On pre-commitment aspects of a time-consistent strategy for a mean-variance investor

F. Cong [a,*], C.W. Oosterlee [a,b]

[a] TU Delft, Delft Institute of Applied Mathematics, HB 03.270, Mekelweg 4, 2628 CD Delft, The Netherlands
[b] CWI-Centrum Wiskunde & Informatica, Amsterdam, The Netherlands

ABSTRACT

In this paper, a link between a time-consistent and a pre-commitment investment strategy is established. We define an implied investment target, which is implicitly contained in a time-consistent strategy at a given time step and wealth level. By imposing the implied investment target at the initial time step on a time-consistent strategy, we form a hybrid strategy which may generate better mean-variance efficient frontiers than the time-consistent strategy. We extend the numerical algorithm proposed in Cong and Oosterlee (2016b) to solve constrained time-consistent mean-variance optimization problems. Since the time-consistent and the pre-commitment strategies generate different terminal wealth distributions, time-consistency is not always inferior to pre-commitment.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Since Markowitz's pioneering work (Markowitz, 1952) on a single-period investment model, the mean-variance portfolio optimization problem has become a very attractive topic in academic and industrial research and development. One way to extend Markowitz's work is to consider multi-period mean-variance optimization.

Dynamic mean-variance optimization is not a trivial task, as in general the Bellman dynamic programming principle (Bellman, 1957) should be applied to this kind of path-dependent optimization problem. Due to the nonlinearity of the variance operator, however, the mean-variance problem cannot be solved in this manner.

In Zhou and Li (2000) and Li and Ng (2000), the authors introduced an embedding technique, by which the original mean-variance problem was formulated as a tractable linear-quadratic (LQ) problem. Instead of pursuing an optimal balance between profit and risk, an investor then designs an investment strategy to minimize the difference between her wealth and a pre-determined investment target. In Basak and Chabakauri (2010), the optimal strategy for the LQ problem is termed *pre-commitment* strategy. Related research on the pre-commitment mean-variance portfolio optimization problem has been performed in Li et al. (2002), Zhu et al. (2004), Bielecki et al. (2005), Fu et al. (2010) and Cui et al. (2014a), where realistic constraints were considered in the dynamic optimization process. Moreover, since target-based features of the pre-

---

* Corresponding author.
 E-mail address: f.cong@tudelft.nl (F. Cong).

commitment strategy fit well to pension plan investment, some discussion regarding this strategy arose in the insurance management literature as well, see Haberman and Vigna (2002), Gerrard et al. (2004) and Vigna (2014).

However, as noticed in Basak and Chabakauri (2010), after an initial time step, a real world investor will be on only one of many possible stochastic paths and an investor with a determined risk aversion attitude may then have an incentive to depart from the pre-commitment strategy. Basak and Chabakauri (2010) emphasize that an optimal dynamic strategy should be a *time-consistent* strategy, i.e. given an optimal investment policy for the complete mean-variance problem, a sub-policy should also constitute the optimal policy for the corresponding sub-problem. As spotted in Wang and Forsyth (2011, 2012) and Chen et al. (2013), introducing the time-consistency condition requires casting additional constraints on the controls and the mean-variance efficient frontier generated by a time-consistent strategy is thus usually lower than the frontier generated by a pre-commitment strategy. In the case of discrete and continuous re-balancing, Czichowsky (2013) developed a time-consistent formulation of mean-variance efficiency in a semimartingale framework. Extensions of the time-consistent mean-variance strategy have been made by modeling the risk aversion parameter as a state-dependent variable, see Hu et al. (2012), Cui et al. (2014b, 2015), and Björk et al. (2014). The strategy of a dynamic mean-variance investor is then defined to be *time-consistent in efficiency*, which does not necessarily meet the time-consistency conditions but may yield higher mean-variance efficient frontiers. Despite some discussion on the pre-commitment and the time-consistent mean-variance strategies in the operational research literature, the link between these strategies is mainly at the stage of problem formulation: a time-consistent strategy can be obtained by casting a time-consistency constraint on a pre-commitment strategy.

In this paper, the point-of-departure is to focus on the relation between the two strategies. Since for the pre-commitment as well as for the time-consistent strategy the optimal allocations for a mean-variance investor are available, we can establish a link between the former, which is driven by a target, and the latter, driven by a risk aversion parameter. In the case where a pre-commitment investor adopts the same asset allocations as a time-consistent investor, a one-to-one function between the target and the risk aversion parameter can be established. This function gives information on how the risk aversion attitude of a pre-commitment investor changes and how much wealth a time-consistent investor desires during the dynamic investment process. We call the target for the time-consistent investor *the implied target*.

During an investment period, the implied target of a time-consistent investor varies while the target of a pre-commitment investor remains fixed. This makes the distributions of their potential terminal wealth significantly different. Reflecting on this, we define a hybrid strategy by introducing a fixed target into the time-consistent strategy. This hybrid mean-variance strategy is dual to a mean-variance strategy which minimizes the shortfall to a target. Introduction of a target into the mean-variance formulation essentially makes a strategy semi-self-financing. As explained in Cui et al. (2012) and Dang and Forsyth (2016), introducing semi-self-financing into a pre-commitment strategy can improve the mean-variance efficient frontiers.

In principle, any target higher than the risk-free target could be a reasonable target for the hybrid strategy. However, if the target is too high, a hybrid investor will behave like the time-consistent investor. In this paper, we choose the target to be equal to the *implied target at the initial time step* of a time-consistent investor. With this setting, we find that the mean-variance frontier generated by the hybrid strategy is higher than the one generated by the time-consistent strategy. In the constrained case, the mean-variance frontier generated by the hybrid strategy is often almost identical to the one generated by the pre-commitment strategy. In this case the hybrid investor behaves more like a pre-commitment than a time-consistent investor.

Since different strategies may generate different terminal wealth distributions, the mean-variance efficient frontier may not tell us all about the performance of a strategy. When considering *partial variance*, for example, we find that the time-consistent strategy is not always inferior to the pre-commitment strategy. If we only consider downside risk, the time-consistent strategy and the pre-commitment strategy perform comparably. Moreover, in terms of the potential to achieve higher returns, the time-consistent strategy may perform even better.

In Cong and Oosterlee (2016b), we proposed an efficient simulation-based algorithm for solving the pre-commitment mean-variance problem with constraints on the controls. In the unconstrained case, the optimal pre-commitment strategy was shown to be identical to the strategy taken by a "myopic investor", who designs her optimal strategy without taking future optimality into account and therefore simply utilizes a forward strategy. In this paper, we will extend the numerical algorithm, based on *a forward followed by a backward iteration* proposed in Cong and Oosterlee (2016b), to the time-consistent problem. Solving the constrained time-consistent mean-variance problem with a simulation-based approach has, to our knowledge, not yet been performed previously. Since time-consistency is not only a constraint but also a condition, Basak and Chabakauri (2010) handle the variance operator by writing it as the difference of two variables being the mean of the square of the variable and the square of the variable's mean. With such modifications in our algorithm, highly satisfactory results for the constrained time-consistent problem can again be obtained. In the unconstrained case the optimal strategy of a time-consistent investor can be duplicated by a myopic investor as well.

This paper is organized as follows. In Section 2, we describe the dynamic mean-variance problem, the pre-commitment and the time-consistent strategies. Section 3 has its focus on the time-consistent strategy. We derive an analytic formula for the time-consistent optimal allocation in Section 3.1. Reflecting on this, we compare the time-consistent and the pre-commitment strategies in Section 3.2. The hybrid strategy, which is constructed by inserting a target in the time-consistent strategy, is introduced in Section 3.3. Section 3.4 introduces the partial variance for evaluating the mean-variance strategies. Our numerical algorithm is presented in Section 4 and following that in Section 5 numerical tests are performed. We conclude in the last section.

2. Problem formulation

This section describes the dynamic portfolio optimization problem for a defined contribution pension plan. We assume that the financial market is defined on a probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$ with finite time horizon $[0, T]$. The state space Ω is the set of all realizations of the financial market within the time horizon $[0, T]$. $\mathcal{F}$ is the sigma algebra of events till time T, i.e. $\mathcal{F} = \mathcal{F}_T$. We assume that the filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ is generated by the price processes of the financial market and augmented with the null sets of $\mathcal{F}$.

For convenience, we consider a portfolio consisting of one risk-free and one risky asset. To extend the analysis to a problem with more than one risky asset is feasible. We assume that the re-balancing dates for the portfolio are equidistantly distributed and that the total number of re-balancing opportunities before terminal time T is equal to M. The time step Δt between two re-balancing dates is thus $\frac{T}{M}$, and the portfolio can be traded at times $t \in [0, \Delta t, ..., T - \Delta t]$. At each trading time t, an investor decides the trading strategy to maximize the expectation of the terminal wealth and to minimize the investment risk. Formally, the investor's problem is given by

$$J_t(W_t) = \max_{\{x_s\}_{s=t}^{T-\Delta t}} [\mathbb{E}[W_T | W_t] - \lambda \cdot \text{Var}[W_T | W_t]], \tag{1}$$

subject to the wealth restriction:

$$W_{s+\Delta t} = W_s \cdot (x_s R_s^e + R_f) + C \cdot \Delta t, \quad s = t, t+\Delta, ..., T - \Delta t. \tag{2}$$

Here x_s denotes the asset allocation of the investor's wealth in the risky asset in the time period $[s, s+\Delta t)$. The admissible investment strategy x_t is assumed to be adapted. R_f is the return of the risk-free asset in one time step, which is assumed to be constant for simplicity, and R_s^e is the excess return of the risky asset during $[s, s+\Delta t)$. We assume that the excess returns $\{R_t^e\}_{t=0}^{T-\Delta t}$ are statistically independent.[1] $C \cdot \Delta t$ stands for a contribution of the investor in the portfolio during $[s, s+\Delta t)$, and a negative C can be interpreted as a constant withdrawal of the investor from the portfolio. The risk aversion attitude of the investor is denoted by λ, which is a trade-off factor between maximizing the profit and minimizing the risk. $J_t(W_t)$ is termed the *value function*, which measures the investor's investment opportunities at time t with wealth W_t.

The difficulty of solving the mean-variance optimization problem in Eq. (1) is caused by the nonlinearity of conditional variances, $\text{Var}[\text{Var}[W_T | \mathcal{F}_t] | \mathcal{F}_s] \neq \text{Var}[W_T | \mathcal{F}_s], s \leq t$. To tackle this problem, the original mean-variance equations can be transformed, as was done in Zhou and Li (2000), Li and Ng (2000) and Wang and Forsyth (2010). The following theorem supports the transformation.

Theorem 2.1. *If $\{x_s^*\}_{s=t}^{T-\Delta t}$ is the optimal control for the problem defined in Eq. (1), then $\{x_s^*\}_{s=t}^{T-\Delta t}$ is also the optimal control for the following problem*:

$$\min_{\{x_s\}_{s=t}^{T-\Delta t}} \left[\mathbb{E}[(W_T - \gamma)^2 | W_t] \right] \tag{3}$$

where $\gamma = \frac{1}{2\lambda} + \mathbb{E}_{x^*}[W_T | W_t]$. Here the operator $\mathbb{E}_{x^*}[\cdot]$ denotes the expectation of the investor's terminal wealth if she invests according to the optimal strategy $\{x_s^*\}_{s=t}^{T-\Delta t}$.

Proof. See Li and Ng (2000).

Based on this theorem, the original mean-variance problem was embedded into a tractable auxiliary LQ problem. The investment strategy corresponding to this LQ problem is called the pre-commitment strategy. A pre-commitment investor always pursues a pre-determined target and thus invests according to the pre-determined strategy. However, this also implies that the investor does not stay with the initial risk aversion attitude.

If an investor has a *time-consistent risk aversion*, then the investment strategy is termed the time-consistent strategy as in Basak and Chabakauri (2010). Mathematically, the time-consistent strategy can be defined as follows:

Definition 2.2 (*Time-consistent strategy*). At time t, the time-consistent optimal control $\{x_s^{*tc}\}_{s=t}^{T-\Delta t}$ is defined by the optimal control for Eq. (1) with an additional time-consistency condition requiring that $\{x_s^{*tc}\}_{s=\tau}^{T-\Delta t}$ also constitutes a time-consistent optimal control for:

$$J_\tau(W_\tau) = \max_{\{x_s\}_{s=\tau}^{T-\Delta t}} [\mathbb{E}[W_T | W_\tau] - \lambda \cdot \text{Var}[W_T | W_\tau]], \tag{4}$$

for $\tau = t + \Delta t, t + 2 \cdot \Delta t, ..., T - \Delta t$.

In real-life applications, various constraints can be cast on the asset allocations. For example, the no bankruptcy constraint, which requires that the asset allocation does not lead to bankruptcy of the investor, and the leverage constraint, which forms bounds on the asset allocation.

---

[1] For a risky asset with dynamics following geometric Brownian motion or a Lévy process, the independence structure is valid. For some problems where the asset returns are directly defined, for example, as model-free data, the independence assumption is also correct. For a VAR or GARCH model, this assumption is, however, not satisfied.

Remark 2.3. In this paper "no-bankruptcy constraints" on asset allocations indicate that an investor has zero probability of bankruptcy. However, since the return of a risky asset can be any real number, prescribing a zero probability of bankruptcy is very strict. Instead, we consider a no-bankruptcy constraint with $1 - 2 \cdot \Omega\%$ certainty, see also Cong and Oosterlee (2016b). At time t, for an investor with wealth W_t, the constraint on the asset allocation x_t reads:

$$\frac{-C\Delta t - W_t \cdot R_f}{W_t \cdot R_t^{e,1-\Omega}} \le x_t \le \frac{-C\Delta t - W_t \cdot R_f}{W_t \cdot R_t^{e,\Omega}}, \tag{5}$$

$R_t^{e,\Omega}$ and $R_t^{e,1-\Omega}$ respectively denote the Ω- and the $(1-\Omega)$-quantiles of the excess return R_t^e.

When we consider the bounded leverage constraint, we set upper and lower bounds for the asset allocation. The bounded leverage constraint reads:

$$x_{\min} \le x_t \le x_{\max}.$$

Remark 2.4. The authors of Björk et al. (2014) introduced a state-dependent risk aversion parameter, which can be determined by the amount of wealth in a portfolio. Formally, the state-dependent risk aversion parameter reads:

$$\Lambda(W_t) = \frac{\theta}{W_t},$$

where θ is a fixed constant.

The state-dependent risk aversion parameter implies that an investor will be less risk averse when she has a large amount of money and more risk averse when she is almost bankrupt. This may be reasonable for a real-world investor, but it also leads to some problems. For example, when an investor is close to bankruptcy, applying the risk-averse strategy indicates that it is hardly possible to get back to sufficient wealth. On the other hand, with a large amount of money, she is less risk-averse and willing to "gamble", which may lead to significant losses. To sum up, although this way of forming the risk aversion parameter seems reasonable, the corresponding investment strategy has "easy-to-lose" and "hard-to-recover" features. Therefore, as shown in Wang and Forsyth (2011) and confirmed in our numerical tests, this strategy will generate a lower efficient frontier than the plain time-consistent strategy.

In the following section, we will analytically derive the optimal unconstrained multi-period time-consistent strategy. Besides, reflecting on our findings in Cong and Oosterlee (2016b) about the pre-commitment strategy, we will make a thorough discussion comparing the time-consistent strategy and the pre-commitment one.

3. The time-consistent strategy

3.1. The optimal asset allocation

Under the following condition, we can derive an analytic expression for the asset allocations in the time-consistent strategy.

Condition 3.1. *The asset allocations at each time step are unconstrained.*

The formulation of the asset allocations for the time-consistent strategy is described in the following theorem. Since it is not easy to derive the value function of the time-consistent strategy (because of the nonlinearity of the variance operator), our proof is based on the essential properties of an "optimal control".

Theorem 3.2. *Assuming that x_t^{*tc} is the optimal asset allocation for the time-consistent mean-variance optimization problem, it can be calculated by*:

$$x_t^{*tc} = \frac{\mathbb{E}[R_t^e]}{2 \cdot \lambda \cdot W_t \cdot R_f^{(T-t)/\Delta t - 1} \cdot \text{Var}[R_t^e]}, \tag{6}$$

where $t = 0, \Delta t, \ldots, T - \Delta t$.

Proof. We consider the dynamic programming technique for solving the time-consistent problem. At time $T - \Delta t$, since the terminal condition at time T is known, the time-consistent strategy will certainly generate the asset allocation as shown in Eq. (6). That is:

$$x_{T-\Delta t}^{*tc} = \frac{\mathbb{E}[R_{T-\Delta t}^e]}{2 \cdot \lambda \cdot W_{T-\Delta t} \cdot \text{Var}[R_{T-\Delta t}^e]}. \tag{7}$$

Subsequently, we consider the optimization problem at time $T - 2\Delta t$, where

$$x_{T-2\Delta t}^{*tc} = \arg \max_{x_{T-2\Delta t}} [\mathbb{E}[W_T | W_{T-2\Delta t}] - \lambda \cdot \text{Var}[W_T | W_{T-2\Delta t}]]. \tag{8}$$

Note that the wealth $W_{T-\Delta t}$ and W_T respectively have the form:

$$W_{T-\Delta t} = W_{T-2\Delta t} \cdot (x_{T-2\Delta t} R_{T-2\Delta t}^e + R_f) + C \cdot \Delta t,$$
$$W_T = W_{T-\Delta t} \cdot (x_{T-\Delta t}^{*tc} R_{T-\Delta t}^e + R_f) + C \cdot \Delta t.$$

Allocation $x_{T-\Delta t}^{*tc}$ indicates that the optimal allocation will be chosen at time step $T - \Delta t$. The formulation for the terminal wealth W_T then reads:

$$
\begin{aligned}
W_T &= \left(W_{T-2\Delta t} \cdot (x_{T-2\Delta t} R_{T-2\Delta t}^e + R_f) + C \cdot \Delta t\right) \cdot x_{T-\Delta t}^{*tc} R_{T-\Delta t}^e \\
&\quad + \left(W_{T-2\Delta t} \cdot (x_{T-2\Delta t} R_{T-2\Delta t}^e + R_f) + C \cdot \Delta t\right) \cdot R_f + C \cdot \Delta t \\
&= \frac{\mathbb{E}[R_{T-1}^e] R_{T-1}^e}{2 \cdot \lambda \cdot \mathrm{Var}[R_{T-1}^e]} + W_{T-2\Delta t} \cdot (x_{T-2\Delta t} R_{T-2\Delta t}^e + R_f) \cdot R_f + C \cdot \Delta t \cdot R_f + C \cdot \Delta t \\
&= K_{T-2\Delta t} + W_{T-2\Delta t} \cdot R_{T-2\Delta t}^e \cdot R_f \cdot x_{T-2\Delta t},
\end{aligned}
\tag{9}
$$

where the second equation is based on the property of the optimal control as shown in Eq. (7): $x_{T-\Delta t}^{*tc}$ multiplied by $W_{T-\Delta t}$ will yield a constant. In the last line in Eq. (9), we use $K_{T-2\Delta t}$ to denote all terms that are not related to $x_{T-2\Delta t}$. Since we perform optimization with respect to control variable $x_{T-2\Delta t}$ and returns at different time steps are independent, knowledge of the details of $K_{T-2\Delta t}$ is not needed when solving our optimization problem.

We insert Eq. (9) into the mean-variance formulation and obtain:

$$
\begin{aligned}
&\mathbb{E}[W_T | W_{T-2\Delta t}] - \lambda \cdot \mathrm{Var}[W_T | W_{T-2\Delta t}] \\
&= K_{T-2\Delta t} + W_{T-2\Delta t} \cdot \mathbb{E}[R_{T-2\Delta t}^e] \cdot R_f \cdot x_{T-2\Delta t} - \lambda \cdot W_{T-2\Delta t}^2 \cdot \mathrm{Var}[R_{T-2\Delta t}^e] \cdot (R_f)^2 \cdot x_{T-2\Delta t}^2.
\end{aligned}
$$

Solving the first-order conditions with respect to $x_{T-2\Delta t}$ gives us:

$$
x_{T-2\Delta t}^{*tc} = \frac{\mathbb{E}[R_{T-2\Delta t}^e]}{2 \cdot \lambda \cdot W_{T-2\Delta t} \cdot R_f \cdot \mathrm{Var}[R_{T-2\Delta t}^e]},
\tag{10}
$$

which is identical to the allocation shown in Eq. (6), when $t = T - 2\Delta t$.

The allocations for other time steps can be proved by backward induction in a similar fashion.   □

The asset allocations for the time-consistent strategy shown in Eq. (6) can be explained in a straight-forward way. Consider a myopic investor who only manages the portfolio for the coming period and chooses a risk-free policy afterwards. Then, the optimal time-consistent strategy is identical to the strategy taken by this myopic investor. We will discuss this in some more detail in Section 4.1.

When the periodic contribution C is not included in the portfolio, a general solution for an unconstrained time-consistent multi-period portfolio management problem has been established in the literature, for example in Proposition 5 in Basak and Chabakauri (2010) and in Lemma 3.2 in Czichowsky (2013). However, we notice that introducing periodic contributions does not affect the optimal asset allocations of a time-consistent investor. Meanwhile, as discussed in Cong and Oosterlee (2016b), a pre-commitment investor takes the contribution amount into account when the investment strategy is determined. Knowledge of the periodic contribution is essential for the pre-commitment strategy.

In the remaining analysis part of this paper, we restrict ourselves to the situation where there is no periodic contribution, i.e. $C = 0$. However, extending the analysis to the case $C \neq 0$ is possible.

Remark 3.3. As shown in Eq. (6), the optimal asset allocation in a time-consistent strategy depends on the first and second moments of the excess return rather than on the complete distribution. However, when we use these asset allocations to generate the mean-variance pairs, the distribution of the excess return has an impact. The same holds for the pre-commitment strategy, see Dang et al. (2015) for related discussion.

3.2. Pre-commitment versus time-consistent strategy

By combining our findings in Section 3.1 with those in Cong and Oosterlee (2016b), we can establish a link between the pre-commitment and the time-consistent strategy for the dynamic mean-variance optimization problem. In fact, the pre-commitment strategy can be seen as a strategy consistent with an investment target but not with a risk aversion attitude, while the time-consistent strategy appears as a strategy consistent with a risk aversion attitude but not with an investment target.

According to Cong and Oosterlee (2016b), if a pre-commitment investor has an investment target γ, the optimal control at state (t, W_t) in the unconstrained case is equal to[2]:

$$
x_t^{*pc} = \frac{\left(\gamma - W_t \cdot (R_f)^{(T-t)/\Delta t}\right) \mathbb{E}[R_t^e]}{W_t \cdot (R_f)^{(T-t)/\Delta t - 1} \mathbb{E}[(R_t^e)^2]},
\tag{11}
$$

where the target γ uniquely determines the asset allocation when other parameters are fixed and vice versa.

---

[2] When the periodic contribution is taken into account, the unconstrained asset allocation for the pre-commitment strategy reads:

$$
x_t^{*pc} = \frac{\left(\gamma - W_t \cdot (R_f)^{(T-t)/\Delta t} - C \cdot \Delta t \cdot \frac{1 - R_f^{(T-t)/\Delta t}}{1 - R_f}\right) \mathbb{E}[R_t^e]}{W_t \cdot (R_f)^{(T-t)/\Delta t - 1} \mathbb{E}[(R_t^e)^2]},
$$

where the periodic contribution obviously has an impact on the pre-commitment strategy.

Suppose that a time-consistent investor has risk aversion λ, at state (t, W_t) the time-consistent optimal asset allocation then reads:

$$x_t^{*tc} = \frac{\mathbb{E}[R_t^e]}{2 \cdot \lambda \cdot W_t \cdot R_f^{(T-t)/\Delta t - 1} \cdot \text{Var}[R_t^e]},$$

where the asset allocation is also uniquely determined by risk aversion parameter λ.

At state (t, W_t), if a pre-commitment investor has the same optimal allocation as x_t^{*tc}, because the allocation x_t^{*pc} and the investment target γ form a one-to-one function, as shown in Eq. (11), we can calculate an *implied investment target* γ_t^{imp} by solving:

$$\frac{\left(\gamma_t^{\text{imp}} - W_t \cdot (R_f)^{(T-t)/\Delta t}\right) \mathbb{E}[R_t^e]}{W_t \cdot (R_f)^{(T-t)/\Delta t - 1} \mathbb{E}[(R_t^e)^2]} = \frac{\mathbb{E}[R_t^e]}{2 \cdot \lambda \cdot W_t \cdot R_f^{(T-t)/\Delta t - 1} \cdot \text{Var}[R_t^e]}.$$

This gives us:

$$\gamma_t^{\text{imp}} = W_t R_f^{(T-t)/\Delta t} + \frac{\mathbb{E}[(R_t^e)^2]}{2\lambda \text{Var}[R_t^e]}. \tag{12}$$

The reasoning for using an "implied investment target" is similar to that for the well-known "implied volatility" in option pricing problems. An option can be priced by an involved model, but the Black–Scholes model is utilized for generating the implied volatility, which reflects the volatility in the market. In our case, the asset allocation corresponds to the "option price" and the time-consistent strategy to the "involved model". The "implied investment target" informs us about the target hidden in the strategy of a time-consistent investor. A similar idea is introduced in Cui et al. (2012), where an induced risk aversion is established for a pre-commitment investor.

In general, it is hard to say whether being consistent with a target or being consistent with a risk aversion attitude is best. Fixing a criterion, either the investment target or the risk aversion parameter, essentially means that the other criterion will vary during an investment period. According to Eq. (12), at time t, for a wealthy investor with a fixed risk aversion λ, the optimal asset allocation indicates that a higher investment target is desired. On the contrary, in the case of lower wealth values, the implied investment target will be lower, which indicates that a time-consistent investor will focus on a lower terminal wealth. In short, a time-consistent investor has a varying target whereas a pre-commitment investor has a fixed target throughout the investment process.

Adopting a target-oriented strategy essentially keeps a pre-commitment investor away from risky strategies and, of course, from potential profits when the amount of wealth is sufficiently large. In a similar situation, a time-consistent investor will act and perform differently, i.e. when the amount of wealth is large, she desires to gain more due to the consistent risk aversion throughout the investment period. This investor is thus never satisfied with a predetermined target.

The differences between the time-consistent and the pre-commitment strategies also lead to differences in the distributions of the corresponding terminal wealth profiles. For a pre-commitment investor, the terminal wealth follows a distribution with a fat left-tail and a large portion of the potential terminal wealth is located very close to the target. For a time-consistent investor, the distribution of terminal wealth is usually rather symmetric. An illustrative plot is presented in Fig. 1.

Remark 3.4. If we consider the time-consistency in efficiency, as proposed in Cui et al. (2012), Eq. (12) indicates that an investor should have a target that is no less than the risk-free target to guarantee that an investment strategy is rational, i.e. the risk aversion parameter should be non-negative.

3.3. Hybrid strategy: time-consistent with a determined target

Although time-consistency appears to be an important element of a real-life investment strategy, taking time-consistency into account is like casting additional constraints on the asset allocations and a time-consistent strategy therefore generates typically lower efficient frontiers than the pre-commitment strategy (Basak and Chabakauri, 2010; Wang and Forsyth, 2011).

As discussed in Section 3.2, an essential difference between the time-consistent and the pre-commitment strategy is the notion of *being satisfied with a target, or not*. In this section we design a *hybrid strategy* which combines features of the time-consistent and the pre-commitment strategy. Our aim is straight-forward: we impose intermediate investment targets on a time-consistent strategy and once the target at an intermediate time step is achieved the investor is forced to take a risk-free investment strategy at subsequent time steps. Following this hybrid strategy, a time-consistent investor is satisfied with a target and thus behaves like a pre-commitment investor when the wealth is sufficient. On the other hand, we still assume that the investor has a time-consistent risk aversion when the amount of wealth does not reach the target. Therefore, in this scenario the investor still acts in a time-consistent way.

Mathematically, intermediate targets can be introduced to a time-consistent strategy in a similar way as done in Dang and Forsyth (2016), where intermediate targets are introduced to a pre-commitment strategy. The definition of the hybrid strategy is as follows.
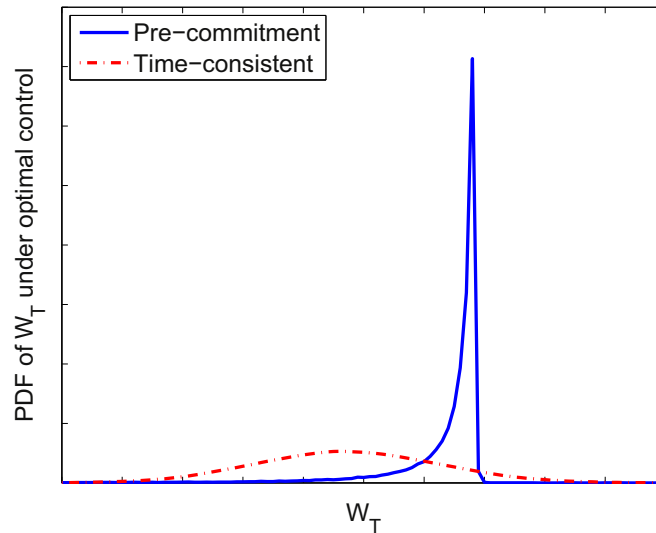
Fig. 1. Comparison of the probability density functions of the terminal wealth, which is achieved by respectively performing unconstrained pre-commitment and time-consistent strategies. For both strategies, the investment horizon, the initial wealth and the return process are the same. By carefully choosing investment target γ and trade-off parameter λ, we make them generate a terminal wealth with the same variance, but the distributions of the terminal wealth appear to be very different.

Definition 3.5 (*Hybrid strategy*). Assume that a terminal investment target θ_T is given. At time t, the hybrid optimal control $\{x_s^{*hy}\}_{s=t}^{T-\Delta t}$ is defined by the optimal control for

$$J_t^{hy}(W_t) = \max_{\{x_s\}_{s=t}^{T-\Delta t}} [\mathbb{E}[\min(W_T, \theta_T)|W_t] - \lambda \cdot \mathrm{Var}[\min(W_T, \theta_T)|W_t]],$$

with an additional condition requiring that $\{x_s^{*hy}\}_{s=\tau}^{T-\Delta t}$ also constitutes an optimal control for:

$$J_\tau^{hy}(W_\tau) = \max_{\{x_s\}_{s=\tau}^{T-\Delta t}} [\mathbb{E}[\min(W_T, \theta_T)|W_\tau] - \lambda \cdot \mathrm{Var}[\min(W_T, \theta_T)|W_\tau]],$$

for $\tau = t + \Delta t, t + 2 \cdot \Delta t, ..., T - \Delta t$.

Due to the target-based feature, the optimal asset allocations of a hybrid strategy are known immediately in some situations. If we define intermediate targets $\{\theta_s\}_{s=t}^{T-\Delta t}$ by

$$\theta_s = \frac{\theta_T}{(R_f)^{(T-s)/\Delta t}}, \quad s = t, t + \Delta t, ..., T - \Delta t,$$

the optimal asset allocation $\{x_s^{*hy}\}_{s=t}^{T-\Delta t}$ for the hybrid strategy satisfies:

$$x_s^{*hy} = 0, \quad \text{if } W_s > \theta_s, \tag{13}$$

for $s = t, t + \Delta t, ..., T - \Delta t$.

Since the hybrid strategy can be interpreted as imposing time-consistent constraints on a pre-commitment strategy when the wealth level is low, the hybrid strategy typically generates lower efficient frontiers than the pre-commitment strategy. However, as the hybrid strategy can also be regarded as one that removes the time-consistency requirements in the time-consistent strategy when the wealth level is high, the hybrid strategy should generally generate better efficient frontiers than the time-consistent strategy.

A suitable target? Introducing an investment target for a pre-commitment strategy is not new, see, for example, Cui et al. (2012) and Dang and Forsyth (2016). Prescribing an investment target in a time-consistent strategy has however, to our knowledge, not yet been pursued. In this setting, the question is how to define a suitable target θ_T. When the investment target is very high, it will be difficult to reach it and the hybrid strategy will just be the same as the plain time-consistent strategy. When the investment target is low, for example the risk-free target, the investor will always follow the risk-free strategy.

In this paper, we propose the following approach to generate an investment target in the case where returns of risky assets are identically distributed.[3] We introduce a tuning parameter k in the *implied investment target* at time $t=0$ and form

---

[3] When the returns at different time steps are not identically distributed, our proposed way to generate the target may be questionable and the conclusions regarding the choice of k in Eq. (14) may be incorrect.

the terminal target by:

$$\theta_T = W_0 R_f^{T/\Delta t} + k \cdot \frac{\mathbb{E}[(R_0^e)^2]}{2\lambda \text{Var}[R_0^e]}. \tag{14}$$

When $k=1$, Eq. (14) generates the implied investment target at the initial time step. For larger k-values, it will be increasingly difficult to reach the investment target and the hybrid strategy will be similar to a time-consistent strategy. In the numerical section, we will also numerically analyze the impact of different k-values for the targets in the hybrid strategy. We found that generating the target using $k=1$ improves the efficient frontiers of the time-consistent strategy. In some cases, the efficient frontier generated by this hybrid strategy is almost identical to the one generated by the pre-commitment strategy. Introducing "targets" in this fashion makes the hybrid investor resemble more a pre-commitment than a time-consistent investor. However, if the predetermined target is too high for a hybrid investor, she will just follow the time-consistent strategy.

Minimizing mean-variance of shortfall: The value function of a hybrid strategy reads:

$$J_t^{\text{hy}}(W_t) = \max_{\{x_s\}_{s=t}^{T-\Delta t}} \left[\mathbb{E}[h(W_T)|W_t] - \lambda \cdot \text{Var}[h(W_T)|W_t] \right], \tag{15}$$

where function $h(u)$ is defined by:

$$h(u) = \min(u, \theta_T).$$

Since we have

$$\min(W_T, \theta_T) = \min(W_T - \theta_T, 0) + \theta_T = -\max(\theta_T - W_T, 0) + \theta_T,$$

and if we define $\hat{h}(u) = \max(\theta_T - u, 0)$, the optimal control for value function (15) can be calculated in a duality form:

$$\arg \max_{\{x_s\}_{s=t}^{T-\Delta t}} \left[\mathbb{E}[h(W_T)|W_t] - \lambda \cdot \text{Var}[h(W_T)|W_t] \right]$$

$$= \arg \min_{\{x_s\}_{s=t}^{T-\Delta t}} \left[\mathbb{E}[\hat{h}(W_T)|W_t] + \lambda \cdot \text{Var}[\hat{h}(W_T)|W_t] \right].$$

The duality form stands for an optimal control of minimizing the shortfall, which is defined by $\max(\theta_T - W_T, 0)$. A similar way of formulating the problem with the shortfall is proposed in Dang and Forsyth (2014).

3.4. Evaluation with partial variance

Being satisfied with a target may be a double-edged sword. A pre-commitment investor is not exposed to the risk of loosing money when she is sufficiently rich but also gives up the chance to gain higher profits. As mentioned in Section 3.2, this makes time-consistent and pre-commitment strategies different in terms of the distributions of the terminal wealth. Although it is reported in Wang and Forsyth (2011) that the pre-commitment strategy always generates higher mean-variance efficient frontiers than the time-consistent strategy, a portfolio manager may worry about the composition of the variance. All paths that are either lower or higher than the expectation contribute to the variance, but only the lower part is undesired since it reflects downside risk, i.e. potential losses. In this section, we discuss *partial variance* as another measure to evaluate the performance of a portfolio management strategy.

Partial variance: For a random variable L, one way to measure the downside risk is to use the Lower Partial Variance (LPV), defined by:

$$\text{LPV}(L) = \mathbb{E}\left[\left(\min\left(L - \mathbb{E}[L], 0\right) \right)^2 \right].$$

Meanwhile, we can also define the Upper Partial Variance (UPV) by:

$$\text{UPV}(L) = \mathbb{E}\left[\left(\max\left(L - \mathbb{E}[L], 0\right) \right)^2 \right].$$

Notice that the variance of L is composed of LPV and UPV:

$$\text{var}(L) = \text{LPV}(L) + \text{UPV}(L).$$

4. A simulation-based algorithm

In this section, we propose a simulation-based numerical algorithm for generating the time-consistent mean-variance portfolio optimization solution. Our algorithm consists of two steps: a forward approximation and a backward programming iteration.

4.1. Forward iteration: the multi-stage strategy

In Section 3.1 we mentioned that in the unconstrained case a time-consistent investor chooses the same asset allocations as a myopic investor. We use the name *multi-stage* investor here, because the investor solves the problem in several separate stages. The optimal asset allocation for a multi-stage investor at state (t, W_t), i.e. time t and wealth W_t, reads:

$$x_t^{*ms} = \arg\max_{x_t} \left[\mathbb{E}\left[W_t \cdot (x_t R_t^e + R_f) \prod_{s=t+\Delta t}^{T-\Delta t} (R_f) | W_t \right] \right.$$
$$\left. - \lambda \cdot \mathrm{Var}\left[W_t \cdot (x_t R_t^e + R_f) \prod_{s=t+\Delta t}^{T-\Delta t} (R_f) | W_t \right] \right]. \tag{16}$$

Solving the static mean-variance optimization problem as shown in Eq. (16) *in a forward fashion* yields the optimal control for the multi-stage strategy, which reads:

$$x_t^{*ms} = \frac{\mathbb{E}[R_t^e]}{2 \cdot \lambda \cdot W_t \cdot R_f^{(T-t)/\Delta t - 1} \cdot \mathrm{Var}[R_t^e]}. \tag{17}$$

If we consider the value function of the time-consistent strategy, then the optimal control at time t can be defined by:

$$x_t^{*tc} = \arg\max_{x_t} \left[\mathbb{E}\left[W_t \cdot (x_t R_t^e + R_f) \prod_{s=t+\Delta t}^{T-\Delta t} (x_s^{*tc} R_s^e + R_f) | W_t \right] \right.$$
$$\left. - \lambda \cdot \mathrm{Var}\left[W_t \cdot (x_t R_t^e + R_f) \prod_{s=t+\Delta t}^{T-\Delta t} (x_s^{*tc} R_s^e + R_f) | W_t \right] \right], \tag{18}$$

where $\{x_s^{*tc}\}_{s=t+\Delta t}^{T-\Delta t}$ denote the future optimal controls in the time-consistent strategy.

As shown in Eqs. (16) and (18), the value function of the multi-stage strategy is clearly different from the value function of the time-consistent strategy. However, the optimal controls in Eqs. (6) and (17) for these two strategies are identical in the unconstrained case, i.e. $x_t^{*tc} = x_t^{*ms}$ for any t.

Similar to in Cong and Oosterlee (2016b), we consider the multi-stage strategy as a sub-optimal solution for a constrained optimization problem.

The multi-stage strategy can be derived in a forward fashion. As time proceeds, new information becomes available and a new optimal multi-stage solution is obtained for the new state. Therefore, the multi-stage strategy is very efficient for generating a sub-optimal solution for a constrained time-consistent problem and it constitutes a forward solution in our algorithm.

Remark 4.1. In order to generate a forward solution for a hybrid strategy, we can combine the multi-stage strategy with the constraint as in Eq. (13). However, even in the unconstrained case, this only constitutes a sub-optimal solution for the hybrid case.

4.2. Backward recursive programming iteration

By using the multi-stage strategy, we can solve the time-consistent problem in a forward manner by Monte-Carlo simulation. However, in a general situation, the multi-stage strategy is usually not the optimal solution to the dynamic optimization problem. In order to find the optimal solution, we need to consider a backward dynamic programming solution. In this section, we present an approach to perform a backward recursive iteration based on the solution of the multi-stage forward iteration.

Decomposing the variance operator: As mentioned, a problem for the efficient solution of the dynamic mean-variance optimization problem is caused by the nonlinearity of the variance operator. To deal with this problem, as in Basak and Chabakauri (2010) and Wang and Forsyth (2011), we consider the conditional mean of the terminal wealth and the conditional mean of the terminal wealth square, respectively, as two separate value functions. Then, at time t, we have:

$$U_t(W_t) = \mathbb{E}^*[U_T(W_T)|W_t], \quad V_t(W_t) = \mathbb{E}^*[V_T(W_T)|W_t],$$

with terminal conditions

$$U_T(W_T) = W_T, \quad V_T(W_T) = W_T^2. \tag{19}$$

Here, $\mathbb{E}^*[U_T(W_T)|W_t]$ and $\mathbb{E}^*[V_T(W_T)|W_t]$ indicate the expectations of the terminal value functions when the time-consistent optimal control is cast on the portfolio in period $[t, T]$.

Remark 4.2. When we implement this algorithm for generating hybrid strategies, we only need to modify the terminal conditions. In that case, the terminal conditions $U_T^{hy}(W_T)$ and $V_T^{hy}(W_T)$ read:

$$U_T^{hy}(W_T) = \min(W_T, \theta_T), \quad V_T^{hy}(W_T) = (\min(W_T, \theta_T))^2.$$

The value function $J_t(W_t)$ as shown in Eq. (1) can then be rewritten as:

$$
\begin{aligned}
J_t(W_t) &= \max_{\{x_s\}_{s=t}^{T-\Delta t}} [\mathbb{E}[W_T|W_t] - \lambda \cdot \mathrm{Var}[W_T|W_t]] \\
&= \mathbb{E}^*[W_T|W_t] - \lambda \cdot \left(\mathbb{E}^*[W_T^2|W_t] - (\mathbb{E}^*[W_T|W_t])^2 \right) \\
&= \mathbb{E}^*[\mathbb{E}^*[W_T|W_{t+\Delta t}]|W_t] \\
&\quad - \lambda \cdot \left(\mathbb{E}^*[\mathbb{E}^*[W_T^2|W_{t+\Delta t}]|W_t] - (\mathbb{E}^*[\mathbb{E}^*[W_T|W_{t+\Delta t}]|W_t])^2 \right) \\
&= \mathbb{E}^*[U_{t+\Delta t}(W_{t+\Delta t})|W_t] \\
&\quad - \lambda \cdot \left(\mathbb{E}^*[V_{t+\Delta t}(W_{t+\Delta t})|W_t] - (\mathbb{E}^*[U_{t+\Delta t}(W_{t+\Delta t})|W_t])^2 \right) \\
&= \max_{x_t} [\mathbb{E}[U_{t+\Delta t}(W_{t+\Delta t})|W_t] \\
&\quad - \lambda \cdot \left(\mathbb{E}[V_{t+\Delta t}(W_{t+\Delta t})|W_t] - (\mathbb{E}[U_{t+\Delta t}(W_{t+\Delta t})|W_t])^2 \right)].
\end{aligned}
$$

Assume that we have obtained the value functions $U_{t+\Delta t}(W_{t+\Delta t})$ and $V_{t+\Delta t}(W_{t+\Delta t})$, then the optimization at time t can be regarded as a static optimization problem. By combining the optimal control x_t^* and the future value functions $U_{t+\Delta t}(W_{t+\Delta t})$ and $V_{t+\Delta t}(W_{t+\Delta t})$, we can calculate value functions $U_t(W_t)$ and $V_t(W_t)$ and proceed with the backward recursive programming iteration.

Backward programming algorithm: The basic idea of backward recursive programming is to turn a globally non-smooth optimization problem into several locally smooth optimization problems. In Cong and Oosterlee (2016b), the pre-commitment mean-variance optimization problem was solved with a similar algorithm. The main difference here is that we have two value functions $U_t(W_t)$ and $V_t(W_t)$ in the backward programming iteration instead of a single value function. The algorithm for solving the time-consistent mean-variance problem is defined as follows:

- *Step* 1. *Initiation*:
  Generate an initial guess, for example using the multi-stage strategy, of optimal asset allocations $\{\tilde{x}_t\}_{t=0}^{T-\Delta t}$ and simulate paths of optimal wealth values $\{W_t(i)\}_{i=1}^{N}, t=0, ..., T$. At terminal time T, we have the determined functions $U_T(W_T)$ and $V_T(W_T)$.
  The following three steps are subsequently performed, recursively, backward in time, at $t = T - \Delta t, ..., \Delta t, 0$.
- *Step* 2. *Solution*:
  Bundle paths into B partitions. In practice, this can be performed by first reordering the paths based on their associated wealth values and then partitioning the reordered paths into bundles such that each bundle contains a similar number of paths and the paths inside a bundle have similar wealth levels at time t.

  Denote the wealth values associated to the paths in the bundle by $\left\{W_t^b(i)\right\}_{i=1}^{N_B}$, with N_B being the number of paths in the bundle. Within each bundle, we perform the following steps:
  ○ For the paths in the bundle, we have the wealth values $\left\{W_{t+\Delta t}^b(i)\right\}_{i=1}^{N_B}$ and continuation values $\left\{U_{t+\Delta t}^b(i)\right\}_{i=1}^{N_B}$ and $\left\{V_{t+\Delta t}^b(i)\right\}_{i=1}^{N_B}$ at time $t+\Delta t$. Functions $f_{t+\Delta t}^b(\cdot)$ and $g_{t+\Delta t}^b(\cdot)$, that locally satisfy $U_{t+\Delta t}^b = f_{t+\Delta t}^b\left(W_{t+\Delta t}^b\right)$ and $V_{t+\Delta t}^b = g_{t+\Delta t}^b\left(W_{t+\Delta t}^b\right)$, respectively, can be determined by *regression*.[4]

  ○ For all paths in the bundle, with $f_{t+\Delta t}^b\left(W_{t+\Delta t}^b\right)$ and $g_{t+\Delta t}^b\left(W_{t+\Delta t}^b\right)$ determined, we solve the optimization problem

  $$
  \max_{x_t} \left[\mathbb{E}\left[f_{t+\Delta t}^b(W_{t+\Delta t})|W_t\right] - \lambda \cdot \left(\mathbb{E}\left[g_{t+\Delta t}^b(W_{t+\Delta t})|W_t\right] - \left(\mathbb{E}\left[f_{t+\Delta t}^b(W_{t+\Delta t})|W_t\right] \right)^2 \right) \right]
  $$

  by solving the first-order conditions. In this way, we get new asset allocations $\left\{\hat{x}_t^b(i)\right\}_{i=1}^{N_B}$.

  ○ Since the wealth values $\left\{W_t^b(i)\right\}_{i=1}^{N_B}$ and allocations $\left\{\hat{x}_t^b(i)\right\}_{i=1}^{N_B}$ are now known, by *regression* we can also compute new continuation values $\left\{\hat{U}_t^b(i)\right\}_{i=1}^{N_B}$ and $\left\{\hat{V}_t^b(i)\right\}_{i=1}^{N_B}$,
  $$
  \hat{U}_t^b(i) := \mathbb{E}\left[U_{t+\Delta t}(W_{t+\Delta t})|W_t = W_t^b(i), x_t = \hat{x}_t^b(i)\right],
  $$
  $$
  \hat{V}_t^b(i) := \mathbb{E}\left[V_{t+\Delta t}(W_{t+\Delta t})|W_t = W_t^b(i), x_t = \hat{x}_t^b(i)\right].
  $$

  By combining them, we obtain the value function $\left\{\hat{J}_t^b(i)\right\}_{i=1}^{N_B}$.
- *Step* 3. *Update*:
  For the paths in a bundle, based on a previous approximation, $\left\{\tilde{x}_t^b\right\}_{i=1}^{N_B}$, for the asset allocations, by *regression* we can

---

[4] Regression refers to the technique of approximating the target function by a truncated basis function expansion, where the expansion coefficients are determined by minimizing the approximation error in least squares sense.

calculate the corresponding continuation values $\left\{\tilde{J}_t^b(i)\right\}_{i=1}^{N_B}$. If, for the i-th path, $\tilde{J}_t^b(i) < \hat{J}_t^b(i)$, we choose $\hat{x}_t^b(i)$ as the updated

allocation. Otherwise, we retain the previous allocation. We denote the updated allocations by $\left\{x_t^b(i)\right\}_{i=1}^{N_B}$.

- *Step* 4. *Evolve*:

Once the updated allocations $\{x_t^b(i)\}_{i=1}^{N_B}$ are obtained, again by *regression* we calculate the "updated" $\left\{U_t^b(i)\right\}_{i=1}^{N_B}$ and

$\left\{V_t^b(i)\right\}_{i=1}^{N_B}$ values, and the backward recursion proceeds.

After one backward recursive calculation, we have updated the asset allocations at each time step. To further improve the asset allocations, we can apply the backward procedure iteratively. After several iterations, we will obtain highly satisfactory results.

The purpose of performing bundling is to reduce the bias caused by solving the first-order conditions for a non-smooth optimization problem (due to the constraints). In the unconstrained case, starting with any forward solution, we will end up with a satisfactory result after one backward iteration. Therefore, in this scenario, bundling is not required. When the value functions are highly non-smooth, a large number of bundles is required. According to our experience, 20 bundles are in general sufficient for a variety of problems.

In the algorithm, at each time step and inside each bundle, four regression steps are performed. Three of them are employed for calculating the value function. Since the value function is used to evolve information between time steps, an approximation error will accumulate due to recursion. In general, we can deal with this problem by using a large number of simulations, which is however expensive. In our numerical approach we use the so-called "regress-later technique", as applied in Jain and Oosterlee (2015) and Cong and Oosterlee (2016a, 2016b).

When we perform the (local) regression, polynomials up to order two are considered as basis functions, that is, in the local bundle, $U_t(W_t)$ and $V_t(W_t)$ are approximated by:

$$U_t(W_t) \approx f_t^b(W_t) = \alpha_t^b(2)W_t^2 + \alpha_t^b(1)W_t + \alpha_t^b(0),$$
$$V_t(W_t) \approx g_t^b(W_t) = \beta_t^b(2)W_t^2 + \beta_t^b(1)W_t + \beta_t^b(0).$$

Here $\left\{\alpha_t^b(i)\right\}_{i=0}^2$ and $\left\{\beta_t^b(i)\right\}_{i=0}^2$ denote the regression parameters inside the bundle.

We choose polynomials up to order two as the basis functions for two reasons. The first is that regression with respect to polynomials up to order two should be sufficient for approximating value functions in small-sized, local domains. The other reason is due to the use of the "regress-later technique" in our algorithm. When we implement this technique, one condition is that the conditional expectations of the basis functions should be known analytically, see Jain and Oosterlee (2015) and Cong and Oosterlee (2015). If we choose other basis functions, the conditional expectations may not be directly available and therefore the accuracy of the "regress-later technique" may be affected negatively.

Since the budget constraint reads $W_t = W_{t-\Delta t} \cdot (x_{t-\Delta t} R_{t-\Delta t}^e + R_f)$ and we have $\mathbb{E}[U_t(W_t)|W_{t-\Delta t}]^2$ in the formulation of $J_{t-\Delta t}(W_{t-\Delta t})$, value function $J_{t-\Delta t}(W_{t-\Delta t})$ in the local domain will become a fourth-order function of $x_{t-\Delta t}$. To solve the optimality with respect to $x_{t-\Delta t}$, we implement the Newton–Raphson method with the solution from the previous iteration as the initial guess. In our numerical tests, the Newton–Raphson method always achieves a convergent solution within three iterations.

Remark 4.3. In the unconstrained case, it is sufficient to parameterize $U_t(W_t)$ by first-order polynomials. However, according to our experiments, in the constrained case approximating $U_t(W_t)$ by second-order polynomials yields more stable and satisfactory results.

Remark 4.4. The forward and backward numerical iterations described in Sections 4.1 and 4.2 can also be used for the time-consistent problem with a state-dependent risk aversion, as in Remark 2.4. The only difference is that, instead of keeping a fixed risk aversion parameter, we choose a state-dependent risk aversion parameter for the value functions and calculate the corresponding optimal strategy. Since this state-dependent time-consistent strategy is only valid when the wealth is positive, which implies a constrained scenario, the forward strategy can only generate a rough approximation for the asset allocations. However, after using the backward recursive updating iterations, we obtain satisfactory results.

5. Numerical tests

In this section, we utilize our algorithm for solving time-consistent multi-period mean-variance portfolio management problems. We will first check the convergence of the backward recursive programming step in our algorithm. Then, our numerical algorithm will be tested on time-consistent problems with different constraints. Following that, we will perform a test with the hybrid strategy proposed in Section 3.3, analyzing the choice of targets. In the last test, the performance of different mean-variance portfolio management strategies will be compared. In the numerical section, we use the following short notations: "TC" for a time-consistent strategy, "PC" for a pre-commitment strategy, "hybrid" for the hybrid strategy proposed in Section 3.3 and "TCSD" for a time-consistent strategy with state-dependent risk aversion.

Table 1
Parameter setting (from Wang and Forsyth (2011)) used in the tests.

Risk-free rate, r_f	0.03
Market price of risk, ξ	0.33
Volatility, σ	0.15
Periodic contribution, C	0.1
Investment duration, T (years)	20
Rebalancing opportunities, M	40
Initial wealth, W_0	1

Table 2
The convergence of the backward recursive iterations. The multi-stage strategy is chosen as the initial solution.

Risk aversion	$\lambda = 0.05$		$\lambda = 0.25$	
	$E_0^*(W_T)$ (s.e.)	$\text{Std}_0^*(W_T)$ (s.e.)	$E_0^*(W_T)$ (s.e.)	$\text{Std}_0^*(W_T)$ (s.e.)
Forward iteration	13.17 (0.04)	9.60 (0.04)	8.49 (0.01)	2.87 (0.01)
After one backward iteration	12.89 (0.04)	9.03 (0.04)	8.30 (0.01)	2.76 (0.01)
After two backward iterations	12.87 (0.04)	8.97 (0.04)	8.28 (0.01)	2.75 (0.01)
After three backward iterations	12.87 (0.04)	8.97 (0.04)	8.28 (0.01)	2.75 (0.01)

5.1. Setup of numerical tests

We perform the tests in the case of one risky asset and one risk-free asset in the portfolio. We choose geometric Brownian motion as the dynamics of the risky asset. We assume that the log-returns of the risky asset are governed by volatility σ and mean $r_f + \xi \cdot \sigma$. Here r_f is the log-return of the risk-free asset and ξ the market price of risk.

In our numerical tests, we choose the sample size to be 50,000 and the number of bundles to be 20 for the backward recursive iteration. In case the forward solutions are not optimal, we always present the solutions obtained after three backward recursive iterations. We use a common random seed to generate Monte-Carlo paths for one run of the algorithm (including one forward and several backward processes). To ensure that the choice of random seed does not introduce a bias, we consider 20 different random seeds in all tests.

When we take the "no bankruptcy" constraint, Eq. (5), into account, we choose parameter Ω to be sufficiently small, $\Omega = 10^{-8}$, which ensures that the undesired event will happen with a very low probability. When we consider the "bounded leverage" constraint, we always take $[x_{\min}, x_{\max}] = [0, 1.5]$.

The model parameters used in our tests are shown in Table 1. With this setup for the tests, both the forward approximation step and one backward iteration in our algorithm can be performed within a few seconds.

Remark 5.1. We also performed tests using the Merton jump-diffusion model (Merton, 1976) with rare and negative jumps as the asset dynamics. However, when we compare the performance of different strategies, we do not get different conclusions from using geometric Brownian motion. Therefore, we do not report the results of testing the jump problem in this paper. For more details about how the efficient frontier may change when jumps come in, we refer the readers to Dang and Forsyth (2014, 2016).

5.2. Numerical results

Convergence of backward recursive iteration: We first analyze numerically the convergence of the backward recursive iterations. We consider two test cases with bounded leverage constraints on the asset allocations. In the first case, we choose the multi-stage strategy as the initial solution and perform a backward recursive iteration. Although the multi-stage strategy does not generate the optimal solution, it should constitute a reasonable initial guess to start the backward iteration. In the second case, we construct the initial solution by constant asset allocations. In general, this kind of initial solution is questionable and the backward iteration process is essential for accurate constrained solutions.

According to Table 2, the solution obtained after the backward iteration is different from the forward solution. The most significant improvement is achieved by the first backward iteration. After the second backward iteration, the convergent solution is obtained.

Fig. 2 shows that, starting with inferior (constant) initial guesses, we can reach convergent results after three backward iterations. This is an indication that the backward iteration is highly efficient and robust in this test case.
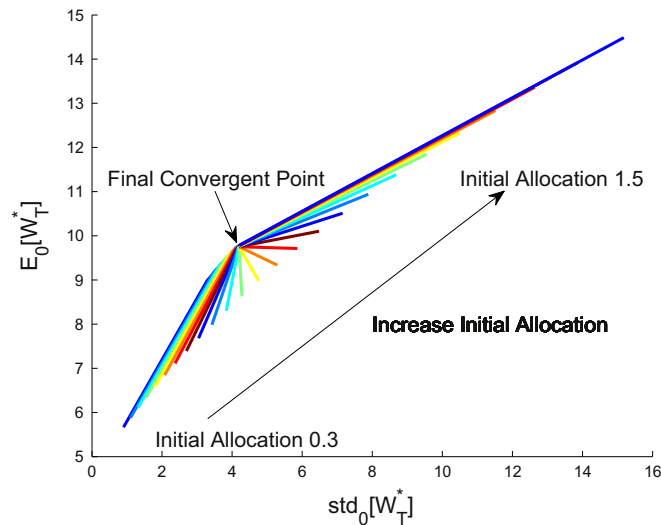
Fig. 2. Risk aversion $\lambda = 0.15$, constant asset allocations, varying from 0.3 to 1.5, as the (inaccurate) initial solution. After three backward recursive iterations, convergent results are obtained.
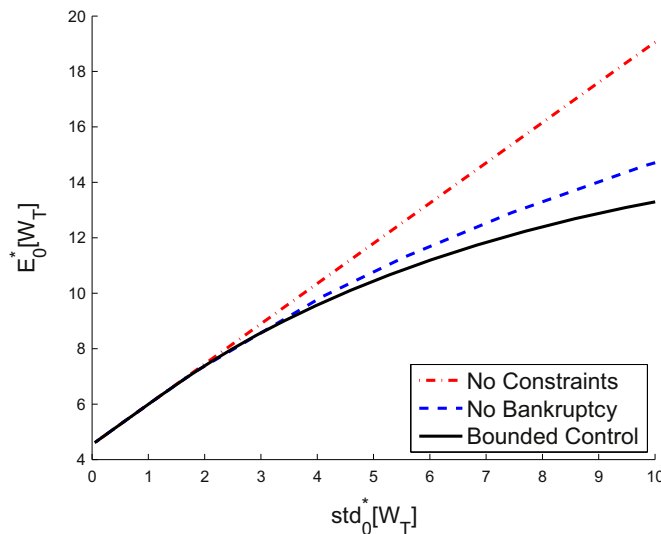

Fig. 3. Casting different constraints on the time-consistent strategy.

Time consistent strategy with constraints: We consider the time-consistent strategy with different constraints. The efficient frontiers are shown in Fig. 3. In the case of a bounded leverage constraint, we have as a reference value $[\text{std}_0^*[W_T], \mathbb{E}_0^*[W_T]] = [8.175, 12.661]$ (Wang and Forsyth, 2011) for the continuous time-consistent re-balancing problem. With our algorithm, if we set the re-balancing opportunities to $M = 240$, we obtain the mean-variance pair $[\text{std}_0^*[W_T], \mathbb{E}_0^*[W_T]] = [8.175, 12.582]$. Our algorithm for the time-consistent problem thus performs highly satisfactory.

Hybrid strategy with different investment targets: As mentioned in Section 3.3, the choice of investment target has a significant impact on the efficient frontiers generated by the hybrid strategy. In this test we consider different choices of the value k in Eq. (14), generating targets in the hybrid strategy.

As shown in Fig. 4, when $k \approx 1$, the corresponding efficient frontiers generated by the hybrid strategy are close to the one generated by the pre-commitment strategy for an optimization problem with bounded leverage constraint. When k increases, the efficient frontiers generated by the hybrid strategy resemble the one generated by the plain time-consistent strategy. In the following numerical tests, when we compare the hybrid strategy to other mean-variance strategies, we will form the target with $k = 1$.

Comparing different mean-variance strategies: First, we compare the mean-variance efficient frontiers for different strategies. We find that in all scenarios the pre-commitment strategy generates the highest efficient frontier and the time-consistent strategy with state-dependent risk aversion as in Remark 2.4 gives the lowest frontier. The hybrid strategy generates similar efficient frontiers to the pre-commitment strategy, especially when constraints are imposed. Even though
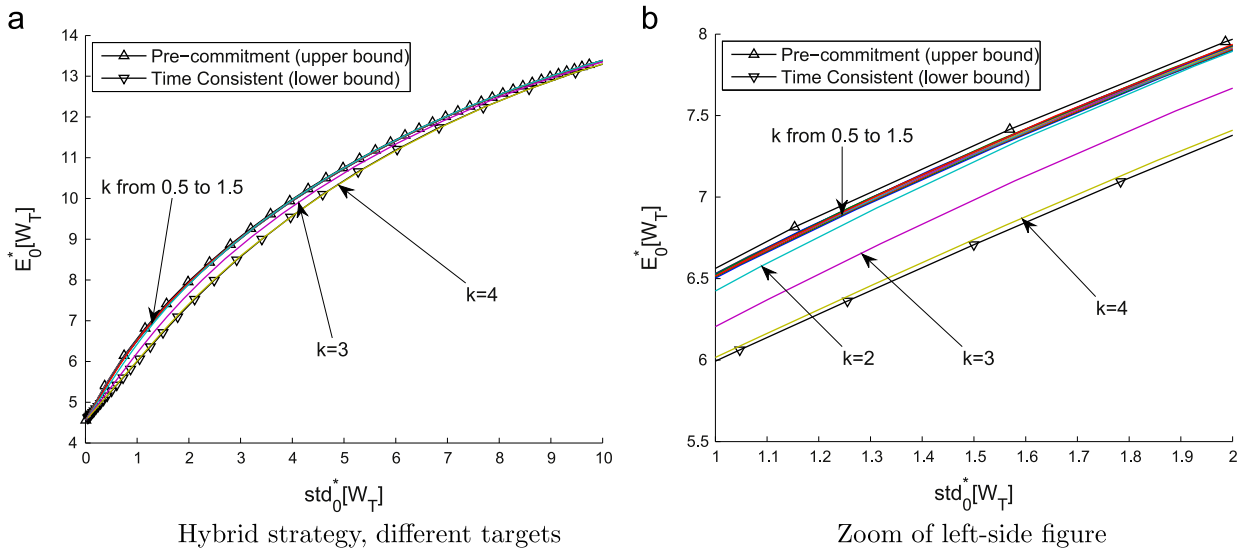
Fig. 4. The influence of choosing different targets for the hybrid strategy. Bounded leverage constraint $[x_{min}, x_{max}] = [0, 1.5]$ is imposed on the asset allocations.
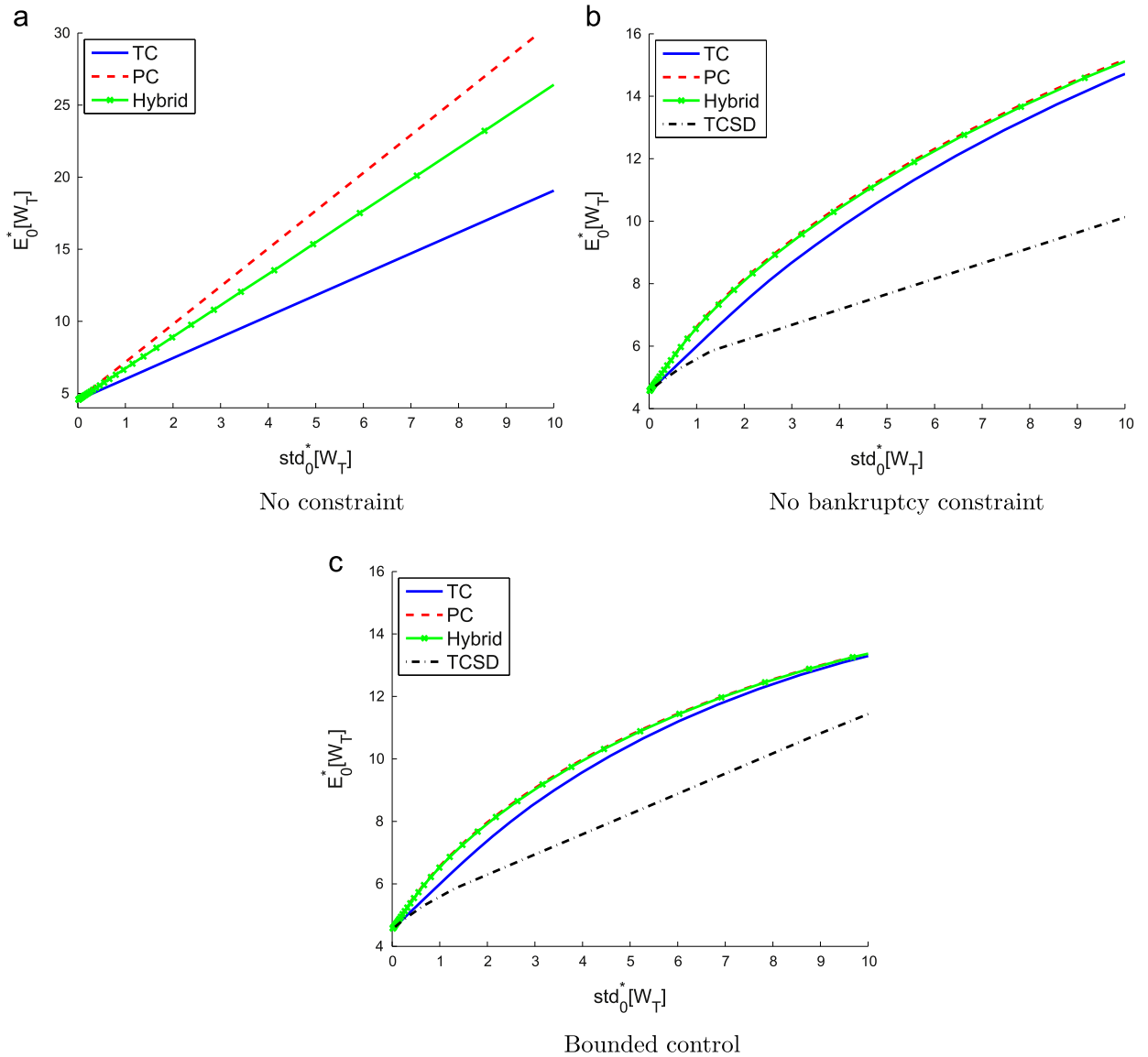

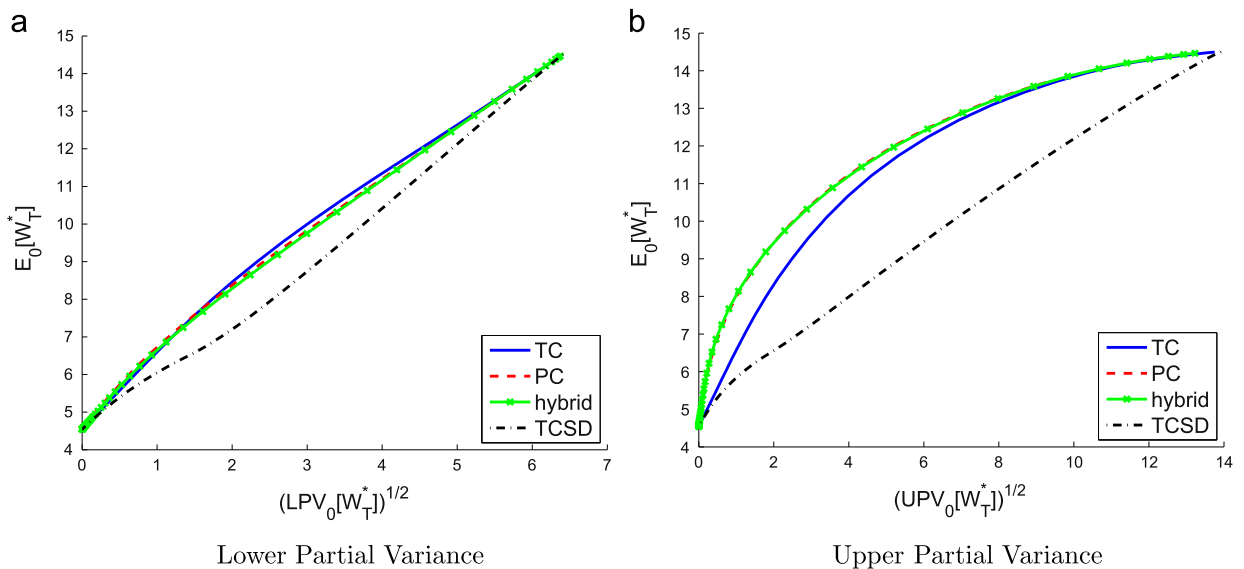Fig. 5. Comparing the mean-variance frontiers of different strategies.

Fig. 6. For different strategies, we compare the Lower Partial Variance and the Upper Partial Variance. Bounded leverage constraint $[x_{min}, x_{max}] = [0, 1.5]$ is imposed on the asset allocations.

the time-consistent strategy generates lower efficient frontiers, the difference between the time-consistent and the pre-commitment frontiers is not significant (Fig. 5).

Fig. 6 shows that if we consider the mean-LPV frontier, it is difficult to say whether the pre-commitment strategy is superior to the time-consistent strategy. The time-consistent strategy usually generates a lower mean-variance efficient frontier than the pre-commitment strategy, but the main reason is that it yields a lower mean-UPV frontier. The UPV indicates the potential for an investor to achieve more wealth than the expectation, so a high UPV is not a problem. The time-consistent strategy with state-dependent risk aversion has the lowest mean-LPV frontier.

By examining how the risk is distributed, we find that the time-consistent strategy may be a better choice than the pre-commitment strategy for multi-period mean-variance optimization. In order to achieve the same level of mean return, the time-consistent strategy is not more risky than the pre-commitment strategy. On the other hand, the time-consistent strategy is more likely to generate a higher amount of wealth than expected and thus exhibits a higher potential than the pre-commitment strategy.

6. Conclusion

In this paper, we focused on the time-consistent mean-variance optimal asset allocation problem, and compared the results to those obtained by the so-called pre-commitment strategy.

In our analysis, we found that in the unconstrained case the risk aversion parameter of a time-consistent investor and the optimal asset allocation form a one-to-one function. At a given time step, the optimal asset allocation is uniquely determined by the risk aversion attitude. In Cong and Oosterlee (2016b) we have reported similarly that, at a given time step, the optimal asset allocation of a pre-commitment investor is uniquely determined by the investment target. Based on this, we established a link between the time-consistent and the pre-commitment strategies. We assumed that at a given time step the time-consistent investor had the same asset allocation as the pre-commitment investor and a one-to-one relation between the risk aversion attitude of the time-consistent investor and the investment target of the pre-commitment investor was obtained. For a time-consistent investor, we called this function the "implied investment target".

We find that the (implied) investment target of a time-consistent investor and that of a pre-commitment investor vary in different ways through the dynamic management process. The investment target of a time-consistent investor varies over time, while a pre-commitment investor always has a fixed target. We defined a hybrid strategy by introducing a fixed target into the time-consistent strategy. In our test cases, when we chose the "implied investment target" at the initial time step as the target, the hybrid strategy generates a better mean-variance efficient frontier than the time-consistent strategy. In the constrained case, the efficient frontier generated by this hybrid strategy is very similar to that generated by the pre-commitment strategy.

We also extended the numerical simulation-based algorithm, which was proposed in Cong and Oosterlee (2016b) for the pre-commitment mean-variance problem, to solving time-consistent mean-variance problems. With some modifications of the algorithm, we could utilize it to achieve highly satisfactory results for the time-consistent problem as well as for the hybrid problem. Our proposed algorithm appears to be robust and very efficient.

"Focusing on a fixed target" often gives completely different distributions of the terminal wealth from time-consistent and pre-commitment strategies. The time-consistent strategy often yields a lower mean-variance frontier, but the main reason is that the time-consistent strategy has a lower mean-UPV frontier (which does not indicate more risk). If we consider the mean-LPV frontier, the time-consistent, the pre-commitment and the hybrid strategies exhibit similar performance. Reflecting on this, we conclude that the time-consistent strategy is not always inferior to the pre-commitment strategy. Requiring time-consistency on a mean-variance strategy is reasonable from the perspective of real-life applications. Meanwhile, it does not introduce additional downside risk to the strategy.

Acknowledgments

References

Basak, S., Chabakauri, G., 2010. Dynamic mean-variance asset allocation. Rev. Financ. Stud. hhq028.

Bellman, R., 1957. Dynamic Programming. Princeton Univ Press, Princeton, NJ, USA.

Bielecki, T.R., Jin, H., Pliska, S.R., Zhou, X., 2005. Continuous-time mean-variance portfolio selection with bankruptcy prohibition. Math. Financ. 15 (2), 213–244.

Björk, T., Murgoci, A., Zhou, X., 2014. Mean-variance portfolio optimization with state-dependent risk aversion. Math. Financ. 24 (1), 1–24.

Chen, Z.-p., Li, G., Guo, J.-e., 2013. Optimal investment policy in the time consistent mean-variance formulation. Insur.: Math. Econ. 52 (2), 145–156.

Cong, F., Oosterlee, C.W., 2015. Pricing bermudan options under Merton jump-diffusion asset dynamics. Int. J. Comput. Math. 92 (12), 2406–2432.

Cong, F., Oosterlee, C.W., 2016a. Accurate and robust numerical methods for the dynamic portfolio management problem. Comput. Econ., 1–26.

Cong, F., Oosterlee, C.W., 2016b. Multi-period mean-variance portfolio optimization based on monte-carlo simulation. J. Econ. Dyn. Control 64 (1), 23–38.

Cui, X., Gao, J., Li, X., Li, D., 2014a. Optimal multi-period mean-variance policy under no-shorting constraint. Eur. J. Oper. Res. 234 (2), 459–468.

Cui, X., Li, X., Li, D., Shi, Y., 2014b. Time consistent behavior portfolio policy for dynamic mean-variance formulation. Available at SSRN 2480299, http://dx.doi.org/10.2139/ssrn.2410632.

Cui, X., Li, D., Wang, S., Zhu, S., 2012. Better than dynamic mean-variance: time inconsistency and free cash flow stream. Math. Financ. 22 (2), 346–378.

Cui, X., Xu, L., Zeng, Y., 2015. Continuous time mean-variance portfolio optimization with piecewise state-dependent risk aversion. Optim. Lett., 1–11.

Czichowsky, C., 2013. Time-consistent mean-variance portfolio selection in discrete and continuous time. Financ. Stoch. 17 (2), 227–271.

Dang, D.-M., Forsyth, P.A., 2014. Continuous time mean-variance optimal portfolio allocation under jump diffusion: an numerical impulse control approach. Numer. Methods Partial Differ. Equ. 30 (2), 664–698.

Dang, D.-M., Forsyth, P.A., 2016. Better than pre-commitment mean-variance portfolio allocation strategies: a semi-self-financing Hamilton–Jacobi–Bellman equation approach. Eur. J. Oper. Res. 250 (3), 827–841.

Dang, D.-M., Forsyth, P.A., Vetzal, K., 2015. The 4% strategy revisited: a pre-commitment optimal mean-variance approach in wealth management. Quant. Financ. 1-17, Available at SSRN 2571530, http://dx.doi.org/10.1080/14697688.2016.1205211.

Fu, C., Lari-Lavassani, A., Li, X., 2010. Dynamic mean-variance portfolio selection with borrowing constraint. Eur. J. Oper. Res. 200 (1), 312–319.

Gerrard, R., Haberman, S., Vigna, E., 2004. Optimal investment choices post-retirement in a defined contribution pension scheme. Insur.: Math. Econ. 35 (2), 321–342.

Haberman, S., Vigna, E., 2002. Optimal investment strategies and risk measures in defined contribution pension schemes. Insur.: Math. Econ. 31 (1), 35–69.

Hu, Y., Jin, H., Zhou, X.Y., 2012. Time-inconsistent stochastic linear-quadratic control. SIAM J. Control Optim. 50 (3), 1548–1572.

Jain, S., Oosterlee, C.W., 2015. The stochastic grid bundling method: efficient pricing of Bermudan options and their greeks. Appl. Math. Comput. 269 (1), 412–431.

Li, D., Ng, W.-L., 2000. Optimal dynamic portfolio selection: multiperiod mean-variance formulation. Math. Financ. 10 (3), 387–406.

Li, X., Zhou, X., Lim, A.E., 2002. Dynamic mean-variance portfolio selection with no-shorting constraints. SIAM J. Control Optim. 40 (5), 1540–1555.

Markowitz, H.M., 1952. Portfolio selection. J. Financ. 7 (1), 77–91.

Merton, R.C., 1976. Option pricing when underlying stock returns are discontinuous. J. Financ. Econ. 3 (1), 125–144.

Vigna, E., 2014. On efficiency of mean-variance based portfolio selection in defined contribution pension schemes. Quant. Financ. 14 (2), 237–258.

Wang, J., Forsyth, P.A., 2010. Numerical solution of the Hamilton–Jacobi–Bellman formulation for continuous time mean variance asset allocation. J. Econ. Dyn. Control 34 (2), 207–230.

Wang, J., Forsyth, P.A., 2011. Continuous time mean variance asset allocation: a time-consistent strategy. Eur. J. Oper. Res. 209 (2), 184–201.

Wang, J., Forsyth, P.A., 2012. Comparison of mean variance like strategies for optimal asset allocation problems. Int. J. Theor. Appl. Financ. 15 (2).

Zhou, X., Li, D., 2000. Continuous-time mean-variance portfolio selection: a stochastic LQ framework. Appl. Math. Optim. 42 (1), 19–33.

Zhu, S.-S., Li, D., Wang, S.-Y., 2004. Risk control over bankruptcy in dynamic portfolio selection: a generalized mean-variance formulation. IEEE Trans. Autom. Control 49 (3), 447–457.