Optimal Investment with S-shaped Utility and Trading and Value at Risk Constraints: an Application to Defined Contribution Pension Plan

Yinghui Dong* and Harry Zheng†

Abstract

In this paper we investigate an optimal investment problem under loss aversion (S-shaped utility) and with trading and Value-at-Risk (VaR) constraints faced by a defined contribution (DC) pension fund manager. We apply the concavification and dual control method to solve the problem and derive the closed-form representation of the optimal terminal wealth in terms of a controlled dual state variable. We propose a simple and effective algorithm for computing the initial dual state value, the Lagrange multiplier and the optimal terminal wealth. Theoretical and numerical results show that the VaR constraint can significantly impact the distribution of the optimal terminal wealth and may greatly reduce the risk of losses in bad economic states due to loss aversion.

2010 MSC: 90C46, 49L20

Keywords: Control, S-shaped utility, trading constraint, value-at-risk constraint, defined contribution pension plan

1 Introduction

Most literature on investment problems focus on maximizing the expectation of a smooth utility of terminal wealth. Loss aversion, first proposed by Kahneman and Tversky (1979) within the framework of prospect theory (PT), is defined over gains and losses in wealth relative to a predefined reference point, rather than in terms of changes in the absolute level of total wealth itself. Every investor has a reference point that defines relative 'losses' and 'gains'. Tversky and Kahneman (1992) also demonstrate the loss aversion and risk seeking behaviour by an asymmetric S-shaped utility function, convex in the domain of losses and concave in the domain of gains. The martingale method is mainly used to solve the S-shaped utility maximization problem when the market is complete. For example, Berkelaar et al. (2004) derive the optimal investment strategies with two utility functions under loss aversion in a continuous-time case. He and Kou (2018) investigate the S-shaped utility maximization under a minimum guarantee. Dong and Zheng (2019) include

*Department of Mathematics and Physics, Suzhou University of Science and Technology, Suzhou 215009, P. R. China. Email: dongyinghui1030@163.com

†Department of Mathematics, Imperial College, London SW7 2AZ, UK. Email: h.zheng@imperial.ac.uk

both short-selling and portfolio insurance (PI) constraints in the model and apply the dual control method to solve the corresponding constrained optimization problem.

Value-at-Risk (VaR), defined as the worst expected loss given a pre-set confidence level, is a quantile measure that controls the tail risk of the terminal wealth. Due to its prominence in current regulatory frameworks for banks (Basel II) as well as for insurance companies (Solvency II), VaR plays an important role in pensions, insurance companies and other financial institutions. A VaR measurement is easy to interpret as it is based on the examination of the percentiles of the distribution, summarising the downside risk of an institution. Furthermore, VaR can be used to facilitate risk aggregation for quantifying all types of risks. VaR-based risk management (VaR-RM) has a convenient property that it nests the PI-based risk management (PI-RM). It is commonly used by practitioners and regulators due to these advantages, although it ignores the magnitude of losses when they do occur. The problem of concave utility (mainly CRRA utility) maximization under a VaR constraint has been studied in the literature. For example, Basak and Shapiro (2001) derive the optimal investment strategies to manage VaR risk by a martingale method. Boyle and Tian (2007) generalize the VaR constraint to the case where the wealth must exceed a stochastic, but hedgeable, benchmark with a given probability. Kraft and Steffensen (2013) impose intertemporal VaR constraints and characterize the optimal terminal wealth by a linear combination of European call and put options. Recently, nonconcave utility maximization without VaR constraints has been considered by many researchers, see Bernard et al. (2015), El Karoui et al. (2005), Bichuch and Sturm (2014), Guan and Liang (2016), Chen et al. (2017), Chen et al. (2019) and He and Kou (2018).

The aforementioned references all assume that the market is complete, which is equivalent to the existence of a unique pricing kernel. The martingale approach is commonly used to solve optimal investment problems as one may first find the optimal terminal wealth by solving a simplified static optimization problem and then find the replicating feasible trading strategy (the optimal control) with the martingale representation theorem, see, for example, Pliska (1986), Cox and Huang (1989), and Karatzas et al. (1986). Cvitanic and Karatzas (1992) and Xu and Shreve (1992) generalize the martingale method to facilitate portfolio constraints. He and Zhou (2011) investigate a concave utility maximization problem in an incomplete market model with infinitely many pricing kernels and show that there exists a unique pricing kernel, called the minimal pricing kernel, in the presence of closed convex cone control constraints and one may use it to solve the constrained problem in the same way as in the complete market case. The optional decomposition theorem (see Follmer and Kramkov (1997)) ensures the existence of a feasible replicating trading strategy. To actually find the replicating trading strategy, one may first find the optimal wealth process by computing the conditional expectation of the optimal terminal wealth with the minimal pricing kernel and then derive the stochastic differential equation (SDE) for the optimal wealth process with Ito's formula. The key step in this procedure is to find explicitly the optimal wealth process, which is in general difficult when the utility function is not a simple strictly concave function (see Chen et al. (2019)), let alone any additional constraints on the terminal wealth.

In this paper, we investigate an S-shaped utility maximization problem under trading and VaR constraints. The wealth process is not self-financing and has additional cash income. The model is well placed for studying the optimal investment of defined contribution (DC) pension schemes which are popular and important in the pension systems of many countries. In a DC pension plan,

the member contributes part of the salary to the plan and bears the financial risk. The retirement benefit is mainly determined by the performance of its fund portfolios before retirement. It is therefore essential for pension fund managers to find optimal investment strategies for the members during the accumulation phase to build sufficient funds on retirement. The optimal DC pension investment has attracted extensive research, see Boulier et al. (2001), Cairns et al. (2006), Zhang et al. (2007), Zhang and Ewald (2010), Yao et al. (2013), Zeng et al. (2018) and Blake et al. (2013, 2014). This paper sheds some new light on the optimal pension plan investment in the presence of S-shaped utility and VaR and control constraints and covers the constrained portfolio choice problem of self-financing wealth process as a special case with zero cash income. This paper is also more complex in mathematics than Dong and Zheng (2019) in the sense that the latter only involves solving the binding budget constraint equation whereas the former needs to tackle the combination of the budget constraint and the VaR constraint simultaneously.

It is a highly nontrivial task to solve a VaR-constrained optimization problem. We circumvent the difficulty by introducing a Lagrange multiplier to reflect the bindingness of VaR constraint, which leads to a utility maximization problem with a non-concave discontinuous utility function. We apply the concavification and the dual control method, together with the pathwise differentiation and likelihood ratio method, to solve the resulting VaR constraint-free maximization problem with a fixed Lagrange multiplier and give the explicit characterizations of the optimal portfolio and wealth processes in terms of the optimal dual control and state processes. To find the correct Lagrange multiplier that makes the optimal solution of the unconstrained problem the same one for the constrained problem, we need to solve two fully coupled nonlinear equations (binding budget and VaR constraints) with Lagrange multiplier and initial dual state value as variables, again a highly nontrivial problem. We overcome the difficulty by finding the explicit relation of the Lagrange multiplier and the initial dual value from the binding VaR constraint and then solve the binding budget constraint equation to find the unique initial dual state value. This methodology may open a way for solving S-shaped utility maximization with multiple VaR and other constraints, which may be more effective in risk management. The key to our success is the dual control method which is effective in solving portfolio optimization problems with control constraints as it relates the original stochastic optimal control problem to a dual problem which may be relatively easier to solve, see Xu and Shreve (1992) and Bian et al. (2011). Unlike He and Zhou (2011), we derive the explicit representation of the optimal wealth process and investment strategy in terms of the dual value function, its derivatives, and the optimal dual state process, which makes possible either to find their closed form expressions or to compute their numerical values with the Monte Carlo simulation method. It is virtually impossible to solve the problem in this paper directly in its primal form. We propose a simple and effective numerical algorithm for finding the initial dual state value, the Lagrange multiplier and the optimal terminal wealth.

Our theoretical and numerical results show that the loss aversion as well as trading and VaR constraints have significant impact on the optimization problem. It is well known that, a VaR constraint leads to heavier losses in bad market scenarios than in the case of no VaR constraint under a smooth concave utility. Loss aversion leads the manager to be risk averse in the gain domain and risk-seeking in the loss domain. We find that introducing a VaR constraint under an S-shaped utility may not lead to more losses than in the case of no VaR constraint. On the contrary, a VaR constraint may strictly improve the risk management for bad economic states. It is

crucially important to choose a reasonable confidence level and protection level in a VaR constraint. If the confidence level is too low or the protection level is too high, then more states need to be insured against and the optimal terminal wealth becomes less volatile which results in a relatively low expected terminal wealth. If the confidence level is too high or the protection level is too low, then a VaR constraint would not well protect the members' benefits.

The main contribution of this paper is that we extend the application of the dual control method for a continuous concave utility in Bian et al. (2011) to a discontinuous nonconcave utility (due to VaR constraint and S-shaped utility), solve the optimal investment problem with S-shaped utility and trading and VaR constraints, prove the existence and uniqueness of the optimal solution, characterize explicitly the optimal wealth and the Lagrange multiplier, and identify the joint impact of the S-shaped utility and VaR and control constraints on the distribution of the optimal terminal wealth. To the best of the authors' knowledge, we are the first in the literature to have completely solved the aforementioned constrained optimization problem based on the concavification and dual control method.

The rest of the paper is organized as follows. In Section 2 we formulate a DC pension investment problem with S-shaped utility and trading and VaR constraints, and convert the constrained optimization problem (2.7) into an equivalent unconstrained one (2.8) coupled with the feasibility and complementary slackness condition (2.9). In Section 3 we apply the concavification and dual control method to solve the unconstrained nonconcave discontinuous optimization problem (2.8) and characterize explicitly the optimal solution and optimal control in Propositions 3.1-3.3 and Theorem 3.7. In Section 4 we state the main result of the paper, Theorem 4.1, on the existence and uniqueness of the optimal solution for the constrained optimization problem (2.7) and give a constructive proof which leads to a simple and effective algorithm to compute the initial dual state variable and Lagrange multiplier. In Section 5 we present a numerical example with short-selling constraints and discuss the impact of VaR constraint on the distribution of the optimal terminal wealth. Section 6 concludes. The appendix contains a technical lemma that is used in constructing the concave envelope of a nonconcave discontinuous function, the proofs of Propositions 3.1-3.3 and explicit expressions for the optimal wealth and portfolio processes.

2 The investment problem for a DC pension fund

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered complete probability space with the filtration $\mathbb{F} := \{\mathcal{F}_t | 0 \leq t \leq T\}$ being the natural filtration generated by an n-dimensional standard Brownian motion $W(t) = (W_1(t), \cdots, W_n(t))^\top$, where $W_1(t), \cdots, W_n(t)$ are independent and $a^\top$ is the transpose of a, and satisfying the usual conditions. The pension fund starts at time 0 and the retirement time is T.

Let the financial market consist of $n+1$ traded securities: one riskless savings account $S_0(t)$ and n risky assets $S_i(t), i = 1, \cdots, n$. The riskless savings account evolves as

$$dS_0(t) = rS_0(t)dt, \tag{2.1}$$

where r is a riskless interest rate. The price processes of the n risky assets are modelled by

$$dS(t) = \text{diag}(S(t))(\mu dt + \sigma dW(t)), \tag{2.2}$$

where $S(t) = (S_1(t), \cdots, S_n(t))^\top$, $\text{diag}(S(t))$ is an $n \times n$ matrix with diagonal elements $S_i(t)$ and all other elements 0, $\mu = (\mu_1, \cdots, \mu_n)^\top$ is a constant vector representing the stock growth rate

with $\mu_i > r, i = 1, \cdots, n$, and $\sigma = (\sigma_{ij})$ is an $n \times n$ nonsingular constant matrix representing the volatility and correlation information of $S(t)$.

In recent years, DC pension plans have become increasingly popular in the pension market due to the prevailing trend of longer lifespans and the development of the equity markets. In a DC pension plan, its members contribute proportions of their salaries to the pension plan before retirement time T. Let $c(t) > 0$ denote the aggregated amount of money contributed at time t of a cohort of fund participants. Note that individual members' contribution rates may be random and fluctuate over the time, but for a large pension fund, the accumulated cash income is in general deterministic and stable. Furthermore, in the labor market, the average salary of employees and the contribution rate often steadily increase in the long run. We therefore assume $c(t)$ is a deterministic, nondecreasing function. Assume that there are no transaction costs or taxes in the financial market. The pension account is endowed with an initial endowment $x_0 \geq 0$. The wealth process $X^\pi(t)$ satisfies the following controlled SDE:

$$dX^\pi(t) = (rX^\pi(t) + \pi^\top(t)\sigma\xi)dt + \pi^\top(t)\sigma dW(t) + c(t)dt,\ t \geq 0, \tag{2.3}$$

with initial condition $X^\pi(0) = x_0$, where $\pi(t) = (\pi_1(t), \cdots, \pi_n(t))^\top$ and $\pi_i(t)$ is the amount of wealth invested in the ith risky asset for $i = 1, \cdots, n$, $\xi = \sigma^{-1}(\mu - r\mathbf{1})$ is the market price of risk vector and $\mathbf{1}$ is a vector with all components 1. We next define the set of admissible trading strategies.

Definition 2.1. *Let K be a closed convex cone. A portfolio strategy $\pi = (\pi_1, \cdots, \pi_n)^\top$ is said to be admissible if it is a progressively measurable, $\mathcal{F}$-adapted process which satisfies $E[\int_0^T \|\pi(t)\|^2 dt] < \infty$, $\pi(t) \in K$, a.s., and there exists a unique strong solution $X^\pi(t)$ to (2.3). The set of all admissible portfolio strategies is denoted by $\mathcal{A}$.*

Most literature on DC pension plan focus on maximizing the expectation of a smooth concave utility of terminal wealth. Kahneman and Tversky (1979) claim that people tend to be risk averse to gains and risk seeking to losses and make decisions relative to some reference levels rather than absolute values directly. Based on experiments, Kahneman and Tversky (1979) propose an S-shaped utility function (see (5.1)) to characterize different behaviors of people over gains and losses relative to a reference point. There has been some research on DC pension plan with S-shaped utilities, see Guan and Liang (2016) and Chen et al. (2017). Here we extend the underlying power utilities in the S-shaped utility to a more general convex-concave utility. Since the purpose of a pension plan is to provide adequate income for its members after retirement, the primary goal of a DC pension fund manager is to find the optimal investment strategies under loss aversion and trading and VaR constraints.

Let θ be a reference point, which is chosen in advance. Consider a utility function defined by:

$$U(x) = \begin{cases} -\infty, & x < 0, \\ -U_2(\theta - x), & 0 \leq x < \theta, \\ U_1(x - \theta), & x \geq \theta, \end{cases} \tag{2.4}$$

where U_1 and U_2 are two strictly increasing, strictly concave, continuously differentiable, real-valued functions defined on $[0, \infty)$ satisfying

$$\lim_{x \to +\infty} U_1(x) = +\infty, \lim_{x \to +\infty} U_1'(x) = 0, U_i(0) = 0, \lim_{x \to 0^+} U_i'(x) = +\infty, \tag{2.5}$$

for $i = 1, 2$ and U_1 satisfies the asymptotic elasticity condition

$$\lim_{x \to +\infty} \frac{xU_1'(x)}{U_1(x)} < 1. \tag{2.6}$$

Note that U is convex when x is less than θ (in the domain of losses) and concave when x is greater than θ (in the domain of gains) which gives an S-shaped graph.

Condition (2.5) ensures that the strictly decreasing function U_1' has a strictly decreasing inverse $I_1 : (0, \infty) \to (0, \infty)$, that is,

$$U_1'(I_1(y)) = y, \forall y > 0, \ I_1(U_1'(x)) = x, \forall x > 0.$$

Let $L > 0$ be a given level. In order to provide a downside protection, the pension manager is to find the optimal investment strategy to maximize the expected utility of the wealth at time T under a VaR constraint:

$$\begin{cases} \max\limits_{\pi \in \mathcal{A}} E[U(X^\pi(T))], \\ s.t.\ X^\pi(t) \text{ satisfies (2.3)}, \\ \ P(X^\pi(T) \geq L) \geq 1 - \varepsilon, \end{cases} \tag{2.7}$$

where $0 \leq \varepsilon \leq 1$ is a given constant. The VaR constraint requires that the probability of the terminal wealth above the level L is at least $1 - \varepsilon$. When $\varepsilon = 1$, the constraint is not binding. When $\varepsilon = 0$, (2.7) recovers the case of portfolio insurance, see Basak (1995).

Remark 2.2. *People often make decisions relative to some reference levels. Thus, the choice of the reference point θ is crucial. If θ is set too low, then it may lead to little utility after retirement. To protect the members' benefits, we can set the protection level L higher than θ. If θ is set too high, then the manager is likely to take a great risk to try to attain the reference point due to his risk-seeking behavior in the loss domain. To prevent too much risk of DC pension plan, we can set the protection level L lower than θ.*

We may use the Lagrange multiplier method to solve problem (2.7). Define

$$\tilde{U}_\lambda(x) = U(x) + \lambda 1_{\{x \geq L\}},$$

where $\lambda \geq 0$ is a Lagrange multiplier to be determined. Note that $\tilde{U}_\lambda$ is a nonconcave discontinuous function which jumps upwards by the amount λ at the point $x = L$ and then continues in parallel to U afterwards.

Consider the following VaR constraint-free version of problem (2.7):

$$\begin{cases} \max\limits_{\pi \in \mathcal{A}} E[\tilde{U}_\lambda(X^\pi(T))], \\ s.t.\ X^\pi(t) \text{ satisfies (2.3)}. \end{cases} \tag{2.8}$$

If we can find the optimal solution $X^{\pi^*, \lambda^*}(T)$ of problem (2.8) for some admissible control π^* and nonnegative constant λ^* such that

$$\begin{cases} P(X^{\pi^*, \lambda^*}(T) \geq L) \geq 1 - \varepsilon, \\ \lambda^* \big(P(X^{\pi^*, \lambda^*}(T) \geq L) - 1 + \varepsilon \big) = 0, \end{cases} \tag{2.9}$$

then $X^{\pi^*, \lambda^*}(T)$ is the optimal solution of problem (2.7), π^* is the optimal control and λ^* is the Lagrange multiplier. This is verified in the next result.

Lemma 2.3. *Assume that there exists a nonnegative constant λ^* such that $X^{\pi^*,\lambda^*}(T)$ solves problem (2.8) and satisfies condition (2.9). Then, $X^{\pi^*,\lambda^*}(T)$ solves problem (2.7).*

Proof. Assume that $X^{\pi^*,\lambda^*}(T)$ is the optimal solution to problem (2.8) with condition (2.9). As $X^{\pi^*,\lambda^*}(T)$ is feasible for problem (2.7), we have

$$E[U(X^{\pi^*,\lambda^*}(T))] \leq \max_{\pi\in\mathcal{A}} E[U(X^{\pi}(T))],$$

where $X^{\pi}(T)$ is an arbitrary feasible solution for problem (2.7). On the other hand, since $X^{\pi^*,\lambda^*}(T)$ solves (2.8), we have

$$E[\tilde{U}_{\lambda *}(X^{\pi}(T))] \leq E[\tilde{U}_{\lambda *}(X^{\pi^*,\lambda^*}(T))],$$

that is,

$$E[U(X^{\pi}(T))] + \lambda^* P(X^{\pi}(T) \geq L) \leq E[U(X^{\pi^*,\lambda^*}(T))] + \lambda^* P(X^{\pi^*,\lambda^*}(T) \geq L).$$

Since $X^{\pi^*,\lambda^*}(T)$ satisfies condition (2.9) and $X^{\pi}(T)$ satisfies the VaR constraint, we have

$$E[U(X^{\pi}(T))] \leq E[U(X^{\pi^*,\lambda^*}(T))] + \lambda^*\big(1-\varepsilon - P(X^{\pi}(T) \geq L)\big) \leq E[U(X^{\pi^*,\lambda^*}(T))],$$

which implies

$$\max_{\pi\in\mathcal{A}} E[U(X^{\pi}(T))] \leq E[U(X^{\pi^*,\lambda^*}(T))].$$

Therefore, $X^{\pi^*,\lambda^*}(T)$ solves problem (2.7). □

When the financial market is incomplete, explicit solutions for portfolio choice problems are in general unavailable since there are many pricing kernels. However, if the incompleteness is due to closed convex cone control constraints, one can use the dual control method to find a particular one among all pricing kernels, which is the so-called minimum pricing kernel (see He and Zhou (2011)), and then solve the problem in the same way as in the complete market case with the help of this particular pricing kernel and the optional decomposition theorem (see Follmer and Kramkov (1997)). For continuous concave utility functions, Bian et al. (2011) use the dual control method to solve the utility maximization problem in the presence of closed convex cone control constraints. Since $\tilde{U}_{\lambda}$ is a discontinuous nonconcave utility function, we cannot directly apply the results in Bian et al. (2011) to solve problem (2.8). However, thanks to the concavification technique (see Carpenter (2000)), we can work on the concave envelope of $\tilde{U}_{\lambda}$ and apply the results in Bian et al. (2011) to solve problem (2.8), which is discussed in detail in the next section.

3 Solving unconstrained optimization problem (2.8)

For a fixed $\lambda \geq 0$, the dual function of $\tilde{U}_{\lambda}$ is defined by:

$$V_{\lambda}(y) = \sup_{x\geq 0}\{\tilde{U}_{\lambda}(x) - xy\}, y > 0. \tag{3.1}$$

Denote by f^c the concave envelope of a function f with domain D, that is, f^c is the smallest concave function that is greater than or equal to f, defined by

$$f^c(x) := \inf\{g(x) : D \to R|\ g \text{ is a concave function}, g(t) \geq f(t), \forall t \in D\}, \forall x \in D.$$

Since U is not concave, we first derive the concave envelope of U. To simplify the formulation, we introduce the notation

$$c_x := U_1'(x-\theta), x > \theta.$$

Note that c_x is a decreasing function of $x > \theta$ and is the slope of the tangent line to the curve $U_1(x-\theta)$ at point $x > \theta$.

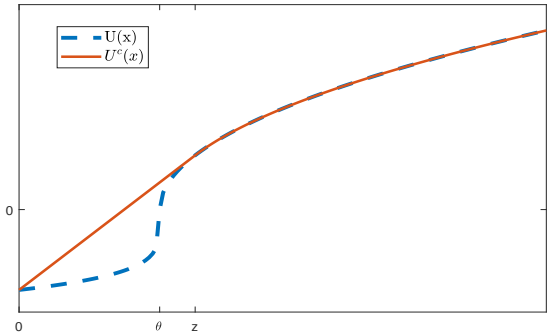

Figure 1: Concave envelope of $U(x), U^c(x)$

Let z be the tangent point of the straight line starting at $(0, -U_2(\theta))$ to the curve $U_1(x-\theta), x \geq \theta$. Simple calculus, using (2.5) and (2.6), shows that there exists a unique solution $z > \theta$ to the equation

$$U_1(x-\theta) + U_2(\theta) - xU_1'(x-\theta) = 0. \tag{3.2}$$

The concave envelope of U is given by (see Figure 1)

$$U^c(x) = \begin{cases} -\infty, & x < 0, \\ c_z x - U_2(\theta), & 0 \leq x < z, \\ U_1(x-\theta), & x \geq z, \end{cases} \tag{3.3}$$

where z is the solution to (3.2), see Carpenter (2000).

Consider the concavified version of problem (3.1):

$$V_\lambda^c(y) = \sup_{x \geq 0}\{\tilde{U}_\lambda^c(x) - xy\}, y > 0. \tag{3.4}$$

From Lemma 2.9 of Reichlin (2013), we have, for $y > 0$,

$$V_\lambda(y) = V_\lambda^c(y) = \tilde{U}_\lambda^c(x^{*,\lambda}(y)) - x^{*,\lambda}(y)y = \tilde{U}_\lambda(x^{*,\lambda}(y)) - x^{*,\lambda}(y)y, \tag{3.5}$$

where $x^{*,\lambda}(y)$ solves both (3.1) and (3.4). For $\lambda \geq 0$ and $L \geq \theta$, denote by

$$k_\lambda = \frac{U(L) + \lambda + U_2(\theta)}{L}. \tag{3.6}$$

Note that k_λ is the slope of the straight line linking points $(0, -U_2(\theta))$ and $(L, U(L)+\lambda)$ and k_λ depends on λ as well as parameters L, θ and utility functions U_1, U_2. The next three results characterize $\tilde{U}_\lambda^c$ and $x^{*,\lambda}(y)$ for three cases: $L \geq z$, $\theta \leq L < z$ and $L < \theta$.

Proposition 3.1. *Let $L \geq z$ with z determined by (3.2) and k_λ be defined by (3.6). Then we have $k_\lambda \geq k_0 \geq c_L$ and $c_z \geq k_0 \geq c_L$. For $y > 0$,*

Case I: If $k_\lambda > c_z$, then

$$\tilde{U}_\lambda^c(x) = \begin{cases} -\infty, & x < 0, \\ k_\lambda x - U_2(\theta), & 0 \leq x < L, \\ U_1(x-\theta) + \lambda, & x \geq L, \end{cases} \tag{3.7}$$

and

$$x^{*,\lambda}(y) = \begin{cases} \theta + I_1(y), & y < c_L, \\ L, & c_L \leq y < k_\lambda, \\ 0, & y \geq k_\lambda. \end{cases} \tag{3.8}$$

Case II: If $c_L \leq k_\lambda \leq c_z$, then

$$\tilde{U}_\lambda^c(x) = \begin{cases} -\infty, & x < 0, \\ c_z x - U_2(\theta), & 0 \leq x < z, \\ U_1(x-\theta), & z \leq x < L_0, \\ c_{L_0}(x-L) + U_1(L-\theta) + \lambda, & L_0 \leq x < L, \\ U_1(x-\theta) + \lambda, & x \geq L, \end{cases} \tag{3.9}$$

and

$$x^{*,\lambda}(y) = \begin{cases} \theta + I_1(y), & y < c_L, \\ L, & c_L \leq y < c_{L_0}, \\ \theta + I_1(y), & c_{L_0} \leq y < c_z, \\ 0, & y \geq c_z, \end{cases} \tag{3.10}$$

where L_0 is the unique solution in the interval $[z, L]$ of the equation

$$U_1(x-\theta) - U_1(L-\theta) - \lambda - (x-L)U_1'(x-\theta) = 0. \tag{3.11}$$

In particular, if $\lambda = 0$, then $L_0 = L$, $\tilde{U}_\lambda^c(x)$ is the same as $U^c(x)$ given by (3.3) and

$$x^{*,0}(y) = \begin{cases} \theta + I_1(y), & y < c_z, \\ 0, & y \geq c_z. \end{cases} \tag{3.12}$$

Proof. See Appendix. □

It is clear that $x^{*,0}(y)$ in (3.12) is bounded above by $x^{*,\lambda}(y)$ in (3.10) that is bounded above by $x^{*,\lambda}(y)$ in (3.8). For $k_\lambda > c_z$, the concave envelope $\tilde{U}_\lambda^c$ is linear in $[0, L)$ and coincides with $\tilde{U}_\lambda$ in $(-\infty, 0) \cup [L, \infty)$. We will later see that the utility changing point L plays an essential role in the optimal terminal wealth in this case. For $k_\lambda \leq c_z$, the concave envelope $\tilde{U}_\lambda^c$ is linear in $[0, z)$ and $[L_0, L)$, respectively, and coincides with $\tilde{U}_\lambda$ in $(-\infty, 0) \cup (z, L_0) \cup [L, \infty)$. In this case, the optimal terminal wealth is determined by the utility changing points z, L_0 and L.

Proposition 3.2. *Let $\theta \le L < z$ with z determined by (3.2) and k_λ be defined by (3.6). Then we have $k_0 < c_L$. For $y > 0$,*

Case I: If $k_\lambda \ge c_L$, then $\tilde{U}^c_\lambda(x)$ and $x^{*,\lambda}(y)$ are given by (3.7) and (3.8), respectively.

Case II: If $k_\lambda < c_L$, then

$$\tilde{U}^c_\lambda(x) = \begin{cases} -\infty, & x < 0, \\ c_{z_0} x - U_2(\theta), & 0 \le x < z_0, \\ U_1(x-\theta) + \lambda, & x \ge z_0, \end{cases} \tag{3.13}$$

and

$$x^{*,\lambda}(y) = \begin{cases} \theta + I_1(y), & y < c_{z_0}, \\ 0, & y \ge c_{z_0}, \end{cases} \tag{3.14}$$

where z_0 is the unique solution in the interval $(L, z]$ of the equation

$$U_1(x-\theta) + U_2(\theta) + \lambda - x U_1'(x-\theta) = 0. \tag{3.15}$$

In particular, if $\lambda = 0$, then $z_0 = z$, $\tilde{U}^c_\lambda(x)$ and $x^{*,0}(y)$ are given by (3.3) and (3.12), respectively.

Proof. See Appendix. □

We now turn to the case $L < \theta$. Let $\tilde{z}$ be the tangent point of the straight line starting at $(L, -U_2(\theta - L))$ to the curve $U_1(x-\theta), x \ge \theta$. Simple calculus shows that there exists a unique solution $\theta < \tilde{z} < z$ to the equation

$$U_1(x-\theta) + U_2(\theta - L) - (x - L) U_1'(x-\theta) = 0. \tag{3.16}$$

Proposition 3.3. *Let $L < \theta < \tilde{z} < z$ with $\tilde{z}, z$ determined by (3.16) and (3.2) and k_λ be defined by (3.6). Then we have $k_0 < c_z < c_{\tilde{z}}$. For $y > 0$,*

Case I: If $k_\lambda > c_{\tilde{z}}$, then

$$\tilde{U}^c_\lambda(x) = \begin{cases} -\infty, & x < 0, \\ k_\lambda x - U_2(\theta), & 0 \le x < L, \\ c_{\tilde{z}}(x - L) - U_2(\theta) + \lambda, & L \le x < \tilde{z}, \\ U_1(x-\theta) + \lambda, & x \ge \tilde{z}, \end{cases} \tag{3.17}$$

and

$$x^{*,\lambda}(y) = \begin{cases} \theta + I_1(y), & y < c_{\tilde{z}}, \\ L, & c_{\tilde{z}} \le y < k_\lambda, \\ 0, & y \ge k_\lambda. \end{cases} \tag{3.18}$$

Case II: If $k_\lambda \le c_{\tilde{z}}$, then $\tilde{U}^c_\lambda(x)$ and $x^{*,\lambda}(y)$ are given by (3.13) and (3.14) with z_0 replaced by $\tilde{z}_0$, where $\tilde{z}_0$ is the unique solution in the interval $[\tilde{z}, z]$ of the equation (3.15).

In particular, if $\lambda = 0$, then $\tilde{z}_0 = z$, $\tilde{U}^c_\lambda(x)$ and $x^{*,0}(y)$ are given by (3.3) and (3.12), respectively.

Proof. See Appendix. □

For $\theta \leq L < z$ and $L < \theta$, the concave envelope is also linear on the intervals where the concavification is needed and coincides with the curves of $\tilde{U}_\lambda$ in other regions. We will later see that the optimal terminal wealth for $\theta \leq L < z$ and $L < \theta$ is still determined by the utility changing points.

The wealth process X^π given by (2.3) is not self-financing due to the contribution term in a DC pension plan. Therefore, the problem (2.8) is not a classical utility maximization problem with control constraints. To apply the existing results on the optimization problem with control constraints, we introduce an auxiliary process as follows:

$$\tilde{X}^\pi(t) = X^\pi(t) + C(t), \tag{3.19}$$

where

$$C(t) = \int_t^T c(s)e^{-r(s-t)}ds \tag{3.20}$$

is the discounted value at time t of total pension contribution from t to T. Using (2.3), we have

$$d\tilde{X}^\pi(t) = (r\tilde{X}^\pi(t) + \pi^\top(t)\sigma\xi)dt + \pi^\top(t)\sigma dW(t), \tilde{X}^\pi(0) = \tilde{x}_0 \geq 0, \tag{3.21}$$

with $\tilde{x}_0 = x_0 + C(0) = x_0 + \int_0^T c(s)e^{-rs}ds$. The total wealth of the pension account $\tilde{X}^\pi(t)$ (the current wealth $X^\pi(t)$ plus the present value of future contributions $C(t)$) satisfies

$$\tilde{X}^\pi(T) = X^\pi(T), \quad \tilde{X}^\pi(t) \geq 0, \text{for all} \quad t \in [0, T].$$

Zhang and Ewald (2010) state that in the presence of a positive endowment stream, the current wealth $X^\pi(t)$ is allowed to be negative, provided the present value of future contributions is large enough to offset such a negative value. In this setup, the manager makes his investment decisions based on not only the current wealth, but also the present value of the future cash income. In theory x_0 can be negative as long as $\tilde{x}_0$ is nonnegative, however, the pension account is always endowed with a positive initial fund x_0. The optimization problem (2.8) is equivalent to the following problem:

$$\begin{cases} \max\limits_{\pi \in \mathcal{A}} E[\tilde{U}_\lambda(\tilde{X}^\pi(T))], \\ s.t.\ \tilde{X}^\pi(t) \text{ satisfies (3.21)}. \end{cases} \tag{3.22}$$

Define the value functions of the primal problem and its concavified version by

$$u_\lambda(t, \tilde{x}) = \max_{\pi \in \mathcal{A}} E[\tilde{U}_\lambda(\tilde{X}^\pi(T))|\tilde{X}^\pi(t) = \tilde{x}], \tag{3.23}$$

and

$$u_\lambda^c(t, \tilde{x}) = \max_{\pi \in \mathcal{A}} E[\tilde{U}_\lambda^c(\tilde{X}^\pi(T))|\tilde{X}^\pi(t) = \tilde{x}]. \tag{3.24}$$

Reichlin (2013) investigates the relationship between $u_\lambda^c(t, \tilde{x})$ and $u_\lambda(t, \tilde{x})$ and gives the following result.

Theorem 3.4. *(Reichlin (2013), Theorem 5.1) Assume that $u_\lambda(t, \tilde{x})$ and $u_\lambda^c(t, \tilde{x})$ are given by (3.23) and (3.24). Then it holds that $u_\lambda(t, \tilde{x}) = u_\lambda^c(t, \tilde{x})$.*

Theorem 3.4 states that the terminal wealth that is optimal for the concavified objective function is also optimal for the original objective function since it never takes on values where the two utility functions disagree. We next use the dual control method to solve the optimization problems $u_\lambda(t,\tilde{x})$ and $u_\lambda^c(t,\tilde{x})$. First, we define the dual control set.

Definition 3.5. *Let $\tilde{K}$ be the positive polar cone of K, i.e., $\tilde{K} = \{p : p^\top v \geq 0 \text{ for all } v \in K\}$. A dual control process is a progressively measurable, $\mathcal{F}$-adapted process $\nu = (\nu_1, \cdots, \nu_n)^\top$ which satisfies $E[\int_0^T \|\nu(t)\|^2 dt] < \infty$ and $\nu(t) \in \tilde{K}$ a.s. for all t. We denote the set of all dual control processes by $\mathcal{A}_0$.*

For $\nu \in \mathcal{A}_0$, define the dual process

$$dY^\nu(t) = Y^\nu(t)(-rdt - (\sigma^{-1}\nu(t) + \xi)^\top dW(t)), Y^\nu(0) = y_0.$$

Consider the dual minimization problem

$$\min_{\nu \in \mathcal{A}_0} E[V_\lambda^c(Y^\nu(T))].$$

The dual value function is defined by

$$v_\lambda(t,y) = \min_{v \in \mathcal{A}_0} E[V_\lambda^c(Y^\nu(T))|Y^\nu(t) = y].$$

The dual HJB equation is given by

$$\begin{cases} \frac{\partial v_\lambda}{\partial t}(t,y) - ry\frac{\partial v_\lambda}{\partial y}(t,y) + \frac{1}{2}y^2 \min_{\nu \in \tilde{K}} \|\xi + \sigma^{-1}\nu\|^2 \frac{\partial^2 v_\lambda}{\partial y^2}(t,y) = 0, y > 0, t < T, \\ v_\lambda(T,y) = V_\lambda(y). \end{cases} \tag{3.25}$$

Here we have used $V_\lambda(y) = V_\lambda^c(y)$ for $y > 0$. There exists a unique minimizer $\hat{\nu} \in \tilde{K}$ for convex quadratic function

$$f(\nu) = \|\xi + \sigma^{-1}\nu\|^2 \tag{3.26}$$

over $\nu \in \tilde{K}$, see Xu and Shreve (1992). Denote by

$$\hat{\xi} = \xi + \sigma^{-1}\hat{\nu}. \tag{3.27}$$

Under the condition $\hat{\xi} \neq 0$, the solution to (3.25) is given by

$$v_\lambda(t,y) = E[V_\lambda(Y^{\hat{\nu}}(T))|Y^{\hat{\nu}}(t) = y], \tag{3.28}$$

where $Y^{\hat{\nu}}(s) = y\frac{H^{\hat{\nu}}(s)}{H^{\hat{\nu}}(t)}$ for $t \leq s \leq T$ and

$$H^{\hat{\nu}}(t) = \exp\left(-\Big(r + \frac{\|\hat{\xi}\|^2}{2}\Big)t - \hat{\xi}^\top W(t)\right) \tag{3.29}$$

is the state-price density process in a fictitious market (see Cox and Huang (1989)) and the minimal pricing kernel (see He and Zhou (2011)). If there is no limitation on the trading strategy, then $\hat{\nu} = 0$ and $H^{\hat{\nu}}(t)$ is exactly the pricing kernel in the Black-Scholes complete market.

Remark 3.6. *If the portfolio shares are restricted to lie in a closed convex set K with $0 \in K$, then similar to Bian et al. (2011), we can define the dual process $Y^\nu(t)$ satisfying the SDE*

$$dY^\nu(t) = Y^\nu(t)(-\big(r + \delta_K(\nu)\big)dt - (\sigma^{-1}\nu + \xi)^\top dW(t)), Y^\nu(0) = y_0,$$

where

$$\delta_K(z) \doteq \sup_{\pi \in K}\{-\pi^\top z\}$$

is the support function of the set $-K$. Then the dual HJB equation is given by

$$\begin{cases} \frac{\partial v_\lambda}{\partial t}(t,y) + \min\limits_{\nu}\big\{ -(r + \delta_K(\nu))y\frac{\partial v_\lambda}{\partial y}(t,y) + \frac{1}{2}y^2\|\xi + \sigma^{-1}\nu\|^2\frac{\partial^2 v_\lambda}{\partial y^2}(t,y)\big\} = 0, y > 0, t < T, \\ v_\lambda(T,y) = V_\lambda(y). \end{cases} \tag{3.30}$$

In the case of $\lambda = 0$ (no VaR constraint) and U is a power or log utility, we can find a closed form solution of (3.30) by the homothetic property of U for any closed convex set K. For general utility with VaR constraint, it is highly difficult to solve (3.30) for general closed convex set K as (3.30) is a fully nonlinear PDE. However, if K is a closed convex cone, then $\delta_K(z) = 0$ if $z \in \tilde{K}$ and ∞ otherwise. Thus, equation (3.30) becomes (3.25) which allows an explicit solution. Equation (3.25) is consistent with (15.5) in Cvitanic and Karatzas (1992) for a strictly concave and continuously differentiable utility. To explicitly solve the optimization problem with an S-shaped utility and trading constraints, we only consider in this paper the case that K is a closed convex cone. We show the relation between the S-shaped utility maximization and the dual problem by using the concavification technique and apply the results in Bian et al. (2011) to solve the problem (2.8) for fixed Lagrange multiplier λ.

Bian et al. (2011) and Xu and Shreve (1992) give the relationship between the optimization problem $u^c_\lambda(t,\tilde{x})$ and the dual optimization problem. Combining with Theorem 3.4, we have the following result.

Theorem 3.7. *(Bian et al. (2011)) Assume that $\hat{\xi} \neq 0$, $v_\lambda(t,y)$ is given by (3.28) and conditions (2.5) and (2.6) hold, then we have, for $0 \leq t < T$,*

$$u_\lambda(t,\tilde{x}) = v_\lambda(t, y(t,\tilde{x})) + \tilde{x}y(t,\tilde{x}), \tilde{x} \geq 0,$$

where $y = y(t,\tilde{x})$ satisfies the equation

$$\frac{\partial v_\lambda}{\partial y}(t,y) + \tilde{x} = 0. \tag{3.31}$$

The optimal feedback control is given by

$$\pi^{*,\lambda}(t,\tilde{x}) = (\sigma^\top)^{-1}\hat{\xi}y(t,\tilde{x})\frac{\partial^2 v_\lambda}{\partial y^2}(t, y(t,\tilde{x})), \tag{3.32}$$

and $\pi^{*,\lambda}(t,\tilde{x}) \in K$. Furthermore, starting with the initial wealth $\tilde{x}_0$ at time 0, the optimal wealth process is given by

$$\tilde{X}^{\pi^*,\lambda}(t) = -\frac{\partial v_\lambda}{\partial y}(t, y_0 H^{\hat{\nu}}(t)), \tag{3.33}$$

where y_0 is the solution to the equation

$$\frac{\partial v_\lambda}{\partial y}(0,y_0) + \tilde{x}_0 = 0. \tag{3.34}$$

Note that $V_\lambda(y)$ is continuous for $y > 0$ and continuously differentiable for $y > 0$ except at finitely many points. Using (3.5) and pathwise differentiation, we have

$$\tilde{X}^{\pi^*,\lambda}(t) = -\frac{\partial v_\lambda}{\partial y}(t,y) = E\left[x^{*,\lambda}(Y^{\hat{\nu}}(T))\frac{Y^{\hat{\nu}}(T)}{y}\Big|Y^{\hat{\nu}}(t) = y\right], \tag{3.35}$$

where $x^{*,\lambda}(y)$ is defined in (3.8), (3.10) for $L \geq z$, (3.8), (3.14) for $\theta \leq L < z$, and (3.18), (3.14) for $L < \theta$, respectively. Equation (3.35) implies that once the unique pricing kernel $H^{\hat{\nu}}(t)$ is found, we can solve the unconstrained optimization problem (2.8) in the same way as in the complete market, which involves solving a terminal static optimization problem first to obtain the optimal terminal wealth and then deriving the optimal wealth process by computing the conditional expectation of the optimal terminal wealth under the pricing kernel $H^{\hat{\nu}}(t)$.

As $x^{*,\lambda}(y)$ is not continuous everywhere for $y > 0$, we cannot use the pathwise differentiation, but we may apply the likelihood ratio method (see Broadie and Glasserman (1996)) to obtain

$$\frac{\partial^2 v_\lambda}{\partial y^2}(t,y) = -E\left[x^{*,\lambda}\big(Y^{\hat{\nu}}(T)\big)\frac{Y^{\hat{\nu}}(T)\Big(\ln(\frac{Y^{\hat{\nu}}(T)}{y}) + \beta(t)\Big)}{\alpha(t)y^2}\Big|Y^{\hat{\nu}}(t) = y\right], \tag{3.36}$$

where

$$\alpha(t) = \|\hat{\xi}\|^2(T-t), \beta(t) = \Big(r - \frac{\|\hat{\xi}\|^2}{2}\Big)(T-t). \tag{3.37}$$

Since $Y^{\hat{\nu}}(T)$ is a lognormal variable, some lengthy but straightforward calculations will lead to closed-form expressions for $\pi^{*,\lambda}(t,\tilde{x})$, $\tilde{X}^{\pi^*,\lambda}(t)$ in (3.32) and (3.33). Once $\tilde{X}^{\pi^*,\lambda}(t)$ is derived, the optimal wealth process $X^{\pi^*,\lambda}(t)$ can be easily obtained from (3.19).

4 Solving constrained optimization problem (2.7)

In the previous section, we have applied the dual control method to solve the unconstrained optimization problem (2.8) for every fixed $\lambda \geq 0$. We now show there exists a $\lambda^* \geq 0$ such that condition (2.9) holds. Applying Theorem 3.7, we can easily find the optimal wealth process $X^{\pi^*,\lambda^*}(t)$ and the optimal investment strategy $\pi^{*,\lambda^*}(t)$ for $0 \leq t \leq T$, which are presented in Proposition A.1 in the appendix. We next state the main result of the paper.

Theorem 4.1. *Assume $\hat{\nu} \in \tilde{K}$ is the unique minimizer of (3.26), $\hat{\xi} \neq 0$, and*

$$x_0 + C(0) > E[LH^{\hat{\nu}}(T)1_{\{H^{\hat{\nu}}(T) < H^*\}}],$$

where H^* solves

$$P(H^{\hat{\nu}}(T) > H^*) = \varepsilon, \tag{4.1}$$

for a given $0 \leq \varepsilon \leq 1$. Then there exists a unique $\lambda^* \geq 0$ such that $X^{\pi^*,\lambda^*}(T)$ is the optimal solution of unconstrained problem (2.8) and satisfies condition (2.9). Therefore, $X^{\pi^*,\lambda^*}(T)$ is the optimal solution of VaR constrained problem (2.7) and λ^* is the corresponding Lagrange multiplier.

Proof. We first consider the case $L \geq z$, where z is the solution to (3.2). For a fixed $\lambda \geq 0$, we can obtain from equation (3.35) that the optimal terminal wealth is given by

$$X^{\pi^{*,\lambda}}(T) = \tilde{X}^{\pi^{*,\lambda}}(T) = x^{*,\lambda}(y_0 H^{\hat{\nu}}(T)), \tag{4.2}$$

where $x^{*,\lambda}(y)$ is defined in (3.8), (3.10) and (3.12), and y_0 is determined by the binding budget constraint

$$E[X^{\pi^{*,\lambda}}(T)H^{\hat{\nu}}(T)] = x_0 + C(0), \tag{4.3}$$

If we can find a unique solution (y_0, λ^*) to equations (2.9) and (4.3), then $X^{\pi^{*,\lambda^*}}(T)$ is the solution to the problem (2.7). To solve equations (2.9) and (4.3), let H^* be defined by (4.1). We now choose λ and check the bindingness of the VaR constraint by comparing the solution to the optimal terminal wealth without VaR constraint with the threshold of the VaR constraint H^*. If $\lambda = 0$, then we have $x^{*,\lambda}(y)$ in (4.2) is defined by (3.12) and the optimal terminal wealth without VaR constraint, denoted by $X^{\pi^*}(T)$, is given by

$$X^{\pi^*}(T) = (\theta + I_1(y_0 H^{\hat{\nu}}(T)))1_{\{H^{\hat{\nu}}(T) < \frac{c_z}{y_0}\}}, \tag{4.4}$$

where y_0 is determined by (4.3) with $X^{\pi^{*,\lambda}}(T)$ replaced by $X^{\pi^*}(T)$. Note that for any $\omega \in \Omega$, $y_0 \to X^{\pi^*}(T)$ is a decreasing function of y_0 since I_1 is strictly decreasing. Then $V(y_0) = E\big[H^{\hat{\nu}}(T)X^{\pi^*}(T)\big]$ is continuous and strictly decreasing in y_0. Furthermore, for any $\omega \in \Omega$, we have $\lim\limits_{y_0 \to 0^+} X^{\pi^*}(T) = \infty$ and $\lim\limits_{y_0 \to \infty} X^{\pi^*}(T) = 0$, which yields

$$\lim_{y_0 \to 0^+} V(y_0) = \infty, \lim_{y_0 \to \infty} V(y_0) = 0 < x_0 + C(0).$$

Thus, there exists a unique solution y_0 to equation (4.3).

If $H^* \leq \frac{c_L}{y_0}$, then

$$P(X^{\pi^*}(T) \geq L) = P(H^{\hat{\nu}}(T) \leq \frac{c_L}{y_0}) \geq P(H^{\hat{\nu}}(T) \leq H^*) = 1 - \varepsilon.$$

We can choose $\lambda^* = 0$ as $X^{\pi^*}(T)$ naturally satisfies the VaR constraint and maximizes problem (2.7).

If $H^* > \frac{c_L}{y_0}$, then the VaR constraint is binding and it should hold that $\lambda > 0$, which implies that $X^{\pi^{*,\lambda}}(T)$ satisfies

$$P(X^{\pi^{*,\lambda}}(T) \geq L) = 1 - \varepsilon. \tag{4.5}$$

Then we can choose λ as a function of y_0 by using (4.1) and (2.9).

(i) If $\frac{c_L}{y_0} < H^* \leq \frac{c_z}{y_0}$, then $X^{\pi^{*,\lambda}}(T) = x^{*,\lambda}(y_0 H^{\hat{\nu}}(T))$ takes a four-region form, where $x^{*,\lambda}(y)$ is defined in (3.10), that is,

$$X^{\pi^{*,\lambda}}(T) = (\theta + I_1(y_0 H^{\hat{\nu}}(T)))\Big(1_{\{H^{\hat{\nu}}(T) < \frac{c_L}{y_0}\}} + 1_{\{H^* \leq H^{\hat{\nu}}(T) < \frac{c_z}{y_0}\}}\Big) + L1_{\{\frac{c_L}{y_0} \leq H^{\hat{\nu}}(T) < H^*\}}. \tag{4.6}$$

Define L_0 by the relation $H^* = \frac{c_{L_0}}{y_0}$, which is to ensure (4.5) holds. Since $c_L < c_{L_0} \leq c_z$, we have $z \leq L_0 < L$. Define

$$\lambda = U(L_0) - U(L) - U'(L_0)(L_0 - L) \hat{=} g_1(y_0).$$

Since $\frac{d}{dx}(U(x)-U(L)-U'(x)(x-L)) = -U''(x)(x-L) < 0$ for $\theta < x < L$, we conclude that $\lambda > U(L)-U(L)-U'(L)(L-L) = 0$.

(ii) If $H^* > \frac{c_z}{y_0}$, then $X^{\pi^*,\lambda}(T) = x^{*,\lambda}(y_0 H^{\hat{\nu}}(T))$ takes a three-region form, where $x^{*,\lambda}(y)$ is defined in (3.8), that is,

$$X^{\pi^*,\lambda}(T) = (\theta + I_1(y_0 H^{\hat{\nu}}(T)))1_{\{H^{\hat{\nu}}(T) < \frac{c_L}{y_0}\}} + L1_{\{\frac{c_L}{y_0} \le H^{\hat{\nu}}(T) < H^*\}}. \tag{4.7}$$

Define k_λ by the relation $H^* = \frac{k_\lambda}{y_0}$, which is again to ensure (4.5) holds. We have $k_\lambda > c_z$. Define

$$\lambda = k_\lambda L - U(L) - U_2(\theta) \hat{=} g_2(y_0).$$

It is easy to check $\lambda = L(k_\lambda - \frac{U_1(L-\theta)+U_2(\theta)}{L}) \ge L(k_\lambda - c_z) > 0$.

Therefore, the multiplier λ can be chosen as a function of y_0:

$$\lambda = g_1(y_0)1_{\{\frac{c_L}{y_0} < H^* \le \frac{c_z}{y_0}\}} + g_2(y_0)1_{\{H^* > \frac{c_z}{y_0}\}} \hat{=} g(y_0).$$

It remains to show that there is a unique root y_0 to (4.3). Note that $V_1(y_0) = E[H^{\hat{\nu}}(T)X^{\pi^*,\lambda}(T)]$ is continuous and strictly decreasing in y_0. Furthermore, for any $\omega \in \Omega$, we have $\lim_{y_0 \to 0^+} X^{\pi^*,\lambda}(T) = \infty$ and $\lim_{y_0 \to \infty} X^{\pi^*,\lambda}(T) = L1_{\{H^{\hat{\nu}}(T) < H^*\}}$, which yields

$$\lim_{y_0 \to 0^+} V_1(y_0) = \infty, \lim_{y_0 \to \infty} V_1(y_0) = E[H^{\hat{\nu}}(T)L1_{\{H^{\hat{\nu}}(T) < H^*\}}] < x_0 + C(0).$$

Thus, there exists a unique solution y_0 to equation (4.3). We conclude that $X^{\pi^*,\lambda^*}(T)$ solves problem (2.7), $\lambda^* = g(y_0)$ is the Lagrange multiplier and y_0 is the unique solution of equation (4.3).

We next consider the case $\theta \le L < z$. The proof is similar. For a fixed $\lambda \ge 0$, the optimal terminal wealth is $X^{\pi^*,\lambda}(T) = x^{*,\lambda}(y_0 H^{\hat{\nu}}(T))$, where $x^{*,\lambda}(y)$ is defined in (3.8), (3.14), (3.12) and y_0 is determined by (4.3). It remains to find the unique solution (y_0, λ^*) to equations (2.9) and (4.3). If $H^* \le \frac{c_z}{y_0}$, then $X^{\pi^*}(T)$ naturally achieves the VaR constraint and the multiplier λ^* is 0. If $H^* > \frac{c_z}{y_0}$, the VaR constraint is binding and we choose λ as a function of y_0.

(i) If $\frac{c_z}{y_0} < H^* \le \frac{c_L}{y_0}$, then $X^{\pi^*,\lambda}(T) = x^{*,\lambda}(y_0 H^{\hat{\nu}}(T))$ takes a two-region form, where $x^{*,\lambda}(y)$ is defined in (3.14), that is,

$$X^{\pi^*,\lambda}(T) = (\theta + I_1(y_0 H^{\hat{\nu}}(T)))1_{\{H^{\hat{\nu}}(T) < H^*\}}. \tag{4.8}$$

Define z_0 by the relation $H^* = \frac{c_{z_0}}{y_0}$, which is to ensure (4.5) holds. Since $c_z < c_{z_0} < c_L$, we have $L < z_0 < z$. Define

$$\lambda = z_0 U'(z_0) - U_2(\theta) - U(z_0) \hat{=} \hat{g}_1(y_0).$$

Since $\frac{d}{dx}(xU'(x) - U(x)) = U''(x)x < 0$ for $x > \theta$, we have that $\lambda > zU'(z) - U_2(\theta) - U(z) = 0$.

(ii) If $H^* > \frac{c_L}{y_0}$, then $X^{\pi^*,\lambda}(T)$ is the same as (4.7) and $\lambda = g_2(y_0)$.

Similarly, we can prove that equation (4.3) has a unique root with $X^{\pi^*,\lambda}(T)$ given by (4.4), (4.7), (4.8) and

$$\lambda \hat{=} \hat{g}(y_0) := \hat{g}_1(y_0)1_{\{\frac{c_z}{y_0} < H^* \le \frac{c_L}{y_0}\}} + g_2(y_0)1_{\{H^* > \frac{c_L}{y_0}\}}.$$

Therefore, when the VaR constraint is binding, $X^{\pi^*,\lambda}(T)$ solves problem (2.7), $\lambda^* = \hat{g}(y_0)$ is the Lagrange multiplier and y_0 is the unique solution of equation (4.3).

Finally, we consider the case $L < \theta$. For a fixed $\lambda \geq 0$, the optimal terminal wealth is $X^{\pi^{*,\lambda}}(T) = x^{*,\lambda}(y_0 H^{\hat{\nu}}(T))$, where $x^{*,\lambda}(y)$ is defined in (3.8), (3.14), (3.18) and y_0 is determined by (4.3). It remains to find the unique solution (y_0, λ^*) to equations (2.9) and (4.3). If $H^* \leq \frac{c_z}{y_0}$, then $X^{\pi^*}(T)$ naturally achieves the VaR constraint and the multiplier λ^* is 0. If $H^* > \frac{c_z}{y_0}$, the VaR constraint is binding and we choose λ as a function of y_0.

(i) If $\frac{c_z}{y_0} < H^* \leq \frac{c_{\tilde{z}}}{y_0}$, then $X^{\pi^{*,\lambda}}(T)$ is given by (4.8). Define $\tilde{z}_0$ by the relation $H^* = \frac{c_{\tilde{z}_0}}{y_0}$, which is to ensure (4.5) holds. Since $c_z < c_{\tilde{z}_0} \leq c_{\tilde{z}}$, we have $\tilde{z} \leq \tilde{z}_0 < z$. Define

$$\lambda = \tilde{z}_0 U'(\tilde{z}_0) - U_2(\theta) - U(\tilde{z}_0) \hat{=} \tilde{g}_1(y_0).$$

Since $\frac{d}{dx}(xU'(x) - U(x)) = U''(x)x < 0$ for $x > \theta$, we have that $\lambda > zU'(z) - U_2(\theta) - U(z) = 0$.

(ii) If $H^* > \frac{c_{\tilde{z}}}{y_0}$, then $X^{\pi^{*,\lambda}}(T) = x^{*,\lambda}(y_0 H^{\hat{\nu}}(T))$ takes a three-region form, where $x^{*,\lambda}(y)$ is defined in (3.18), that is,

$$X^{\pi^{*,\lambda}}(T) = (\theta + I_1(y_0 H^{\hat{\nu}}(T)))1_{\{H^{\hat{\nu}}(T) < \frac{c_{\tilde{z}}}{y_0}\}} + L1_{\{\frac{c_{\tilde{z}}}{y_0} \leq H^{\hat{\nu}}(T) < H^*\}}. \tag{4.9}$$

Define $\lambda \hat{=} g_2(y_0) = k_\lambda L - U(L) - U_2(\theta) = k_\lambda L + U_2(L - \theta) - U_2(\theta)$. It is easy to check $\lambda = L(k_\lambda - \frac{U_2(\theta) - U_2(L-\theta)}{L}) \geq L(k_\lambda - c_{\tilde{z}}) > 0$. Similarly, we can prove that equation (4.3) has a unique root with $X^{\pi^{*,\lambda^*}}(T)$ given by (4.4),(4.8), (4.9) and

$$\lambda \hat{=} \tilde{g}(y_0) := \tilde{g}_1(y_0)1_{\{\frac{c_z}{y_0} < H^* \leq \frac{c_{\tilde{z}}}{y_0}\}} + g_2(y_0)1_{\{H^* > \frac{c_{\tilde{z}}}{y_0}\}}.$$

Therefore, when the VaR constraint is binding, $X^{\pi^{*,\lambda^*}}(T)$ solves problem (2.7), $\lambda^* = \tilde{g}(y_0)$ is the Lagrange multiplier and y_0 is the unique solution of equation (4.3). □

Remark 4.2. *Similar to Basak and Shapiro (2001), from (4.4), (4.6), (4.7) and (4.8), it is easy to see that if $x_0 + C(0) < E[LH^{\hat{\nu}}(T)1_{\{H^{\hat{\nu}}(T) < H^*\}}]$, then the optimization problem (2.7) is infeasible. For $x_0 + C(0) = E[LH^{\hat{\nu}}(T)1_{\{H^{\hat{\nu}}(T) < H^*\}}]$, there is only one admissible solution $X^{\pi^{*,\lambda^*}}(T) = L1_{\{H^{\hat{\nu}}(T) < H^*\}}$. In particular, if $\varepsilon = 0$, then $H^* = \infty$ and $X^{\pi^{*,\lambda^*}}(T) = L = e^{rT}(x_0 + C(0))$, which implies that one should only invest in the riskless savings account. The assumption $x_0 + C(0) > E[LH^{\hat{\nu}}(T)1_{\{H^{\hat{\nu}}(T) < H^*\}}]$ ensures there is a set of nontrivial admissible solutions.*

The constrained and unconstrained optimal terminal wealth $X^{\pi^{*,\lambda^*}}(T)$ can be expressed as a function of the state price density at maturity $H^{\hat{\nu}}(T)$ and the initial dual state value y_0, which is determined via the budget constraint. For the unconstrained optimal terminal wealth, when the state price density $H^{\hat{\nu}}(T)$ is relatively low, the optimal terminal wealth is similar to the smooth utility; when $H^{\hat{\nu}}(T)$ increases above a critical value of the price density $\frac{c_z}{y_0}$, that is, the boundary point of the bad-states region, the optimal terminal wealth drops to 0 since the loss aversion states a risk-seeking preference in the the domain of losses. To attain the VaR constraint, when $H^{\hat{\nu}}(T)$ is in the intermediate-states region, which is $\{\max\{\frac{c_L}{y_0}, \frac{c_z}{y_0}\} \leq H^{\hat{\nu}}(T) < H^*\}$ for $L \geq \theta$ or $\{\frac{c_z}{y_0} \leq H^{\hat{\nu}}(T) < H^*\}$ for $L < \theta$, the optimal terminal wealth is equal to the protection level of VaR constraint or is similar to the smooth utility. $X^{\pi^{*,\lambda^*}}(T)$ may take a two-, three- or four-region form for $L \geq z$ and may take a two- or three-region form for $\theta \leq L < z$ and $L < \theta$ according to the relative position of critical values of the state price, H^*, $\frac{c_L}{y_0}$ and $\frac{c_z}{y_0}$ for $L \geq \theta$ or H^*, $\frac{c_z}{y_0}$ and $\frac{c_{\tilde{z}}}{y_0}$ for $L < \theta$, which are the boundary points of the bad-, intermediate- and good-states regions. In all

cases, the optimal terminal wealth takes zero when the state price increases above the boundary point of the bad-states region. Once the exact form is determined, $X^{\pi^*,\lambda^*}(T)$ is assigned to different wealth levels corresponding to the partition of the terminal market states with boundary points H^*, $\frac{c_L}{y_0}$ and $\frac{c_z}{y_0}$ for $L \geq \theta$ or H^*, $\frac{c_z}{y_0}$ and $\frac{c_{\tilde{z}}}{y_0}$ for $L < \theta$. The values of y_0 in different forms of $X^{\pi^*,\lambda^*}(T)$ are different. Denote by $y_0^2, y_0^3, y_0^4, \tilde{y}_0^2, \tilde{y}_0^3$ the values of y_0 determined by (4.4), (4.6), (4.7), (4.8) and (4.9), respectively. It is easy to conclude that $y_0^2 < y_0^4 < y_0^3$, $y_0^2 < \tilde{y}_0^2 < y_0^3$ and $y_0^2 < \tilde{y}_0^2 < \tilde{y}_0^3$ from the budget constraint. With the help of these observations, we can design a simple algorithm to find the optimal terminal wealth and the Lagrange multiplier.

Algorithm for finding the optimal terminal wealth and Lagrange multiplier.

Step 0 Find $z > \theta$ from equation (3.2). If $L \geq z$, go to Step 1; if $\theta \leq L < z$, go to Step 1'; if $L < \theta$, go to Step 1".

Step 1 Let $X^{\pi^*,\lambda^*}(T)$ be given by (4.7) and compute $y_0 = y_0^3$ from (4.3). If $y_0^3 > \frac{c_z}{H^*}$, then $X^{\pi^*,\lambda^*}(T)$ is the optimal terminal wealth with the Lagrange multiplier $\lambda^* = g_2(y_0^3)$, stop. If $y_0^3 \leq \frac{c_z}{H^*}$, go to Step 2.

Step 1' Let $X^{\pi^*,\lambda^*}(T)$ be given by (4.7) and compute $y_0 = y_0^3$ from (4.3). If $y_0^3 > \frac{c_L}{H^*}$, then $X^{\pi^*,\lambda^*}(T)$ is the optimal terminal wealth with the Lagrange multiplier $\lambda^* = g_2(y_0^3)$, stop. If $y_0^3 \leq \frac{c_L}{H^*}$, go to Step 2'.

Step 1" Find $\theta < \tilde{z} < z$ from equation (3.16). Let $X^{\pi^*,\lambda^*}(T)$ be given by (4.9) and compute $y_0 = \tilde{y}_0^3$ from (4.3). If $\tilde{y}_0^3 > \frac{c_{\tilde{z}}}{H^*}$, then $X^{\pi^*,\lambda^*}(T)$ is the optimal terminal wealth with the Lagrange multiplier $\lambda^* = g_2(\tilde{y}_0^3)$, stop. If $\tilde{y}_0^3 \leq \frac{c_{\tilde{z}}}{H^*}$, go to Step 2'.

Step 2 Let $X^{\pi^*,\lambda^*}(T)$ be given by (4.6) and compute $y_0 = y_0^4$ from (4.3). If $y_0^4 > \frac{c_L}{H^*}$, then $X^{\pi^*,\lambda^*}(T)$ is the optimal terminal wealth with the Lagrange multiplier $\lambda^* = g_1(y_0^4)$, stop. If $y_0^4 \leq \frac{c_L}{H^*}$, go to Step 3.

Step 2' Let $X^{\pi^*,\lambda^*}(T)$ be given by (4.8) and compute $y_0 = \tilde{y}_0^2$ from (4.3). If $\tilde{y}_0^2 > \frac{c_z}{H^*}$, then $X^{\pi^*,\lambda^*}(T)$ is the optimal terminal wealth with the Lagrange multiplier $\lambda^* = \hat{g}_1(\tilde{y}_0^2)$, stop. If $\tilde{y}_0^2 \leq \frac{c_z}{H^*}$, go to Step 3.

Step 3 Let $X^{\pi^*}(T)$ by given by (4.4). Then $X^{\pi^*}(T)$ naturally satisfies the VaR constraint and is the optimal terminal wealth with the Lagrange multiplier $\lambda^* = 0$, stop.

Since λ^* can be expressed as a function of y_0 and λ^* is related to H^*, which is determined by ε, we use the superscript ε in place of λ^* in $X^{\pi^*,\lambda^*}(T)$ in the following analysis, that is, the optimal terminal wealth is written as $X^{\pi^*,\varepsilon}(T)$.

Remark 4.3. *When the utility is a smooth concave increasing function, we can obtain the optimal wealth process with the VaR constraint from Theorem 4.1 by setting $\theta = 0$, which results in $z = 0$ and $c_z = \infty$. Guan and Liang (2016) apply the martingale method to derive it in a complete market. If there are only two risky assets and the constraint set K is the whole space, then the optimal terminal wealth is given by (4.6) with $\theta = 0$ and $\hat{\nu} = (0, 0)^\top$, which is the same as (3.20) in Guan and Liang (2016).*

Remark 4.4. *From Theorem 4.1 we can conclude that if $\varepsilon = 0$, then $H^* = \infty$, which implies that all the bad states are insured against and the VaR constraint becomes the PI constraint. For $L \geq \theta$, the optimal terminal wealth takes a two-region form given by (4.7) with $H^* = \infty$, that is*

$$X^{\pi^{*,0}}(T) = (\theta + I_1(y_0 H^{\hat{\nu}}(T)))1_{\{H^{\hat{\nu}}(T) < \frac{c_L}{y_0}\}} + L1_{\{H^{\hat{\nu}}(T) \geq \frac{c_L}{y_0}\}}. \tag{4.10}$$

For $L < \theta$, the optimal terminal wealth also takes a two-region form given by (4.9) with $H^* = \infty$, that is

$$X^{\pi^{*,0}}(T) = (\theta + I_1(y_0 H^{\hat{\nu}}(T)))1_{\{H^{\hat{\nu}}(T) < \frac{c_z}{y_0}\}} + L1_{\{H^{\hat{\nu}}(T) \geq \frac{c_z}{y_0}\}}. \tag{4.11}$$

If $\varepsilon = 1$, then $H^* = 0$ and $\lambda = 0$, which implies that the VaR constraint vanishes. The optimal terminal wealth $X^{\pi^{*,1}}(T)$ is the same as (4.4), that is,

$$X^{\pi^{*,1}}(T) = X^{\pi^*}(T). \tag{4.12}$$

Similar to the optimal terminal wealth under a smooth concave utility in Basak and Shapiro (2001), in contrast to $X^{\pi^*}(T)$, $X^{\pi^{*,\varepsilon}}(T)$ is not modified in the good- and bad-states regions and $X^{\pi^{*,\varepsilon}}(T)$ is greater than or equals to the bound L in the intermediate-states region $\{\max\{\frac{c_L}{y_0}, \frac{c_z}{y_0}\} \leq H^{\hat{\nu}}(T) < H^*\}$ for $L \geq \theta$ or $\{\frac{c_z}{y_0} \leq H^{\hat{\nu}}(T) < H^*\}$ for $L < \theta$ in order to achieve the VaR constraint.

We next analyze how ε and L impact the optimal terminal wealth. Following Proposition 1 of Basak and Shapiro (2001), we can deduce from the budget constraint (4.3) that for a fixed ε, y_0 increases in L and the intermediate-states region grows at the expense of the good-states region. Accordingly, to attain a higher protection level, the optimal terminal wealth in the good-states region decreases. Similarly, for a fixed L, y_0 decreases in ε. With ε decreasing, the intermediate-states region $\{\max\{\frac{c_L}{y_0}, \frac{c_z}{y_0}\} \leq H^{\hat{\nu}}(T) < H^*\}$ for $L \geq \theta$ or $\{\frac{c_z}{y_0} \leq H^{\hat{\nu}}(T) < H^*\}$ for $L < \theta$ enlarges as more states need to be insured against, while the good-states region $\{H^{\hat{\nu}}(T) < \max\{\frac{c_L}{y_0}, \frac{c_z}{y_0}\}\}$ for $L \geq \theta$ or $\{H^{\hat{\nu}}(T) < \frac{c_z}{y_0}\}$ for $L < \theta$ and the bad-states region $\{H^{\hat{\nu}}(T) \geq H^*\}$ both shrink. Furthermore, to meet the VaR constraint, the optimal terminal wealth in the good-states region must decrease.

As shown in Theorem 4.1, the constrained optimal terminal wealth $X^{\pi^{*,\varepsilon}}(T)$ takes different forms depending on the relative position of critical values of state price H^*, $\frac{c_L}{y_0}$ and $\frac{c_z}{y_0}$ for $L \geq \theta$ or H^*, $\frac{c_z}{y_0}$ and $\frac{c_{\tilde{z}}}{y_0}$ for $L < \theta$. We now analyze how ε determines the form of $X^{\pi^{*,\varepsilon}}(T)$ for $L \geq z$. Denote by

$$\varepsilon^* = P(X^{\pi^*}(T) < L) = P(H^{\hat{\nu}}(T) > \frac{c_L}{y_0}),$$

where $X^{\pi^*}(T)$ is the unconstrained optimal terminal wealth given by (4.4) and y_0 is determined by the binding budget constraint (4.3). When $\varepsilon \geq \varepsilon^*$, $X^{\pi^*}(T)$ naturally satisfies the VaR constraint and therefore, $X^{\pi^{*,\varepsilon}}(T)$ is the same as $X^{\pi^*}(T)$. When $\varepsilon < \varepsilon^*$, the VaR constraint is binding and the uninsured loss-states region $\{H^{\hat{\nu}}(T) \geq H^*\}$ may take a one- or two-region form. More precisely, the states in the region $\{H^* \leq H^{\hat{\nu}}(T) < \max\{H^*, \frac{c_z}{y_0}\}\}$ have wealth in $[z, L)$ and the states in the region $\{H^{\hat{\nu}}(T) \geq \max\{H^*, \frac{c_z}{y_0}\}\}$ have wealth 0. Let

$$\bar{\varepsilon}^* = P(X^{\pi^{*,\varepsilon}}(T) < L) = P(H^{\hat{\nu}}(T) > \frac{c_z}{y_0}),$$

where $X^{\pi^*,\varepsilon}(T)$ is given by (4.7) with $H^* = \frac{c_z}{y_0}$ and y_0 is determined by the binding budget constraint (4.3). When $\varepsilon = \overline{\varepsilon}^*$, all the states with wealth in $[z, L)$ are insured against and the manager leaves all the states with wealth 0 uninsured. For $\varepsilon \in (\overline{\varepsilon}^*, \varepsilon^*)$, we have $H^* \in (\frac{c_L}{y_0}, \frac{c_z}{y_0})$, which implies that some states with wealth in $[z, L)$ and all the states with wealth 0 are left uninsured, so $X^{\pi^*,\varepsilon}(T)$ takes a two-region form in the loss-states region and takes a four-region form given by (4.6). For $\varepsilon \in (0, \overline{\varepsilon}^*]$, we have $H^* \in [\frac{c_z}{y_0}, \infty)$, which implies that the manager chooses to leave some states with wealth 0 uninsured, so $X^{\pi^*,\varepsilon}(T)$ only takes 0 in the loss-states region and takes a three-region form given by (4.7). From the above analysis, for a given $\varepsilon \in [0, 1]$, we can specify the exact form of $X^{\pi^*,\varepsilon}(T)$ without resorting to the algorithm.

One disadvantage of the VaR constraint under a smooth concave utility is that it leads to higher losses in bad market scenarios than in the case of no VaR constraint, see Basak and Shapiro (2001). Loss aversion leads $X^{\pi^*}(T)$ to be 0 in the region $\{H^{\hat{\nu}}(T) \geq \frac{c_z}{y_0}\}$. For a relatively large ε, a VaR constraint brings higher losses for those states in the region $(H^*, \frac{c_z}{y_0})$, where the optimal terminal wealth takes $\theta + I_1(y_0 H^{\hat{\nu}}(T))$. However, if ε is small enough such that $H^* \geq \frac{c_z}{y_0}$, then we have $X^{\pi^*,\varepsilon}(T) \geq X^{\pi^*}(T)$ in the region where $X^{\pi^*}(T) \leq L$. Therefore, a VaR constraint with a relatively small ε strictly improves risk management in bad economic states.

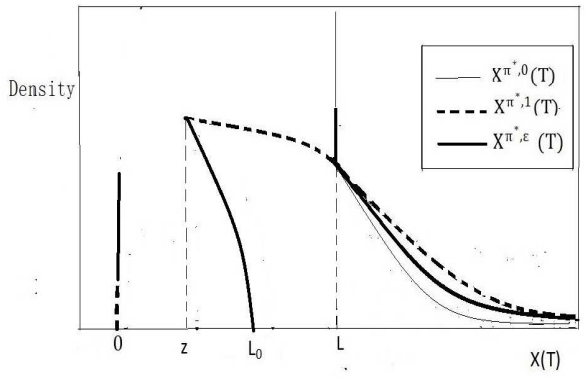

Figure 2: Probability density of optimal terminal wealth with a relatively large ε

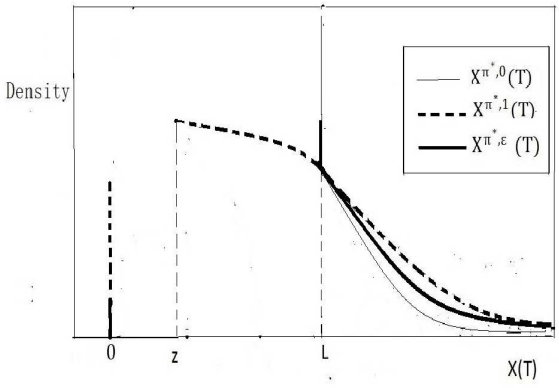

Figure 3: Probability density of optimal terminal wealth with a relatively small ε

Figures 2 and 3 depict the distributions of $X^{\pi^*,\varepsilon}(T)$ for $L \geq z$ with relatively large and small ε. The distribution of the optimal terminal wealth is not continuous. There is a probability mass at L when the VaR constraint is binding and there is a probability mass at 0 for $\varepsilon > 0$. For a relatively large ε (see Figure 2), states with wealth 0 have a higher probability than that in case of no VaR constraint and there are some states with wealth less than that in case of no VaR constraint, which implies that a VaR constraint leads to more losses. However, for a relatively small ε (see Figure 3), there are no states with wealth between (z, L) and states with wealth 0 have a lower probability than that in case of no VaR constraint, which implies that the VaR constrained optimal terminal wealth dominates the unconstrained one for bad economic states.

Similarly, for $\theta \leq L < z$, denote by

$$\varepsilon_1^* = P(X^{\pi^*}(T) < L) = P(H^{\hat{\nu}}(T) > \frac{c_z}{y_0}),$$

where y_0 is determined by the binding budget constraint (4.3). When $\varepsilon \geq \varepsilon_1^*$, $X^{\pi^*}(T)$ given by

(4.4) naturally satisfies the VaR constraint and $X^{\pi^*,\varepsilon}(T) = X^{\pi^*}(T)$. Let

$$\overline{\varepsilon}_1^* = P(X^{\pi^*,\varepsilon}(T) < L) = P(H^{\hat{\nu}}(T) > \frac{c_L}{y_0}),$$

where $X^{\pi^*,\varepsilon}(T)$ is given by (4.8) with $H^* = \frac{c_L}{y_0}$ and y_0 is determined by the binding budget constraint (4.3). For $\varepsilon \in [\overline{\varepsilon}_1^*, \varepsilon_1^*)$, we have $H^* \in (\frac{c_z}{y_0}, \frac{c_L}{y_0}]$ and the pension manager insures against the region $\{\frac{c_z}{y_0} \le H^{\hat{\nu}}(T) < H^*\}$ by letting $X^{\pi^*,\varepsilon}(T) = \theta + I_1(y_0 H^{\hat{\nu}}(T))$, so $X^{\pi^*,\varepsilon}(T)$ takes a two-region form given by (4.8). For $\varepsilon \in (0, \overline{\varepsilon}_1^*)$, we have $H^* \in (\frac{c_L}{y_0}, \infty)$ and the pension manager insures against the regions $\{\frac{c_z}{y_0} \le H^{\hat{\nu}}(T) < \frac{c_L}{y_0}\}$ and $\{\frac{c_L}{y_0} \le H^{\hat{\nu}}(T) < H^*\}$ by letting $X^{\pi^*,\varepsilon}(T) = \theta + I_1(y_0 H^{\hat{\nu}}(T))$ and $X^{\pi^*,\varepsilon}(T) = L$, respectively, so $X^{\pi^*,\varepsilon}(T)$ takes a three-region form given by (4.7).

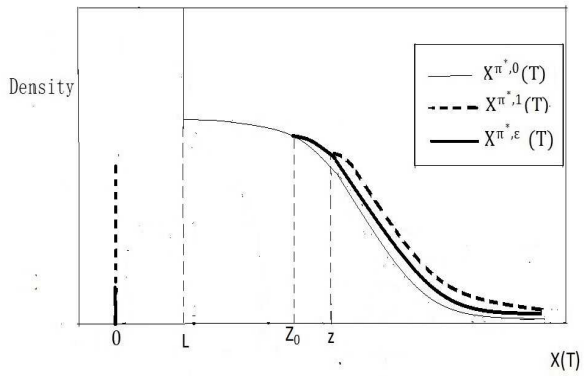

Figure 4: Probability density of optimal terminal wealth with a relatively large ε

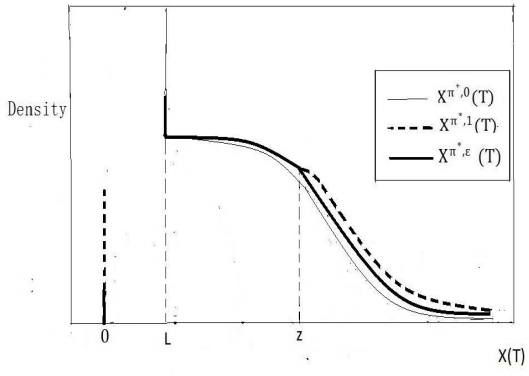

Figure 5: Probability density of optimal terminal wealth with a relatively small ε

Figures 4 and 5 depict the distributions of $X^{\pi^*,\varepsilon}(T)$ for $\theta \le L < z$ with relatively large and small ε. The distribution of the optimal terminal wealth is not continuous. There is a probability mass at 0 for $\varepsilon > 0$ and there is a probability mass at L for a relatively small ε. Comparing with the case $L \ge z$, a VaR constraint with any $\varepsilon > 0$ under the case $\theta \le L < z$ provides a genuine improvement of the risk management for the loss states. Furthermore, relative to the unconstrained optimal terminal wealth, a VaR constraint shifts the distribution of good states to the left.

Finally, for a given $\varepsilon \in [0, 1]$ in the case $L < \theta$, we identify the exact form of $X^{\pi^*,\varepsilon}(T)$. When $\varepsilon \ge \varepsilon_1^*$, we have

$$P(X^{\pi^*}(T) \ge L) = P(H^{\hat{\nu}}(T) \le \frac{c_z}{y_0}) = 1 - \varepsilon_1^* \ge 1 - \varepsilon,$$

which implies $X^{\pi^*}(T)$ given by (4.4) naturally satisfies the VaR constraint and $X^{\pi^*,\varepsilon}(T) = X^{\pi^*}(T)$. Let

$$\tilde{\varepsilon}_1^* = P(X^{\pi^*,\varepsilon}(T) < L) = P(H^{\hat{\nu}}(T) > \frac{c_{\tilde{z}}}{y_0}),$$

where $X^{\pi^*,\varepsilon}(T)$ is given by (4.8) with $H^* = \frac{c_{\tilde{z}}}{y_0}$ and y_0 is determined by the binding budget constraint (4.3). For $\varepsilon \in [\tilde{\varepsilon}_1^*, \varepsilon_1^*)$, we have $H^* \in (\frac{c_z}{y_0}, \frac{c_{\tilde{z}}}{y_0}]$ and the pension manager insures against the region $\{\frac{c_z}{y_0} \le H^{\hat{\nu}}(T) < H^*\}$ by letting $X^{\pi^*,\varepsilon}(T) = \theta + I_1(y_0 H^{\hat{\nu}}(T))$, so $X^{\pi^*,\varepsilon}(T)$ takes a two-region form given by (4.8). For $\varepsilon \in (0, \tilde{\varepsilon}_1^*)$, we have $H^* \in (\frac{c_{\tilde{z}}}{y_0}, \infty)$ and the pension manager insures against the

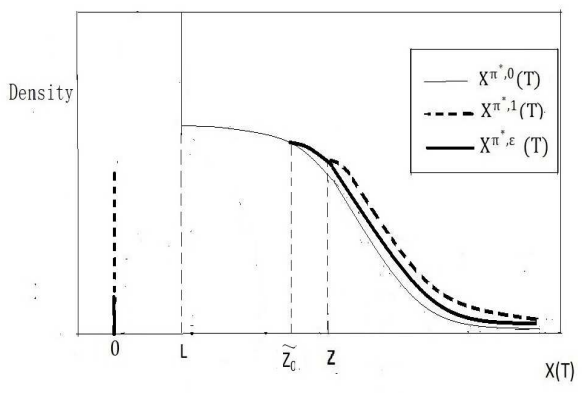

Figure 6: Probability density of optimal terminal wealth with a relatively large ε

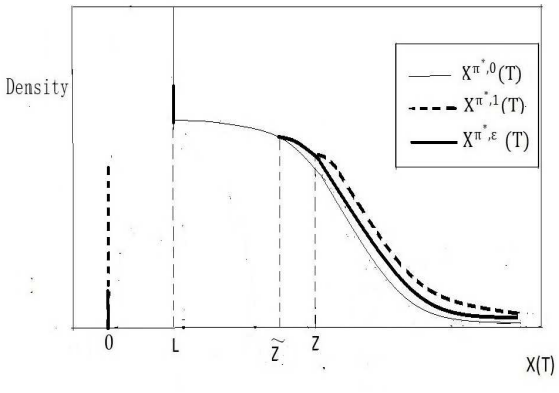

Figure 7: Probability density of optimal terminal wealth with a relatively small ε

regions $\{\frac{c_z}{y_0} \le H^{\hat{\nu}}(T) < \frac{c_{\widetilde{z}}}{y_0}\}$ and $\{\frac{c_{\widetilde{z}}}{y_0} \le H^{\hat{\nu}}(T) < H^*\}$ by letting $X^{\pi^*,\varepsilon}(T) = \theta + I_1(y_0 H^{\hat{\nu}}(T))$ and $X^{\pi^*,\varepsilon}(T) = L$, respectively, so $X^{\pi^*,\varepsilon}(T)$ takes a three-region form given by (4.9).

Figures 6 and 7 depict the distributions of $X^{\pi^*,\varepsilon}(T)$ for $L < \theta$ with relatively large and small ε. Similar to the case $\theta \le L < z$, a VaR constraint with any $\varepsilon > 0$ under the case $L < \theta$ also provides a genuine improvement of the risk management for the loss states.

5 Numerical analysis

In this section, we do some numerical calculations to investigate the influence of a VaR constraint on the optimal terminal wealth.

Assume that

$$U(x) = \begin{cases} -A(\theta - x)^{\gamma_1}, & x < \theta, \\ (x-\theta)^{\gamma}, & x \ge \theta, \end{cases} \tag{5.1}$$

where $A > 1, 0 < \gamma, \gamma_1 < 1$, and (5.1) is the S-shaped utility defined in Kahneman and Tversky (1979). Assume that the financial market consists of three tradable assets, whose price processes are modelled by (2.1)-(2.2) with $n = 2$, and $\sigma_{11} = \sigma_1, \sigma_{12} = 0$, $\sigma_{21} = \rho\sigma_2, \sigma_{22} = \sqrt{1-\rho^2}\sigma_2$, ρ is a correlation coefficient, and $\xi = (\vartheta_{S_1}, \frac{\vartheta_{S_2} - \rho\vartheta_{S_1}}{\sqrt{1-\rho^2}})^\top$, where $\vartheta_{S_1} = \frac{\mu_1 - r}{\sigma_1} > 0, \vartheta_{S_2} = \frac{\mu_2 - r}{\sigma_2} > 0$ are the Sharpe ratios of risky assets S_1 and S_2, respectively. Let $K = [0,\infty)^2$, which means short selling is not allowed. The positive polar cone of K is given by $\tilde{K} = [0,\infty)^2$. To compute $\hat{\xi}$ in (3.27), we need to find the minimizer $\hat{\nu} \in \tilde{K}$ of (3.26). The Kuhn-Tucker optimality condition implies there exists a Lagrange multiplier $u = (u_1, u_2)^\top$ such that $\hat{\nu}$ and u satisfy the following set of equations:

$$\begin{cases} \frac{2}{\sigma_1}(\frac{\hat{\nu}_1}{\sigma_1} + \vartheta_{S_1}) - \frac{2\rho}{(1-\rho^2)\sigma_1}(\frac{\hat{\nu}_2}{\sigma_2} + \vartheta_{S_2} - \rho(\frac{\hat{\nu}_1}{\sigma_1} + \vartheta_{S_1})) - u_1 = 0, \\ \frac{2}{(1-\rho^2)\sigma_2}(\frac{\hat{\nu}_2}{\sigma_2} + \vartheta_{S_2} - \rho(\frac{\hat{\nu}_1}{\sigma_1} + \vartheta_{S_1})) - u_2 = 0, \\ u_i\hat{\nu}_i = 0, u_i \ge 0, \hat{\nu}_i \ge 0, i = 1, 2. \end{cases}$$

Solving the above equation gives

$$\hat{\xi} = \begin{cases} \left(\vartheta_{S_1}, \frac{\vartheta_{S_2} - \rho\vartheta_{S_1}}{\sqrt{1-\rho^2}}\right)^\top, & \vartheta_{S_2} > \rho\vartheta_{S_1}, \vartheta_{S_1} > \rho\vartheta_{S_2}, \\ (\vartheta_{S_1}, 0)^\top, & \rho\vartheta_{S_1} \geq \vartheta_{S_2}, \\ (\rho\vartheta_{S_2}, \sqrt{1-\rho^2}\vartheta_{S_2})^\top, & \rho\vartheta_{S_2} \geq \vartheta_{S_1}. \end{cases}$$

Note that when $K = (-\infty, \infty)^2$, we have $\tilde{K} = \{(0,0)^\top\}$ and $\hat{\xi} = \xi$, which is different from the case $K = [0, \infty)^2$. Therefore, the trading constraint impacts the optimal terminal wealth through the pricing kernel $H^{\hat{\nu}}(T)$.

For all numerical computations, the benchmark data used are the following: $r = 0.02, c(t) = 0.1, x_0 = 35, L = 80, \mu_1 = 0.06, \mu_2 = 0.065, \sigma_1 = 0.3, \sigma_2 = 0.4, \rho = 0.5, T = 40, \theta = 40, A = 2.25, \gamma_1 = 0.2, \gamma = 0.4$. From (3.2), we have $z = 45.3$.

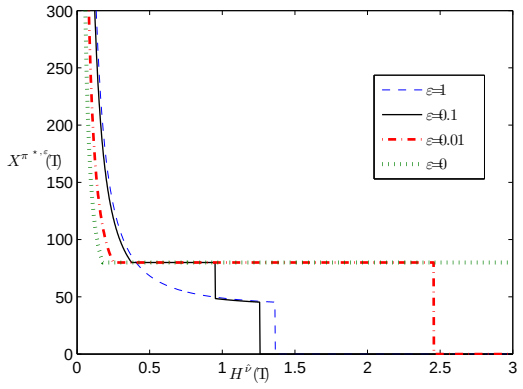

Figure 8: $X^{\pi^*,\varepsilon}(T)$ versus $H^{\hat{\nu}}(T)$ for different ε with $L > z$

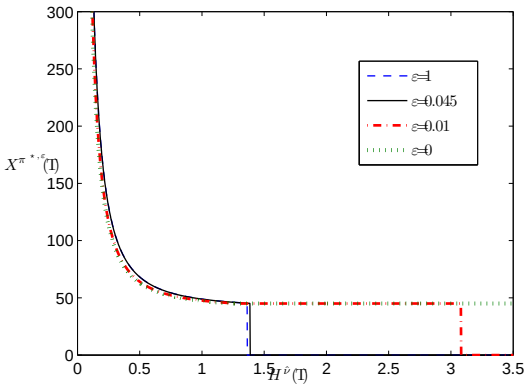

Figure 9: $X^{\pi^*,\varepsilon}(T)$ versus $H^{\hat{\nu}}(T)$ for different ε with $\theta \leq L < z$

Figure 8 shows the optimal terminal value $X^{\pi^*,\varepsilon}(T)$ versus $H^{\hat{\nu}}(T)$ for different ε with $L = 80$. We see that $X^{\pi^*,\varepsilon}(T)$ takes a two-, three- or four-region form according to the value of ε. When ε decreases, the intermediate-states region enlarges while the good- and bad-states regions both shrink. In order to meet the protection level in the intermediate-states region, the VaR constraint leads to a decrease in the optimal terminal wealth of good states. We also note that $X^{\pi^*}(T) = X^{\pi^*,1}(T)$ is dominated by $X^{\pi^*,0.01}(T)$ in the region where $X^{\pi^*}(T) < L$, while $X^{\pi^*,0.1}(T)$ is dominated by $X^{\pi^*}(T)$ in the region where $X^{\pi^*,0.1}(T) < L$, which numerically confirms the result presented in Section 4: when $L \geq z$, a VaR constraint with a relatively small ε can reduce risk exposure in bad market conditions whereas with a relatively large ε can incur heavier losses.

Figure 9 shows the optimal terminal value $X^{\pi^*,\varepsilon}(T)$ versus $H^{\hat{\nu}}(T)$ for different ε with $L = 45$. As shown in Theorem 4.1, $X^{\pi^*,\varepsilon}(T)$ takes a two- or three-region form according to the value of ε in case of $\theta \leq L < z$. It is observed that similar to the case $L > z$, the VaR constraint leads to a decrease in the optimal terminal wealth of good economic states. Furthermore, unlike the case $L \geq z$, $X^{\pi^*}(T)$ in the bad-states region is dominated by the constrained optimal terminal wealth $X^{\pi^*,\varepsilon}(T)$, which is consistent with the result presented in Section 4: when $\theta \leq L < z$, a VaR constraint can improve risk management for the bad-economic states.

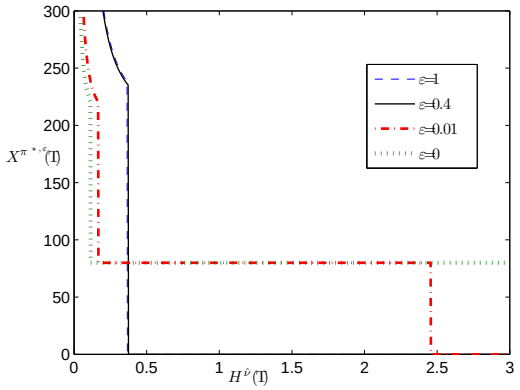

Figure 10: $X^{\pi^{*,\varepsilon}}(T)$ versus $H^{\hat{\nu}}(T)$ for different ε with $L < \theta$

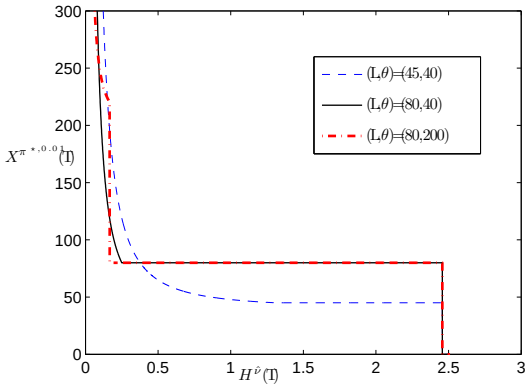

Figure 11: $X^{\pi^{*,0.01}}(T)$ versus $H^{\hat{\nu}}(T)$ for different (L, θ)

Figure 10 plots the optimal terminal value $X^{\pi^{*,\varepsilon}}(T)$ versus $H^{\hat{\nu}}(T)$ for different ε with $L = 80$ and $\theta = 200$. We observe that, as shown in Theorem 4.1, $X^{\pi^{*,\varepsilon}}(T)$ takes a two- or three-region form according to the value of ε in case of $L < \theta$. It is noted that similar to the case $L \geq \theta$, the VaR constraint also leads to a decrease in the optimal terminal wealth of good economic states. Furthermore, we can observe that, as pointed out in Section 4, a VaR constraint can also provide a genuine improvement of the risk management for the loss states in case of $L < \theta$.

Figure 11 presents the optimal terminal wealth $X^{\pi^{*,0.01}}(T)$ versus $H^{\hat{\nu}}(T)$ for different L and θ. We observe that the bad-states region $\{H^{\hat{\nu}}(T) \geq H^*\}$ for different L remains unchanged, due to H^* only depending on ε and the distribution of $H^{\hat{\nu}}(T)$, that the intermediate region increases with L at the expense of the good-states region as more states need to be insured against, and that $X^{\pi^{*,0.01}}(T)$ in the good-states region decreases to attain a higher L in the intermediate region. It is noted that $X^{\pi^{*,0.01}}(T)$ with a lower L is dominated by that with a higher L in the region where both the optimal terminal wealth value are below the higher protection level, which implies that a higher protection level leads to a lower left tail risk for a small ε. We can also see that for $(L, \theta) = (80, 200)$, the manager is not sure to obtain this relatively high reference point. Therefore, the manager takes a more conservative allocation strategy to attain the VaR constraint, which illustrates that a VaR constraint can well protect the members' benefits when the reference point is too high.

Table 1 lists some probabilities, expectations, standard deviations, conditional expectations and quantile values at low end and high end of $X^{\pi^{*,\varepsilon}}(T)$ for different ε with $L = 80$ and different (L, θ) with $\varepsilon = 0.01$. Numerical results further illustrate different forms of the optimal terminal wealth according to the value of ε presented in Section 4: for $\varepsilon \geq \varepsilon^* = 0.366$, the VaR constraint is not binding and $X^{\pi^{*,\varepsilon}}(T) = X^{\pi^*}(T)$ takes either 0 or in (z, ∞); for $0.061 = \varepsilon_1^* < \varepsilon < 0.366$, $X^{\pi^{*,\varepsilon}}(T)$ takes either 0 or L or in (z, L) or in (L, ∞); for $0 < \varepsilon \leq 0.061$, $X^{\pi^{*,\varepsilon}}(T)$ takes either 0 or L or in (L, ∞); and for $\varepsilon = 0$, $X^{\pi^{*,0}}(T)$ takes either L or in (L, ∞). We observe that for a fixed θ, the probability $P(X^{\pi^{*,\varepsilon}}(T) = L)$ decreases in ε and increases in L, which is in line with what has been discussed in Figures 8 and 11: when ε decreases or L increases, more states need to be insured against and therefore the intermediate-states region grows at the expense of unmodified regions.

Table 1: (conditional) expectations, standard deviations, quantile values and probabilities

(ε, L)	ε $(L = 80)$				(L, θ) $(\varepsilon = 0.01)$		
	1	0.1	0.01	0	(45,40)	(80,40)	(80,200)
mean	248.68	227.28	147.20	111.85	224.78	111.85	130.01
std dev	627.21	547.53	275.04	150.61	550.56	271.04	108.29
0.1 quantile	49.67	48.46	80	80	48.5	80	80
0.9 quantile	506.39	448.09	247.44	155.71	449.78	247.44	253.60
$P(X^{\pi^*,\varepsilon}(T) = 0)$	0.047	0.056	0.01	0	0.01	0.01	0.01
$P(X^{\pi^*,\varepsilon}(T) \in (z, L))$	0.319	0.044	0	0	0	0	0
$P(X^{\pi^*,\varepsilon}(T) = L)$	0	0.3	0.567	0.719	0.041	0.567	0.73
$P(X^{\pi^*,\varepsilon}(T) > L)$	0.634	0.6	0.423	0.281	0.949	0.423	0.26
$E(X^{\pi^*,\varepsilon}(T) \mid X^{\pi^*,\varepsilon}(T) > L)$	392.18	335.18	240.59	193.25	234.97	240.59	275.11

We observe from Table 1 that the expectation and the standard deviation decrease in L and increase in ε, respectively, which implies that in order to achieve the VaR constraint with a smaller ε or a higher protection level L, the manager takes more prudent strategies such that $X^{\pi^*,\varepsilon}(T)$ becomes less volatile. The quantile value at high end decreases in L and increases in ε, consistent with the observations from Figures 8 and 11: the optimal terminal wealth in good states decreases in L and increases in ε. For $\varepsilon = 0.01$, the quantile value at low end increases with L, since the optimal terminal wealth with a lower L in the loss states is dominated by that with a higher L for a relatively low ε, which has been revealed in Figure 11. However, there is no monotonicity in ε for the quantile value at low end since a VaR constraint leads to more losses for a relatively high ε. We note that different from the expectation, a stronger protection (with a higher L or a smaller ε) does not necessarily lead to a decrease in the conditional expectation. For $L = 80$, a smaller ε leads to a lower right tail risk and a smaller conditional expectation, consistent with the observation from Figure 8: the optimal terminal wealth with a lower ε in good states is dominated by that with a higher ε. However, for $\varepsilon = 0.01$, the conditional expectation increases with L. The reason is that as observed from Figure 11, changing L does not lead to monotone changes in the optimal terminal wealth in good states. We can also see that for a fixed $L = 80$, as explained in Figure 11, a relatively high reference point θ leads the manager to invest much more money in the riskless asset since the manager is not sure to achieve it. Therefore, the optimal terminal wealth with $\theta = 200$ has a less standard deviation, a lower probability $P(X^{\pi^*,\varepsilon}(T) > L)$ and a higher probability $P(X^{\pi^*,\varepsilon}(T) = L)$. However, the expectation and the 0.9 quantile value are not a monotonic function of θ. This is due to two effects: the first effect is that $X^{\pi^*,\varepsilon} = \theta + I_1(y_0 H^{\hat{\nu}}(T))$ is an increasing function of θ and a decreasing function of y_0. The second effect is that y_0 increases with θ.

Figure 12 represents the relationship between the constrained and unconstrained expectations of the optimal terminal wealth and σ_1. We can see that the unconstrained expectation is greater than the constrained one, in line with the observations in Table 1, which is due to the fact that the VaR constraint makes the manager to invest less in the risky assets. Both expectations are decreasing functions of σ_1. The reason is that when the volatility σ_1 increases, the Sharpe ratio decreases, which leads to less investment in the first risky asset and overall lower expectation of

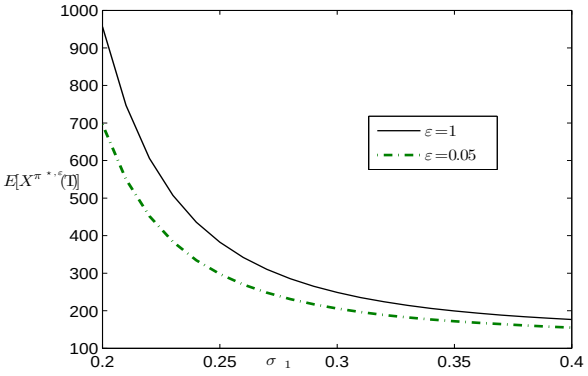

Figure 12: $E[X^{\pi^{*,\varepsilon}}(T)]$ versus σ_1 for different (ε, L)

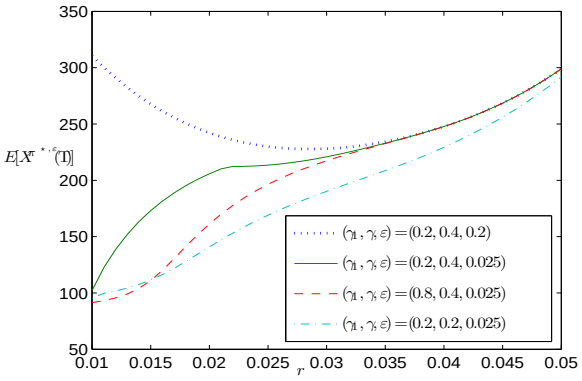

Figure 13: $E[X^{\pi^{*,\varepsilon}}(T)]$ versus r for different $(\gamma_1, \gamma, \varepsilon)$

the optimal terminal wealth, consistent with economic intuition.

Figure 13 displays the impact of r on the expectation of the optimal terminal wealth for different γ_1, γ and ε. The parameters γ_1 and γ control the curvature of the S-shaped utility in the domain of losses and gains, respectively. We can see that the expectation decreases with γ_1 while increases with γ. These two parameters show diversified characteristics of the investor's behavior. Specifically, the manager has different attitudes towards risk in different scenarios. The effect of greater risk seeking in the domain of losses, measured by a lower γ_1, leads to more investment in risky assets, whereas the effect of higher risk aversion in the domain of gains, measured by a lower γ, leads to less investment in risky assets. The relation of $E[X^{\pi^{*,\varepsilon}}(T)]$ and r is more complicated. An increase in r is good for investment in riskless assets but bad for investment in risky assets due to decreasing Sharpe ratio. The overall impact of r on $E[X^{\pi^{*,\varepsilon}}(T)]$ is determined jointly by these two opposite effects, depending on the level of VaR constraint and other parameters. For example, in the case of $\varepsilon = 0.05$, a strong VaR constraint, $E[X^{\pi^{*,0.025}}(T)]$ increases with r, but in the case of $\varepsilon = 0.2$, a weak VaR constraint, $E[X^{\pi^{*,0.2}}(T)]$ decreases first and then increases with r, displaying a "smile" shaped relation, see Figure 13.

6 Conclusions

In this paper, we investigate the optimal portfolio selection problem under loss aversion and with trading and VaR constraints. The analysis is applied to study the optimal asset allocation for a DC pension plan. We solve the problem in two steps: First, we solve the unconstrained optimization problem (2.8) with a fixed Lagrange multiplier. By using a concavification technique and a dual control method, we find the unique pricing kernel in the presence of closed convex cone control constraints and derive the optimal terminal wealth by solving a terminal static optimization problem in the same way as in the complete market case. We also derive the representations of the optimal wealth process and the optimal trading strategy in terms of the dual value function, its derivatives, and the optimal dual state process via the relationship between the primal and dual optimization problems. Second, we solve the constrained optimization problem (2.7) by finding the solutions of two coupled nonlinear equations (binding budget and VaR constraints). We propose a simple

algorithm to constructively compute the initial dual state value and the Lagrange multiplier which are used to find the optimal terminal wealth. Theoretical and numerical results show that the VaR constraint has significant impact on the distribution of optimal terminal wealth and may provide an effective improvement in bad states due to loss aversion. There remain many open questions for optimal allocation of a DC pension plan, for example, the DC contribution rate $c(t)$ may be modelled by a stochastic process, the financial market may include some credit-related products, the S-shaped utility maximization may have multiple terminal and intertemporal VaR constraints. We leave these and other questions for future research.

A Appendix

We give a useful result which is used in constructing the concave envelope of $\tilde{U}_\lambda$.

Lemma A.1. Let $0 \le z_1 \le z_2 < z_3$ be given constants and f be right continuous on $[z_1, \infty)$ satisfying

1. f is concave on intervals $[z_1, z_2)$ and (z_3, ∞),
2. $f(x) \le k(x - z_2) + f(z_2)$ on $[z_1, z_3]$ with $k = \frac{f(z_3)-f(z_2)}{z_3-z_2} \ge f'_+(z_3) > 0$,

where f'_+ is the right derivative of f at z_3. Then the concave envelope of f is given by

$$f^c(x) = \begin{cases} f(x), & z_1 \le x < z_2, \\ k(x - z_2) + f(z_2), & z_2 \le x < z_3, \\ f(x), & x \ge z_3. \end{cases} \tag{A.1}$$

In particular, for $z_1 = z_2$, the concave envelope of f is given by

$$f^c(x) = \begin{cases} k(x - z_1) + f(z_1), & z_1 \le x < z_3, \\ f(x), & x \ge z_3. \end{cases} \tag{A.2}$$

Proof. By definition $f^c \ge f$. Let g be concave with $g \ge f$. Then $g \ge f^c$ on $[z_1, z_2) \cup [z_3, \infty)$. Assume $x = uz_2 + (1 - u)z_3 \in (z_2, z_3)$ for $u \in (0, 1)$. By concavity of g, we have

$$g(x) \ge ug(z_2) + (1 - u)g(z_3) \ge uf(z_2) + (1 - u)(k(z_3 - z_2) + f(z_2)) = f(z_2) + k(x - z_2) = f^c(x).$$

It remains to prove that f^c is concave. Let $z_1 \le x_0 < x_1 < \infty$ and $x_u = ux_0 + (1 - u)x_1$ with $u \in (0, 1)$. It is easy to conclude that if $x_1 \le z_2$, or $x_0 \ge z_3$, and or $z_2 \le x_0 < x_1 \le z_3$, then

$$f^c(x_u) \ge uf^c(x_0) + (1 - u)f^c(x_1).$$

If $z_2 \le x_0 \le z_3 \le x_1$, then we have

$$f^c(x_0) = k(x_0 - z_2) + f(z_2), f^c(z_3) = k(z_3 - z_2) + f(z_2). \tag{A.3}$$

Note that by concavity

$$f^c(x_1) = f(x_1) \le f'_+(z_3)(x_1 - z_3) + f(z_3) \le k(x_1 - z_3) + f(z_3) = k(x_1 - z_2) + f(z_2). \tag{A.4}$$

Equations (A.3) and (A.4) imply that the slope of the line through $(z_3, f^c(z_3))$ and $(x_1, f^c(x_1))$ is less than the slope of the line through $(x_0, f^c(x_0))$ and $(x_1, f^c(x_1))$, that is,

$$\frac{f^c(x_1) - f^c(x_0)}{x_1 - x_0} \geq \frac{f^c(x_1) - f^c(z_3)}{x_1 - z_3}. \tag{A.5}$$

If $x_u \in (z_2, z_3)$, then

$$\begin{aligned} f^c(x_u) &= k(x_u - z_2) + f(z_2) = k(u(x_0 - z_2) + (1-u)(x_1 - z_2)) + f(z_2) \\ &= uf^c(x_0) + (1-u)(k(x_1 - z_2) + f(z_2)) \geq uf^c(x_0) + (1-u)f^c(x_1), \end{aligned}$$

where the last inequality follows from (A.4).

If $x_u \in (z_3, \infty)$, then

$$\begin{aligned} f^c(x_u) &\geq \frac{f^c(x_1) - f^c(z_3)}{x_1 - z_3}(x_u - x_1) + f^c(x_1) \\ &\geq \frac{f^c(x_1) - f^c(x_0)}{x_1 - x_0}(x_u - x_1) + f^c(x_1) \\ &= uf^c(x_0) + (1-u)f^c(x_1). \end{aligned}$$

where the first inequality holds since $f^c(x)$ is concave on $[z_3, \infty)$, and the second inequality follows from (A.5). Therefore, we can conclude that f^c is concave on $[z_2, \infty)$.

If $z_1 \leq x_0 \leq z_2 \leq x_1$, then we have

$$f^c(x_0) \leq k(x_0 - z_2) + f^c(z_2), f^c(z_2) < f^c(x_1) \leq k(x_1 - z_2) + f^c(z_2),$$

which implies that the slope of the line through $(x_0, f^c(x_0))$ and $(x_1, f^c(x_1))$ is less than the slope of the line through $(x_0, f^c(x_0))$ and $(z_2, f^c(z_2))$, and is greater than the slope of the line through $(z_2, f^c(z_2))$ and $(x_1, f^c(x_1))$, that is,

$$\frac{f^c(x_1) - f^c(z_2)}{x_1 - z_2} \leq \frac{f^c(x_1) - f^c(x_0)}{x_1 - x_0} \leq \frac{f^c(z_2) - f^c(x_0)}{z_2 - x_0}. \tag{A.6}$$

If $x_u \in (z_2, x_1)$, then

$$\begin{aligned} f^c(x_u) &\geq \frac{f^c(x_1) - f^c(z_2)}{z_2 - x_1}(x_u - x_1) + f^c(x_1) \\ &\geq \frac{f^c(x_1) - f^c(x_0)}{z_2 - x_1}(x_u - x_1) + f^c(x_1) \\ &= uf^c(x_0) + (1-u)f^c(x_1), \end{aligned}$$

where the first inequality holds from the fact that f^c is concave on $[z_2, \infty)$, and the second inequality follows from (A.6).

If $x_u \in (x_0, z_2)$, then by concavity

$$\begin{aligned} f^c(x_u) &\geq \frac{f^c(z_2) - f^c(x_0)}{z_2 - x_0}(x_u - x_0) + f^c(x_0) \\ &\geq \frac{f^c(x_1) - f^c(x_0)}{x_1 - x_0}(x_u - x_0) + f^c(x_0) \\ &= uf^c(x_0) + (1-u)f^c(x_1), \end{aligned}$$

which concludes the proof. □

Proof of Proposition 3.1

Proof. Since $L \geq z$, we have $c_z \geq k_0 = \frac{U_1(L-\theta)+U_2(\theta)}{L} \geq c_L$, which yields $k_\lambda \geq k_0 \geq c_L$ for $\lambda \geq 0$.

Case I: When $k_\lambda > c_z$, we have that $\tilde{U}_\lambda(x) \leq k_\lambda x - U_2(\theta)$ for $0 \leq x \leq L$ and $\tilde{U}_\lambda(x)$ is concave on (L, ∞). Then Lemma A.1 yields the concave envelope of $\tilde{U}_\lambda$ given by (3.7)(see Figure 14). It is easy to obtain the superdifferential of $\tilde{U}^c_\lambda(x)$ as follows:

$$(\tilde{U}^c_\lambda)'(x) = \begin{cases} [k_\lambda, \infty), & x = 0, \\ \{k_\lambda\}, & 0 < x < L, \\ [c_L, k_\lambda), & x = L, \\ \{\tilde{U}'_\lambda(x)\}, & x > L. \end{cases} \tag{A.7}$$

Then we can find the point $x^{*,\lambda}(y) \in \{x | \tilde{U}_\lambda(x) = \tilde{U}^c_\lambda(x)\}$ solving both (3.1) and (3.4) for which 0 is in the superdifferential of $\tilde{U}^c_\lambda(x) - xy$ given by (3.8).

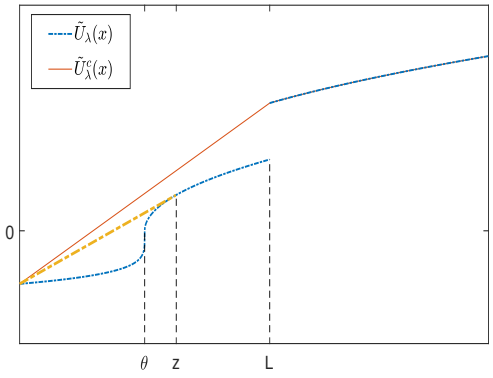

Figure 14: Concave envelope of $\tilde{U}_\lambda$,$\tilde{U}^c_\lambda$ for $k_\lambda > c_z$

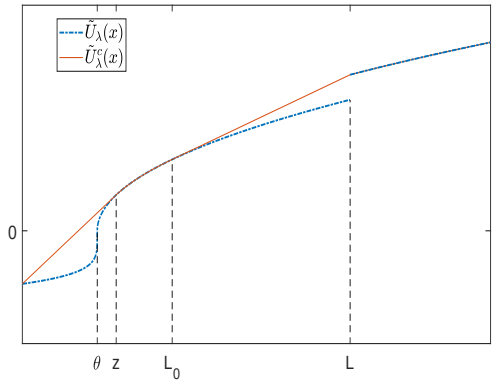

Figure 15: Concave envelope of $\tilde{U}_\lambda$,$\tilde{U}^c_\lambda$ for $k_\lambda \leq c_z$

Case II: When $c_L \leq k_\lambda \leq c_z$, we let L_0 be the tangent point of the straight line starting at $(L, U_1(L-\theta)+\lambda)$ to the curve $U_1(x-\theta), \theta \leq x \leq L$ (see Figure 15). Straightforward calculation shows that there exists a unique solution L_0 in the interval $[z, L]$ to the equation (3.11). Lemma A.1 gives the concave envelope of $\tilde{U}_\lambda$ represented by (3.9). Similar to deriving (3.8), one can easily find $x^{*,\lambda}(y)$ given by (3.10). □

Proof of Proposition 3.2

Proof. Since $\theta \leq L < z$, we have $k_0 = \frac{U_1(L-\theta)+U_2(\theta)}{L} < c_L$.

Case I: If $k_\lambda \geq c_L$, then the expressions for $\tilde{U}^c_\lambda(x)$ and $x^{*,\lambda}(y)$ are the same as those in Case I of Proposition 3.1(see Figure 16).

Case II: If $k_\lambda < c_L$, then we let z_0 be the tangent point of the straight line starting at $(0, -U_2(\theta))$ to the curve $U_1(x-\theta)+\lambda, x \geq \theta$ (see Figure 17). It is easy to verify that there exists a unique solution $L < z_0 \leq z$ to the equation (3.15) and $\tilde{U}_\lambda(x) \leq c_{z_0} x - U_2(\theta)$ for $0 \leq x < z_0$. From Lemma A.1, the concave envelope of $\tilde{U}_\lambda$ is given by (3.13). Then one can easily derive $x^{*,\lambda}(y)$ given by (3.14) by using the same arguments as in deriving (3.8). □

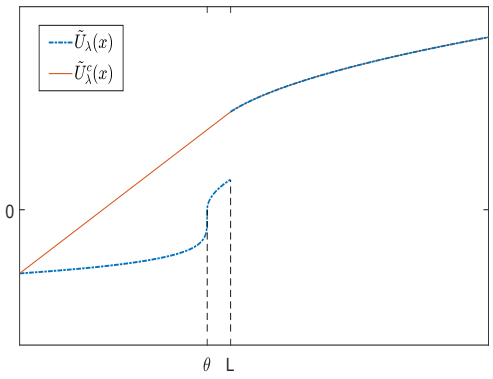

Figure 16: Concave envelope of $\tilde{U}_\lambda$,$\tilde{U}^c_\lambda$ for $k_\lambda \geq c_L$

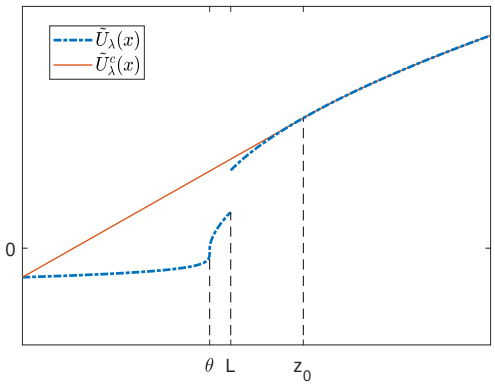

Figure 17: Concave envelope of $\tilde{U}_\lambda$,$\tilde{U}^c_\lambda$ for $k_\lambda < c_L$

Proof of Proposition 3.3

Proof. Since $L < \theta < \tilde{z} < z$, we have $k_0 < c_z < c_{\tilde{z}}$.

Case I: If $k_\lambda > c_{\tilde{z}}$, then Lemma A.1 yields that $\tilde{U}^c_\lambda(x)$ given by (3.17) (see Figure 18). Similar to deriving (3.8), we can find the maximizer $x^{*,\lambda}(y) \in \{x|\tilde{U}_\lambda(x) = \tilde{U}^c_\lambda(x)\}$ given by (3.18) solving both for which 0 is in the superdifferential of $\tilde{U}^c_\lambda(x) - xy$ given by (3.8).

Case II: If $k_\lambda \leq c_{\tilde{z}}$, then we let $\tilde{z}_0$ be the tangent point of the straight line starting at $(0, -U_2(\theta))$ to the curve $U_1(x-\theta)+\lambda, x \geq \theta$. It is easy to verify that there exists a unique solution $z \leq \tilde{z}_0 \leq \tilde{z}$ to (3.15) and $\tilde{U}_\lambda(x) \leq c_{\tilde{z}_0} x - U_2(\theta)$ for $0 \leq x < \tilde{z}_0$. From Lemma A.1, the concave envelope of $\tilde{U}_\lambda$ is given by (3.13) (see Figure 19). Then from Case II in Proposition 3.2, the maximizer $x^{*,\lambda}(y)$ is given by (3.14).

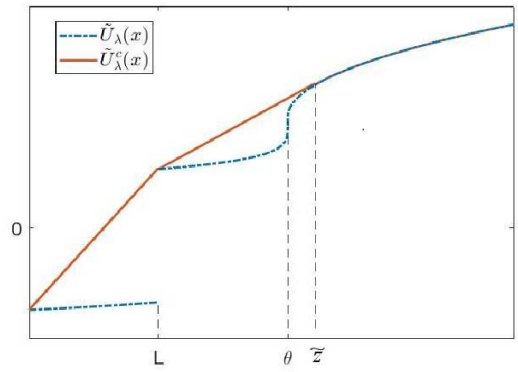

Figure 18: Concave envelope of $\tilde{U}_\lambda$,$\tilde{U}^c_\lambda$ for $k_\lambda \geq c_{\tilde{z}}$

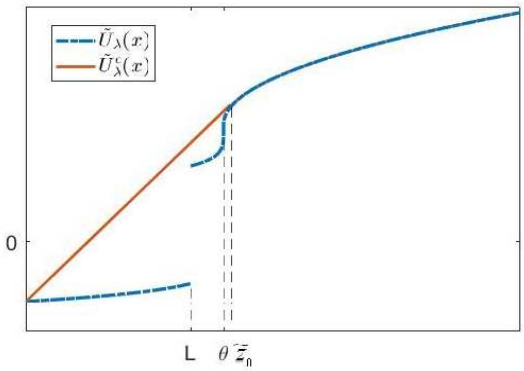

Figure 19: Concave envelope of $\tilde{U}_\lambda$,$\tilde{U}^c_\lambda$ for $k_\lambda < c_{\tilde{z}}$

□

The optimal wealth and portfolio processes

Using Theorem 3.7, we can easily derive the optimal wealth process $X^{\pi^{*,\varepsilon}}$ and the optimal control process $\pi^{*,\varepsilon}$. The expressions are different but similar for $L \geq z$, $\theta \leq L < z$ and $L < \theta$. To simplify notations and avoid repetitions, we denote Φ the cumulative standard normal distribution function, φ the standard normal density function, and define the following functions:

$$\begin{aligned} h(u,w,t) &\equiv \int_{-\infty}^{d_1(u,w,t)} (\theta + I_1(we^{\eta(x,t)}))e^{\eta(x,t)}\varphi(x)dx \\ \overline{h}(u,w,t) &\equiv \int_{-\infty}^{d_1(u,w,t)} (\theta + I_1(we^{\eta(x,t)}))e^{\eta(x,t)}(x-1)\varphi(x)dx \\ \psi(u,w,t) &\equiv (\alpha(t)-1)\Phi(d_3(u,w,t)) - \varphi(d_3(u,w,t)), \end{aligned}$$

where $d_3(l,w,t) = (\ln(l/w) + \beta(t))/\alpha(t)$, $\alpha(t), \beta(t)$ are defined by (3.37), $\eta(t,x) = \alpha(t)(x-1) - \beta(t)$, $d_1(l,w,t) = d_3(l,w,t) + 1$. Denote by $L(t) = Le^{-r(T-t)}$. We can now give an explicit characterization of the optimal wealth process $X^{\pi^{*,\varepsilon}}$ and the optimal control process $\pi^{*,\varepsilon}$.

Proposition A.1. *Assume that $\hat{\nu} \in \tilde{K}$ is the unique minimizer of (3.26), $\hat{\xi} \neq 0$, z is the solution to (3.2), and $x_0 + C(0) > E[LH^{\hat{\nu}}(T)1_{\{H^{\hat{\nu}}(T) \leq H^*\}}]$, where H^* solves (4.1) for $0 \leq \varepsilon \leq 1$. Then for the optimization problem (2.7), the optimal wealth $X^{\pi^{*,\varepsilon}}(t)$ and the optimal control $\pi^{*,\varepsilon}(t)$ at time t for $0 \leq t \leq T$ are given by*

$$\begin{aligned} X^{\pi^{*,\varepsilon}}(t) &= A(y_0, y_0 H^{\hat{\nu}}(t), t) \\ \pi^{*,\varepsilon}(t) &= -(\sigma^\top)^{-1}\hat{\xi}B(y_0, y_0 H^{\hat{\nu}}(t), t) \in K, \end{aligned}$$

where y_0 is determined by the equation

$$A(y_0, y_0, 0) = x_0$$

and A and B are given by the following:

1. *If $L \geq z$, then*

$$\begin{aligned} A(y_0,w,t) &= 1_{\{H^* \leq \frac{c_L}{y_0}\}} h(c_z,w,t) + 1_{\{\frac{c_L}{y_0} < H^* < \frac{c_z}{y_0}\}} \Big(h(c_z,w,t) - h(y_0 H^*, w, t) \Big) \\ &\quad + 1_{\{H^* > \frac{c_L}{y_0}\}} \Big(L(t)(\Phi(d_3(y_0H^*,w,t)) - \Phi(d_3(c_L,w,t))) + h(c_L,w,t) \Big) - C(t) \\ B(y_0,w,t) &= 1_{\{H^* \leq \frac{c_L}{y_0}\}} \overline{h}(c_z,w,t) + + 1_{\{\frac{c_L}{y_0} < H^* < \frac{c_z}{y_0}\}} \Big(\overline{h}(c_z,w,t) - \overline{h}(y_0 H^*, w, t) \Big) \\ &\quad + 1_{\{H^* > \frac{c_L}{y_0}\}} \Big(L(t)(\psi(y_0H^*,w,t) - \psi(c_L,w,t)) + \overline{h}(c_L,w,t) \Big). \end{aligned}$$

2. *If $\theta \leq L < z$, then*

$$\begin{aligned} A(y_0,w,t) &= 1_{\{H^* \leq \frac{c_z}{y_0}\}} h(c_z,w,t) + 1_{\{H^* \geq \frac{c_L}{y_0}\}} L(t) \Big(\Phi(d_3(y_0H^*,w,t)) - \Phi(d_3(c_L,w,t)) \Big) \\ &\quad + 1_{\{H^* > \frac{c_z}{y_0}\}} h(\min\{c_L, y_0 H^*\}, w, t) - C(t) \\ B(y_0,w,t) &= 1_{\{H^* \leq \frac{c_z}{y_0}\}} \overline{h}(c_z,w,t) + 1_{\{H^* \geq \frac{c_L}{y_0}\}} L(t) \Big(\psi(y_0H^*,w,t) - \psi(c_L,w,t) \Big) \\ &\quad + 1_{\{H^* > \frac{c_z}{y_0}\}} \overline{h}(\min\{c_L, y_0 H^*\}, w, t). \end{aligned}$$

3. If $L < \theta$, then

$$\begin{aligned} A(y_0,w,t) &= 1_{\{H^*\le \frac{c_z}{y_0}\}}h(c_z,w,t) + 1_{\{H^*>\frac{c_z}{y_0}\}}L(t)\Big(\Phi(d_3(y_0H^*,w,t)) - \Phi(d_3(c_{\tilde{z}},w,t))\Big) \\ &\quad + 1_{\{H^*>\frac{c_z}{y_0}\}}h(\min\{c_{\tilde{z}},y_0H^*\},w,t) - C(t) \\ B(y_0,w,t) &= 1_{\{H^*\le \frac{c_z}{y_0}\}}\overline{h}(c_z,w,t) + 1_{\{H^*>\frac{c_z}{y_0}\}}L(t)\Big(\psi(y_0H^*,w,t) - \psi(c_{\tilde{z}},w,t)\Big) \\ &\quad + 1_{\{H^*>\frac{c_z}{y_0}\}}\overline{h}(\min\{c_{\tilde{z}},y_0H^*\},w,t). \end{aligned}$$

Acknowledgments The authors are grateful to five anonymous reviewers whose constructive criticisms, comments and suggestions have helped to improve the paper of the previous two versions. The first author expresses sincere gratitude to the department of mathematics of Imperial College for hosting her one year academic visit during which this paper was written. The research of Yinghui Dong was supported by the NSF of Jiangsu Province (Grant No. BK20170064), the NNSF of China (Grant No. 11771320), QingLan Project and the scholarship of Jiangsu Overseas Visiting Scholar Program.

References

[1] Basak, S. (1995). A general equilibrium model of portfolio insurance. *Review of Financial Studies*, 8, 1059-1090.

[2] Basak, S., & Shapiro, A. (2001). Value-at-risk-based risk management: optimal policies and asset prices. *Review of Financial Studies*, 14, 371-405.

[3] Berkelaar, A. B., Kouwenberg, R., & Post, T. (2004). Optimal portfolio choice under loss aversion. *Review of Economics and Statistics*, 86, 973-987.

[4] Bernard, C., He, X., Yan, J. A., & Zhou, X. Y. (2015). Optimal insurance design under rank-dependent expected utility. *Mathematical Finance*, 25, 154-186.

[5] Bian, B.J., Miao, S., & Zheng, H. (2011). Smooth value functions for a class of nonsmooth utility maximization problems. *SIAM Journal on Financial Mathematics*, 2, 727-747.

[6] Bichuch, M., & Sturm, S. (2014). Portfolio optimization under convex incentive schemes. *Finance and Stochastics*, 18, 873-915.

[7] Blake, D., Wright, D., & Zhang, Y. M. (2013). Target-driven investing: optimal investment strategies in defined contribution pension plans under loss aversion. *Journal of Economic Dynamics and Control*, 37, 195-209.

[8] Blake, D., Wright, D., & Zhang, Y. M. (2014). Age-dependent investing: optimal funding and investment strategies in defined contribution pension plans when members are rational life cycle financial planners. *Journal of Economic Dynamics and Control*, 38, 105-124.

[9] Boulier, J. F., Huang, S. J., & Taillard, G. (2001). Optimal management under stochastic interest rates: the case of a protected defined contribution pension fund. *Insurance Mathematics and Economics*, 28, 173-189.

[10] Boyle, P., & Tian, W. (2007). Portfolio management with constraints. *Mathematical Finance*, 17, 319-344.

[11] Broadie, M., & Glasserman, P. (1996). Estimating security price derivatives using simulation. *Management Science*, 42, 269-285.

[12] Cairns, A. J., Blake, D., & Dowd, K. (2006). Stochastic lifestyling: optimal dynamic asset allocation for defined contribution pension plans. *Journal of Economic Dynamics and Control*, 30, 843-877.

[13] Carpenter, J. N. (2000). Does option compensation increase managerial risk appetite?. *Journal of Finance*, 55, 2311-2331.

[14] Chen, A., Hieber, P., & Nguyen, T. (2019). Constrained non-concave utility maximization: An application to life insurance contracts with guarantees. *European Journal of Operational Research*, 273, 1119-1135.

[15] Chen, Z., Li, Z. F., Zeng, Y., & Sun, J. Y. (2017). Asset allocation under loss aversion and minimum performance constraint in a DC pension plan with inflation risk. *Insurance Mathematics and Economics*, 75, 137-150.

[16] Cox, J., & Huang, C. F. (1989). Optimun consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49, 33-83.

[17] Cvitanic, J.,& Karatzas, I. (1992). Convex duality in constrained portfolio optimization. *The Annals of Applied Probability*, 2(4), 767-818.

[18] Dong, Y., & Zheng, H. (2019). Optimal investment of DC pension plan under short-selling constraints and portfolio insurance. *Insurance Mathematics and Economics*, 85, 47-59.

[19] El Karoui, N., Jeanblanc, M., & Lacoste, V. (2005). Optimal portfolio management with American capital guarantee. *Journal of Economic Dynamics and Control*, 29, 449-468.

[20] Follmer, H., & Kramkov, D. (1997). Optional decompositions under constraints. *Probability Theory and Related Fields*, 109, 1-25.

[21] Guan, G. H., & Liang, Z. X. (2016). Optimal management of DC pension plan under loss aversion and value-at-risk constraints. *Insurance Mathematics and Economics*, 69, 224-237.

[22] He, X. D., & Kou, S. (2018). Profit sharing in hedge funds. *Mathematical Finance*, 28, 50-81.

[23] He, X. D., & Zhou, X. Y. (2011). Portfolio choice via quantiles. *Mathematical Finance*, 21, 203-231.

[24] Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47, 263-292.

[25] Karatzas, I., Lehoczky, J. P., Sethi, S. P., & Shreve, S. E. (1986). Explicit solution of a general consumption/investment problem. *Mathematics of Operations Research*, 11, 261-294.

[26] Kraft, H., & Steffensen, M. (2013) A dynamic programming approach to constrained portfolios. *European Journal of Operational Research*, 229, 453-461.

[27] Pliska, S. (1986). A stochastic calculus model of continuous trading: optimal portfolios. *Mathematics of Operations Research*, 11, 371-382.

[28] Reichlin, C. (2013). Utility maximization with a given pricing measure when the utility is not necessarily concave. *Mathematics and Financial Economics*, 7, 531-556.

[29] Tversky, A., & Kahneman, D.(1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297-323.

[30] Xu, G. L., & Shreve, S. E. (1992). A duality method for optimal consumption and investment under short-selling prohibition. II. constant market coefficients. *Annals of Applied Probability*, 2, 314-328.

[31] Yao, H. X., Yang, Z., & Chen, P. (2013). Markowitz's mean-variance defined contribution pension fund management under inflation: a continuous-time model. *Insurance Mathematics and Economics*, 53, 851-863.

[32] Zeng, Y., Li, D. P., Chen, Z., & Yang, Z. (2018) Ambiguity aversion and optimal derivative-based pension investment with stochastic income and volatility. *Journal of Economic Dynamics and Control*, 88, 70-103.

[33] Zhang, A. H., Korn, R., & Ewald, C. O. (2007). Optimal management and inflation protection for defined contribution pension plans. *Blatter der DGVFM*, 28, 239-258.

[34] Zhang, A. H., & Ewald, C. O. (2010). Optimal investment for a pension fund under inflation risk. *Mathematical Methods of Operations Research*, 71, 353-369.