical section. For more details see Consigli et al. (2017). These are the decision variable of the model:

$x_{i,n}^+$ investment in node n , of asset i ;
 $x_{i,n}^-$ selling in node n , of asset i ;
 $x_{i,n}$ holding in node n , of asset i ;
 $z_n = z_n^+ - z_n^-$ cash account in node n .

Inventory balance constraints: The inventory balance constraints affect the time evolution of each asset, and the sum of the holding in each asset determines the wealth w_n of the pension fund:

$$\begin{aligned} x_{i,0} &= \hat{x}_i + x_{i,0}^+ - x_{i,0}^- \quad \forall i, \\ w_0 &= \sum_i x_{i,0}. \end{aligned} \quad (12)$$

where \hat{x}_i is the initial allocation, and

$$\begin{aligned} x_{i,n} &= x_{i,n-} (1 + \rho_{i,n}) - x_{i,n}^- + x_{i,n}^+ \quad \forall i, n, \\ w_n &= \sum_i x_{i,n} \quad \forall n. \end{aligned} \quad (13)$$

where $\rho_{i,n}$ is the price return of asset i at node n .

Cash balance constraints: We consider cash outflows due to: liability payments (L_n), negative interest (ζ_n^-) on cash account deficits, corporate taxes (T_n), buying decisions ($x_{i,n}^+$) and operating and human resource costs (O_n); cash inflows are due to: insurance premiums (R_n), selling decisions ($x_{i,n}^-$), interest (ζ_n^+) on cash account surpluses and the unexpected sponsor contribution (Φ_n). Given an initial cash balance \hat{z} :

$$\hat{z} + \sum_i x_{i,0}^- - \sum_i x_{i,0}^+ + z_0^+ - z_0^- = 0. \quad (14)$$

the following cash balance constraints become:

$$\begin{aligned} & z_n^+ (1 + \zeta_n^+) - z_n^- (1 + \zeta_n^-) - z_n^+ + z_n^- \\ & + \sum_i x_{i,n}^- - \sum_i x_{i,n}^+ \\ & + R_n - L_n - O_n - T_n + \Phi_n = 0, \forall n. \end{aligned} \quad (15)$$

Bounds on the asset portfolio: Rebalancing decisions are typically constrained to maintain a sufficient portfolio diversification:

$$l_i w_n \leq x_{i,n} \leq u_i w_n, \forall i, n \quad (16)$$

where l_i and u_i are the lower and upper bounds on holdings in asset i with respect to the current portfolio position. A maximum turnover (γ) will limit portfolio rebalancing from one stage to the next.

$$\sum_i x_{i,n}^- + \sum_i x_{i,n}^+ \leq \gamma \cdot \left[\sum_i x_{i,n-} \cdot (1 + \rho_{i,n}) \right], \forall n. \quad (17)$$

For the definition of the variable in the objective function we need also to compute:

- $Y_{1,n}$: a joint measure of the ALM risk and of the liquidity gap: we compute the mismatch between the average duration of the assets and the duration of the liability (ALM risk) and we associate a liquidity coefficient to each asset to compute the average liquidity of the portfolio;
- $Y_{2,n}$: a measure of the return adjusted by the risk: we consider the wealth w_n of the portfolio and we penalize it with the riskiness of the portfolio measured by specific risk coefficients associated to each asset;
- $Y_{3,n}$: the cumulative of the sponsor unexpected contribution Φ_n : we simply compute the sum of Φ_n over the ancestor nodes of n ;
- $Y_{4,n}$: the difference between the DBO D_n and the portfolio value w_n .

B Scenario generation

We generate the scenarios starting from the historical series of quarterly price and income returns for all the assets. We assume that each process of price return can be described as a linear regression of all asset price returns and of two main macroeconomic variables: Gross Domestic Product (GDP) and Consumer Price Index (CPI). Regressors are initially included with 5 lags and we also assume that they can be included with lag 0 in a hierarchical sense, that means that according to the order in Table 1 variable $i = 1, \dots, n$ can depend on variables $j = 1, \dots, i - 1$ at the same time, so that, excluding simultaneous equations, we obtain the following general model for the price return of asset i :

$$\begin{aligned} \rho_{i,t} = & \beta_{i,0} + \sum_{j=1}^{i-1} \beta_{i,j,0} \rho_{j,t} + \sum_{j=1}^n \sum_{l=1}^5 \beta_{i,j,l} \rho_{j,t-l} \\ & + \sum_{v=1,2}^5 \sum_{l=1}^5 \gamma_{i,v,l} w_{v,t-l} + \epsilon_{i,t}, \quad i = 1, \dots, n \end{aligned}$$

where $\rho_{i,t}$ is the price return of the asset i at time t , $w_{v,t}$ is the price return of the macroeconomic variable v at time t , $\beta_{i,j,l}$ and $\gamma_{i,v,l}$ are the coefficients to estimate, and $\epsilon_{i,t}$ is the error term. The same process is needed also to determine the regression for the macroeconomic variables:

$$w_{v,t} = \alpha_{v,0} + \sum_{i=1}^n \sum_{l=0}^5 \alpha_{v,i,l} \rho_{i,t-l} + \sum_{j=1,2}^5 \sum_{l=1}^5 \kappa_{v,j,l} w_{j,t-l} + \xi_{v,t}, \quad v = 1, 2$$

where $\alpha_{v,i,l}$ and $\kappa_{v,j,l}$ are the coefficients to estimate, and $\xi_{v,t}$ is the error term. Then, for each linear regression, we proceed iteratively. At each step we estimate the β , γ , α and κ coefficients, we remove the most non-significant one and we estimate again until all coefficients are statistically significant. Finally, if the associated R^2 statistics is large enough, we assume that the model is reliable to be used for further estimations, otherwise we estimate only the regression $\rho_{i,t} = \beta_{i,0} + \epsilon_{i,t}$ (or $w_{v,t} = \alpha_{v,0} + \xi_{v,t}$ for the macroeconomic variables) and we assume that the underlying process is a geometric Brownian motion having $\mu = \beta_{i,0}$ and $\sigma = \sigma(\epsilon_{i,t})$ (or $\mu = \alpha_{v,0}$ and $\sigma = \sigma(\xi_{v,t})$ for the macroeconomic variables). Once the models have been defined for all assets and macroeconomic variables, we adopt a Monte Carlo approach to generate the tree nodal values for price returns. For all assets, except the

cash, we proceed in a similar way to define the tree nodal values of the income returns. For more details see Moriggia et al. (2019).

References

- Armbruster, B., & Luedtke, J. (2015). Models and formulations for multivariate dominance-constrained stochastic programs. *IIE Transactions*, *47*(1), 1–14.
- Bertocchi, M., Consigli, G., D’Ecclesia, R., et al. (2013). Euro bonds: Markets, infrastructure and trends. World 611 Scientific Books.
- Bradley, S. P., & Crane, D. B. (1972). A dynamic model for bond portfolio management. *Management Science*, *19*(2), 139–151.
- Bradley, S.P., & Crane, D.B. (1980). Managing a bank bond portfolio over time. In: Dempster M.A.H. (Ed) *Stochastic Programming*. Academic Press, pp. 449–471.
- Cariño, D. R., & Ziemba, W. T. (1998). Formulation of the Russell-Yasuda Kasai financial planning model. *Operations Research*, *46*(4), 433–449.
- Cariño, D. R., Kent, T., Myers, D. H., et al. (1994). The Russell-Yasuda Kasai model: An asset/liability model for a Japanese insurance company using multistage stochastic programming. *Interfaces*, *24*(1), 29–49.
- Cariño, D. R., Myers, D. H., & Ziemba, W. T. (1998). Concepts, technical issues, and uses of the Russell-Yasuda Kasai financial planning model. *Operations Research*, *46*(4), 450–462.
- Consigli, G., & Dempster, M. A. H. (1998). The CALM stochastic programming model for dynamic asset-liability management. *Worldwide Asset and Liability Modeling*, *10*, 464.
- Consigli, G., & Dempster, M. A. H. (1998). Dynamic stochastic programming for asset-liability management. *Annals of Operations Research*, *81*, 131–162.
- Consigli, G., & Moriggia, V. (2014). Applying stochastic programming to insurance portfolios stress-testing. *Quantitative Finance Letters*, *2*(1), 7–13.
- Consigli, G., Iaquina, G., Moriggia, V., et al. (2012). Retirement planning in individual asset-liability management. *IMA Journal of Management Mathematics*, *23*(4), 365–396.
- Consigli, G., Moriggia, V., Benincasa, E., et al. (2017) Optimal multistage defined-benefit pension fund management. In: Consigli G, Stefani S, Zambruno G (Eds) *Recent Advances in Commodity and Financial Modeling: Quantitative methods in Banking, Finance, Insurance, Energy and Commodity markets*. Springer’s International Series in Operations Research and Management Science. Springer, Cham.
- Consigli, G., Moriggia, V., & Vitali, S. (2020). Long-term individual financial planning under stochastic dominance constraints. *Annals of Operations Research*, *292*, 973–1000.
- Dempster, M. A. H., & Ireland, A. M. (1988). MIDAS: An expert debt management advisory system. *Data, Expert Knowledge and Decisions* (pp. 116–127). Springer.
- Dempster, M. A. H., & Ireland, A. M. (1989). Object-oriented model integration in MIDAS (Manager’s Intelligent Debt Advisory System). In: *Proceedings of the Twenty-Second Annual Hawaii International Conference on System Sciences. Vol. III: Decision Support and Knowledge Based Systems Track*, IEEE Computer Society, pp. 612–620.
- Dempster, M. A. H., & Ireland, A. M. (1991). Object-oriented model integration in a financial decision support system. *Decision Support Systems*, *7*(4), 329–340.
- Dempster, M. A. H., Germano, M., Medova, E. A., et al. (2003). Global asset liability management. *British Actuarial Journal*, *9*(01), 137–195.
- Dentcheva, D., & Ruszczyński, A. (2003). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, *14*(2), 548–566.
- Dentcheva, D., & Ruszczyński, A. (2004). Semi-infinite probabilistic optimization: First-order stochastic dominance constrain. *Optimization*, *53*(5–6), 583–601.
- Dentcheva, D., & Ruszczyński, A. (2009). Optimization with multivariate stochastic dominance constraints. *Mathematical Programming*, *117*, 111–127. <https://doi.org/10.1007/s10107-007-0165-x>
- Dentcheva, D., & Wolfhagen, E. (2015). Optimization with multivariate stochastic dominance constraints. *SIAM Journal on Optimization*, *25*(1), 564–588.
- Dentcheva, D., & Wolfhagen, E. (2016). Two-stage optimization problems with multivariate stochastic order constraints. *Mathematics of Operations Research*, *41*(1), 1–22.
- Dentcheva, D., Martinez, G., & Wolfhagen, E. (2016). Augmented Lagrangian methods for solving optimization problems with stochastic-order constraints. *Operations Research*, *64*(6), 1451–1465.
- Dert, C. L. (1998). A dynamic model for asset liability management for defined benefit pension funds. *Worldwide Asset and Liability Modeling*, *10*, 501–536.

- Dupačová, J., & Kopa, M. (2012). Robustness in stochastic programs with risk constraints. *Annals of Operations Research*, 200(1), 55–74.
- Dupačová, J., & Kopa, M. (2014). Robustness of optimal portfolios under risk and stochastic dominance constraints. *European Journal of Operational Research*, 234(2), 434–441.
- Geyer, A., & Ziemba, W. T. (2008). The Innovest Austrian pension fund financial planning model InnoALM. *Operations Research*, 56(4), 797–810.
- Hadar, J., & Russell, W. R. (1969). Rules for ordering uncertain prospects. *The American Economic Review*, 59(1), 25–34.
- Hanoch, G., & Levy, H. (1969). The Efficiency Analysis of Choices Involving Risk. *The Review of Economic Studies*, 36(3), 335–346. <https://doi.org/10.2307/2296431>
- Kopa, M., & Post, T. (2015). A general test for SSD portfolio efficiency. *OR Spectrum*, 37(3), 703–734.
- Kopa, M., Moriggia, V., & Vitali, S. (2018). Individual optimal pension allocation under stochastic dominance constraints. *Annals of Operations Research*, 260(1–2), 255–291.
- Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science*, 50(10), 1390–1406.
- Kusy, M. I., & Ziemba, W. T. (1986). A bank asset and liability management model. *Operations Research*, 34(3), 356–376.
- Levy, H. (2006). Stochastic dominance: Investment decision making under uncertainty. Springer.
- Luedtke, J. (2008). New formulations for optimization under stochastic dominance constraints. *SIAM Journal on Optimization*, 19(3), 1433–1450.
- Moriggia, V., Kopa, M., & Vitali, S. (2019). Pension fund management with hedging derivatives, stochastic dominance and nodal contamination. *Omega*, 87, 127–141.
- Muller, A., & Stoyan, S. (2002). *Comparison methods for stochastic models and risks*. Wiley.
- Mulvey, J. M. (1994). An asset-liability investment system. *Interfaces*, 24(3), 22–33.
- Mulvey, J. M. (1994b). Financial planning via multi-stage stochastic programs. *Mathematical Programming: State of the Art*
- Mulvey, J. M., Gould, G., & Morgan, C. (2000). An asset and liability management system for Towers Perrin-Tillinghast. *Interfaces*, 30(1), 96–114.
- Nielsen, S. S., & Zenios, S. A. (1996). A stochastic programming model for funding single premium deferred annuities. *Mathematical Programming*, 75(2), 177–200.
- Pflug, G. C., & Świetanowski, A. (1999). Dynamic asset allocation under uncertainty for pension fund management. *Control and Cybernetics*, 28, 755–777.
- Post, T. (2003). Empirical tests for stochastic dominance efficiency. *The Journal of Finance*, 58(5), 1905–1931.
- Quirk, J. P., & Saposnik, R. (1962). Admissibility and measurable utility functions. *The Review of Economic Studies*, 29, 140–146.
- Vitali, S., & Moriggia, V. (2021). Pension fund management with investment certificates and stochastic dominance. *Annals of Operations Research*, 299, 273–292. <https://doi.org/10.1007/s10479-020-03855-7>
- Vitali, S., Moriggia, V., & Kopa, M. (2017). Optimal pension fund composition for an Italian private pension plan sponsor. *Computational Management Science*, 14(1), 135–160.
- Yang, X., Gondzio, J., & Grothey, A. (2010). Asset liability management modelling with risk control by stochastic dominance. *Journal of Asset Management*, 11(2), 73–93.
- Zenios, S. A. (1995). Asset/liability management under uncertainty for fixed-income securities. *Annals of Operations Research*, 59(1), 77–97.
- Zenios, S. A., & Ziemba, W. T. (2006). *Handbook of Asset and Liability Management: Theory and Methodology*. North-Holland Finance Handbook Series (Vol. 1). Elsevier.
- Zenios, S. A., & Ziemba, W. T. (2007). *Handbook of Asset and Liability Management: Applications and case studies*. North-Holland Finance Handbook Series (Vol. 2). Elsevier.
- Ziemba, W. T. (2007). In S. A. Zenios & W. T. Ziemba (Eds.), *Handbook of Asset and Liability Management: Applications and case studies*. The Russell-Yasuda, InnoALM and related models for pensions, insurance companies and high net worth individuals. North-Holland Finance Handbook Series (pp. 861–962). Elsevier.
- Ziemba, W. T., & Mulvey, J. M. (Eds.). (1998). *Worldwide asset and liability modeling* (Vol. 10). Cambridge University Press.