This is the Pre-Published Version.
This is the accepted manuscript of the following article: Xue Dong He, Zhaoli Jiang (2021) Mean-Variance Portfolio Selection with Dynamic Targets for Expected Terminal Wealth. Mathematics of Operations Research 47(1):587-615, which has been published in final form at https://doi.org/10.1287/moor.2021.1142.

Submitted to *Mathematics of Operations Research*
manuscript (Please, provide the manuccript number!)

Authors are encouraged to submit new papers to INFORMS journals by means of a style file template, which includes the journal title. However, use of a template does not certify that the paper has been accepted for publication in the named journal. INFORMS journal templates are for the exclusive purpose of submitting to an INFORMS journal and should not be used to distribute the papers in print or online or to submit the papers to another publication.

Mean-Variance Portfolio Selection with Dynamic Targets for Expected Terminal Wealth

Xue Dong He
Room 505, William M.W. Mong Engineering Building, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, xdhe@se.cuhk.edu.hk

Zhaoli Jiang
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, zljiang@se.cuhk.edu.hk

In a market that consists of multiple stocks and one risk-free asset whose mean return rates and volatility are deterministic, we study a continuous-time mean-variance portfolio selection problem in which an agent is subject to a constraint that the expectation of her terminal wealth must exceed a target and minimizes the variance of her terminal wealth. The agent can revise her expected terminal wealth target dynamically to adapt to the change of her current wealth, and we consider the following three targets: (i) the agent's current wealth multiplied by a target expected gross return rate, (ii) the risk-free payoff of the agent's current wealth plus a premium, and (iii) a weighted average of the risk-free payoff of the agent's current wealth and a pre-set aspiration level. We derive the so-called equilibrium strategy in closed form for each of the three targets and find that the agent effectively minimizes the variance of the instantaneous change of her wealth subject to a certain constraint on the expectation of the instantaneous change of her wealth.

Key words: portfolio selection, dynamic mean-variance analysis, time inconsistency, equilibrium strategies
MSC2000 subject classification: Primary: 91G10; secondary: 93E20
OR/MS subject classification: Primary: Finance: portfolio; secondary: Dynamic programming/optimal control: applications
History:

---

1. Introduction Expected utility maximization and mean-variance analysis are two of the most commonly used approaches to portfolio selection. In the former, an investor maximizes the expected utility of her wealth at the end of a certain investment period. In the latter, an investor minimizes the variance of her wealth at the end of a given period subject to a constraint that the expectation of the gross return in the same period must be as high as a given target. See for instance Merton [38] for the former approach and Markowitz [37] for the latter.

One advantage of the mean-variance analysis over the expected utility maximization is that the target for the expected gross return is directly elicitable by asking investors simple and understandable questions, whereas the utility function of an investor is not. In addition, the measure used to assess investment risk, namely the variance of terminal wealth, is relatively easier to understand than utility functions by common folk. One disadvantage of the mean-variance analysis is that, when applied to portfolio selection in multiple periods or in continuous time, it leads to time-inconsistent behavior. More precisely, at each time the so-called pre-committed strategy of an agent, which minimizes the variance of the gross return of the investment with the expectation of the gross return exceeding a target that is set by the agent at that time, no longer minimizes

the variance of the gross return or meets the target for the expected gross return at future time. Thus, the objectives of the agent's selves at different times do not align with each other.

One possible remedy for the time inconsistency is to force the agent's future selves to follow the pre-committed strategy that is set up today with the help of certain commitment device. For mean-variance portfolio selection problems, pre-committed strategies have been extensively studied in the literature; see for instance Li and Ng [34] for the discrete-time setting and Richardson [45], Bajeux-Besnainou and Portait [2], Zhou and Li [50], Bielecki et al. [8], and Lim and Zhou [36] for the continuous-time setting. The Lagrange dual method is commonly applied in the above works to find the pre-committed strategy. More precisely, suppose today is t and the end of the investment horizon is T. The pre-committed strategy for the agent today is then the optimal solution to the following optimization problem:

$$\begin{aligned} \min_{\boldsymbol{\pi}} \quad & \mathrm{var}_t(X^{\boldsymbol{\pi}}(T)) \\ \text{subject to } & \mathbb{E}_t\left[X^{\boldsymbol{\pi}}(T)\right] \geq L_t, \end{aligned}$$

where $X^{\boldsymbol{\pi}}(T)$ stands for the agent's wealth at T under strategy $\boldsymbol{\pi}$, $\mathbb{E}_t[X^{\boldsymbol{\pi}}(T)]$ and $\mathrm{var}_t(X^{\boldsymbol{\pi}}(T))$ stand for the expectation and variance of $X^{\boldsymbol{\pi}}(T)$, respectively, and L_t is the agent's target today for the expected terminal wealth. Note that this formulation is the same as minimizing the variance of the gross return subject to a constraint on the expected gross return because the gross return is simply the agent's wealth at T divided by her current wealth. This formulation, referred to as the *constraint formulation*, is equivalent to the following *penalty formulation*

$$\min_{\boldsymbol{\pi}} \ -\mathbb{E}_t\left[X^{\boldsymbol{\pi}}(T)\right] + \big(\Gamma_t/2\big)\mathrm{var}_t(X^{\boldsymbol{\pi}}(T))$$

for certain Γ_t that represents the agent's risk aversion degree, in that the optimal strategies, which are the agent's pre-committed strategies, are the same in these two formulations. Therefore, in the literature these two formulations are not differentiated when pre-committed strategies are sought.

Another remedy for time inconsistency is to consider the agent's selves at different future times as different players and then to seek an equilibrium strategy. More precisely, at each time the agent can control herself in a small time period only, so at that time the agent chooses a strategy that can be implemented in this time period in order to optimize her objective which also depends on the action of the future selves whom her today's self cannot control. Such an idea dates back to Strotz [46], and it should be the choice of a rational agent who has no control of her future selves and is sophisticated enough to foresee the action of her future selves.

To find the equilibrium strategy, we first need to model the agent's preferences at each time of the investment horizon. In the mean-variance analysis, we need to model how the agent's expected terminal wealth target L_t or risk aversion degree Γ_t dynamically adapts to the evolution of time and the agent's wealth. In the literature, the study of the equilibrium strategy in the mean-variance analysis focuses exclusively on the penalty formulation and assumes some particular forms for the risk aversion degree Γ_t. For example, Basak and Chabakauri [3] and Pun [44] assume Γ_t to be a constant. Björk et al. [11], Bensoussan et al. [7], and Bensoussan et al. [6] assume Γ_t to be inversely proportional to the agent's current wealth. Kryger et al. [32] consider a general class of portfolio selection problems, and one of the examples considered therein is the mean-variance problem under the penalty formulation with $1/\Gamma_t$ to be an affine function of the agent's current wealth. Cui et al. [17] and Cui, Li, Li and Shi [16] consider a form of $1/\Gamma_t$ that is piecewise linear in the agent's current wealth. Bayraktar et al. [4] consider equilibrium stopping strategies under mean-variance and mean-standard-deviation criteria with a constant risk aversion degree Γ_t.

Using the penalty formulation in the study of mean-variance equilibrium strategies, however, loses the advantage of the mean-variance analysis over the expected utility maximization, because similar to utility functions, the risk aversion parameter Γ_t is neither directly elicitable nor easily

understandable. In particular, it is unclear whether the models of the risk aversion degree used in the aforementioned studies of mean-variance analysis under the penalty formulation are reasonable, and it is also unknown which of these models represents investors' risk attitude most accurately.

To overcome the above difficulties, in the present paper, we propose a notion of equilibrium strategies for mean-variance analysis under the constraint formulation and then solve the equilibrium strategy explicitly when the target for expected terminal wealth L_t takes some reasonable, flexible forms. Thus, our approach restores the advantage of the mean-variance analysis over the expected utility maximization, as L_t is taken directly as an input in our approach. In addition, it is possible to test in experiments whether investors' target for expected terminal wealth takes the forms suggested in the present paper and to directly estimate the target, whereas it seems difficult to do so for the risk aversion parameter Γ_t in the penalty formulation.

More precisely, we consider a dynamic mean-variance portfolio selection problem in continuous time, in which an agent with certain investment horizon has a target for her expected terminal wealth at each time and tries to minimize the risk of her investment that is represented by the variance of her terminal wealth. The agent can invest in multiple risky stocks whose mean return rates and volatility are deterministic and in a risk-free asset whose return rate is also deterministic. The market can be incomplete in that the dimension of the Brownian motion that drives asset prices can be larger than the number of the stocks and in that we allow for portfolio constraints. We propose a notion of equilibrium strategies that is new in the literature to take into account the constraint on the expected terminal wealth.

We consider three plausible models of the agent's target for terminal wealth. In the first model, at each time the agent attempts to achieve a given target for the expected gross return rate of her investment. In other words, the agent's target for the expected terminal wealth is proportional to her current wealth. In the second model, the agent's expected terminal wealth target is equal to the risk-free payoff of her current wealth plus a premium that is independent of her current wealth. In the third model, the agent's expected terminal wealth target is a weighted average of the risk-free payoff of the agent's current wealth and an aspiration level pre-set by the agent.

For each of the three models, we solve the equilibrium strategy in closed form and prove uniqueness in a large set of feasible strategies. Thanks to the setting of deterministic asset mean return rates and volatility, the equilibrium strategy obtained in the present paper is a myopic one in that it minimizes the variance of the instantaneous change of the agent's wealth subject to a constraint on the expectation of the same change, although the objective in the mean-variance problem faced by the agent concerns the agent's terminal wealth and thus is non-myopic. Mathematically, the equilibrium strategy can be obtained by solving a quadratic program. Note, however, that the constraint on the expectation of the instantaneous wealth change varies with the choice of the expected terminal wealth target used by the agent, leading to a variety of strategies. More precisely, for the first target, the equilibrium strategy is to invest a wealth-independent proportion of the agent's wealth in the stocks. For the second target, the equilibrium strategy is to invest a wealth-independent dollar amount in the stocks. For the third target, the equilibrium strategy is to invest in the stocks a dollar amount that is proportional to the distance between the agent's current wealth and the discounted value of the aspiration level.

When seeking pre-committed strategies, the constraint formulation and the penalty formulation are not differentiated because we can find a one-to-one mapping between the expected terminal wealth target L_t in the former and the risk aversion degree Γ_t in the latter by matching the pre-committed strategies in these two formulations. One can then model or elicit L_t from investors' behavior, employ the above mapping to obtain Γ_t, and apply Γ_t in the penalty formulation to derive an equilibrium strategy. Such a solution, however, does not work properly. Indeed, as we will illustrate in the present paper, with L_t and Γ_t chosen such that the pre-committed strategies in above two formulations are same, the equilibrium strategy in the constraint formulation leads to

strictly more investment in stocks than (and thus differs from) the one in the penalty formulation. Thus, in the study of mean-variance equilibrium strategies, the constraint formulation is more intuitive and flexible to take into account investors' risk attitude as represented by their target for expected terminal wealth.

To summarize, the contribution of our work is three-fold: First, we propose a general framework to study equilibrium strategies for mean-variance analysis under the constraint formulation. Second, assuming the mean return rates and volatility of the assets to be deterministic and taking portfolio constraints into account, we derive closed-form formulae for the equilibrium strategy for three types of targets for the expected terminal wealth. Third, we show that the equilibrium strategy in the constraint formulation implies more investment in risky stocks than (and thus differs from) the one in the penalty formulation, assuming that the expected terminal wealth target in the former and the risk aversion parameter in the latter are chosen so that the pre-committed strategies in these two formulations are the same.

Let us emphasize that our framework is the first of its kind to study time-inconsistent problems with dynamic constraints on expected values of the state process, and as the first step to build such a framework, we consider the case in which the mean return rates and volatility of the assets are deterministic. The presence of the dynamic targets for expected terminal wealth in our model leads to new challenges compared to the studies of mean-variance analysis under the penalty formulation in the literature. In particular, we cannot use the extended HJB equation developed by Björk et al. [9] and some other similar tools in Kryger et al. [32] and He and Jiang [25]. To address these challenges, we build a connection between an equilibrium strategy and a constrained quadratic program. Then, we study the existence and uniqueness of the equilibrium strategy by looking into some properties of the quadratic program. This method is new in the literature.

Following most literature on time-inconsistent stochastic portfolio selection, such as Ekeland and Pirvu [24], Björk et al. [11], Bensoussan et al. [7], Bensoussan et al. [6], and Kryger et al. [32], we focus on the setting of deterministic asset mean returns and volatility. Our framework, however, can also be applied beyond this setting. Indeed, in a follow-up paper, He and Jiang [26] apply this framework to study a mean-variance portfolio selection problem with a dynamic expected terminal wealth target in a stochastic volatility model and derive the equilibrium portfolio strategy in closed form up to solving a Ricatti equation. In their derivation of the equilibrium strategy, besides dealing with some new challenges due to stochastic volatility, He and Jiang [26] also use the idea of solving a constrained quadratic program as we apply in the present paper. It is worth mentioning that with penalty formulation, equilibrium mean-variance portfolio selection in a stochastic volatility market has not been completely solved in the literature.[1] Thus, the results in He and Jiang [26] prove the effectiveness and general applicability of our framework.

To our best knowledge, the only two papers in the literature that discussed to some extent equilibrium strategies under the constraint formulation are Cui, Li and Li [14] and Dai et al. [19]. In their Appendix A9, Cui, Li and Li [14] consider a discrete-time mean-variance portfolio problem

[1] With a constant risk aversion parameter, Basak and Chabakauri [3] derive an equilibrium strategy in closed form, but this strategy implies investment of a wealth-independent dollar amount in stocks and thus is not appealing. Dai et al. [19] consider an unconventional mean-variance model with stochastic volatility and derive an equilibrium strategy. In their model, the authors consider log returns of the agent's investments, which is in contrast to the settings of most mean-variance works in the literature wherein gross returns are used. With this unconventional setup, the authors derive the equilibrium strategy in a stochastic volatility market in closed form. Li et al. [35] consider a mean-variance portfolio selection problem with stochastic volatility in the penalty formulation. The authors, however, heuristically derive a set of equations only, and do not prove the existence of the solution to these equations or the existence of an equilibrium portfolio strategy. In a general semi-martingale setting, Czichowsky [18] studies time-consistent open-loop investment strategies under mean-variance criteria in both discrete-time and continuous-time settings, but the notion of open-loop investment strategies therein differs from ours; see further discussions in Footnote 4 in the following.

by assuming that the asset returns in different periods are independent of each other, that the target for the expected terminal wealth is the same for the agent at different times, and that the expected terminal wealth of the agent has to be exactly equal to the target, i.e., the constraint on the expected terminal wealth is an equality constraint rather than an inequality one. Our model differs from theirs in three main respects. First, we consider portfolio selection in continuous time while they use a discrete-time setting. The notion of equilibrium strategies in a discrete-time, finite-period setting can be defined naturally by backward induction (see e.g., Laibson [33] and O'Donoghue and Rabin [39]), while the definition of equilibrium strategies in continuous time is much more involved (see e.g., early discussion by Pollak [43] and Peleg and Yaari [42]). Second, assuming the target for the expected terminal wealth to be a constant, as in Cui, Li and Li [14], might not be reasonable because investors tend to adjust, e.g., increase, the target when their current wealth becomes higher. In contrast, we consider three types of target for the expected terminal wealth, which are flexible enough to accommodate a variety of patterns of investors' target. Third, modeling the target for the expected terminal wealth as an inequality constraint, as in the present paper, is more reasonable than modeling it as an equality constraint as in Cui, Li and Li [14], because a higher return without additional risk is always preferred.[2] Technically, the inequality constraint is more difficult to deal with than the equality one because with the former, one has to take into account the possibility that this inequality constraint is not binding.

Dai et al. [19] consider the mean-variance analysis with the penalty formulation, but the carrier of the mean and variance in the analysis is the log return rate of the agent's investment instead of the gross return rate (equivalently, terminal wealth) that was used by Markowitz [37] and by most works in the literature. The authors discussed at the end of Section 2.1 therein the constraint formulation of their problem and showed in Theorem 2.1 that if one sets the target for the expected log return in the constraint formulation to be the expectation of the log return of the equilibrium strategy in the penalty formulation, the equilibrium strategy in the constraint formulation is the same as the one in the penalty formulation. The present paper differs from theirs mainly in two respects. First, as consistent with the literature, we consider gross return rates, whereas Dai et al. [19] consider log returns. Second, Dai et al. [19] did not solve the equilibrium strategy in the constraint formulation. In addition, they can solve the equilibrium strategy in the penalty formulation only when the risk aversion parameter Γ_t is a constant. As a result, the only case in which they obtain the equilibrium strategy in the constraint formulation is when the target for the expected log return is the one induced by a constant risk aversion parameter in the penalty formulation. By contrast, we solve the equilibrium strategy in the constraint formulation directly for three types of target for the expected terminal wealth.

Finally, there are also some studies of mean-variance problems without using the notion of equilibrium strategies. For example, Pedersen and Peskir [40, 41] consider dynamically optimal strategies for mean-variance problems in the Black-Scholes market. These strategies are essentially naive strategies taken by agents who are never aware of the time inconsistency of their plans and thus at each time implement their pre-committed strategies set up at that time for a small time period only.[3] Chen [13] also study naive strategies for mean-variance portfolio selection. Czichowsky

[2] This also marks the difference of our model from the endogenous habit formation mean-variance model in Kryger and Steffensen [31] and Kryger et al. [32]. In that model, the authors consider an objective that involves a term $\mathbb{E}_t[(X(T)-K_t)^2]$, where K_t depends on the agent's wealth at time t. In our model, we consider the minimization of variance $\mathrm{var}_t(X(T)) = \mathbb{E}_t\big[(X(T)-\mathbb{E}_t[X(T)])^2\big]$ subject to a constraint that $\mathbb{E}_t[X(T)] \geq L_t$. Because this constraint is an inequality rather than an equality, we cannot write the objective function in our model as $\mathbb{E}_t\big[(X(T)-L_t)^2\big]$. Thus, our model is different from the endogenous habit formation mean-variance model.

[3] In Pedersen and Peskir [40], a stopping strategy is dynamically optimal if and only if its stopping region is contained in the stopping region of the naive stopping strategy; see Definition 2 therein. In Pedersen and Peskir [41], an investment strategy is dynamically optimal if and only if it is a naive strategy as long as the pre-committed strategy is unique at each time and state; see Definition 2 therein.

[18] and Hu et al. [27, 28] study equilibrium strategies for mean-variance problems under the penalty formulation, but their notion of equilibrium strategies is different from ours and from most works of time inconsistency in the literature.[4] Cui et al. [15], Trybuła and Zawisza [48], Černỳ [12], and Strub and Li [47] exploit the non-monotonicity of the mean-variance preferences and show that in some market settings, one can extract free cash flows without reducing the mean-variance value.

The remainder of the paper is organized as follows: In Section 2, we propose the dynamic mean-variance portfolio selection model, and define the notion of equilibrium strategies. When the agent's target for the expected terminal wealth takes three reasonable forms, we solve the equilibrium strategies explicitly. In Section 3, we compare the equilibrium strategy to the pre-committed strategy and to the equilibrium strategy in the penalty formulation. Section 4 concludes. The Appendix contains all proofs (Appendix C), a summary of some results about pre-committed strategies in the literature that are relevant to the present paper (Appendix A), some useful lemmas (Appendix B), generalization of our model to account for additional portfolio constraints (Appendix D), and discussion of two slightly different notions of equilibrium strategies (Appendix E).

2. Mean-Variance Portfolio Selection with Dynamic Targets

In this section, we formulate the mean-variance problem with a general target for expected terminal wealth and derive equilibrium strategies in closed form when the target takes three reasonable, flexible forms.

2.1. Market

Consider a market that consists of a risk-free asset and m risky stocks, and these assets are traded continuously in time. Denote the price of the risk-free asset and the price of stock i at time t as $S_0(t)$ and $S_i(t)$, respectively, $i=1,\dots,m$, and assume

$$dS_0(t)=S_0(t)r(t)dt,\quad dS_i(t)=S_i(t)\big[\big(b_i(t)+r(t)\big)dt+\sum_{j=1}^{d}\sigma_{ij}(t)dW_j(t)\big],\ i=1,\dots,m,$$

where $W(t):=\big(W_1(t),...,W_d(t)\big)^\top$, $t\geq 0$ is a standard, d-dimensional Brownian motion. Here and hereafter, denote the transpose of any matrix A as $A^\top$ and the Euclidean norm of a vector a as $\|a\|$. The risk-free rate $r(t)$, mean excess return rate vector $b(t):=\big(b_1(t),\dots,b_m(t)\big)^\top$, and volatility matrix $\sigma(t):=\big(\sigma_{i,j}(t)\big)$ satisfy the following assumption that will be in force throughout the paper.

Assumption 1. *$r(t)$, $b(t)$, and $\sigma(t)$, $t\geq 0$ are deterministic functions of t, right-continuous and bounded in t. Moreover, the following two non-degeneracy conditions hold: (i) $b(t)\neq 0$ for any $t\geq 0$ and (ii) there exists $\delta>0$ such that $\sigma(t)\sigma(t)^\top-\delta I$ is positive semi-definite for all $t\geq 0$, where I stands for the n-dimensional identity matrix.*

By Assumption 1, we have

$$\rho(t):=b(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}b(t)>0,\quad t\geq 0. \tag{2.1}$$

Note that a modification of $r(t)$, $b(t)$, and $\sigma(t)$ in a zero-measure set of t does not change the dynamics of the asset prices. Thus, assuming these functions to be right-continuous is equivalent to assuming them to have right-continuous modifications. In the rest of the paper, by assuming a

[4] They consider open-loop controls so that when the agent takes an alternative portfolio today, which changes her wealth tomorrow, e.g., from \$100 to \$200, the dollar amount that she invests in the stock tomorrow does not change. In contrast, in our notion, the agent's self tomorrow will follow a given feedback portfolio strategy, which is a mapping from time and wealth levels to dollar amounts invested in the stock. As a result, when the agent's wealth tomorrow changes due to a different portfolio taken today, her dollar amount invested in the stock tomorrow also changes.

function of t to be right-continuous, we actually assume the function to have a right-continuous modification, and we always refer to this modification.

Suppose at each time t an agent invests $\pi_i(t)$ dollars in stock i, $i=1,\dots,m$ and the remaining of her wealth in the risk-free asset. Then, the dynamics of the agent's wealth, denoted as $X(t)$, follow

$$dX(t)=\big(r(t)X(t)+\pi(t)^\top b(t)\big)dt+\pi(t)^\top\sigma(t)dW(t),\quad t\geq 0, \tag{2.2}$$

where $\pi(t):=\big(\pi_1(t),\dots,\pi_m(t)\big)^\top$ is referred to as the agent's *portfolio*.

2.2. Dynamic Mean-Variance Portfolio Selection We consider an agent who has a planning horizon with end date T. The agent sets a target for the expectation of her wealth at terminal time T, denoted as $X(T)$, to achieve and then minimizes the variance of $X(T)$. The agent may face some portfolio constraints, such as the no-shorting constraint. We model these portfolio constraints by $Q\pi\geq 0$, where π stands for the agent's dollar amount invested in risky stocks and Q is an n-by-m matrix. Assume that at each time t before T, the agent's target for expected terminal wealth is $L(t,x)$ for some *deterministic* function L, where x refers to the agent's wealth at t.

Due to time-inconsistency, in general there does not exist a strategy that is optimal for the agent at every time. We follow the classical literature on time-inconsistent problems, dating back to Strotz [46] and recently revisited by Ekeland and Lazrak [21, 22, 23] and Björk and Murgoci [10], to consider *equilibrium strategies*: The agent's selves at different times are viewed as different players in a sequential game. To this end, we assume that the agent considers feedback strategies only; i.e., the agent's portfolio $\pi(t)=\boldsymbol{\pi}(t,X(t))$, $t\in[0,T)$ for some deterministic, measurable function $\boldsymbol{\pi}(t,x)$ on $[0,T)\times\mathbb{R}$ such that the wealth equation (2.2) has a unique strong solution with $\mathbb{E}[X(T)^2]<+\infty$. Note that given a feedback strategy, the agent's wealth process is Markovian, so the distribution of $X(T)$ conditional on the information at time t depends on the agent's wealth at time t only.

The agent's wealth level x_0 at the initial time 0 is given and assumed to be positive. We can formulate the agent's decision problem at each time $t\in[0,T)$ formally as follows:

$$\begin{cases}\min\limits_{\boldsymbol{\pi}} & \mathrm{var}_t(X^{\boldsymbol{\pi}}(T))\\ \text{subject to} & dX^{\boldsymbol{\pi}}(s)=\big[r(s)X^{\boldsymbol{\pi}}(s)+b(s)^\top\boldsymbol{\pi}(s,X^{\boldsymbol{\pi}}(s))\big]ds\\ & \qquad +\boldsymbol{\pi}(s,X^{\boldsymbol{\pi}}(s))^\top\sigma(s)dW(s),\ s\in[t,T),\ X^{\boldsymbol{\pi}}(t)=x_t,\\ & \mathbb{E}_t[X^{\boldsymbol{\pi}}(T)]\geq L(t,x_t),\\ & Q\boldsymbol{\pi}(s,X^{\boldsymbol{\pi}}(s))\geq 0,\ s\in[t,T),\end{cases} \tag{2.3}$$

where $X^{\boldsymbol{\pi}}$ denotes the wealth process generated by strategy $\boldsymbol{\pi}$ and $\mathrm{var}_t(X^{\boldsymbol{\pi}}(T))$, the variance of $X^{\boldsymbol{\pi}}(T)$, measures the risk of the investment strategy $\boldsymbol{\pi}$. Here and hereafter, we denote by $\mathbb{E}_t[X]$ and $\mathrm{var}_t(X)$ the expectation and variance, respectively, of a random payoff X conditional on the information at time t; moreover, when time 0 is referred to, we simply drop the subscript. Note that besides the dynamics of the agent's wealth, there are two constraints in (2.3), which represent the expected terminal wealth target and portfolio constraints, respectively. A feedback strategy $\boldsymbol{\pi}$ is *feasible* if (i) the resulting wealth equation has a unique, strong solution $X^{\boldsymbol{\pi}}$, (ii) $\mathbb{E}\left[\int_0^T\big(\boldsymbol{\pi}(s,X^{\boldsymbol{\pi}}(s))\big)^2ds\right]<+\infty$, and (iii) all constraints in (2.3) are met for any $t\in[0,T)$. Note that condition (ii) ensures that $X^{\boldsymbol{\pi}}(T)$ is square integrable because b, σ, and r are bounded.

Let $\hat{\boldsymbol{\pi}}(t,x)$, $(t,x)\in[0,T)\times\mathbb{R}$ be a given feasible feedback strategy. For $t\in[0,T)$, $\epsilon\in(0,T-t)$, and $\pi\in\mathbb{R}^m$, denote

$$\boldsymbol{\pi}_{t,\epsilon,\pi}(s,y):=\begin{cases}\pi, & s\in[t,t+\epsilon),y\in\mathbb{R},\\ \hat{\boldsymbol{\pi}}(s,y), & s\notin[t,t+\epsilon),y\in\mathbb{R},\end{cases} \tag{2.4}$$

which represents a new strategy in which the time-t self of the agent, with wealth level x at that time, implements portfolio π for a time period of length ϵ and follows strategy $\hat{\boldsymbol{\pi}}$ afterwards. In addition, for each $t\in[0,T]$, we denote by $\mathbb{X}_t$ the set of *reachable states* at time t; mathematically, $\mathbb{X}_t$ is defined to be the union of the interior of the support and the atoms of the distribution of $X^{\hat{\boldsymbol{\pi}}}(t)$ conditional on the information at time 0. For each $t\in[0,T)$ and $x\in\mathbb{R}$, denote

$$\Pi_{t,x}^{\hat{\boldsymbol{\pi}}}:=\{\pi\in\mathbb{R}^m \mid \text{there exists } \bar{\epsilon}\in(0,T-t), \text{ such that } \boldsymbol{\pi}_{t,\epsilon,\pi} \text{ is feasible, } \forall\epsilon\in(0,\bar{\epsilon}]\},$$

which represents the dollar amount π the agent can choose to invest in the stocks in an infinitesimally small period $[t,t+\epsilon)$ such that the resulting strategy $\boldsymbol{\pi}_{t,\epsilon,\pi}$, in which the agent follows $\hat{\boldsymbol{\pi}}$ after time $t+\epsilon$, is feasible (in particular, meets the expected terminal wealth target).

DEFINITION 1. Let $\hat{\boldsymbol{\pi}}$ be a feasible feedback strategy. $\hat{\boldsymbol{\pi}}$ is an *equilibrium strategy* of the agent if for any $t\in[0,T)$, $x\in\mathbb{X}_t$, and $\pi\in\Pi_{t,x}^{\hat{\boldsymbol{\pi}}}$ with $\pi\neq\hat{\boldsymbol{\pi}}(t,x)$, there exists $\epsilon_0\in(0,T-t)$ such that

$$\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T))-\mathrm{var}_t(X^{\hat{\boldsymbol{\pi}}}(T))\geq 0,\quad \forall\epsilon\in(0,\epsilon_0]. \tag{2.5}$$

The above definition stipulates that $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy if and only if at any time t and in any reachable state x at that time, the variance of the terminal wealth becomes larger if the agent at that time switches to any *alternative* portfolio π that still leads to a feasible strategy (i.e., $\pi\in\Pi_{t,x}^{\hat{\boldsymbol{\pi}}}$ with $\pi\neq\hat{\boldsymbol{\pi}}(t,x)$), assuming that the agent can commit herself and thus implement the alternative portfolio only in an infinitesimally small time period. As such, the agent is unwilling to deviate from $\hat{\boldsymbol{\pi}}$ at any time, implying that $\hat{\boldsymbol{\pi}}$ is an equilibrium in a game with the players to be the agent's selves at different times.

It should be emphasized that due to the constraint arising from the expected terminal wealth target, we have to take into consideration the feasibility of the new strategy that is obtained by perturbing a given strategy in a small time period before examining whether the risk of this new strategy is lower, and this is taken into account in the above definition by introducing the set $\Pi_{t,x}^{\hat{\boldsymbol{\pi}}}$. This marks one of the main differences of our definition of equilibrium strategies from those in the literature, and ours is the first of its kind, because no such constraints have ever been considered in the literature on time-inconsistent problems in continuous-time settings. As mentioned in the introduction, the reason why we introduce constraints on expected terminal wealth is because the constraint formulation is advantageous over the penalty formulation in that the expected wealth target is easier to elicit than the risk aversion degree.

Note that a necessary condition for (2.5) to hold is following first-order condition

$$\liminf_{\epsilon\downarrow 0}\frac{\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T))-\mathrm{var}_t(X^{\hat{\boldsymbol{\pi}}}(T))}{\epsilon}\geq 0. \tag{2.6}$$

This type of first-order conditions is used by Björk and Murgoci [10] and all follow-up papers in the literature to study time-inconsistent problems, leading to so-called weak equilibrium: $\hat{\boldsymbol{\pi}}$ is a weak equilibrium strategy if for any $t\in[0,T)$, $x\in\mathbb{X}_t$, and $\pi\in\Pi_{t,x}^{\hat{\boldsymbol{\pi}}}$, (2.6) holds. As first noted by Björk et al. [9] and later exemplified by Huang and Zhou [29] and He and Jiang [25], the first-order condition (2.6) does not necessarily infer (2.5), the latter being a natural condition for equilibrium strategies from the game-theoretic point of view. He and Jiang [25] shows that the notion of equilibrium as employed in the present paper, referred to as regular equilibrium, improves the classical notion of weak equilibrium in that the former implies the latter and that the agent may find taking a very different strategy from a certain weak equilibrium is better off. Thus, we chose to follow He and Jiang [25] to use the notion of regular equilibrium in the present paper. Note, however, that the results in the present paper remain the same if we use the notion of weak equilibrium instead, and this is the case due to the particular setting of our model and might not

hold in general. See Huang and Zhou [29] and He and Jiang [25] for detailed discussions of how to define equilibrium strategies in continuous-time time-inconsistent problems, and see also Bayraktar et al. [5] for a related study of optimal stopping problems. In Appendix E we further discuss this issue and show how it affects the results in the present paper.

2.3. Three Choices of Expected Wealth Targets It is reasonable to assume that $L(t,x) \geq xe^{\int_t^T r(s)ds}$ for any $t \in [0,T)$ and $x \in \mathbb{R}$. Then, the following proposition shows that when approaching the terminal time T, we have to set the target to be the wealth level at that time.

PROPOSITION 1. *Suppose that $L(t,x) \geq xe^{\int_t^T r(s)ds}$ for any $t \in [0,T)$ and $x \in \mathbb{R}$, that $L(t,x)$ is continuous in $x \in \mathbb{R}$ for any $t \in [0,T)$, and that $L(T,x) := \lim_{t \uparrow T} L(t,x)$ exists for any $x \in \mathbb{R}$ and is continuous in x. If there exists a feasible strategy $\boldsymbol{\pi}$ with corresponding wealth process $X^{\boldsymbol{\pi}}$, then $L(T,x) = x$ for any x in the support of $X^{\boldsymbol{\pi}}(T)$ viewed at time 0.*

In the following, we consider three choices of the target L. The first choice is $L(t,x)/x = M(t)$ for some deterministic function M; i.e., the agent's target for the expected gross return rate at each time is independent of her wealth level at that time. Because of Proposition 1, we must have $M(T) = 1$. Then, by assuming $M(t)$ to be differentiable in t, we can write this target in the following form:

$$L(t,x) = xe^{\int_t^T [r(s)+\psi(s)]ds}, \quad t \in [0,T), x \in \mathbb{R} \tag{2.7}$$

for some deterministic function ψ.

The second choice of the target is $L(t,x) = xe^{\int_t^T r(s)ds} + h(t)$ for some deterministic function h that takes nonnegative values; i.e., the agent's target is the risk-free return of her current wealth plus a premium. Because of Proposition 1, we must have $h(T) = 0$. Then, by assuming $h(t)$ to be differentiable in t, we can write this target in the following form:

$$L(t,x) = \frac{1}{\gamma}(e^{\int_t^T \psi(s)ds} - 1) + xe^{\int_t^T r(s)ds}, \quad t \in [0,T), x \in \mathbb{R} \tag{2.8}$$

for some $\gamma > 0$ and some deterministic function ψ.

The third choice of the target is $L(t,x) = w(t)xe^{\int_t^T r(s)ds} + (1-w(t))\xi$ for some $\xi \geq x_0 e^{\int_0^T r(s)ds}$ and some deterministic function w that takes values in $[0,1]$; i.e., the agent's target at each time is a weighted average of the risk-free payoff of her current wealth and a pre-set aspiration level ξ. Because of Proposition 1, we must have $w(T) = 1$, so by assuming $w(t)$ to be differentiable in t, we can write this target in the following form:

$$L(t,x) = e^{-\int_t^T \psi(s)ds} xe^{\int_t^T r(s)ds} + \left(1 - e^{-\int_t^T \psi(s)ds}\right)\xi, \quad t \in [0,T), x \in \mathbb{R} \tag{2.9}$$

for some deterministic function ψ. When $\psi = \rho$ and the market is complete (i.e., $m = d$ and there no portfolio constraints), Theorem 2 in the Appendix A shows that the agent is time consistent, i.e., the pre-committed strategy of the agent at initial time 0 remains optimal for her future selves. This observation has also been made by Cui et al. [15] in Propositions 4.2 and 4.3 therein and by Karnam et al. [30] in Section 2.1 therein.

2.4. Equilibrium Strategy for Mean-Variance Problems

ASSUMPTION 2. *For any $t \in [0,T)$, the set $\{y \in \mathbb{R}^m \mid b(t)^\top y > 0, \quad Qy \geq 0\}$ is nonempty.*

Assumption 2 implies that for any $t\in[0,T)$ and any $N>0$, there exists $y\in\mathbb{R}^m$ such that $b(t)^\top y>N$ and $Qy\geq 0$. This assumption implies that the following optimization problem is feasible for any $t\in[0,T)$:

$$\begin{cases}\min\limits_{v\in\mathbb{R}^m} & \frac{1}{2}v^\top\sigma(t)\sigma(t)^\top v\\ \text{subject to } & b(t)^\top v\geq\psi(t),\ Qv\geq 0,\end{cases}\tag{2.10}$$

where ψ is the one in (2.7), (2.8), or (2.9), depending on which of the three targets is referred to. Consequently, problem (2.10) has a unique solution, denoted as $v^*(t)$. Assume the following:

ASSUMPTION 3. *ψ is nonnegative, right-continuous, and integrable in $t\in[0,T)$.*

When it becomes closer to the terminal time (i.e., when $T-t$ is smaller), the agent has a shorter time period to invest in the stock markets where she can gain positive expected excess returns, so it is reasonable to assume that the target for the expected terminal wealth becomes less aggressive in the sense that the target becomes closer to the risk-free payoff. The above suggests that fixing x, $L(t,x)/(xe^{\int_t^T r(s)ds})$ is decreasing with respect to t when L is given by (2.7), that $L(t,x)-xe^{\int_t^T r(s)ds}$ is decreasing with respect to t when L is given by (2.8), and that $1-e^{-\int_t^T\psi(s)ds}$ is decreasing with respect to t when L is given by (2.9); consequently, we need ψ to be nonnegative, as required in Assumption 3.

By Assumptions 1 and 3, we can conclude from the standard stability result for quadratic programming that v^* is right-continuous on $[0,T)$; see e.g., Daniel [20]. We also make the following assumption:

ASSUMPTION 4. $\int_0^T\|v^*(t)\|^2dt<+\infty$.

Assumption 4 is purely technical and holds for reasonable parameter setups; see Lemma 1 in the Appendix.

THEOREM 1. *Suppose that Assumptions 1-4 hold. Then, the following are true:*
(i) Suppose that L is given by (2.7). Then,

$$\hat{\boldsymbol{\pi}}(t,x)=xv^*(t),\quad t\in[0,T),x\in\mathbb{R}\tag{2.11}$$

is an equilibrium strategy. Moreover, any equilibrium strategy $\hat{\boldsymbol{\pi}}$ such that $\hat{\boldsymbol{\pi}}(t,x)/x$ is independent of $x\in\mathbb{R}$ and right-continuous, square-integrable in $t\in[0,T)$ must be given by (2.11).
(ii) Suppose that L is given by (2.8). Then,

$$\hat{\boldsymbol{\pi}}(t,x)=\frac{1}{\gamma}e^{\int_t^T(\psi(s)-r(s))ds}v^*(t),\quad t\in[0,T),x\in\mathbb{R}\tag{2.12}$$

is an equilibrium strategy. Moreover, any equilibrium strategy $\hat{\boldsymbol{\pi}}$ that is independent of $x\in\mathbb{R}$ and right-continuous, square-integrable in $t\in[0,T)$ must be given by (2.12).
(iii) Suppose that L is given by (2.9), and the aspiration level $\xi\geq x_0e^{\int_0^T r(s)ds}$. Denote $x^*(t)=\xi e^{-\int_t^T r(s)ds}$ for every $t\in[0,T)$. If the following holds:

$$e^{-\int_t^T\|\sigma(s)^\top v^*(s)\|^2ds}\geq 1-\frac{\psi(t)}{\rho(t)},\quad t\in[0,T),\tag{2.13}$$

then,

$$\hat{\boldsymbol{\pi}}(t,x)=v^*(t)\big(x^*(t)-x\big),\quad t\in[0,T),x\in\mathbb{R}\tag{2.14}$$

is an equilibrium strategy. On the other hand, any equilibrium strategy $\hat{\boldsymbol{\pi}}$ *such that* $\hat{\boldsymbol{\pi}}(t,x)/(x^*(t)-x)$ *is independent of* $x\in\mathbb{R}$ *and right-continuous, square-integrable in* $t\in[0,T)$ *must be given by* (2.14) *if there exists* $\delta_0>0$ *such that*

$$e^{-\int_t^T \|\sigma(s)^\top v^*(s)\|^2 ds} \geq 1-\frac{\psi(t)}{\rho(t)}+\delta_0, \quad t\in[0,T). \tag{2.15}$$

To find the equilibrium strategy, we only need to solve, for each fixed time t, the static optimization problem (2.10), which is a standard quadratic optimization problem and thus can be solved efficiently in general. Furthermore, without portfolio constraint, the equilibrium strategy is in closed form because we can solve $v^*(t)$ explicitly in this case.

Let us give an intuitive explanation for the reason why the equilibrium strategy can be solved from (2.10). To this end, suppose that $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy and that the expected wealth constraint is binding under this strategy at any time, i.e., that $\mathbb{E}_t[X^{\hat{\boldsymbol{\pi}}}(T)]=L(t,X^{\hat{\boldsymbol{\pi}}}(t))$ for all $t\in[0,T)$. Now, fix a time t with $X^{\hat{\boldsymbol{\pi}}}(t)=x$. The time-$t$ self of the agent can implement her strategy π in an infinitesimally small investment period $[t,t+\epsilon)$ and she believes that her future selves will follow the strategy $\hat{\boldsymbol{\pi}}$ after this period and thus her terminal wealth is governed by the strategy $\boldsymbol{\pi}_{t,\epsilon,\pi}$ as defined by (2.4). By the iterated expectation law and recalling the fact that $\hat{\boldsymbol{\pi}}$ binds the expected wealth constraint for future selves of the agent, we conclude

$$\mathbb{E}_t[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)]=\mathbb{E}_t[\mathbb{E}_{t+\epsilon}[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)]]=\mathbb{E}_t[L(t+\epsilon,X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon))].$$

The expected wealth target for the time-t self of the agent becomes

$$\mathbb{E}_t[L(t+\epsilon,X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon))]-L(t,x)\geq 0.$$

Dividing both sides of the above inequality by ϵ and sending ϵ to zero, we obtain the following constraint:

$$L_t(t,x)+L_x(t,x)(r(t)x+b(t)^\top\pi)+\frac{1}{2}L_{xx}(t,x)\|\sigma(t)^\top\pi\|^2\geq 0.$$

For each of the three targets (2.7), (2.8), and (2.9), we have $L_{xx}\equiv 0$, so the above constraint simply becomes

$$b(t)^\top\pi\geq K(t,x), \quad K(t,x):=-\left(\frac{L_t(t,x)}{L_x(t,x)}+r(t)x\right). \tag{2.16}$$

In other words, the agent wants the expected instantaneous wealth change in excess of the risk-free rate to be as high as $K(t,x)$.

On the other hand, when L is given by (2.7), recalling $\hat{\boldsymbol{\pi}}$ as defined in (2.11), one can see that $Z^{\boldsymbol{\pi}_{t,\epsilon,\pi}}_{t+\epsilon,T}:=X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)/X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$ is independent of the information at time $t+\epsilon$ because $\boldsymbol{\pi}_{t,\epsilon,\pi}=\hat{\boldsymbol{\pi}}$ in the period $[t+\epsilon,T]$. By the conditional variance formula, we have

$$\begin{aligned}
\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)) &= \mathbb{E}_t[\mathrm{var}_{t+\epsilon}(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T))]+\mathrm{var}_t(\mathbb{E}_{t+\epsilon}[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)])\\
&= \mathbb{E}_t\left[\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)\right)^2\right]\mathrm{var}_{t+\epsilon}\left(Z^{\boldsymbol{\pi}_{t,\epsilon,\pi}}_{t+\epsilon,T}\right)+\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)\right)\left(\mathbb{E}_{t+\epsilon}\left[Z^{\boldsymbol{\pi}_{t,\epsilon,\pi}}_{t+\epsilon,T}\right]\right)^2\\
&= \left\{\mathrm{var}_{t+\epsilon}\left(Z^{\boldsymbol{\pi}_{t,\epsilon,\pi}}_{t+\epsilon,T}\right)+\left(\mathbb{E}_{t+\epsilon}\left[Z^{\boldsymbol{\pi}_{t,\epsilon,\pi}}_{t+\epsilon,T}\right]\right)^2\right\}\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)\right)\\
&\quad +\mathrm{var}_{t+\epsilon}\left(Z^{\boldsymbol{\pi}_{t,\epsilon,\pi}}_{t+\epsilon,T}\right)\left(\mathbb{E}_t\left[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)\right]\right)^2.
\end{aligned}$$

Thus, to minimize $\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T))$, the time-$t$ self of the agent only needs (i) to minimize $\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)\right)$, which is approximately equal to $\pi^\top\sigma(t)\sigma(t)^\top\pi\epsilon$ from the wealth dynamics, and (ii) to minimize $\mathbb{E}_t\left[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)\right]$. The two minimization problems (i) and (ii) happen to align with

each other because a portfolio with low risk tends to have low growth as well. Therefore, the time-t self of the agent chooses her optimal portfolio by minimizing $\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)\right)$ subject to (2.16), i.e., by solving the following optimization problem:

$$\begin{cases} \min\limits_{\pi\in\mathbb{R}^m} & \frac{1}{2}\pi^\top\sigma(t)\sigma(t)^\top\pi \\ \text{subject to } & b(t)^\top\pi \geq K(t,x),\ Q\pi\geq 0. \end{cases} \tag{2.17}$$

Note that $K(t,x)=\psi(t)x$ when L is given by (2.7) and the agent's wealth remains positive under the equilibrium strategy, so the optimal solution to (2.17) is $\hat{\boldsymbol{\pi}}(t,x)$ as given by (2.11).

When L is given by (2.8), the equilibrium strategy $\hat{\boldsymbol{\pi}}$ is independent of wealth level. This is intuitive because at each time, the distance between the risk-free payoff of the current wealth and the target expected wealth level, $L(t,x)-xe^{\int_t^T r(s)ds}$, is independent of the current wealth level x. In addition, because $\hat{\boldsymbol{\pi}}$ as in (2.12) is a deterministic function of t, $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$ and $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)-X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$ are independent and thus

$$\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T))=\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon))+\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)-X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)).$$

Therefore, minimizing $\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T))$ is equivalent to minimizing $\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon))$, the latter is equal to $\pi^\top\sigma(t)\sigma(t)^\top\pi\epsilon$. Therefore, the time-$t$ self of the agent chooses his optimal portfolio by solving (2.17). Note that $K(t,x)=(1/\gamma)e^{\int_t^T(\psi(s)-r(s))ds}\psi(t)$ when L is given by (2.8), so the optimal solution to (2.17) is given by (2.12).

When L is given by (2.9), the equilibrium strategy is to invest in risky stocks an amount that is proportional to $x^*(t)-x$, the distance between the current wealth level x and the discounted value of the aspiration level ξ. This is intuitive because the closer the current wealth level is to the aspiration level, the less risk the agent should take. Consider the time-t self of the agent who can implement portfolio π in a small time period $[t,t+\epsilon)$. If the agent invests more in risky stocks, i.e., chooses a larger π, then both the variance and the mean of $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$ become larger. Note that when $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$ becomes larger, the agent wealth at time $t+\epsilon$ becomes closer to the aspiration level, so she will take less risk afterwards, leading to a lower variance of $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)-X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$. Thus, whether more investment in risky stocks by the time-t self of the agent leads to a larger variance of $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)$, which is contributed by both the variance of $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$ and the variance of $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)-X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$, depends on whether the resulting increment in the variance of $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$ dominates the resulting increment in the mean of $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)$. Under condition (2.13), the overall effect is an increase in the variance of $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)$, so minimizing $\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right)$ is equivalent to minimizing $\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(t+\epsilon)\right)$. Thus, the optimal strategy of the time-t self of the agent is the optimal solution to (2.17). Note that when L is given by (2.9), $K(t,x)=(x^*(t)-x)\psi(t)$. Thus, $\hat{\boldsymbol{\pi}}(t,x)$ is the optimal solution to (2.17) if and only if it is given by (2.14).

Note that condition (2.15) implies condition (2.13). When $\rho(t)\leq\psi(t), t\in[0,T)$, condition (2.15) holds automatically. In general, this condition holds when T is small. To see it more clearly, we consider the case in which there are no portfolio constraints and r, b, σ, and ψ are constants. In this case, we have $v^*=(\sigma\sigma^\top)^{-1}b\psi/\rho$, so condition (2.15) becomes

$$e^{-(\psi^2/\rho)T}>1-(\psi/\rho). \tag{2.18}$$

We can show that (2.18) holds for any $\psi\geq 0$ when $T<2.455/\rho$. Let us consider the case of a single stock and use the following parameter values taken from Björk et al. [11]: $r=0.04$, $\sigma=0.2$, $b=0.08$ and thus $\rho=b^2/\sigma^2=0.16$. Then, condition (2.18) holds for any $\psi\geq 0$ if $T<15.34$.

The following proposition shows that when there are no portfolio constraints, condition (2.13) is also necessary for $\hat{\boldsymbol{\pi}}$ as defined in (2.14) to be an equilibrium strategy.

PROPOSITION 2. *Suppose that Assumptions 1–4 hold, L is given by (2.9), and there are no portfolio constraints, i.e,. $Q=0$. Suppose that condition (2.13) does not hold. Then, $\hat{\boldsymbol{\pi}}$ as given by (2.14) is not an equilibrium strategy.*

2.5. Discussions

2.5.1. Connection with Merton's Portfolio If there are no portfolio constraints, the equilibrium strategy for the first target is given as $\hat{\boldsymbol{\pi}}(t,x) = x\frac{\psi(t)}{\rho(t)}(\sigma(t)\sigma(t)^\top)^{-1}b(t), t\in[0,T), x\in\mathbb{R}$. On the other hand, Merton's portfolio, which is the optimal portfolio in the expected utility framework with a power utility function, is given by $x(1/\delta)\big(\sigma(t)\sigma(t)^\top\big)^{-1}b(t), t\in[0,T), x\in\mathbb{R}$, where $\delta>0$ is the relative risk aversion degree (RRAD), assuming the market to be complete (i.e., $m=d$). Thus, if b, r, σ, and ψ are constants, then the portfolio rule in our model is the same as Merton's portfolio if and only if $\delta=\rho/\psi$.

Dai et al. [19] consider mean-variance analysis with the penalty formulation, with the carrier of the mean and variance to be the log return rate of the agent's investment. Assuming the risk aversion parameter $\bar{\gamma}$ in their model to be a positive constant and the market to be complete with constant b, r, and σ, they find that the equilibrium strategy in their model is the same as Merton's portfolio if and only if the RRAD $\delta=1+\bar{\gamma}$. Thus, the model in Dai et al. [19] can only mimic Merton's portfolio for investors with RRAD larger than one, while the equilibrium strategy (2.11) can mimic Merton's portfolio for any investor because ψ in our model can be any nonnegative number.

2.5.2. Life-Cycle Asset Allocation Professionals and fund managers usually provide the following advice for life-cycle asset allocation (such as investment for retirement): the longer the investment horizon, the larger percentage of wealth invested in equities; see for instance Ameriks and Zeldes [1]. Following this advice, many investment companies offer "life-cycle" funds in which the proportion of wealth of an investor in equities automatically reduces as the investor ages. This advice and the demand for "life-cycle" funds are not consistent with Merton's portfolio in which the proportion of wealth invested in equities should remain constant over time.

The equilibrium strategy (2.11) under the target (2.7) is consistent with the above investment advice. To see this, consider the case of no portfolio constraints and constant market parameters b, r, and σ. Then, we have $\hat{\boldsymbol{\pi}}(t,x)/x = \big((\sigma\sigma^\top)^{-1}b/\rho\big)\psi(t), t\in[0,T), x>0$. Therefore, when the agent's target for the instantaneous excess mean return rate $\psi(t)$ is decreasing in t, the proportion of wealth invested in the stocks, $\hat{\boldsymbol{\pi}}(t,x)/x$, is also decreasing in t, which is consistent with the advice of investing less when ages become larger.

3. Comparison of Different Strategies

In this section, we compare the equilibrium strategies under different expected terminal wealth targets, and compare them with other strategies in the literature.

3.1. Comparison of Equilibrium Strategies under Different Targets

In the following numerical example, we assume that there is only one stock, i.e., $m=d=1$, that there are no portfolio constraints, i.e., $Q=0$, and that r, b, and σ are constants. We use the parameter values taken by Björk et al. [11]: $r=0.04$, $\sigma=0.2$, $b=0.08$, and thus $\rho=b^2/\sigma^2=0.16$. We set $x_0=1$, $T=1$, and $\psi=0.06$ for all of the three expected terminal wealth targets. We set $\gamma=x_0e^{-rT}=0.961$ in the target (2.8) and $\xi=x_0e^{rT}(e^{\psi T}-e^{-\psi T})/(1-e^{-\psi T})=2.146$ in the target (2.9) so that all the three targets have the same value at the initial time 0. For fixed time $s=0.5T$, we plot the proportion of wealth invested in the stock as a function of the wealth level at time s under the equilibrium strategies when the agent set the dynamic targets (2.7), (2.8), and (2.9), and they are represented by dashed line, solid line, and triangular line, respectively, in the left panel of Figure 1. We plot the same in the right panel of Figure 1 for $s=0.9T$. We can observe that although the three dynamic targets have the same value at the initial time 0, the strategy taken by the agent under these three targets are completely different because these targets differ from each other after time 0. In particular, the proportion of wealth invested in the stock under the target (2.7) is

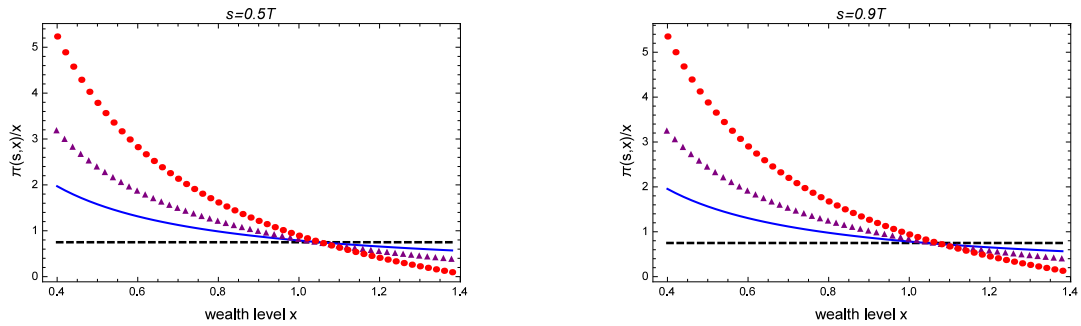

FIGURE 1. Comparison of the equilibrium strategies and the pre-committed strategy for problem (2.3) with the same expected terminal wealth targets at time 0. Set $m=d=1$, $Q=0$, $r=0.04$, $\sigma=0.2$, $b=0.08$, $\psi=0.06$, $x_0=1$, and $T=1$. Set $\gamma=x_0e^{-rT}=0.961$ in the target (2.8) and $\xi=x_0e^{rT}(e^{\psi T}-e^{-\psi T})/(1-e^{-\psi T})=2.146$ in the target (2.9) so that these two targets have the same value as target (2.7) at time 0. The proportion of wealth invested in the stock as a function of the wealth level at certain time s under the equilibrium strategies for targets (2.7), (2.8), are (2.9) are plotted in dashed lines, solid lines, and triangular lines, respectively, where s is set to be $0.5T$ and $0.9T$ in the left and right panels, respectively. The dotted lines plot the pre-committed strategy of the time-0 self of the agent with the same value of expected terminal wealth target at time 0.

independent of the wealth level because in this target, the expected return rate the agent wants to achieve is independent of the current wealth level. For the other two targets, the proportion of wealth invested in the stock is decreasing in the wealth level. This is because in those two targets, the expected return rate the agent wants to achieve, namely, $L(t,x)/x$, is decreasing with respect to the current wealth level.

We also compare our equilibrium strategies with the so-called pre-committed strategy. Fixing $t\in[0,T)$, a pre-committed strategy for the time-t self of the agent who faces problem (2.3) is a strategy $\boldsymbol{\pi}_{\mathrm{pc},t}(s,y)$, $s\in[t,T), y\in\mathbb{R}$ that minimizes the objective function in (2.3) with constraints therein. Assuming the market to be complete, i.e., assuming that $m=d$ and that there are no portfolio constraints, the pre-committed strategy has been obtained by Zhou and Li [50] without the no-bankruptcy constraint; see Theorem 2 in the Appendix for details.

Following the above example, we compute the pre-committed strategy of the time-0 self of the agent. Again we set the target at time 0 to be the same as above. The proportion of wealth invested in the stock at time s as a function of wealth at that time under the pre-committed strategy is plotted in dotted lines in Figure 1, with the left-panel representing the case $s=0.5T$ and the right-panel representing the case $s=0.9T$. As mentioned in Section 2.3, the pre-committed strategy happens to be the same as the equilibrium strategy under the target (2.9) with $\psi=\rho$. Thus, under the pre-committed strategy, at each time t the agent effectively sets the target for the terminal wealth to be $L(t,x)$ in (2.9) with $\psi=\rho$, and for this target $L(t,x)/x$ is decreasing in x. This explains why the proportion investment in the stock under the pre-committed strategy is decreasing with respect to the current wealth level.

3.2. Comparison with Equilibrium Strategies in the Penalty Formulation Basak and Chabakauri [3], Björk and Murgoci [10], and Björk et al. [11] consider equilibrium strategies for an agent whose preferences at each time t are represented by a linear combination of the mean and variance, evaluated at time t, of the agent's terminal wealth. Motivated by their work, we consider the following problem:

$$\begin{cases} \min\limits_{\boldsymbol{\pi}} & -\mathbb{E}_t[X^{\boldsymbol{\pi}}(T)]+\frac{\Gamma(t,x_t)}{2}\mathrm{var}_t(X^{\boldsymbol{\pi}}(T)) \\ \text{subject to} & dX^{\boldsymbol{\pi}}(s)=X^{\boldsymbol{\pi}}(s)\big[\big(r(s)+b(s)^\top\boldsymbol{\pi}(s,X^{\boldsymbol{\pi}}(s))\big)ds \\ & \qquad +\boldsymbol{\pi}(s,X^{\boldsymbol{\pi}}(s))^\top\sigma(s)dW(s)\big],\ s\in[t,T),\ X^{\boldsymbol{\pi}}(t)=x_t, \\ & Q\boldsymbol{\pi}(s,X^{\boldsymbol{\pi}}(s))\geq 0,\ s\in[t,T] \end{cases} \tag{3.1}$$

for some deterministic, positive function Γ that represents the agent's risk aversion degree.

It is well known that with a fixed time t and fixed wealth level x_t at that time, problem (2.3) and problem (3.1) are equivalent when one looks for pre-committed strategies: A strategy is a pre-committed strategy of the time-t self of the agent for problem (2.3) if and only if it is an optimal one for problem (3.1) with a carefully chosen risk aversion parameter $\Gamma(t, x_t)$. We want to study whether the equilibrium strategies for problem (2.3) and for problem (3.1) are the same if the pre-committed strategies are the same for these two problems at any time.

We rely on the existing literature to solve the equilibrium strategy for problem (3.1). However, the equilibrium strategy for problem (3.1) has been solved in the literature only for specific choices of Γ. Thus, we need to (i) solve the equilibrium strategy for problem (3.1) given an Γ used in the literature and (ii) establish a link between the expected wealth target $L(t, x_t)$ in problem (2.3) and risk aversion parameter $\Gamma(t, x_t)$ in problem (3.1).

Because the portfolio constraint is not present in the models in the literature, we set $Q = 0$ in the following discussion.

3.2.1. Comparison with Björk et al. [11] When there is only one stock, all market parameters are constants, and $\Gamma(t, x_t) = \gamma/x_t$ for some constant $\gamma > 0$, Björk et al. [11] derive a closed-form equilibrium strategy. Therefore, in the following we assume r, b, and σ to be constants, and set $m = d = 1$, $\Gamma(t, x_t) = \gamma/x_t$ for some constant $\gamma > 0$. The equilibrium strategy for (3.1) is given as in Theorem 4.6 of Björk et al. [11]. Moreover, the proof of Theorems 4.6 and 4.7 of Björk et al. [11] shows that this strategy is unique among all strategies of investing a continuous-in-time, wealth-independent proportion of the agent's wealth in the stock.

Next, we need to connect the risk aversion parameter $\Gamma(t, x_t) = \gamma/x_t$ in problem (3.1) to the expected wealth target $L(t, x_t)$ in problem (2.3). A natural way is to find L such that the pre-committed strategies for these two problems are the same at any time $t \in [0, T)$. Proposition 3-(i) in the Appendix shows that L takes the form (2.7), and ψ can be found in closed-form and proved to satisfy Assumptions 3 and 4. Thus, we can apply Theorem 1 to compute the equilibrium strategy with this L and compare this strategy to the equilibrium strategy for problem (3.1).

We use the parameter values taken by Björk et al. [11]: $r = 0.04$, $\sigma = 0.2$, $b = 0.08$, and thus $\rho = b^2/\sigma^2 = 0.16$. In addition, we set $T = 1$. We consider $\Gamma(t, x_t) = \gamma/x_t$ for four values of γ: 0.2, 1, 3, and 10, corresponding respectively to the upper-left, upper-right, bottom-left, and bottom-right panels of Figure 2. In each of the panels, the equilibrium strategy, i.e., the proportion of wealth invested in the stock as a function of time t, for problem (3.1) is represented by the dotted line. Then, we compute ψ from γ by Proposition 3-(i) in the Appendix and compute the equilibrium strategy (2.11). We use the solid line to represent the proportion of wealth invested in the stock driven by strategy (2.11). One can see that these two strategies are different: the strategy for problem (2.3) always invests more in the stock than the strategy for (3.1). Consequently, because the risk-premium of the stock, b, is positive, the expected return rate of the equilibrium strategy for problem (3.1) fails to reach the expected excess return target ψ.

Note that under the equilibrium strategy solved by Björk et al. [11], the wealth process is a geometric Brownian motion and thus is nonnegative. Thus, to facilitate the comparison of the equilibrium strategies for problem (3.1) and for problem (2.3), we can also consider to derive expected target $L(t, x_t)$ in problem (2.3) from $\Gamma(t, x_t) = \gamma/x_t$ in problem (3.1) by matching the pre-committed strategies for these two problems with the no-bankruptcy constraint. Such L takes the form (2.7) and ψ is derived in Proposition 4 in the Appendix. We can verify numerically that the resulting ψ satisfies Assumptions 3 and 4. In each panel of Figure 2, we plot the equilibrium strategy (2.11) with this ψ in the dashed line. We can see that this strategy also invests more in the stock than the equilibrium strategy for problem (3.1).

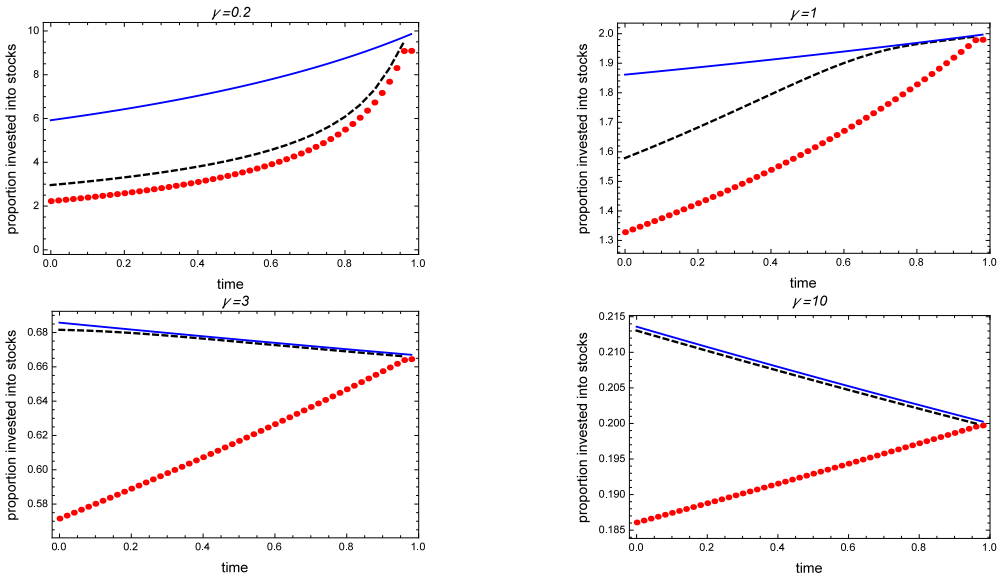

FIGURE 2. Equilibrium strategies for the mean-variance problem (3.1) and for the mean-variance problem (2.3). We set $T=1$, $r=0.04$, $\sigma=0.2$, $b=0.08$, and thus $\rho=b^2/\sigma^2=0.16$. We set $\Gamma(t,x_t)$ in (3.1) to be γ/x_t for four values of γ: 0.2, 1, 3, and 10, corresponding respectively to the upper-left, upper-right, bottom-left, and bottom-right panels. In each panel, the equilibrium strategy, i.e., the proportion of wealth invested in the stock as a function of time t, for problem (3.1) is represented by the dotted line. The equilibrium strategy for problem (2.3), with ψ computed from γ by Proposition 4 in the Appendix, is represented by the dashed line. The equilibrium strategy for problem (2.3), with ψ computed from γ by Proposition 3-(i) in the Appendix, is represented by the solid line.

3.2.2. Comparison with Basak and Chabakauri [3] In various market settings, Basak and Chabakauri [3] derive a closed-form equilibrium strategy for problem (3.1) with $\Gamma(t,x_t)=\gamma$ for some constant $\gamma>0$. In particular, when the mean excess return rates and volatility of the assets are deterministic, the equilibrium strategy is given as $\hat{\boldsymbol{\pi}}_{\rm BC}(t)=\frac{1}{\gamma}e^{-\int_t^T r(s)ds}(\sigma(t)\sigma(t)^\top)^{-1}b(t)$; see Proposition 6 of Basak and Chabakauri [3]. This strategy is also unique among all wealth-independent strategies that are right-continuous and square-integrable in time $t\in[0,T)$; see Footnote 8 of Basak and Chabakauri [3] and Section 7.1 in Björk and Murgoci [10]. We then compute $L(t,x_t)$ such that the pre-committed strategies for (3.1) and for (2.3) are the same. It turns out that L is given by (2.8) with $\psi=\rho$; see Proposition 3-(ii) in the Appendix. Consequently, Theorem 1-(ii) yields that the equilibrium strategy for (2.3) is

$$\hat{\boldsymbol{\pi}}(t)=\frac{1}{\gamma}e^{\int_t^T(\rho(s)-r(s))ds}(\sigma(t)\sigma(t)^\top)^{-1}b(t)=e^{\int_t^T\rho(s)ds}\hat{\boldsymbol{\pi}}_{\rm BC}(t).$$

Thus, the equilibrium strategy for (2.3) differs from that for (3.1): The former invests more in risky stocks than the latter.

3.2.3. Discussion about the Different Formulations The previous examples show that even if the pre-committed strategies for the two different formulations of mean-variance portfolio selection problems, namely, problem (2.3) and problem (3.1), are the same at any time, the equilibrium strategies for these two formulations can be different. A general question is whether we can find a pair of L in (2.3) and Γ in (3.1) such that the equilibrium strategies for these two problems are the same. We already showed that such a pair cannot be found by matching the pre-committed strategies for these two problems. One possible approach is to solve the equilibrium strategy for (3.1) with a given Γ, compute the expected terminal wealth associated with this strategy at each time, and use this expectation as the target $L(t,x_t)$ in problem (2.3). Let us emphasize that for this approach to work, we first have to define and solve the equilibrium strategies for the mean-variance

problem with the constraint formulation, and this, however, has never been done in the literature and is main goal of the present paper.

More importantly, to use the formulation (3.1) to guide portfolio selection, one needs to estimate the risk aversion parameter Γ, which is not directly accessible. Solving Γ from the expected wealth target, which is easy to access, by matching the pre-committed strategies in these two different formulations, leads to different equilibrium strategies for these two formulations. By contrast, as the target for the expected return is more elicitable, a more natural problem should be (2.3) rather than (3.1). Consequently, problem (2.3) is a more intuitive formulation than (3.1) to solve mean-variance portfolio selection problems.

4. Conclusion One classical approach to portfolio choice is mean-variance analysis: an agent minimizes the variance of her wealth at a given terminal time subject to a constraint that the expected terminal wealth exceeds a target. The mean-variance framework is appealing to many investors because the investment risk in the analysis, which is the variance of terminal wealth, is relatively easier to interpret compared to utility functions in the classical expected utility theory. This framework, however, leads to time inconsistency when applied to multi-period or continuous-time portfolio selection, in that the agent's optimal strategy of future investment set up today might no longer be optimal at certain future time. The so-called pre-committed strategies, which stand for agent's optimal strategies at the initial time, are largely studied in the literature, but such strategies are not implementable unless the agent has unrealistically strong self control.

In this paper, we proposed a notion of equilibrium strategies for the mean-variance analysis in a continuous-time setting. In such strategies, the agent's selves at different times act as different players in a sequential game, and the definition differs from those in the literature in that we had to take into account the constraint on the expected terminal wealth.

We focused on a market with multiple risky stocks and one risk-free asset whose mean return rates and volatility are deterministic, and portfolio constraints can be present in the market. We considered three types of dynamic targets for the expected terminal wealth: (i) the agent's current wealth multiplied by a deterministic gross return rate, (ii) the risk-free payoff of the agent's current wealth plus a premium, and (iii) and a weighted average of the risk-free payoff of the agent's current wealth and a pre-set aspiration level. We solved the equilibrium strategies for these three targets in closed form. With the first target, the strategy is to invest a wealth-independent proportion of current wealth in the stocks. With the second target, the strategy is to invest a wealth-independent dollar amount in the stocks. With the third target, the strategy is to invest in the stocks a dollar amount that is proportional to the distance from the agent's current wealth to the discounted value of the aspiration level.

Equilibrium strategies for mean-variance portfolio selection have been studied in the literature in a different formulation from ours. In all those studies, the agent minimizes a weighted average of the mean and variance of the terminal wealth, where the weight for the variance represents the agent's risk aversion degree. Moreover, the equilibrium strategies can be derived only for very specific forms of risk aversion degrees. Our formulation is more appealing because we take the target for expected terminal wealth, which is easier to understand and elicit than the risk aversion degree, as a model input directly. Moreover, we are able to solve equilibrium strategies for three reasonable and flexible forms of targets for expected terminal wealth, while it is unclear whether the specific forms of risk aversion degrees used in the literature are consistent with investors' risk attitude. We also showed that even if we choose the target for expected terminal wealth in our formulation and the risk aversion degree in the formulation used in the literature in a way that the resulting pre-committed strategies are the same, the equilibrium strategies in these two formulations are still different.

Acknowledgments. The present paper is a spin-out of an earlier version titled "Dynamic Mean-Risk Asset Allocation". The authors thank comments from seminar and conference participants at University of Toronto, Fudan University, Tianjin University, the 3rd PKU-NUS Annual International Conference on Quantitative Finance and Economics at Beijing, the 2018 INFORMS International Meeting at Taipei, and the 10th World Congress of the Bachelier Society at Dublin, 2018 INFORMS Annual Meeting at Phoenix, the Sixth Asian Quantitative Finance Conference at Guangzhou. The authors are especially grateful to Xunyu Zhou for his constructive comments on the paper. The authors acknowledge financial support from the General Research Fund of the Research Grants Council of Hong Kong SAR (Project No. 14200917).

Appendix A: Pre-committed Strategies

A.1. Pre-committed Strategies We consider the pre-committed strategy for problem (2.3) at time t when the market is complete. The strategy has been obtained by Zhou and Li [50] when bankruptcy is allowed, and by Bielecki et al. [8] when bankruptcy is not allowed. For readers' convenience, we reproduce their results in our setting.

THEOREM 2. *Suppose that Assumption 1 holds, that $m = d$, and that $Q = 0$. Fix time t and $x_t > 0$ and consider the portfolio selection (2.3) in which the agent chooses the optimal portfolio $\boldsymbol{\pi}^*_{\mathrm{pc},t}(s,y)$, $s \in [t,T)$, $y \in \mathbb{R}$ such that $\mathrm{var}_t(X^{\boldsymbol{\pi}}(T))$ is minimized subject to the constraint $\mathbb{E}_t[X^{\boldsymbol{\pi}}(T)] \geq L(t,x_t)$ and to the wealth equation. Assume $L(t,x_t) \geq x_t e^{\int_t^T r(s)ds}$. Then, the optimal pre-committed strategy is given by*

$$\begin{aligned}\boldsymbol{\pi}^*_{\mathrm{pc},t}(s,y) &= \left(\sigma(s)\sigma(s)^\top\right)^{-1} b(s) \\ &\quad \times \left[\left(\frac{L(t,x_t) - x_t e^{\int_t^T (r(z)-\rho(z))dz}}{1 - e^{-\int_t^T \rho(z)dz}}\right) e^{-\int_s^T r(z)dz} - y\right],\ s \in [t,T), y \in \mathbb{R}. \end{aligned} \tag{A.1}$$

Consequently, $\boldsymbol{\pi}^*_{\mathrm{pc},t}$ is independent of t and x_t if and only if L is given by (2.9) for some $\xi \geq x_0 e^{\int_0^T r(z)dz}$ and with $\psi = \rho$. Moreover,

$$\begin{aligned}&\mathbb{E}_s\left[X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(T) \mid X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(s) = y\right] = y e^{\int_s^T (r(z)-\rho(z))dz} \\ &+ \left(1 - e^{-\int_s^T \rho(z)dz}\right)\left(\frac{L(t,x_t) - x_t e^{\int_t^T (r(z)-\rho(z))dz}}{1 - e^{-\int_t^T \rho(z)dz}}\right),\ s \in [t,T), y \in \mathbb{R}. \end{aligned} \tag{A.2}$$

Proof. Fixing time $t \in [0,T)$, the pre-committed strategy for problem (3.1) has been solved by Zhou and Li [50], and it is given by

$$\boldsymbol{\pi}^*_{\mathrm{pc},t}(s,y) = \left(\sigma(s)\sigma(s)^\top\right)^{-1} b(s)\left[\left(\frac{e^{\int_t^T \rho(z)dz}}{\Gamma(t,x_t)} + x_t e^{\int_t^T r(z)dz}\right) e^{-\int_s^T r(z)dz} - y\right],\ s \in [t,T), y \in \mathbb{R};$$

see equation (5.12) therein. Then, a standard argument based on the Lagrange dual theory shows that $\boldsymbol{\pi}^*_{\mathrm{pc},t}$ is also the optimal pre-committed strategy for (2.3) with $L(t,x_t) = \mathbb{E}_t\left[X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(T)\right]$. It then follows from equations (6.5) and (6.6) in Zhou and Li [50] that

$$\begin{aligned}L(t,x_t) &= \mathbb{E}_t\left[X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(T)\right] \\ &= x_t e^{\int_t^T (r(z)-\rho(z))dz} + \left(1 - e^{-\int_t^T \rho(z)dz}\right)\left(\frac{1}{\Gamma(t,x_t)} e^{\int_t^T \rho(z)dz} + x_t e^{\int_t^T r(z)dz}\right). \end{aligned} \tag{A.3}$$

Thus, the optimal pre-committed strategy must be given by (A.1).

It is straightforward to see that $\boldsymbol{\pi}^*_{\mathrm{pc},t}$ is independent of t and x_t if and only if L is given by (2.9) for some constant ξ and with $\psi = \rho$. Because we assume $L(t,x_t) \geq x_t e^{\int_t^T r(z)dz}$ for any $t \in [0,T)$, we must have $\xi \geq x_0 e^{\int_0^T r(z)dz}$.

Finally, a direct calculation yields that (A.2). □

THEOREM 3. *Suppose that Assumption 1 holds, that $m = d$, and that $Q = 0$. Fix time t and $x_t > 0$ and consider the portfolio selection problem (2.3) in which the agent chooses the optimal portfolio $\boldsymbol{\pi}^*_{\mathrm{pc},t}(s,y)$, $s \in [t,T)$, $y \geq 0$ such that $\mathrm{var}_t(X^{\boldsymbol{\pi}}(T))$ is minimized subject to the constraint $\mathbb{E}_t[X^{\boldsymbol{\pi}}(T)] \geq L(t,x_t)$, to the no-bankruptcy constraint $X^{\boldsymbol{\pi}}(s) \geq 0, s \in [t,T]$, and to the wealth equation. Assume $L(t,x_t) > x_t e^{\int_t^T r(s)ds}$. Denote $z_t := L(t,x_t)$ and define $\lambda_t > 0$ and $\mu_t > 0$ by the following equations*

$$\begin{cases} \lambda_t \Phi\left(\frac{\ln(\lambda_t/\mu_t) + \int_t^T \left(r(s) - \frac{1}{2}\rho(s)\right)ds}{\sqrt{\int_t^T \rho(s)ds}}\right) - \mu_t e^{-\int_t^T (r(s)-\rho(s))ds} \\ \quad \times \Phi\left(\frac{\ln(\lambda_t/\mu_t) + \int_t^T \left(r(s) - \frac{3}{2}\rho(s)\right)ds}{\sqrt{\int_t^T \rho(s)ds}}\right) = x_t e^{\int_t^T r(s)ds}, \\ \lambda_t \Phi\left(\frac{\ln(\lambda_t/\mu_t) + \int_t^T \left(r(s) + \frac{1}{2}\rho(s)\right)ds}{\sqrt{\int_t^T \rho(s)ds}}\right) - \mu_t e^{-\int_t^T r(s)ds} \\ \quad \times \Phi\left(\frac{\ln(\lambda_t/\mu_t) + \int_t^T \left(r(s) - \frac{1}{2}\rho(s)\right)ds}{\sqrt{\int_t^T \rho(s)ds}}\right) = z_t. \end{cases} \tag{A.4}$$

For each $y > 0$, define $\eta_t(s,y)$ to be the solution to the following equation for η:

$$y = \lambda_t \Phi\left(-d_{t,-}(s,\eta)\right) e^{-\int_s^T r(z)dz} - \Phi\left(-d_{t,+}(s,\eta)\right) \mu_t e^{-\int_s^T (2r(z)-\rho(z))dz} \eta, \tag{A.5}$$

where

$$d_{t,+}(s,\eta) := \frac{\ln\eta + \ln(\mu_t/\lambda_t) + \int_s^T \left(-r(z) + \frac{3}{2}\rho(z)\right)dz}{\sqrt{\int_s^T \rho(z)dz}},$$
$$d_{t,-}(s,\eta) := d_{t,+}(s,\eta) - \sqrt{\int_s^T \rho(z)dz}, \quad s \in [t,T), \eta > 0.$$

Then,

$$\boldsymbol{\pi}^*_{\mathrm{pc},t}(s,y) = -(\sigma(s)\sigma(s)^\top)^{-1} b(s) \left[y - \lambda_t e^{-\int_s^T r(z)dz} \Phi\left(-d_{t,-}(s,\eta(s,y))\right)\right], \; s \in [t,T), y > 0. \tag{A.6}$$

Moreover,

$$\begin{aligned} \mathbb{E}_s\left[X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(T) \mid X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(s) = y\right] &= \lambda_t \Phi\left(-d_{t,\times}(s,\eta_t(s,y))\right) \\ &\quad - e^{-\int_s^T r(z)dz} \mu_t \eta_t(s,y) \Phi\left(-d_{t,-}(s,\eta_t(s,y))\right), \quad s \in [t,T), y > 0, \end{aligned} \tag{A.7}$$

where $d_{t,\times}(s,\eta) := d_{t,-}(s,\eta) - \sqrt{\int_s^T \rho(z)dz}$.

Proof. Define

$$\tilde{\eta}_t(s) := e^{-\frac{1}{2}\int_t^s \rho(z)dz - \int_t^s (\sigma(z)^{-1}b(z))^\top dW(z)}, \quad \eta_t(s) := e^{-\int_t^s r(z)dz} \tilde{\eta}_t(s), \quad s \in [t,T]. \tag{A.8}$$

It follows from Theorem 5.1, Theorem 7.1, and Proposition 7.1 in Bielecki et al. [8] that $\boldsymbol{\pi}^*_{\mathrm{pc},t}$ as given by (A.6) is the optimal strategy of the agent at time t when we replace the wealth target constraint $\mathbb{E}_t[X^{\boldsymbol{\pi}}(T)] \geq z_t$ with an equality constraint, i.e., with $\mathbb{E}_t[X^{\boldsymbol{\pi}}(T)] = z_t$. Moreover,

$$\eta_t(s) = \eta_t(s, X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(s)), \quad s \in [t,T].$$

Lemma 6.1 in Bielecki et al. [8] shows that $\boldsymbol{\pi}^*_{\mathrm{pc},t}$ is still optimal with the original wealth target constraint $\mathbb{E}_t\left[X^{\boldsymbol{\pi}}(T)\right] \geq z_t$.

Finally, we compute $\mathbb{E}_s\left[X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(T)\right]$. Theorem 4.1 of Bielecki et al. [8] shows that $X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(T) = \left(\lambda_t - \mu_t\eta_t(T)\right)^+$, so straightforward calculation yields that

$$\mathbb{E}_s\left[X^{\boldsymbol{\pi}^*_{\mathrm{pc},t}}(T)\right] = \lambda_t\Phi\left(-d_{t,\times}(s,\eta_t(s))\right) - e^{-\int_s^T r(z)dz}\mu_t\eta_t(s)\Phi\left(-d_{t,-}(s,\eta_t(s))\right),$$

i.e., (A.7) holds. □

A.2. A Connection between L and Γ

Proposition 3. *Suppose that Assumption 1 holds, that $m = d$, and that $Q = 0$. Consider the pre-committed strategies for problem (2.3) and problem (3.1).*

(i) Suppose $\Gamma(t,x_t) = \gamma/x_t$ for some constant $\gamma > 0$. Then, the pre-committed strategies for problem (2.3) and problem (3.1) are the same for any time $t \in [0,T)$ and any wealth level $x_t > 0$ if and only if L is given by (2.7) with

$$\psi(t) = \frac{\frac{1}{\gamma}\left[e^{\int_t^T \rho(s)ds}(\rho(t) - r(t)) + r(t)\right]}{\frac{1}{\gamma}\left(e^{\int_t^T \rho(s)ds} - 1\right) + e^{\int_t^T r(s)ds}}, \quad t \in [0,T). \tag{A.9}$$

Moreover, when r and ρ are constant, $\psi(t) > 0$ for any $t \in [0,T)$ if $T \leq 1/r$.

(ii) Suppose $\Gamma(t,x_t) = \gamma$ for some constant $\gamma > 0$. Then, the pre-committed strategies for problem (2.3) and problem (3.1) are the same for any time $t \in [0,T)$ and any wealth level $x_t \in \mathbb{R}$ if and only if L is given by (2.8) with $\psi = \rho$.

Proof. The proof of Theorem 2 yields that fixing time $t \in [0,T)$, the optimal pre-committed strategies for problem (2.3) and for problem (3.1) are the same if and only if (A.3) holds. Consequently, when $\Gamma(t,x_t) = \gamma/x_t$, we obtain $L(t,x_t) = x_t\left[\frac{1}{\gamma}\left(e^{\int_t^T \rho(s)ds} - 1\right) + e^{\int_t^T r(s)ds}\right]$. In other words, L is given by (2.7) with ψ in (A.9). Note that when r and ρ are constant, $\psi(t) > 0$ for any $t \in [0,T)$ if and only if $(1-\rho/r)e^{\rho T} - 1 < 0$, and a sufficient condition is that $rT \leq 1$. When $\Gamma(t,x_t) = \gamma$, it is straightforward to see that L is given by (2.8) with $\psi = \rho$. □

Proposition 4. *Suppose Assumption 1 holds and consider problems (2.3) and (3.1) subject to the no-bankruptcy constraint. Suppose $Q = 0$ and $m = d$. Suppose $\Gamma(t,x_t) = \gamma/x_t$ for some constant $\gamma > 0$. Then, the pre-committed strategies for these two problems are the same at any time $t \in [0,T)$ and any wealth level $x_t > 0$ if and only if L is given by (2.7), with $e^{\int_t^T [r(s)+\psi(s)]ds} = \tilde{M}_t - (1/\gamma)$ for any $t \in [0,T)$, where $\tilde{M}_t$, together with certain $\nu_t > 0$, is the solution to the following equation:*

$$\begin{cases} \left(\tilde{M}_t/\nu_t\right)\Phi\left(\frac{\ln(\tilde{M}_t/\nu_t) + \int_t^T\left(r(s) - \frac{1}{2}\rho(s)\right)ds}{\sqrt{\int_t^T \rho(s)ds}}\right) - e^{-\int_t^T (r(s)-\rho(s))ds} \\ \qquad \times\Phi\left(\frac{\ln(\tilde{M}_t/\nu_t) + \int_t^T\left(r(s) - \frac{3}{2}\rho(s)\right)ds}{\sqrt{\int_t^T \rho(s)ds}}\right) = e^{\int_t^T r(s)ds}/\nu_t, \\ \left(\tilde{M}_t/\nu_t\right)\Phi\left(\frac{\ln(\tilde{M}_t/\nu_t) + \int_t^T\left(r(s) + \frac{1}{2}\rho(s)\right)ds}{\sqrt{\int_t^T \rho(s)ds}}\right) - e^{-\int_t^T r(s)ds} \\ \qquad \times\Phi\left(\frac{\ln(\tilde{M}_t/\nu_t) + \int_t^T\left(r(s) - \frac{1}{2}\rho(s)\right)ds}{\sqrt{\int_t^T \rho(s)ds}}\right) = \left(\tilde{M}_t/\nu_t\right) - (1/\gamma)/\nu_t. \end{cases} \tag{A.10}$$

Proof. Fix $t \in [0,T)$ and $x_t > 0$. The analysis in Bielecki et al. [8], especially Theorem 4.1 therein, shows that fixing $\lambda_t > 0$ and letting μ_t be solved from the first equation of (A.4), the optimal solution to the following problem:

$$\begin{array}{ll} \min\limits_{\boldsymbol{\pi}} & \mathbb{E}_t\left[X^{\boldsymbol{\pi}}(T)^2\right] - 2\lambda_t \mathbb{E}_t\left[X^{\boldsymbol{\pi}}(T)\right] \\ \text{subject to} & dX^{\boldsymbol{\pi}}(s) = \left[r(s)X^{\boldsymbol{\pi}}(s) + b(s)^\top \boldsymbol{\pi}(s, X^{\boldsymbol{\pi}}(s))\right]ds \\ & \quad + \boldsymbol{\pi}(s, X^{\boldsymbol{\pi}}(s))^\top \sigma(s) dW(s),\ s \in [t,T),\ X^{\boldsymbol{\pi}}(t) = x_t, \\ & X^{\boldsymbol{\pi}}(s) \geq 0, s \in [t,T]. \end{array} \tag{A.11}$$

is given by (A.6). To highlight its dependence on λ_t, we denote this optimal solution by $\bar{\boldsymbol{\pi}}^*_{\mathrm{pc},\lambda_t,t}$. Then, $z_t := \mathbb{E}_t\left[X^{\bar{\boldsymbol{\pi}}^*_{\mathrm{pc},\lambda_t,t}}(T)\right]$ is given by the second equation of (A.4). Because $\bar{\boldsymbol{\pi}}^*_{\mathrm{pc},\lambda_t,t}$ is given by (A.6), it is the optimal solution to the following problem:

$$\begin{array}{ll} \min\limits_{\boldsymbol{\pi}} & \mathrm{var}_t\left(X^{\boldsymbol{\pi}}(T)\right) \\ \text{subject to} & dX^{\boldsymbol{\pi}}(s) = \left[r(s)X^{\boldsymbol{\pi}}(s) + b(s)^\top \boldsymbol{\pi}(s, X^{\boldsymbol{\pi}}(s))\right]ds \\ & \quad + \boldsymbol{\pi}(s, X^{\boldsymbol{\pi}}(s))^\top \sigma(s) dW(s),\ s \in [t,T),\ X^{\boldsymbol{\pi}}(t) = x_t, \\ & \mathbb{E}_t\left[X^{\boldsymbol{\pi}}(T)\right] \geq z_t,\ X^{\boldsymbol{\pi}}(s) \geq 0, s \in [t,T]. \end{array} \tag{A.12}$$

Now, consider the pre-committed strategy for problem (3.1) with the no-bankruptcy constraint and with $\Gamma(t,x_t) = \gamma/x_t$ for some constant $\gamma > 0$. Following the proof of Theorem 2.1 of Bielecki et al. [8], this strategy must exist and we denote it as $\boldsymbol{\pi}^*_{\mathrm{pc},\gamma,t}$. Then, Theorem 3.1 of Zhou and Li [50] shows that $\boldsymbol{\pi}^*_{\mathrm{pc},\gamma,t}$ must be the optimal solution to (A.11) with $\lambda_t = \frac{x_t}{\gamma} + \mathbb{E}_t\left[X^{\boldsymbol{\pi}^*_{\mathrm{pc},\gamma,t}}(T)\right]$, i.e., $\boldsymbol{\pi}^*_{\mathrm{pc},\gamma,t} = \bar{\boldsymbol{\pi}}^*_{\mathrm{pc},\lambda_t,t}$. Consequently, $\boldsymbol{\pi}^*_{\mathrm{pc},\gamma,t}$ is also optimal to problem (A.12). Moreover, $z_t = \mathbb{E}_t\left[X^{\boldsymbol{\pi}^*_{\mathrm{pc},\gamma,t}}(T)\right] = \mathbb{E}_t\left[X^{\bar{\boldsymbol{\pi}}^*_{\mathrm{pc},\lambda_t,t}}(T)\right]$ and $\lambda_t = \frac{x_t}{\gamma} + z_t$ satisfy (A.4). Consequently, denoting $M_t := z_t/x_t$, $\tilde{M}_t := \lambda_t/x_t = (1/\gamma) + M_t$, and $\nu_t := \mu_t/x_t$, we obtain (A.10). Note that we must have $z_t > x_t e^{\int_t^T r(s)ds}$, i.e., $\tilde{M}_t > e^{\int_t^T r(s)ds} + 1/\gamma$, because $\lambda_t > 0$. □

Appendix B: Some Useful Lemmas Given an n-by-m matrix Q and an n-dimensional column vector $q \leq 0$, consider the following:

Assumption 5. *For any $t \in [0,T)$, the set $\{y \in \mathbb{R}^m \mid b(t)^\top y > \psi(t), Qy \geq q\}$ is nonempty.*

With the above assumption, we have $\psi(t) < \bar{\beta}(t) := \sup_{u \in \mathbb{R}^m : Qu \geq q} b(t)^\top u, t \in [0,T)$.

Lemma 1. *Suppose Assumptions 1, 3 and 5 hold. For any $t \in [0,T)$, consider the optimal solution $u^*(t)$ to*

$$\begin{cases} \min\limits_{u \in \mathbb{R}^m} & \frac{1}{2} u^\top \sigma(t)\sigma(t)^\top u \\ \textit{subject to} & b(t)^\top u \geq \psi(t), \\ & Qu \geq q. \end{cases} \tag{B.1}$$

(i) Suppose $Q = q = 0$. Then, $\int_0^T \|u^*(t)\|^2 dt < +\infty$ if $\int_0^T \frac{\psi(t)^2}{\rho(t)} dt < +\infty$.

(ii) Suppose b and σ are constants. Then, $\int_0^T \|u^*(t)\|^2 dt < +\infty$ if $\int_0^T \psi(t)^2 dt < +\infty$.

Proof of Lemma 1 According to Assumption 1, we have $\|\sigma(t)^\top u^*(t)\|^2 \geq \delta \|u^*(t)\|^2$. Thus, we only need to prove that $\int_0^T \|\sigma(t)^\top u^*(t)\|^2 dt < +\infty$.

We consider (i) first, in which case $Q = q = 0$. Then, we have $u^*(t) = \frac{\psi(t)}{\rho(t)} (\sigma(t)\sigma(t)^\top)^{-1} b(t)$. Consequently,

$$\int_0^T \|\sigma(t)^\top u^*(t)\|^2 dt = \int_0^T \frac{\psi(t)^2}{\rho(t)} dt$$

Thus, $\int_0^T \|u^*(t)\|^2 dt < +\infty$ if $\int_0^T \frac{\psi(t)^2}{\rho(t)} dt < +\infty$.

Next, we consider (ii), in which case b and σ are constants. Then, $\bar{\beta}$ is also a constant. If $\bar{\beta} < +\infty$, then there exists $u \in \mathbb{R}^m$ with $Qu \geq q$ such that $b^\top u = \bar{\beta}$. Consequently, the following optimization problem

$$\begin{cases} \min\limits_{u \in \mathbb{R}^m} & \frac{1}{2} u^\top \sigma \sigma^\top u \\ \text{subject to } & b^\top u \geq \bar{\beta},\ Qu \geq q \end{cases}$$

is feasible and thus its optimal value $\bar{U} < +\infty$. From $\psi(t) < \bar{\beta}$, we have $\frac{1}{2}\|\sigma^\top u^*(t)\|^2 \leq \bar{U} < +\infty$. As a result, $\int_0^T \|\sigma^\top u^*(t)\|^2 dt \leq 2\bar{U}T < +\infty$.

If $\bar{\beta} = +\infty$, we first show that the set $\mathcal{P} := \{v \in \mathbb{R}^m \mid Qv \geq 0, b^\top v > 0\}$ is nonempty. For the sake of contradiction, suppose $\mathcal{P} = \emptyset$. Then, Farkas Lemma yields that there exists $v_0 \in \mathbb{R}^m$ such that $-Q^\top v_0 = b$ and $v_0 \geq 0$. For any u such that $Qu \geq q$, we have $v_0^\top (Qu - q) \geq 0$, which implies that $b^\top u \leq -v_0^\top q$. Consequently, we have $\bar{\beta} \leq -v_0^\top q < +\infty$, which is a contraction. Thus, $\mathcal{P} \neq \emptyset$. Now, fix $v \in \mathcal{P}$ and consider $u(t) := \frac{\psi(t)}{b^\top v} v$. Because $Qu(t) \geq 0 \geq q$ and $b^\top u(t) = \psi(t)$, $u(t)$ is feasible to problem (B.1). As a result, $\|\sigma^\top u^*(t)\|^2 \leq \|\sigma^\top u(t)\|^2 = \|\frac{\sigma^\top v}{b^\top v}\|^2 \psi(t)^2$. Thus $\int_0^T \|u^*(t)\|^2 dt < +\infty$ if $\int_0^T \psi(t)^2 dt < +\infty$. □

Lemma 2. *Suppose Σ is a positive definite, m-by-m matrix, $\varphi \in \mathbb{R}$, $b \in \mathbb{R}^m$ with $b \neq 0$, Q is an n-by-m matrix, and $q \in \mathbb{R}^n$ with $q \leq 0$. Define $\rho := b^\top \Sigma^{-1} b$.*

(i) Let $\hat{y}$ be the optimal solution to

$$\begin{cases} \min\limits_{y \in \mathbb{R}^m} & \frac{1}{2} y^\top \Sigma y + \varphi b^\top y \\ \textit{subject to } & Qy \geq q. \end{cases} \tag{B.2}$$

Then, we have $b^\top \hat{y} \leq \max\{-\varphi, 0\}\rho$.

(ii) Fix $d \geq 0$ and assume $\varphi \geq -d/\rho$ and $\{y \in \mathbb{R}^m \mid b^\top y \geq d, Qy \geq q\}$ to be nonempty. Then, the optimal solution to

$$\begin{array}{ll} \min\limits_{y \in \mathbb{R}^m} & \frac{1}{2} y^\top \Sigma y + \varphi b^\top y \\ \textit{subject to } & b^\top y \geq d,\ Qy \geq q. \end{array} \tag{B.3}$$

is the same as the optimal solution to

$$\begin{array}{ll} \min\limits_{y \in \mathbb{R}^m} & \frac{1}{2} y^\top \Sigma y \\ \textit{subject to } & b^\top y \geq d,\ Qy \geq q, \end{array} \tag{B.4}$$

and the optimal solution, denoted as $\hat{y}$, must satisfy $b^\top \hat{y} = d$.

Proof of Lemma 2 (i) Note that $y = 0$ is feasible to (B.2), so its optimal solution, $\hat{y}$, must exist. When $\varphi = 0$, it is obvious that $\hat{y} = 0$, so $b^\top \hat{y} = 0$. When $\varphi > 0$, because $y = 0$ is feasible, we conclude from the optimality of $\hat{y}$ that $\frac{1}{2}\hat{y}^\top \Sigma \hat{y} + \varphi b^\top \hat{y} \leq 0$. Consequently, we must have $b^\top \hat{y} \leq 0$.

In the following, we consider the case in which $\varphi < 0$. Recall the following Karush-Kuhn-Tucker (KKT) conditions:

$$\begin{cases} \Sigma \hat{y} = -\varphi b + Q^\top \nu, \\ \nu^\top (Q\hat{y} - q) = 0, \\ \nu \geq 0, Q\hat{y} \geq q. \end{cases} \tag{B.5}$$

Multiplying by $\nu^\top Q \Sigma^{-1}$ on both sides of the first equation of (B.5) and recalling the second equation of (B.5), we conclude

$$-\varphi \nu^\top Q \Sigma^{-1} b + \nu^\top Q \Sigma^{-1} Q^\top \nu = \nu^\top Q \hat{y} = \nu^\top q \leq 0,$$

where the inequality is the case because $\nu \geq 0$ and $q \leq 0$. Then, because $\nu^\top Q\Sigma^{-1}Q^\top \nu \geq 0$ and $\varphi < 0$, we conclude that $\nu^\top Q\Sigma^{-1}b \leq 0$. Next, multiplying by $b^\top\Sigma^{-1}$ on both sides of the first equation of (B.5), we conclude

$$b^\top \hat{y} = -\varphi\rho + b^\top\Sigma^{-1}Q^\top\nu \leq -\varphi\rho, \tag{B.6}$$

where the inequality is the case because $\nu^\top Q\Sigma^{-1}b \leq 0$.

(ii) Because $\{y \mid b^\top y \geq d, Qy \geq q\}$ is nonempty, (B.3) is feasible and thus admits a unique optimal solution, denoted as $\hat{y}$. When $\varphi \geq 0$, we claim that $b^\top\hat{y} = d$. For the sake of contradiction, suppose it is not the case. Then, $\hat{y} \neq 0$ because $b^\top\hat{y} > d \geq 0$. In addition, for sufficiently small $\lambda \in (0,1)$, $(1-\lambda)\hat{y}$ is feasible to (B.3) because $b^\top\hat{y} > d$ and $q \leq 0$. Moreover, the objective function value of $(1-\lambda)\hat{y}$ is strictly lower than that of $\hat{y}$ because $\hat{y} \neq 0$, $\varphi \geq 0$, and $b^\top\hat{y} > d \geq 0$. Thus, we arrive at a contradiction.

When $\varphi \in [-d/\rho, 0)$, recall the following KKT conditions:

$$\begin{cases} \Sigma\hat{y} = (\theta - \varphi)b + Q^\top\nu, \\ \theta[b^\top\hat{y} - d] = 0, \\ \nu^\top(Q\hat{y} - q) = 0, \\ \theta \geq 0, \nu \geq 0, b^\top u \geq d, Q\hat{y} \geq q. \end{cases} \tag{B.7}$$

If $\theta > 0$, it follows from the second equality of (B.7) that $b^\top\hat{y} = d$. If $\theta = 0$, the set of conditions in (B.7) contain those in (B.5), so we conclude that (B.6) holds. Together with $\varphi \geq -d/\rho$, we conclude $b^\top\hat{y} \leq d$. Recalling that $b^\top\hat{y} \geq d$ by the feasibility of $\hat{y}$, we immediately conclude that $b^\top\hat{y} = d$.

The above analysis shows that $b^\top\hat{y} = d$. As a result, $\hat{y}$ is optimal to (B.3) if and only if it is optimal to the following problem:

$$\begin{array}{ll} \min\limits_{y\in\mathbb{R}^m} & \frac{1}{2}y^\top\Sigma y + \varphi b^\top y \\ \text{subject to } b^\top y = d, & Qy \geq q. \end{array} \tag{B.8}$$

It is obvious that the optimal solution to (B.8) is invariant to φ. Consequently, the optimal solution to (B.3) is invariant to $\varphi \in [-d/\rho, +\infty)$. In particular, (B.3) and (B.4) have the same optimal solution. □

LEMMA 3. *Fix $t \geq 0$, $x \in \mathbb{R}$, and $T > t$. Consider a deterministic function $\mathbf{a}(s,y)$, $s \geq t, y \in \mathbb{R}$ such that $\lim_{s\downarrow t}\big(\mathbf{a}(s,x) - \mathbf{a}(t,x)\big) = 0$ and that there exists $\bar{\epsilon} \in (0, T-t)$, $L > 0$ such that $\|\mathbf{a}(s,y) - \mathbf{a}(s,y')\| \leq L|y - y'|$ for all $s \in [t, t+\bar{\epsilon}]$ and $y, y' \in \mathbb{R}$. Consider $\hat{\boldsymbol{\pi}}(s,y) := u(s)y + v(s)$, $s \geq t, y \in \mathbb{R}$ for two deterministic functions u and v that are square-integrable in $s \in [t,T]$. For each $\epsilon \in (0, \bar{\epsilon})$, define $\boldsymbol{\pi}_1^\epsilon(s,y) := \mathbf{a}(s,y)\mathbf{1}_{s\in[t,t+\epsilon)} + \hat{\boldsymbol{\pi}}(s,y)\mathbf{1}_{s\geq t+\epsilon}$ and $\boldsymbol{\pi}_2^\epsilon(s,y) := \mathbf{a}(t,x)\mathbf{1}_{s\in[t,t+\epsilon)} + \hat{\boldsymbol{\pi}}(s,y)\mathbf{1}_{s\geq t+\epsilon}$, $s \geq t$, $y \in \mathbb{R}$, and consider the following two processes for $s \geq t$:*

$$dX^{\boldsymbol{\pi}_i^\epsilon}(s) = r(s)X^{\boldsymbol{\pi}_i^\epsilon}(s)ds + b(s)^\top\boldsymbol{\pi}_i^\epsilon(s, X^{\boldsymbol{\pi}_i^\epsilon}(s))ds + \boldsymbol{\pi}_i^\epsilon(s, X^{\boldsymbol{\pi}_i^\epsilon}(s))^\top\sigma(s)dW(s),\ X^{\boldsymbol{\pi}_i^\epsilon}(t) = x, \quad i = 1,2,$$

where W is a d-dimensional standard Brownian motion and $r(s)$, $b(s)$, and $\sigma(s)$ are bounded, deterministic functions of s. Denote $Y^\epsilon(s) := X^{\boldsymbol{\pi}_1^\epsilon}(s) - X^{\boldsymbol{\pi}_2^\epsilon}(s)$, $s \geq t$. Then, $\lim_{\epsilon\downarrow 0}\mathbb{E}_t[Y^\epsilon(T)/\epsilon] = 0$, $\lim_{\epsilon\downarrow 0}\mathbb{E}_t\left[Y^\epsilon(T)X^{\boldsymbol{\pi}_i^\epsilon}(T)/\epsilon\right] = 0$, and $\lim_{\epsilon\downarrow 0}\mathbb{E}_t[(|Y^\epsilon(T)|/\sqrt{\epsilon})^p] = 0$ for any $p \geq 1$. Consequently, $\lim_{\epsilon\downarrow 0}\big(\mathrm{var}_t(Y^\epsilon(T))/\epsilon\big) = \lim_{\epsilon\downarrow 0}\big(\mathrm{cov}_t(Y^\epsilon(T), X^{\boldsymbol{\pi}_i^\epsilon}(T))/\epsilon\big) = 0$, where cov_t stands for the covariance of two random variables.

Proof of Lemma 3 We fix $K > 0$ to be a uniform bound for r, b, and σ. Note that $\mathbf{a}(s,x)$ is square integrable in $s \in [t, t+\epsilon]$ because $\lim_{s\downarrow t}\big(\mathbf{a}(s,x) - \mathbf{a}(t,x)\big) = 0$. Then, Theorem 6.3 in Chapter 1 of Yong and Zhou [49] shows that for any $p \geq 1$, there exists a constant $K_p > 0$ such that

$$\mathbb{E}_t\left[|X^{\boldsymbol{\pi}_i^\epsilon}(s) - x|^p\right] \leq K_p(1 + x^p)|s-t|^{p/2}, \quad \forall \epsilon \in (0, T-t), s \in [t, t+\epsilon], \quad i = 1,2. \tag{B.9}$$

Denote $\delta_{t,s} := |\mathbf{a}(s,x) - \mathbf{a}(t,x)|$, then $\lim_{s\downarrow t} \delta_{t,s} = 0$. Recall that

$$\begin{aligned} dY^\epsilon(s) &= r(s)Y^\epsilon(s)ds + b(s)^\top \big(\mathbf{a}(s, X^{\boldsymbol{\pi}_1^\epsilon}(s)) - \mathbf{a}(t,x)\big)ds \\ &\quad + \big(\mathbf{a}(s, X^{\boldsymbol{\pi}_1^\epsilon}(s)) - \mathbf{a}(t,x)\big)^\top \sigma(s)dW(s), \quad s \in [t, t+\epsilon],\ Y^\epsilon(t) = 0. \end{aligned} \tag{B.10}$$

Note that for any $p \geq 1$,

$$\begin{aligned} &\max_{z\in[t,t+\epsilon]} \left| \int_t^z b(s)^\top \big(\mathbf{a}(s, X^{\boldsymbol{\pi}_1^\epsilon}(s)) - \mathbf{a}(t,x)\big)ds \right|^p \leq \left[K \int_t^{t+\epsilon} \left(L \left| X^{\boldsymbol{\pi}_1^\epsilon}(s) - x \right| + \delta_{t,s} \right) ds \right]^p \\ &\leq K^p \epsilon^{p-1} 2^{p-1} \left(L^p \int_t^{t+\epsilon} \left| X^{\boldsymbol{\pi}_1^\epsilon}(s) - x \right|^p ds + \int_t^{t+\epsilon} \delta_{t,s}^p ds \right). \end{aligned} \tag{B.11}$$

where the second inequality is because $(|a|+|b|)^p \leq 2^{p-1}(|a|^p + |b|^p), a, b \in \mathbb{R}, p \geq 1$, and because the holder inequality. By the Burkholder-Davis-Gundy inequality, for any $p \geq 2$, there exists $C_p > 0$ such that

$$\begin{aligned} &\mathbb{E}_t \left[\max_{z\in[t,t+\epsilon]} \left| \int_t^z \big(\mathbf{a}(s, X^{\boldsymbol{\pi}_1^\epsilon}(s)) - \mathbf{a}(t,x)\big)^\top \sigma(s)dW(s) \right|^p \right] \\ &\leq C_p \mathbb{E}_t \left[\left(\int_t^{t+\epsilon} \left\| \big(\mathbf{a}(s, X^{\boldsymbol{\pi}_1^\epsilon}(s)) - \mathbf{a}(t,x)\big)^\top \sigma(s) \right\|^2 ds \right)^{p/2} \right] \\ &\leq K^p C_p \mathbb{E}_t \left[\left(\int_t^{t+\epsilon} \big(L|X^{\boldsymbol{\pi}_1^\epsilon}(s) - x| + \delta_{t,s}\big)^2 ds \right)^{p/2} \right] \\ &\leq K^p C_p \epsilon^{p/2-1} \mathbb{E}_t \left[\int_t^{t+\epsilon} \big(L|X^{\boldsymbol{\pi}_1^\epsilon}(s) - x| + \delta_{t,s}\big)^p ds \right] \\ &\leq K^p C_p \epsilon^{p/2-1} 2^{p-1} \left\{ \int_t^{t+\epsilon} \big(L^p \mathbb{E}_t[|X^{\boldsymbol{\pi}_1^\epsilon}(s) - x|^p] + \delta_{t,s}^p\big) ds \right\}. \end{aligned} \tag{B.12}$$

where the third inequality is from the holder inequality. Combining (B.9)-(B.12) and recalling Gronwall's inequality, we immediately conclude that for fixed $p \geq 2$, $\lim_{\epsilon\downarrow 0} \mathbb{E}_t[(|Y^\epsilon(t+\epsilon)|/\sqrt{\epsilon})^p] = 0$. This limit is still valid for any $p \in [1,2)$ because the L^p norm of a random variable is increasing in $p \in [1, +\infty)$.

By (B.10) and (B.11), we have, for any $z \in (t, t+\epsilon]$,

$$\begin{aligned} |\mathbb{E}_t[Y^\epsilon(z)]| &\leq \left| \mathbb{E}_t \left[\int_t^z r(s)Y^\epsilon(s)ds \right] \right| + \left| \mathbb{E}_t \left[\int_t^z b(s)^\top \big(\mathbf{a}(s, X^{\boldsymbol{\pi}_1^\epsilon}(s)) - \mathbf{a}(t,x)\big)ds \right] \right| \\ &\leq K \int_t^z |\mathbb{E}_t[Y^\epsilon(s)]|ds + K \left(L \int_t^{t+\epsilon} \left| X^{\boldsymbol{\pi}_1^\epsilon}(s) - x \right| ds + \int_t^{t+\epsilon} \delta_{t,s} ds \right). \end{aligned}$$

Then, (B.9) and Gronwall's inequality yield that $\lim_{\epsilon\downarrow 0} \mathbb{E}_t[Y^\epsilon(t+\epsilon)/\epsilon] = 0$.

Because $\hat{\boldsymbol{\pi}}(s,y) = u(s)y + v(s)$, we have

$$dY^\epsilon(s) = \big(r(s) + b(s)^\top u(s)\big)Y^\epsilon(s)ds + Y^\epsilon(s)u(s)^\top \sigma(s)dW(s), \quad s \in [t+\epsilon, T], \tag{B.13}$$

which implies

$$Y^\epsilon(T) = Y^\epsilon(t+\epsilon)Z_{t+\epsilon,T}, \quad Z_{t+\epsilon,T} := e^{\int_{t+\epsilon}^T \left(r(s) + b(s)^\top u(s) - \frac{1}{2}\|\sigma(s)^\top u(s)\|^2\right)ds + \int_{t+\epsilon}^T u(s)^\top \sigma(s)dW(s)}.$$

Because r, b, and σ are bounded and because u is square-integrable, for any $p \geq 1$, $\mathbb{E}_t[Z_{t+\epsilon,T}^p]$ is bounded uniformly in $\epsilon \in [0, \bar{\epsilon}]$. In addition, $Z_{t+\epsilon,T}$ is independent of $Y^\epsilon(t+\epsilon)$. Consequently,

$$\limsup_{\epsilon\downarrow 0} |\mathbb{E}_t[Y^\epsilon(T)/\epsilon]| = \limsup_{\epsilon\downarrow 0} \big(|\mathbb{E}_t[Y^\epsilon(t+\epsilon)/\epsilon]| \cdot \mathbb{E}_t[Z_{t+\epsilon,T}]\big) = 0.$$

Moreover, for any $p \geq 1$,

$$\limsup_{\epsilon \downarrow 0} \mathbb{E}_t[(|Y^\epsilon(T)|/\sqrt{\epsilon})^p] = \limsup_{\epsilon \downarrow 0} \left(\mathbb{E}_t[(|Y^\epsilon(t+\epsilon)|/\sqrt{\epsilon})^p] \cdot \mathbb{E}_t[Z^p_{t+\epsilon,T}]\right) = 0.$$

Next, define $X^{\hat{\boldsymbol{\pi}}}$ by the following dynamics:

$$dX^{\hat{\boldsymbol{\pi}}}(s) = r(s)X^{\hat{\boldsymbol{\pi}}}(s)ds + b(s)^\top \hat{\boldsymbol{\pi}}(s, X^{\hat{\boldsymbol{\pi}}}(s))ds + \hat{\boldsymbol{\pi}}(s, X^{\hat{\boldsymbol{\pi}}}(s))^\top \sigma(s) dW(s).$$

A standard argument of Feynman-Kac type shows that

$$\mathbb{E}_s\left[Z_{s,T} X^{\hat{\boldsymbol{\pi}}}(T)\right] = h(s)X^{\hat{\boldsymbol{\pi}}}(s) + g(s), \quad s \in [t,T],$$

where h and g are two deterministic functions satisfying the following equations:

$$\begin{cases} h'(s) + \left[2\big(r(s) + b(s)^\top u(s)\big) + u(s)^\top \sigma(s)\sigma(s)^\top u(s)\right] h(s) = 0, \quad h(T) = 1, \\ g'(s) + \big(r(s) + b(s)^\top v(s)\big) g(s) + \big(b(s)^\top v(s) + u(s)^\top \sigma(s)\sigma(s)^\top v(s)\big) h(s) = 0, \quad g(T) = 0. \end{cases}$$

Because $\hat{\boldsymbol{\pi}}(s,y) = \boldsymbol{\pi}^\epsilon_i(s,y)$ for $s \in [t+\epsilon, T)$, we have $\mathbb{E}_{t+\epsilon}\left[Z_{t+\epsilon,T} X^{\boldsymbol{\pi}^\epsilon_i}(T)\right] = h(t+\epsilon)X^{\boldsymbol{\pi}^\epsilon_i}(t+\epsilon) + g(t+\epsilon)$. As a result,

$$\begin{aligned} &\mathbb{E}_t\left[Y^\epsilon(T) X^{\boldsymbol{\pi}^\epsilon_i}(T)\right] = \mathbb{E}_t\left[\mathbb{E}_{t+\epsilon}\left[Y^\epsilon(t+\epsilon) Z_{t+\epsilon,T} X^{\boldsymbol{\pi}^\epsilon_i}(T)\right]\right] \\ &= h(t+\epsilon)\mathbb{E}_t\left[Y^\epsilon(t+\epsilon) X^{\boldsymbol{\pi}^\epsilon_i}(t+\epsilon)\right] + g(t+\epsilon)\mathbb{E}_t\left[Y^\epsilon(t+\epsilon)\right] \\ &= h(t+\epsilon)\mathbb{E}_t\left[Y^\epsilon(t+\epsilon)\big(X^{\boldsymbol{\pi}^\epsilon_i}(t+\epsilon) - x\big)\right] + \big(g(t+\epsilon) + x h(t+\epsilon)\big)\mathbb{E}_t\left[Y^\epsilon(t+\epsilon)\right]. \end{aligned} \tag{B.14}$$

We have

$$\left|\mathbb{E}_t\left[Y^\epsilon(t+\epsilon)\big(X^{\boldsymbol{\pi}^\epsilon_i}(t+\epsilon) - x\big)/\epsilon\right]\right| \leq \left(\mathbb{E}_t\left[\big(Y^\epsilon(t+\epsilon)/\sqrt{\epsilon}\big)^2\right]\right)^{1/2} \left(\mathbb{E}_t\left[\Big(\big(X^{\boldsymbol{\pi}^\epsilon_i}(t+\epsilon) - x\big)/\sqrt{\epsilon}\Big)^2\right]\right)^{1/2},$$

which implies $\lim_{\epsilon\downarrow 0} \mathbb{E}_t\left[Y^\epsilon(t+\epsilon)\big(X^{\boldsymbol{\pi}^\epsilon_i}(t+\epsilon) - x\big)/\epsilon\right] = 0$ due to the bound (B.9) and the limit $\lim_{\epsilon\downarrow 0} \mathbb{E}_t[(|Y^\epsilon(t+\epsilon)|/\sqrt{\epsilon})^2] = 0$ that we already proved. Consequently, because $\lim_{\epsilon\downarrow 0} \mathbb{E}_t[Y^\epsilon(t+\epsilon)/\epsilon] = 0$ and because $h(s)$ and $g(s)$ are continuous in $s \in [t,T]$, we conclude from (B.14) that $\lim_{\epsilon\downarrow 0} \mathbb{E}_t\left[Y^\epsilon(T) X^{\boldsymbol{\pi}^\epsilon_i}(T)/\epsilon\right] = 0$.

Finally, because $\lim_{\epsilon\downarrow 0} \mathbb{E}_t[Y^\epsilon(T)/\epsilon] = \lim_{\epsilon\downarrow 0} \mathbb{E}_t[(Y^\epsilon(T)/\sqrt{\epsilon})^2] = 0$, we immediately conclude that $\lim_{\epsilon\downarrow 0} \big(\mathrm{var}_t(Y^\epsilon(T))/\epsilon\big) = 0$. Because $\lim_{\epsilon\downarrow 0} \mathbb{E}_t\left[Y^\epsilon(T) X^{\boldsymbol{\pi}^\epsilon_i}(T)/\epsilon\right] = 0$, $\lim_{\epsilon\downarrow 0} \mathbb{E}_t[Y^\epsilon(T)/\epsilon] = 0$, and $\mathbb{E}_t\left[X^{\boldsymbol{\pi}^\epsilon_i}(T)\right]$ is uniformly bounded in ϵ, we conclude that $\lim_{\epsilon\downarrow 0} \big(\mathrm{cov}_t(Y^\epsilon(T), X^{\boldsymbol{\pi}^\epsilon_i}(T))/\epsilon\big) = 0$ □

Appendix C: Proofs

Proof of Proposition 1. Denote by $Z(t)$, $t \in [0,T]$ the right-continuous version of the martingale $\mathbb{E}_t[X^{\boldsymbol{\pi}}(T)]$, $t \in [0,T]$. By the martingale representation theorem, $Z(t)$ is continuous in t. Because $\boldsymbol{\pi}$ is a feasible strategy, we have $Z(t) = \mathbb{E}_t[X^{\boldsymbol{\pi}}(T)] \geq L(t, X^{\boldsymbol{\pi}}(t))$. Sending t to T and using the continuity of $Z(t)$ in t, we obtain $X^{\boldsymbol{\pi}}(T) = Z(T) \geq L(T, X^{\boldsymbol{\pi}}(T)) \geq X^{\boldsymbol{\pi}}(T)$. The above, together with continuity of $L(T,x)$ in x, implies that $L(T,x) = x$ for any x in the support of $X^{\boldsymbol{\pi}}(T)$. □

Proof of Theorem 1. (i) The case when L is given by (2.7).

We first consider any feasible strategy $\hat{\boldsymbol{\pi}}$ such that $\hat{\boldsymbol{\pi}}(t,x) = x\hat{\mathbf{u}}(t)$ for any $t \in [0,T)$ and $x \in \mathbb{R}$, where $\hat{\mathbf{u}}$ is right-continuous and square-integrable in $t \in [0,T)$, and derive a sufficient and necessary condition for this strategy to be an equilibrium one. For fixed $t \in (0,T]$, if $\hat{\mathbf{u}}(s) = 0$ for all $s \in [0,t)$, we have $X^{\hat{\boldsymbol{\pi}}}(t) = x_0 e^{\int_0^t r(s)ds} > 0$. If there exists $s \in [0,t)$, such that $\hat{\mathbf{u}}(s) \neq 0$, then from the right-continuity of $\hat{\mathbf{u}}$ and $x_0 > 0$, one can see that the set of reachable states $\mathbb{X}_t$ under strategy $\hat{\boldsymbol{\pi}}$ is given by $(0, +\infty)$. Consequently, for any $t \in [0,T)$, $\mathbb{X}_t$ is contained in $(0,+\infty)$ and has at least one positive element.

Now, we fix any $t\in[0,T)$ and $x\in\mathbb{X}_t$, then $x>0$. Consider any $\pi\in\mathbb{R}^m$, $\epsilon\in(0,T-t)$, and $\boldsymbol{\pi}_{t,\epsilon,\pi}$ as defined in (2.4). Define

$$\boldsymbol{\pi}'_{t,\epsilon,\pi}(s,y):=\begin{cases}(\pi/x)y, & s\in[t,t+\epsilon),y\in\mathbb{R},\\ \hat{\boldsymbol{\pi}}(s,y), & s\notin[t,t+\epsilon),y\in\mathbb{R}.\end{cases}\tag{C.1}$$

Note that $\boldsymbol{\pi}'_{t,\epsilon,\pi}(s,y)$ is proportional to y. Then, defining $\mathbf{u}_{t,\epsilon,\pi}(s)=\boldsymbol{\pi}'_{t,\epsilon,\pi}(s,y)/y, s\in[t,T)$, we have

$$X^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(T)=xe^{\int_t^T\left(r(s)+b(s)^\top\mathbf{u}_{t,\epsilon,\pi}(s)-\frac{1}{2}\|\sigma(s)^\top\mathbf{u}_{t,\epsilon,\pi}(s)\|^2\right)ds+\int_t^T\mathbf{u}_{t,\epsilon,\pi}(s)^\top\sigma(s)dW(s)}.\tag{C.2}$$

Straightforward calculation yields

$$\mathbb{E}_t[X^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(T)]=xe^{\int_t^T(r(s)+b(s)^\top\hat{\mathbf{u}}(s))ds}e^{\int_t^{t+\epsilon}(b(s)^\top\pi/x-b(s)^\top\hat{\mathbf{u}}(s))ds},\tag{C.3}$$

$$\begin{aligned}\mathrm{var}_t[X^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(T)]&=x^2e^{2\int_t^T(r(s)+b(s)^\top\hat{\mathbf{u}}(s))ds}e^{2\int_t^{t+\epsilon}(b(s)^\top\pi/x-b(s)^\top\hat{\mathbf{u}}(s))ds}\\&\quad\times\left[e^{\int_t^{t+\epsilon}\left[\|\sigma(s)^\top\pi/x\|^2-\|\sigma(s)^\top\hat{\mathbf{u}}(s)\|^2\right]ds}e^{\int_t^T\|\sigma(s)^\top\hat{\mathbf{u}}(s)\|^2ds}-1\right].\end{aligned}\tag{C.4}$$

With the help of Lemma 3, we conclude that

$$\lim_{\epsilon\downarrow0}\frac{\mathbb{E}_t\left[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right]-\mathbb{E}_t\left[X^{\hat{\boldsymbol{\pi}}}(T)\right]}{\epsilon}=\lim_{\epsilon\downarrow0}\frac{\mathbb{E}_t\left[X^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(T)\right]-\mathbb{E}_t\left[X^{\hat{\boldsymbol{\pi}}}(T)\right]}{\epsilon},\tag{C.5}$$

$$\lim_{\epsilon\downarrow0}\frac{\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right)-\mathrm{var}_t\left(X^{\hat{\boldsymbol{\pi}}}(T)\right)}{\epsilon}=\lim_{\epsilon\downarrow0}\frac{\mathrm{var}_t\left(X^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(T)\right)-\mathrm{var}_t\left(X^{\hat{\boldsymbol{\pi}}}(T)\right)}{\epsilon}.\tag{C.6}$$

Consequently, combining (C.5) with (C.3) and combining (C.6) with (C.4), we have

$$\lim_{\epsilon\downarrow0}\frac{\mathbb{E}_t\left[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right]-\mathbb{E}_t\left[X^{\hat{\boldsymbol{\pi}}}(T)\right]}{\epsilon}=xe^{\int_t^T(r(s)+b(s)^\top\hat{\mathbf{u}}(s))ds}\left(b(t)^\top\pi/x-b(t)^\top\hat{\mathbf{u}}(t)\right),\tag{C.7}$$

$$\begin{aligned}\lim_{\epsilon\downarrow0}\frac{\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right)-\mathrm{var}_t\left(X^{\hat{\boldsymbol{\pi}}}(T)\right)}{\epsilon}&=2x^2e^{2\int_t^T(r(s)+b(s)^\top\hat{\mathbf{u}}(s))ds}e^{\int_t^T\|\sigma(s)^\top\hat{\mathbf{u}}(s)\|^2ds}\\&\quad\times\left[\frac{1}{2}\left(\|\sigma(t)^\top\pi/x\|^2-\|\sigma(t)^\top\hat{\mathbf{u}}(t)\|^2\right)+\left(1-e^{-\int_t^T\|\sigma(s)^\top\hat{\mathbf{u}}(s)\|^2ds}\right)\left(b(t)^\top\pi/x-b(t)^\top\hat{\mathbf{u}}(t)\right)\right].\end{aligned}\tag{C.8}$$

Because $\mathbb{E}_t[X^{\hat{\boldsymbol{\pi}}}(T)]=xe^{\int_t^T\left(r(s)+b(s)^\top\hat{\mathbf{u}}(s)\right)ds}$ and because $x>0$, the feasibility of $\hat{\boldsymbol{\pi}}$ implies that

$$\int_t^T b(s)^\top\hat{\mathbf{u}}(s)ds\geq\int_t^T\psi(s)ds,\quad t\in[0,T).\tag{C.9}$$

As a result, one can see from (C.3), (C.5), and (C.7) that a sufficient condition for $\pi\in\Pi^{\hat{\boldsymbol{\pi}}}_{t,x}$ with $\pi\neq\hat{\boldsymbol{\pi}}(t,x)$ is the following:

$$\begin{cases}\pi\neq\hat{\boldsymbol{\pi}}(t,x),\ Q\pi\geq0, & \text{if } \int_t^T b(s)^\top\hat{\mathbf{u}}(s)ds>\int_t^T\psi(s)ds,\\ \pi\neq\hat{\boldsymbol{\pi}}(t,x),\ Q\pi\geq0,\ b(t)^\top\pi/x>b(t)^\top\hat{\mathbf{u}}(t), & \text{if } \int_t^T b(s)^\top\hat{\mathbf{u}}(s)ds=\int_t^T\psi(s)ds.\end{cases}\tag{C.10}$$

Similarly, a necessary condition for $\pi\in\Pi^{\hat{\boldsymbol{\pi}}}_{t,x}$ with $\pi\neq\hat{\boldsymbol{\pi}}(t,x)$ is the following:

$$\begin{cases}\pi\neq\hat{\boldsymbol{\pi}}(t,x),\ Q\pi\geq0, & \text{if } \int_t^T b(s)^\top\hat{\mathbf{u}}(s)ds>\int_t^T\psi(s)ds,\\ \pi\neq\hat{\boldsymbol{\pi}}(t,x),\ Q\pi\geq0,\ b(t)^\top\pi/x\geq b(t)^\top\hat{\mathbf{u}}(t), & \text{if } \int_t^T b(s)^\top\hat{\mathbf{u}}(s)ds=\int_t^T\psi(s)ds.\end{cases}\tag{C.11}$$

On the other hand, we observe from (C.8) that a sufficient condition for (2.5) to hold for some $\epsilon_0 \in (0, T-t)$ is that

$$\frac{1}{2}\left(\|\sigma(t)^\top \pi/x\|^2 - \|\sigma(t)^\top \hat{\mathbf{u}}(t)\|^2\right) + \varphi(t;\hat{\mathbf{u}})\left(b(t)^\top \pi/x - b(t)^\top \hat{\mathbf{u}}(t)\right) > 0, \tag{C.12}$$

where $\varphi(t;\hat{\mathbf{u}}) := 1 - e^{-\int_t^T \|\sigma(s)^\top \hat{\mathbf{u}}(s)\|^2 ds} \in [0,1)$. Similarly, a necessary condition for (2.5) to hold for some $\epsilon_0 \in (0, T-t)$ is that

$$\frac{1}{2}\left(\|\sigma(t)^\top \pi/x\|^2 - \|\sigma(t)^\top \hat{\mathbf{u}}(t)\|^2\right) + \varphi(t;\hat{\mathbf{u}})\left(b(t)^\top \pi/x - b(t)^\top \hat{\mathbf{u}}(t)\right) \geq 0. \tag{C.13}$$

We claim that a sufficient and necessary condition for $\hat{\boldsymbol{\pi}}$ to be an equilibrium strategy is the following: For each $t \in [0,T)$, $\hat{\mathbf{u}}(t)$ is the optimal solution to

$$\begin{cases} \min\limits_{u \in \mathbb{R}^m} & \frac{1}{2} u^\top \sigma(t)\sigma(t)^\top u + \varphi(t;\hat{\mathbf{u}})\left(b(t)^\top u - b(t)^\top \hat{\mathbf{u}}(t)\right), \\ \text{subject to} & \left(b(t)^\top u - b(t)^\top \hat{\mathbf{u}}(t)\right)\mathbf{1}_{\int_t^T b(s)^\top \hat{\mathbf{u}}(s) ds = \int_t^T \psi(s) ds} \geq 0, \\ & Qu \geq 0. \end{cases} \tag{C.14}$$

We prove the sufficiency first. Suppose that each $t \in [0,T)$, $\hat{\mathbf{u}}(t)$ is the optimal solution to (C.14). Because the objective function of problem (C.14) is strictly convex, its optimal solution must be unique. As a result, (C.12) holds for any $x > 0$ and π that satisfies (C.11). This implies that for any $t \in [0,T)$, $x \in \mathbb{X}_t$, and $\pi \in \Pi^{\hat{\boldsymbol{\pi}}}_{t,x}$ with $\pi \neq \hat{\boldsymbol{\pi}}(t,x)$, (2.5) holds for certain $\epsilon_0 \in (0,T)$, i.e., that $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy.

We then prove the necessity. Suppose $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy. Recall that for each fixed $t \in [0,T)$ and $x \in \mathbb{X}_t$, a necessary condition for (2.5) to hold for certain $\epsilon_0 \in (0,T)$ is that (C.13) holds and a sufficient condition for $\pi \in \Pi^{\hat{\boldsymbol{\pi}}}_{t,x}$ with $\pi \neq \hat{\boldsymbol{\pi}}(t,x)$ is that (C.10) holds. As a result, we have $f(t,\hat{\mathbf{u}}(t)) \leq f(t,\pi/x)$ for any π that satisfies (C.10), where $f(t,u)$ denotes the objective function in (C.14). Thus, to complete the proof of the necessity, we only need to show that for any $t \in [0,T)$ with $\int_t^T b(s)^\top \hat{\mathbf{u}}(s) ds = \int_t^T \psi(s) ds$ and any u_* such that $b(t)^\top u_* = b(t)^\top \hat{\mathbf{u}}(t)$, $f(t,\hat{\mathbf{u}}(t)) \leq f(t,u_*)$. By Assumption 2, there exists $v \in \mathbb{R}^m$ such that $b(t)^\top v > b(t)^\top \hat{\mathbf{u}}(t)$ and $Qv \geq 0$. Then, for any $\lambda \in (0,1)$, $u_\lambda := \lambda v + (1-\lambda)u_*$ satisfies $b(t)^\top u_\lambda > b(t)^\top \hat{\mathbf{u}}(t)$ (which implies $u_\lambda \neq \hat{\mathbf{u}}(t)$) and $Qu_\lambda \geq 0$. Because $f(t,\hat{\mathbf{u}}(t)) \leq f(t,\pi/x)$ for any π that satisfies (C.10), we conclude that $f(t,\hat{\mathbf{u}}(t)) \leq f(t,u_\lambda)$. Sending λ to 0, we immediately conclude $f(t,\hat{\mathbf{u}}(t)) \leq f(t,u_*)$.

Next, we prove that $\hat{\boldsymbol{\pi}}$ as defined in (2.11) is an equilibrium strategy. This follows immediately from the above sufficient and necessary condition, from the equality $\int_t^T b(s)^\top \hat{\mathbf{u}}(s) ds = \int_t^T \psi(s) ds$ for any $t \in [0,T)$, and from Lemma 2-(ii).

Finally, we prove that any equilibrium strategy $\hat{\boldsymbol{\pi}}$ such that $\hat{\boldsymbol{\pi}}(t,x)/x$ is independent of $x \in \mathbb{R}$ and right-continuous, square-integrable in $t \in [0,T)$ must be given by (2.11). Indeed, consider any such strategy $\hat{\boldsymbol{\pi}}$ and recall that $\hat{\mathbf{u}}(t) := \hat{\boldsymbol{\pi}}(t,1)$, $t \in [0,T)$. We first prove that $b(t)^\top \hat{\mathbf{u}}(t) = \psi(t)$ for all $t \in [0,T)$. For the sake of contradiction, suppose it is not the case. Then, by (C.9), there must exist $t \in [0,T)$ such that $b(t)^\top \hat{\mathbf{u}}(t) > \psi(t)$. We also have $\int_t^T b(s)^\top \hat{\mathbf{u}}(s) ds > \int_t^T \psi(s) ds$ for the same t, because $b(s)^\top \hat{\mathbf{u}}(s) > \psi(s)$ for $s \in [t, t+\epsilon_0]$ for some $\epsilon_0 \in (0, T-t)$ by the right-continuity of b, $\hat{\mathbf{u}}$, and ψ, and because (C.9) holds. Then, because $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy, $\hat{\mathbf{u}}(t)$ must be the optimal solution to (C.14) and thus the optimal solution to (B.2) with $\Sigma = \sigma(t)\sigma(t)^\top$, $b = b(t)$, $q = 0$, and $\varphi = \varphi(t;\hat{\mathbf{u}})$. Because $\varphi(t;\hat{\mathbf{u}}) \geq 0$, we conclude from Lemma 2-(i) that $b(t)^\top \hat{\mathbf{u}}(t) \leq 0$, which contradicts $b(t)^\top \hat{\mathbf{u}}(t) > \psi(t)$ and Assumption 3.

We already showed that $b(t)^\top \hat{\mathbf{u}}(t) = \psi(t)$ for all $t \in [0,T)$. Then, for any fixed $t \in [0,T)$, $\hat{\mathbf{u}}(t)$ is the optimal solution to (C.14) and thus the optimal solution to (B.3) with $\Sigma = \sigma(t)\sigma(t)^\top$, $b = b(t)$, $d = \psi(t)$, $q = 0$, and $\varphi = \varphi(t;\hat{\mathbf{u}})$. Because $\varphi(t;\hat{\mathbf{u}}) \geq 0$, Lemma 2-(ii) shows that $\hat{\mathbf{u}}(t)$ is the optimal solution to (2.10), i.e., $\hat{\boldsymbol{\pi}}$ is given by (2.11).

(ii) The case when L is given by (2.8).

Consider any feasible strategy $\hat{\boldsymbol{\pi}}$ that is independent of wealth level x and right-continuous, square-integrable in time $t \in [0,T)$. Fix any $t \in [0,T)$ and $x \in \mathbb{X}_t$ and consider any $\pi \in \mathbb{R}^m$, $\epsilon \in (0, T-t)$, and $\boldsymbol{\pi}_{t,\epsilon,\pi}$ as defined in (2.4). Then,

$$\begin{aligned} X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T) = e^{\int_t^T r(s)ds}\Bigg\{ x + \left[\int_{t+\epsilon}^T e^{-\int_t^s r(z)dz} b(s)^\top \hat{\boldsymbol{\pi}}(s) ds + \int_t^{t+\epsilon} e^{-\int_t^s r(z)dz} b(s)^\top \pi ds\right] \\ + \left[\int_{t+\epsilon}^T e^{-\int_t^s r(z)dz} \hat{\boldsymbol{\pi}}(s)^\top \sigma(s) dW(s) + \int_t^{t+\epsilon} e^{-\int_t^s r(z)dz} \pi^\top \sigma(s) dW(s)\right]\Bigg\}, \end{aligned}$$

which leads to

$$\mathbb{E}_t\left[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right] = e^{\int_t^T r(s)ds}\left\{ x + \int_{t+\epsilon}^T e^{-\int_t^s r(z)dz} b(s)^\top \hat{\boldsymbol{\pi}}(s) ds + \int_t^{t+\epsilon} e^{-\int_t^s r(z)dz} b(s)^\top \pi ds\right\}, \tag{C.15}$$

$$\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right) = e^{2\int_t^T r(s)ds}\left\{ \int_{t+\epsilon}^T e^{-2\int_t^s r(z)dz} \|\sigma(s)^\top \hat{\boldsymbol{\pi}}(s)\|^2 ds + \int_t^{t+\epsilon} e^{-2\int_t^s r(z)dz} \|\sigma(s)^\top \pi\|^2 ds\right\}. \tag{C.16}$$

As a result,

$$\begin{aligned} \lim_{\epsilon\downarrow 0} \frac{\mathbb{E}_t\left[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right] - \mathbb{E}_t\left[X^{\hat{\boldsymbol{\pi}}}(T)\right]}{\epsilon} &= e^{\int_t^T r(s)ds} b(t)^\top\big(\pi - \hat{\boldsymbol{\pi}}(t)\big), \\ \lim_{\epsilon\downarrow 0} \frac{\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right) - \mathrm{var}_t\left(X^{\hat{\boldsymbol{\pi}}}(T)\right)}{\epsilon} &= e^{2\int_t^T r(s)ds}\big(\|\sigma(t)^\top \pi\|^2 - \|\sigma(t)^\top \hat{\boldsymbol{\pi}}(t)\|^2\big). \end{aligned}$$

Because $\hat{\boldsymbol{\pi}}$ is feasible, $\mathbb{E}_t\left[X^{\hat{\boldsymbol{\pi}}}(T)\right]$ hits the target $L(t,x)$ at any time $t \in [0,T)$ and any $x \in \mathbb{X}_t$, which is equivalent to the following conditions:

$$H(t;\hat{\boldsymbol{\pi}}) := e^{\int_t^T r(s)ds}\int_t^T e^{-\int_t^s r(z)dz} b(s)^\top \hat{\boldsymbol{\pi}}(s) ds - \big(e^{\int_t^T \psi(s)ds} - 1\big)/\gamma \geq 0, \quad t \in [0,T). \tag{C.17}$$

Straightforward calculation yields

$$H(T;\hat{\boldsymbol{\pi}}) = 0, \quad H_t(t;\hat{\boldsymbol{\pi}}) := \frac{\partial}{\partial t} H(t;\hat{\boldsymbol{\pi}}) = -e^{\int_t^T r(s)ds} b(t)^\top \hat{\boldsymbol{\pi}}(t) + \psi(t) e^{\int_t^T \psi(s)ds}/\gamma.$$

Then, the same proof as the one of case (i) in this theorem implies that $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy if and only if for any $t \in [0,T)$, $\hat{\boldsymbol{\pi}}(t)$ is the optimal solution to

$$\begin{aligned} \min_{\pi \in \mathbb{R}^m} \quad & \tfrac{1}{2}\pi^\top \sigma(t)\sigma(t)^\top \pi \\ \text{subject to} \quad & \big(b(t)^\top \pi - b(t)^\top \hat{\boldsymbol{\pi}}(t)\big)\mathbf{1}_{H(t;\hat{\boldsymbol{\pi}})=0} \geq 0,\ Q\pi \geq 0. \end{aligned} \tag{C.18}$$

Next, we prove that $\hat{\boldsymbol{\pi}}$ as defined in (2.12) is an equilibrium strategy. Note that $\hat{\boldsymbol{\pi}}(t)$ is square-integrable in t due to Assumptions 1, 3, and 4. In addition, for this $\hat{\boldsymbol{\pi}}$, we have $H(t;\hat{\boldsymbol{\pi}}) = 0$, $t \in [0,T)$, so it is feasible. Moreover, because $v^*(t)$ is the optimal solution to (2.10), we immediately conclude that $\hat{\boldsymbol{\pi}}(t)$ is the optimal solution to (C.18). Consequently, $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy.

Finally, we prove that any equilibrium strategy that is independent of the agent's wealth level x and right-continuous, square-integrable in time t must be the one given by (2.12). Indeed, consider any such strategy $\hat{\boldsymbol{\pi}}$. We first prove that $H_t(t;\hat{\boldsymbol{\pi}}) = 0$ for all $t \in [0,T)$. For the sake of contradiction, suppose it is not the case. Then, by (C.17) and the observation $H(T;\hat{\boldsymbol{\pi}}) = 0$, there must exist $t \in [0,T)$ such that $H_t(t;\hat{\boldsymbol{\pi}}) < 0$. We also have $H(t;\hat{\boldsymbol{\pi}}) > 0$ for the same t, because $H_t(s;\hat{\boldsymbol{\pi}}) < 0$ for $s \in [t, t+\epsilon_0]$ for some $\epsilon_0 \in (0, T-t)$ by the right-continuity of $\hat{\boldsymbol{\pi}}$, b, ψ, and r, and because (C.17)

holds. Then, because $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy, it must be the optimal solution to (C.18) and thus the optimal solution to (B.2) with $\Sigma=\sigma(t)\sigma(t)^\top$, $b=b(t)$, $q=0$, and $\varphi=0$. Then, we conclude from Lemma 2-(i) that $b(t)^\top\hat{\boldsymbol{\pi}}(t)\le 0$, which contradicts $H_t(t;\hat{\boldsymbol{\pi}})<0$ because ψ is nonnegative.

We already showed that $H_t(t;\hat{\boldsymbol{\pi}})=0$ for all $t\in[0,T)$, i.e., $b(t)^\top\hat{\boldsymbol{\pi}}(t)=\psi(t)e^{\int_t^T(\psi(s)-r(s))ds}/\gamma$ for all $t\in[0,T)$. Consequently, $H(t;\hat{\boldsymbol{\pi}})=0$ and thus $\hat{\boldsymbol{\pi}}(t)$ is the optimal solution to (B.3) with $\Sigma=\sigma(t)\sigma(t)^\top$, $b=b(t)$, $q=0$, $\varphi=0$, and $d=b(t)^\top\hat{\boldsymbol{\pi}}(t)=\psi(t)e^{\int_t^T(\psi(s)-r(s))ds}/\gamma$ for any $t\in[0,T)$. As a result, $\hat{\boldsymbol{\pi}}$ must be given by (2.12).

(iii) The case when L is given by (2.9).

Consider any feasible strategy $\hat{\boldsymbol{\pi}}$ such that $\hat{\boldsymbol{\pi}}(t,x)=\hat{\mathbf{v}}(t)(x^*(t)-x)$ for any $t\in[0,T)$ and $x\in\mathbb{R}$, where $\hat{\mathbf{v}}$ is right-continuous and square-integrable in $t\in[0,T)$. Then, we have

$$d\big(x^*(t)-X^{\hat{\boldsymbol{\pi}}}(t)\big)=\big(x^*(t)-X^{\hat{\boldsymbol{\pi}}}(t)\big)\left[\big(r(t)-b(t)^\top\hat{\mathbf{v}}(t)\big)dt-\hat{\mathbf{v}}(t)^\top\sigma(t)dW(t)\right].$$

Therefore, $\mathbb{X}_t\subseteq(-\infty,x^*(t))$ and contains at least one element x such that $x<x^*(t)$.

Fix any $t\in[0,T)$ and $x\in\mathbb{X}_t$ with $x<x^*(t)$. Consider any $\pi\in\mathbb{R}^m$, $\epsilon\in(0,T-t)$, and $\boldsymbol{\pi}_{t,\epsilon,\pi}$ as defined in (2.4). Define

$$\boldsymbol{\pi}'_{t,\epsilon,\pi}(s,y):=\begin{cases}\big(\pi/(x^*(t)-x)\big)(x^*(s)-y), & s\in[t,t+\epsilon),y\in\mathbb{R},\\ \hat{\boldsymbol{\pi}}(s,y), & s\in[t+\epsilon,T),y\in\mathbb{R}\end{cases}\tag{C.19}$$

and denote the corresponding wealth process as $X^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}$. Then, Lemma 3 shows that the derivatives of the mean and variance of $X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)$ with respect to ϵ are the same as those of $X^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(T)$ at $\epsilon=0$. On the other hand, denote $v:=\pi/(x^*(t)-x)$ and $y:=x^*(t)-x$. Define $Y^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(s):=x^*(s)-X^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(s)$, $Y^{\hat{\boldsymbol{\pi}}}(s):=x^*(s)-X^{\hat{\boldsymbol{\pi}}}(s)$, $\mathbf{v}_{t,\epsilon,\pi}(s):=-v\mathbf{1}_{s\in[t,t+\epsilon)}-\hat{\mathbf{v}}(s)\mathbf{1}_{s\in[t+\epsilon,T)}$, $s\in[t,T]$. Then,

$$Y^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(T)=ye^{\int_t^T\left(r(s)+b(s)^\top\mathbf{v}_{t,\epsilon,\pi}(s)-\frac{1}{2}\|\sigma(s)^\top\mathbf{v}_{t,\epsilon,\pi}(s)\|^2\right)ds+\int_t^T\mathbf{v}_{t,\epsilon,\pi}(s)^\top\sigma(s)dW(s)}.\tag{C.20}$$

Similar to the proof of case (i) in this theorem, and recalling (C.7) and (C.8), we conclude

$$\begin{aligned}
&\lim_{\epsilon\downarrow0}\frac{\mathbb{E}_t\left[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right]-\mathbb{E}_t\left[X^{\hat{\boldsymbol{\pi}}}(T)\right]}{\epsilon}=-\lim_{\epsilon\downarrow0}\frac{\mathbb{E}_t\left[Y^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(T)\right]-\mathbb{E}_t\left[Y^{\hat{\boldsymbol{\pi}}}(T)\right]}{\epsilon}\\
=&(x^*(t)-x)e^{\int_t^T\left(r(s)-b(s)^\top\hat{\mathbf{v}}(s)\right)ds}(b(t)^\top v-b(t)^\top\hat{\mathbf{v}}(t)),\\
&\lim_{\epsilon\downarrow0}\frac{\mathrm{var}_t\left(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)\right)-\mathrm{var}_t\left(X^{\hat{\boldsymbol{\pi}}}(T)\right)}{\epsilon}=\lim_{\epsilon\downarrow0}\frac{\mathrm{var}_t\left(Y^{\boldsymbol{\pi}'_{t,\epsilon,\pi}}(T)\right)-\mathrm{var}_t\left(Y^{\hat{\boldsymbol{\pi}}}(T)\right)}{\epsilon},\\
=&2(x^*(t)-x)^2e^{2\int_t^T(r(s)-b(s)^\top\hat{\mathbf{v}}(s))ds}e^{\int_t^T\|\sigma(s)^\top\hat{\mathbf{v}}(s)\|^2ds}\left[\frac{1}{2}\left(\|\sigma(s)^\top v\|^2-\|\sigma(s)^\top\hat{\mathbf{v}}(s)\|^2\right)\right.\\
&\left.-\left(1-e^{-\int_t^T\|\sigma(s)^\top\hat{\mathbf{v}}(s)\|^2ds}\right)\left(b(t)^\top v-b(t)^\top\hat{\mathbf{v}}(t)\right)\right].
\end{aligned}$$

On the other hand, because $\hat{\boldsymbol{\pi}}$ is feasible, $\mathbb{E}_t\left[X^{\hat{\boldsymbol{\pi}}}(T)\right]$ hits the target $L(t,x)$ at any time $t\in[0,T)$ and any $x\in\mathbb{X}_t$, which is equivalent to the following conditions:

$$\bar{H}(t;\hat{\mathbf{v}}):=\int_t^T b(s)^\top\hat{\mathbf{v}}(s)ds-\int_t^T\psi(s)ds\ge0,\quad t\in[0,T).\tag{C.21}$$

Straightforward calculation yields

$$\bar{H}(T;\hat{\mathbf{v}})=0,\quad \bar{H}_t(t;\hat{\mathbf{v}}):=\frac{\partial}{\partial t}\bar{H}_t(t;\hat{\mathbf{v}})=\psi(t)-b(t)^\top\hat{\mathbf{v}}(t).$$

As a result, the same argument as the one in the proof of case (i) in this theorem yields that a sufficient and necessary condition for a feasible strategy $\hat{\boldsymbol{\pi}}(t,x)=\hat{\mathbf{v}}(t)(x^*(t)-x)$ to be an equilibrium one is the following: For each $t\in[0,T)$, $\hat{\mathbf{v}}(t)$ is the optimal solution to

$$\begin{array}{ll}\min\limits_{v\in\mathbb{R}^m} & \frac{1}{2}v^\top\sigma(t)\sigma(t)^\top v+\varphi(t;\hat{\boldsymbol{\pi}})b(t)^\top v\\ \text{subject to} & \big(b(t)^\top v-b(t)^\top\hat{\mathbf{v}}(t)\big)\mathbf{1}_{\bar{H}(t;\hat{\boldsymbol{\pi}})=0}\geq 0,\ Qv\geq 0,\end{array}\tag{C.22}$$

where

$$\varphi(t;\hat{\boldsymbol{\pi}}):=-\left(1-e^{-\int_t^T\|\sigma(s)^\top\hat{\mathbf{v}}(s)\|^2ds}\right)\in(-1,0].\tag{C.23}$$

Now, suppose $\hat{\boldsymbol{\pi}}$ is given by (2.14), i.e., $\hat{\mathbf{v}}(t)$ is the optimal solution to (2.10) for any $t\in[0,T)$. Then, $\bar{H}(t;\hat{\mathbf{v}})=0$ for any $t\in[0,T)$. Consequently, we conclude from Lemma 2-(ii) and from condition (2.13) that $\hat{\mathbf{v}}(t)$ is the optimal solution to (C.22) for any $t\in[0,T)$; i.e., $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy.

Finally, suppose $\hat{\boldsymbol{\pi}}(t,x)=\hat{\mathbf{v}}(t)(x^*(t)-x)$ is an equilibrium strategy, which implies that $\hat{\mathbf{v}}(t)$ is the optimal solution to (C.22) for any $t\in[0,T)$, and we prove that $\hat{\mathbf{v}}(t)=v^*(t)$ for all $t\in[0,T)$. For the sake of contradiction, suppose it is not the case. Denote

$$\begin{aligned}t^*&=\inf\{t\in[0,T)\,|\,\hat{\mathbf{v}}(s)=v^*(s)\quad\forall s\in[t,T)\},\\ t_b^*&=\inf\{t\in[0,T)\,|\,b(s)^\top\hat{\mathbf{v}}(s)=b(s)^\top v^*(s)\quad\forall s\in[t,T)\}\end{aligned}$$

with $\inf\emptyset:=T$. Because $\hat{\mathbf{v}}(t)\neq v^*(t)$ for some $t\in[0,T)$, we must have $t^*>0$.

We claim that $t^*=t_b^*$. It is clear that $t_b^*\leq t^*$, so for the sake of contradiction, suppose $t_b^*<t^*$. By definition, for any $t>t_b^*$, $b(t)^\top\hat{\mathbf{v}}(t)=b(t)^\top v^*(t)=\psi(t)$ and thus $H(t;\hat{\mathbf{v}})=0$; consequently, $\hat{\mathbf{v}}(t)$ must the optimal solution to

$$\begin{array}{ll}\min\limits_{v\in\mathbb{R}^m} & \frac{1}{2}v^\top\sigma(t)\sigma(t)^\top v+\varphi(t;\hat{\boldsymbol{\pi}})b(t)^\top v\\ \text{subject to} & b(t)^\top v\geq b(t)^\top\hat{\mathbf{v}}(t)=\psi(t),\ Qv\geq 0.\end{array}\tag{C.24}$$

From the definition of t^*, we conclude that for any $t\in(t_b^*,t^*)$,

$$\varphi(t;\hat{\boldsymbol{\pi}})=-\left(1-e^{-\int_t^T\|\sigma(s)^\top\hat{\mathbf{v}}(s)\|^2ds}\right)=-\left(1-e^{-\int_t^{t^*}\|\sigma(s)^\top\hat{\mathbf{v}}(s)\|^2ds}e^{-\int_{t^*}^T\|\sigma(s)^\top v^*(s)\|^2ds}\right).$$

Condition (2.15) yields that $\frac{\psi(t)}{\rho(t)}-\left(1-e^{-\int_t^T\|\sigma(s)^\top v^*(s)\|^2ds}\right)\geq\delta_0>0$. By the continuity of indefinite integrals, there exists $\epsilon_0\in(0,t^*)$ such that for any $t\in(t^*-\epsilon_0,t^*)$, $\left|\varphi(t;\hat{\boldsymbol{\pi}})+\left(1-e^{-\int_t^T\|\sigma(s)^\top v^*(s)\|^2ds}\right)\right|\leq\frac{\delta_0}{2}$ and thus $\frac{\psi(t)}{\rho(t)}+\varphi(t;\hat{\boldsymbol{\pi}})\geq\frac{\delta_0}{2}>0,\forall t\in(t_b^*\vee(t^*-\epsilon_0),t^*)$. Consequently, we conclude from Lemma 2-(ii) that $\hat{\mathbf{v}}(t)$, which is the optimal solution to (C.24), must be equal to $v^*(t)$ for any $t\in(t_b^*\vee(t^*-\epsilon_0),t^*)$. This contradicts the definition of t^*, so we must have $t^*=t_b^*$.

Now, the definition of t^* implies that $\bar{H}_t(t;\hat{\mathbf{v}})=\bar{H}(t;\hat{\mathbf{v}})=0$ for any $t\in(t^*,T)$. Because $t^*=t_b^*$, the definition of t_b^* and (C.21) imply that for any $\epsilon\in(0,t^*)$, there exists $t_\epsilon\in(t^*-\epsilon,t^*)$ such that $\bar{H}_t(t_\epsilon;\hat{\mathbf{v}})=\psi(t_\epsilon)-b(t_\epsilon)^\top\hat{\mathbf{v}}(t_\epsilon)=b(t_\epsilon)^\top v^*(t_\epsilon)-b(t_\epsilon)^\top\hat{\mathbf{v}}(t_\epsilon)<0$ and, consequently, by (C.21), $\bar{H}(t_\epsilon;\hat{\mathbf{v}})>0$. As a result, $\hat{\mathbf{v}}(t_\epsilon)$, which is the optimal solution to (C.22) with $t=t_\epsilon$, must solve

$$\begin{array}{ll}\min\limits_{v\in\mathbb{R}^m} & \frac{1}{2}v^\top\sigma(t_\epsilon)\sigma(t_\epsilon)^\top v+\varphi(t_\epsilon;\hat{\boldsymbol{\pi}})b(t_\epsilon)^\top v\\ \text{subject to} & Qv\geq 0.\end{array}\tag{C.25}$$

Then, Lemma 2-(i) yields that

$$b(t_\epsilon)^\top\hat{\mathbf{v}}(t_\epsilon)\leq\max\big(-\varphi(t_\epsilon;\hat{\boldsymbol{\pi}}),0\big)\rho(t_\epsilon)=-\varphi(t_\epsilon;\hat{\boldsymbol{\pi}})\rho(t_\epsilon).\tag{C.26}$$

Because $\epsilon \in (0, t^*)$ is arbitrary, we can choose a sufficiently small $\epsilon > 0$ such that for any $t \in (t^* - \epsilon, t^*)$, $\left|\varphi(t;\hat{\boldsymbol{\pi}}) + \left(1 - e^{-\int_t^T \|\sigma(s)^\top v^*(s)\|^2 ds}\right)\right| \le \frac{\delta_0}{2}$ and thus $\frac{\psi(t)}{\rho(t)} + \varphi(t;\hat{\boldsymbol{\pi}}) \ge \frac{\delta_0}{2} > 0$. In combination with (C.26), we conclude that $b(t_\epsilon)^\top \hat{\mathbf{v}}(t_\epsilon) \le \psi(t_\epsilon)$, which contradicts $\bar{H}_t(t_\epsilon; \hat{\mathbf{v}}) = \psi(t_\epsilon) - b(t_\epsilon)^\top \hat{\mathbf{v}}(t_\epsilon) < 0$.

Thus, we have $t^* = 0$; i.e., $\hat{\mathbf{v}}(t) = v^*(t)$ for any $t \in [0, T)$. □

Proof of Proposition 2. For the sake of contradiction, suppose $\hat{\boldsymbol{\pi}}$ is an equilibrium strategy. Because (2.13) does not hold, there exists $t \in [0, T)$ such that $e^{-\int_t^T \|\sigma(s)^\top v^*(s)\|^2 ds} < 1 - \frac{\psi(t)}{\rho(t)}$, which leads to $\varphi(t;\hat{\boldsymbol{\pi}}) < -\psi(t)/\rho(t)$, where φ is defined by (C.23). Note that the proof of Theorem 1-(iii) shows that $v^*(t)$ is the optimal solution to (C.22) with $Q = 0$. On the other hand, because $v^*(t)$ is the optimal solution to (2.10) with $Q = 0$, we have $b(s)^\top v^*(s) = \psi(s)$, $s \in [0, T)$. Consequently, (C.21) implies that $\bar{H}(t;\hat{\boldsymbol{\pi}}) = 0$, so $v^*(t)$ must be the optimal solution to (C.24) with $Q = 0$. Now, consider the minimization of $\frac{1}{2} v^\top \sigma(t)\sigma(t)^\top v + \varphi(t;\hat{\boldsymbol{\pi}}) b(t)^\top v$ in v, and straightforward calculation yields that the minimizer is $\bar{v}(t) := -(\sigma(t)\sigma(t))^{-1} b(t)\varphi(t;\hat{\boldsymbol{\pi}})$. It is easy to see that $b(t)^\top \bar{v}(t) = -\rho(t)\varphi(t;\hat{\boldsymbol{\pi}}) > \psi(t)$, showing that $\bar{v}(t)$ is the optimal solution to (C.24) with $Q = 0$. In other words, $v^*(t) = \bar{v}(t)$ and thus $b(t)^\top v^*(t) > \psi(t)$. On the other hand, because $v^*(t)$ is the optimal solution to (2.10) with $Q = 0$, we must have $b(t)^\top v^*(t) = \psi(t)$, which is a contradiction. Thus, $\hat{\boldsymbol{\pi}}$ cannot be an equilibrium strategy. □

Appendix D: Generalization of Portfolio Constraints for the First Target In this section, assuming that L is given by (2.7), we consider the following problem

$$\begin{cases} \min\limits_{\boldsymbol{\pi}} & \mathrm{var}_t(X^{\boldsymbol{\pi}}(T)) \\ \text{subject to} & dX^{\boldsymbol{\pi}}(s) = \big[r(s)X^{\boldsymbol{\pi}}(s) + b(s)^\top \boldsymbol{\pi}(s, X^{\boldsymbol{\pi}}(s))\big] ds \\ & \qquad + \boldsymbol{\pi}(s, X^{\boldsymbol{\pi}}(s))^\top \sigma(s) dW(s),\ s \in [t, T),\ X^{\boldsymbol{\pi}}(t) = x_t, \\ & \mathbb{E}_t[X^{\boldsymbol{\pi}}(T)] \ge L(t, x_t), \\ & Q\boldsymbol{\pi}(s, X^{\boldsymbol{\pi}}(s)) \ge q X^{\boldsymbol{\pi}}(s),\ s \in [t, T) \end{cases} \tag{D.1}$$

that generalizes (2.3) by introducing a more general constraint on portfolios: $Q\boldsymbol{\pi}(s, X^{\boldsymbol{\pi}}(s)) \ge qX^{\boldsymbol{\pi}}(s)$. Here, q is an n-dimensional column vector and we assume that $q \le 0$ so that $\pi(t) = 0, t \in [0, T)$, which stands for the strategy of always investing all money in the risk-free asset, meets the portfolio constraint. This constraint can accommodate many commonly observed portfolio constraints in the market, such as the no-shorting and the no-borrowing constraints.

Assumption 6. *$\int_0^T \|u^*(t)\|^2 dt < +\infty$, where $u^*(t)$ is the optimal solution to (B.1).*

Theorem 4. *Suppose that L is given by (2.7) and Assumptions 1, 3, 5, and 6 hold. Then,*

$$\hat{\boldsymbol{\pi}}(t, x) = x u^*(t), \quad t \in [0, T), x \in \mathbb{R}, \tag{D.2}$$

where $u^*(t)$ is the optimal solution to (B.1), is an equilibrium strategy for problem (D.1). Moreover, any equilibrium strategy $\hat{\boldsymbol{\pi}}$ such that $\hat{\boldsymbol{\pi}}(t, x)/x$ is independent of $x \in \mathbb{R}$ and right-continuous, square-integrable in $t \in [0, T)$ must be given by (D.2).

Proof. The proof is the same as that for case (i) of Theorem 1. □

Appendix E: Different Notions of Equilibrium Strategies In the following, we present two alternative definitions of equilibrium strategies for our problem and examine whether they affect the main results.

E.1. Strong Equilibrium

DEFINITION 2. A feasible feedback strategy $\hat{\boldsymbol{\pi}}$ for problem (2.3) is an equilibrium strategy of the agent if for any $t\in[0,T)$, $x\in\mathbb{X}_t$, and $\pi\in\Pi^{\hat{\boldsymbol{\pi}}}_{t,x}$, there exists $\epsilon_0\in(0,T-t)$, such that (2.5) holds.

In the above definition, the portfolio π that the time t-self of the agent can implement in an infinitesimally period can be $\hat{\boldsymbol{\pi}}(t,x)$, while such possibility is excluded in Definition 1 in the main text. Following Huang and Zhou [29] and He and Jiang [25], Definition 2 is referred to as *strong equilibria* and Definition 1 is referred to as *regular equilibria*. In the following, we examine whether the strategy we derived in the main text is a strong equilibrium. As the following two propositions show, whether the answer is affirmative depends on which target for the expected terminal wealth is used. This is consistent with the finding in He and Jiang [25] that strong equilibrium strategies do not exist in general.

A function g defined on an interval I is *non-oscillating* at $t\in I$ if there exists $\epsilon_t>0$ such that either $g(s)<g(t),\forall s\in(t,t+\epsilon_t)\cap I$, or $g(s)>g(t),\forall s\in(t,t+\epsilon_t)$, or $g(s)=g(t),\forall s\in(t,t+\epsilon_t)$.

PROPOSITION 5. *Consider problem (2.3) with expected terminal wealth target (2.8) and recall $\hat{\boldsymbol{\pi}}(t)$ as defined in (2.12). For each $t\in[0,T)$, define $g(s):=b(s)^\top\hat{\boldsymbol{\pi}}(t)-b(s)^\top\hat{\boldsymbol{\pi}}(s), s\in[t,T)$ and assume g to be non-oscillating at t. Then, Theorem 1-(ii) still holds when the notion of equilibrium strategies is given by Definition 2.*

Proof. It is obvious that the uniqueness of equilibrium strategies under Definition 1 implies the uniqueness of equilibrium strategies under Definition 2, so we only need to prove that $\hat{\boldsymbol{\pi}}$ as defined by (2.12) is an equilibrium strategy under Definition 2. To this end, we only need to prove that for any $t\in[0,T)$ and wealth level $x\in\mathbb{R}$ such that

$$\mathbb{E}_t\left[X^{\boldsymbol{\pi}_{t,\epsilon,\hat{\boldsymbol{\pi}}(t)}}(T)\right]\geq\frac{1}{\gamma}(e^{\int_t^T\psi(s)ds}-1)+xe^{\int_t^T r(s)ds},\quad\forall\epsilon\in(0,\epsilon_1] \tag{E.1}$$

for certain $\epsilon_1>0$, there exists $\epsilon_0\in(0,\epsilon_1]$ such that $\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\hat{\boldsymbol{\pi}}(t)}}(T))-\mathrm{var}_t(X^{\hat{\boldsymbol{\pi}}}(T))\geq 0,\forall\epsilon\in(0,\epsilon_0]$.

By (C.15) and the observation that $b(s)^\top\hat{\boldsymbol{\pi}}(s)=\frac{1}{\gamma}e^{\int_s^T(\psi(z)-r(z))dz}b(s)^\top v^*(s)=\frac{\psi(s)}{\gamma}e^{\int_s^T(\psi(z)-r(z))dz}$ for any $s\in[0,T)$, straightforward calculation yields that (E.1) is equivalent to

$$\int_t^{t+\epsilon}e^{-\int_t^s r(z)dz}b(s)^\top(\hat{\boldsymbol{\pi}}(t)-\hat{\boldsymbol{\pi}}(s))ds\geq 0,\quad\forall\epsilon\in(0,\epsilon_1]. \tag{E.2}$$

Because $g(s)=b(s)^\top(\hat{\boldsymbol{\pi}}(t)-\hat{\boldsymbol{\pi}}(s)),s\in[t,T)$ is non-oscillating at $s=t$, (E.2) holds for some $\epsilon_1>0$ if and only if there exists $\epsilon_2>0$ such that $b(s)^\top\hat{\boldsymbol{\pi}}(t)\geq b(s)^\top\hat{\boldsymbol{\pi}}(s),\forall s\in(t,t+\epsilon_2]$. Without loss of generality, we set $\epsilon_2=\epsilon_1$ in the following.

Now, for any $s\in(t,t+\epsilon_1]$, because $b(s)^\top\hat{\boldsymbol{\pi}}(t)\geq b(s)^\top\hat{\boldsymbol{\pi}}(s)=\frac{\psi(s)}{\gamma}e^{\int_s^T(\psi(z)-r(z))dz}$, $\hat{\boldsymbol{\pi}}(t)\gamma e^{-\int_s^T(\psi(z)-r(z))dz}$ is feasible to problem (2.10) with t replaced by s. Consequently, we have $\|\sigma(s)^\top\hat{\boldsymbol{\pi}}(t)\gamma e^{-\int_s^T(\psi(z)-r(z))dz}\|^2\geq\|\sigma(s)^\top v^*(s)\|^2$ and then $\|\sigma(s)^\top\hat{\boldsymbol{\pi}}(t)\|^2\geq\|\sigma(s)^\top\hat{\boldsymbol{\pi}}(s)\|^2,\forall s\in(t,t+\epsilon_1]$. Then, we immediately conclude from (C.16) that $\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\hat{\boldsymbol{\pi}}(t)}}(T))-\mathrm{var}_t(X^{\hat{\boldsymbol{\pi}}}(T))\geq 0,\forall\epsilon\in(0,\epsilon_1]$. □

PROPOSITION 6. *Consider problem (2.3). Suppose that Assumption 1 holds, that r, b, and σ are constants, and that $r>0$. Suppose that L is given by (2.7) with $\psi(s)=\psi_0e^{-rs}/\big(1+\psi_0(1-e^{-rs})/r\big),s\in[0,T]$ for some constant $\psi_0>0$, and that there are no portfolio constraints, i.e,. $Q=0$. Then, $\hat{\boldsymbol{\pi}}$ as given by (2.11) is not an equilibrium strategy under Definition 2.*

Proof. Because $Q=0$, we have $v^*(t)=\psi(t)(\sigma\sigma^\top)^{-1}b/\rho$, $t\in[0,T)$. Recalling (2.4), direct calculation shows that for any $t\in[0,T)$, $x\in\mathbb{X}_t$, $\pi\in\mathbb{R}^m$, and $\epsilon\in[0,T-t)$, we have

$$\begin{aligned}X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)=e^{r\epsilon}\left\{x+\left[\int_t^{t+\epsilon}e^{-r(s-t)}b^\top\pi ds\right]+\left[\int_t^{t+\epsilon}e^{-r(s-t)}\pi^\top\sigma dW(s)\right]\right\}\\ \times e^{\int_{t+\epsilon}^T\left(r+b^\top v^*(s)-\frac{1}{2}\|\sigma^\top v^*(s)\|^2\right)ds+\int_{t+\epsilon}^T v^*(s)^\top\sigma dW(s)}\end{aligned} \tag{E.3}$$

and for $\boldsymbol{\pi} \in \{\boldsymbol{\pi}_{t,\epsilon,\pi}, \hat{\boldsymbol{\pi}}\}$,

$$\mathbb{E}_t[X^{\boldsymbol{\pi}}(T)] = \mathbb{E}_t\left[X^{\boldsymbol{\pi}}(t+\epsilon)\right]\mathbb{E}_t\left[Z_{t+\epsilon,T}\right], \tag{E.4}$$

$$\mathrm{var}_t(X^{\boldsymbol{\pi}}(T)) = \mathbb{E}_{t+\epsilon}\left[\left(Z_{t+\epsilon,T}\right)^2\right]\mathrm{var}_t\left(X^{\boldsymbol{\pi}}(t+\epsilon)\right) + \mathrm{var}_{t+\epsilon}\left(Z_{t+\epsilon,T}\right)\left(\mathbb{E}_t\left[X^{\boldsymbol{\pi}}(t+\epsilon)\right]\right)^2, \tag{E.5}$$

where $Z_{t+\epsilon,T} := e^{\int_{t+\epsilon}^T \left(r + b^\top v^*(s) - \frac{1}{2}\|\sigma^\top v^*(s)\|^2\right)ds + \int_{t+\epsilon}^T v^*(s)^\top \sigma dW(s)}$.

Now, consider time 0 and $\pi = \hat{\boldsymbol{\pi}}(0,x_0) = x_0\psi_0(\sigma\sigma^\top)^{-1}b/\rho$. Direct calculation yields that for any $\epsilon \in [0,T)$, $\mathbb{E}\left[X^{\boldsymbol{\pi}_{0,\epsilon,\pi}}(\epsilon)\right] = e^{r\epsilon}\left(x_0 + \int_0^\epsilon e^{-rs}b^\top \pi ds\right) = \mathbb{E}\left[X^{\hat{\boldsymbol{\pi}}}(\epsilon)\right]$, where the second equality is the case due to our specific choice of ψ. The above, together with (E.4), implies that $\mathbb{E}\left[X^{\boldsymbol{\pi}_{0,\epsilon,\pi}}(T)\right] = \mathbb{E}\left[X^{\hat{\boldsymbol{\pi}}}(T)\right]$. Consequently, $\pi \in \Pi^{\hat{\boldsymbol{\pi}}}_{0,x_0}$. In addition, for any $\epsilon \in [0,T)$, we have

$$\mathrm{var}\left(X^{\boldsymbol{\pi}_{0,\epsilon,\pi}}(\epsilon)\right) = x_0^2\frac{\psi_0^2}{2r\rho}\left(e^{2r\epsilon}-1\right), \quad \mathrm{var}\left(X^{\hat{\boldsymbol{\pi}}}(\epsilon)\right) = x_0^2 e^{2\int_0^\epsilon [r+\psi(s)]ds}\left(e^{\int_0^\epsilon \psi(s)^2/\rho ds}-1\right).$$

Defining $f(\epsilon) := e^{-2r\epsilon}\left[\mathrm{var}\left(X^{\hat{\boldsymbol{\pi}}}(\epsilon)\right) - \mathrm{var}\left(X^{\boldsymbol{\pi}_{0,\epsilon,\pi}}(\epsilon)\right)\right]/x_0^2$, $\epsilon \in [0,T)$. Then, we have $f(0) = 0$ and

$$f'(\epsilon) = \left(e^{\int_0^\epsilon \psi(s)^2/\rho ds}-1\right)\left(2\psi(\epsilon)e^{2\int_0^\epsilon \psi(s)ds} + \frac{\psi_0^2}{\rho}e^{-2r\epsilon}\right) > 0, \quad \epsilon \in (0,T).$$

As a result, for all $\epsilon \in (0,T)$, we have $\mathrm{var}\left(X^{\hat{\boldsymbol{\pi}}}(\epsilon)\right) - \mathrm{var}\left(X^{\boldsymbol{\pi}_{0,\epsilon,\pi}}(\epsilon)\right) > 0$. Plugging the above inequality into (E.5) and recalling that $\mathbb{E}\left[X^{\boldsymbol{\pi}_{0,\epsilon,\pi}}(\epsilon)\right] = \mathbb{E}\left[X^{\hat{\boldsymbol{\pi}}}(\epsilon)\right]$, we conclude that $\mathrm{var}\left(X^{\hat{\boldsymbol{\pi}}}(T)\right) - \mathrm{var}\left(X^{\boldsymbol{\pi}_{0,\epsilon,\pi}}(T)\right) > 0$. Thus, (2.5) does not hold for $t = 0$ and $\pi = \hat{\boldsymbol{\pi}}(0,x_0) \in \Pi^{\hat{\boldsymbol{\pi}}}_{0,x_0}$, so $\hat{\boldsymbol{\pi}}$ as given by (2.11) is not an equilibrium strategy under Definition 2. □

E.2. Nonconstant Alternative Portfolios For each $t \in [0,T)$, denote by $\mathbf{D}_t$ the set of portfolio strategies $\mathbf{a}$ such that (i) $\|\mathbf{a}(s,y) - \mathbf{a}(s,y')\| \le L|y-y'|, \forall y,y' \in \mathbb{R}, s \in [t,t+\bar{\epsilon}]$ for some $\bar{\epsilon} \in (0,T-t)$ and $L > 0$ and (ii) that $\lim_{s\downarrow t}\left(\mathbf{a}(s,x) - \mathbf{a}(t,x)\right) = 0, \forall x \in \mathbb{R}$. Given a feasible feedback strategy $\hat{\boldsymbol{\pi}}(t,x)$, $(t,x) \in [0,T)\times\mathbb{R}$ for problem (2.3), for each $t \in [0,T)$, $\epsilon \in (0,T-t)$, and $\mathbf{a} \in \mathbf{D}_t$, denote

$$\boldsymbol{\pi}_{t,\epsilon,\mathbf{a}}(s,y) := \begin{cases} \mathbf{a}(s,y), & s \in [t,t+\epsilon), y \in \mathbb{R}, \\ \hat{\boldsymbol{\pi}}(s,y), & s \notin [t,t+\epsilon), y \in \mathbb{R}. \end{cases} \tag{E.6}$$

For each $t \in [0,T)$ and $x \in \mathbb{R}$, denote

$$\bar{\Pi}^{\hat{\boldsymbol{\pi}}}_{t,x} := \{\mathbf{a} \in \mathbf{D}_t \mid \text{there exists } \bar{\epsilon} \in (0,T-t) \text{ such that } \boldsymbol{\pi}_{t,\epsilon,\mathbf{a}} \text{ is feasible }, \ \forall \epsilon \in (0,\bar{\epsilon}]\}.$$

Definition 3. Let $\hat{\boldsymbol{\pi}}$ be a feasible feedback strategy for problem (2.3). $\hat{\boldsymbol{\pi}}$ is an *equilibrium strategy* of the agent if for any $t \in [0,T)$, $x \in \mathbb{X}_t$, and $\mathbf{a} \in \bar{\Pi}^{\hat{\boldsymbol{\pi}}}_{t,x}$ with $\mathbf{a}(t,x) \neq \hat{\boldsymbol{\pi}}(t,x)$, there exists $\epsilon_0 \in (0,T-t)$ such that $\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\mathbf{a}}}(T)) - \mathrm{var}_t(X^{\hat{\boldsymbol{\pi}}}(T)) \ge 0$, $\forall \epsilon \in (0,\epsilon_0]$.

Definition 3 differs from Definition 1 in that the alternative strategy $\mathbf{a}$ implemented by the time-t self of the agent in an infinitesimally small time period can be a nonconstant. The following proposition shows that the use of Definition 3 for the notion of equilibrium strategies does not change the results in the present paper.

Proposition 7. *Theorem 1 and Theorem 4 still hold if the notion of equilibrium strategies is given by Definition 3.*

Proof. We consider case (i) in Theorem 1 only, as the other cases can be treated similarly. Because the notion of equilibrium strategies under Definition 3 allows for a larger class of alternative strategies taken by the agent at certain time than those under Definition 1, the uniqueness result in case (i) of Theorem 1 still holds. Thus, we only need to prove the existence.

Consider strategy $\hat{\boldsymbol{\pi}}$ given by (2.11). Fix $t \in [0,T)$, and $x \in \mathbb{R}$. Consider any $\mathbf{a} \in \mathbf{D}_t$ and denote $\pi := \mathbf{a}(t,x)$. Define $\boldsymbol{\pi}_{t,\epsilon,\pi}$ and $\boldsymbol{\pi}_{t,\epsilon,\mathbf{a}}$ as in (2.4) and in (E.6), respectively. Then, Lemma 3 shows that

$$\lim_{\epsilon \downarrow 0} \frac{\mathbb{E}_t[X^{\boldsymbol{\pi}_{t,\epsilon,\mathbf{a}}}(T)] - \mathbb{E}_t[X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T)]}{\epsilon} = 0, \quad \lim_{\epsilon \downarrow 0} \frac{\mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\mathbf{a}}}(T)) - \mathrm{var}_t(X^{\boldsymbol{\pi}_{t,\epsilon,\pi}}(T))}{\epsilon} = 0.$$

Consequently, the proof of case (i) in Theorem 1 yields that $\hat{\boldsymbol{\pi}}$ as defined in (2.11) is an equilibrium strategy under Definition 3. □

References

[1] Ameriks, J. and Zeldes, S. P. [2004]. How do household portfolio shares vary with age, *Technical report*, Columbia University.

[2] Bajeux-Besnainou, I. and Portait, R. [1998]. Dynamic asset allocation in a mean-variance framework, *Management Science* **44**(11-part-2): S79–S95.

[3] Basak, S. and Chabakauri, G. [2010]. Dynamic mean-variance asset allocation, *Review of Financial Studies* **23**(8): 2970–3016.

[4] Bayraktar, E., Zhang, J. and Zhou, Z. [2019]. Time consistent stopping for the mean-standard deviation problem—the discrete time case, *SIAM Journal on Financial Mathematics* **10**(3): 667–697.

[5] Bayraktar, E., Zhang, J. and Zhou, Z. [2020]. Equilibrium concepts for timeinconsistent stopping problems in continuous time, *Mathematical Finance* **Forthcoming**.

[6] Bensoussan, A., Wong, K. C. and Yam, S. C. P. [2019]. A paradox in time-consistency in the mean–variance problem?, *Finance and Stochastics* **23**(1): 173–207.

[7] Bensoussan, A., Wong, K. C., Yam, S. C. P. and Yung, S.-P. [2014]. Time-consistent portfolio selection under short-selling prohibition: From discrete to continuous setting, *SIAM Journal on Financial Mathematics* **5**(1): 153–190.

[8] Bielecki, T. R., Jin, H., Pliska, S. R. and Zhou, X. Y. [2005]. Continuous-time mean-variance portfolio selection with bankruptcy prohibition, *Mathematical Finance* **15**(2): 213–244.

[9] Björk, T., Khapko, M. and Murgoci, A. [2017]. On time-inconsistent stochastic control in continuous time, *Finance and Stochastics* **21**(2): 331–360.

[10] Björk, T. and Murgoci, A. [2010]. A general theory of Markovian time inconsistent stochastic control problems. SSRN:1694759.

[11] Björk, T., Murgoci, A. and Zhou, X. Y. [2014]. Mean-variance portfolio optimization with state dependent risk aversion, *Mathematical Finance* **24**(1): 1–24.

[12] Černỳ, A. [2020]. Semimartingale theory of monotone mean–variance portfolio allocation, *Mathematical Finance* **30**(3): 1168–1178.

[13] Chen, L. [2020]. *Continuous-Time and Distributionally Robust Mean-Variance Models*, PhD thesis, Columbia University.

[14] Cui, X., Li, D. and Li, X. [2017]. Mean-variance policy for discrete-time cone-constrained markets: time consistency in efficiency and the minimum-variance signed supermartingale measure, *Mathematical Finance* **27**(2): 471–504.

[15] Cui, X., Li, D., Wang, S. and Zhu, S. [2012]. Better than dynamic mean-variance: Time inconsistency and free cash flow stream, *Mathematical Finance* **22**(2): 346–378.

[16] Cui, X., Li, X., Li, D. and Shi, Y. [2017]. Time consistent behavioral portfolio policy for dynamic mean–variance formulation, *Journal of the Operational Research Society* **68**(12): 1647–1660.

[17] Cui, X., Xu, L. and Zeng, Y. [2016]. Continuous time mean-variance portfolio optimization with piecewise state-dependent risk aversion, *Optimization Letters* **10**(8): 1681–1691.

[18] Czichowsky, C. [2013]. Time-consistent mean-variance portfolio selection in discrete and continuous time, *Finance and Stochastics* **17**(2): 227–271.

[19] Dai, M., Jin, H., Kou, S. and Xu, Y. [2019]. A dynamic mean-variance analysis for log returns, *Management Science* **forthcoming**.

[20] Daniel, J. W. [1973]. Stability of the solution of definite quadratic programs, *Mathematical Programming* **5**(1): 41–53.

[21] Ekeland, I. and Lazrak, A. [2006]. Being serious about non-commitment: subgame perfect equilibrium in continuous time. arXiv:math/0604264.

[22] Ekeland, I. and Lazrak, A. [2008]. Equilibrium policies when preferences are time inconsistent. arXiv:0808.3790.

[23] Ekeland, I. and Lazrak, A. [2010]. The golden rule when preferences are time inconsistent, *Mathematics and Financial Economics* **4**(1): 29–55.

[24] Ekeland, I. and Pirvu, T. [2008]. Investment and consumption without commitment, *Mathematics and Financial Economics* **2**(1): 57–86.

[25] He, X. D. and Jiang, Z. [2019]. On the equilibrium strategies for time-inconsistent problems in continuous time. SSRN:3308274.

[26] He, X. D. and Jiang, Z. [2020]. Dynamic mean-variance efficient fractional kelly portfolios in a stochastic volatility model. SSRN:3670621.

[27] Hu, Y., Jin, H. and Zhou, X. Y. [2012]. Time-inconsistent stochastic linear-quadratic control, *SIAM Journal on Control and Optimization* .

[28] Hu, Y., Jin, H. and Zhou, X. Y. [2017]. Time-inconsistent stochastic linear-quadratic control: Characterization and uniqueness of equilibrium, *SIAM Journal on Control and Optimization* **55**(2): 1261–1279.

[29] Huang, Y.-J. and Zhou, Z. [2019]. Strong and weak equilibria for time-inconsistent stochastic control in continuous time, *Mathematics of Operations Research* **forthcoming**. arXiv:1809.09243.

[30] Karnam, C., Ma, J., Zhang, J. et al. [2017]. Dynamic approaches for some time-inconsistent optimization problems, *Annals of Applied Probability* **27**(6): 3435–3477.

[31] Kryger, E. M. and Steffensen, M. [2010]. Some solvable portfolio problems with quadratic and collective objectives. SSRN:1577265.

[32] Kryger, E., Nordfang, M.-B. and Steffensen, M. [2020]. Optimal control of an objective functional with non-linearity between the conditional expectations: solutions to a class of time-inconsistent portfolio problems, *Mathematical Methods of Operations Research* **91**(3): 405–438.

[33] Laibson, D. [1997]. Golden eggs and hyperbolic discounting, *Quarterly Journal of Economics* **112**(2): 443–477.

[34] Li, D. and Ng, W. L. [2000]. Optimal dynamic portfolio selection: Multiperiod mean-variance formulation, *Mathematical Finance* **10**(3): 387–406.

[35] Li, S., Luong, C., Angkola, F. and Wu, Y. [2016]. Optimal asset portfolio with stochastic volatility under the mean-variance utility with state-dependent risk aversion, *Journal of Industrial & Management Optimization* **12**(4): 1521–1533.

[36] Lim, A. E. and Zhou, X. Y. [2002]. Mean-variance portfolio selection with random parameters in a complete market, *Mathematics of Operations Research* **27**(1): 101–120.

[37] Markowitz, H. [1952]. Portfolio selection, *Journal of Finance* **7**(1): 77–91.

[38] Merton, R. C. [1969]. Lifetime portfolio selection under uncertainty: the continuous-time case, *Review of Economics and Statistics* **51**(3): 247–257.

[39] O'Donoghue, T. and Rabin, M. [1999]. Doing it now or later, *American Economic Review* **89**(1): 103–124.

[40] Pedersen, J. L. and Peskir, G. [2016]. Optimal mean-variance selling strategies, *Mathematics and Financial Economics* **10**(2): 203–220.

[41] Pedersen, J. L. and Peskir, G. [2017]. Optimal mean-variance portfolio selection, *Mathematics and Financial Economics* **11**(2): 137–160.

[42] Peleg, B. and Yaari, M. E. [1973]. On the existence of a consistent course of action when tastes are changing, *Review of Economic Studies* **40**(3): 391–401.

[43] Pollak, R. A. [1968]. Consistent planning, *Review of Economic Studies* **35**(2): 201–208.

[44] Pun, C. S. [2018]. Time-consistent mean-variance portfolio selection with only risky assets, *Economic Modelling* **75**: 281–292.

[45] Richardson, H. R. [1989]. A minimum variance result in continuous trading portfolio optimization, *Management Science* **35**(9): 1045–1055.

[46] Strotz, R. H. [1955-1956]. Myopia and inconsistency in dynamic utility maximization, *Review of Economic Studies* **23**(3): 165–180.

[47] Strub, M. S. and Li, D. [2020]. A note on monotone mean-variance preferences for continuous processes, *Operations Research Letters* .

[48] Trybuła, J. and Zawisza, D. [2019]. Continuous-time portfolio choice under monotone mean-variance preferences—stochastic factor case, *Mathematics of Operations Research* **44**(3): 966–987.

[49] Yong, J. and Zhou, X. Y. [1999]. *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer.

[50] Zhou, X. Y. and Li, D. [2000]. Continuous time mean-variance portfolio selection: A stochastic LQ framework, *Applied Mathematics and Optimization* **42**(1): 19–33.