On the number of stages in multistage stochastic programs

Giovanni Pantuso, Trine K. Boomsma
Department of Mathematical Sciences, University of Copenhagen,
Universitetsparken 5, 2100 Copenhagen, Denmark
gp@math.ku.dk, trine@math.ku.dk

February 14, 2019

Abstract

Multistage stochastic programs serve to make sequential decisions conditional on a gradual realization of some stochastic process. The progressive arrival of new information divides a problem into stages. As the number of stages increases, however, the applicability of stochastic programming is halted by the curse of dimensionality. A commonly used approximation is obtained by replacing the random parameters by their expected values, which results in a problem with fewer stages. The present paper suggests metrics to support the choice of the number of stages: The value of an extended Expected Value (EV) problem which accounts for stochastic approximations, the Expected result of using the Expected Value solution (EEV), the Value of the Stochastic Solution (VSS), and the values of two Wait-and-See (WS) approximations. We show that for linear minimization problems with random right-hand-side, the value of the EV problem provides a lower bound that improves with the number of stages. The same holds for the WS approximations for general linear minimization problems and with one of the approximations improving the EV bound. In contrast, we show that the upper bound provided by the EV solution may not improve with the number of stages. Finally, we define the marginal benefit of including an additional stage in an EV approximation and demonstrate how to use it in a heuristic. We apply our approach to a hedging problem and confirm that increasing the number of stages does not necessarily foster better decisions.

1 Introduction

Stochastic programming is a well-recognized approach to decision-making under uncertainty. It assumes that one or more parameters of a decision problem can be described by random variables with a known distribution. In its most simple version, some decisions must made independent of the realization of uncertainty, whereas other (recourse) decisions can adapt. The division of decisions produces two *stages* and the resulting problem is known as a two-stage stochastic recourse program. An extended version serves to make sequential decisions conditional on a gradual realization of some stochastic process. The arrivals of new information divides the problem into multiple stages and the problem is referred to as a multistage stochastic recourse program (MSRP). For an introduction to stochastic programming, see Kall and Wallace (1994), Birge and Louveaux (1997), King and Wallace (2012), Pflug and Pichler (2014).

Problems with such a structure arise, for instance, in asset/liability management Cariño et al. (1994); Mulvey and Shetty (2004), capacity expansion Ahmed et al. (2003), energy systems Wallace and Fleten (2003); Fleten and Kristoffersen (2008); Kristoffersen and Fleten (2010), water resource management Li et al. (2006), sports Pantuso (2017), air transportation Alonso et al. (2000), and maritime transportation Pantuso et al. (2015); Bakkehaug et al. (2014). Stages often represent the points in time at which new information arrives, e.g. the realization of future returns, demand and/or supply, or market prices.

We adopt the following formulation of a MSRP:

$$RP = \min \Big\{ \mathbb{E}_\xi \Big[z(x_1, \ldots, x_T, \xi) := \sum_{\tau=1}^{T} c_\tau x_\tau \Big] \Big| A_{\tau-1,\tau} x_{\tau-1} + A_{\tau,\tau} x_\tau = b_\tau,\ x_\tau \in X_\tau, \\ x_\tau \text{ is } \mathcal{F}_\tau \text{ -measurable},\ \tau = 1, \ldots, T \Big\}, \tag{1}$$

where $1, \ldots, T$ are stages and $\xi := \{\xi_t : t = 1, \ldots, T\}$ is a stochastic process defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with the filtration $\mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_T$. Here, ξ_τ is a stochastic vector collating the random components of the parameters $c_\tau \in \mathbb{R}^{n_\tau}, b_\tau \in \mathbb{R}^{m_\tau}, A_{\tau-1,\tau} \in \mathbb{R}^{m_\tau} \times \mathbb{R}^{n_{\tau-1}}$ and $A_{\tau,\tau} \in \mathbb{R}^{m_\tau} \times \mathbb{R}^{n_\tau}$ (the objective function coefficients, the coefficients of the right-hand-sides and the constraint matrices), $\xi^\tau := (\xi_1, \ldots, \xi_\tau)$ and $\mathcal{F}_\tau$ is the σ-algebra generated by ξ^τ, $\tau = 1, \ldots, T$. $\mathcal{F}_\tau$ represents the available information in stage t. We assume that ξ_1 is known in stage 1[1]. With $\mathcal{F}_1 := \{\Omega, \emptyset\}$ and $\mathcal{F}_T := \mathcal{F}$, no further information has arrived in the first stage whereas all information is available in the last stage. Also, $x := (x_1, \ldots, x_T)$ is a vector of (stochastic) decisions and we assume $X_\tau \subseteq \mathbb{R}^{n_\tau}$ to be polyhedral subsets. Note that the constraints $A_{\tau-1,\tau} x_{\tau-1} + A_{\tau,\tau} x_\tau = b_\tau, x_\tau \in X_\tau$ mean $A_{\tau-1,\tau}(\omega) x_{\tau-1}(\omega) + A_{\tau,\tau}(\omega) x_\tau(\omega) = b_\tau(\omega), x_\tau(\omega) \in X_\tau$ for $\mathbb{P} - a.a.\ \omega \in \Omega$. The requirement that x_τ is $\mathcal{F}_\tau$ -measurable ensures that x_τ depends only on information available in stage τ, i.e., is *nonanticipative*. In almost all practical applications the distribution of ξ is discrete, inducing a so-called scenario tree. The assumption of a discrete distribution makes the MSRP a large-scale linear program.

Despite the development of solution methods for MSRPs (see, e.g., Birge (1985), Løkketangen and Woodruff (1996), Lulli and Sen (2004), Escudero et al. (2009), and Pantuso et al. (2015)), an accurate representation of the discrete distribution and/or the inclusion of all natural stages make many real-life problems computationally intractable. There exists a number of methods for scenario tree reduction, including clustering methods, such as Rockafellar and Wets (1991), Dupacová et al. (2003), Heitsch and Römisch (2003), Kovacevic and Pichler (2015) and Pflug and Pichler (2016). However, even for MSRPs with a very coarse discretization of uncertainty, the major challenge is often the number of stages. Many problems have an infinite planning horizon and, consequently, an infinite number of stages. For these problems it is necessary to rely on a significant length of the planning horizon, i.e., on a representative number of stages. Even for many finite-horizon problems the inclusion of all stages may likewise pose serious tractability issues. In any case, a typical workaround is to reduce the number of stages. Here, we do this by assuming a deterministic future from a certain stage until the end of the planning horizon (see also the discussion in Powell (2014)).

Multistage problems are often approximated by two-stage problems in which future uncertainty is fully disclosed in the second stage. Alternatively, the number of stages is chosen arbitrarily or determined by the largest problem one can solve. It is relevant for decision-makers to ask: *What is the value of solving the original problem compared to the approximation?* or, stated differently: *What are the costs of approximating the original problem by reducing the number of stages?* Yet, possibly more

[1] We define $A_{0,1}$ to be a matrix of zeros such that the constraints of stage 1 are $A_{1,1} x_1 = b_1$.

importantly: *What is the benefit of including an additional stage in the approximation?* The present paper aims to provide a quantitative answer to these questions.

The existing literature already suggests metrics for evaluating the benefit of solving the original problem. Given a two-stage stochastic program, the *Value of the Stochastic Solution* (VSS), gives the value of solving the stochastic program rather than its Expected Value problem (EV), obtained by replacing the random parameters of the second stage by their expected values, see Birge (1982). VSS is computed by fixing the first-stage variables of the original Recourse Problem (RP) to their values in the EV solution and comparing the objective function value to the optimal value of RP. Expected values are also used in the context of problems with multiple time horizons, e.g., in such a way that a tree branches at the weekly level and conditional expectations are used to approximate the uncertainty at the daily level (see, e.g., Werner et al. (2013) and Kaut et al. (2014) for stochastic programs with multiple horizons). Alternative strategies for evaluating the quality of an expected value solution are based on fixing variables at value zero or fixing variables according to their reduced costs in the EV solution, cf. Crainic et al. (2018). For minimization problems, the fixing of the EV solution produces an upper bound on the optimal value of RP, making VSS non-negative. For linear RPs with random right-hand-side only, the optimal value of the EV problem provides a lower bound on the RP. The EV bound, however, may not be very tight. The Wait-and-See (WS) problem obtained by relaxing non-anticipativity, gives a lower bound that is as least as good as the EV bound and applies for general minimization problems. For proofs, see e.g. Birge (1982). The VSS can also be computed for multistage programs, as shown by Escudero et al. (2007), who use a *rolling-horizon* version. As for two-stage problems, this version provides the benefit of solving a MSRP rather than using the EV solution in every stage. The computation, however, is slightly more complex than for two-stage problems. The EV is solved in a rolling-horizon framework to account for the fact that in a multistage problem decisions are made sequentially and conditional on the arrivals of new information. Other deterministic rolling-horizon approximations to a MSRP are discussed by Maggioni et al. (2014), e.g. using a reference scenario, and bounding procedures are suggested by Maggioni and Pflug (2016), Sandikçi and Özaltin (2017), and Maggioni and Pflug (2019). Maggioni et al. (2016) extend these to stochastic approximations using a subset of reference scenarios.

In this paper, we consider multistage approximations obtained by reducing the number of stages in the original problem – but without reverting to a deterministic problem – and we evaluate the benefits of using the solutions in a rolling-horizon procedure. With an infinite planning horizon, approximations would naturally have the same number of representative stages, even if implemented in a rolling-horizon framework. Given a finite-horizon problem and a corresponding data set, however, we aim to evaluate approximations using the same data set. Thus, in a rolling-horizon procedure, the number of stages of our approximations decreases in every iteration. We introduce the following metrics: The expected result of using the EV solution (EEV), the value of the EV problem and the values of two WS approximations. The EV problem can be seen as a stochastic extension of the deterministic approximation by Escudero et al. (2007) and the EEV can be viewed as an extension of the rolling-horizon value of the expected value scenario by Maggioni et al. (2014). We show that for linear minimization problems with random right-hand-side, the EV problem provides a lower bound that improves with the number of stages and thereby also improves the bound of Escudero et al. (2007). The same holds for the WS approximations, however, for general linear minimization problems and with one of the approximations improving the EV bound. In contrast, the upper bound provided by the EV solution may not improve with the number of stages. Our metrics are used to support the choice of the number of stages in a MSRP. More specifically, we extend the concept of VSS to account for stochastic approximations and we define the Marginal Stage Value (MSV) as the value of including an additional stage in an approximation. Furthermore, we demonstrate how to use the MSV in a simple heuristic. The computation of these metrics are illustrated for a case study of a hedging problem.

The contributions of this paper can be summarized as follows:

- We suggest the concepts of *the Expected Value (EV) problem, the Expected result of using the Expected Value solution (EEV), the Value of the Stochastic Solution (VSS)* and the values of two *Wait-and-See* (WS) approximations for MSRP, extending the deterministic approximations for two-stage problems to stochastic approximations for multistage problems obtained by reducing the number of stages,
- We define the *Marginal Stage Value (MSV)* to quantify the benefit of including an additional stage in an approximation to a MSRP and use it to develop a heuristic for choosing the number of stages,
- We prove the validity of bounds provided by the EV problem and the two WS approximations and assess their tightness as a function of the number of stages,
- We illustrate that, contrary to intuition, increasing the number of stages of the approximation does not necessarily foster better solutions.

The organization of the paper is the following. In Section 2 we introduce the value of the Expected Value problem and the Expected result of using the Expected Value solution and we illustrate the computation of the two metrics. In Section 3 we provide bounds for the linear MSRP, including two Wait-and-See bounds. In Section 4 we extend the Value of the Stochastic Solution and introduce the Marginal Stage Value. In Section 5 we apply the approach to a hedging problem and, finally, we draw conclusions in Section 6.

2 The Expected Value problem and the Expected result of using the Expected Value solution

The aim of this section is to extend the concepts of the Expected Value (EV) problem and the Expected result of using the Expected Value solution (EEV) from two-stage to multistage stochastic programming.

Consider the original T-stage problem (1). For $1 \leq T' \leq T$, the idea is to solve a number of T'-stage stochastic approximations obtained by replacing the random parameters of the remaining stages by their expected values and to evaluate the value of using the corresponding solutions in the original problem.

We start by defining the stage-t recourse problem

$$RP^t(x_1, \dots, x_{t-1}, \xi^{t-1}) = \min \Big\{ \mathbb{E}_{\xi^{t,T}|\xi^{t-1}} \Big[z^t(x_t, \dots, x_T, \xi^{t,T}) := \sum_{\tau=t}^{T} c_\tau x_\tau \Big] \Big| A_{\tau-1,\tau} x_{\tau-1} + A_{\tau,\tau} x_\tau = b_\tau,\ x_\tau \in X_\tau,\ x_\tau \text{ is } \mathcal{F}_\tau \text{ -measurable},\ \tau = t, \dots, T \Big\},$$

where $t, \dots, T$ are the stages and $\xi^{t,T} := (\xi_t, \dots, \xi_T)$. This is a subproblem of the original problem with $T - t + 1$ stages. The stage-1 problem is in fact the original problem. We denote an optimal solution to the stage-t problem by $x^{t,T} := x^{t,T}(x_1, \dots, x_{t-1}, \xi^{t-1})$, where $x^{t,T} = (x_t^{t,T}, \dots, x_T^{t,T})$. To simplify the analysis, we assume complete recourse such that we can ignore feasibility issues. Thus, the stage-t problem is feasible for any choice of $(x_1, \dots, x_{t-1})$.

We proceed to define the T'-stage approximation of the stage-t problem

$$EV^{t,T'}(x_1, \dots, x_{t-1}, \xi^{t-1}) = \min \Big\{ \mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}} \Big[z^t(x_t, \dots, x_T, \bar{\xi}^{t,t+T'-1}) := \sum_{\tau=t}^{T} c_\tau x_\tau \Big] \Big| \\ A_{\tau-1,\tau} x_{\tau-1} + A_{\tau,\tau} x_\tau = b_\tau,\ x_\tau \in X_\tau,\ \tau = t, \dots, T, \\ x_\tau \text{ is } \mathcal{F}_\tau \text{ -measurable},\ \tau = t, \dots, t+T'-1, \\ x_\tau \text{ is } \mathcal{F}_{t+T'-1} \text{ -measurable},\ \tau = t+T', \dots, T \Big\},$$

where $\bar{\xi}^{t,t+T'-1} := (\xi_t, \dots, \xi_{t+T'-1}, \mathbb{E}_{\xi_{t+T'}|\xi^{t,t+T'-1}}[\xi_{t+T'}], \dots, \mathbb{E}_{\xi_T|\xi^{t,t+T'-1}}[\xi_T])$. This problem has T' stages. Hence, the T'-stage approximation of the stage-t problem is obtained by replacing the stochastic vector $\xi^{t,T}$ by $\bar{\xi}^{t,t+T'-1}$, such that the components of $t+T', \dots, T$ become deterministic conditional on $\xi^{t,t+T'-1}$. The expectation with respect to $\xi^{t,T}$ is therefore replaced by the expectation with respect to $\xi^{t,t+T'-1}$. Likewise, the condition "x_τ is $\mathcal{F}_\tau$ -measurable" is replaced by "x_τ is $\mathcal{F}_{t+T'-1}$ -measurable" for $\tau \geq t+T'$. We denote an optimal solution to the T'-stage approximation of the stage-t problem by $\bar{x}^{t,t+T'-1} := \bar{x}^{t,t+T'-1}(x_1, \dots, x_{t-1}, \xi^{t-1})$ (indicating the stages at which the random variables are not replaced by their expectation), where $\bar{x}^{t,t+T'-1} = (\bar{x}_t^{t,t+T'-1}, \dots, \bar{x}_T^{t,t+T'-1})$. Note that $EV^{t,T'} = RP^t$ and $\bar{x}^{t,t+T'-1} = x^{t,T}$ for $t \geq T - T' + 1$, that is, the T'-stage approximation is the same as the stage-t problem when the number of stages is less than or equal to T'. In the following, we refer interchangeably to $EV^{t,T'}$ as the optimal value of the T'-stage approximation and the optimization problem itself.

In the two-stage case, $EV^{1,1}$ corresponds to the *Expected Value* (EV) problem, see Birge (1982). For multistage problems, $EV^{t,1}$ corresponds to the optimal value of the average scenario model by Escudero et al. (2007), and so, we extend their deterministic approximations to include stochastic problems. As it will be clear from the following section, we thereby improve the lower bounds provided by the approximations. Also, using the terminology of Powell (2014), an approximation in our context corresponds to a *stochastic lookahead model* but without the reduction of the planning horizon, and is similar to a *stage aggregation lookahead model*, although possibly with more than two stages.

In a multistage problem, decisions are made sequentially and conditional on the arrivals of new information. To account for this, we use the approximations in a rolling-horizon framework. In particular, upon reaching a given stage, new information arrives and conditional decisions are made on the basis of a new approximation. We store only the first component of the solution until the approximation is the same as the original problem and we store all components of the solution. More formally, we solve T'-stage approximations of the stage-t problems for $t = 1, \dots, T - T'$, recalling that the T'-stage approximations of the stage-t problems for $t = T - T' + 1, \dots, T$ are in fact the same as the stage-t problems. For $t = 1, \dots, T - T'$, the vector ξ_t becomes known and we solve the T'-stage approximations of the stage-t problems to obtain the optimal solutions $\bar{x}^{t,t+T'-1}$. We let $\bar{x}_t^{t,t+T'-1}$ be the first component and store only this component of the solution. For $t = T - T' + 1$, the vector $\xi_{T-T'+1}$ likewise becomes known and we solve the stage-$T - T' + 1$ problem to obtain the optimal solution $x^{T-T'+1,T}$. We store all components of this solution. The result is a solution $(\bar{x}_1^{1,T'}, \dots, \bar{x}_{T-T'}^{T-T',T-1}, x_{T-T'+1}^{T-T'+1,T}, \dots, x_T^{T-T'+1,T})$. The procedure is as follows:

Procedure 1 *Rolling-horizon T'-stage approximations*

Step 1: Solve $(EV^{1,T'})$ with $\bar{\xi}^{1,T'} = (\xi_1, \ldots, \xi_{T'}, \mathbb{E}_{\xi_{1+T'}|\xi^{1,T'}}[\xi_{1+T'}], \ldots, \mathbb{E}_{\xi_T|\xi^{1,T'}}[\xi_T])$, where ξ_1 is known, and let an optimal solution be $\bar{x}^{1,T'} := \bar{x}^{1,T'}(\xi^{1,1})$. Store $\bar{x}_1^{1,T'}$.

Step t: For all $(\bar{x}_1^{1,T'}, \ldots, \bar{x}_{t-1}^{t-1,t+T'-2}, \xi^{t-1})$, solve $(EV^{t,T'})$ with $\xi^{t,t+T'-1} = (\xi_t, \ldots, \xi_{t+T'-1}, \mathbb{E}_{\xi_{t+T'}|\xi^{t,t+T'-1}}[\xi_{t+T'}], \ldots, \mathbb{E}_{\xi_T|\xi^{t,t+T'-1}}[\xi_T])$, where ξ_t is known, and let an optimal solution be $\bar{x}^{t,t+T'-1} := \bar{x}^{t,t+T'-1}(\bar{x}_1^{1,T'}, \ldots, \bar{x}_{t-1}^{t-1,t+T'-2}, \xi^{1,t})$. Store $\bar{x}_t^{t,t+T'-1}$. Let $t := t+1$.

Step $T - T' + 1$: For all $(\bar{x}_1^{1,T'}, \ldots, \bar{x}_{T-T'}^{T-T',T-1}, \xi^{1,T-T'})$, solve $(RP^{T-T'+1})$ with $\xi^{T-T'+1,T} = (\xi_{T-T'+1}, \ldots, \xi_T)$, where $\xi_{T-T'+1}$ is known, and let an optimal solution be $x^{T-T'+1,T} := x^{T-T'+1,T}(\bar{x}_1^{1,T'}, \ldots, \bar{x}_{t-1}^{T-T',T'}, \xi^{1,T-T'+1})$. Store $x^{T-T'+1,T}$.

Example 1 Consider a two-stage approximation of the four-stage problem in Figure 1, for which the discrete stochastic process is represented by a scenario tree. Hence, $T = 4$ and $T' = 2$. To simplify notation, assume that $\xi_1, \ldots, \xi_4$ are mutually independent. Then, the procedure is:

Step 1: Solve the two-stage approximation $(EV^{1,2})$ rooted in the first-stage node and with $\bar{\xi}^{1,2} = (\xi_1, \xi_2, \mathbb{E}[\xi_3], \mathbb{E}[\xi_4])$. Let an optimal solution be $\bar{x}^{1,2} := \bar{x}^{1,2}(\xi_1)$.

Step 2: For ξ_1 and corresponding $\bar{x}_1^{1,2}$, solve $(EV^{2,2})$ rooted in a second-stage node and with $\bar{\xi}^{2,3} = (\xi_2, \xi_3, \mathbb{E}[\xi_4])$ and let an optimal solution be $\bar{x}^{2,3} := \bar{x}^{2,3}(\bar{x}_1^{1,2}, \xi_1, \xi_2)$.

Step 3: For all (ξ_1, ξ_2) and corresponding $(\bar{x}_1^{1,2}, \bar{x}_2^{2,2})$, solve (RP^3) rooted in a third-stage node with $\xi^{3,4} = (\xi_3, \xi_4)$ and let an optimal solution be $x^{3,4} := x^{3,4}(\bar{x}_1^{1,2}, \bar{x}_2^{2,3}, \xi_1, \xi_2, \xi_3)$.

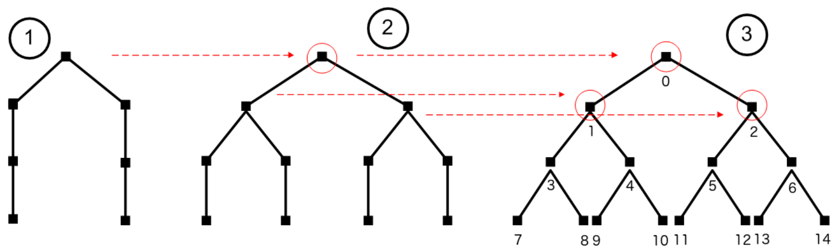

Figure 1: Rolling-horizon two-stage approximations of a four-stage problem. The numbers in circles are the steps of the rolling-horizon procedure. The four-stage scenario tree on the right illustrates the original unfolding of the random parameters and numbers beneath the node are indices.

It may be instructive to index the nodes of the scenario tree. In Figure 1, we index the nodes of right-most four-stage scenario tree by $0, \ldots, 14$. The rolling-horizon two-stage approximations work as follows. In Step 1, ξ_1 is known. We consider the first-stage node and form a two-stage scenario tree rooted in this node by replacing the random variables in stages 3 and 4 by their expected values. We solve the two-stage approximation and store its first-stage decision, $\bar{x}_0$. In Step 2, we fix $\bar{x}_0$ and both ξ_1 and ξ_2 are known. For each second-stage node n (i.e., each realization of the second-stage random variables) we form a two-stage scenario sub-tree rooted in node n by replacing the random variables in stage 4 by their expected values. These sub-trees are conditional on the information available in their respective second-stage nodes, including their history. We solve the two-stage approximations and store the decisions of the second stage, $\bar{x}_n, n = 1, 2$. Note that these decisions are, in general, different from those of the original four-stage problem. Finally, in Step 3, we fix $\bar{x}_n, n = 0, 1, 2$ and ξ_1, ξ_2 and ξ_3 are now known. For each third-stage node n, the scenario sub-tree rooted in node n is two-stage, and so, we solve the corresponding subproblem of the original problem and store the decisions $x_n, n = 3, \ldots, 14$. The result is a solution $(\bar{x}_0, \bar{x}_1, \bar{x}_2, x_3, \ldots, x_{14})$.

To evaluate the quality the T'-stage approximations, we consider the value of using the solution $(\bar{x}_1^{1,T'}, \ldots, \bar{x}_{T-T'}^{T-T',T-1}, x_{T-T'+1}^{T-T'+1,T}, \ldots, x_T^{T-T'+1,T})$ in the original problem. We therefore define

$$EEV^{t,T'}(x_1, \ldots, x_{t-1}, \xi^{t-1}) = \mathbb{E}_{\xi^{t,T}|\xi^{t-1}}\Big[z^t(\bar{x}_t^{t,T'+t-1}, \ldots, \bar{x}_{T-T'}^{T-T',T-1}, x_{T-T'+1}^{T-T'+1,T}, \ldots, x_T^{T-T'+1,T}, \xi^{t,T})\Big],$$

where $EEV^{t,T'} = RP^t$ for $t \geq T - T' + 1$. In the two-stage case, $EEV^{1,1}$ corresponds to the *Expected result of using the Expected Value solution* (EEV), see Birge (1982). For multistage problems, $EEV^{1,1}$ is equivalent to the Rolling-Horizon Value of the Reference Scenario (RHVRS) introduced by Maggioni et al. (2014), using the mean as the reference scenario. As such, we extend the rolling-horizon value from a deterministic reference problem to more advanced stochastic approximations. Notice also that, consistently with Powell (2014), approximate solutions are evaluated in the original objective function.

3 Bounds

This section is devoted to upper and lower bounds for the MSRP. As for two-stage stochastic programming, the multistage extensions of the EV problem and the EEV provide bounds on the optimal value of the original problem. Since the EV bound may not be very tight, we also suggest two WS approximations and assess their tightness.

We start with the Expected Value problem. For linear MSRP with random right-hand-sides we show that the optimal value of a T'-stage approximation, $EV^{1,T'}$, is a lower bound on that of the original problem, RP^1. The more stages in the approximation, the better this lower bound.

Proposition 1 Assume that the objective function coefficients and the constraint matrices are fixed such that ξ_τ collates the components of b_τ. Let $1 \leq t \leq T$. For all $(x_1, \ldots, x_{t-1}, \xi^{t-1})$,

$$EV^{t,T'}(x_1, \ldots, x_{t-1}, \xi^{t-1}) \leq EV^{t,T'+1}(x_1, \ldots, x_{t-1}, \xi^{t-1})$$

for $T' = 1, \ldots, T-1$ and, in particular, $EV^{t,T'}(x_1, \ldots, x_{t-1}, \xi^{t-1}) \leq EV^{t,T-t+1}(x_1, \ldots, x_{t-1}, \xi^{t-1}) = RP^t(x_1, \ldots, x_{t-1}, \xi^{t-1})$ for $T' \leq T - t + 1$.

Proof: For ease of exposition, we omit the constraints in the optimization problems below. For the optimal solution $\bar{x}^{t,t+T'}$ to the $T'+1$-approximation $EV^{t,T'+1}$ and for all $\xi^{t,t+T'}$ conditional on $\xi^{t,t+T'-1}$

$$\mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\big[z^t(\bar{x}_t^{t,t+T'}, \ldots, \bar{x}_T^{t,t+T'}, \bar{\xi}^{t,t+T'})\big] \geq \min_{x_t,\ldots,x_T} \mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\big[z^t(x_t, \ldots, x_T, \bar{\xi}^{t,t+T'})\big]$$

and, hence,

$$EV^{t,T'+1}(x_1,\ldots,x_{t-1},\xi^{t-1}) = \mathbb{E}_{\xi^{t,t+T'}|\xi^{t,t+T'-1}}\Big[\mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\big[z^t(\bar{x}_t^{t,t+T'},\ldots,\bar{x}_T^{t,t+T'},\bar{\xi}^{t,t+T'})\big]\Big]$$
$$\geq \mathbb{E}_{\xi^{t,t+T'}|\xi^{t,t+T'-1}}\Big[\min_{x_t,\ldots,x_T}\ \mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\big[z^t(x_t,\ldots,x_T,\bar{\xi}^{t,t+T'})\big]\Big].$$

The value function of the minimization problem is convex in its right-hand-side $\xi_{t+T'}$, and hence, by Jensen's inequality

$$\mathbb{E}_{\xi^{t,t+T'}|\xi^{t,t+T'-1}}\Big[\min_{x_t,\ldots,x_T}\ \mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\big[z^t(x_t,\ldots,x_T,\bar{\xi}^{t,t+T'})\big]\Big]$$
$$\geq \min_{x_t,\ldots,x_T}\ \mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\Big[z^t(x_t,\ldots,x_T,\mathbb{E}_{\xi^{t,t+T'}|\xi^{t,t+T'-1}}[\bar{\xi}^{t,t+T'}])\Big] = EV^{t,T'}(x_1,\ldots,x_{t-1},\xi^{t-1}).$$

□

Note that Proposition 1 holds for problems with random right-hand sides, as is also the case for two-stage problems, see Birge (1982). With randomess only in the right-hand side, the convexity of the second-stage value function allows for the application of Jensen's inequality.

We now define two other T'-stage approximations of the stage-t problem. The first is obtained by relaxing non-anticipativity in stages $t,\ldots,T-T'$:

$$\begin{aligned}\overline{WS}^{t,T'}(x_1,\ldots,x_{t-1},\xi^{t-1}) =&\mathbb{E}_{\xi^{t,T}|\xi^{T-T'+1,T}}\Big[\min\Big\{\mathbb{E}_{\xi^{T-T'+1,T}|\xi^{t-1}}\big[z^t(x_t,\ldots,x_T,\xi^{t,T}) := \sum_{\tau=1}^{T} c_\tau x_\tau\big]\Big|\\ &A_{\tau-1,\tau}x_{\tau-1}+A_{\tau,\tau}x_\tau = b_\tau,\ x_\tau\in X_\tau,,\ \tau=t,\ldots,T\\ &x_\tau \text{ is } \mathcal{F}_{T-T'+1}\text{ -measurable},\ \tau=t,\ldots,T-T',\\ &x_\tau \text{ is } \mathcal{F}_\tau\text{ -measurable},\ \tau=T-T'+1,\ldots,T\Big\}\Big].\end{aligned}$$

The second is obtained by relaxing non-anticipativity in stages $t+T'-1,\ldots,T$:

$$\begin{aligned}WS^{t,T'}(x_1,\ldots,x_{t-1},\xi^{t-1}) =&\min\Big\{\mathbb{E}_{\xi^{t,T}|\xi^{t-1}}\big[z^t(x_t,\ldots,x_T,\xi^{t,T}) := \sum_{\tau=1}^{T} c_\tau x_\tau\big]\Big|\\ &A_{\tau-1,\tau}x_{\tau-1}+A_{\tau,\tau}x_\tau = b_\tau,\ x_\tau\in X_\tau,,\ \tau=t,\ldots,T\\ &x_\tau \text{ is } \mathcal{F}_\tau\text{ -measurable},\ \tau=t,\ldots,t+T'-2,\\ &x_\tau \text{ is } \mathcal{F}_T\text{ -measurable},\ \tau=t+T'-1,\ldots,T\Big\}.\end{aligned}$$

We denote optimal solutions to the two T'-stage approximations by $\bar{x}^{T-T'+1,T}$ and $x^{t,t+T'-2}$, respectively (indicating the stages at which non-anticipativity is not relaxed). Note that $\overline{WS}^{t,T'} = WS^{t,T'} = RP^t$ for $t \geq T-T'+1$. In the two-stage case, $\overline{WS}^{1,1}$ and $WS^{1,1}$ correspond to the *Wait-and-See* (WS) problem, see Birge (1982).

Example 2 Consider the four-stage problem of Example 1. Figure 2 illustrates the three-stage approximations, $EV^{1,3}, \overline{WS}^{1,3}$ and $WS^{1,3}$.

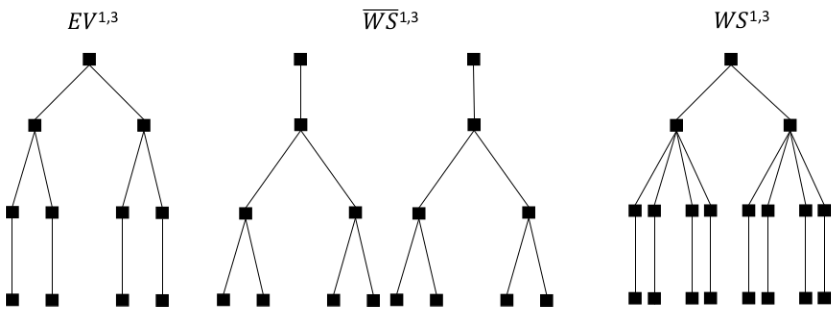

Figure 2: Three-stage approximations of the four-stage problem in Figure 1.

For the general MSRP, the optimal values of the T'-stage approximation, $\overline{WS}^{1,T'}$ and $WS^{1,T'}$, are likewise lower bounds on that of the original problem, RP^1. Also, the more stages in the approximations, the better these lower bounds.

Proposition 2 Let $1 \leq t \leq T$. For all $(x_1, \ldots, x_{t-1}, \xi^{t-1})$,

$$\overline{WS}^{t,T'}(x_1, \ldots, x_{t-1}, \xi^{t-1}) \leq \overline{WS}^{t,T'+1}(x_1, \ldots, x_{t-1}, \xi^{t-1})$$

for $T' = 1, \ldots, T-1$ and, in particular, $\overline{WS}^{t,T'}(x_1, \ldots, x_{t-1}, \xi^{t-1}) \leq \overline{WS}^{t,T-t+1}(x_1, \ldots, x_{t-1}, \xi^{t-1}) = RP^t(x_1, \ldots, x_{t-1}, \xi^{t-1})$ for $T' \leq T - t + 1$. The same holds with $\overline{WS}^{t,T'}$ replaced by $WS^{t,T'}$.

Proof: Observe that $\overline{WS}^{t,T'}$ and $WS^{t,T'}$ are obtained from $\overline{WS}^{t,T'+1}$ and $WS^{t,T'+1}$ by relaxing non-anticipativity in stages $T - T'$ and $t + T' - 1$, respectively.

The lower bound $EV^{t,T'}$ is readily available from Procedure 1, as opposed to $\overline{WS}^{t,T'}$ and $WS^{t,T'}$ that involve solving (stochastic) approximations. Whereas $\overline{WS}^{t,T'}$ is computed by solving a number of subproblems that are significantly smaller than RP^t, the problem $WS^{t,T'}$ has the same size as RP^t (although a more simple structure). However, the lower bound provided by $WS^{t,T'}$ is guaranteed to be at least as good as $EV^{t,T'}$.

Proposition 3 Assume that the objective function coefficients and the constraint matrices are fixed such that ξ_τ collates the components of b_τ. Let $1 \leq t \leq T$. For all $(x_1, \ldots, x_{t-1}, \xi^{t-1})$,

$$EV^{t,T'}(x_1, \ldots, x_{t-1}, \xi^{t-1}) \leq WS^{t,T'}(x_1, \ldots, x_{t-1}, \xi^{t-1})$$

for $T' = 1, \ldots, T-1$.

Proof: For ease of exposition, we omit the constraints in the optimization problems below. For the optimal solution $x^{t,t+T'-2}$ to the T'-stage approximation $WS^{t,T'}$ and for all $\xi^{t,t+T'}$ conditional on $\xi^{t,t+T'-1}$

$$\mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\left[z^t(x_t^{t,t+T'-2}, \ldots, x_T^{t,t+T'-2}, \xi^{t,T})\right] \geq \min_{x_t,\ldots,x_T} \mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\left[z^t(x_t, \ldots, x_T, \xi^{t,T})\right]$$

and, hence,

$$WS^{t,T'}(x_1,\dots,x_{t-1},\xi^{t-1}) = \mathbb{E}_{\xi^{t+T',T}|\xi^{t,t+T'-1}}\Big[\mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\big[z^t(x_t^{t,t+T'-2},\dots,x_T^{t,t+T'-2},\xi^{t,T})\big]\Big]$$
$$\geq \mathbb{E}_{\xi^{t+T',T}|\xi^{t,t+T'-1}}\Big[\min_{x_t,\dots,x_T}\ \mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\big[z^t(x_t,\dots,x_T,\xi^{t,T})\big]\Big],$$

where the first equality uses that $\mathbb{E}_{\xi^{t,T}|\xi^{t,t+T'-1}}[\cdot] = \mathbb{E}_{\xi^{t+T',T}|\xi^{t,t+T'-1}}[\cdot]$. The minimization problem is convex in its right-hand-side $\xi^{t,T}$, and hence, by Jensen's inequality

$$\mathbb{E}_{\xi^{t+T',T}|\xi^{t,t+T'-1}}\Big[\min_{x_t,\dots,x_T}\ \mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\big[z^t(x_t,\dots,x_T,\xi^{t,T})\big]\Big]$$
$$\geq \min_{x_t,\dots,x_T}\ \mathbb{E}_{\xi^{t,t+T'-1}|\xi^{t-1}}\Big[z^t(x_t,\dots,x_T,\mathbb{E}_{\xi^{t+T',T}|\xi^{t,t+T'-1}}[\xi^{t,T}])\Big] = EV^{t,T'}(x_1,\dots,x_{t-1},\xi^{t-1}),$$

where the last equality follows since $\mathbb{E}_{\xi^{t+T',T}|\xi^{t,t+T'-1}}[\cdot] = \mathbb{E}_{\bar{\xi}^{t+T',T}|\xi^{t,t+T'-1}}[\cdot]$ □

We proceed with the Expected value of using the Expected Value solution. For the general MSRP, the optimal solutions to the stage-T' approximations obtained by Procedure 1 provide an upper bound, $EEV^{1,T'}$, on the optimal value of the original problem, RP^1:

Proposition 4 Let $1 \leq t, T' \leq T$. For all $(x_1,\dots,x_{t-1},\xi^{t-1})$,

$$RP^t(x_1,\dots,x_{t-1},\xi^{1,t}) \leq EEV^{t,T'}(x_1,\dots,x_{t-1},\xi^{1,t}).$$

Proof: Observe that $EEV^{t,T'}$ and RP^t are obtained by solving the same MSRP, with the exception that the variables of stages $1,\dots,T-T'$ are fixed in the computation of $EEV^{t,T'}$.

Unfortunately, examples show that the upper bound does not necessarily improve with the number of stages, see Section 5. That is, increasing the number of stages does not necessarily produce better solutions. We do, however, obtain the following:

Corollary 1 Let $1 \leq t, T' \leq T$. For all $(x_1,\dots,x_{t-1},\xi^{t-1})$,

$$WS^{t,T'}(x_1,\dots,x_{t-1},\xi^{t-1}) \leq RP^t(x_1,\dots,x_{t-1},\xi^{t-1}) \leq \min_{\tau=1,\dots,T'} EEV^{t,\tau}(x_1,\dots,x_{t-1},\xi^{t-1}).$$

Proof: By Propositions (2) and (4), $WS^{t,T'} \leq RP^t \leq EEV^{t,\tau}$ for all $\tau = 1,\dots,T'$.

If the original T-stage problem is hard to solve or even intractable these bounds become useful. The computations of $EV^{1,T'}, \bar{W}S^{1,T'}, WS^{1,T'}$ and $EEV^{1,T'}$ involve the solution to problems with at most T' stages, which are hopefully significantly easier to solve than the original T-stage problem.

4 The Marginal Stage Value

The Expected result of using the Expected Value solution can be used to support the choice of the number of stages in an approximation. For this purpose, we introduce the following:

Definition 1. *The Value of the Stochastic Solution (VSS) to the stage-t problem compared to a T'-stage approximation is the quantity:*

$$VSS^{t,T'} = \min_{\tau=1,\dots,T'} EEV^{t,\tau}(x_1,\dots,x_{t-1},\xi^{t-1}) - RP^t$$

Definition 2. *The Marginal Stage Value (MSV) for a T'-stage approximation is the quantity:*

$$MSV^{t,T'} = VSS^{t,T'-1} - VSS^{t,T'} = \min_{\tau=1,\dots,T'-1} EEV^{t,\tau} - \min_{\tau=1,\dots,T'} EEV^{t,\tau}$$

It follows from Definition 1 and Definition 2 that $VSS^{t,T'} \geq 0$ and $MSV^{t,T'} \geq 0$. Moreover, $MSV^{t,T'} > 0$ if and only if $EEV^{t,T'} < EEV^{t,\tau}$ for $\tau = 1,\dots,T'-1$, i.e., the solution to the T'-stage approximation improves the objective function value of the original problem.

We propose the following heuristic for choosing the number of stages, using the MSV. Let $M \in \mathbb{R}_+$ and $m \in \mathbb{Z}_+$ be some pre-specified threshold and iteration counter, respectively. The procedure is then:

Procedure 2 *Heuristic for choosing the number of stages*

Step 1: Use the rolling-horizon τ-stage approximation and compute $EEV^{1,\tau}$ for $\tau = 1,\dots,m+1$. Let $MSV^{1,1} = 0$ and compute $MSV^{1,\tau}$ for $\tau = 2,\dots,m+1$.

Step T': If $MSV^{1,\tau} < M$ for $\tau = T',\dots,T'+m-1$, stop. Use the rolling-horizon $T'-1$-stage approximation. Otherwise, use the rolling-horizon $T'+m$-stage approximation, compute $EEV^{1,T'+m}$ and $MSV^{1,T'+m}$. Let $T' := T'+1$.

The idea is that we solve approximations with an increasing number of stages. In step T', we compute the rolling-horizon T'-stage approximation, compute $EEV^{1,T'}$ and update $MSV^{1,T'}$. If the approximation fails to improve the objective function value of the original problem by at least M during m subsequent iterations, we stop the heuristic.

5 Tests on a portfolio replication problem

In this section we illustrate the rolling-horizon solution of the EV problems, the solutions of the two WS approximations, and the computation of the EEV, VSS and MSV for a problem of portfolio replication, see e.g. Dempster and Thompson (2002). Rather than solving a problem of real-life size and detail, the purpose of our computational study is to illustrate how the approximations and bounds relate to the original problem and its optimal value. To do this, we use a problem formulation and instances that allow for optimal solutions within acceptable time.

An investor aims to track, i.e. replicate, the value of a target portfolio by allocating an initial budget B to a set of investment instruments $\mathcal{I}$ and re-allocate at discrete points in time $t = 1,\dots,T$ such as to minimize the average absolute tracking error over time. Such problem arises, for example, in the course of hedging in an incomplete market. The value of the target portfolio is random and represented by a one-dimentional stochastic process $\{\tilde{G}_t, t = 1,\dots,T\}$. The returns of the investment instruments are likewise random and follow a (possibly correlated) multi-dimentional stochastic process $\{\tilde{R}_{it}, i \in \mathcal{I}, t = 1,\dots,T\}$.

Table 1: Return characteristics of the investment assets.

Asset	Mean	StD	Skewness	Kurtosis	Risk Premium	Vol. Clumping	Mean Reversion
Cash	4.33	0.94	0.8	2.62	0.0	0.3	0.2
Bonds	5.91	0.82	0.49	2.39	0.0	0.3	0.2
D. Stocks	7.61	13.38	-0.75	2.93	0.3	0.3	0.0
F. Stocks	8.09	15.70	-0.74	2.97	0.3	0.3	0.0

We formulate the tracking problem as a linear T-stage stochastic recourse problem. Assuming discrete distributions, the realizations of the stochastic processes can be represented by a scenario tree on a set of nodes $\mathcal{N}$. Each node $n \in \mathcal{N}$ is characterized by a realization G_n of the target value and realizations $R_{in}, i \in \mathcal{I}$ of the asset returns. For $t = 1, \ldots, T$, we let $\mathcal{N}_t$ be the set of nodes at stage t. Moreover, 0 denotes the root node with probability $p_0 = 1$ and, for $n \in \mathcal{N} \backslash \{0\}$, $a(n)$ denotes the predecessor of node n and p_n denotes the probability, with $\sum_{n \in \mathcal{N}_t} p_n = 1, t = 1, \ldots, T$ and $\sum_{m:a(m)=n} p_m = p_n, n \in \mathcal{N} \backslash \mathcal{N}_T$. The decision variables x_{in} represent the amount allocated to asset i in node $n \in \mathcal{N}$ and y_{in}^+ and y_{in}^- represent excess and deficit of the target, respectively.

The linear MSRP is given by:

$$\min \frac{1}{T} \sum_{n \in \mathcal{N}} p_n (y_n^- + y_n^+) \tag{2a}$$

$$\text{s.t.} \sum_{i \in \mathcal{I}} x_{i0} \leq B, \tag{2b}$$

$$\sum_{i \in \mathcal{I}} x_{in} - \sum_{i \in \mathcal{I}} R_{in} x_{i,a(n)} = 0, \qquad n \in \mathcal{N}_t, t = 2, \ldots, T, \tag{2c}$$

$$\sum_{i \in \mathcal{I}} x_{in} - y_n^+ + y_n^- = G_n, \qquad n \in \mathcal{N}_t, t = 2, \ldots, T \tag{2d}$$

$$x_{in} \geq 0, \qquad i \in \mathcal{I}, n \in \mathcal{N}, \tag{2e}$$

$$y_{in}^+, y_{in}^- \geq 0, \qquad i \in \mathcal{I}, n \in \mathcal{N}. \tag{2f}$$

The objective function (2a) accumulates the average expected absolute tracking error over the planning horizon. Constraint (2b) ensures that no more than the initial budget is allocated to assets, constraints (2c) state that reallocation is possible at every stage, and (2d) that the investor is to track, by means of its portfolio, the value of the target portfolio. Constraints (2e) - (2f) ensure non-negativity of the decision variables.

We use the investment instruments described in Høyland and Wallace (2001). Four asset classes are considered, namely *Bonds*, *Cash*, *Domestic Stocks* and *Foreign Stocks*. The random returns of the assets are described by their first four marginal moments and their correlations as reported in Table 1 and Table 2, respectively. In addition to their moments, the asset returns are characterized by their risk premium, volatility clumping and mean reversion (see Høyland and Wallace (2001) for details), also reported in Table 1. We generate returns by means of the moment-matching heuristic introduced by Høyland et al. (2003). The idea behind the moment-matching heuristic is as follows: First, a number of discrete univariate random variables are constructed, each having a given specification of the first four moments. Then, the random variables are transformed to a random vector having a given correlation matrix. However, since this transformation distorts the third and fourth marginal

Table 2: Correlations between the returns of the investment assets.

	Bonds	D. Stocks	Cash	F. Stocks
Cash	1.00	0.60	-0.20	-0.10
Bonds		1.00	-0.30	-0.20
D. Stocks			1.00	0.60
F. Stocks				1.00

moments, the heuristic starts out with specifications of these two moments different from their targets, and such that final marginals have the correct moments. For further details see Høyland et al. (2003).

For each realization at a given stage, the heuristic is used to generate conditional distributions for the following stage, characterized by their four marginal moments and correlations, and taking into account that mean and standard deviation are stage-dependent due to volatility clumping and mean reversion. Furthermore, we assume an initial budget corresponding to the initial value of the target portfolio, i.e. $B = G_0 = 55$ thousand dollars. For every stage $t = 2, \ldots, T$, realized values of the target portfolio, G_n, $n \in \mathcal{N}_t$, are sampled from a triangular distribution, $T[G_0-\kappa, G_0, G_0+\kappa]$, where $G_0 = 55$ thousand dollars represents the mode and $G_0 - \kappa$ and $G_0 + \kappa$ the lower and upper limits, respectively. We test the values of $\kappa = 10$ and 15 when the target mode and limits are constant and $\kappa = 10$ when the target grows at the same rate as the average asset return, that is 6.485%. The instances are also made available online[2].

We begin by computing the EV, WS, EEV, VSS and MSV for the special case of the tracking problem with fixed asset returns. This setting allows us to obtain a version of (2) with random right-hand-side only, and thus, to illustrate the validity of the EV bound. We let $R_{in} = M_i$ for each $i \in \mathcal{I}$ and $n \in \mathcal{N}$, where M_i is the mean return of asset i reported in Table 1. For this case, the scenario tree is obtained by branching five times per stage, with each branch holding a target value sampled from the underlying distribution. Consequently, the resulting scenario trees count 5^{T-1} scenarios, for $T = 4, 5, 6$. Table 3 shows the results for the case with fixed asset returns and $\kappa = 10$, Table 5 with $\kappa = 10$ and a target growth of 6.485% per stage, and Table 4 shows the results with $\kappa = 15$.

Table 3: Results for the tracking problem with fixed asset returns and $\tilde{G}_t \approx T[45, 55, 65]$, $t = 2, \ldots, T$, varying the number of stages T in the original problem and T' in the approximation. $VSS^{1,T'}[\%]$ represents the percentage increase in objective value generated by the approximation and is computed as $100 \times VSS^{1,T'}/RP^1$. $MSV^{1,T'}[\%]$ represents the percentage increase in the value of the stochastic solution by including stage T' and is computed as $100 \times MSV^{1,T'}/VSS^{1,T'-1}$. The remaining values are expressed in thousands of dollars.

T	T'	$EV^{1,T'}$	$WS^{1,T'}$	$\overline{WS}^{1,T'}$	RP^1	$EEV^{1,T'}$	$VSS^{1,T'}$	$VSS^{1,T'}[\%]$	$MSV^{1,T'}$	$MSV^{1,T'}[\%]$
4	1	9.00	13.41	11.66	13.72	14.03	0.31	2.25	0.00	0.00
	2	11.82	13.52	11.95		14.16	0.31	2.25	0.00	0.00
	3	13.26	13.72	12.34		13.72	0.00	0.00	0.31	100.00
	4	13.72	13.72	13.72		13.72	0.00	0.00	0.00	0.00
5	1	14.10	18.97	17.15	19.25	19.34	0.09	0.45	0.00	0.00
	2	15.87	19.13	17.39		20.75	0.09	0.45	0.00	0.00
	3	18.08	19.15	18.06		19.80	0.09	0.45	0.00	0.00
	4	18.98	19.25	18.75		19.27	0.01	0.06	0.08	86.82

[2] The instances are available at the following address https://sid.erda.dk/sharelink/ecTlLZXJDP.

	5	19.25	19.25	19.25		19.25	0.00	0.00	0.01	100.00
6	1	20.78	26.07	24.31	26.27	26.90	0.63	2.39	0.00	0.00
	2	21.05	26.15	24.53		26.65	0.38	1.44	0.25	39.73
	3	23.66	26.22	24.84		26.70	0.38	1.44	0.00	0.00
	4	25.54	26.22	25.32		26.37	0.10	0.38	0.28	73.28
	5	26.16	26.27	25.90		26.29	0.02	0.06	0.09	84.42
	6	26.27	26.27	26.27		26.27	0.00	0.00	0.02	100.00

We observe from Table 3 that the lower bound provided by the expected value problem $EV^{1,T'}$ improves with the number of stages T' as proven in Proposition 1 and that the improvements are significant. However, the bounds provided by including only the first stages are not very tight. The two-stage approximations, for example, result in gaps of 17-24% between the upper and lower bounds, whereas the three-stage approximations have gaps of 3-11%. To obtain gaps below 5% for this specific problem, one should include 70-80% of the stages in the original problem.

The lower bounds provided by the wait-and-see approximations $WS^{1,T'}, \overline{WS}^{1,T'}$ likewise improve with the number of stages T' as proven in Proposition 2. We also confirm that the bound $WS^{1,T'}$ is always at least as good as $EV^{1,T'}$. In fact, when using the bound $WS^{1,T'}$, $T' = 1$ results in gaps of 1.9-4.4% between the upper and lower bounds, whereas $T' = 2$ produces gaps of 1.1-7.8% (the reason that the gap does not necessarily decrease with the number of stages is that the upper bound does not always decrease, see below). Moreover, in this specific application, the last few stages are more or less redundant, with gaps around 0.1% for $T' = T - 1$ and $T = 5$ or 6 and a gap of 0.6% for $T' = T - 2$ and $T = 6$.

While the size of the $WS^{1,T'}$ approximation is comparable to that of the original problem [3], $\overline{WS}^{1,T'}$ consists of a number subproblems that are typically significantly smaller. As an example, see Figure 2. The $WS^{1,3}$ approximation includes the same number of leaf nodes/scenarios as the four-stage problem in Figure 1 (though the non-anticipativity constraints in the third stage are relaxed), whereas $\overline{WS}^{1,3}$ consists of two stochastic but smaller subproblems. The tightness of this bound is slightly better than $EV^{1,T'}$ when including only the first stages in the approximations and otherwise comparable, with the gaps of the two-stage and three-stage approximations being 8-16% and 7-10%, respectively.

In contrast to the lower bounds, the upper bounds do not always improve with the number of stages. In Table 3, we notice that $EEV^{1,2} > EEV^{1,1}$ for $T = 4$ and 5, meaning that there is a loss from including the second stage in the deterministic problem and implying that $MSV^{1,2} = 0$. For $T = 5$, we likewise observe that $MSV^{1,3} = 0$, and thus, the deterministic problem is also at least as good as the three-stage approximation. Even though the quality of the solutions generally improves with the number of stages, the progression is often erratic. For $T = 6$, one gains from including the second and fourth stage, but not the third. This erratic behavior likewise applies to problems with longer planning horizons, as seen in Table 9 in Appendix A. For the $T = 15$, the solution improves with the inclusion of the second, fourth, sixth and then ninth and tenth stage, whereas nothing is gained from the addition of the fifth stage to a four-stage approximation. Thus, contrary to intuition, including a few more stages in the approximation does not necessarily foster better solutions.

In spite of an erratic progression, we observe that $MSV^{1,1} = MSV^{1,2} = 0$ whereas $EEV^{1,2} > EEV^{1,3}$ and thus $MSV^{1,3} > 0$ for $T = 4$, $MSV^{1,1} = MSV^{1,2} = MSV^{1,3} = 0$ whereas $EEV^{1,3} > EEV^{1,4}$ and thus $MSV^{1,4} > 0$ for $T = 5$, and $MSV^{1,1}, MSV^{1,2}, MSV^{1,4}, MSV^{1,5} > 0$ for $T = 6$. We conclude that for this specific problem, one should either exclude the last few stages of the original

[3] Assuming a scenario formulation is used instead of the node formulation in (2), the difference is the number of non-anticipativity constraints.

problem, while facing an accompanying computational effort, or include only the first few stages, accepting a lower quality of the resulting solution. In particular, one should either use a deterministic approximation for $T = 4$ and 5 and a two-stage approximation for $T = 6$ (use the heuristic with $M = 0.01$ and $m = 1$) or approximations with $T' - 1$ stages for all $T = 4, 5$ and 6 (use the heuristic with $M = 0.05$ and $m = 3$).

Table 4: Results for the tracking problem with fixed asset returns and $\tilde{G}_t \approx T[40, 55, 70]$, $t = 2, \ldots, T$, varying the number of stages T in the original problem and T' in the approximation. $VSS^{1,T'}[\%]$ represents the percentage increase in objective value generated by the approximation and is computed as $100 \times VSS^{1,T'}/RP^1$. $MSV^{1,T'}[\%]$ represents the percentage increase in the value of the stochastic solution by including stage T' and is computed as $100 \times MSV^{1,T'}/VSS^{1,T'-1}$. The remaining values are expressed in thousands of dollars.

T	T'	$EV^{1,T'}$	$WS^{1,T'}$	$\overline{WS}^{1,T'}$	RP^1	$EEV^{1,T'}$	$VSS^{1,T'}$	$VSS^{1,T'}[\%]$	$MSV^{1,T'}$	$MSV^{1,T'}[\%]$
4	1	8.94	17.19	14.44	17.70	18.11	0.40	2.28	0.00	0.00
	2	13.98	17.45	14.96		19.10	0.40	2.28	0.00	0.00
	3	16.53	17.70	15.76		18.06	0.36	2.01	0.05	11.93
	4	17.70	17.70	17.70		17.70	0.00	0.00	0.36	100.00
5	1	14.15	23.50	20.75	24.16	24.18	0.02	0.06	0.00	0.00
	2	17.30	23.82	21.27		25.90	0.02	0.06	0.00	0.00
	3	21.09	23.92	22.31		24.82	0.02	0.06	0.00	0.00
	4	23.31	24.16	22.89		24.16	0.00	0.00	0.02	100.00
	5	24.16	24.16	24.16		24.16	0.00	0.00	0.00	0.00
6	1	20.88	31.02	28.59	31.61	32.79	1.18	3.74	0.00	0.00
	2	22.22	31.18	29.05		33.01	1.18	3.74	0.00	0.00
	3	26.26	31.37	29.53		33.61	1.18	3.74	0.00	0.00
	4	29.74	31.42	30.35		31.80	0.19	0.61	0.99	83.73
	5	30.97	31.61	30.90		31.62	0.01	0.04	0.18	93.33
	6	31.61	31.61	31.61		31.61	0.00	0.00	0.01	100.00

We proceed by analyzing the impact of more uncertainty. When the variance of the target increases, tracking the target becomes more difficult, and the tracking error becomes larger, as evidenced by a higher RP^1 in Table 4 than in Table 3. In many cases, this raises the $VSS^{1,T'}$ and the marginal stage value for T' with $MSV^{1,T'} > 0$. However, it may not be the same T' for which $MSV^{1,T'} > 0$ in Table 4 as in Table 3. In particular, no improvement is obtained by including the second stage for any T. Furthermore, for $T = 4$, the inclusion of the fourth stage provides a significant improvement to the quality of the solution. Thus, the more uncertainty in the form of an increase in variance, the more stages may have to be included in an approximation.

Table 5: Results for the tracking problem with fixed asset returns and $\tilde{G}_t \approx T[45(1+\xi)^{t-1}, 55(1+\xi)^{t-1}, 65(1+\xi)^{t-1}]$, $t = 2, \ldots, T$, and $\xi = 6.485\%$, varying the number of stages T in the original problem and T' in the approximation. $VSS^{1,T'}[\%]$ represents the percentage increase in objective value generated by the approximation and is computed as $100 \times VSS^{1,T'}/RP^1$. $MSV^{1,T'}[\%]$ represents the percentage increase in the value of the stochastic solution by including stage T' and is computed as $100 \times MSV^{1,T'}/VSS^{1,T'-1}$. The remaining values are expressed in thousands of dollars.

T	T'	$EV^{1,T'}$	$WS^{1,T'}$	$\overline{WS}^{1,T'}$	RP^1	$EEV^{1,T'}$	$VSS^{1,T'}$	$VSS^{1,T'}[\%]$	$MSV^{1,T'}$	$MSV^{1,T'}[\%]$
4	1	0.00	9.19	7.95	10.63	10.82	0.20	1.86	0.00	0.00

	2	4.80	9.84	9.00		10.66	0.04	0.33	0.16	82.01
	3	7.88	10.63	9.65		10.67	0.04	0.33	0.00	0.00
	4	10.63	10.63	10.63		10.63	0.00	0.00	0.04	100.00
5	1	0.00	11.78	10.70	14.34	14.62	0.28	1.96	0.00	0.00
	2	4.81	12.73	11.80		14.43	0.09	0.63	0.19	67.91
	3	8.13	13.55	12.86		14.38	0.04	0.30	0.05	52.94
	4	11.58	14.34	13.40		14.35	0.00	0.02	0.04	92.71
	5	14.34	14.34	14.34		14.34	0.00	0.00	0.00	0.00
6	1	0.00	14.68	13.66	18.26	18.67	0.41	2.24	0.00	0.00
	2	4.81	15.52	14.85		18.45	0.19	1.04	0.22	53.42
	3	8.15	16.60	15.86		18.33	0.07	0.36	0.12	65.16
	4	11.87	17.40	16.84		18.26	0.00	0.00	0.07	100.00
	5	15.25	18.26	17.33		18.26	0.00	0.00	0.00	0.00
	6	18.26	18.26	18.26		18.26	0.00	0.00	0.00	0.00

In Table 5, we investigate growth in the random target, assuming the initial target distribution remains unchanged but its parameters grow over time at the same rate as the average asset return. Thus, the mean and standard deviation grow at this rate. When comparing to the results in Table 3, we would expect the RP to decrease as a result of the same growth in asset returns and the mean target, decreasing the tracking error. This may also reduce the VSS. At the same time, we would expect the VSS to increase due to a larger variance of the target and thereby a larger tracking error. For $T = 4$ and 6, the result is smaller $VSS^{1,T'}$ and $MSV^{1,T'} > 0$ for most T', whereas for $T = 5$, $VSS^{1,T'}$ and $MSV^{1,T'} > 0$ are mostly larger.

As in Table 3, the quality of the solutions generally increases with the number of stages. In Table 5, however, the improvements are less erratic. The same behavior is confirmed for longer planning horizons, as shown by the results in Table 10 in Appendix A.

Comparing Table 5 with Table 3 we observe that one should include one more stage in the approximation, i.e. a two-stage approximation for $T = 4$ and 5 and a three-stage approximation for $T = 6$. It could likewise make sense to include two more stages. However, there are no gains in including many more.

We finally consider the general case of the tracking problem with random asset returns, and hence, obtain a version of (2) with stochastic right-hand-side and recourse matrix. In this case, target values are again sampled from the underlying triangular distribution, while asset returns are generated using the moment-matching heuristic by Høyland et al. (2003). We assume that target values and asset returns are uncorrelated. The scenario tree is obtained by branching 10 times per stage, i.e. by generating 10 asset and target realizations of the conditional distribution in a node. We obtain scenario trees counting 10^{T-1} scenarios, for $T = 4, 5, 6$. Table 6 shows the results for $\kappa = 10$ and Table 7 with $\kappa = 10$ and a target growth of 6.485% per stage.

Table 6: Results for the tracking problem with random asset returns and $\tilde{G}_t \approx T[45, 55, 65]$, $t = 2, \ldots, T$, varying the number of stages T in the original problem and T' in the approximation. $VSS^{1,T'}[\%]$ represents the percentage increase in objective value generated by the approximation and is computed as $100 \times VSS^{1,T'}/RP^1$. $MSV^{1,T'}[\%]$ represents the percentage increase in the value of the stochastic solution by including stage T' and is computed as $100 \times MSV^{1,T'}/VSS^{1,T'-1}$. The remaining values are expressed in thousands of dollars.

T	T'	RP^1	$EEV^{1,T'}$	$VSS^{1,T'}$	$VSS^{1,T'}[\%]$	$MSV^{1,T'}$	$MSV^{1,T'}[\%]$
	1		18.76	3.22	20.74	0.00	0.00
4		15.54					

	2		15.68	0.14	0.91	3.08	95.64
	3		15.55	0.01	0.09	0.13	90.47
	4		15.54	0.00	0.00	0.01	100.00
5	1	22.26	25.19	2.93	13.15	0.00	0.00
	2		22.90	0.64	2.85	2.29	78.30
	3		22.36	0.10	0.45	0.54	84.38
	4		22.28	0.02	0.08	0.08	83.00
	5		22.26	0.00	0.00	0.02	100.00
6	1	29.69	34.36	4.67	15.74	0.00	0.00
	2		30.13	0.44	1.50	4.23	90.49
	3		29.91	0.23	0.76	0.22	48.93
	4		29.83	0.14	0.47	0.09	38.53
	5		29.72	0.03	0.12	0.11	75.47
	6		29.69	0.00	0.00	0.03	100.00

Table 7: Results for the tracking problem with random asset returns and $\tilde{G}_t \approx T[45(1+\xi)^{t-1}, 55(1+\xi)^{t-1}, 65(1+\xi)^{t-1}]$, $t = 2, \ldots, T$, and $\xi = 6.485\%$, varying the number of stages T in the original problem and T' in the approximation. $VSS^{1,T'}[\%]$ represents the percentage increase in objective value generated by the approximation and is computed as $100 \times VSS^{1,T'}/RP^1$. $MSV^{1,T'}[\%]$ represents the percentage increase in the value of the stochastic solution by including stage T' and is computed as $100 \times MSV^{1,T'}/VSS^{1,T'-1}$. The remaining values are expressed in thousands of dollars.

T	T'	RP^1	$EEV^{1,T'}$	$VSS^{1,T'}$	$VSS^{1,T'}[\%]$	$MSV^{1,T'}$	$MSV^{1,T'}[\%]$
4	1	12.16	14.93	2.76	22.72	0.00	0.00
	2		12.28	0.12	0.99	2.64	95.64
	3		12.18	0.02	0.14	0.10	85.42
	4		12.16	0.00	0.00	0.02	100.00
5	1	16.10	20.14	4.04	25.08	0.00	0.00
	2		16.29	0.18	1.14	3.86	95.47
	3		16.16	0.06	0.38	0.12	66.67
	4		16.11	0.01	0.05	0.05	86.95
	5		16.10	0.00	0.00	0.01	100.00
6	1	20.50	25.79	5.28	25.76	0.00	0.00
	2		20.81	0.31	1.49	4.98	94.20
	3		20.64	0.13	0.64	0.17	56.91
	4		20.62	0.11	0.55	0.02	14.41
	5		20.51	0.00	0.02	0.11	95.94
	6		20.50	0.00	0.00	0.00	0.00

Results for the case with random asset returns are shown in Table 6 and Table 7. When comparing to the cases with fixed asset returns, we observe a higher marginal stage value in the second stage. This suggests that deterministic approximations do not perform very well (note also the high $VSS^{1,1}$ compared to the previous cases), whereas the solution to a two-stage approximation already significantly reduces the VSS (for approximations with two or more stages $VSS^{1,T'}$ is always below 3%). Using the heuristic with $M = 1$ and $m = 1$, one should indeed solve a two-stage approximation. Finally, contrary to the previous cases, the quality of the solutions increases consistently with the number of stages. In fact, the $MSV^{1,T'} > 0$ for almost all T', implying that including an additional stage in the approximation always provides a marginal improvement to the solution. With fixed targets, the uncertainty is only in the future and its impact is less, the more distant this future is. With random returns, however, uncertainty is present and has an impact throughout the whole planning horizon, which can explain why increasing the number of stages leads to more stable improvements.

Finally, we stress that our computations of VSS and MSV aim at good here-and-now decisions. Note that only the first-stage solution of an MSRP is implementable. In the majority of the practical applications, the remaining stages of the scenario tree serve to approximate the uncertainty and the remaining solutions of the problem help to make sensible first-stage decisions. In general, the appropriate number of stages in an approximation to a MSRP should be driven by their impact on first-stage decision-making. Consistent with this, we only implement the first-stage components of the solutions obtained during our rolling-horizon procedure. As an example, consider the four-stage problem in Table 7. The optimal value of the recourse problem is 12.16 thousand dollars, whereas the result of using the solution to the $EV^{1,1}$ problem is 14.93 thousand dollar (recall that the $EV^{1,1}$ is the deterministic problem obtained by replacing the scenario tree by its mean). The corresponding first-stage solutions are shown in Table 8. It can be observed that the RP^1 allocates the budget to bonds and domestic stocks, with a small holding of international stocks. The deterministic problem suggests a heavy investment in domestic stocks, with a minor holding of cash reserves. The two solutions differ in their tracking error but not least in their allocation strategy. On the other hand, in Table 7 we also notice that the VSS is significantly reduced by using the solution to the $EV^{1,2}$ problem (that is, with the two-stage approximation of the four-stage problem). In particular, the RP^1 is 12.16 and $EEV^{1,2}$ is 12.28. The improvement of the EEV can be explained by the first-stage solution of $EV^{1,2}$ in Table 8. The allocation strategy of the two-stage approximation is consistent with that of RP^1: both problems now allocate the budget to bonds and domestic stocks, though the amounts allocated are still different.

Table 8: First stage solution to the RP^1, $EV^{1,1}$ and $EV^{1,2}$ problems for the four-stage case in Table 7.

Problem	Bonds	Cash	D. Stocks	I. Stocks
RP^1	32.31	0.00	17.76	4.92
$EV^{1,1}$	0.00	7.56	47.43	0.00
$EV^{1,2}$	43.81	0.00	11.18	0.00

6 Conclusions

This paper extends the following metrics from two-stage stochastic recourse problems to multistage problems: The Value of the Expected Value problem (EV), the values of two Wait-and-See approximations and the Expected result of using the Expected Value solution (EEV). We show that these quantities provide upper and lower bounds for the optimal value of MSRPs. Moreover, we illustrate how these bounds can be used to support the choice of the number of stages in a multistage stochastic programming problem, with the aim to make good here-and-now decisions. In particular, we generalize the Value of the Stochastic Solution (VSS) to account for stochastic approximations, use it to define the Marginal Stage Value (MSV) and demonstrate how to utilize this in a heuristic.

We illustrate the approach for a portfolio replication problem. Our case study shows that the lower bounds significantly increase with the number of stages, and that the solutions to stochastic approximations in general improve as the number of stages increases. However, the improvements are sometimes erratic. In fact, we illustrate that increasing the number of stages may produce worse solutions. For some problem instances, a few stages are sufficient to obtain solutions of good quality. For other instances, however, the number of stages can only be reduced by one or two without compromising quality. In spite of this, even a small reduction in the number of stages may save substantial computational effort for a MSRP (or simply facilitate a solution to a problem that would otherwise be intractable).

References

S. Ahmed, A. King, and G. Parija. A multi-stage stochastic integer programming approach for capacity expansion under uncertainty. *Journal of Global Optimization*, 26(1):3–24, 2003.

A. Alonso, L. F. Escudero, and M. T. Ortuño. A stochastic 0-1 program based approach for the air traffic flow management problem. *European Journal of Operational Research*, 120(1):47 – 62, 2000.

R. Bakkehaug, E. S. Eidem, K. Fagerholt, and L. M. Hvattum. A stochastic programming formulation for strategic fleet renewal in shipping. *Transportation Research Part E: Logistics and Transportation Review*, 72:60 – 76, 2014.

J. R. Birge. The value of the stochastic solution in stochastic linear programs with fixed recourse. *Mathematical Programming*, 24(1):314–325, 1982.

J. R. Birge. Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research*, 33(5):989–1007, 1985.

J. R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer, New York, 1997.

D. R. Cariño, T. Kent, D. H. Myers, C. Stacy, M. Sylvanus, A. L. Turner, K. Watanabe, and W. T. Ziemba. The Russell-Yasuda Kasai Model: An Asset/Liability Model for a Japanese Insurance Company Using Multistage Stochastic Programming. *Interfaces*, 24(1):29–49, 1994.

G. Crainic, F. Maggioni, G. Perboli, and W. Rei. Reduced cost-based variable fixing in two-stage stochastic programming. *Annals of Operations Research*, pages 1–37, 2018.

M. Dempster and G. Thompson. *Dynamic portfolio replication using stochastic programming*. Cambridge University Press, 2002.

J. Dupacová, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming. *Mathematical Programming*, 95(3):493–511, 2003.

L. Escudero, A. Garín, M. Merino, and G. Pérez. The value of the stochastic solution in multistage problems. *TOP*, 15(1):48–64, 2007.

L. F. Escudero, A. Garín, M. Merino, and G. Pérez. BFC-MSMIP: an exact branch-and-fix coordination approach for solving multistage stochastic mixed 0–1 problems. *TOP*, 17(1):96–122, 2009.

S.-E. Fleten and T. K. Kristoffersen. Short-term hydropower production planning by stochastic programming. *Computers & Operations Research*, 35(8):2656 – 2671, 2008.

H. Heitsch and W. Römisch. Scenario reduction algorithms in stochastic programming. *Computational optimization and applications*, 24(2-3):187–206, 2003.

K. Høyland and S. W. Wallace. Generating scenario trees for multistage decision problems. *Management Science*, 47(2):295–307, 2001.

K. Høyland, M. Kaut, and S. W. Wallace. A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24(2–3):169–185, 2003.

P. Kall and S. W. Wallace. *Stochastic programming*. John Wiley and Sons Ltd, Chichester, 1994.

M. Kaut, K. T. Midthun, A. S. Werner, A. Tomasgard, L. Hellemo, and M. Fodstad. Multi-horizon stochastic programming. *Computational Management Science*, 11(1-2):179–193, 2014.

A. J. King and S. W. Wallace. *Modeling with stochastic programming*. Springer New York, New York, 2012.

R. M. Kovacevic and A. Pichler. Tree approximation for discrete time stochastic processes: a process distance approach. *Annals of operations research*, 235(1):395–421, 2015.

T. K. Kristoffersen and S.-E. Fleten. Stochastic programming models for short-term power generation scheduling and bidding. In E. Bjørndal, M. Bjørndal, P. M. Pardalos, and M. Rönnqvist, editors, *Energy, Natural Resources and Environmental Economics*, Energy Systems, pages 187–200. Springer Berlin Heidelberg, 2010.

Y. Li, G. Huang, and S. Nie. An interval-parameter multi-stage stochastic programming model for water resources management under uncertainty. *Advances in Water Resources*, 29(5):776 – 789, 2006.

A. Løkketangen and D. L. Woodruff. Progressive hedging and tabu search applied to mixed integer (0, 1) multistage stochastic programming. *Journal of Heuristics*, 2(2):111–128, 1996.

G. Lulli and S. Sen. A branch-and-price algorithm for multistage stochastic integer programming with application to stochastic batch-sizing problems. *Management Science*, 50(6):786–796, 2004.

F. Maggioni and G. C. Pflug. Bounds and Approximations for Multistage Stochastic Programs. *SIAM Journal on Optimization*, 26(1):831–855, 2016.

F. Maggioni and G. C. Pflug. Guaranteed bounds for general nondiscrete multistage risk-averse stochastic optimization programs. *SIAM Journal on Optimization*, 29(1):454–483, 2019.

F. Maggioni, E. Allevi, and M. Bertocchi. Bounds in multistage linear stochastic programming. *Journal of Optimization Theory and Applications*, 163(1):200–229, 2014.

F. Maggioni, E. Allevi, and M. Bertocchi. Monotonic bounds in multistage mixed-integer linear stochastic programming. *Computational Management Science*, 13(3):423–457, 2016.

J. M. Mulvey and B. Shetty. Financial planning via multi-stage stochastic optimization. *Computers & Operations Research*, 31(1):1 – 20, 2004.

G. Pantuso. The football team composition problem: a stochastic programming approach. *Journal of Quantitative Analysis in Sports*, 13(3):113–129, 2017.

G. Pantuso, K. Fagerholt, and S. W. Wallace. Solving hierarchical stochastic programs: application to the maritime fleet renewal problem. *INFORMS Journal of Computing*, 27(1):89–102, 2015.

G. Pflug and A. Pichler. *Multistage stochastic optimization*. Springer series in Operations Research and Financial Engineering. Springer, 2014.

G. C. Pflug and A. Pichler. From empirical observations to tree models for stochastic optimization: Convergence properties. *SIAM Journal on Optimization*, 26(3):1715–1740, 2016.

W. B. Powell. Clearing the jungle of stochastic optimization. In A. Newman and J. Leung, editors, *Bridging Data and Decisions*, pages 109–137. Informs, 2014.

R. T. Rockafellar and R. J.-B. Wets. Scenarios and Policy Aggregation in Optimization under Uncertainty. *Mathematics of Operations Research*, 16(1):119–147, 1991.

B. Sandikçi and O. Y. Özaltin. A scalable bounding method for multistage stochastic programs. *SIAM Journal on Optimization*, 27(3):1772–1800, 2017.

S. W. Wallace and S.-E. Fleten. Stochastic programming models in energy. In R. A. and A. Shapiro, editors, *Stochastic Programming*, Handbooks in Operations Research and Management Science, pages 637 – 677. Elsevier, 2003.

A. S. Werner, A. Pichler, K. T. Midthun, L. Hellemo, and A. Tomasgard. Risk measures in multi-horizon scenario trees. In *Handbook of Risk Management in Energy Production and Trading*, pages 177–201. Springer, 2013.

A Additional Results

We provide additional results for problems with longer planning horizons, namely $T = 10, 12$ and 15 stages, for the case with fixed asset returns and for approximations with up to $T' = 10$ stages. Table 9 shows the results for $\kappa = 10$ and Table 10 with $\kappa = 10$ and a target growth of 6.485% per stage. To be able to solve the original problem, the scenario trees are obtained by sampling two realizations of the random target per stage. The resulting scenario trees count 2^{T-1} scenarios, with $T = 10, 12, 15$.

Table 9: Results for the tracking problem with fixed asset returns and $\tilde{G}_t \approx T[45, 55, 65]$, $t = 2, \ldots, T$, varying the number of stages T in the problem and T' in the approximation. $VSS^{1,T'}[\%]$ represents the percentage increase in objective value generated by the approximation and is computed as $100 \times VSS^{1,T'}/RP^1$. $MSV^{1,T'}[\%]$ represents the percentage increase in the VSS by including stage T' and is computed as $100 \times MSV^{1,T'}/VSS^{1,T'-1}$. The remaining values are expressed in thousands of dollars.

T	T'	$EV^{1,T'}$	$WS^{1,T'}$	$\overline{WS}^{1,T'}$	RP^1	$EEV^{1,T'}$	$VSS^{1,T'}$	$VSS^{1,T'}[\%]$	$MSV^{1,T'}$	$MSV^{1,T'}[\%]$
	1	56.83	61.27	59.86		61.59	0.20	0.33	0.00	0.00
	2	57.15	61.32	59.86		61.48	0.08	0.14	0.12	58.52
	3	57.42	61.39	59.87		61.58	0.08	0.14	0.00	0.00
	4	58.07	61.39	59.96		61.45	0.06	0.10	0.02	29.26
10	5	59.52	61.39	60.26	61.39	61.39	0.00	0.00	0.06	100.00
	6	61.17	61.39	60.64		61.39	0.00	0.00	0.00	0.00
	7	61.22	61.39	61.03		61.39	0.00	0.00	0.00	0.00
	8	61.37	61.39	61.06		61.39	0.00	0.00	0.00	0.00
	9	61.39	61.39	61.39		61.39	0.00	0.00	0.00	0.00
	10	61.39	61.39	61.39		61.39	0.00	0.00	0.00	0.00
	1	83.53	87.54	86.28		88.04	0.37	0.43	0.00	0.00
	2	83.76	87.64	86.28		87.79	0.11	0.13	0.26	69.29
12					87.67					

	3	83.93	87.66	86.28		87.70	0.03	0.03	0.09	76.80
	4	84.22	87.67	86.30		87.70	0.03	0.03	0.00	0.00
	5	84.84	87.67	86.37		87.68	0.01	0.01	0.02	64.76
	6	86.44	87.67	86.63		87.67	0.00	0.00	0.01	98.02
	7	87.43	87.67	87.06		87.68	0.00	0.00	0.00	0.00
	8	87.49	87.67	87.34		87.67	0.00	0.00	0.00	0.00
	9	87.66	87.67	87.55		87.67	0.00	0.00	0.00	0.00
	10	87.67	87.67	87.65		87.67	0.00	0.00	0.00	0.00
	1	133.97	138.55	136.90		138.90	0.20	0.15	0.00	0.00
	2	134.06	138.64	136.90		138.85	0.15	0.11	0.05	25.93
	3	134.15	138.69	136.90		138.87	0.15	0.11	0.00	0.00
	4	134.17	138.70	136.91		138.82	0.12	0.09	0.03	20.36
15	5	134.40	138.70	136.92	138.70	138.83	0.12	0.09	0.00	0.00
	6	134.78	138.70	137.00		138.73	0.03	0.02	0.09	76.76
	7	135.60	138.70	137.31		139.00	0.03	0.02	0.00	0.00
	8	137.21	138.70	137.82		139.08	0.03	0.02	0.00	0.00
	9	138.52	138.70	138.16		138.72	0.02	0.01	0.01	35.61
	10	138.55	138.70	138.31		138.70	0.00	0.00	0.02	90.36

Table 10: Results for the tracking problem with fixed asset returns and $\tilde{G}_t \approx T[45(1+\xi)^{t-1}, 55(1+\xi)^{t-1}, 65(1+\xi)^{t-1}]$, $t = 2, \ldots, T$, and $\xi = 6.485\%$, varying the number of stages T in the problem and T' in the approximation. $VSS^{1,T'}[\%]$ represents the percentage increase in objective value generated by the approximation and is computed as $100 \times VSS^{1,T'}/RP^1$. $MSV^{1,T'}[\%]$ represents the percentage increase in the VSS by including stage T' and is computed as $100 \times MSV^{1,T'}/VSS^{1,T'-1}$. The remaining values are expressed in thousands of dollars.

T	T'	$EV^{1,T'}$	$WS^{1,T'}$	$\overline{WS}^{1,T'}$	RP^1	$EEV^{1,T'}$	$VSS^{1,T'}$	$VSS^{1,T'}[\%]$	$MSV^{1,T'}$	$MSV^{1,T'}[\%]$
	1	0.03	20.56	20.31		27.72	1.66	6.37	0.00	0.00
	2	5.12	20.63	21.56		26.67	0.61	2.35	1.05	63.17
	3	7.27	21.06	22.52		26.13	0.07	0.26	0.54	89.02
	4	9.67	21.82	23.29		26.06	0.00	0.00	0.07	99.95
	5	12.93	22.38	24.01		26.06	0.00	0.00	0.00	0.00
10	6	16.14	23.16	24.54	26.06	26.06	0.00	0.00	0.00	0.00
	7	18.42	24.05	25.24		26.06	0.00	0.00	0.00	0.00
	8	21.39	24.83	25.83		26.06	0.00	0.00	0.00	0.00
	9	23.57	26.06	25.86		26.06	0.00	0.00	0.00	0.00
	10	26.06	26.06	26.06		26.06	0.00	0.00	0.00	0.00
	1	0.03	26.21	25.96		35.62	2.45	7.38	0.00	0.00
	2	5.12	26.28	26.98		34.28	1.12	3.36	1.33	54.40
	3	7.27	26.69	27.77		33.27	0.10	0.32	1.01	90.59
	4	9.67	27.43	28.79		33.18	0.01	0.02	0.10	93.16
	5	12.93	27.96	29.66		33.17	0.00	0.00	0.01	100.00
12	6	16.24	28.73	30.40	33.17	33.17	0.00	0.00	0.00	0.00
	7	18.69	29.55	31.12		33.17	0.00	0.00	0.00	0.00
	8	22.19	30.41	31.64		33.17	0.00	0.00	0.00	0.00
	9	25.57	31.39	32.55		33.17	0.00	0.00	0.00	0.00
	10	29.12	32.10	32.94		33.17	0.00	0.00	0.00	0.00
	1	5.97	45.74	45.49		59.56	3.37	5.99	0.00	0.00
	2	11.06	45.80	47.35		58.10	1.91	3.40	1.46	43.33
	3	14.63	46.22	48.37		56.48	0.29	0.51	1.62	85.05
	4	18.68	46.96	49.37		56.26	0.06	0.12	0.22	77.29
15	5	22.24	47.48	50.11	56.19	56.20	0.01	0.02	0.05	84.38

6	25.68	48.25	50.83	56.19	0.00	0.00	0.01	83.58
7	28.38	49.05	51.82	56.19	0.00	0.00	0.00	0.00
8	31.96	49.96	52.68	56.19	0.00	0.00	0.00	0.00
9	35.36	50.95	53.41	56.19	0.00	0.00	0.00	0.00
10	38.99	51.98	54.13	56.19	0.00	0.00	0.00	0.00