UTILITY-DEVIATION-RISK PORTFOLIO SELECTION

K. C. WONG*†, S. C. P. YAM ‡, AND H. ZHENG †

Abstract. We here provide a comprehensive study of the utility-deviation-risk portfolio selection problem. By considering the first-order condition for the corresponding objective function, we first derive the necessary condition that the optimal terminal wealth satisfying two mild regularity conditions solves for a primitive static problem, called *Nonlinear Moment Problem*. We then illustrate the application of this general necessity result by revisiting the non-existence of the optimal solution for the mean-semivariance problem. Secondly, we establish an alternative version of the verification theorem serving as the sufficient condition that the solution, which satisfies another mild condition different from that for necessity, of the Nonlinear Moment Problem is the optimal terminal wealth of the original utility-deviation-risk portfolio selection problem. We then apply this general sufficiency result to revisit the various well-posed mean-risk problems already known in the literature, and to also establish the existence of the optimal solutions for both utility-downside-risk and utility-strictly-convex-risk problems under the assumption that the underlying utility satisfies the Inada Condition. To the best of our knowledge, the positive answers to the latter two problems have long been absent in the literature. In particular, the existence result in the utility-downside-risk problem is in contrast to the well-known non-existence of an optimal solution for the mean-downside-risk problem. As a corollary, the existence result in utility-semivariance problem allows us to utilize the semivariance as a proper risk measure in the classical portfolio management paradigm.

Key words. Nonlinear Moment Problem; Deviation risk function; Downside deviation risk; Portfolio selection.

AMS subject classifications. 91G10, 90C46, 93E20

1. Introduction.

Since its first introduction in Markowitz [25], the portfolio selection problem has become one of the most important research topics in finance. Expected utility and mean-variance are two common criteria for evaluating portfolio performance. For example, Merton [30] and Samuelson [37] investigated utility maximization problems in continuous time and multiperiod settings respectively, by formulating them as a stochastic control problem. The advantage of that formulation allows a direct application of dynamic programming or via HJB by invoking the inherent tower property; see [10, 30, 39] for details. In addition to the use of the Dynamic Programming Principle, the martingale method can be applied to solve for this utility maximization problem in a complete market, where the existence of the optimal solution can be shown by using duality method, and then utilize the Clark-Ocone formula to seek for the optimal weight; see [9, 15, 16, 19, 35].

Apart from utility optimization, many scholars use the mean-variance criteria for evaluating the portfolio performance. For example, Markowitz [26] and Merton [31] aimed to minimize the variance of the portfolio return subject to a constraint on the expected return of the terminal wealth, and they also established the efficient frontier. The advantage of using mean-variance criteria is due to its relative computational simplicity and convenience in selling in bulk to accommodate market demand; indeed,

---

*Department of Statistics and Actuarial Science, The University of Hong Kong, Hong Kong (wongkc89@hku.hk). The first author was supported by The University of Hong Kong, and the present work constitutes a part of his work for his postgraduate dissertation.

†Department of Mathematics, Imperial College London, London SW7 2AZ, UK (w.kwok-chuen13@imperial.ac.uk, h.zheng@imperial.ac.uk).

‡Department of Statistics, The Chinese University of Hong Kong, Hong Kong (scpyam@sta.cuhk.edu.hk). The second author was supported by The Hong Kong RGC GRF 14301015 with the project title: “Advance in Mean Field Theory” and The Chinese University of Hong Kong Direct Grant 2015/2016 Project ID: 4053141.

different consumers possess different utilities towards return, but due to the limitation of resources available, it is more convenient to sell a uniform package which can cater for the needs of most people. Levy and Markowitz [22] showed that the optimal portfolio in utility maximization can be approximated by the mean-variance efficient frontier over ranges of commonly used utilities, return rates and volatilities. Hence, the mean-variance portfolio can basically entertain the almost optimal satisfaction of common consumers. Further studies support this approximation; for instances, see [20, 27, 29, 36].

Due to the nonlinear nature of the square function of the expectation of the terminal wealth involved in the variance, an immediate application of dynamic programming principle is not viable, which results that the analytic research in mean-variance portfolio optimization is used to mainly focused on single-period models at the first stage. The embedding technique developed by Li and Zhou [23] broke the ice by converting the mean-variance problems under both continuous time and multiperiod settings into the canonical linear-quadratic stochastic control problems. From that point on, more complicated mean-variance problems have also been investigated, in work such as [4, 6, 7, 24].

Variance is not the only risk measure commonly adopted in the portfolio selection problem. Jin et al. [14] consider a general convex risk function of the deviation of the terminal payoff from its own mean, by following the Lagrangian approach as proposed in [4], to characterize the optimal terminal payoff, and then they applied the Clark-Ocone formula to determine the optimal portfolio weights. Besides, they also studied the mean-downside-risk problem and established the non-existence of an optimal solution by showing that the optimal value function is unattainable by any admissible control. The downside-risk measure can remedy the common criticism on incurring penalty on the upside return which happens in the use of variance. Markowitz [28] also claims that "semivariance (an example of downside risk measure) seems more plausible than variance as a measure of risk since it is concerned only with adverse deviations". In contrast to continuous time models, Jin et al. [13] solved for the single-period mean-semivariance portfolio selection problem. After that, the study on the optimization problem subject to downside risk measure has been absent until the recent study by Cao et al. [5], in which they showed that mean-lower-partial-moments problem possesses a positive solution if we impose a uniform upper bound on the terminal payoff. For the relevant literature in connection with downside risk measure and semivariance, see also [12, 32, 33, 38]. Apart from using deviation risk measure, He et al. [11] studied continuous-time mean-risk portfolio choice problems with general risk measures including VaR, CVaR, and law-invariant coherent risk measures.

Turning back to reality, a number of financial crises have been observed frequently over recent decades, so tighter government regulations have been enforced in the financial market. On the other hand, the intensive competition in the market pushes any old-fashioned profitable strategies to the edge; all of these urge most companies to provide more tailor-made investment products in order to maintain their profit margins. A uniform package such as the mean-variance portfolio mentioned above can barely satisfy the demand of sophisticated investors nowadays, and a definitive answer to utility maximization with minimal risk is eagerly sought. Nevertheless, before our present work, the solution to this most relevant optimization problem has still been long absent in the literature.

In this article, we first provide a comprehensive study of utility-risk[1] portfolio selection problems: we suggest that the objective function of portfolio selection is not simply the expected value of a certain functional of the terminal payoff, but it also deals with the deviation risk caused by the underlying portfolio. Our proposed problem follows the recent trend of embedding various risk management criteria into the utility maximization framework. Such risk-monitoring mechanisms reduce the drawback caused by the ambitious investment strategy in pure utility maximization problems, which could lead to higher risk of potential pecuniary loss (see [42]). To name a few along this direction, Basak and Shapiro [2] first suggested implementing a Value-at-Risk (VaR) constraint into the portfolio optimization due to the prevailing regulation on VaR limitation. Some researches in [8, 21, 41] further turn the VaR limitation from a static constraint to a dynamic one in various utility-optimization problems. Besides, Zheng [42] studied the efficient frontier problem of both maximizing the expected utility of the terminal wealth and minimizing the conditional VaR of any potential loss. To the best of our knowledge, our present work is the first attempt to apply risk management to utility maximization subject to the deviation risk measure.

More precisely, we model the objective function as the difference of deviation risk (function of the deviation of the terminal payoff from its own mean) from the utility (concave increasing function of the terminal payoff) as in (2.2). We first follow the same idea as in [4] and [14] to convert our dynamic optimization problem into an equivalent static problem. By considering the first-order condition for the objective function, we can obtain a primitive static problem, called the *Nonlinear Moment Problem*, which characterizes the optimal terminal wealth with respect to the respective necessity and sufficiency results (Sections 3.1 and 4.1), which are fundamentally different, and not equivalent to each other. For necessity, the optimal terminal wealth satisfying two mild regularity conditions (Conditions 3.1 (i) and (ii)) solves for the Nonlinear Moment Problem; while for sufficiency, the solution of the Nonlinear Moment Problem that satisfies Condition 4.1 serves as the optimal terminal wealth. Note that this Nonlinear Moment Problem includes a variational inequality (3.1) with a set of constraints (3.2)-(3.4) involving the expectation of some nonlinear functions of the optimal terminal wealth and its own mean, or the "mean-field term" in the context of mean-field type control theory. The formulation of the Nonlinear Moment Problem is motivated by the mean-field approach developed in [3], in which the authors studied the classical mean-variance problem with the aid of a novel mean-field type HJB equation. Note that the same static problem may be obtained via the formal Lagrangian multiplier approach as in [4] and [14].

With the aid of the Nonlinear Moment Problem, our necessity conditions warrant an alternative deduction of the non-existence result of the mean-semivariance problem, first considered in [14]. On the other hand, for the application of the sufficiency conditions together with the Nonlinear Moment Problem, we replicate the explicit construction of the optimal solutions of various well-posed mean-risk problems in the existing literature. Furthermore, the novelty of our new approach allows us to establish new existence result for the optimal solutions for a variety of utility-risk problems, especially the utility-downside-risk (in Section 4.2) and the utility-strictly-convex-risk problems (in Section 4.3), in which the underlying utility satisfies the

[1] Jin et al. [14] call their problem formulation mean-risk problem, though they only consider deviation risk measure in their work, so to avoid ambiguity, we use a single word "risk" to stand for the deviation risk measure in the remaining of this paper.

common Inada Condition. To the best of our knowledge, these problems have not been considered so far before our work. Note that by the sufficiency result in Theorem 4.2, we can conclude that there exists an optimal solution for the utility-downside-risk problem including utility-semivariance problem, and this result is in contrast to [14], in which they find that the continuous-time mean-downside-risk problem possesses no optimal solution at all. As a consequence, the possibility of using semivariance as a natural risk measure in portfolio selection can now be legitimately implemented.

The determination of an optimal portfolio subject to the semivariance constraint plays a crucial role in the daily capital budgeting management. This semivariance risk measure is a reasonable one because it penalizes the downside loss deviation but not the upside profit. In the meantime, using the expected payoff or using a concave utility function of portfolio reward are the common means to measure the agent's satisfaction of the portfolio payoff. Hence, the portfolio management using either mean-semivariance or utility-semivariance criteria is crucial in both academia and industry. However, under the mean-semivariance setting, even though it has been shown that this problem has an optimal solution under the single-period setting, but no optimal solution exists in the continuous-time paradigm. In this article, we shall study the continuous-time utility-semivariance problem and provide a positive solution which settles the alternative standing problem in portfolio management. In addition to utility-semivariance setting, we actually consider the problem under a more general utility-deviation-risk framework. To tackle this general utility-risk problem, we convert it to the Nonlinear Moment Problem whose solution is directly related to the existence of optimal solution of the former. To the best of our knowledge, our newly proposed approach, via the Nonlinear Moment Problem, is crucial and it cannot be replaced by the standard approaches commonly encountered in the existing literature, such as (1) the direct approach (by constructing the optimal solution from an optimizing sequence) and (2) the indirect approach (by applying convex analysis through the conjugate functions); for details, we provide a comprehensive discussion on their infeasibility in Section A in Appendix. Alternatively, our problem can also be tackled using the common Lagrangian multiplier approach, however it will still eventually lead to exactly the same Nonlinear Moment Problem; the detail shall be demonstrated in Appendix A.3. Hence, the resolution of our Nonlinear Moment Problem is indispensable for establishing the existence of an optimal solution for our present utility-risk problem.

In this paper, we first introduce the problem formulation in Section 2 and convert our continuous-time utility-risk problem into an equivalent static formulation as stated in Theorem 2.5. In Section 3, we derive the necessary condition that the optimal terminal wealth satisfying two mild regularity conditions, Conditions 3.1 (i) and (ii), solves for the Nonlinear Moment Problem in Theorem 3.2. We then apply this necessity result to revisit the non-existence result for the mean-semivariance problem. In Section 4, we establish the verification theorem (Theorem 4.2), serving as the sufficient condition that the solution of the Nonlinear Moment Problem satisfying Condition 4.1 serves as the optimal terminal wealth. We then apply the sufficiency result to establish the existence of the corresponding optimal solutions for utility-downside-risk and utility-strictly-convex-risk problems in Sections 4.2 and 4.3 respectively; the technical proofs are deferred to the Appendix. Finally, we apply the Nonlinear Moment Problem to establish the sufficient condition for the existence of an optimal solution of mean-risk problem in Section 4.4. Such sufficient condition can be used to revisit the various well-known mean risk problems such as mean-weighted-power-risk (Example 4.18) (which includes mean-weighted-variance and mean-variance as special cases, also see Remark 4.19) and mean-exponential-risk problems (Example 4.20).

2. Problem Setting. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a fixed complete probability space, over which $W(t) = (W_1(t), \ldots, W_m(t))^t$ denotes $m-$dimensional standard Brownian motion; M^t denotes the transpose of a matrix M. We adopt the same market modelling setting as in Jin et al. [14]. Define $\mathcal{F}_t := \sigma(W(s) : s \le t)$. Suppose that the market has one riskless money account with price process $B(t)$ and m risky assets with the joint price process, $S(t) := (S_1(t), \ldots, S_m(t))^t$, such that the pair $(B(t), S(t))$ satisfies the following equations:

$$\begin{cases} dB(t) & = r(t)B(t)dt, \quad B(0) = b_0 > 0, \\ dS_k(t) & = \mu_k(t)S_k(t)\,dt + S_k(t)\sum_{j=1}^m \sigma_{kj}(t)dW_j(t), \quad S_k(0) = s_k > 0, k = 1, \ldots, m, \end{cases}$$

where $r(t)$ is the riskless interest-rate, $\mu_k(t)$ and $\sigma_k(t) := (\sigma_{k1}(t), \ldots, \sigma_{km}(t))$ are respectively the appreciation rate and volatility of the k-th risky asset, all assumed to be uniformly bounded. We also assume that the volatility matrix of assets $\sigma(t) := (\sigma_{kj}(t))_{m\times m}$ is uniformly elliptic, so that $\sigma(t)\sigma(t)^t \ge \delta I$ for some $\delta > 0$, so **the market is complete** and $(\sigma(t))^{-1}$ exists for all t.

Let $\pi(t) := (\pi_1(t), \ldots, \pi_m(t))^t$, where $\pi_k(t)$ is the money amount invested in the k-th risky asset of the portfolio at time t. The dynamics of controlled wealth process is:

$$dX^\pi(t) = (r(t)X^\pi(t) + \pi(t)^t\alpha(t))dt + \pi(t)^t\sigma(t)dW(t), \quad X^\pi(0) = x_0 > 0, \tag{2.1}$$

where $\alpha(t) := (\alpha_1(t), \ldots, \alpha_m(t))^t$ and $\alpha_k(t) := \mu_k(t) - r(t)$ for any $k \in \{1, \ldots, m\}$. The objective functional is:

$$J(\pi) := \mathbb{E}[U(X^\pi(T))] - \gamma\mathbb{E}[D\left(\mathbb{E}[X^\pi(T)] - X^\pi(T)\right)], \tag{2.2}$$

where the terminal time T is finite and $\gamma > 0$ denotes the risk-aversion coefficient. We denote by U a utility function such that $U : Dom(U) \to \mathbb{R}$ is strictly increasing, concave and continuously differentiable in the interior; here the domain of U, $\mathcal{D} := Dom(U)$, is a convex set in $\mathbb{R}$. Define the lower end point of the domain $\mathcal{D}$, $K := \inf(\mathcal{D}) \in [-\infty, \infty)$. For completeness, we extend the definition of U over $\mathbb{R}$ so that $U(x) = -\infty$ for $x \in \mathbb{R}/\mathcal{D}$ and $U'(K) := \lim_{x\downarrow K} U'(x)$. Here the function $D : \mathbb{R} \to \mathbb{R}_+$ stands for a risk function which measures the deviation of the random return from its own expectation. We assume that D is non-negative, convex and continuously differentiable.

For any given $p \ge 1$, denote $\mathcal{L}^p := \left\{Z|\, \|Z\|_p := \mathbb{E}[|Z|^p]^{\frac{1}{p}} < \infty\right\}$ and $\mathcal{L}^\infty := \{Z|\, \|Z\|_\infty := \sup_{\omega\in\Omega} |Z(\omega)| < \infty\}$. Define $\mathcal{H}^2$ to be the class of all $\mathcal{F}_t$-adapted processes π, equipped with a norm $\|\pi\|^2_{\mathcal{H}^2} := \mathbb{E}\left[\int_0^T \pi(t)^t\pi(t)dt\right] < \infty$.

Definition 2.1. *We define the class of all admissible controls $\pi \in \mathcal{A}$ as follows:*

$$\mathcal{A} := \left\{\pi \in \mathcal{H}^2 \middle| X^\pi(T) \in \mathcal{X}\right\},$$

where $\mathcal{X}$ is the class of all admissible terminal wealths, such that

$$\mathcal{X} := \{X \in \mathcal{L}^2 | X \in \mathcal{F}_T, X \in \mathcal{D} \ a.s., U(X) \in \mathcal{L}^1, D(\mathbb{E}[X] - X) \in \mathcal{L}^1\}.$$

Note that, for every admissible terminal wealth, both its expected utility and expected deviation risk are well-defined. It is clear that $\mathcal{X}$ is a convex subspace of $\mathcal{L}^2$[2]. For any admissible control π, we have $X^\pi \in \mathcal{H}^2$ and $X^\pi(t) \in \mathcal{L}^2$ for any $t \in [0,T]$ by Theorems 1.2 and 2.1 in [40].

Under the above settings, our utility risk problem can be stated as follows:

PROBLEM 2.2.

Maximize $J(\pi)$,

subject to $\pi \in \mathcal{A}$ *and* $(X^\pi(\cdot), \pi(\cdot))$ *satisfies (2.1) with initial wealth* x_0.

We define $\xi(t)$ as

$$\xi(t) := \exp\left(-\int_0^t \left(r(s)ds + \frac{1}{2}\alpha(s)^t \left(\sigma(s)\sigma(s)^t\right)^{-1} \alpha(s)ds + \alpha(s)^t \left(\sigma(s)^t\right)^{-1} dW(s)\right)\right).$$

By applying Itô's formula to $\xi(t)X^\pi(t)$, it is clear that $\xi(t)$ is the pricing kernel. Denote $\xi := \xi(T) \in \mathcal{L}^p$ for any $p \geq 1$. Hence, for a given initial condition $X^\pi(0) = x_0$, $\mathbb{E}[\xi X^\pi(T)] = x_0$ for any $\pi \in \mathcal{A}$. If $x_0 < \mathbb{E}[\xi] K$, $\mathcal{A}$ is empty[3]. If $x_0 = \mathbb{E}[\xi] K$, even when $\mathcal{A}$ is non-empty, all such $\pi \in \mathcal{A}$ will give the same terminal wealth, $X^\pi(T) = K$ a.s.[4], so no actual optimization is required, thus the corresponding problem becomes trivial. In the rest of this paper, based on this observation, we only consider our problem under this natural assumption:

ASSUMPTION 2.3. *The initial wealth* x_0*, the lower end point of* $\mathcal{D}$*,* $K \in [-\infty, \infty)$*, and pricing kernel* $\xi := \xi(T)$ *altogether satisfy:*

$$x_0 > \mathbb{E}[\xi] K.$$

Note that if we choose U to be linear and D to be quadratic, i.e. $U(x) = x$ and $D(x) = x^2$, then Problem 2.2 reduces to the classical mean-variance problem. If we only choose U to be linear, then Problem 2.2 reduces to the mean-risk problem as in [14]; in particular, if we alternatively choose $D(x) = ax_+ + bx_-$, then Problem 2.2 reduces to the mean-weighted-variance problem. If we just set D to be a convex function with $D(x) = 0$ for $x \leq 0$, Problem 2.2 is to maximize utility and minimize the downside risk of terminal wealth; its resolution will be established in Subsection 4.2.

Since our market is complete, all $\mathcal{L}^2$-integrable and $\mathcal{F}_T$-measurable terminal wealth can be attained by an admissible control, in the light of Martingale Representation Theorem. Our dynamic utility-risk optimization problem 2.2 can be converted into the following static optimization problem:

Define $\Psi : \mathcal{X} \to \mathbb{R}$ such that

$$\Psi(X) := \mathbb{E}[U(X)] - \gamma \mathbb{E}[D(\mathbb{E}[X] - X)]. \tag{2.3}$$

[2]Note that U is increasing and D is convex, and hence we have

$$|U(\theta x + (1-\theta)y)| \leq |U(x)| + |U(y)| \qquad \text{for all } x, y \in \mathcal{D} \text{ and } \theta \in [0,1];$$
$$0 \leq D(\theta x + (1-\theta)y) \leq \theta D(x) + (1-\theta)D(y) \text{ for all } x, y \in \mathbb{R} \text{ and } \theta \in [0,1].$$

From which, the claim follows.

[3]If $\mathcal{A}$ is non-empty, we have $x_0 = \mathbb{E}[\xi X^\pi(T)] \geq \mathbb{E}[\xi] K$ since $X^\pi(T) \in \mathcal{X}$ for any $\pi \in \mathcal{A}$, which implies $X^\pi(T) \geq K$ a.s.

[4]If there exists $\pi \in \mathcal{A}$ such that $\mathbb{P}[X^\pi(T) > K] > 0$, then $x_0 = \mathbb{E}[\xi X^\pi(T)] > \mathbb{E}[\xi] K$.

PROBLEM 2.4.

$$\begin{aligned}&\textit{Maximize } \Psi(X),\\&\textit{subject to } X\in\mathcal{X} \textit{ and } \mathbb{E}[\xi X]=x_0.\end{aligned}\tag{2.4}$$

Then, the optimal solution of Problem 2.4 is the optimal terminal wealth of Problem 2.2:

THEOREM 2.5 (Theorem 2.1 in [4] and Theorem 2.1 in [14]). *If $\pi(t)$ is optimal for Problem 2.2, then $X^\pi(T)$ is optimal for Problem 2.4. Conversely, if $X\in\mathcal{X}$ is optimal for Problem 2.4, there exists $\pi\in\mathcal{A}$ such that $X^\pi(T)=X$ and π is optimal for Problem 2.2.*

Note that the maximization in Problem 2.4 is confined to the set $\mathcal{X}$, so that the solution obtained in Problem 2.4 is an admissible terminal wealth in Problem 2.2. Our present paper aims to establish an admissible terminal wealth $X\in\mathcal{X}$ that maximizes $\Psi(X)$ under rather general scenarios, including those not yet covered in the existing literature.

3. Necessary Condition.

3.1. Maximum Principle. We first introduce the following two very mild technical conditions:

CONDITION 3.1.

- **(i)** *Both $U'(Z)\in\mathcal{L}^1$ and $D'\left(\mathbb{E}[Z]-Z\right)\in\mathcal{L}^1$.*
- **(ii)** *There exists $\delta>0$ such that $D\left(\mathbb{E}\left[Z\right]-Z-\delta\right)\in\mathcal{L}^1$ and $D\left(\mathbb{E}\left[Z\right]-Z+\delta\right)\in\mathcal{L}^1$.*

To show the necessity for optimality, we assume that the optimal solution of Problem 2.4, $\hat{X}\in\mathcal{X}$, satisfies Conditions 3.1 (i) and (ii). Now, it is necessary for $\hat{X}$ to solve for the following auxiliary static problem, we call it the *Nonlinear Moment Problem*:

THEOREM 3.2 (Nonlinear Moment Problem). *If $\hat{X}$ is the optimal solution of Problem 2.4 satisfying Conditions 3.1 (i) and (ii), then it is necessary that there exist constants $Y,M,R\in\mathbb{R}$ such that the quadruple $\left(\hat{X},Y,M,R\right)$ solves for the following variational inequality:*

$$\begin{cases}Y\xi=U'(\hat{X})-\gamma R+\gamma D'\left(M-\hat{X}\right), & \textit{a.s. on } \{\hat{X}>K\},\\ Y\xi\geq U'(\hat{X})-\gamma R+\gamma D'\left(M-\hat{X}\right), & \textit{a.s. on } \{\hat{X}=K\},\end{cases}\tag{3.1}$$

subject to the nonlinear moment constraints

$$\mathbb{E}[\xi\hat{X}]=x_0,\tag{3.2}$$

$$\mathbb{E}[\hat{X}]=M,\tag{3.3}$$

$$\mathbb{E}\left[D'\left(M-\hat{X}\right)\right]=R.\tag{3.4}$$

Note that if $\mathbb{P}[\hat{X}=K]=0$, the variational inequality (3.1) is reduced to an equality.

To prove the necessity, we first apply the first-order conditions as stated in Proposition 3.3. Next, we make use of Proposition 3.3 to give a preliminary result for characterizing the optimal solution of Problem 2.4, $\hat{X}$, in Lemma 3.5: if we can find a random variable, Z, as described in Lemma 3.5, then it is necessary that $\hat{X}$ has to satisfy the variational inequality (3.1). Finally, in Proposition 3.9, we construct such a Z.

3.1.1. Proof of Theorem 3.2. Let $\hat{X} \in \mathcal{X}$ be an optimal solution of Problem 2.4. We define $\Gamma : \mathcal{L}^2 \times \mathcal{D} \to \mathbb{R}$ by

$$\Gamma(X, x) := U'(x) - \gamma\mathbb{E}\left[D'\left(\mathbb{E}[X] - X\right)\right] + \gamma D'\left(\mathbb{E}[X] - x\right). \tag{3.5}$$

For simplicity of notation, in the rest of the paper, we shall denote the random variable $\Gamma\left(\hat{X}, \hat{X}\right)$ by $\hat{\Gamma}$.

Proposition 3.3. *If $\hat{X}$ is optimal for Problem 2.4 satisfying Condition 3.1 (i), then*

$$\mathbb{E}\left[\tilde{X}\hat{\Gamma}\right] \leq 0,$$

for all $\tilde{X} \in \Theta := \left\{Z \in \mathcal{L}^\infty \,|\, \mathbb{E}[Z\xi] = 0 \text{ and } \hat{X} + Z \in \mathcal{X}\right\}$.

For any $\tilde{X} \in \Theta$, by the convexity of $\mathcal{X}$, $\hat{X} + \theta\tilde{X} \in \mathcal{X}$ for all $0 < \theta < 1$. The directional derivative of $\Psi(X)$ is

$$\frac{d}{d\theta}\Psi(\hat{X} + \theta\tilde{X})\bigg|_{\theta=0} = \frac{d}{d\theta}\left(\mathbb{E}\left[U(\hat{X} + \theta\tilde{X})\right] - \gamma\mathbb{E}\left[D\left(\mathbb{E}[\hat{X} + \theta\tilde{X}] - (\hat{X} + \theta\tilde{X})\right)\right]\right)\bigg|_{\theta=0}.$$

Before we proceed on the proof of Proposition 3.3, we first justify the interchange of the order of differentiation and taking expectation of the above expression. To this end, we need the following lemma (whose proof is postponed to Appendix B):

Lemma 3.4. *Given two random variables $\hat{X} \in \mathcal{X}$ and $\tilde{X} \in \mathcal{L}^2$ such that $\hat{X} + \tilde{X} \in \mathcal{X}$,*

$$\begin{aligned}
&\lim_{\theta \downarrow 0} \mathbb{E}\left[\frac{U(\hat{X} + \theta\tilde{X}) - \gamma D\left(\mathbb{E}\left[\hat{X} + \theta\tilde{X}\right] - (\hat{X} + \theta\tilde{X})\right) - \left(U(\hat{X}) - \gamma D\left(\mathbb{E}\left[\hat{X}\right] - \hat{X}\right)\right)}{\theta}\right] \\
&= \mathbb{E}\left[\lim_{\theta \downarrow 0} \frac{U(\hat{X} + \theta\tilde{X}) - \gamma D\left(\mathbb{E}\left[\hat{X} + \theta\tilde{X}\right] - (\hat{X} + \theta\tilde{X})\right) - \left(U(\hat{X}) - \gamma D\left(\mathbb{E}\left[\hat{X}\right] - \hat{X}\right)\right)}{\theta}\right].
\end{aligned}$$

Proof of Proposition 3.3. By Lemma 3.4, the chain rule and Condition 3.1 (i), we have

$$\begin{aligned}
\frac{d}{d\theta}\Psi(\hat{X} + \theta\tilde{X})\bigg|_{\theta=0} &= \mathbb{E}\left[U'(\hat{X})\tilde{X}\right] - \gamma\mathbb{E}\left[D'\left(\mathbb{E}[\hat{X}] - \hat{X}\right)\left(\mathbb{E}[\tilde{X}] - \tilde{X}\right)\right] \\
&= \mathbb{E}\left[\tilde{X}\hat{\Gamma}\right]. \qquad (3.6)
\end{aligned}$$

Our claim follows by the first-order necessary condition for optimality. ☐

To characterize the optimal solution $\hat{X}$, we first have the following lemma:

Lemma 3.5. *Given that $\hat{X}$ is optimal for Problem 2.4 satisfying Condition 3.1 (i), if there exists a random variable, $Z \in [0, 1]$, such that the following three items hold:*

$$\begin{cases} Z > 0 & a.s. \ on \left\{\hat{X} > K\right\}, \\ Z = 0 & a.s. \ on \left\{\hat{X} = K\right\}, \end{cases} \tag{3.7}$$

$$Z\xi\hat{\Gamma} \in \mathcal{L}^1, \quad Z\left(\hat{\Gamma} - Y\xi\right) \in \mathcal{L}^\infty, \text{ where } Y := \frac{\mathbb{E}\left[Z\xi\hat{\Gamma}\right]}{\mathbb{E}[Z\xi^2]}, \text{ and} \tag{3.8}$$

$$\hat{X} + Z\left(\hat{\Gamma} - Y\xi\right) \in \mathcal{X}, \tag{3.9}$$

then it is necessary that $\hat{\Gamma}$ defined in (3.5) satisfies the following algebraic structure:

$$\begin{cases} \hat{\Gamma} = Y\xi & \text{a.s. on} \left\{\hat{X} > K\right\}, \\ \hat{\Gamma} \le Y\xi & \text{a.s. on} \left\{\hat{X} = K\right\}. \end{cases} \tag{3.10}$$

Proof. We split our proof into two parts: (i) $\hat{\Gamma} = Y\xi$ a.s. on $\left\{\hat{X} > K\right\}$, and (ii) $\hat{\Gamma} \le Y\xi$ a.s. on $\left\{\hat{X} = K\right\}$.

(i) Take

$$\tilde{X} = Z\left(\hat{\Gamma} - Y\xi\right), \tag{3.11}$$

(3.8) and (3.9) warrants that $\tilde{X} \in \mathcal{L}^\infty$ with $\hat{X} + \tilde{X} \in \mathcal{X}$ and

$$\mathbb{E}[\tilde{X}\xi] = \mathbb{E}\left[Z\xi\hat{\Gamma}\right] - Y\mathbb{E}[Z\xi^2] = 0.$$

By Proposition 3.3 and the fact that $Y = \mathbb{E}\left[Z\xi\hat{\Gamma}\right]/\mathbb{E}[Z\xi^2]$,

$$0 \ge \mathbb{E}\left[\tilde{X}\hat{\Gamma}\right] = \mathbb{E}\left[Z\left(\hat{\Gamma} - Y\xi\right)\hat{\Gamma}\right] - Y\mathbb{E}\left[Z\xi\hat{\Gamma}\right] + Y^2\left[Z\xi^2\right] = \mathbb{E}\left[Z\left(\hat{\Gamma} - Y\xi\right)^2\right].$$

(3.7) ensures that $Z\left(\hat{\Gamma} - Y\xi\right)^2 \ge 0$, and therefore $\mathbb{E}\left[Z\left(\hat{\Gamma} - Y\xi\right)^2\right] = 0$, which implies that $Z\left(\hat{\Gamma} - Y\xi\right)^2 = 0$ a.s. By (3.7), on $\left\{\hat{X} > K\right\}$, $Z > 0$, hence $\hat{\Gamma} = Y\xi$ a.s. on $\left\{\hat{X} > K\right\}$.

(ii) Assume the contrary that $\mathbb{P}\left[\mathbb{I}\left\{\hat{X} = K\right\}\left(\hat{\Gamma} - Y\xi\right) > 0\right] > 0$. Consider

$$\tilde{X} = k\mathbb{I}\left\{\hat{X} = K\right\} - \frac{\min\{\hat{X} - K, 1\}}{2}$$

where $k := \mathbb{E}\left[\left(\min\{\hat{X} - K, 1\}\right)\xi\right] \Big/ \left(2\mathbb{E}\left[\mathbb{I}\left\{\hat{X} = K\right\}\xi\right]\right) > 0$ in light of the required feasibility of $\mathcal{X}$ and our interest being only on non-trivial setting. We have

$$\hat{X} + \tilde{X} = \begin{cases} \max\left\{\frac{\hat{X}+K}{2}, \hat{X} - \frac{1}{2}\right\}, & \text{if } \hat{X} > K, \\ K + k, & \text{if } \hat{X} = K. \end{cases}$$

Obviously, $\tilde{X} \in \mathcal{L}^\infty$, $K \le \hat{X} + \tilde{X} \in \mathcal{D}$ a.s. and $\mathbb{E}\left[\tilde{X}\xi\right] = 0$.

Since U is monotonic,

$$|U(\hat{X}+\tilde{X})| \le |U(K+k)| + \left| U\left(\max\left\{\frac{\hat{X}+K}{2}, \hat{X}-\frac{1}{2}\right\}\right)\right| \le |U(K+k)|+|U(K)|+|U(\hat{X})|.$$

As $\mathbb{P}[\hat{X}=K]>0$ and $\mathbb{P}[\hat{X}=K]\cdot|U(K)| \le \mathbb{E}[|U(\hat{X})|]$ is finite, they prevent $U(K)$ from taking $-\infty$. Clearly, $K+k \in \mathcal{D}$, and so $U(K+k)$ is finite. Since $\hat{X} \in \mathcal{X}$, we also have $U(\hat{X}) \in \mathcal{L}^1$. These three claims altogether imply that $U(\hat{X}+\tilde{X}) \in \mathcal{L}^1$.
Note that

$$\begin{aligned}
\mathbb{E}\left[\hat{X}+\tilde{X}\right] &= \mathbb{E}\left[(K+k)\,\mathbb{I}[\hat{X}=K]\right] + \mathbb{E}\left[\max\left\{\frac{\hat{X}+K}{2}, \hat{X}-\frac{1}{2}\right\}\left(1-\mathbb{I}[\hat{X}=K]\right)\right] \\
&= k\mathbb{P}[\hat{X}=K] + \mathbb{E}\left[\max\left\{\frac{\hat{X}+K}{2}, \hat{X}-\frac{1}{2}\right\}\right],
\end{aligned}$$

we then establish the upper bound of $D\left(\mathbb{E}\left[\hat{X}+\tilde{X}\right]-\left(\hat{X}+\tilde{X}\right)\right)$ into three different cases: (1) $\hat{X}=K$, (2) $K<\hat{X}<K+1$, and (3) $\hat{X}\ge K+1$.

(1) $D\left(\mathbb{E}\left[\hat{X}+\tilde{X}\right]-\left(\hat{X}+\tilde{X}\right)\right) = D\left(\mathbb{E}\left[\hat{X}+\tilde{X}\right]-K-k\right)$ is a finite constant.

(2) Note that

$$\begin{aligned}
\mathbb{E}\left[\frac{\hat{X}+K}{2}\right]-\frac{\hat{X}+K}{2} &\le \mathbb{E}\left[\max\left\{\frac{\hat{X}+K}{2}, \hat{X}-\frac{1}{2}\right\}\right]-\frac{\hat{X}+K}{2} \\
&\le \mathbb{E}\left[\max\left\{\frac{\hat{X}+K}{2}, \hat{X}-\frac{1}{2}\right\}\right]-K.
\end{aligned}$$

By convexity of D, we have

$$\begin{aligned}
& D\left(\mathbb{E}\left[\hat{X}+\tilde{X}\right]-\left(\hat{X}+\tilde{X}\right)\right) = D\left(k\mathbb{P}\left[\hat{X}=K\right]+\mathbb{E}\left[\max\left\{\frac{\hat{X}+K}{2}, \hat{X}-\frac{1}{2}\right\}\right]-\frac{\hat{X}+K}{2}\right) \\
&\le D\left(k\mathbb{P}\left[\hat{X}=K\right]+\mathbb{E}\left[\max\left\{\frac{\hat{X}+K}{2}, \hat{X}-\frac{1}{2}\right\}\right]-K\right) + D\left(k\mathbb{P}\left[\hat{X}=K\right]+\mathbb{E}\left[\frac{\hat{X}+K}{2}\right]-\frac{\hat{X}+K}{2}\right) \\
&\le C + D\left(k\mathbb{P}\left[\hat{X}=K\right]+\frac{\mathbb{E}\left[\hat{X}\right]-\hat{X}}{2}\right) \le C + D\left(k\mathbb{P}\left[\hat{X}=K\right]+\frac{\mathbb{E}\left[\hat{X}\right]-K}{2}\right) + D\left(\frac{\mathbb{E}\left[\hat{X}\right]-\hat{X}}{2}\right) \\
&\le C + D(0) + D\left(\mathbb{E}\left[\hat{X}\right]-\hat{X}\right), \text{ where } C \text{ is a constant.}
\end{aligned}$$

(3) By the same arguments as that for case (ii), we have

$$\begin{aligned}
D\left(\mathbb{E}\left[\hat{X}+\tilde{X}\right]-\left(\hat{X}+\tilde{X}\right)\right) &\le D\left(k\mathbb{P}\left[\hat{X}=K\right]+\mathbb{E}\left[\max\left\{\frac{\hat{X}+K}{2}, \hat{X}-\frac{1}{2}\right\}\right]-K\right) \\
&\quad + D\left(k\mathbb{P}\left[\hat{X}=K\right]+\mathbb{E}\left[\hat{X}\right]-K\right) + D\left(\mathbb{E}\left[\hat{X}\right]-\hat{X}\right).
\end{aligned}$$

Since $\hat{X} \in \mathcal{X}$, we have $D(\mathbb{E}[\hat{X}] - \hat{X}) \in \mathcal{L}^1$, so $D(\mathbb{E}[\hat{X} + \tilde{X}] - (\hat{X} + \tilde{X})) \in \mathcal{L}^1$. Finally,

$$
\begin{aligned}
\mathbb{E}\left[\hat{\Gamma}\tilde{X}\right] &= k\mathbb{E}\left[\mathbb{I}\left\{\hat{X} = K\right\}\hat{\Gamma}\right] - \frac{\mathbb{E}\left[\left(\min\left\{\hat{X} - K, 1\right\}\right)\hat{\Gamma}\right]}{2} \\
&= k\mathbb{E}\left[\mathbb{I}\left\{\hat{X} = K\right\}\hat{\Gamma}\right] - \frac{\mathbb{E}\left[\left(\min\left\{\hat{X} - K, 1\right\}\right)Y\xi\right]}{2} \\
&= k\left(\mathbb{E}\left[\mathbb{I}\left\{\hat{X} = K\right\}\hat{\Gamma}\right] - \mathbb{E}\left[\mathbb{I}\left\{\hat{X} = K\right\}Y\xi\right]\right) > 0,
\end{aligned}
\tag{3.12}
$$

where the second equality follows because we have shown that $\hat{\Gamma} = Y\xi$ when $\hat{X} > K$ and the third equality follows because $k = \mathbb{E}\left[\left(\min\left\{\hat{X} - K, 1\right\}\right)\xi\right] \Big/ \left(2\mathbb{E}\left[\mathbb{I}\left\{\hat{X} = K\right\}\xi\right]\right)$. (3.12) violates Proposition 3.3, this implies that $\mathbb{P}\left[\mathbb{I}\left\{\hat{X} = K\right\}\left(\hat{\Gamma} - Y\xi\right) > 0\right] > 0$ leads to a contradiction. We have

$$
\mathbb{P}\left[\mathbb{I}\left\{\hat{X} = K\right\}\left(\hat{\Gamma} - Y\xi\right) \leq 0\right] = 1.
$$

Therefore, the complete characterization as specified in (3.10) now follows. ☐

The overall necessity claim will be accomplished if the explicit construction of Z as described in the hypothesis in Lemma 3.5 can be obtained. Even the nature of such Z appears to be complicated and uncommon in the literature, we shall devote the remaining part of this subsection to the establishment of its existence.

In order to satisfy (3.9), $\tilde{X}$ expressed in terms of Z as in (3.11) needs to be bounded so that the deviation of $U(\hat{X} + \tilde{X})$ from $U(\hat{X})$ is less than some constant, say 1 for simplicity, almost surely. To warrant this, we need the following lemma:

Lemma 3.6. *There exists $\delta^U : int(\mathcal{D}) \to (0, 1]$ such that for any $x_0 \in int(\mathcal{D})$,*

$$
|U(x) - U(x_0)| \leq 1, \forall x \in \mathcal{D} \text{ such that } |x - x_0| < \delta^U(x_0).
$$

Proof. See Appendix. ☐

We shall make use of δ^U defined in Lemma 3.6 to construct Z so that $U(\hat{X} + \tilde{X}) \in \mathcal{L}^1$, where $\tilde{X}$ in terms of Z is given in (3.11). Beforehand, for any $y \in (0, \infty)$, define a random variable $Z_y \in [0, 1]$ by:

$$
Z_y := \begin{cases} 0, & \text{if } \hat{X} = K, \\ 1, & \text{if } \hat{X} > K \text{ and } \hat{\Gamma} = y\xi, \\ \min\left\{\frac{\min\left\{\delta^U(\hat{X}), \frac{1}{2}(\hat{X} - K), \frac{\delta}{2}\right\}}{|\hat{\Gamma} - y\xi|}, 1\right\}, & \text{otherwise.} \end{cases}
\tag{3.13}
$$

First, we show that (3.8) is satisfied for any $y \in (0, \infty)$:

Lemma 3.7. *For any $y \in (0, \infty)$, we have $Z_y\xi\hat{\Gamma} \in \mathcal{L}^1$ and $Z_y\left(\hat{\Gamma} - y\xi\right) \in \mathcal{L}^\infty$.*

Proof. By the definition of Z_y in (3.13), $\left|Z_y\left(\hat{\Gamma} - y\xi\right)\right| \leq \delta$, so $Z_y\left(\hat{\Gamma} - y\xi\right) \in \mathcal{L}^\infty$. Since

$$
\left|Z_y\xi\hat{\Gamma}\right| \leq \left|Z_y\xi\left(\hat{\Gamma} - y\xi\right)\right| + \left|y\xi^2 Z_y\right| \leq \delta\xi + y\xi^2,
$$

we have $Z_y\xi\hat{\Gamma} \in \mathcal{L}^1$. ☐

Next, we want to ensure one can find a y so that $y = \frac{\mathbb{E}\left[\xi\hat{\Gamma}Z_y\right]}{\mathbb{E}[\xi^2 Z_y]}$ is satisfied.

LEMMA 3.8. *Define $f : (0, \infty) \to \mathbb{R}$ by*

$$f(y) := y\mathbb{E}\left[\xi^2 Z_y\right] - \mathbb{E}\left[\xi\hat{\Gamma}Z_y\right].$$

There is a root $y^* \in (0, \infty)$ such that $f(y^*) = 0$.

Proof. See Appendix. ∎

PROPOSITION 3.9. *Suppose the optimal solution of Problem 2.4, $\hat{X}$, satisfies Conditions 3.1 (i) and (ii). There exists a random variable $Z \in [0, 1]$ satisfying (3.7)-(3.9) in Lemma 3.5.*

Proof. We shall verify that $Z := Z_{y^*}$ with Z_y as defined in (3.13) and y^* obtained in Lemma 3.8 satisfies (3.7)-(3.9). Note that δ^U and δ^D in Lemma 3.6 only take positive values no matter what the corresponding arguments are; in particular, according to (3.13), when $\hat{X} > K$, $Z > 0$. Therefore, Z satisfies (3.7).

Note that $y^* = \mathbb{E}\left[Z\xi\hat{\Gamma}\right]/\mathbb{E}[Z\xi^2]$ by Lemma 3.8. By Lemma 3.7, we have $Z\xi\hat{\Gamma} \in \mathcal{L}^1$ and $Z\left(\hat{\Gamma} - y^*\xi\right) \in \mathcal{L}^\infty$, thus (3.8) is satisfied.

By a simple calculation under the third case in (3.13), $\left|Z\left(\hat{\Gamma} - y^*\xi\right)\right| \leq \frac{1}{2}\left(\hat{X} - K\right)$, thus we have

$$\hat{X} + Z\left(\hat{\Gamma} - \frac{\mathbb{E}\left[Z\xi\hat{\Gamma}\right]}{\mathbb{E}[Z\xi^2]}\xi\right) \geq \hat{X} - \frac{1}{2}\left(\hat{X} - K\right) = \frac{1}{2}\left(\hat{X} + K\right) \in \mathcal{D} \text{ a.s.}$$

Since $\left|Z\left(\hat{\Gamma} - y^*\xi\right)\right| \leq \delta^U\left(\hat{X}\right)$, by a direct application of Lemma 3.6 (a), we have $\left|U\left(\hat{X} + Z\left(\hat{\Gamma} - y^*\xi\right)\right) - U(\hat{X})\right| \leq 1$, and thus

$$\left|U\left(\hat{X} + Z\left(\hat{\Gamma} - \frac{\mathbb{E}\left[Z\xi\hat{\Gamma}\right]}{\mathbb{E}[Z\xi^2]}\xi\right)\right)\right| \leq \left|U(\hat{X})\right| + 1$$

. Hence, $U\left(\hat{X} + Z\left(\hat{\Gamma} - \frac{\mathbb{E}\left[Z\xi\hat{\Gamma}\right]}{\mathbb{E}[Z\xi^2]}\xi\right)\right) \in \mathcal{L}^1$.

Similarly, since we also have $\left|\mathbb{E}\left[Z\left(\hat{\Gamma} - y^*\xi\right)\right] - Z\left(\hat{\Gamma} - y^*\xi\right)\right| \leq \delta$, we have

$$\begin{aligned}
&D\left(\mathbb{E}\left[\hat{X} + Z\left(\hat{\Gamma} - y^*\xi\right)\right] - \left(\hat{X} + Z\left(\hat{\Gamma} - y^*\xi\right)\right)\right) \\
&= D\left(\mathbb{E}\left[\hat{X}\right] - \hat{X} - Z\left(\hat{\Gamma} - y^*\xi\right) + \mathbb{E}\left[Z\left(\hat{\Gamma} - y^*\xi\right)\right]\right) \\
&\leq D\left(\mathbb{E}\left[\hat{X}\right] - \hat{X} - \delta\right) + D\left(\mathbb{E}\left[\hat{X}\right] - \hat{X} + \delta\right).
\end{aligned}$$

Hence, by Condition 3.1 (ii), $D\left(\mathbb{E}\left[\hat{X} + Z\left(\hat{\Gamma} - y^*\xi\right)\right] - \left(\hat{X} + Z\left(\hat{\Gamma} - y^*\xi\right)\right)\right) \in \mathcal{L}^1$.

We can now conclude $\hat{X} + Z\left(\hat{\Gamma} - y^*\xi\right)$ satisfies all the admissibility conditions of $\mathcal{X}$, and hence Z satisfies (3.9). ∎

In summary, by Proposition 3.9, we have Z_{y^*} satisfying (3.7)-(3.9) in Lemma 3.5. By Lemma 3.5, it is necessary that $\hat{\Gamma}$ in terms of $\hat{X}$ as in (3.5) satisfies the following algebraic structure:

$$\begin{cases} \hat{\Gamma} = Y\xi & \text{a.s. on } \left\{\hat{X} > K\right\}, \\ \hat{\Gamma} \leq Y\xi & \text{a.s. on } \left\{\hat{X} = K\right\}, \end{cases}$$

where $Y = \mathbb{E}\left[Z\xi\hat{\Gamma}\right]/\mathbb{E}\left[Z\xi^2\right]$ and Z is obtained in Proposition 3.9. Now, by setting $M := \mathbb{E}[\hat{X}]$ and $R := \mathbb{E}[D'(M - \hat{X})]$ together with the constraint $\mathbb{E}[\hat{X}\xi] = x_0$ given in Problem 2.4, the claim described in Theorem 3.2 follows.

Remark 3.10 (Comments on the Proof of Theorem 3.2).

We sincerely thank the anonymous referee for his/her suggestion on an inspiring alternative proof of Theorem 3.2 in a special case when the domain of the utility is the whole real line ($\mathcal{D} = \mathbb{R}$). We now streamline his/her arguments as follows:

To avoid technical difficulties, all random variables are supposed to be integrable. Together with $\mathcal{D} = \mathbb{R}$, $\mathcal{X}$ becomes a collection of all real-valued random variables. By Proposition 3.3, for an arbitrary $\tilde{X} \in \Theta := \{Z | \mathbb{E}[Z\xi] = 0\}$, which is the collection of all plausible perturbation of $\hat{X}$, we have $\mathbb{E}\left[\tilde{X}\hat{\Gamma}\right] \leq 0$ and $\mathbb{E}\left[-\tilde{X}\hat{\Gamma}\right] \leq 0$. Thus, we conclude that, for any $\tilde{X} \in \Theta$,

$$\mathbb{E}\left[\tilde{X}\hat{\Gamma}\right] = 0 \tag{3.14}$$

and therefore, $\hat{\Gamma}$ is orthogonal to Θ.

Define $y := \frac{\mathbb{E}[\hat{\Gamma}\xi]}{\mathbb{E}[\xi^2]} \in \mathbb{R}$, we have

$$\mathbb{E}[(\hat{\Gamma} - y\xi)\xi] = \mathbb{E}[\hat{\Gamma}\xi] - y\mathbb{E}[\xi^2] = 0, \tag{3.15}$$

thus $\hat{\Gamma} - y\xi \in \Theta$. By (3.14), we have

$$\mathbb{E}[(\hat{\Gamma} - y\xi)\hat{\Gamma}] = 0. \tag{3.16}$$

Further applying (3.15) and (3.16), we have

$$\mathbb{E}[(\hat{\Gamma} - y\xi)^2] = \mathbb{E}[(\hat{\Gamma} - y\xi)\hat{\Gamma}] - y\mathbb{E}[(\hat{\Gamma} - y\xi)\xi] = 0,$$

which concludes that $\hat{\Gamma} = y\xi$ a.s., and the necessity result in Theorem 3.2 follows.

More rigorously, we now revert to discuss on the integrability conditions which we assumed. With such consideration, $\mathcal{X}$ and Θ are confined as the following:

$$\begin{aligned} &\mathcal{X} := \{X \in \mathcal{L}^2 | X \in \mathcal{F}_T, X \in \mathcal{D} \text{ a.s.}, U(X) \in \mathcal{L}^1, D(\mathbb{E}[X] - X) \in \mathcal{L}^1\}; \\ &\Theta := \{Z \in \mathcal{L}^\infty | \mathbb{E}[Z\xi] = 0, \hat{X} + Z \in \mathcal{X}\}. \end{aligned}$$

To conclude that $\hat{\Gamma}$ is orthogonal to Θ, we need $-\tilde{X} \in \Theta$ for any $\tilde{X} \in \Theta$ but it is not apparent because $U(\hat{X} - \tilde{X})$ may not be integrable. The integrability of $U(\hat{X} - \tilde{X})$ can be warranted if $\lim_{x\to-\infty} \frac{U(x)}{x} < \infty$. On the other hand, to have $\hat{\Gamma} - y\xi \in \Theta$, we need both (3.15) and $\hat{\Gamma} - y\xi \in \mathcal{L}^\infty$, but the validity of the latter is usually not immediate either.

Furthermore, the domain of the utility may not be necessarily the whole real line in general. For instance, the domains of power utility and logarithm utility, which are commonly considered in literature, are usually only the positive half real line. Also, *the existence results of optimal solution of utility-downside-risk and utility-strictly-convex-risk problems, which will be considered in Sections 4.2 and 4.3 respectively, require the assumption that the utility function satisfies the Inada conditions, under which the domain of the utility is certainly not the whole real line. In these cases,* $\hat{X} - \tilde{X} \in \mathcal{D}$ *a.s. may not hold for any* $\tilde{X} \in \Theta$. *Thus, the inequality* $\mathbb{E}\left[-\tilde{X}\hat{\Gamma}\right] \leq 0$ *may also not hold for all* $\tilde{X} \in \Theta$, *and hence we cannot have* $\mathbb{E}\left[\tilde{X}\hat{\Gamma}\right] = 0$ *for any* $\tilde{X} \in \Theta$.

3.2. Application to the Mean-Semivariance Problem. In this subsection, we take $U(x) = x$, $D(x) = \frac{1}{2}x_+^2$. Then $D'(x) = x_+$. We revisit the non-existence result first obtained in [14] via our Theorem 3.2.

THEOREM 3.11. *There is no optimal solution for the continuous-time mean-semivariance problem.*

Proof. Assume the contrary, that there exists an admissible optimal control $\hat{\pi}$; then its corresponding optimal terminal wealth $\hat{X} \in \mathcal{L}^2$ solves Problem 2.4 by Theorem 2.5. Since D and D' are bounded by quadratic and linear functions respectively, it is clear that $\hat{X}$ satisfies Conditions 3.1 (i) and (ii). Hence, by Theorem 3.2, it is necessary that there exist constants $Y, M, R \in \mathbb{R}$ such that the quadruple $(\hat{X}, Y, M, R)$ solves for the following Nonlinear Moment Problem:

$$Y\xi + \gamma R - 1 = \gamma\left(M - \hat{X}\right)_+ \text{ a.s.,} \tag{3.17}$$

subject to the constraints: $\mathbb{E}\left[\xi\hat{X}\right] = x_0$, $\mathbb{E}\left[\hat{X}\right] = M$ and $\mathbb{E}\left[\left(M - \hat{X}\right)_+\right] = R$.

Firstly, by taking expectation on the both sides of (3.17), we immediately have $Y = 1/\mathbb{E}[\xi] > 0$. If $\gamma R - 1 \geq 0$, then by (3.17), $\gamma\left(M - \hat{X}\right)_+ > 0$ a.s., and hence $\mathbb{E}[\hat{X}] < M$ which is in conflict with the constraint $\mathbb{E}[\hat{X}] = M$. If $\gamma R - 1 < 0$, there exists some $\xi_0 > 0$ such that $\gamma R - 1 + Y\xi < 0$ for all $\xi \in (0, \xi_0)$, which contradicts the positivity of the right hand side in (3.17). Thus, the nonlinear moment problem has no solution. We conclude that mean-semivariance problem does not admit an optimal solution. ∎

REMARK 3.12. *The mean-semivariance problem has been investigated in [14]. The authors considered the semivariance minimization problem with a fixed mean, and showed that this problem does not have an optimal solution except for the trivial case in which the mean is equal to the terminal wealth, which is the initial wealth accumulated at riskless interest rate. The nonexistence was proven in their work by showing that the optimal value function is non-attainable. The constrained optimization problem in [14] and in Problem 2.2 are equivalent for suitable values of mean and risk aversion parameter. The trivial riskless solution becomes optimal in Problem 2.2 only when* $\gamma = \infty$. *For* $\gamma < \infty$, *the riskless strategy is dominated by another strategy attaining* $\frac{x_0}{\mathbb{E}[\xi]} + \theta\left(1 - \frac{\mathbb{E}[\xi]}{\mathbb{E}[\xi^2]}\xi\right)$ *as the corresponding terminal wealth for sufficiently small values of* θ[5].

4. Sufficient Condition.

[5] Because the mean of $\frac{x_0}{\mathbb{E}[\xi]} + \theta\left(1 - \frac{\mathbb{E}[\xi]}{\mathbb{E}[\xi^2]}\xi\right)$ is greater than $\frac{x_0}{\mathbb{E}[\xi]}$ in the order of $O(\theta)$ while the semivariance of $\frac{x_0}{\mathbb{E}[\xi]} + \theta\left(1 - \frac{\mathbb{E}[\xi]}{\mathbb{E}[\xi^2]}\xi\right)$ is of the order $O(\theta^2)$, therefore $\frac{x_0}{\mathbb{E}[\xi]} + \theta\left(1 - \frac{\mathbb{E}[\xi]}{\mathbb{E}[\xi^2]}\xi\right)$ has a greater objective value for sufficiently small θ.

4.1. Verification Theorem. We first introduce the following technical condition:

CONDITION 4.1. *Both* $U'(Z) \in \mathcal{L}^2$ *and* $D'(\mathbb{E}[Z]-Z) \in \mathcal{L}^2$.

In this subsection, we aim to show that the solution of the Nonlinear Moment Problem satisfying Condition 4.1 is optimal terminal wealth of Problem 2.2. There is a fundamental difference between the necessary condition in Theorem 3.2 and the sufficient condition in the next theorem. Conditions 3.1 (i) and (ii) are needed for the optimal terminal wealth satisfying the Nonlinear Moment Problem in the necessity result, while Condition 4.1 is required for the sufficiency.

THEOREM 4.2. *Suppose that there exists* $\hat{X} \in \mathcal{X}$ *satisfying Condition 4.1 and there exist constants* $Y, M, R \in \mathbb{R}$ *so that the quadruple* $\left(\hat{X}, Y, M, R\right)$ *solves for the Nonlinear Moment Problem (3.1)-(3.4). Then,* $\hat{X}$ *is optimal for Problem 2.4, and it is also the optimal terminal wealth of Problem 2.2.*

REMARK 4.3. *Theorem 4.2 boils the optimal control problem 2.2 down to a static problem. Suppose that there exists an implicit function* $I(m,y) \in \mathbb{R}$ *satisfying:*

$$U'(I(m,y)) + \gamma D'(m - I(m,y)) = y, \quad \text{for any } (m,y). \tag{4.1}$$

Then the Nonlinear Moment Problem (3.1)-(3.4) will be solved by $(\max\{I(M, \gamma R + Y\xi), K\}, Y, M, R)$, *where the constants* Y, M *and* R *satisfy the following system of nonlinear equations:*

$$\mathbb{E}[\xi \max\{I(M, \gamma R + Y\xi), K\}] = x_0, \tag{4.2}$$

$$\mathbb{E}[\max\{I(M, \gamma R + Y\xi), K\}] = M, \tag{4.3}$$

$$\mathbb{E}\left[D'\left(M - \max\{I(M, \gamma R + Y\xi), K\}\right)\right] = R. \tag{4.4}$$

After we verify that $\max\{I(M, \gamma R + Y\xi), K\}$ *belongs to* $\mathcal{X}$ *and also satisfies Condition 4.1,* $\max\{I(M, \gamma R + Y\xi), K\}$ *is the optimal solution for Problem 2.4.*

Proof of Theorem 4.2.

Let $(\hat{X}, Y, M, R)$ be the solution of Nonlinear Moment Problem (3.1)-(3.4) and $\tilde{X} \in \mathcal{L}^2$ be an arbitrary random variable such that $\hat{X} + \tilde{X}$ is admissible for Problem 2.4, i.e. $\hat{X} + \tilde{X} \in \mathcal{X}$ and $\mathbb{E}\left[\xi\left(\hat{X} + \tilde{X}\right)\right] = x_0$. By (3.2), we have $\mathbb{E}\left[\xi\tilde{X}\right] = 0$. By Lemma 3.4, the chain rule and and under our hypothesis that $\hat{X}$ satisfies (3.1) and Condition 4.1, we have

$$\begin{aligned}
\left.\frac{d}{d\theta}\Psi(\hat{X} + \theta\tilde{X})\right|_{\theta=0} &= \mathbb{E}\left[U'(\hat{X})\tilde{X}\right] - \gamma\mathbb{E}\left[D'\left(\mathbb{E}[\hat{X}] - \hat{X}\right)\left(\mathbb{E}[\tilde{X}] - \tilde{X}\right)\right] \\
&= \mathbb{E}\left[\tilde{X}\left(U'(\hat{X}) - \gamma\mathbb{E}\left[D'\left(\mathbb{E}[\hat{X}] - \hat{X}\right)\right] + \gamma D'\left(\mathbb{E}[\hat{X}] - \hat{X}\right)\right)\right] \\
&\leq \mathbb{E}\left[\tilde{X}(Y\xi)\right] = 0.
\end{aligned} \tag{4.5}$$

The last inequality follows from the fact that $\hat{X}$ satisfies (3.1), and $\tilde{X} \geq 0$ whenever $\hat{X} = K$ due to the admissibility of $\hat{X} + \tilde{X} \in \mathcal{X}$ which demands that $\hat{X} + \tilde{X} \geq K$.

By the concavity of U and convexity of D, it is clear that $\Psi(\hat{X} + \theta\tilde{X}) \geq (1-\theta)\Psi(\hat{X}) + \theta\Psi(\hat{X} + \tilde{X})$ for any $\theta \in (0,1]$. Then

$$\Psi\left(\hat{X}\right) \geq \Psi\left(\hat{X} + \tilde{X}\right) - \frac{\Psi\left(\hat{X} + \theta\tilde{X}\right) - \Psi(\hat{X})}{\theta}. \tag{4.6}$$

By (4.5), $\lim_{\theta\downarrow 0}\frac{\Psi\left(\hat{X}+\theta\tilde{X}\right)-\Psi(\hat{X})}{\theta} = \frac{d}{d\theta}\Psi(\hat{X}+\theta\tilde{X})\big|_{\theta=0} \leq 0$. After taking limits on both sides of (4.6), $\Psi\left(\hat{X}\right) \geq \Psi\left(\hat{X}+\tilde{X}\right)$, hence $\hat{X}$ is optimal for Problem 2.4. By Theorem 2.5, We can now conclude that $\hat{X}$ is the optimal terminal wealth of Problem 2.2. ▯

In the next three subsections, we apply Theorem 4.2 to establish the existence of optimal solutions for different utility-risk frameworks: (i) Utility-Downside-Risk, (ii) Utility-Strictly-Convex-Risk, and (iii) Mean-Risk. In particular, the positive answers to the first two problems have long been absent in the literature.

4.2. Application to the Utility-Downside-Risk Problem. In this subsection, we take $\mathcal{D}=[0,\infty)$. We assume that $U:[0,\infty)\to[0,\infty)$ is strictly concave, and U and $D:\mathbb{R}\to[0,\infty)$ are continuously differentiable. We consider D to be a downside risk function, so D is positive and strictly convex on $(0,\infty)$ and $D(x)=0$ for $x\leq 0$. Thus, we have $D'(x)>0$ when $x>0$ and $D'(x)=0$ when $x\leq 0$. In this proposed model, the payoff greater than its mean will not be penalized, and only the downside risk would be taken into account. Moreover, we assume that U and D satisfy the following conditions:

$$U'(0)=\infty, U'(\infty)=0 \text{ and } D'(\infty)=\infty. \tag{4.7}$$

Thus any utility functions satisfying the Inada conditions can be covered. Note that this formulation can cover the utility-semivariance problem, its positive answer has a substantial contrast to the nonexistence of an optimal solution to the mean-semivariance problem. We further make the following assumption on the utility function:

Assumption 4.4. *There exists $k_0>0$ so that the inverse of the first-order derivative of U, $(U')^{-1}$, satisfies $(U')^{-1}(k_0\xi)\in\mathcal{L}^2$.* [6]

According to Remark 4.3, we first find an implicit function satisfying (4.1), then the Nonlinear Moment Problem (3.1)-(3.4) can be reduced into a nonlinear programming problem (4.2)-(4.4).

Proposition 4.5. *There exists an implicit function $I:\mathbb{R}\times(0,\infty)\to(0,\infty)$ satisfying:*

$$U'(I(m,y))+\gamma D'(m-I(m,y))-y=0, \quad \text{for any } (m,y)\in\mathbb{R}\times(0,\infty). \tag{4.8}$$

Moreover, this function I possesses the following regularities:

- **(a)** **(i)** *For each $m\in\mathbb{R}$, $I(m,y)$ is strictly decreasing in y on $(0,\infty)$.*
  - **(ii)** *For each $y\in(0,\infty)$, $I(m,y)$ is strictly increasing in m on $\{m\in\mathbb{R}|\, y\geq U'(m)\}$; $I(m,y)=(U')^{-1}(y)\in(0,\infty)$ for all $m\in\{m\in\mathbb{R}|\, y\leq U'(m)\}$.*
- **(b)** *$I(m,y)$ is jointly continuous in $(m,y)\in\mathbb{R}\times(0,\infty)$.*

Proof. See Appendix C.1. ▯

Since the implicit function I never takes value in the boundary of $\mathcal{D}$, so we now look for numbers Y, M and R that solve the following system of equations as described in Remark 4.3:

$$\mathbb{E}[\xi I\left(M,\gamma R+Y\xi\right)]=x_0; \tag{4.9}$$

$$\mathbb{E}[I\left(M,\gamma R+Y\xi\right)]=M; \tag{4.10}$$

$$\mathbb{E}\left[D'\left(M-I\left(M,\gamma R+Y\xi\right)\right)\right]=R. \tag{4.11}$$

[6] This assumption can be satisfied if there exist $\beta\in(0,1)$ and $\gamma>1$ such that $U'(\beta y)\leq\gamma U'(y)$ for all $y>0$, and this condition has been adopted in [42].

Proposition 4.6. *There exist numbers $Y, M, R \in (0, \infty)$ such that the system of nonlinear equations of (4.9)-(4.11) is satisfied. Thus, $(I(M, \gamma R + Y\xi), Y, M, R)$ is a solution of the system of Equations (3.1)-(3.4), where I is given in Proposition 4.5.*

We shall solve for roots Y, M and R one by one via applying the intermediate value theorem successively. We shall provide the main idea here and the technical details will be collected in Appendix C.

Lemma 4.7. *Given $Y, M \in (0, \infty)$, there exists a unique $R = R_{Y,M} \in (0, D'(M))$ satisfying*

$$\mathbb{E}\left[D'\left(M - I\left(M, \gamma R + Y\xi\right)\right)\right] = R; \tag{4.12}$$

or equivalently by (4.8):

$$\mathbb{E}\left[U'\left(I\left(M, \gamma R + Y\xi\right)\right)\right] = Y\mathbb{E}[\xi]. \tag{4.13}$$

Furthermore, $R_{Y,M}$ is strictly increasing in M for a fixed Y and is also strictly increasing in Y for a fixed M.

Lemma 4.8. *Given $Y \in (0, \infty)$ and $R_{Y,M}$ as specified for each $M \in (0, \infty)$ in Lemma 4.7, there exists a unique $M = M_Y \in (0, \infty)$ such that*

$$\mathbb{E}[I\left(M, \gamma R_{Y,M} + Y\xi\right)] = M. \tag{4.14}$$

Furthermore, M_Y is strictly decreasing in Y.

Lemma 4.9. *Given $R_{Y,M}$ and M_Y as specified in Lemmas 4.7 and 4.8 respectively for each $Y, M \in (0, \infty)$, there exists a (not necessarily unique) $Y^* \in (0, \infty)$ such that*

$$\mathbb{E}[\xi I\left(M_Y, \gamma R_{Y,M_Y} + Y\xi\right)] = x_0. \tag{4.15}$$

The existence result of these three Lemmas can be verified via the application of the intermediate value theorem. The technical details for the proof of these lemmas are similar; they will be provided in the Appendix for completeness.

Proof of Proposition 4.6. According to Lemmas 4.7, 4.8 and 4.9, the triple $(Y^*, M_{Y^*}, R_{Y^*, M_{Y^*}})$ solves the system of nonlinear equations in (4.9)-(4.11). ◻

Next, we shall verify that $\hat{X} = I(M, \gamma R + Y\xi)$, where I is given in Proposition 4.5 and the numbers Y, M and R are warranted in Proposition 4.6, belongs to $\mathcal{X}$ and satisfies Condition 4.1. Then, the optimal terminal wealth can be found by Theorem 4.2:

Theorem 4.10. *$\hat{X} = I(M, \gamma R + Y\xi)$ is an optimal terminal wealth to the utility-downside-risk problem, where I is given in Proposition 4.5 and the numbers Y, M and R are warranted in Proposition 4.6.*

Proof. According to Proposition 4.5, $I(m, y)$ is finite on $\mathbb{R} \times (0, \infty)$ and is strictly decreasing in y for a fixed m. We have $0 \leq I(M, Y\xi + \gamma R) \leq I(M, \gamma R) < \infty$, thus $U(I(M, Y\xi + \gamma R))$ and $D(M - I(M, Y\xi + \gamma R))$ are both uniformly bounded. Hence, $X = I(M, Y\xi + \gamma R) \in \mathcal{X}$. Since D' is increasing, $D'(M - I(M, \gamma R)) \leq D'(M - I(M, Y\xi + \gamma R)) \leq D'(M)$, and hence $D'(M - I(M, Y\xi + \gamma R))$ is uniformly bounded and in $\mathcal{L}^2$. Furthermore, by (4.8), $U'(I(M, Y\xi + \gamma R)) = Y\xi + \gamma R - \gamma D'(M - I(M, Y\xi + \gamma R))$, which is in $\mathcal{L}^2$, hence $\hat{X} = I(M, Y\xi + \gamma R)$ satisfies Condition 4.1.

With (Y, M, R) as warranted in Proposition 4.6, $(I(M, Y\xi + \gamma R), Y, M, R)$ solves the Nonlinear Moment Problem (3.1)-(3.4). Then, by Theorem 4.2, $\hat{X} = I(M, Y\xi + \gamma R)$

is an optimal solution to Problem 2.4 with downside risk function D. Finally, by Theorem 2.5, $\hat{X} = I(M, Y\xi + \gamma R)$ is an optimal terminal wealth of utility-downside-risk problem. □

From the construction, we can see that the optimal terminal wealth is actually uniformly bounded; its proof together with the financial motivation will be included in Appendix D.

Remark 4.11.

To show the existence of optimal solution to our present utility-risk problem from Theorem 4.2, we first characterize the optimal terminal wealth in terms of an implicit function and a solution of a system of equations from Nonlinear Moment Problem, as described in Remark 4.3. Then, we show that the implicit function and the solution of equation system exist in Propositions 4.5 and 4.6 through applications of the intermediate value theorem.

4.3. Application to the Utility-Strictly-Convex-Risk Problem. In this subsection, we take $\mathcal{D} = [0,\infty)$. We assume that $U : [0,\infty) \to [0,\infty)$ is strictly concave and continuously differentiable, while $D : \mathbb{R} \to [0,\infty)$ is strictly convex and continuously differentiable. Moreover, we assume that U and D satisfy (4.7) and $D'(-\infty) = -\infty$. Thus any utility functions satisfying the Inada conditions can be covered.

We can establish the existence of the solution of the nonlinear moment problem in (3.1) by using the same approach as in Subsection 4.2. Since most derivations are similar, we only indicate here the major differences from the last subsection.

Proposition 4.12.

There exists an implicit function $I : \mathbb{R}^2 \to (0,\infty)$ satisfying:

$$U'(I(m,y)) + \gamma D'(m - I(m,y)) - y = 0, \quad \textit{for any } (m,y) \in \mathbb{R}^2. \tag{4.16}$$

Moreover, this function I possesses the following regularities:

(a) **(i)** *For each $y \in \mathbb{R}$, $I(m,y)$ is strictly increasing in m.*
(ii) *For each $m \in \mathbb{R}$, $I(m,y)$ is strictly decreasing in y.*

(b) *$I(m,y)$ is jointly continuous in $(m,y) \in \mathbb{R}^2$.*

Proof. See Appendix C.5. □

Proposition 4.13. *There exist constants $Y, M \in (0,\infty)$ and $R \in \mathbb{R}$ satisfying a system of nonlinear equations in (4.9)-(4.11). Thus, $(I(M, \gamma R + Y\xi), Y, M, R)$ is the solution of system of Equations (3.1)-(3.4), where I is given in Proposition 4.12.*

Proof. The approach is again the same as that of Proposition 4.6. Major changes will be demonstrated in Appendix C.6. □

Using the same argument as in Section 4.2, we can draw the same existence conclusion:

Theorem 4.14. *$\hat{X} = I(M, \gamma R + Y\xi)$ is an optimal terminal wealth of the utility-strictly-convex-risk problem, where I is specified in Proposition 4.12 and the numbers Y, M and R are warranted in Proposition 4.13.*

Furthermore, if we specify risk function D to be the square function, i.e. $D(x) = x^2$, and hence variance of the terminal payoff is the risk measure concerned, then the solution of the Nonlinear Moment Problem in Theorem 4.2 is unique:

Proposition 4.15. *There exists a unique set of numbers $Y, M \in (0,\infty)$ and $R \in \mathbb{R}$ such that a system of nonlinear equations in (4.9)-(4.11) is satisfied. Thus, $X = I(M, \gamma R + Y\xi)$, where I is a function defined in Proposition 4.12, is the unique optimal terminal wealth of the utility-variance problem.*

Proof. Now, $D'(x) = 2x$, then $R_{Y,M_Y} = 0$ for all Y by (4.10) and (4.11). Since M_Y is strictly decreasing in Y, by Proposition 4.12 (b), $I\left(M_Y, \gamma R_{Y,M_Y} + Y\xi\right) = I\left(M_Y, Y\xi\right)$ is strictly decreasing in Y. Because ξ is absolute continuous with no point mass and its support is $\mathbb{R}$, hence $\mathbb{E}\left[\xi I\left(M_Y, \gamma R_{Y,M_Y} + Y\xi\right)\right]$ is strictly decreasing in Y. Therefore, Y^* obtained in Proposition 4.13 is unique. Thus, (4.9)-(4.11) is uniquely solved by $(Y^*, M_{Y^*}, 0)$.

By remark 4.3, $\left(I\left(M_{Y^*}, Y^*\xi\right), Y^*, M_{Y^*}, 0\right)$ solve the Nonlinear Moment Problem. By Theorems 2.5 and 4.2, the second assertion follows. □

4.4. Application to the Mean-Risk Problem. In this subsection, we assume the utility function to be linear, i.e. $U(x) = x$, and we set $\mathcal{D} = \mathbb{R}$. Our Problem 2.2 reduces to a mean-risk optimization problem:

$$\max_{\pi\in\mathcal{A}} \mathbb{E}[X^\pi(T)] - \gamma\mathbb{E}[D\left(\mathbb{E}[X^\pi(T)] - X^\pi(T)\right)]. \tag{4.17}$$

As the Inada conditions in (4.7) do not hold in this case, the method developed in the previous subsection cannot be directly translated here. Suppose that there is an inverse function for the first-order derivative of risk function, $I_2 := (D')^{-1}$. The Nonlinear Moment Problem (3.1) corresponding to (4.17) can be simplified as follows:

$$Y\xi = 1 - \gamma R + \gamma D'\left(M - \hat{X}\right), \tag{4.18}$$

where the numbers $Y, M, R \in \mathbb{R}$ satisfy

$$\mathbb{E}[\xi\hat{X}] = x_0, \tag{4.19}$$

$$\mathbb{E}[\hat{X}] = M, \tag{4.20}$$

$$\mathbb{E}\left[D'\left(M - \hat{X}\right)\right] = R. \tag{4.21}$$

In accordance with Theorem 4.2, we are going to show that the reduced Nonlinear Moment Problem (4.18) admits a solution, so that the corresponding $\hat{X}$ will be an optimal terminal wealth for the mean-risk problem (4.17).

THEOREM 4.16. *If there exists a unique $R \in \mathbb{R}$ so that:*

$$I_2\left(R + \frac{\xi}{\gamma\mathbb{E}[\xi]} - \frac{1}{\gamma}\right) \in \mathcal{L}^2 \qquad \textit{and} \tag{4.22}$$

$$\mathbb{E}\left[I_2\left(R + \frac{\xi}{\gamma\mathbb{E}[\xi]} - \frac{1}{\gamma}\right)\right] = 0, \tag{4.23}$$

then by setting

$$\hat{X} := M - I_2\left(R + \frac{\xi Y}{\gamma} - \frac{1}{\gamma}\right), \tag{4.24}$$

$$M := \frac{x_0}{\mathbb{E}[\xi]} + \frac{\mathbb{E}\left[\xi I_2\left(R + \frac{\xi Y}{\gamma} - \frac{1}{\gamma}\right)\right]}{\mathbb{E}[\xi]}, \tag{4.25}$$

$$Y := \frac{1}{\mathbb{E}[\xi]}, \tag{4.26}$$

together with R, they will solve the reduced Nonlinear Moment Problem (4.18)-(4.21).

Proof. Condition (4.22) guarantees that X defined in (4.24) is in $\mathcal{L}^2$ and $\mathbb{E}\left[\xi I_2\left(R+\frac{\xi Y}{\gamma}-\frac{1}{\gamma}\right)\right]$ is finite, by the Cauchy-Schwarz inequality. It is clear that Condition 4.1 is satisfied. Since D is convex,

$$D\left(I_2\left(R+\frac{\xi Y}{\gamma}-\frac{1}{\gamma}\right)\right)\le D(0)+\left(R+\frac{\xi Y}{\gamma}-\frac{1}{\gamma}\right)I_2\left(R+\frac{\xi Y}{\gamma}-\frac{1}{\gamma}\right),$$

hence $I_2\left(R+\frac{\xi Y}{\gamma}-\frac{1}{\gamma}\right)\in\mathcal{X}$. (4.18)–(4.21) can be verified easily by direct substitution of (4.24)–(4.26). □

In Theorem 4.16, we solve the mean-risk optimization problem for any risk function satisfying (4.22) and (4.23). (4.23) can be obviously satisfied when I_2 is both continuous and coercive. Note that the uniqueness of R can be warranted by the strict convexity of D and (4.22) can be satisfied when I_2 is of polynomial growth.

REMARK 4.17. *In [14], they studied the same mean-risk optimization problem by using the Lagrangian approach, and they also formulated the problem as follows:*

$$\min D(\mathbb{E}[X(T)]-X(T)),\ \textit{subject to } \mathbb{E}[X(T)]=z.$$

This problem is equivalent to (4.17) for appropriate relationship between γ and z. The work [14] shows that if the mean-risk problem has a solution, the optimal terminal wealth $X=z-I_2(\mu\xi-\lambda)$, where λ and μ satisfy the equations

$$\begin{aligned}\mathbb{E}[I_2(\mu\xi-\lambda)]&=0,\\ \mathbb{E}[\xi I_2(\mu\xi-\lambda)]&=z\mathbb{E}[\xi]-x_0.\end{aligned}$$

For any z such that there exists $\gamma>0$ satisfying

$$z=\frac{x_0}{\mathbb{E}[\xi]}+\frac{\mathbb{E}\left[\xi I_2\left(R+\frac{\xi}{\gamma\mathbb{E}[\xi]}-\frac{1}{\gamma}\right)\right]}{\mathbb{E}[\xi]},$$

if we set

$$\mu=\frac{1}{\gamma\mathbb{E}[\xi]},\lambda=\frac{1}{\gamma}-R,$$

where R is as obtained in (4.23), the solution in [14] can then be recovered.

EXAMPLE 4.18 (Mean-Weighted-Power-Risk Function case).

Consider

$$D(x)=\frac{a}{2}\frac{x_+^{\rho+1}}{\rho+1}-\frac{b}{2}\frac{x_-^{\rho+1}}{\rho+1}$$

for $\rho>0$ and $a\ge b>0$. $a\ge b$ means that the risk incurred when the return is less than the expectation will be greater than that when the return is greater than the expectation. Now, $D'(x)=ax_+^\rho-bx_-^\rho$, and $I_2(x)=\frac{1}{a}x_+^{\frac{1}{\rho}}-\frac{1}{b}x_-^{\frac{1}{\rho}}$. To verify (4.22), we consider two cases: (i) $\rho\le 2$ and (ii) $\rho>2$ respectively.

(i) If $\rho\le 2$, by Minkowski's inequality, for any $R\in\mathbb{R}$,

$$\begin{aligned}\mathbb{E}\left[\left(I_2\left(\frac{1}{\gamma\mathbb{E}[\xi]}\xi+R-\frac{1}{\gamma}\right)\right)^2\right]&\le\mathbb{E}\left[\frac{1}{b^2}\left|\frac{1}{\gamma\mathbb{E}[\xi]}\xi+R-\frac{1}{\gamma}\right|^{\frac{2}{\rho}}\right]\\&\le\frac{1}{b^2}\left(\mathbb{E}\left[\left|\frac{\xi}{\gamma\mathbb{E}[\xi]}\right|^{\frac{2}{\rho}}\right]^{\frac{\rho}{2}}+\left|R-\frac{1}{\gamma}\right|\right)^{\frac{2}{\rho}}.\end{aligned}$$

It is clear that $\mathbb{E}\left[\xi^k\right]$ *is bounded for any* $k \in \mathbb{R}$, *and therefore* $I_2\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right) \in \mathcal{L}^2$, *i.e. (4.22) is satisfied.*

(ii) If $\rho > 2$, *for any* $R \in \mathbb{R}$,

$$\begin{aligned}
&\mathbb{E}\left[\left(I_2\left(\frac{1}{\gamma\mathbb{E}[\xi]}\xi + R - \frac{1}{\gamma}\right)\right)^2\right] \\
&\le \frac{1}{b^2}\left(\frac{2^{\frac{1}{\rho}}}{\gamma^{\frac{2}{\rho}}\left(\mathbb{E}[\xi]\right)^{\frac{2}{\rho}}}\mathbb{E}\left[\xi^{\frac{2}{\rho}}\right] + \frac{2}{p}\left(\frac{2}{\gamma^2\left(\mathbb{E}[\xi]\right)^2}\right)^{\frac{1}{\rho}-1}\left(R - \frac{1}{\gamma}\right)^2\mathbb{E}\left[\xi^{2\left(\frac{1}{\rho}-1\right)}\right]\right),
\end{aligned}$$

by concavity of $x^{\frac{1}{\rho}}$. *By the fact that* $\mathbb{E}\left[\xi^k\right]$ *is bounded for any* $k \in \mathbb{R}$, *(4.22) is satisfied.*

Note that the expression

$$I_2\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right) = \frac{1}{a}\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)_+^{\frac{1}{\rho}} - \frac{1}{b}\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)_-^{\frac{1}{\rho}}$$

is increasing in R *and* $\mathcal{L}^1$*-integrable by Jensen's inequality, for all* $R \in \mathbb{R}$. *Thus,* $\mathbb{E}\left[I_2\left(R + \frac{\xi}{\gamma\mathbb{E}[\xi]} - \frac{1}{\gamma}\right)\right]$ *is continuous in* R *by the Dominated Convergence Theorem. It is not difficult to use the Monotone Convergence Theorem to show that* $\mathbb{E}\left[I_2\left(R + \frac{\xi}{\gamma\mathbb{E}[\xi]} - \frac{1}{\gamma}\right)\right]$ *is coercive in the sense that*

$$\lim_{R\to-\infty}\mathbb{E}\left[I_2\left(R + \frac{\xi}{\gamma\mathbb{E}[\xi]} - \frac{1}{\gamma}\right)\right] = -\infty, \qquad \lim_{R\to\infty}\mathbb{E}\left[I_2\left(R + \frac{\xi}{\gamma\mathbb{E}[\xi]} - \frac{1}{\gamma}\right)\right] = \infty.$$

By the intermediate value theorem, there exists a unique $R \in \mathbb{R}$ *so that (4.23) and (4.22) are satisfied. The solution of the mean-weighted-power-risk problem is:*

$$\begin{aligned}
\hat{X} = &\frac{1}{\mathbb{E}[\xi]}\left(x_0 + \mathbb{E}\left[\frac{\xi}{a}\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)_+^{\frac{1}{\rho}} - \frac{\xi}{b}\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)_-^{\frac{1}{\rho}}\right]\right) \\
&-\frac{1}{a}\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)_+^{\frac{1}{\rho}} + \frac{1}{b}\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)_-^{\frac{1}{\rho}},
\end{aligned} \tag{4.27}$$

where R *is the unique root of the equation*

$$\mathbb{E}\left[\frac{1}{a}\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)_+^{\frac{1}{\rho}} - \frac{1}{b}\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)_-^{\frac{1}{\rho}}\right] = 0. \tag{4.28}$$

Remark 4.19.

If $\rho = 1$, *this mean-weighted-power-risk model becomes the mean-weighted-variance one, studied in [14]. The results in [14] can be recovered by choosing* $\mu = \frac{1}{\gamma\mathbb{E}[\xi]}$, $\lambda = \frac{1}{\gamma} - R$ *where* $\frac{1}{\gamma}$ *is selected such that*

$$z = \frac{1}{\mathbb{E}[\xi]}\left(x_0 + \mathbb{E}\left[\frac{\xi}{a}\left(R + \frac{\xi}{\gamma\mathbb{E}[\xi]} - \frac{1}{\gamma}\right)_+ - \frac{\xi}{b}\left(R + \frac{\xi}{\gamma\mathbb{E}[\xi]} - \frac{1}{\gamma}\right)_-\right]\right).$$

If $a = b = 1$, this mean-weighted-variance model further becomes the classical mean-variance setting. We can easily get that $R = 0$ from (4.28). Then we can recover the following solution:

$$\hat{X} = \frac{x_0}{\mathbb{E}[\xi]} + \frac{1}{\gamma}\left(\frac{\mathbb{E}[\xi^2]}{(\mathbb{E}[\xi])^2} - \frac{\xi}{\mathbb{E}[\xi]}\right).$$

This result can coincide with the solution on P.226–227 in [4] by choosing

$$\mu = \frac{1}{\gamma\mathbb{E}[\xi]}, \lambda = \frac{x_0}{\mathbb{E}[\xi]} + \frac{\mathbb{E}[\xi^2]}{\gamma\left(\mathbb{E}[\xi]\right)^2}, \text{ where } \frac{1}{\gamma} = \frac{\mathbb{E}[\xi]\left(z\mathbb{E}[\xi] - x_0\right)}{Var[\xi]}.$$

Example 4.20 (Mean-Exponential-Risk Function Case). *We further revisit another example found in [14]. Consider the exponential risk function $D(x) = e^x$. Then $D'(x) = e^x$ and $I_2(x) = \ln x$ for $x > 0$.*

Proposition 4.21. *Mean-Exponential-Risk Problem possesses an optimal solution if and only if $\gamma \geq \exp\left(\mathbb{E}\left[\ln\left(\frac{\xi}{\mathbb{E}[\xi]}\right)\right]\right)$*[7]*. Furthermore, if the problem possesses an optimal solution, the optimal terminal wealth is*

$$\hat{X} = \frac{1}{\mathbb{E}[\xi]}\left(x_0 + \mathbb{E}\left[\xi\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)\right]\right) - \ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right), \tag{4.29}$$

where $R \in [\frac{1}{\gamma}, \infty)$ is the unique root of the equation:

$$\mathbb{E}\left[\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)\right] = 0.$$

Proof. See Appendix C.7. □

5. Conclusion. In this paper, we studied the utility risk portfolio selection problem. We derived the Nonlinear Moment Problem in (3.1)-(3.4), whose solution can completely characterize the optimal terminal wealth by the necessity and sufficiency results in Theorems 3.2 and 4.2 respectively. The nonexistence of optimal solution for the mean-semivariance problem can be revisited by the application of Theorem 3.2. Furthermore, we applied Theorem 4.2 to establish the existence of optimal solutions for the utility-downside-risk and utility-strictly-convex-risk problems. Their resolutions have long been missing in the literature, and the positive answer in utility-downside-risk problem is in big contrast to the negative answer in mean-downside-risk problem; with our present result, we can now use semivariance as a proper risk measure in portfolio selection. Finally, we established the sufficient condition for the Nonlinear Moment Problem through which the existence of optimal solution of mean-risk problem can be ensured.

6. Acknowledgments. We express our sincere gratitude to the editor and anonymous referees for their very useful suggestions and inspiring comments which much enrich our article.We are also very grateful to our colleague, Nicholas Bingham, for his suggestions on enhancing the presentation of our paper. The second author expresses his sincere gratitude to the hospitality of the Hausdorff Center for Mathematics of the University of Bonn for his fruitful stay in the Hausdorff Trimester Program with title, "Stochastic Dynamics in Economics and Finance".

[7] $\mathbb{E}\left[\ln\left(\frac{\mathbb{E}[\xi]}{\xi}\right)\right]$ is known to be the Kullback-Leibler Divergence (relative entropy) from $\mathbb{P}$ to $\mathbb{Q}$, the risk neutral measure.

REFERENCES

[1] J. Aubin and I. Ekeland, *Applied nonlinear analysis*, John Wiley and Sons, 1984.
[2] S. Basak and A. Shapiro, *Value-at-risk-based risk management: optimal policies and asset prices*, Rev. Financial Stud. 14 (2001), pp. 371–405.
[3] A. Bensoussan, K. C. Wong, and S. C. P. Yam, *Mean-variance precommitment policies revisited via a mean-field technique*, Recent Advances In Financial Engineering 2012, World Scientific, Singapore, pp. 177–198, 2014.
[4] T. R. Bielecki, H. Jin, S. R. Pliska, and X. Y. Zhou , *Continuous time mean-variance portfolio selection with bankruptcy prohibition*, Math. Finance, 15 (2005), pp. 213–244.
[5] X. Cao, J. Gao, D. Li, and K. Zhou, *Dynamic mean-LPM and mean-CVaR portfolio optimization in continuous-time*, working paper. http://arxiv.org/abs/1402.3464
[6] P. Chen and H. Yang, *Markowitzs mean-variance asset-liability management with regime switching: A multi-period model*, Appl. Math. Finance, 18 (2011), pp. 29–50.
[7] P. Chen, H. Yang, and G. Yin, *Markowitzs mean-variance asset-liability management with regime switching: A continuous-time model*, Insurance Math. Econom., 43 (2008), pp. 456–465.
[8] D. Cuoco, H. He, and S. Isaenko, *Optimal dynamic trading strategies with risk limits*, Oper. Res., 56 (2008), pp. 358–368.
[9] G. Deelstra, H. Pham, and N. Touzi, *Dual formulation of the utility maximization problem under transaction costs*, Ann. Appl. Probab., 11 (2001), pp. 1353–1383.
[10] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, Springer, New York, 2006.
[11] X. D. He, H. Jin, and X. Y. Zhou, *Dynamic portfolio choice when risk is measured by weighted VaR*, Math. Oper. Res., forthcoming. http://dx.doi.org/10.1287/moor.2014.0695.
[12] W. W. Hogan and J. M. Warren, *Computation of the efficient boundary in the E-S portfolio selection model*, J. Financial Quantit. Anal., 7 (1972), pp. 1881–1896.
[13] H. Jin, H. Markowitz, and X. Y. Zhou, *A note on semivariance*, Math. Finance, 16(2006), pp. 53–62.
[14] H. Jin, J. Yan, and X. Y. Zhou, *Continuous-time mean-risk portfolio selection.* Ann. Inst. H. Poincaré Probab. Stat., 41 (2005), pp. 559–580.
[15] I. Karatzas, J. P. Lehoczky, and S. E. Shreve, *Optimal portfolio and consumption decisions for a "small investor" on a finite horizon*, SIAM J. Control Optim., 25 (1987), pp. 1557–1586.
[16] I. Karatzas, J. P. Lehoczky, S. E. Shreve, and G. L. Xu, *Martingale and duality methods for utility maximization in an incomplete market*, SIAM J. Control Optim., 29 (1991), pp. 702–730.
[17] I. Karatzas and S. E. Shreve, *Methods of mathematical finance*, Springer-Verlag New York (1998).
[18] D. Kramkov and W. Schachermayer, *The asymptotic elasticity of utility functions and optimal investment in incomplete markets*, Ann. Appl. Probab., 9 (1999), pp. 904–950.
[19] D. Kramkov and W. Schachermayer, *Necessary and sufficient conditions in the problem of optimal investment in incomplete markets*, Ann. Appl. Probab., 13 (2003), pp. 1504–1516.
[20] Y. Kroll, H. Levy, and H. M. Markowitz, *Mean-variance versus direct utility maximization*, J. Finance, 39 (1984), pp. 47–61.
[21] M. Leippold, F. Trojani, and P. Vanini, *Equilibrium impact of value-at-risk regulation*, J. Econom. Dynam. Control, 30 (2006), pp. 1277–1313.
[22] H. Levy and H. M. Markowitz, *Approximating expected utility by a function of mean and variance*, Amer. Econom. Rev., 69 (1979), pp. 308–317.
[23] D. Li and X. Y. Zhou *Continuous-time mean-variance portfolio selection: a stochastic LQ framework*, Appl. Math. Optim., 42 (2000), pp. 19–33.
[24] A. E. B. Lim and X. Y. Zhou, *Mean-variance portfolio selection with random coefficients in a complete market*, Math. Oper. Res., 27 (2002), pp. 101–120.
[25] H. Markowitz, *Portfolio selection*, J. Finance, 7 (1952), pp. 77–91.
[26] H. Markowitz, *The optimization of a quadratic function subject to linear constraints*, Naval Res. Logist., 3 (1956), pp. 111–133.
[27] H. Markowitz, *Portfolio selection: efficient diversification of investments*, Wiley, New York, 1959.
[28] H. Markowitz, *Foundations of portfolio theory*, J. Finance, 46 (1990), pp. 469–477.
[29] H. Markowitz, *Portfolio theory: as I still see it*, Annu. Rev. Financial Econom., 2 (2010), pp. 1–23.
[30] R. C. Merton, *Optimum consumption and portfolio rules in a continuous-time model*, J.

Econom. Theory, 3 (1971), pp. 373–413.
[31] R. C. Merton, *An analytical derivation of the efficient portfolio frontier*, J. Financial Quant. Anal., 7 (1972), pp. 1851–1872.
[32] T. J. Nantell and B. Price, *An analytical comparison of variance and semivariance capital market theories*, J. Financial Quantit. Anal., 14 (1979), pp. 221–242.
[33] D. Nawrocki, *A brief history of downside risk measures*, J. Investing, 8 (1999), pp. 9–25.
[34] H. Pham, *Continuous-time stochastic control and optimization with financial applications*, Springer-Verlag Berlin Heidelberg, 2009.
[35] S. R. Pliska, *A stochastic calculus model of continuous trading: optimal portfolios*, Math. Oper. Res., 11 (1986), pp. 371–384.
[36] L. B. Pulley, *A general mean-variance approximation to expected utility for short holding periods*, J. Financial Quantit. Anal., 16 (1981), pp. 361–373.
[37] P. Samuelson, *Lifetime portfolio selection by dynamic stochastic programming*, Rev. Econom. Statist., 51 (1969), pp. 239–246.
[38] M. C. Steinbach, *Markowitz revisited: mean-variance models in financial portfolio analysis*, SIAM Rev., 43 (2001), pp. 31–85.
[39] N. Touzi, *Stochastic control problems, viscosity solutions, and application to finance*, Lecture Note, 2002.
[40] N. Touzi, *Optimal stochastic control, stochastic target problems, and backward SDE*, Springer, New York, 2013.
[41] K. F. C. Yiu, *Optimal portfolios under a value-at-risk constraint*, J. Econom. Dynam. Control, 28 (2004), pp. 1317–1334.
[42] H. Zheng, *Efficient frontier of utility and conditional VaR*, Math. Methods Oper. Res., 70 (2009), pp. 129–148.

Appendix A. Discussion on the Possible use of Standard Approaches to Utility-Risk Problem.

In the existing literature, there are three standard approaches of tackling the optimization problems: (i) Direct/ Primal method: constructing the optimal solution by considering a weak limit of an optimizing sequence or (ii) Dual method: applying usual convex analysis through the conjugate functions of objective functions.

A.1. Direct Approach. One can find a comprehensive approach from a book chapter of Chapter 7.3.2 in Pham [34].

Assume that the value function is non-degenerate, i.e. the optimal objective functional is finite such that

$$V(x_0) := \sup_{X\in\mathcal{X},\mathbb{E}[\xi X]=x_0} \Psi(X) = \sup_{X\in\mathcal{X},\mathbb{E}[\xi X]\le x_0} \Psi(X) < \infty, \tag{A.1}$$

where Ψ was defined in (2.3) and the equality holds since the functional Ψ is concave and an admissible $X \in \mathcal{X}$ will never be optimal to Problem 2.4 if $\mathbb{E}[\xi X]$ cannot take the largest possible value x_0[8]. This primal method commonly applied in literature is to directly construct the optimal terminal wealth by using the Komlos Theorem; also see [34]. The finiteness of the value function (A.1) implies the existence of a maximizing sequence $\{X_n\} \in \mathcal{X}$ such that $\Psi(X_n) \to V(x_0)$ and $\mathbb{E}[\xi X_n] = x_0$ for all n. To avoid unnecessary technicalities, we further assume that

$$\sup_n \mathbb{E}[|X_n|^2] < \infty; \tag{A.2}$$

also see the discussion in Remark A.3. According to Komlos Theorem, there exists a subsequence $\{Y_k := X_{n_k}\}$ and a limit $Z^* \in \mathcal{L}^2$ such that $Z_k := \frac{1}{k}\sum_{i=1}^k Y_i \to Z^*$ a.s., so that the following facts hold:

$$\mathbb{E}[\xi Z_k] = x_0 \text{ for all } k; \tag{A.3}$$

$$\lim_{k\to\infty} \frac{1}{k}\sum_{i=1}^k \Psi(Y_i) = V(x_0); \tag{A.4}$$

$$\Psi(Z_k) \ge \frac{1}{k}\sum_{i=1}^k \Psi(Y_i). \tag{A.5}$$

We aim to show that Z^* is an optimal terminal wealth, thus we need to show (i) $V(x_0) = \Psi(Z^*)$, (ii) $Z^* \in \mathcal{X}$, and (iii) $\mathbb{E}[Z^*\xi] \le x_0$. The latter statement is clear by using the fact (A.3) and applying Fatou's Lemma since $Z^* \ge 0$ a.s. However, the first two claims are not necessarily immediate; indeed, consider a special case of utility-only maximization, i.e. $D \equiv 0$, for if $\{U(Z_k)\}$ is uniformly integrable, we immediately have $U(Z^*) \in \mathcal{L}^1$ through Fatou's lemma. Meanwhile, by (A.4) and (A.5), we have

$$\begin{aligned}\Psi(Z^*) = \mathbb{E}[U(Z^*)] = \mathbb{E}\left[\lim_{k\to\infty} U(Z_k)\right] &= \lim_{k\to\infty} \mathbb{E}\left[U(Z_k)\right] = \lim_{k\to\infty} \Psi\left(Z_k\right)\\ &\ge \lim_{k\to\infty} \frac{1}{k}\sum_{i=1}^k \Psi(Y_i) = V(x_0).\end{aligned}$$

[8] Indeed, for if $\mathbb{E}[\xi X] < x_0$, there is a constant $\theta > 0$ such that $\mathbb{E}[\xi(X+\theta)] = x_0$. Then, we have $X + \theta \in \mathcal{X}$ and $\Psi(X+\theta) > \Psi(X)$, which means that X cannot be optimal.

In summary, to prove the existence of optimal terminal wealth using the common primal approach, even in the special case of utility-only maximization, there is an outstanding technical issue: to check whether $\{U(Z_k); k = 1, 2, \dots\}$ is uniformly integrable. This issue is not too certain in general; to address this, [34] made the following assumption:

$$\limsup_{x \to \infty} \frac{V(x)}{x} \leq 0. \tag{A.6}$$

However, as also pointed out in [34], (A.6) is hard to check in most practical considerations, because the value function can barely be characterized without the prior knowledge of the optimal solution. One may attempt to use dynamic programming principle to find the HJB equation which characterizes the value function; however, in general, except some common utility function such as power of logarithm, there is no explicit guess solution of the value function for the HJB equation. Even worst, in the presence of deviation risk measure, then the objective function does not admit Tower property, we even cannot apply dynamic programming principle to this utility-risk problem, and no HJB equation can even be obtained in this general case. Instead, one may resort to the use of Legendre transform of U especially when the asymptotic elasticity of U is less than 1; also see [18] and Section 7.3.3 in [34]. Hence, it is unlikely to solve for the present utility-risk problem by using this primal approach.

A.2. Dual Approach. Under this approach, both the existence and characterization of an optimal solution are established through the use of the conjugate function (Legendre transform) of the objective function and utilizing the convex analysis.

Under the case of utility-only maximization:

$$\max_{X \in \mathcal{X}} \mathbb{E}[U(X)] \text{ subject to } \mathbb{E}[\xi X] = x_0,$$

where $U : [0, \infty) \to \mathbb{R}$ is concave, we can utilize the convex analysis over the finite dimensional space and consider the following conjugate function of U:

$$U^*(y) := \sup_{x \in \mathbb{R}} \{U(x) - xy\} \text{ for } y \geq 0. \tag{A.7}$$

Since $U(x) - xy$ is concave, its maximizer is $(U')^{-1}(y)$. The Fenchel's inequality implies that $U(x) \leq U^*(y) + xy$ for any $x, y \geq 0$. Then, we further have

$$\mathbb{E}[U(X)] \leq \mathbb{E}[U^*(y\xi)] + \mathbb{E}[Xy\xi] = \mathbb{E}[U^*(y\xi)] + yx_0, \tag{A.8}$$

for any $\{X \in \mathcal{X} \,|\, \mathbb{E}[\xi X] = x_0\}$ and $y \geq 0$. If we can find X^* and y^* so that the equality in (A.8) holds, then X^* is optimal for (A.7). To achieve this, $X^* = (U')^{-1}(y\xi)$ so that y^* satisfies $\mathbb{E}[\xi X^*] = x_0$. For a more comprehensive details of using dual method in the case of utility-only maximization, one can refer to the book by Karatzas and Shreve [17].

However, under the general utility-risk setting, due to the presence of the expectation $\mathbb{E}[X]$ in the deviation risk term, we cannot have the usual conjugate function of the objective function as in (A.7) over finite finite dimensional space. Hence, we have to use the convex analysis developed over the abstract infinite dimensional space. For the detailed discussion on this dual approach on resolving general convex optimization problem over abstract infinite dimensional space, one can consult the textbook by Aubin and Ekeland [1]. As discussed in Chapter 4 in [1], a sufficient condition for the existence of an optimal solution for the primal problem is verifying the subdifferentiability of the conjugate function of the original objective function at zero, which demands that the dual function is well-defined in the neighborhood of zero.

More specifically, we reformulate our problem as follows:

$$\inf_{X\in\mathcal{X}} \mathcal{J}(X), \tag{A.9}$$

where $\mathcal{J}(X) := \mathcal{U}(X) + \mathcal{V}(\mathcal{A}(X))$, $\mathcal{U} : \mathcal{L}^2 \to \mathbb{R} \cup \{\infty\}$ is defined by

$$\mathcal{U}(X) := \begin{cases} -\mathbb{E}[U(X)] + \gamma\mathbb{E}[D(\mathbb{E}(X) - X)], & X \in \mathcal{X}; \\ \infty, & X \notin \mathcal{X}; \end{cases}$$

and $\mathcal{V} : \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is defined by

$$\mathcal{V}(y) := \begin{cases} 0, & \text{if } y = x_0; \\ \infty, & \text{otherwise}; \end{cases}$$

and the continuous linear functional $\mathcal{A} : \mathcal{L}^2 \to \mathbb{R}$ is given by $\mathcal{A}(X) := \mathbb{E}[X\xi]$. Then, the conjugate functions of $\mathcal{U}$, $\mathcal{V}$, and $\mathcal{A}$ are given by

$$\begin{aligned} \mathcal{U}^*(Y) &:= \sup_{X\in\mathcal{L}^2} \left\{\mathbb{E}[XY] - \mathcal{U}(X)\right\}, \text{ for any } Y \in \mathcal{L}^2; \\ \mathcal{V}^*(p) &:= \sup_{y\in\mathbb{R}} \{py - \mathcal{V}(y)\} = px_0, \text{ for any } p \in \mathbb{R}; \\ \mathcal{A}^*(y) &:= y\xi, \text{ for any } y \in \mathbb{R}. \end{aligned}$$

Under the above formulation, Aubin and Ekeland [1] provided a sufficient condition for the existence of an optimal solution for the primal problem (A.9):

Theorem A.1 (Theorem 1a in Section 4.6, [1]). *Assume that $\mathcal{U}$ and $\mathcal{V}$ are lower-semicontinuous and convex. If*

$$0 \in Int\left(\mathcal{A}^* Dom\left(\mathcal{V}^*\right) + Dom\left(\mathcal{U}^*\right)\right), \tag{A.10}$$

where $Dom\left(\mathcal{V}^*\right) := \{p \in \mathbb{R} \,|\, \mathcal{V}^*(p) < \infty\}$ and $Dom\left(\mathcal{U}^*\right) := \{Y \in \mathcal{L}^2 \,|\, \mathcal{U}^*(Y) < \infty\}$. Then there exists a solution $\overline{X} \in \mathcal{L}^2$ to the problem (A.9).

In the proof of Theorem A.1, Condition (A.10) is essential to establish that $K_\alpha := \{X \in \mathcal{X} \,|\, \mathcal{J}(X) \le \alpha\}$ is weakly bounded and thus weakly relatively compact for every λ. Then, by the lower semicontinuity of $\mathcal{J}$, the existence of any optimal solution of the primal problem (A.9) can be warranted.

With Theorem A.1, the existence of an optimal solution of the primal problem (A.9) is warranted if Condition (A.10) holds. It is obvious that $\mathcal{A}^* Dom\left(\mathcal{V}^*\right) = \{y\xi \,|\, y \in \mathbb{R}\} \subset \mathcal{L}^2$, thus Condition (A.10) is equivalent to that there exists $\delta > 0$ such that, for any $\overline{X} \in \mathcal{L}^2$ with $\left\|\overline{X}\right\|_2 \le \delta$, there exists $\overline{y} \in \mathbb{R}$ such that $\overline{X} - \overline{y}\xi \in Dom\left(\mathcal{U}^*\right)$. However, such Condition (A.10) is still not so apparent in most concrete problems, and we have to study case by case.

Now, we consider a specific example of power utility-semivariance optimization, i.e. $U(x) = \begin{cases} \frac{1}{1-\rho}x^{1-\rho}, & x \ge 0 \\ -\infty, & x < 0 \end{cases}$, where $\rho \in (0, 1)$, and $D(x) = x_+^2$. For a fixed $\delta > 0$ and $y \in \mathbb{R}$, take $\overline{X} = \frac{\delta\xi^{-1}}{\|\xi^{-1}\|_2}$,

$$\mathcal{U}^*\left(\frac{\delta\xi^{-1}}{\|\xi^{-1}\|_2} - y\xi\right) = \sup_{X\in\mathcal{L}^2} \left\{\mathbb{E}\left[\left(\frac{\delta\xi^{-1}}{\|\xi^{-1}\|_2} - y\xi\right) X\right] + \frac{1}{1-\rho}\mathbb{E}\left[X^{1-\rho}\right] - \gamma\mathbb{E}\left[(\mathbb{E}[X] - X)_+^2\right]\right\}.$$

Construct a sequence $X_N := N\mathbb{I}\{\xi < b_N\}$, where b_N is the smallest possible such that $\mathbb{P}\{\xi < b_N\} = \frac{1}{N}$. It is obvious that $\lim_{N\to\infty} b_N = 0$. Hence, we have

$$\begin{aligned}\mathcal{U}^*\left(\frac{\delta\xi^{-1}}{\|\xi^{-1}\|_2} - y\xi\right) &\geq \mathbb{E}\left[\left(\frac{\delta\xi^{-1}}{\|\xi^{-1}\|_2} - y\xi\right)X_N\right] + \frac{1}{1-\rho}\mathbb{E}\left[X_N^{1-\rho}\right] - \gamma\mathbb{E}\left[(\mathbb{E}[X_N] - X_N)_+^2\right] \\ &\geq N\mathbb{E}\left[\left(\frac{\delta\xi^{-1}}{\|\xi^{-1}\|_2} - y\xi\right)\mathbb{I}\{\xi < b_N\}\right] - \gamma\mathbb{E}\left[N^2(\mathbb{P}\{\xi < b_N\} - \mathbb{I}\{\xi < b_N\})_+^2\right] \\ &\geq N\mathbb{E}\left[\left(\frac{\delta}{\|\xi^{-1}\|_2}b_N^{-1} - yb_N\right)\mathbb{I}\{\xi < b_N\}\right] - \gamma \\ &= \left(\frac{\delta}{\|\xi^{-1}\|_2}b_N^{-1} - yb_N\right) - \gamma \to \infty, \text{ as } N \to \infty.\end{aligned}$$

Thus, for any $\delta > 0$, $\frac{\delta\xi^{-1}}{\|\xi^{-1}\|_2} - y\xi \notin \mathit{Dom}\,(\mathcal{U}^*)$ for all $y \in \mathbb{R}$, which means that $0 \notin \mathit{Int}\,(\mathcal{A}^*\mathit{Dom}\,(\mathcal{V}^*) + \mathit{Dom}\,(\mathcal{U}^*))$. Because Condition (A.10) fails to hold, we cannot apply Theorem A.1 in the present case of power utility-semivariance optimization. Hence, Condition (A.10) is apparently too demanding that cannot be applied in our present utility-risk problem.

A.3. Lagrangian Multiplier Approach. The convex analysis with the use of conjugate functions can be alternatively integrated into Lagrangian multiplier approach in portfolio optimization; see Bielecki et al. [4] and Jin et al. [14]. Under the Lagrangian multiplier approach, an equivalent unconstrained optimization problem can be obtained through eliminating the budget constraint. In our framework, the unconstrained problem becomes:

$$\sup_{X\in\mathcal{X}} \tilde{\Psi}_\lambda(X) \tag{A.11}$$

where $\tilde{\Psi}_\lambda(X) := \Psi(X) - \lambda\mathbb{E}[\xi X]$ and Ψ was defined in (2.3). There is a well-known result that links the optimality of the original constrained optimization problem 2.4 with that of the unconstrained Lagrangian problem (A.11):

THEOREM A.2 (Proposition 4.1 in [4]). *Suppose that Problem 2.4 has a solution X^*, then there exists a $\lambda^* \in \mathbb{R}$ such that X^* also solve problem (A.11) when $\lambda = \lambda^*$. Conversely, if X^* solves Problem (A.11) for some λ and X^* satisfies $x_0 = \mathbb{E}[\xi X^*]$, then it must also solves for Problem 2.4.*

Deriving the optimality condition for the Lagrangian formulation (A.11), we can obtain exactly the same Nonlinear Moment Problem in Theorem 3.2. Indeed, assume that Condition 3.1 holds. Suppose $\hat{X} \in \mathcal{X}$ is optimal to Problem 2.4, by Theorem A.2, there exists λ^* such that $\hat{X}$ is optimal to the unconstrained problem $\tilde{\Psi}_\lambda$. The first order optimality condition for $\tilde{\Psi}_\lambda$ gives a similar necessary optimality condition as Proposition 3.3, i.e. it is necessary that

$$\mathbb{E}\left[\tilde{X}\left(\hat{\Gamma} - \lambda\xi\right)\right] \leq 0 \text{ for any } \tilde{X} \in \mathcal{L}^\infty \text{ such that } \hat{X}_\lambda + \tilde{X} \in \mathcal{X}, \tag{A.12}$$

where $\hat{\Gamma} := U'(\hat{X}) - \gamma\mathbb{E}\left[D'\left(\mathbb{E}[\hat{X}] - \hat{X}\right)\right] + \gamma D'\left(\mathbb{E}[\hat{X}] - \hat{X}\right)$. Since $\tilde{X}$ is arbitrary, the following algebraic structure is expected to be satisfied by the optimal solution $\hat{X}$:

$$\begin{cases}\hat{\Gamma} = \lambda\xi & \text{a.s. on } \left\{\hat{X} > K\right\}, \\ \hat{\Gamma} \leq \lambda\xi & \text{a.s. on } \left\{\hat{X} = K\right\}.\end{cases} \tag{A.13}$$

Meanwhile, the Lagrangian multiplier, λ, is chosen such that the budget constraint, $\mathbb{E}[\hat{X}\xi] = x_0$, is satisfied. Then, we can obtain the same necessary result as Theorem 3.2: it is necessary that $\hat{X}$ satisfies the same Nonlinear Moment Problem. To obtain (A.13) from (A.12), a natural method is to pick a suitable perturbation $\tilde{X}$ so that the validity of (A.13) is ensured, and one may consider $\tilde{X} = \hat{\Gamma} - \lambda\xi$ so that (A.12) becomes $\mathbb{E}\left[\left(\hat{\Gamma} - \lambda\xi\right)^2\right] \leq 0$. However, whether both $\hat{\Gamma} - \lambda\xi \in \mathcal{L}^\infty$ and $\hat{X} + \hat{\Gamma} - \lambda\xi \in \mathcal{X}$ are satisfied or not is not apparent. Using the similar argument in Lemma 3.5, if we can construct a random variable $Z \in \mathcal{L}^\infty$ such that the following three items hold:

$$\begin{cases} Z > 0 & \text{a.s. on } \left\{\hat{X} > K\right\}, \\ Z = 0 & \text{a.s. on } \left\{\hat{X} = K\right\}, \end{cases}$$

$$Z\left(\hat{\Gamma} - \lambda\xi\right) \in \mathcal{L}^\infty, \text{ and}$$

$$\hat{X} + Z\left(\hat{\Gamma} - \lambda\xi\right) \in \mathcal{X},$$

we can then obtain (A.13) from (A.12). In particular, the interior case in (A.13) is obtained by setting the perturbation $\tilde{X} = Z\left(\hat{\Gamma} - \lambda\xi\right)$ while the boundary case in (A.13) is obtained by setting the perturbation $\tilde{X} = \mathbb{I}\{\hat{X} = K\}$. Therefore, the remaining claim is to construct the random variable Z which satisfies the aforementioned three items. To achieve this, we can construct Z as the following, similar to (3.13):

$$Z := \begin{cases} 0, & \text{if } \hat{X} = K, \\ 1, & \text{if } \hat{X} > K \text{ and } \hat{\Gamma} = \lambda\xi, \\ \min\left\{\frac{\min\left\{\delta^U(\hat{X}), \frac{1}{2}(\hat{X}-K), \frac{\delta}{2}\right\}}{|\hat{\Gamma}-\lambda\xi|}, 1\right\}, & \text{otherwise,} \end{cases}$$

and δ^U was defined in Lemma 3.6. Finally, using the similar argument in Proposition 3.9, we can show that $Z\left(\hat{\Gamma} - \lambda\xi\right) \in \mathcal{L}^\infty$ and $\hat{X} + Z\left(\hat{\Gamma} - \lambda\xi\right) \in \mathcal{X}$, then the establishment of necessary result via Lagrangian multiplier approach accomplishes. Hence, we can find that the similar argument in Lemma 3.5 and Proposition 3.9 could be used for deriving (A.13). Therefore, the same technical derivation in Proof of Theorem 3.2 could be reused to derive the Nonlinear Moment Problem via Lagrangian multiplier approach.

On the other hand, with Theorem A.2, we can solve the constrained optimization problem 2.4 through the following steps:

1. For each λ, find the maximizer $\hat{X}_\lambda$ for $\tilde{\Psi}_\lambda$;
2. Find a suitable λ^* such that the budget constraint is satisfied, i.e. $\mathbb{E}[\hat{X}_{\lambda^*}\xi] = x_0$;
3. Conclude that the optimal solution for Problem 2.4 is $\hat{X}_{\lambda^*}$.

Applying Lagrangian multiplier approach to the present utility-risk problem, the sufficient condition of optimality coincides with that in Theorem 4.2: Given a solution of NMP $(\hat{X}^*, Y^*, M^*, R^*)$, the validations of (3.1), (3.3), and (3.4) means that $\hat{X}^*$ is the maximizer of $\tilde{\Psi}_{Y^*}$; while the validation of (3.2) warrants that $\hat{X}^*$ satisfies the budget constraint. Thus, $\hat{X}^*$ is optimal to Problem 2.4.

In particular, under the simple case of sole utility maximization, the maximizer of $\tilde{\Psi}_\lambda$ is given by $\hat{X} = (U')^{-1}(\lambda\xi)$, while the Lagrangian multiplier λ is determined by the budget constraint. We only have to solve an equation with an (real) unknown in this special case. Under the general case with the presence of deviation risk measure, since the optimal terminal wealth of the unconstrained problem (A.11) has no explicit form, the determination of the existence of the Lagrangian multiplier is no longer immediate. In order to solve for our general utility-risk problem, we have to tackle with the Nonlinear Moment Problem which can be converted into a problem of system of three equations in three unknowns. Propositions 4.6 and 4.13 have been established to solve for this system of equations, and to the best of our knowledge, the resolution of the system, and so our original utility-risk portfolio problem, are highly non-trivial and did not appear in the existing literature before our work.

REMARK A.3. *Using the primal approach, the $\mathcal{L}^2$-uniform integrability of $\{X_k\}$ in (A.2) is essential for a successful application of the Komlos theorem and guaranteeing that the subsequential limit Z^* is $\mathcal{L}^2$-uniformly integrable. If the objective function admits coercivity in the sense that, for any sequence $\{X_n\}$ in $\{X \in \mathcal{X} \,|\, \mathbb{E}[\xi X] \leq x_0\}$ such that $\|X_n\|_2 \to \infty$,*

$$\lim_{n\to\infty} \Psi(X_n) = -\infty,$$

the $\mathcal{L}^2$-norm of maximizing sequence in Section A.1 has to be bounded, thus the integrability in (A.2) can be ensured[9]. However, under the budget constraint $\mathbb{E}[\xi X] = x_0$, in general, such coercivity for the constrained problem is not immediate to check without the consideration of the conjugate function. Usually, we have to verify the coercivity case by case depending on a specific form of objective function. Nevertheless, even the coercivity can be satisfied, the uniform integrability of $\{U(Z_k)\}$ is a substantial issue in the primal approach.

A usual method is to convert a constrained problem into an unconstrained problem via Lagrangian multiplier method. Hence, we use Theorem A.2 to eliminate the budget constraint, and we turn to consider the coercivity of $\tilde{\Psi}_\lambda$ in the sense that, for any sequence $\{X_n\}$ in $\mathcal{X}$ such that $\|X_n\|_2 \to \infty$,

$$\lim_{n\to\infty} \tilde{\Psi}_\lambda(X_n) = -\infty,$$

which is equivalent to the boundedness of $\{\tilde{\Psi}_\lambda(X) \geq k\}$ under $\mathcal{L}^2$-norm for every k: [10]

We now study the case of power-utility-semivariance optimization by setting

$$\tilde{\Psi}_\lambda(X) := \frac{1}{1-\rho}\mathbb{E}[U(X)] - \gamma\mathbb{E}\left[(\mathbb{E}[X] - X)_+\right] - \lambda\mathbb{E}[\xi X],$$

where $U(x) = \begin{cases} \frac{1}{1-\rho}x^{1-\rho}, & x \geq 0 \\ -\infty, & x < 0 \end{cases}$ for some $\rho \in (0,1)$. Considering the sequence $X_N := N\mathbb{I}\{\xi < b_N\}$, where b_N is the smallest possible such that $p_N := \mathbb{P}\{\xi < b_N\} =$

[9] It is clear that $\Psi(0) > -\infty$. For if the maximizing sequence $\{X_n\}$ in Section A.1 does not satisfy (A.2), i.e. $\sup_n \mathbb{E}[|X_n|^2] = \infty$, then, by the coercivity of Ψ, we have $\lim_{n\to\infty} \Psi(X_n) = -\infty$, which contradicts to the maximizing nature of $\{X_n\}$.

[10] If the boundedness of $\{\tilde{\Psi}_\lambda(X) \geq k\}$ does not hold for some k, there exists a sequence $\{X_n\}$ in $\mathcal{X}$ such that $\lim_{n\to\infty} \|X_n\|_2 = \infty$ and $\tilde{\Psi}_\lambda(X_n) \geq k$ for all n, which violates the coercivity of $\tilde{\Psi}_\lambda$. Conversely, if the coercivity of $\tilde{\Psi}_\lambda$ does not hold, there exists a sequence $\{X_n\}$ in $\mathcal{X}$ such that $\lim_{n\to\infty} \|X_n\|_2 = \infty$ and $\lim_{n\to\infty} \tilde{\Psi}_\lambda(X_n) = M > -\infty$. Then, we can find some k such that $\tilde{\Psi}_\lambda(X_n) \geq k$ for all n, which violates the boundedness of $\{\tilde{\Psi}_\lambda(X) \geq k\}$.

$\frac{1}{N}$, *by using similar calculation as in Section A.2, we have:*

$$\begin{aligned}\tilde{\Psi}_\lambda(X_N) &\geq \frac{1}{1-\rho}N^{1-\rho}p_N - \gamma N^2 p_N^2(1-p_N) - \lambda N\mathbb{E}\left[\xi\mathbb{I}\{\xi < b_N\}\right] \\ &\geq \frac{1}{1-\rho}N^{-\rho} - \gamma - \lambda b_N,\end{aligned}$$

which is bounded from below. Meanwhile, $\{X_n\}$ *is not bounded in* $\mathcal{L}^2$*-norm; indeed, for any* $\delta > 0$,

$$\mathbb{E}\left[X_N^2\right] \geq N^2 p_N = N \to \infty \text{ as } N \to \infty$$

which means that $\tilde{\Psi}_\lambda$ *does not admit the usual coercivity in* $\mathcal{L}^2$*-norm topology for every* $\lambda \in \mathbb{R}$.

Appendix B. Technical proofs in Section 3.

Proof of Lemma 3.4. Since U is concave and D is convex function, so $f(\theta) := U(\hat{X}+\theta\tilde{X}) - \gamma D(\mathbb{E}[\hat{X}+\theta\tilde{X}] - (\hat{X}+\theta\tilde{X}))$ is concave in $\theta > 0$. Thus, for any $\delta > 0$, by concavity of f, $f(\theta) \geq \frac{\delta}{\theta+\delta}f(0) + \frac{\theta}{\theta+\delta}f(\theta+\delta)$, so

$$\frac{f(\theta)-f(0)}{\theta} \geq \frac{f(\theta+\delta)-f(0)}{\theta+\delta}.$$

Hence, $\frac{1}{\theta}\left(U(\hat{X}+\theta\tilde{X}) - \gamma D\left(\mathbb{E}\left[\hat{X}+\theta\tilde{X}\right] - (\hat{X}+\theta\tilde{X})\right) - \left(U(\hat{X}) - \gamma D\left(\mathbb{E}\left[\hat{X}\right] - \hat{X}\right)\right)\right)$ is increasing as θ decreases to 0. Since $\hat{X}$ and $\hat{X}+\tilde{X}$ are admissible terminal wealth, thus $U\left(\hat{X}+\tilde{X}\right) - \gamma D\left(\mathbb{E}\left[\hat{X}+\tilde{X}\right] - (\hat{X}+\tilde{X})\right)$ and $U(\hat{X}) - \gamma D\left(\mathbb{E}\left[\hat{X}\right] - \hat{X}\right)$ are both $\mathcal{L}^1$-integrable because $\hat{X}+\tilde{X}, \hat{X} \in \mathcal{X}$. Hence, this lemma follows from the Monotone Convergence Theorem.
□

Proof of Lemma 3.6

By the Mean Value Theorem, for any $x, x_0 \in \mathcal{D}$,

$$|U(x) - U(x_0)| \leq |x-x_0||U'(\theta)|, \text{ for some } \theta \in (x_0 - |x-x_0|, x_0 + |x-x_0|).$$

As U' is positive and decreasing, $|U'(\theta)| \leq |U'(x_0 - |x-x_0|)|$ for $\theta \in (x_0 - |x-x_0|, x_0 + |x-x_0|)$. Finally, our result follows by choosing

$$\delta^U(x) := \min\left\{\frac{1}{U^*(x)}, \frac{x-K}{2}, 1\right\}, \text{ where } U^*(x) := \begin{cases} U'(x-1), & \text{if } x-2 \in \mathcal{D}, \\ U'(\frac{x+K}{2}), & \text{if } x-2 \notin \mathcal{D}.\end{cases}$$

□

Before we prove Lemma 3.8, we require to show the claim that $\hat{\Gamma} > 0$:

Lemma B.1. *Given* $\hat{X}$ *is optimal for Problem 2.4 satisfying Conditions 3.1 (i) and (ii), it is necessary that* $\hat{\Gamma} > 0$ *almost surely.*

Proof. We consider two cases: (i) $K > -\infty$ and (ii) $K = -\infty$, respectively.

We consider case (i) $K > -\infty$. With the optimal solution $\hat{X}$ of Problem 2.4 , we define $h(x) := \Gamma\left(\hat{X}, x\right)$, which is a decreasing continuous function. Since $U' > 0$ on $\{\hat{X} > K\}$ and U' is decreasing because of concavity of U, so we have $\lim_{x\downarrow K} U'(x) > 0$. Hence,

$$\lim_{x\downarrow K} h(x) > -\gamma\mathbb{E}\left[D'\left(\mathbb{E}\left[\hat{X}\right] - \hat{X}\right)\right] + \gamma D'\left(\mathbb{E}\left[\hat{X}\right] - K\right).$$

By definition, $\hat{X} \geq K$ uniformly, and D' is increasing, thus $D'\left(\mathbb{E}\left[\hat{X}\right] - K\right) \geq \mathbb{E}\left[D'\left(\mathbb{E}\left[\hat{X}\right] - \hat{X}\right)\right]$. So $\lim_{x \to K} h(x) > 0$. Then, there exists $k_0 := \inf\{x > K \mid h(x) \leq 0\} \in (K, \infty]$, i.e., $\hat{\Gamma} \leq 0$ implies $\hat{X} \geq k_0$. If $k_0 = \infty$, it is immediate that $\hat{\Gamma} > 0$ almost surely, so we consider the case that $k_0 < \infty$. Assume the contrary, that $\mathbb{P}\left[\hat{\Gamma} \leq 0\right] > 0$. Consider

$$\tilde{X} := \begin{cases} -\frac{k_0 - K}{2}, & \text{if } \hat{\Gamma} \leq 0, \\ \frac{k_0 - K}{2} \frac{\mathbb{E}\left[\mathbb{I}\left\{\hat{\Gamma} \leq 0\right\}\xi\right]}{\mathbb{E}\left[\mathbb{I}\left\{\hat{\Gamma} > 0\right\}\xi\right]}, & \text{if } \hat{\Gamma} > 0. \end{cases} \tag{B.1}$$

We have $\hat{X} + \tilde{X} > K$ and $\tilde{X} \in \mathcal{L}^\infty$. Since U is concave,

$$U(\hat{X}) + \tilde{X}U'(\hat{X} + \tilde{X}) \leq U(\hat{X} + \tilde{X}) \leq U(\hat{X}) + \tilde{X}U'(\hat{X}).$$

Furthermore, when $\hat{\Gamma} \leq 0$, we have $\tilde{X} < 0$ and $\hat{X} \geq k_0$, then

$$\tilde{X}U'\left(\hat{X} + \tilde{X}\right) \geq \tilde{X}U'\left(k_0 - \frac{k_0 - K}{2}\right);$$

while $\hat{\Gamma} > 0$, we have $\tilde{X} > 0$ and $\hat{X}(T) \leq k_0$, then

$$\tilde{X}U'\left(\hat{X} + \tilde{X}\right) \geq \tilde{X}U'\left(k_0 + \frac{k_0 - K}{2} \frac{\mathbb{E}\left[\mathbb{I}\left\{\hat{\Gamma} \leq 0\right\}\xi\right]}{\mathbb{E}\left[\mathbb{I}\left\{\hat{\Gamma} > 0\right\}\xi\right]}\right).$$

Thus, $U(\hat{X} + \tilde{X}) \in \mathcal{L}^1$. On the other hand, since D is convex,

$$\begin{aligned} 0 &\leq D\left(\mathbb{E}\left[\hat{X} + \tilde{X}\right] - \left(\hat{X} + \tilde{X}\right)\right) \\ &\leq D\left(\mathbb{E}\left[\hat{X}\right] - \hat{X}\right) + \left(\mathbb{E}\left[\tilde{X}\right] - \tilde{X}\right) D'\left(\mathbb{E}\left[\hat{X} + \tilde{X}\right] - \left(\hat{X} + \tilde{X}\right)\right). \end{aligned}$$

Similar to showing $\tilde{X}U'(\hat{X} + \tilde{X})$ being bounded from below, we can show that $\left(\mathbb{E}\left[\tilde{X}\right] - \tilde{X}\right) D'\left(\mathbb{E}\left[\hat{X} + \tilde{X}\right] - \left(\hat{X} + \tilde{X}\right)\right)$ is bounded from above, thus $D\left(\mathbb{E}\left[\hat{X} + \tilde{X}\right] - \left(\hat{X} + \tilde{X}\right)\right) \in \mathcal{L}^1$. Hence, $\hat{X} + \tilde{X} \in \mathcal{X}$. Also, $\mathbb{E}\left[\tilde{X}\xi\right] = 0$, but $\mathbb{E}\left[\tilde{X}\hat{\Gamma}\right] > 0$, which altogether violates Proposition 3.3. So $\mathbb{P}\left[\hat{\Gamma} \leq 0\right] = 0$.

Now, consider the case (ii) $\mathcal{D} = \mathbb{R}$; the approach is similar as in the case (i). Firstly, there exists $k_0 := \inf\{x > -\infty \mid h(x) \leq 0\} \in (-\infty, \infty]$. If $k_0 < \infty$, we assume the contrary, that $\mathbb{P}\left[\hat{\Gamma} \leq 0\right] > 0$, and then as in case (i), we can show that Proposition 3.3 is violated by setting $\tilde{X}$ as follows:

$$\tilde{X} := \begin{cases} -1, & \text{if } \hat{\Gamma} \leq 0, \\ \frac{\mathbb{E}\left[\mathbb{I}\left\{\hat{\Gamma} \leq 0\right\}\xi\right]}{\mathbb{E}\left[\mathbb{I}\left\{\hat{\Gamma} > 0\right\}\xi\right]}, & \text{if } \hat{\Gamma} > 0. \end{cases}$$

□

Proof of Lemma 3.8. For any $y \in (0, \infty)$, by (3.13), we have

$$\left|y\xi^2 Z_y - \xi\hat{\Gamma}Z_y\right| = \left|y\xi - \hat{\Gamma}\right| \cdot |Z_y|\xi \leq \delta\xi. \tag{B.2}$$

By the Dominated Convergence Theorem, f is continuous on $(0,\infty)$. Since $\hat{\Gamma} > 0$ almost surely by Lemma B.1,

$$\lim_{y\downarrow 0} Z_y = \min\left\{\frac{\min\left\{\delta^U\left(\hat{X}\right), \frac{1}{2}\left(\hat{X}-K\right), \frac{\delta}{2}\right\}}{\hat{\Gamma}}, 1\right\} > 0$$

almost surely on $\left\{\hat{X} > K\right\}$. Since $\mathbb{P}\left[\hat{X} > K\right] > 0$, we have, by the Dominated Convergence Theorem,

$$\lim_{y\to 0} f(y) = \mathbb{E}\left[\lim_{y\to 0}\left(y\xi^2 Z_y - \xi\hat{\Gamma} Z_y\right)\right] = -\mathbb{E}\left[\xi\hat{\Gamma}\lim_{y\downarrow 0} Z_y\right] < 0.$$

Note that

$$\lim_{y\to\infty} y\xi^2 Z_y = \xi\min\left\{\delta^U\left(\hat{X}\right), \frac{1}{2}\left(\hat{X}-K\right), \frac{\delta}{2}\right\} > 0 \text{ a.s. on } \left\{\hat{X} > K\right\}$$

and

$$\lim_{y\to\infty} \xi\hat{\Gamma} Z_y = \xi\hat{\Gamma}\lim_{y\to\infty}\min\left\{\frac{\min\left\{\delta^U\left(\hat{X}\right), \frac{1}{2}\left(\hat{X}-K\right), \frac{\delta}{2}\right\}}{|\hat{\Gamma} - y\xi|}, 1\right\} = 0.$$

By applying the Dominated Convergence Theorem and the fact that $\mathbb{P}\left[\hat{X} > K\right] > 0$ under Assumption 2.3, we have:

$$\lim_{y\to\infty} f(y) = \mathbb{E}\left[\lim_{y\to\infty}\left(y\xi^2 Z_y - \xi\hat{\Gamma} Z_y\right)\right] = \mathbb{E}\left[\lim_{y\to\infty} y\xi^2 Z_y\right] > 0,$$

Our claim follows by intermediate value theorem. □

Appendix C. Technical proofs in Section 4.

C.1. Proof of Proposition 4.5. Fix $(m,y) \in \mathbb{R}\times(0,\infty)$. Since U is strictly concave and D is convex, $U'(z) + \gamma D'(m-z) - y$ is strictly decreasing in z. Since U' and D' is continuous, $U'(z) + \gamma D'(m-z) - y$ is continuous in z. Under Assumptions (4.7), we can also easily show that $U'(z) + \gamma D'(m-z) - y$ is coercive in the sense that

$$\lim_{z\to 0} U'(z) + \gamma D'(m-z) - y = \infty, \qquad \lim_{z\to\infty} U'(z) + \gamma D'(m-z) - y = -y < 0.$$

Thus, by the intermediate value theorem and strict monotonicity, for any $(m,y) \in \mathbb{R}\times(0,\infty)$, there exists a unique $I(m,y) \in (0,\infty)$ such that

$$U'(I(m,y)) + \gamma D'(m - I(m,y)) - y = 0.$$

(a) **(i)** For fixed (m,y), $U'(z) + \gamma D'(m-z) - y$ is strictly decreasing in z. When (z,m) is fixed, $U'(z) + \gamma D'(m-z) - y$ is strictly decreasing in y, so $I(m,y)$ is strictly decreasing in y.

(ii) We first claim that $m \geq I(m,y)$ when $y \geq U'(m)$. Assume the contrary, that $m < I(m,y)$. We have $D'(m - I(m,y)) = 0$, and then $y = U'(I(m,y)) \geq U'(m)$, which contradicts to $m < I(m,y)$ as U' is decreasing. Next, we assume another contrary, that there exists m_0, y_0 with $y_0 \leq U'(m_0)$ and $\delta > 0$ such that $I(m_0+\delta, y_0) \leq I(m_0, y_0)$. Then we have $m_0 + \delta - I(m_0+\delta, y_0) > m_0 - I(m_0, y_0) \geq 0$, thus

$$U'(I(m_0+\delta, y_0)) + \gamma D'(m_0+\delta-I(m_0+\delta, y_0)) > U'(I(m_0, y_0)) + \gamma D'(m_0 - I(m_0, y_0)),$$

which contradicts (4.8).

For the second assertion, $y \leq U'(m)$ implies that $m \leq (U')^{-1}(y)$, thus $I(m,y) = (U')^{-1}(y)$ satisfies (4.8), and it is the unique solution by the main result in this proposition.

(b) Fix $(M_0, Y_0) \in \mathbb{R} \times (0,\infty)$. By part (a), for any small enough $\epsilon > 0$,

$$I(M_0 - \epsilon, Y_0 + \epsilon) \leq I(m,y) \leq I(M_0 + \epsilon, Y_0 - \epsilon) \tag{C.1}$$

for any $|(m,y) - (M_0, Y_0)| < \epsilon$.

It is straightforward to show that $\lim_{\epsilon\downarrow 0} I(M_0+\epsilon, Y_0-\epsilon)$ and $I(M_0, Y_0)$ satisfy the same equation in (4.8), so we have $\lim_{\epsilon\downarrow 0} I(M_0+\epsilon, Y_0-\epsilon) = I(M_0, Y_0)$. Similarly, we have $\lim_{\epsilon\downarrow 0} I(M_0-\epsilon, Y_0+\epsilon) = I(M_0, Y_0)$. Applying the sandwich theorem to (C.1), we can conclude that

$$\lim_{(m,y)\to(M_0,Y_0)} I(m,y) = I(M_0, Y_0).$$

C.2. Proof of Lemma 4.7. In this lemma, we prove the followings in order:

(a) $\mathbb{E}[D'\left(M - I\left(M, \gamma R + Y\xi\right)\right) - R]$ is strictly decreasing in R,

(b) $\mathbb{E}[D'\left(M - I\left(M, \gamma R + Y\xi\right)\right)]$ is continuous in R,

(c1) $\lim_{R\to 0} \mathbb{E}[D'\left(M - I\left(M, \gamma R + Y\xi\right)\right) - R] > 0$,

(c2) $\lim_{R\to D'(M)} \mathbb{E}\left[D'\left(M - I\left(M, \gamma R + Y\xi\right)\right) - R\right] < 0$.

In light of (b), (c1) and (c2), then by the intermediate value theorem, there exists $R = R_{Y,M}$ satisfying (4.12) while the uniqueness of $R_{Y,M}$ is guaranteed by (a). Finally, we show that

(d1) $R_{Y,M}$ is strictly increasing in M for fixed Y,

(d2) $R_{Y,M}$ is strictly increasing in Y for fixed M.

For each of the above items:

(a) By Proposition 4.5 (a)(i), $U'\left(I\left(M, \gamma R + Y\xi\right)\right)$ is strictly increasing in R almost surely. By (4.8), $D'\left(M - I\left(M, \gamma R + Y\xi\right)\right) - R = \frac{1}{\gamma}\left(Y\xi - U'\left(I\left(M, \gamma R + Y\xi\right)\right)\right)$ is therefore strictly decreasing in R almost surely. Thus, $\mathbb{E}\left[D'\left(M - I\left(M, \gamma R + Y\xi\right)\right) - R\right]$ is strictly decreasing in R.

(b) Since D' and I are both continuous, so $D'\left(M - I\left(M, \gamma R + Y\xi\right)\right)$ is continuous in R. Hence, the claim follows by an application of the Dominated Convergence Theorem.

(c1) When $Y\xi > U'(M)$, we have $M > I\left(M, Y\xi\right)$, thus

$$D'\left(M - I\left(M, Y\xi\right)\right) > 0, \text{ a.s. on } \left\{Y\xi > U'(M)\right\}.$$

Since $\frac{U'(M)}{Y} \in (0,\infty)$, by the definition of ξ, we have

$$\mathbb{P}[D'(M - I(M, Y\xi)) > 0] \geq \mathbb{P}\left[\xi > \frac{U'(M)}{Y}\right] > 0.$$

By the Dominated Convergence Theorem,

$$\begin{aligned}\lim_{R\to 0}\mathbb{E}\left[D'\left(M-I\left(M,\gamma R+Y\xi\right)\right)-R\right]&=\mathbb{E}\left[\lim_{R\to 0}D'\left(M-I\left(M,\gamma R+Y\xi\right)\right)\right]\\&=\mathbb{E}\left[D'\left(M-I\left(M,Y\xi\right)\right)\right]>0.\end{aligned}$$

(c2) By (4.8), we have $\lim_{R\to D'(M)} I\left(M,\gamma R+Y\xi\right)>0$ almost surely, so

$$\lim_{R\to D'(M)} D'(M-I\left(M,\gamma R+Y\xi\right))<D'(M),\ \text{a.s. }.$$

By the Monotone Convergence Theorem, $\lim_{R\to D'(M)}\mathbb{E}\left[D'(M-I\left(M,\gamma R+Y\xi\right))\right]<D'(M)$, thus this part follows.

(d1) Assume the contrary, that there exists a $M_0\in(0,\infty)$ and a $\delta>0$ such that $R_{Y,M_0}\geq R_{Y,M_0+\delta}$.
By Proposition 4.5 (a)(ii), when $\gamma R+Y\xi\geq U'(M)$, $I\left(M,\gamma R+Y\xi\right)$ is strictly increasing in M, thus $D'(M-I\left(M,\gamma R+Y\xi\right))-R$ is strictly increasing in M on $\{\gamma R+Y\xi\geq U'(M)\}$ by (4.8). When $\gamma R+Y\xi\leq U'(M)$, $D'(M-I\left(M,\gamma R+Y\xi\right))-R=-R$. So given Y and R, $D'(M-I\left(M,\gamma R+Y\xi\right))$ is increasing in M and is strictly increasing on $\{\gamma R+Y\xi\geq U'(M)\}$. Thus we have $\mathbb{E}\left[D'(M-I\left(M,\gamma R+Y\xi\right))-R\right]$ is strictly increasing in M.
On the other hand, in (a), $D'(M-I\left(M,\gamma R+Y\xi\right))-R$ is strictly decreasing in R almost surely. Now,

$$\begin{aligned}0&=\mathbb{E}\left[D'\left(M_0+\delta-I\left(M_0+\delta,\gamma R_{Y,M_0+\delta}+Y\xi\right)\right)-R_{Y,M_0+\delta}\right]\\&\geq\mathbb{E}\left[D'\left(M_0+\delta-I\left(M_0+\delta,\gamma R_{Y,M_0}+Y\xi\right)\right)-R_{Y,M_0}\right]\\&>\mathbb{E}\left[D'\left(M_0-I\left(M_0,\gamma R_{Y,M_0}+Y\xi\right)\right)-R_{Y,M_0}\right]=0,\end{aligned}$$

a contradiction.

(d2) In Proposition 4.5 (a)(i), we have $I\left(M,\gamma R+Y\xi\right)$ strictly decreasing in Y given a fixed point of M and R, thus $\mathbb{E}\left[D'(M-I\left(M,\gamma R+Y\xi\right))-R\right]$ is also strictly increasing in Y almost surely given (M,R). Thus, this part can be verified as in (d1).

C.3. Proof of Lemma 4.8. In this lemma, we prove the following in order:

(a) $\mathbb{E}[I\left(M,\gamma R_{Y,M}+Y\xi\right)]-M$ is strictly decreasing in M,

(b) $\mathbb{E}[I\left(M,\gamma R_{Y,M}+Y\xi\right)]$ is continuous in M,

(c1) $\lim_{M\to 0}\mathbb{E}[I\left(M,\gamma R_{Y,M}+Y\xi\right)-M]>0$,

(c2) $\lim_{M\to\infty}\mathbb{E}\left[M-I\left(M,\gamma R_{Y,M}+Y\xi\right)\right]=\infty$.

By using (b), (c1) and (c2), in accordance with the intermediate value theorem and part (a), there exists a unique $M=M_Y$ satisfying (4.14). Finally, we show that

(d) M_Y is strictly decreasing in Y.

For each of the above items:

(a) By Proposition 4.5 (a)(ii) , $D'\left(M-I\left(M,\gamma R+Y\xi\right)\right)-R=\frac{1}{\gamma}\left(Y\xi-U'\left(I\left(M,\gamma R+Y\xi\right)\right)\right)$ is strictly increasing in M when $\gamma R+Y\xi\geq U'(M)$. Thus, because D' is strictly increasing for positive $M-I\left(M,\gamma R+Y\xi\right)$, $I\left(M,\gamma R+Y\xi\right)-M$ is strictly decreasing in M on $\{\gamma R+Y\xi\geq U'(M)\}$. On the other hand, on $\{\gamma R+Y\xi\leq U'(M)\}$, $I\left(M,\gamma R+Y\xi\right)-M=(U')^{-1}(\gamma R+Y\xi)-M$ is strictly decreasing in M.
By Proposition 4.5 (a)(i), $I\left(M,\gamma R+Y\xi\right)-M$ is strictly decreasing in R. By Lemma 4.7, $R_{Y,M}$ is strictly increasing in M. Thus, for any $\delta>0$, it is

almost surely that:

$$\begin{aligned} I\left(M,\gamma R_{Y,M}+Y\xi\right)-M &> I\left(M,\gamma R_{Y,M+\delta}+Y\xi\right)-M \\ &> I\left(M+\delta,\gamma R_{Y,M+\delta}+Y\xi\right)-(M+\delta) \end{aligned}$$

Thus, $\mathbb{E}\left[I\left(M,\gamma R_{Y,M}+Y\xi\right)\right]-M$ is strictly decreasing in M.

(b) Fix $M_0\in(0,\infty)$. By the continuity of D' and I, it is almost surely that:

$$\begin{aligned} &\lim_{M\downarrow M_0}\left(D'\left(M-I\left(M,\gamma R_{Y,M}+Y\xi\right)\right)-R_{Y,M}\right) \\ &= D'\left(M_0-I\left(M_0,\gamma\lim_{M\downarrow M_0}R_{Y,M}+Y\xi\right)\right)-\lim_{M\downarrow M_0}R_{Y,M} \end{aligned} \tag{C.2}$$

By (4.12), (C.2) and the Dominated Convergence Theorem, we have

$$\begin{aligned} &\mathbb{E}\left[D'\left(M_0-I\left(M_0,\gamma\lim_{M\downarrow M_0}R_{Y,M}+Y\xi\right)\right)\right]-\lim_{M\downarrow M_0}R_{Y,M} \\ =&\lim_{M\downarrow M_0}\left(\mathbb{E}\left[D'\left(M-I\left(M,\gamma R_{Y,M}+Y\xi\right)\right)\right]-R_{Y,M}\right)=0. \end{aligned}$$

By the uniqueness result in Lemma 4.7, we conclude that $\lim_{M\downarrow M_0}R_{Y,M}=R_{Y,M_0}$. Similarly, we have $\lim_{M\uparrow M_0}R_{Y,M}=R_{Y,M_0}$. By continuity of I,

$$\lim_{M\downarrow M_0}I\left(M,\gamma R_{Y,M}+Y\xi\right)=I\left(M_0,\gamma R_{Y,M_0}+Y\xi\right).$$

Similarly, the equality of limits from the opposite side can also be deduced, so

$$\lim_{M\to M_0}I\left(M,\gamma R_{Y,M}+Y\xi\right)=I\left(M_0,\gamma R_{Y,M_0}+Y\xi\right).$$

Finally, our claim follows by the Dominated Convergence Theorem.

(c1) Since $I\left(M,\gamma R_{Y,M}+Y\xi\right)-M$ is decreasing in M by (a), thus

$$\lim_{M\to 0}I\left(M,\gamma R_{Y,M}+Y\xi\right)=\lim_{M\to 0}\left(I\left(M,\gamma R_{Y,M}+Y\xi\right)-M\right).$$

We claim that $\lim_{M\to 0}I\left(M,\gamma R_{Y,M}+Y\xi\right)>0$ almost surely. Assume the contrary, that there exists a sample value of ξ_0 such that $\lim_{M\to 0}I\left(M,\gamma R_{Y,M}+Y\xi_0\right)=0$. Then, we have $\lim_{M\to 0}D'\left(M-I\left(M,\gamma R_{Y,M}+Y\xi_0\right)\right)=D'(0)$ and $\lim_{M\to 0}U'\left(I\left(M,\gamma R_{Y,M}+Y\xi_0\right)\right)=\infty$. But by (4.8), we again have:

$$\begin{aligned} &\lim_{M\to 0}U'\left(I\left(M,\gamma R_{Y,M}+Y\xi_0\right)\right) \\ &=\xi_0 Y+\gamma\lim_{M\to 0}\left(R_{Y,M}-D'\left(M-I\left(M,\gamma R_{Y,M}+Y\xi_0\right)\right)\right) \\ &\le \xi_0 Y+\gamma\lim_{M\to 0}\left(D'(M)-D'\left(M-I\left(M,\gamma R_{Y,M}+Y\xi_0\right)\right)\right) \\ &\le \xi_0 Y+\gamma\left(D'(0)-D'(0)\right)<\infty, \end{aligned}$$

which leads to a contradiction. Hence, $\lim_{M\to 0}I\left(M,\gamma R_{Y,M}+Y\xi\right)>0$ almost surely. Since $I\left(M,\gamma R_{Y,M}+Y\xi\right)-M$ is decreasing in M, by the Monotone Convergence Theorem,

$$\lim_{M\to 0}\mathbb{E}\left[I\left(M,\gamma R_{Y,M}+Y\xi\right)-M\right]>0.$$

(c2) By (4.8), we either have $\lim_{M\to\infty} I\left(M, \gamma R_{Y,M} + Y\xi\right) = \infty$ almost surely or $\lim_{M\to\infty} R_{Y,M} = \infty$.
Assume that $\lim_{M\to\infty} R_{Y,M} < \infty$, thus $\lim_{M\to\infty} I\left(M, \gamma R_{Y,M} + Y\xi\right) = \infty$ almost surely, then by continuity of U',

$$\lim_{M\to\infty} U'\left(I\left(M, \gamma R_{Y,M} + Y\xi\right)\right) = 0 \text{ a.s.} \tag{C.3}$$

Then, by the Dominated Convergence Theorem and (C.3),

$$\lim_{M\to\infty} \mathbb{E}\left[U'(I\left(M, \gamma R_{Y,M} + Y\xi\right))\right] = \mathbb{E}\left[\lim_{M\to\infty} U'(I\left(M, \gamma R_{Y,M} + Y\xi\right))\right] = 0.$$

Hence, we have $\lim_{M\to\infty} \mathbb{E}[U'(I\left(M, \gamma R_{Y,M} + Y\xi\right)) - \xi Y] = -\mathbb{E}[\xi]Y < 0$, this contradicts (4.13) in Lemma 4.7, so we must have $\lim_{M\to\infty} R_{Y,M} = \infty$.
Given a sample ξ_0, assume the contrary, that $\lim_{M\to\infty} M - I\left(M, \gamma R_{Y,M} + Y\xi_0\right) < \infty$, then $\lim_{M\to\infty} D'\left(M - I\left(M, \gamma R_{Y,M} + Y\xi_0\right)\right) < \infty$. By (4.8) and $\lim_{M\to\infty} R_{Y,M} = \infty$, we have $\lim_{M\to\infty} U'\left(I\left(M, \gamma R_{Y,M} + Y\xi_0\right)\right) = \infty$, thus $\lim_{M\to\infty} I\left(M, \gamma R_{Y,M} + Y\xi_0\right) = 0$ and then it results in: $\lim_{M\to\infty}\left(M - I\left(M, \gamma R_{Y,M} + Y\xi_0\right)\right) = \infty$, a contradiction.
Now, we have $\lim_{M\to\infty} M - I\left(M, \gamma R_{Y,M} + Y\xi\right) = \infty$ almost surely. By the Monotone Convergence Theorem,

$$\lim_{M\to\infty} \mathbb{E}\left[M - I\left(M, \gamma R_{Y,M} + Y\xi\right)\right] = \infty > 0.$$

(d) In part (a), we have shown that $\mathbb{E}\left[I\left(M, \gamma R_{Y,M} + Y\xi\right)\right] - M$ is strictly decreasing in M for a fixed Y. On the other hand, in Proposition 4.5 (a)(i) and Lemma 4.7 (d1), we can show that $\mathbb{E}\left[I\left(M, \gamma R_{Y,M} + Y\xi\right)\right] - M$ is strictly decreasing in Y for a fixed M. By (4.14), $\mathbb{E}\left[I\left(M_Y, \gamma R_{Y,M_Y} + Y\xi\right)\right] - M_Y = 0$ for all values of Y, thus, M_Y is strictly decreasing in Y.

C.4. Proof of Lemma 4.9. In this lemma, we prove the following in order:
(a) $\mathbb{E}\left[\xi I\left(M_Y, \gamma R_{Y,M_Y} + Y\xi\right)\right]$ is continuous in Y,
(b1) $\lim_{Y\to 0} \mathbb{E}\left[\xi I\left(M_Y, \gamma R_{Y,M_Y} + Y\xi\right)\right] = \infty$,
(b2) $\lim_{Y\to\infty} \mathbb{E}\left[\xi I\left(M_Y, \gamma R_{Y,M_Y} + Y\xi\right)\right] = 0$.
Then by the intermediate value theorem with (a), (b1) and (b2), there exists Y satisfying (4.15).
For each of the above items:
(a) Fix $Y_0 \in (0, \infty)$. For any $\epsilon > 0$, by Lemmas 4.7 and 4.8, we have

$$R_{Y_0-\epsilon, M_{Y_0+\epsilon}} < R_{Y,M_Y} < R_{Y_0+\epsilon, M_{Y_0-\epsilon}} \text{ for any } Y_0 - \epsilon < Y < Y_0 + \epsilon, \tag{C.4}$$

and $R_{Y_0+\epsilon, M_{Y_0-\epsilon}}$ is increasing in ϵ, we therefore have both the finite existence of $\lim_{\epsilon\downarrow 0} R_{Y_0+\epsilon, M_{Y_0-\epsilon}}$ and $\lim_{\epsilon\downarrow 0} R_{Y_0-\epsilon, M_{Y_0+\epsilon}}$. By Proposition 4.5 (b), I is jointly continuous,

$$\lim_{\epsilon\downarrow 0} I\left(M_{Y_0-\epsilon}, (Y_0+\epsilon)\xi + \gamma R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right) = I\left(\lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}, Y_0\xi + \gamma \lim_{\epsilon\downarrow 0} R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right).$$

Hence,

$$\begin{aligned}&\lim_{\epsilon\downarrow 0}\left(D'\left(M_{Y_0-\epsilon} - I\left(M_{Y_0-\epsilon}, (Y_0+\epsilon)\xi + \gamma R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right)\right) - R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right)\\ &= D'\left(\lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon} - I\left(\lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}, Y_0\xi + \gamma \lim_{\epsilon\downarrow 0} R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right)\right) - \lim_{\epsilon\downarrow 0} R_{Y_0+\epsilon, M_{Y_0-\epsilon}}.\end{aligned}$$

By standard application of the Dominated Convergence Theorem,

$$\begin{aligned}&\mathbb{E}\left[D'\left(\lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon} - I\left(\lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}, Y_0\xi + \gamma \lim_{\epsilon\downarrow 0} R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right)\right) - \lim_{\epsilon\downarrow 0} R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right]\\
&= \lim_{\epsilon\downarrow 0}\mathbb{E}\left[D'\left(M_{Y_0-\epsilon} - I\left(M_{Y_0-\epsilon}, (Y_0+\epsilon)\,\xi + \gamma R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right)\right) - R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right] = 0.\end{aligned}$$

Since R in Lemma 4.7 is uniquely defined in (4.12), thus $R_{Y_0, \lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}} = \lim_{\epsilon\downarrow 0} R_{Y_0+\epsilon, M_{Y_0-\epsilon}}$. Similarly, we have $R_{Y_0, \lim_{\epsilon\downarrow 0} M_{Y_0+\epsilon}} = \lim_{\epsilon\downarrow 0} R_{Y_0-\epsilon, M_{Y_0+\epsilon}}$. By the Dominated Convergence Theorem and (4.14),

$$\begin{aligned}&\mathbb{E}\left[I\left(\lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}, Y_0\xi + \gamma R_{Y_0, \lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}}\right) - \lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}\right]\\
&= \mathbb{E}\left[I\left(\lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}, Y_0\xi + \gamma \lim_{\epsilon\downarrow 0} R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right) - \lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}\right]\\
&= \mathbb{E}\left[\lim_{\epsilon\downarrow 0}\left(I\left(M_{Y_0-\epsilon}, (Y_0+\epsilon)\,\xi + \gamma R_{Y_0+\epsilon, M_{Y_0-\epsilon}}\right) - M_{Y_0-\epsilon}\right)\right] = 0.\end{aligned}$$

Since M in Lemma 4.8 is uniquely defined in (4.14), thus $\lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon} = M_{Y_0}$. Similarly, we have $\lim_{\epsilon\uparrow 0} M_{Y_0-\epsilon} = M_{Y_0}$ and $\lim_{Y\to Y_0} M_Y = M_{Y_0}$. Hence,

$$\lim_{\epsilon\downarrow 0} R_{Y_0+\epsilon, M_{Y_0-\epsilon}} = R_{Y_0, \lim_{\epsilon\downarrow 0} M_{Y_0-\epsilon}} = R_{Y_0, M_{Y_0}} = R_{Y_0, \lim_{\epsilon\downarrow 0} M_{Y_0+\epsilon}} = \lim_{\epsilon\downarrow 0} R_{Y_0-\epsilon, M_{Y_0+\epsilon}}.$$

By (C.4), we have $\lim_{Y\to Y_0} R_{Y, M_Y} = R_{Y_0, M_{Y_0}}$. Then we have

$$\begin{aligned}\lim_{Y\to Y_0} I\left(M_Y, Y\xi + \gamma R_{Y, M_Y}\right) &= I\left(\lim_{Y\to Y_0} M_Y, Y_0\xi + \gamma \lim_{Y\to Y_0} R_{Y, M_Y}\right)\\
&= I\left(M_{Y_0}, Y_0\xi + \gamma R_{Y_0, M_{Y_0}}\right).\end{aligned}$$

Finally, our claim follows from another application of the Dominated Convergence Theorem.

(b1) For an arbitrary a sample value $\xi_0 \in (0, \infty)$. Assume the contrary that

$$\liminf_{Y\to 0} I\left(M_Y, Y\xi_0 + \gamma R_{Y, M_Y}\right) < \infty,$$

then there exists a sequence $\{y_n\}$ with $y_n \to 0$ such that

$$\liminf_{Y\to 0} I\left(M_Y, Y\xi_0 + \gamma R_{Y, M_Y}\right) = \lim_{n\to\infty} I\left(M_{y_n}, y_n\xi_0 + \gamma R_{y_n, M_{y_n}}\right) < \infty.$$

Clearly, $\lim_{n\to\infty} U'\left(I\left(M_{y_n}, y_n\xi_0 + \gamma R_{y_n, M_{y_n}}\right)\right) > 0$. Furthermore, since $U'(I(m,\cdot))$ is increasing, thus for any $\xi > \xi_0$,

$$\liminf_{n\to\infty} U'\left(I\left(M_{y_n}, y_n\xi + \gamma R_{y_n, M_{y_n}}\right)\right) \geq \lim_{n\to\infty} U'\left(I\left(M_{y_n}, y_n\xi_0 + \gamma R_{y_n, M_{y_n}}\right)\right) > 0.$$

By Fatou's Lemma and (4.13),

$$\begin{aligned}\liminf_{n\to\infty} y_n\mathbb{E}[\xi] &= \liminf_{n\to\infty} \mathbb{E}\left[U'\left(I\left(M_{y_n}, y_n\xi + \gamma R_{y_n, M_{y_n}}\right)\right)\right]\\
&\geq \mathbb{E}\left[\liminf_{n\to\infty} U'\left(I\left(M_{y_n}, y_n\xi + \gamma R_{y_n, M_{y_n}}\right)\right)\right]\\
&\geq \lim_{n\to\infty} U'\left(I\left(M_{y_n}, y_n\xi_0 + \gamma R_{y_n, M_{y_n}}\right)\right)\mathbb{P}\left[\xi > \xi_0\right] > 0,\end{aligned}$$

which contradict $\liminf_{n\to\infty} y_n\mathbb{E}[\xi] = 0$. So $\liminf_{Y\to 0} I\left(M_Y, Y\xi_0 + \gamma R_{Y,M_Y}\right) = \infty$, for any $\xi_0 \in (0,\infty)$. Hence, $\liminf_{Y\to 0} \xi I\left(M_Y, Y\xi + \gamma R_{Y,M_Y}\right) = \infty$ almost surely. By Fatou's Lemma,

$$\liminf_{Y\to 0} \mathbb{E}\left[\xi I\left(M_Y, \gamma R_{Y,M_Y} + Y\xi\right)\right] \geq \mathbb{E}\left[\liminf_{Y\to 0} \xi I\left(M_Y, \gamma R_{Y,M_Y} + Y\xi\right)\right] = \infty.$$

(b2) Since M_Y is decreasing in Y as shown in Lemma 4.8, thus $\lim_{Y\to\infty} M_Y$ exists and is finite.
For any $N \in (0,\infty)$, it is clear from its definition that $\lim_{Y\to\infty} R_{N,M_Y} \geq 0$. Therefore, $\lim_{Y\to\infty} I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right) = I\left(\lim_{Y\to\infty} M_Y, N\xi + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right)$ is finite almost surely for any $N \in (0,\infty)$. Since $I\left(\lim_{Y\to\infty} M_Y, N\xi + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right)$ is decreasing in N, then $\lim_{N\to\infty} I\left(\lim_{Y\to\infty} M_Y, N\xi + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right)$ exists and is finite since I is always non-negative.
For if there exists a sample $\xi_0 \in (0,\infty)$ such that

$$\lim_{N\to\infty} I\left(\lim_{Y\to\infty} M_Y, N\xi_0 + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right) > 0, \tag{C.5}$$

then $\lim_{N\to\infty} D'\left(\lim_{Y\to\infty} M_Y - I\left(\lim_{Y\to\infty} M_Y, N\xi_0 + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right)\right) < \infty$ and $\lim_{N\to\infty} U'\left(I\left(\lim_{Y\to\infty} M_Y, N\xi_0 + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right)\right) < \infty$ if (C.5) holds. By (4.8),

$$\begin{aligned}N\xi_0 + \gamma \lim_{Y\to\infty} R_{N,M_Y} &= U'\left(I\left(\lim_{Y\to\infty} M_Y, N\xi_0 + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right)\right)\\ &\quad + D'\left(\lim_{Y\to\infty} M_Y - I\left(\lim_{Y\to\infty} M_Y, N\xi_0 + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right)\right).\end{aligned}$$

Then taking $N \to \infty$ into the both sides, the limit in the left hand side tends to infinity as $\lim_{Y\to\infty} R_{N,M_Y} \geq 0 > -\infty$ for any finite fixed N, while the limit in the right hand side is finite, a contradiction. So

$$\lim_{N\to\infty} I\left(\lim_{Y\to\infty} M_Y, N\xi + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right) = 0, \quad a.s. \tag{C.6}$$

On the other hand, since $I\left(M_Y, Y\xi + \gamma R_{Y,M_Y}\right) \leq I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right)$ for all $N \leq Y < \infty$ and $\lim_{Y\to\infty} I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right)$ exists, we have almost surely, for any N,

$$\begin{aligned}\limsup_{Y\to\infty} I\left(M_Y, Y\xi + \gamma R_{Y,M_Y}\right) &\leq \lim_{Y\to\infty} I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right)\\ &= I\left(\lim_{Y\to\infty} M_Y, N\xi + \gamma \lim_{Y\to\infty} R_{N,M_Y}\right).\end{aligned}$$

By (C.6), $\limsup_{Y\to\infty} I\left(M_Y, Y\xi + \gamma R_{Y,M_Y}\right) = 0$ almost surely. Finally, by reverse Fatou's lemma, since $I\left(M_Y, Y\xi + \gamma R_{Y,M_Y}\right) \leq I\left(M_{k_0}, k_0\xi\right) \leq I\left(M_{k_0}, U'(M_{k_0})\right) + (U')^{-1}(k_0\xi) \in \mathcal{L}^2$ for $Y \geq k_0$, where k_0 is given in Assumption 4.4, we have

$$\begin{aligned}&\limsup_{Y\to\infty} \mathbb{E}\left[\xi I\left(M_Y, Y\xi + \gamma R_{Y,M_Y}\right)\right]\\ &\leq \mathbb{E}\left[\limsup_{Y\to\infty} \xi I\left(M_Y, Y\xi + \gamma R_{Y,M_Y}\right)\right] = 0.\end{aligned} \tag{C.7}$$

C.5. Proof of Proposition 4.12. The proof is essentially the same as the proof of Proposition 4.5 except that $\lim_{Z\to\infty} U'(Z) + \gamma D'(m - Z) - y = -\infty$ for any $(m, y) \in \mathbb{R}^2$.

C.6. Proof of Proposition 4.13.

(i) Since R can take values in $(-\infty, D'(M))$ instead. We have to prove the following in place of that in (c1) in Appendix C.2:

$$\lim_{R\to-\infty} \mathbb{E}[D'\left(M - I\left(M, \gamma R + Y\xi\right)\right) - R] > 0.$$

By (4.16), we have $\lim_{R\to-\infty} I\left(M, \gamma R + Y\xi\right) = \infty$ almost surely. Since U' is continuous, so we have $\lim_{R\to-\infty} Y\xi - U'(I\left(M, \gamma R + Y\xi\right)) = Y\xi$ almost surely by the Inada condition. Since $Y\xi - U'\left(I\left(M, \gamma R + Y\xi\right)\right)$ is strictly increasing as $R \to -\infty$ and $Y\xi - U'\left(I\left(M, \gamma R + Y\xi\right)\right) \geq D'\left(M - I\left(M, 0\right)\right)$ for $R < 0$ by (4.16), then by the Monotone Convergence Theorem,

$$\begin{aligned}\lim_{R\to-\infty} \mathbb{E}\left[D'\left(M - I\left(M, \gamma R + Y\xi\right)\right) - R\right] &= \lim_{R\to-\infty} \mathbb{E}\left[\frac{1}{\gamma}\left(Y\xi - U'(I\left(M, \gamma R + Y\xi\right))\right)\right] \\ &= \frac{1}{\gamma} Y\mathbb{E}\left[\xi\right] > 0.\end{aligned}$$

(ii) In part (d1) in Appendix C.2 , by Proposition 4.12 (a), $I\left(M, \gamma R + Y\xi\right)$ is strictly increasing in M, so $D'(M - I\left(M, \gamma R + Y\xi\right)) - R$ is strictly increasing in M.

(iii) In part (a) in Appendix C.3, by (4.8), $D'(M - I\left(M, \gamma R + Y\xi\right)) - R$ is strictly increasing in M almost surely. Since D' is strictly increasing, $I\left(M, \gamma R + Y\xi\right) - M$ is strictly decreasing in M for fixed Y, R.

(iv) In part (b2) in Appendix C.4, since R can be negative, we no longer have $\lim_{Y\to\infty} R_{N,M_Y} \geq 0$. Instead, we claim that for any $N \in (0, \infty)$, $\lim_{Y\to\infty} R_{N,M_Y} > -\infty$. Assume the contrary, that there exists N such that $\lim_{Y\to\infty} R_{N,M_Y} = -\infty$. Then

$$\lim_{Y\to\infty} \left(\xi N + \gamma R_{N,M_Y}\right) = -\infty \text{ for all } \xi \in (0, \infty). \tag{C.8}$$

Fix a sample $\xi_0 \in (0, \infty)$. Let $\{y_n\}$ be a sequence with $y_n \to \infty$ such that

$$\lim_{n\to\infty} I\left(M_{y_n}, N\xi_0 + \gamma R_{N,M_{y_n}}\right) = \liminf_{Y\to\infty} I\left(M_Y, N\xi_0 + \gamma R_{N,M_Y}\right).$$

If $\lim_{n\to\infty} I\left(M_{y_n}, N\xi_0 + \gamma R_{N,M_{y_n}}\right) < \infty$, then $\lim_{n\to\infty} U'\left(I\left(M_{y_n}, N\xi_0 + \gamma R_{N,M_{y_n}}\right)\right) > 0$ and $\lim_{n\to\infty} D'\left(M_{y_n} - I\left(M_{y_n}, N\xi_0 + \gamma R_{N,M_{y_n}}\right)\right) > -\infty$. With (4.16), they contradict to (C.8). So

$$\liminf_{Y\to\infty} I\left(M_Y, N\xi_0 + \gamma R_{N,M_Y}\right) = \lim_{n\to\infty} I\left(M_{y_n}, N\xi_0 + \gamma R_{N,M_{y_n}}\right) = \infty.$$

Thus, we have $\liminf_{Y\to\infty} I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right) = \infty$ almost surely. Then,

$$\lim_{Y\to\infty} U'\left(I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right)\right) = U'\left(\lim_{Y\to\infty} I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right)\right) = 0.$$

Since $U'\left(I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right)\right) \leq N\xi + \gamma R_{N,M_N} - D'(-I(0, \gamma R_{N,M_N}))$ for $Y > N$, then by the reverse Fatou lemma, we have

$$\begin{aligned}0 &= \mathbb{E}\left[\limsup_{Y\to\infty} U'\left(I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right)\right)\right] \\ &\geq \limsup_{Y\to\infty} \mathbb{E}\left[U'\left(I\left(M_Y, N\xi + \gamma R_{N,M_Y}\right)\right)\right] \\ &= N\mathbb{E}\left[\xi\right] > 0,\end{aligned}$$

which is a contradiction; thus for any $N \in (0,\infty)$, $\lim_{Y\to\infty} R_{N,M_Y} > -\infty$. From here on the rest of the proof is the same as that for part (b2) in Proof for Lemma 4.9 until the last line, $I\left(M_Y, Y\xi + \gamma R_{Y,M_Y}\right) \le I\left(M_1, \gamma \lim_{Y\to\infty} R_{1,M_Y}\right) < \infty$, which is independent of ξ. Hence, the reverse Fatou Lemma still works in (C.7).

C.7. Proof of Proposition 4.21. Firstly, we know that given $x, y > 0$, if $0 < x + y \le 1$, then $|\ln(x+y)| \le |\ln x|$; on the other hand, if $x + y > 1$, then $0 < \ln(x+y) < \ln x + \frac{y}{x}$. Combining, $|\ln(x+y)| \le |\ln(x)| + \left|\frac{y}{x}\right|$. For fixed $R \in \left[\frac{1}{\gamma}, \infty\right)$,

$$
\begin{aligned}
\mathbb{E}\left[\left|\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)\right|^2\right] &\le \mathbb{E}\left[\left(\left|\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]}\right)\right| + \left|\mathbb{E}[\xi]\frac{\gamma R - 1}{\xi}\right|\right)^2\right] \\
&\le \mathbb{E}\left[\left(|\ln\xi| + |\ln(\gamma\mathbb{E}[\xi])| + (\gamma R - 1)\mathbb{E}[\xi]\xi^{-1}\right)^2\right] \\
&\le 3\left(\mathbb{E}\left[|\ln\xi|^2\right] + \ln(\gamma\mathbb{E}[\xi])^2 + (\gamma R-1)^2\mathbb{E}[\xi]^2\mathbb{E}\left[\xi^{-2}\right]\right). \qquad \text{(C.9)}
\end{aligned}
$$

Clearly, by a simple calculation,

$$
\begin{aligned}
&\mathbb{E}\left[|\ln\xi|^2\right] \\
&\le \mathbb{E}\left[2T\int_0^T \left|r(s) + \frac{1}{2}\alpha(s)^t\left(\sigma(s)\sigma(s)^t\right)^{-1}\alpha(s)\right|^2 ds + 2\int_0^T \alpha(s)^t\left(\sigma(s)\sigma(s)^t\right)^{-1}\alpha(s)ds\right] \\
&\le 2T^2\left(r + \frac{C}{2}\right)^2 + 2CT,
\end{aligned}
$$

where C is a constant. With this last result and the boundedness of $\mathbb{E}[\xi^{-2}]$, we can show that (C.9) is bounded, and hence $\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right) \in \mathcal{L}^2$ for any $R \in [\frac{1}{\gamma}, \infty)$.

Now, we consider three different cases: (i) $\gamma > \exp\left(\mathbb{E}\left[\ln\left(\frac{\xi}{\mathbb{E}[\xi]}\right)\right]\right)$, (ii) $\gamma = \exp\left(\mathbb{E}\left[\ln\left(\frac{\xi}{\mathbb{E}[\xi]}\right)\right]\right)$, and (iii) $\gamma < \exp\left(\mathbb{E}\left[\ln\left(\frac{\xi}{\mathbb{E}[\xi]}\right)\right]\right)$.

- **(i)** Obviously, $\mathbb{E}\left[\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)\right]$ is strictly increasing and continuous in $R \in \left(\frac{1}{\gamma}, \infty\right)$. By the Monotone Convergence Theorem, $\lim_{R\to\infty}\mathbb{E}\left[\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)\right] = \infty$. On the other hand, $\lim_{R\to\frac{1}{\gamma}}\mathbb{E}\left[\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]} + R - \frac{1}{\gamma}\right)\right] = \mathbb{E}\left[\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]}\right)\right] < 0$. Hence, by intermediate value theorem, there exist an unique R satisfying (4.22) and (4.23).
- **(ii)** $R = \frac{1}{\gamma}$ is the unique solution satisfying (4.22) and (4.23).
- **(iii)** Assume the contrary that there exists an admissible solution $\hat{X}$ being an optimal terminal wealth for this mean-exponential-risk problem. Since $D = D'$, so $\hat{X}$ satisfies Condition 3.1 (i). It is clear that $\hat{X}$ satisfies Condition 3.1 (ii). Hence, by Theorem 3.2, it is necessary that there exist numbers Y, M, R such that $(\hat{X}, Y, M, R)$ solves the Nonlinear Moment Problem (4.18)-(4.21). By (4.18) and (4.21), we first have

$$
\hat{X} = M - \ln\left(R + \frac{\xi Y}{\gamma} - \frac{1}{\gamma}\right), \quad Y = \frac{1}{\mathbb{E}[\xi]}. \qquad \text{(C.10)}
$$

Given that $\gamma < \exp\left(\mathbb{E}\left[\ln\left(\frac{\xi}{\mathbb{E}[\xi]}\right)\right]\right)$, we have $\ln\left(\frac{\xi}{\gamma\mathbb{E}[\xi]}\right) > \ln[\xi] - \mathbb{E}[\ln[\xi]]$. Taking expectation on both sides of $\hat{X}$ in (C.10), for any $R \in \left[\frac{1}{\gamma}, \infty\right)$, we have

$$\mathbb{E}\left[\hat{X}\right] = M - \mathbb{E}\left[\ln\left(R + \frac{\xi Y}{\gamma} - \frac{1}{\gamma}\right)\right] \geq M - \mathbb{E}\left[\ln\left(\frac{\xi Y}{\gamma}\right)\right]$$
$$> M - \mathbb{E}\left[\ln\left[\xi\right] - \mathbb{E}\left[\ln\left[\xi\right]\right]\right] = M,$$

which contradicts with (4.20). Therefore, there is no solution for the Nonlinear Moment Problem, and hence, this mean-exponential risk problem has no optimal solution.

Appendix D. On the Boundedness of Optimal Wealth.

In the proof of Theorem 4.10, it seems not immediate that the optimal terminal wealth is uniformly bounded even if the risk measure is a downside one. By Theorems 2.5 and 3.2, there exist numbers Y, M and R so that any optimal terminal wealth $\hat{X}$ (satisfying Conditions 3.1 (i) and (ii)) satisfies:

$$\begin{cases} Y\xi = f_{M,R}(\hat{X}), & \text{a.s. on } \{\hat{X} > 0\}, \\ Y\xi \leq f_{M,R}(\hat{X}), & \text{a.s. on } \{\hat{X} = 0\}, \end{cases} \tag{D.1}$$

where $f_{M,R}(x) := U'(x) - \gamma R + \gamma D'(M - x)$. By taking expectation on both sides of (D.1) with some terms being eliminated in accordance with (3.4), we have $Y \geq \mathbb{E}[U'(\hat{X})]/\mathbb{E}[\xi] > 0$. By the definition of I in Proposition 4.5 and the fact that $f_{M,R}$ is decreasing, $f_{M,R}(x) \leq 0$ whenever $x \geq I(M, \gamma R) > 0$, together with the facts that $Y\xi > 0$ and $\hat{X}$ has to satisfy (D.1) a.s., there is no possibility that $\hat{X}$ takes value greater than $I(M, \gamma R)$. In other words, $\hat{X}$ has to be bounded above by the finite deterministic number $I(M, \gamma R)$. Note that the optimal terminal wealth is bounded when the risk function is strictly convex. In particular, the optimal terminal payoff in our utility-risk problem in Theorem 4.10 is countermonotonic with the pricing kernel, which is a commonly found property in the portfolio selection literature.

To motivate the claim of the boundedness of the optimal payoff from a financial perspective, we consider a simple single period example. Based on the previous observation, it is justifiable to simply take the optimal terminal payoff under this example to be also countermonotonic with the pricing kernel.

We suppose that the payoff is a random variable Z with two possible outcomes, 0 and a number $z > 1$, and their respective probabilities are $p_0 := 1 - p_z$ and p_z. Our objective function is

$$J(Z) := \mathbb{E}[U(Z)] - \mathbb{E}[D(\mathbb{E}[Z] - Z)] := \mathbb{E}[Z^\theta] - \mathbb{E}[(\mathbb{E}[Z] - Z)_+^\rho], \tag{D.2}$$

where $\theta < 1$ and $\rho > 1$. There is a budget constraint on the payoff, namely

$$\mathbb{E}[\xi Z] = q_z z = 1, \tag{D.3}$$

where $q_z := \mathbb{E}[\xi \mathbb{I}\{Z = z\}]$. Note that, if a single-period risk-free simple rate is given to be r, $(1+r)q_z$ becomes the risk neutral probability of $Z = z$. We look for the optimal z so that the corresponding payoff Z maximizes (D.2). Since we assume that the payoff Z and the pricing kernel ξ are countermonotonic, there exist a $\xi_0(z) \in (0, \infty)$ such that $\{Z = z\} = \{\xi < \xi_0(z)\}$, thus $q_z = \int_0^{\xi_0(z)} \xi \mathbb{P}[d\xi]$. Then, in order to maintain the budget constraint at the same level, increasing z has to be balanced off with a smaller risk neutral probability q_z, thus $\xi_0(z)$ decreases in z. Therefore, $p_z = \int_0^{\xi_0(z)} \mathbb{P}[d\xi]$ decreases in z. Define $h(z) := \frac{q_z}{p_z} = \frac{\int_0^{\xi_0(z)} \xi \mathbb{P}[d\xi]}{\int_0^{\xi_0(z)} \mathbb{P}[d\xi]}$, simple calculus concludes that $h'(z) < 0$, and therefore $h(z)$ decreases in z. By (D.3), $p_z = \frac{h(z)^{-1}}{z}$ and $\mathbb{E}[Z] = h(z)^{-1}$. Then, we have

(i) $\mathbb{E}[U(Z)] = \mathbb{E}[Z^\theta] = \dfrac{h(z)^{-1}}{z^{1-\theta}}$ and

(ii) $\mathbb{E}[D(\mathbb{E}[Z] - Z)] = \mathbb{E}[(\mathbb{E}[Z] - Z)_+^\rho] = (\mathbb{E}[Z])^\rho(1 - p_z) \geq \delta_0(h(z)^{-1})^\rho$,

for some $\delta_0 > 0$. Next, we consider two cases: (i) $\lim_{z\to\infty} h(z)^{-1} < \infty$ or (ii) $\lim_{z\to\infty} h(z)^{-1} = \infty$.

(i) We have $\lim_{z\to\infty} \mathbb{E}[U(Z)] = \lim_{z\to\infty} \frac{h(z)^{-1}}{z^{1-\theta}} = 0$, i.e. a bounded Z is optimal even in the ordinary utility maximization.

(ii) We have

$$\lim_{z\to\infty} \frac{\mathbb{E}[D(\mathbb{E}[Z] - Z)]}{\mathbb{E}[U(Z)]} \geq \lim_{z\to\infty} \frac{\delta_0(h(z)^{-1})^\rho}{\frac{h(z)^{-1}}{z^{1-\theta}}} \geq \lim_{z\to\infty} \delta_0(h(z)^{-1})^{\rho-1} = \infty,$$

which means that $\mathbb{E}[D(\mathbb{E}[Z] - Z)]$ grows with z faster than $\mathbb{E}[U(Z)]$ due to the diminishing marginal value of utility U and the increasing marginal value of deviation risk $D(\mathbb{E}[Z] - Z)$. As a result, taking an arbitrarily large value in z actually causes a negative effect on the objective value.

In the case that the risk measure penalizes both upside and downside deviation risks, the investor resist to put more proportion of wealth into more risky asset to avoid double penalties due to increased upside and downside deviation risks. Under a downside risk measure, there is no double penalties on deviation risk, so it is intuitively expected that investing in more risky asset is comparatively more encouraging. Nevertheless, our general results and the simple example show that it is not the case. Investing in more risky asset can increase both return and downside deviation risk. On the one hand, since the utility has a diminishing marginal value, any additional gain in a large return can only give a reducing increment on the investor's satisfaction. On the other hand, the additional gain in a large return makes that the convex downside risk measure enlarges the the investor's dissatisfaction on downside deviation risk. As a result, such additional dissatisfaction on risk compensates the very marginal increment in the satisfaction on return. Hence, this intuition motivates the uniform boundedness of optimal payoff.