A Continuous Time Framework for Sequential Goal-Based Wealth Management

October 13, 2023

Abstract

We develop a continuous time framework for sequential goals-based wealth management. A stochastic factor process drives asset price dynamics as well as the client's goal amount and income. We prove the weak dynamic programming principle for the value function of our control problem, which we show to be the unique viscosity solution of the corresponding Hamilton-Jacobi-Bellman equation. We develop an equivalent and computationally efficient representation of the Hamiltonian, which yields the optimal portfolio within a factor-dependent opportunity set defined by the maximum and minimum variance hypersurfaces. Our analysis shows that it is optimal to fund an expiring goal up to the level where the marginal benefit of additional fundedness is exceeded by the opportunity cost of diverting wealth from future goals. An all-or-nothing investor is more risk averse towards an approaching goal deadline if she is well funded, but more risk seeking if she is not on track with upcoming goals, compared to an investor with flexible goals.

1 Introduction

Goals-based wealth management is an investment paradigm centered around the fulfillment of client's consumption goals. Clients specify needs, wishes, wants, and dreams, along with the probabilities they would like these to be attained (e.g. Brunel [2015]). The wealth manager then constructs an investment strategy to meet these goals to the extent possible.

Consumption goals are central in the behavioral portfolio theory of Shefrin and Statman [2000]. This theory states that investors segregate their portfolios into multiple mental accounts, with various levels of fear and aspiration depending on the goals associated with that account. Investors then treat their money with risk-return preferences depending on the goal. They may be more risk averse for higher priority goals and less risk-averse for lower priority goals (Thaler [1985a] and Thaler [1985b]). The communication strategy is client-centered, and focused around terms familiar to households such as investment horizons, target wealth levels and desired probabilities of attaining them. As argued in Das et al. [2010], behavioral research has shown that investors understand the notion of goals better than the concept of utility function. This view is also consistent with a stream of literature in experimental social psychology, which has shown that agents are intrinsically driven by goals in their daily life (see, for instance, Locke and Latham [2002]).

Unlike the classical approach of tracking risks and expected returns of benchmarks via Sharpe-ratio maximization, the wealth manager focuses on meeting the investor's needs and preferences (Chhabra [2005], Nevins [2004]). The risk is related to the probability of failing to meet, or meeting only partially, a set of goals, rather than to the market volatility as in the classical mean-variance optimization framework.

The importance of goal-based investing has been recognized both by the academic community and by the private sector. In the 2020 Q Group panel discussion (see Leibowitz et al. [2016]), Robert Merton emphasized goal-based investing as one of the most important problems in financial engineering for the next decade. He stressed the view that managers will be driven by the idea of greater service by knowing their clients better, understanding their needs, and designing dynamic strategies that achieve their goals.

Because of the intuitive appeal and ability to map communicated information into portfolios which meet the client's goal specifications, the principles of goals-based wealth management are also becoming increasingly predominant in the retail robo-advising industry. For example, investment goals on the Schwab and Betterment platforms range from generating retirement income to major one-time big purchases such as house downpayments or new car purchases. Robo-advising firms let their customers specify their own goals. Anecdotal evidence suggests that clients typically specify at least three or more goals, including retirement, kids' college tuition, vacation, and home purchase. These claims are also supported by data in the Fifth Annual Global Investor Sentiment Survey by Franklin Templeton[1], where

[1]See `https://www.franklintempleton.com/press-releases/news-room/2015/investor-optimism-dominates-in-fifth-annual-global-investor-sentiment-survey-by-franklin-templeton` for details. This is a broad survey where Franklin Templeton compiled responses from over 11,500 investors in 23 countries across the Americas, Africa, Asia Pacific and Europe.

respondents shared their top investment goals. Retirement, preparation for emergencies, and planning a vacation ranked among the top three for US clients. Home purchases and investing in a business were the top goals in Latin America, while planning a vacation was the highest priority in Europe.

We develop a continuous time framework for sequential goals-based wealth management (SGBWM). Our model features a stochastic factor process, whose dynamics captures the evolution of macroeconomic factors, such as inflation, employment levels, and gross domestic product. The factor process drives the asset dynamics in the client's portfolio.[2] We position ourselves from the point of view of a client, who is endowed with an initial amount of wealth. The client also specifies the portion of income (e.g. paychecks) injected into her portfolio at any time period. The client specifies goals she would like to reach throughout the investment horizon, with each goal uniquely characterized by its amount, deadline, and importance weight. Both the client's income process and goal amounts depend on the factor process.

Our framework handles a broad class of client's preference functions with respect to the degree of goal fulfillment. At each goal deadline, we solve a static convex optimization problem to recover the optimal *funding ratio*, i.e., the percentage of goal amount consumed by the client. Between two consecutive goals, the client solves a continuous time stochastic control problem to recover the optimal investment strategy. The investment set includes risky assets whose dynamics are given by a correlated geometric Brownian motion, and a risk-free asset. To maintain consistency with practices of goal-based robo-advisors which do not take short positions but just invest their clients' money, we allow for long-only positions, prohibit borrowing from the risk-free money market account, and impose a maximum volatility constraint on the client's portfolio.

There are methodological contributions in our efforts. The dependence of the portfolio volatility constraint on the stochastic factor process and the non-smoothness of the value function at the goal deadlines make the structure of our control problem unique, and not covered by existing studies. We show that the value function of our control problem satisfies the weak dynamic programming principle (DPP), and can be recovered as the unique viscosity solution of the corresponding Hamilton-Jacobi-Bellman (HJB) equation, which is a fully nonlinear partial differential equation (PDE). By expressing the Hamiltonian in terms of the

[2] This modeling choice is supported by ample empirical evidence of time variation in rates and expected stock returns (see, for instance, Fama and Schwert [1977]), and in stochastic volatility of returns (see Pagan and Schwert [1990]).

standard deviation, mean of minimum-variance portfolio returns, and the covariance of these returns with the stochastic factor process, we provide a novel interpretation of the optimal portfolio strategy. More specifically, when client preferences are linear with respect to the goal's funding ratio, our analysis reveals that the optimal portfolio's mean and variance can be geometrically represented in a manner similar to the well-known *capital asset line* and *tangency portfolio* from classical mean-variance theory.

We investigate numerically how the optimal investment strategies, distribution of goal fundedness, and consumption thresholds change depending on the client's goal and income stream. We consider two types of client's preferences towards goals. The first type includes *flexible* goals, i.e., for which the client has linear preferences and thus extracts value even if the goals are only partially fulfilled at their deadlines. The second type includes *all-or-nothing* goals, where the client extracts full benefit only if the goal is entirely fulfilled, and zero benefit otherwise. Flexible goals include, for instance, financing a vacation. Even if her plan is a vacation which costs an amount of $20,000 five years from now, she might still extract a benefit from a cheaper vacation of, say, $10,000. Examples of all-no-nothing goals include upgrades and renovations. For instance, a client whose goal is to upgrade her car to electric will either be able to implement the upgrade or needs to forfeit her goal if she does not have the required budget at the due date. Similarly, if a client intends to install a new heating system a year from now but does not manage to attain the appropriate budget to cover labor and material costs for this renovation, she would give up on her goal in one year.

Regardless of her preferences, we find that the client is risk-seeking and chooses a portfolio with high volatility if her current wealth is below the critical level for the upcoming goal. Near the goal deadline, such critical level coincides with the goal amount. Prior to the deadline, it is discounted by the client's income and interest rate, and corresponds to the minimum wealth amount which guarantees fulfillment of future goals with certainty. Compared to a client with flexible goals, an all-or-nothing client takes higher risk if she is underfunded.

Our comparative statics analysis highlights the differences in investment strategies and goal allocation for an all-or-nothing client compared to a client with linear preferences. Suppose we leave aside discounting and adjustments for goal deadlines. Both types of clients would focus on the goal with the highest cost-adjusted importance, i.e., with the highest ratio between importance weight and amount, if financial resources are low. A client with linear preferences, when endowed with sufficient resources, would invest towards multiple goals based on their cost-adjusted importance. On the other hand, an all-or-nothing client aims to allocate resources in a manner that increases the total probability of attaining multiple goals by their deadlines. Unlike a client with linear preferences whose marginal benefit from increasing the funding ratio of a goal is constant, an all-or-nothing client would extract nonzero benefit only if the goal is fully funded.

We analyze how the consumption-saving tradeoff depends on goal characteristics and income stream. For a client with flexible goals, we find that a higher income stream simplifies the fulfillment of future goals with high cost-adjusted significance, thereby allowing a larger consumption allocation for upcoming goals. Conversely, for goals with low cost-adjusted importance, i.e., those whose amounts are unlikely to be fully met by their deadline, this dynamic shifts. If the terms of such a goal are relaxed, either by reducing its amount or extending its deadline, the presence of a more substantial income stream might shift the strategy from foregoing the goal to at least partially funding it. As a result, the optimal strategy favors saving for future goals as they become attractive, rather than consuming towards the current goal.

1.1 Related Literature

Our paper contributes to the stream of literature on goal-based wealth management. Deguest and Martellini [2021] provide a comprehensive treatment of operational and implementation challenges of a goal-based investment approached faced by financial advisors.

Our paper is related to Das et al. [2018], Das et al. [2020], who develop a continuous time model of investment, where the optimal trading strategy is chosen to maximize the probability of achieving a single client's goal by a specified deadline. Das et al. [2022] extend the framework in their earlier works to handle multiple competing goals. In their model, goals are managed using portfolios on the efficient frontier. The client provides cost and utility preferences for different levels of goal fulfillment, ranging from full goal attainment to forgoing the goal. While their approach involves discretizing both the portfolio investment strategies and wealth grid, our method characterizes the solution to the stochastic control problem in continuous time. We establish general properties satisfied by the optimal funding ratio and consumption threshold, only resorting to discretization for numerical solutions of the HJB PDE.

Dixon and Alperin [2021] develop a discrete time inverse reinforcement learning framework for goal-based wealth management. Their framework does not feature multiple competing goals, but rather the investor targets a single goal by making investments in the portfolio when he is employed, and then draws from the account when he is retired.

The importance of goal setting has also been highlighted in the study of Gargano and Rossi [2023]. They show that setting goal for specific objectives results in a monthly savings increase by 90% for the average user. They also find that the feasibility of a goal deadline has an important impact on the effectiveness of goal setting; in particular, users who set long-term deadlines are less likely to achieve their goals compared to those who set short-term deadlines.

More broadly, our paper belongs to the stream of literature on asset and liability management (Mulvey and Vladimirou [2012], Dempster et al. [2003], Consigli and Dempster [1998], and Dempster and Medova). These studies utilitize stochastic dyanamic programming to construct the optimal strategy, and their methodologies are not designed to handle multiple goals as in our setting. Different from these studies, our solution technique is based on a mix of static and dynamic optimization. The optimal funding ratios are obtained sequentially at each goal deadline by solving a static optimization problem, whereas the optimal investment strategy in the risky assets is recovered by solving the HJB equation between consecutive goal deadlines.

Cvitanic et al. [2020] propose an objective criterion, which consists of a weighted average of probabilities of achieving target wealth levels and specified loss levels. In their study, all targets have the same deadline and their optimization problem is related to multi-objective decision making. By contrast, in our study, the goals have sequential deadlines, and the key trade-off faced is between immediate consumption and savings for future expenses. Dai et al. [2020] introduce a novel definition of viscosity solution tailored for addressing nonconcave utility maximization problems. Their framework accommodates terminal utility functions that can be either continuous or discontinuous, including all-or-nothing goals. However, their methodology is not directly applicable to our setting, because it does not account for state-dependent constraints and continuous income streams, both of which are integral component of our setting (see also Remark 3.2 for a more in-depth discussion).

The paper proceeds as follows. We present the market model and the client's control problem in Section 2; a complete solution to the SGBWM problem for linear preferences in Section 3; the generalization of our framework to nonlinear funding preferences in Section 4; the comparative statics with respect to client's input in Section 5; run times of the numerical procedure and an analysis of the computational complexity in Section 6; and concluding remarks in Section 7. All technical proofs are delegated to appendices.

2 Model Setup

In this section, we introduce the dynamic multi-goal optimization framework. We specify the dynamics of underlying securities in Section 2.1, the client's inputs in Section 2.2, and the set of admissible controls in Section 2.3.

Throughout the paper, decision variables which are treated as random variables or stochastic processes will be denoted with bold symbols, to distinguish them from the values they can take. All vectors are column vectors by default, unless explicitly stated otherwise. We use $\top$ to denote the transpose operator.

2.1 Market Dynamics

The market consists of N risky assets $S = (S_1, \ldots, S_N)^\top$ and a risk-free asset accruing interest at the rate $r > 0$. The drift and volatility rate processes are driven by a d-dimensional *stochastic factor* process $Y = (Y_1, \ldots, Y_d)^\top$, whose dynamics is given by $Y_0 = y_0 \in \mathbb{R}^d$, and

$$dY_t = \mu_Y(Y_t)dt + \Sigma_Y(Y_t)dW_t, \tag{2.1}$$

where W is an N'-dimensional standard Brownian motion. The drift coefficient $\mu_Y : \mathbb{R}^d \to \mathbb{R}^d$ and diffusion coefficient $\Sigma_Y : \mathbb{R}^d \to \mathbb{R}^{d\times N'}$ are assumed to be Lipschitz continuous.

The dynamics of the investment asset, conditioned on the factor process Y, follows a geometric Brownian motion:

$$dS_{i,t} = S_{i,t}(\mu_i(Y_t)dt + \sigma_i(Y_t)dW_t), \quad i = 1, \ldots, N, \tag{2.2}$$

where $\sigma_i(\cdot)$ is a row vector in $\mathbb{R}^{N'}$. We use $\mathcal{F}_t := \sigma(W_u; \quad u \leq t)$ to denote the natural filtration.

For each $y \in \mathbb{R}^d$, we write $\mu(y) = (\mu_1(y), \ldots, \mu_N(y))^\top$ and $\Sigma(y) = (\sigma_1(y), \ldots, \sigma_N(y))^\top$ for the vector of drift coefficients and the N-by-N' volatility matrix, respectively. We assume $N' \geq N$ and require the covariance matrix $C(y) = \Sigma\Sigma^\top(y)$ to be invertible, which immediately implies that the market is arbitrage-free and thus ensures the existence of a market price of risk vector

$$\Lambda(Y_t) := \Sigma^\top(Y_t)(\Sigma\Sigma^\top(Y_t))^{-1}\left(\mu(Y_t) - r\mathbb{1}_N\right). \tag{2.3}$$

We impose that the coefficients $\mu(\cdot)$ and $\Sigma(\cdot)$ are bounded Lipschitz. The Lipschitz property of μ_Y, Σ_Y and the bounded Lipschitz property of μ, Σ imply the existence of a unique strong solution to the coupled system of SDEs (2.1)-(2.2); see, for instance, Pham [2009].

Because the drift $\mu(\cdot)$ and volatility $\sigma(\cdot)$ are both functions of the continuous factor process Y_t, the market price of risk process $\Lambda(Y_t)$ is predictable in our model.

2.2 Client Information

The client specifies the following input:

- Initial capital x_0: amount of capital that the client deposits into her portfolio at time $t = 0$.

- Contribution rate $I(\cdot) \geq 0$: this function determines the infusion into the client's portfolio at t based on the realization of the factor process Y_t. In particular, $I(Y_t)$ represents the infusion amount to her portfolio at time t.

- Goals. Each of the $M \geq 1$ goals is uniquely characterized by the following triple (G_k, T_k, α_k), $k = 1, 2, \ldots, M$:

  - T_k: the deadline of the k-th goal.
  - $G_k(\cdot) > 0$: this function represents the dollar amount of the k-th goal, which depends on the realization of the factor process at the goal deadline T_k. Specifically, $G_k(Y_{T_k})$ denotes the amount of the k-th goal at its deadline T_k.
  - $\alpha_k \geq 0$: the *importance weight* of the k-th goal. Collectively, we require the weights to add up to one, i.e., $\sum_{k=1}^{M} \alpha_k = 1$.

  Throughout the paper, we refer to the quantity $\alpha_k / G_k(\cdot)$ as the *cost-adjusted importance* of the k-th goal. We assume that there are no concurrent goals, i.e., any two goals have different deadlines and we can sort them in ascending order $0 < T_1 < T_2 < \ldots < T_M$.[3]

- Maximal portfolio volatility $c(\cdot) \geq 0$: this function sets the upper bound on the volatility of any investment portfolio for a given realization of the factor process. In particular, $c(Y_t)$ represents the maximum portfolio volatility at time t.

[3]Concurrent goals can be approximated to a high level of accuracy by making their deadlines very close to each other.

We assume that the functions $I(\cdot), c(\cdot)$ are Lipschitz continuous; $G_k(\cdot)$'s are bounded; and $c(\cdot) > 0$ is bounded away from zero. Observe that the contribution rate, goal amount, and maximal portfolio volatility are stochastic processes because they depend on the factor process Y_t.

2.3 Admissible Controls

We denote by X_t the *wealth* process, and by $\boldsymbol{\pi}_{i,t}$ the fraction of total wealth invested in the i-th risky asset at time t. This implies that $1-\sum_{i=1}^{N} \boldsymbol{\pi}_{i,t}$ is the fraction of wealth invested in the risk-free asset. In-between two goal deadlines, the dynamics of the wealth process is governed by the following stochastic differential equation (SDE)

$$dX_t = I(Y_t)dt + rX_t \left(1 - \sum_{i=1}^{N} \boldsymbol{\pi}_{i,t}\right) dt + \sum_{i=1}^{N} X_t \boldsymbol{\pi}_{i,t}(\mu_i(Y_t)dt + \sigma_i(Y_t)dW_t),$$

i.e., at time t the wealth accrues external income at a rate $I(Y_t)$; the portion of wealth invested in the risk-free asset accumulates interest at a rate r, and the portion of wealth invested in the risky asset is exposed to market risk modeled through a Brownian motion.

We write $\boldsymbol{\pi}_t = (\boldsymbol{\pi}_{1,t}, \ldots, \boldsymbol{\pi}_{N,t})$ as a row vector and denote by $\mathbb{1}_N$ an N-dimensional column vector of all entries equal to 1. The above equation can also be written compactly as

$$dX_t = [I(Y_t) + rX_t + X_t\boldsymbol{\pi}_t(\mu(Y_t) - r\mathbb{1}_N)]\, dt + X_t\boldsymbol{\pi}_t\Sigma(Y_t)dW_t,$$

At time T_k, the k-th goal is consumed; that is, an amount $G_k(Y_{T_k})\boldsymbol{\theta}_k$ of wealth is withdrawn by the client, where $\boldsymbol{\theta}_k$ is a $[0,1]$-valued, $\mathcal{F}_{T_k}$-measurable random variable. We interpret $\boldsymbol{\theta}_k$ as the *funding ratio* for the k-th goal. If $\boldsymbol{\theta}_k = 1$, the k-th goal is fully funded; if $\boldsymbol{\theta}_k < 1$, the k-th goal is partially funded with funding ratio $\boldsymbol{\theta}_k$. Withdrawal leads to a jump in the wealth process at each goal deadline:

$$X_{T_k} = X_{T_k-} - G_k(Y_{T_k})\boldsymbol{\theta}_k.$$

Accounting for all cash flows, including those in-between goal deadlines and those at each deadline, the wealth process follows the dynamics

$$X_t = x_0 + \int_0^t I(Y_s)ds + \int_0^t \left[rX_s + X_s\boldsymbol{\pi}_s(\mu(Y_s) - r\mathbb{1}_N)\right] ds + \int_0^t X_s\boldsymbol{\pi}_s\Sigma(Y_s)dW_s$$
$$- \sum_{k=1}^{M} 1_{\{T_k \le t\}} G_k(Y_{T_k})\boldsymbol{\theta}_k. \tag{2.4}$$

We require $(\boldsymbol{\pi}, \boldsymbol{\theta})$ to be such that the controlled wealth process stays non-negative at all times. In addition, we prohibit short-selling and borrowing, and require the portfolio volatility to be bounded. Define $\Delta := \left\{\pi \in \mathbb{R}_+^N : \pi\mathbb{1}_N \le 1\right\}$ to be the N-dimensional simplex. Then $\boldsymbol{\pi}_t$ takes values in the convex set $\Delta_c(Y_t)$, where $\Delta_c(y) := \Delta \cap \{\pi \in \mathbb{R}_+^N : ||\pi\Sigma(y)|| \le c(y)\}$, $||z||$ denotes the l_2 norm of the vector z, and $\mathbb{R}_+$ denotes the nonnegative real line.[4] Note that the boundedness of $\boldsymbol{\pi}$ automatically guarantees that the diffusion will vanish when $X_t = 0$, so that between goal deadlines the controlled wealth process will not diffuse to bankruptcy. Throughout the paper and whenever needed for clarity, we use $X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}$ and $Y^{t,y}$ to denote the controlled wealth process and factor process with initial condition $(X_t, Y_t) = (x, y)$.

Definition 2.1. *The set of admissible controls for the problem starting at time t with wealth x and factor value y, denoted by $\mathcal{A}_t(x, y)$, consists of the pairs $(\boldsymbol{\pi}, \boldsymbol{\theta})$ where $\boldsymbol{\pi}$ is a progressively measurable, $\Delta_c(Y^{t,y})$-valued process, and $\boldsymbol{\theta}$ is a random vector whose k-th component is $\mathcal{F}_{T_k}$-measurable, $[0, 1]$-valued, and satisfies $G_k(Y_{t_k}^{t,y})\boldsymbol{\theta}_k \le X_{T_k-}^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}$ for all k such that $T_k > t$.*

Whenever we consider the special case of the model without the factor process, we will omit the dependence of the various quantities on y.

3 SGBWM for Linear Preferences

For a fixed $I(\cdot)$, the objective of the robo-advisor is to choose the asset allocation $\boldsymbol{\pi} = (\boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_N)$ and funding ratio $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M)$ to maximize the expected weighted goal fundedness, where the fundedness of each goal is weighted by its importance weight, i.e.,

$$\max_{(\boldsymbol{\pi},\boldsymbol{\theta})} \mathbb{E}\left[\sum_{k=1}^{M} \alpha_k\boldsymbol{\theta}_k \middle| X_0 = x_0, Y_0 = y_0\right], \tag{3.1}$$

[4]The model can also be extended to incorporate user-specified maximal fraction in risky assets, i.e. $\pi\mathbb{1}_N \le \pi_{\max}$, similar to what is done by the robo-advisor firm Betterment.

over the set of strategies $(\boldsymbol{\pi}, \boldsymbol{\theta}) \in \mathcal{A}_0(x_0, y_0)$.

Remark 3.1. *Maximizing the weighted expected fundedness is equivalent to minimizing the weighted expected shortfall of fundedness:* $\sum_{k=1}^{M} \alpha_k \mathbb{E}[1 - \boldsymbol{\theta}_k]$*, where we recall that* $\boldsymbol{\theta}_k$ *is* $[0,1]$*-valued.*

The dynamic optimization problem 3.1 under the self-financing condition (2.4)is a stochastic control problem which can be solved backward in time via dynamic programming. We characterize the value function of the control problem in Section 3.1. We provide a reformulation of the HJB equation, which highlights connections to minimum variance hypersurfaces from mean-variance analysis in Section 3.2. We solve the funding ratio problem faced by the client at each goal deadline in Section 3.3.

3.1 Characterization of the Value Function

The value function of the control problem (3.1) is defined as

$$V(t,x,y) := \sup_{(\boldsymbol{\pi},\boldsymbol{\theta}) \in \mathcal{A}_t(x,y)} \mathbb{E}_{t,x,y}\left[\sum_{k=1}^{M} \alpha_k \boldsymbol{\theta}_k 1_{\{T_k > t\}}\right], \quad (t,x,y) \in [0,T] \times \mathbb{R}_+ \times \mathbb{R}^d, \tag{3.2}$$

where $\mathbb{E}_{t,x,y}[\cdot]$ is a shorthand notation for the conditional expectation $\mathbb{E}[\cdot | X_t = x, Y_t = y]$. It follows directly from the definition that $0 \leq V \leq \sum_{k=1}^{M} \alpha_k 1_{\{T_k > t\}} \leq 1$, and that V is increasing in x. We also have the following concavity result:

Lemma 3.1. *$x \mapsto V(t,x,y)$ is concave for each t and y.*

Proof. This lemma follows immediately from Theorem 4.1 (e) by setting $f_k(\theta_k) = \theta_k$. See Appendix F for the proof of Theorem 4.1. □

Being a bounded concave function, $x \mapsto V(t,x,y)$ is continuous, and possesses left and right derivatives everywhere on the strictly positive real line.

At any time t and for any value y of the factor process, there exists a *safe level* $\mathfrak{s}(t,y)$ for our optimization problem (3.2). This is a level of wealth at which all remaining future goals are fully funded with probability one. To identify the safe level, we begin by considering the safe level process in hindsight $\tilde{s}_t$, i.e., if the client were to know the realization of the factor process Y_s, $s \geq t$. The process $\tilde{s}_t$ is non-adapted, and can be identified recursively via the following procedure: $\tilde{\mathfrak{s}}_{T_M} = 0$ and, for $k = M, M-1, \dots, 1$,

$$\frac{d\tilde{\mathfrak{s}}_t}{dt} = I(Y_t) + r\tilde{\mathfrak{s}}_t, \quad t \in (T_{k-1}, T_k), \quad \tilde{\mathfrak{s}}_{T_k-} = \tilde{\mathfrak{s}}_{T_k} + G_k(Y_{T_k}),$$

The above equation is a pathwise ordinary differential equation (i.e., for each realization of the Y process) whose solution is given by the stochastic process

$$\tilde{\mathfrak{s}}_t = e^{-r(T_k-t)}(\tilde{\mathfrak{s}}_{T_k} + G_k(Y_{T_k})) - \int_t^{T_k} e^{-r(T_k-s)} I(Y_s) ds.$$

It is easily seen that $\tilde{\mathfrak{s}}_t$ is right continuous and satisfies $\tilde{\mathfrak{s}}_t \leq \sum_{k:T_k>t} G_k(Y_{T_k})$. We then define the safe level at time t with factor value y as the supremum of the support for the conditional law $\mathcal{L}$ of $\tilde{\mathfrak{s}}$ given y, i.e., $\mathfrak{s}(t,y) = \sup\left\{supp\left(\mathcal{L}(\tilde{\mathfrak{s}}_t \middle| Y_t = y)\right)\right\}$. Observe that $\mathfrak{s}(t,y)$ is bounded from above because $supp\left(\mathcal{L}(\tilde{\mathfrak{s}}_t \middle| Y_t = y)\right)$ is upper bounded, which is guaranteed by our assumption that $G_k(\cdot)$ is bounded for all k. As we demonstrate later, the safe level plays an important role in understanding the optimal investment and goal funding strategies and their sensitivities to model parameters.

Similar to the safe level process in hindsight, the value function $V(t,x,y)$ can also be computed backward in time. At the last goal maturity, we have the terminal condition $V(T_M, x, y) = 0$, $(x,y) \in \mathbb{R}_+ \times \mathbb{R}^d$. Suppose $V(t,x,y)$ has already been computed for $t \geq T_k$ for some $k \in \{1, \dots, M\}$. We can then compute $V(t,x,y)$, for $t \in [T_{k-1}, T_k)$, using the following procedure:

- **Optimal Funding Ratio at Goal Deadlines.** We first determine $V(T_k-, x, y) = \lim_{t \nearrow T_k} V(t,x,y)$ by solving the static constrained optimization problem:

$$V(T_k-, x, y) = \sup_{\theta_k} \left\{ \alpha_k \theta_k + V\left(T_k, x - G_k(y)\theta_k, y\right) : \theta_k \in [0,1], G_k(y)\theta_k \leq x \right\}. \quad (3.3)$$

  That is, at each goal maturity T_k, one needs to decide how much wealth to allocate to the goal due at T_k and how much to save for future goal liabilities. We refer to the above problem as static, because the optimization is conducted at a fixed instant in time, which corresponds to the goal deadline. We remark that $V(T_k-, x, y)$ is generally a non-smooth function of x. For instance, the solution of the above optimization problem (3.3) at the last goal deadline is $V(T_M-, x, y) = \alpha_M(1 \wedge \frac{x}{G_M(y)})$.

- **Value Function Between Goal Deadlines.** Using heuristic arguments first, we expect the value function to be the solution of the following HJB equation: for $(t, x, y) \in [T_{k-1}, T_k) \times \mathbb{R}_+ \times \mathbb{R}^d$,

$$V_t + \sup_{\pi \in \Delta_c(y)} \left\{ V_x(I(y) + rx + x\pi(\mu(y) - r\mathbb{1}_N)) + x\pi\Sigma\Sigma_Y^\top(y)V_{xy} \right.$$
$$\left. + \frac{1}{2}V_{xx}x^2\|\pi\Sigma(y)\|^2 + \mu_Y^\top(y)V_y + \frac{1}{2}\operatorname{Tr}(\Sigma_Y\Sigma_Y^\top(y)V_{yy}) \right\} = 0 \tag{3.4}$$

with terminal condition $V(T_k-, x, y)$. In the above expression, Tr denotes the trace operator, V_t is the partial derivative with respect to t, V_x is the partial derivative with respect to x, $V_y = (V_{y_1}, \ldots, V_{y_d})^\top$ is the y-gradient vector, $V_{xy} = (V_{xy_1}, \ldots, V_{xy_d})^\top$ is the vector of mixed derivatives with respect to x and the vector y, V_{xx} is the second derivative with respect to x, and V_{yy} is the Hessian matrix in y.

The above procedure is made rigorous in Proposition 3.1. Therein, for each interval $[T_{k-1}, T_k)$ we characterize the value function (3.2) as the unique viscosity solution of the HJB equation (3.4).[5] Notice that although a unique viscosity solution exists on each subinterval $[T_{k-1}, T_k)$, this fact *does not imply* the existence and uniqueness of a unique viscosity solution over the entire interval $[0, T_M]$. The reason is that, in each subinterval $[T_{k-1}, T_k)$, the terminal condition of the PDE corresponds to the function V_{T_k-} as given in (3.3), rather than V_{T_k}. Consequently, one cannot paste the viscosity solution obtained within the subintervals $[T_{k-1}, T_k]$ for $k = 1, \ldots, M$.

While we do not exclude the existence of a stronger notion of solution to the HJB-PDE, this characterization is not needed for our purposes. Our objective is to illustrate the properties of the portfolio allocation strategy $\boldsymbol{\pi}$ and goal funding ratio $\boldsymbol{\theta}$, and a solution in the viscosity sense is sufficient for our numerical scheme.[6]

To avoid the complexity of dealing with the state constraint $X_t \geq 0$ in the viscosity characterization, in Proposition 3.1 and Theorem 4.1 (a)-(d) we extend the domain of our control problem (4.2) to $\overline{\mathbf{S}} := [0, T_M] \times \mathbb{R} \times \mathbb{R}^d$ by allowing negative initial wealth but modifying the wealth dynamics (2.3) to

$$dX_t = \left[I(Y_t) + rX_t + X_t^+\boldsymbol{\pi}_t(\mu(Y_t) - r\mathbb{1}_N)\right]dt + X_t^+\boldsymbol{\pi}_t\Sigma(Y_t)dW_t, \tag{3.5}$$

[5] The viscosity solution is a weaker notion of solution to partial differential equations (PDE), which is not necessarily smooth in the whole domain. We refer to Crandall et al. [1992] for a comprehensive treatment of viscosity solutions, and to Jensen [1988] for pioneering uniqueness results for second-order equations.

[6] Barles and Souganidis [1991] showed that for a nonlinear parabolic PDE verifying the comparison principle, any monotone, stable and consistent numerical scheme converges to the unique viscosity solution.

and requiring $G_k(Y_{T_k})\boldsymbol{\theta}_k \leq X^+_{T_k-}$ for all k, where $x^+ = \max(0, x)$ denotes the positive part of x. Such an extension does not change the value functions in the original domain $\overline{\mathbf{S}}_+ := [0, T_M] \times \mathbb{R}_+ \times \mathbb{R}^d$, because the boundedness of $\boldsymbol{\pi}$ ensures that $X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}} \geq 0$ if $x \geq 0$.

Proposition 3.1. *For each $k \in \{1, \ldots, M\}$,*

$$g_k(x, y) := \sup_{\theta_k} \left\{ \alpha_k \theta_k + V\left(T_k, x - G_k(y)\theta_k, y\right) : \theta_k \in [0, 1], G_k(y)\theta_k \leq x^+ \right\}$$

is a continuous function of x and y that equals to $V(T_k-, x, y)$. Moreover, V is continuous on $\mathbf{S}_k := [T_{k-1}, T_k) \times \mathbb{R} \times \mathbb{R}^d$, and is the unique viscosity solution with polynomial growth to the nonlinear Cauchy problem:

$$-v_t + F(x, y, v_x, v_y, v_{xy}, v_{xx}, v_{yy}) = 0, \quad (t, x, y) \in \mathbf{S}_k, \tag{3.6}$$

$$v(T_k, x, y) = g_k(x, y), \quad (x, y) \in \mathbb{R} \times \mathbb{R}^d, \tag{3.7}$$

where $K := \sup_{(t,y) \in [0,T_M] \times \mathbb{R}^d} |\mathfrak{s}(t, y)| < \infty$, $\rho(x) := x^+ \wedge (2K - x)^+$, and

$$\begin{aligned} F(x, y, p, q, g, A, B) &:= -p(I(y) + rx) - \mu_Y^\top(y) q - \frac{1}{2} \operatorname{Tr}(\Sigma_Y \Sigma_Y^\top(y) B) \\ &- \sup_{\pi \in \Delta_c(y)} \left\{ p\rho(x)\pi(\mu(y) - r\mathbb{1}_N) + \rho(x)\pi\Sigma\Sigma_Y^\top(y) g + \frac{1}{2} A\rho^2(x) \|\pi\Sigma(y)\|^2 \right\}. \end{aligned} \tag{3.8}$$

Note that the HJB equation specified by (3.6)-(3.8) *reduces to equation* (3.4) *with terminal condition* (3.3) *in the restricted domain $[T_{k-1}, T_k) \times \mathbb{R}_+ \times \mathbb{R}^d$.*

Proof. This proposition is a special case of parts (a), (c), (d) of Theorem 4.1, proven in Appendix F. We set $f_k(\theta) = \theta$ for all $k = 1, \ldots, M$ therein. □

Observe that the value function V is not continuous in time at the goal deadlines. The viscosity solution $v(\cdot, x, y)$ is continuous at the goal deadline T_k, but its value $g_k(x, y)$ is equal to $V(T_k-, x, y)$, i.e., to the left limit of the value function and not to $V(T_k, x, y)$. The nonlinear PDE in the previous proposition does not admit an explicit solution. In Appendix A, we describe a procedure for computing V when the factor process Y is not present.

Remark 3.2. *Between each pair of consecutive goals, the sequential goal-based wealth management problem can be related to the stochastic control problem considered in Dai et al.*

[2020] if the value function just before the deadline of the next goal given in Eq. (F.1) *is regarded as the terminal utility function in Dai et al. [2020]. However, it is important to note that the methodology introduced in their paper is not directly applicable to our framework for the following reasons:*

- *State-dependent controls. Our set of admissible controls* $\boldsymbol{\pi}_t$*, depends on the current state* Y_t*, i.e.,* $\boldsymbol{\pi}_t \in \Delta_c(Y_t)$*. This is in contrast to their setting, where constraints on* π *are state-independent.*

- *Income stream. In our model, the income stream* $I(Y_t)$ *accrues continuously over time to the wealth process. Unlike their framework, the continuous income stream prevents us from expressing the wealth process dynamics in a geometric form.*

- *Assumptions on coefficients: In their model extension which includes stochastic factors (see their online Appendix), the drift and volatility coefficients of both the stock price and stochastic factor process are assumed to be twice continuously differentiable. Our analysis is conducted under the weaker regularity condition that drift and volatility coefficients of the stock price and stochastic factor process are, respectively, bounded Lipschitz and Lipschitz continuous.*

3.2 Relation to the Minimum Variance and Tangency Portfolios

In this section, we illustrate how the solution to the goal-based optimization problem can be connected with analogous concepts from mean-variance analysis, such as the capital asset line and the tangency portfolio.

To begin, observe that the HJB equation (3.4) involves a pointwise optimization over the set $\Delta_c(y) \subseteq \mathbb{R}^N$. However, the dependence on $\pi \in \Delta_c(y)$ is only via the triple

$$\Phi(\pi; y) := (\|\pi\Sigma(y)\|, r + \pi(\mu(y) - r\mathbb{1}_N), \pi\Sigma\Sigma_Y^\top(y)), \tag{3.9}$$

where the first two components are the standard deviation and mean of the portfolio return, and the third component is the covariance between the portfolio return and the factor process. Therefore, instead of optimizing over $\Delta_c(y)$, we can optimize over

$$\mathcal{P}_c(y) := \{\Phi(\pi; y) : \pi \in \Delta_c(y)\} \subseteq \mathbb{R}^{2+d}.$$

The set $\mathcal{P}_c(y)$ is referred to as the *opportunity set* given the factor value y. It lives in a subspace of $\mathbb{R}^{2+d}$ of dimension $2 + d'(y)$, where $d'(y) := \mathrm{rank}(\Sigma\Sigma_Y^\top(y)) \le \min(N, d)$.

Let us next introduce some notations. For each $y \in \mathbb{R}^d$, define a set valued map $\mathcal{S}_c(y) : \mathbb{R}^{1+d} \to 2^{\mathbb{R}}$ which maps a mean-covariance pair to the corresponding set of possible volatilities, i.e.,

$$\mathcal{S}_c(y)(m, \nu) := \{s \in \mathbb{R}_+ : (s, m, \nu) \in \mathcal{P}_c(y)\} \subseteq [0, c(y)], \tag{3.10}$$

where $\nu = (\nu_1, \ldots, \nu_d)$ is a row vector. The domain of $\mathcal{S}_c(y)$, denoted by $dom(\mathcal{S}_c(y))$, represents the set of feasible mean-covariance pairs. Define $s_{\min}(m, \nu; y, c) := \inf \mathcal{S}_c(y)(m, \nu)$ with the convention that $\inf \emptyset = \infty$. By Lemma C.1, $s_{\min}(\cdot; y, c)$ is a convex function, and $dom(\mathcal{S}_c(y)) = \{(m, \nu) \in \mathbb{R}^{1+d} : s_{\min}(m, \nu; y, c) < \infty\}$ is a convex set.

Proposition 3.2. *The HJB equation* (3.4) *can be rewritten as*

$$\begin{aligned} &V_t + I(y)V_x + \mu_Y^\top(y)V_y + xV_x\hat{m}(V_x, V_{xy}, xV_{xx}, y) + x\hat{\nu}(V_x, V_{xy}, xV_{xx}, y)V_{xy} \\ &\quad + \frac{1}{2}x^2 V_{xx}\hat{s}^2(V_x, V_{xy}, xV_{xx}, y) + \frac{1}{2}\mathrm{Tr}(\Sigma_Y\Sigma_Y^\top(y)V_{yy}) = 0, \end{aligned} \tag{3.11}$$

where $(\hat{s}, \hat{m}, \hat{\nu})(a, g, A, y)$ *is a pointwise optimizer of*

$$\mathcal{H}_c(a, g, A, y) = \sup_{(s,m,\nu)\in\mathcal{P}_c(y)} \left\{ am + \nu g + \frac{1}{2}As^2 \right\}, \quad (a, g, A, y) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d. \tag{3.12}$$

- *If* $A < 0$, *we have that*

$$\mathcal{H}_c(a, g, A, y) = Ah_c\left(\frac{a}{A}, \frac{g}{A}, y\right), \tag{3.13}$$

where

$$\begin{aligned} h_c(\gamma_1, \gamma_2, y) &= \inf_{(m,\nu)\in dom(\mathcal{S}_c(y))} \left\{ \gamma_1 m + \nu\gamma_2 + \frac{1}{2}s_{\min}^2(m, \nu; y, c) \right\} && (3.14) \\ &= \inf_{(s,m,\nu)\in\Gamma_c(y)} \left\{ \gamma_1 m + \nu\gamma_2 + \frac{1}{2}s^2 \right\}, \quad \gamma_1 \in \mathbb{R}, \gamma_2 \in \mathbb{R}^d, && (3.15) \end{aligned}$$

and $\Gamma_c(y) := \{(s_{\min}(m, \nu; y, c), m, \nu) : (m, \nu) \in dom(\mathcal{S}_c(y))\}$ *is a continuous hypersurface of dimension* $(1 + d'(y))$ *contained in* $\mathcal{P}_c(y)$.

- *If $A = 0$, we have that*

$$\mathcal{H}_c(a, g, A, y) = \sup_{(m,\nu)\in\mathcal{E}(dom(\mathcal{S}_c(y)))} \{am + \nu g\}, \tag{3.16}$$

where $\mathcal{E}(dom(\mathcal{S}_c(y)))$ denotes the set of extreme points of the $(1 + d'(y))$-dimensional convex set $dom(\mathcal{S}_c(y))$.

Proof. See Appendix C. □

Remark 3.3. *We refer to $\Gamma_c(y)$ as the minimum variance hypersurface which is the higher dimensional analogue of the familiar notion of minimum variance curve in modern portfolio theory. If $d'(y) = 0$, which occurs if there is no factor process or if the Brownian motions driving the factor process are independent from those driving the risky asset prices, then $\Gamma_c(y)$ reduces to a curve in the (s, m)-plane. Its upper branch is known as the efficient frontier. This is a straight line segment emanating from the risk-free portfolio $(0, r)$ and tangent to the left boundary of the risky-asset-only opportunity set (a.k.a. the Markowitz bullet). If short selling is allowed, the Markowitz bullet is a hyperbola. Eliminating short selling shrinks the risky-asset-only opportunity set and lowers the efficient frontier.*

We next provide a geometric interpretation for the optimizer in (3.15). Without a factor process, the term $\nu\gamma_2 = 0$. Dropping the y and γ_2 variables, we have $h_c(\gamma_1) = \inf_{(s,m)\in\Gamma_c} \left\{\gamma_1 m + \frac{1}{2}s^2\right\}$, with $\gamma_1 \in \mathbb{R}$. The unique optimizer for $h_c(\gamma_1)$ can be found by plotting Γ_c in the (s^2, m)-plane and examining the family of parallel lines with slope $-1/(2\gamma_1)$ that intersect or are tangent to Γ_c, where $\gamma_1 = 0$ corresponds to vertical lines. The optimizer is given by the point of intersection or tangency of the line having the smallest horizontal intercept, and the optimal value is equal to half of the horizontal intercept. If $\gamma_1 = 0$, the objective function collapses to only minimizing the portfolio variance, and thus the minimum variance portfolio is the risk-free portfolio $(0, r)$.

If the factor process is present, the optimizer becomes harder to visualize. However, the analogy is to plot $\Gamma_c(y)$ in the (s^2, m, ν)-space and examine the family of hyperplanes with normal vector $(1/2, \gamma_1, \gamma_2)$ that touch $\Gamma_c(y)$. The one with the smallest s^2-intercept is optimal in the sense that the point of intersection or tangency is the optimizer, and half of its s^2-intercept is the optimal value.

3.3 Optimal Funding Ratios

In this section, we characterize the optimal funding ratios at the goal deadlines. Let $\Theta_k(x, y)$ be the set of optimizers of the static optimization problem (3.3). Similar to the proof of the continuity of g_k in Proposition 3.1, Berge's maximum theorem also implies that $\Theta_k(x, y)$ is an upper hemicontinous set-valued map with non-empty compact values. To avoid potential non-uniqueness of the optimal funding ratio, we specify a tie breaking rule that favors early consumption. This amounts to choosing $\theta_k^*(x, y) := \max \Theta_k(x, y)$. By Theorem 17.30 in Aliprantis and Border [2006], $\theta_k^*(x, y) = \arg\max_{\theta \in \Theta_k(x,y)} \theta$ is upper semicontinuous, and thus, defines an optimal funding ratio in feedback form.

Let $x > 0$ and $y \in \mathbb{R}^d$ be fixed. Slater's condition obviously holds for the static optimization problem (3.3). Denote by $\partial_x^- V$ and $\partial_x^+ V$ the left and right derivative of V with respect to x, which exist everywhere by concavity. Any $\theta_k \in \Theta_k(x, y)$ is characterized by the generalized version of the Karush-Kuhn-Tucker (KKT) conditions[7]:

(Stationarity)

$$0 = -\alpha_k + vG_k(y) + s - t + \lambda G_k(y) \text{ for some } v \in [\partial_x^+ V, \partial_x^- V](T_k, x - G_k\theta_k, y);$$

(Primal feasibility) $0 \le \theta_k \le 1, \quad G_k(y)\theta_k \le x'$

(Dual feasibility) $s, t, \lambda \ge 0;$

(Complementary Slackness) $s(1 - \theta_k) = 0, \quad t\theta_k = 0, \quad \lambda(x - G_k(y)\theta_k) = 0.$

The KKT condition leads to the following characterization of θ_k^*.

Proposition 3.3. *The largest optimizer of the optimization problem specified by* (3.3) *is given in feedback form by*

$$\theta_k^*(x, y) = \inf \left\{ \theta \in [0, x/G_k(y)] : \partial_x^+ V(T_k, x - G_k(y)\theta, y) > \frac{\alpha_k}{G_k(y)} \right\} \wedge 1 \wedge \frac{x}{G_k(y)},$$

with the convention that $\inf \emptyset = \infty$*, where* $u \wedge w$ *denotes the minimum of* u *and* w*.*

Proof. It is obtained from Theorem 4.1 (f) by setting $f_k(\theta) = \theta$ for all $k = 1, \ldots, M$. □

Proposition 3.3 captures the fundamental tradeoff between allocating wealth towards upcoming goals versus saving for future goal liabilities. For each y, the quantity $\alpha_k/G_k(y)$

[7] We refer to Theorem 3.34 in Ruszczynski [2006] for details. We also remark that the KKT conditions are sufficient conditions for optimality.

is the marginal benefit of allocating wealth to the k-th goal, while $\partial_x^+ V(T_k, x - G_k(y)\theta_k, y)$ is the marginal benefit of reinvesting wealth to meet future goal liabilities, given that an amount of $G_k(y)\theta_k$ has already been allocated to the k-th goal. The optimal solution is to increase θ_k from zero until one of the following three conditions is met: (i) goal k is fully funded, i.e., $\theta_k = 1$, (ii) the budget constraint is tight, i.e., $\theta_k G_k(y) = x$, and (iii) the marginal benefit of funding the current goal is smaller than that of saving for future goal liabilities, i.e. $\alpha_k/G_k(y) < \partial_x^+ V(T_k, x - G_k(y)\theta_k, y)$.

We can further show that the optimal funding ratio $\theta_k^*(x, y)$ characterized by Proposition 3.3 is piecewise linear in x with at most two change points. The first change point (if it exists) determines when to switch from saving to consuming for the k-th goal; the second change point occurs when the k-th goal has just been fully funded.

Corollary 3.1. *Fix $y \in \mathbb{R}^d$. The largest optimizer $\theta_k^*(\cdot, y)$ is piecewise linear and given by*

$$\theta_k^*(x, y) = \begin{cases} 0, & \text{if } 0 \leq x \leq b_k(y), \\ 1 \wedge \frac{x - b_k(y)}{G_k(y)}, & \text{if } x > b_k(y), \end{cases}$$

where $b_k(y) = \sup\{x \geq 0 : \partial_x^+ V(T_k, x, y) > \alpha_k/G_k(y)\} \vee 0$ with the convention that $\sup \emptyset = -\infty$, and where $u \vee w$ denotes the maximum of u and w.

Proof. See Appendix B. □

In view of Corollary 3.1, we can interpret $b_k(\cdot)$ as the consumption threshold for the k-th goal, i.e, the level of wealth beyond which it is optimal to consume towards the k-th goal, given the realized factor process at the goal deadline. We can further analyze the relation between b_k and the client-dependent parameters. First, we expect the safe level to be an upper bound for this threshold: if the current wealth equals the safe level, there is no further benefit in saving extra units of wealth toward future liabilities, and thus it is optimal to raise the fundedness of the current goal. Second, it is optimal to allocate some wealth towards future goals if future income streams are small, and some of the upcoming goals have high cost-adjusted importance, for instance high importance weights or low amounts. The following proposition formalizes this intuition.

Proposition 3.4. *Fix $y \in \mathbb{R}^d$ and let $b_k(\cdot)$ be defined as in Corollary 3.1. The following statements hold.*

(i) $b_k(y) \leq \mathfrak{s}(T_k, y)$;

(ii) If $\alpha_k/G_k(y) \leq \max_{j>k} \mathbb{E}[\alpha_j/G_j(Y_{T_j}^{T_k,y})]$ and $\mathbb{E}[\int_{T_k}^{T_M} I(Y_s^{T_k,y})ds]$ is sufficiently small, then $b_k(y) > 0$, where we recall that $Y^{T_k,y}$ denotes the process Y starting at y at time T_k.

Proof. The statements are obtained from Theorem 4.1 (h) by setting $f_k(\theta) = \theta$ for all k. $\square$

4 SGBWM for General Preferences

In this section, we consider a broader class of preferences in funding ratios. Suppose the client specifies her satisfaction rate $u_{k,\theta} \in [0,1]$ over a small set of funding ratios θ for the k-th goal. The polar cases $u_{k,\theta} = 0$ and $u_{k,\theta} = 1$ indicate that the client is, respectively, fully dissatisfied or completely satisfied with the funding ratio θ. The wealth manager would then fit a smooth function f_k to the client's specified satisfaction rates u_k, so as to satisfy $f_k(\theta) = u_{k,\theta}$. With this type of preferences, we can capture situations where the marginal benefit from goal fundedness is non-constant. If $f_k(\theta) = \theta$, we recover the linear preference setting analyzed in Section 3. Formally, the optimization criterion is now given by

$$\max_{\boldsymbol{\pi},\boldsymbol{\theta}} \mathbb{E}\left[\sum_{k=1}^{M} \alpha_k f_k(\boldsymbol{\theta}_k)\right]. \tag{4.1}$$

We expect f_k to be an increasing function, because a higher funding ratio would bring more satisfaction to the client.

Remark 4.1. *Our extended goal based optimization framework nests the setting of Das et al. [2022] as a special case. In that paper, each goal can have a finite number of "partial goals", which each have specified costs c_k and corresponding satisfaction rates, u_k. This corresponds to making $f_k(\cdot)$ a right-continuous, increasing step function.*

4.1 Value Function, HJB Equations, and Optimal Funding Ratios

In this section, we generalize the results from previous sections (specifically Lemma 3.1, Proposition 3.1, Proposition 3.3, Corollary 3.1 and Proposition 3.4) to deal with a broader class of functions f_k's. As before, define the value function

$$V(t,x,y) := \sup_{(\boldsymbol{\pi},\boldsymbol{\theta})\in\mathcal{A}_t(x,y)} \mathbb{E}_{t,x,y}\left[\sum_{k=1}^{M} \alpha_k f_k(\boldsymbol{\theta}_k) 1_{\{T_k>t\}}\right] \tag{4.2}$$

on the extended domain $\overline{\mathbf{S}} = [0, T_M] \times \mathbb{R} \times \mathbb{R}^d$, where the stochastic factor process Y_t is specified by (2.1) and the wealth process X_t between consecutive goal deadlines is governed by the dynamics (3.5), and satisfies $X_{T_k} = X_{T_k-} - G_k(Y_{T_k})\boldsymbol{\theta}_k$ at each deadline T_k. For a locally bounded function $g : \mathcal{O} \to \mathbb{R}$, we denote by

$$g^*(z) := \limsup_{z' \to z, z' \in \mathcal{O}} g(z) \quad \text{and} \quad g_*(z) := \liminf_{z' \to z, z \in \mathcal{O}} g(z) \tag{4.3}$$

the upper and lower semicontinuous envelopes of g in $\mathcal{O}$.

Theorem 4.1. *Suppose $f_k : [0,1] \to [0,1]$ is continuous for all $k = 1, \dots, M$. Then for each $k \in \{1, \dots, M\}$, the following statements hold in the extended domain $\overline{\mathbf{S}}$:*

(a) The function

$$g_k(x, y) := \sup_{\theta_k} \{\alpha_k f_k(\theta_k) + V(T_k, x - G_k(y)\theta_k, y) : \theta_k \in [0,1], G_k(y)\theta_k \le x^+\} \tag{4.4}$$

is a continuous function of x and y.

(b) The largest maximizer $\theta_k^*(x, y)$ of the static optimization problem in part (a) exists, and is upper semicontinuous.

(c) Restricting V to $\mathbf{S}_k := [T_{k-1}, T_k) \times \mathbb{R} \times \mathbb{R}^d$, we have $V^*(T_k, x, y) = V_*(T_k, x, y) = g_k(x, y)$. In particular, $V(T_k-, x, y) = g_k(x, y)$.

(d) V is continuous on $\mathbf{S}_k$, and is the unique viscosity solution without polynomial growth to the nonlinear Cauchy problem:

$$-v_t + F(x, y, v_x, v_y, v_{xy}, v_{xx}, v_{yy}) = 0, \quad (t, x, y) \in \mathbf{S}_k, \tag{4.5}$$

$$v(T_k, x, y) = g_k(x, y), \quad (x, y) \in \mathbb{R} \times \mathbb{R}^d, \tag{4.6}$$

where the function F has been specified in Eq. (3.8).

Assume additionally that all f_k's are differentiable, concave and increasing. Then we further have the following claims in the original domain $\overline{\mathbf{S}}_+$:

(e) V is concave in x.

(f) $\theta_k^*(x, y) = \inf\left\{\theta \in [0, x/G_k(y)] : \partial_x^+ V(T_k, x - G_k(y)\theta, y) > \frac{\alpha_k}{G_k(y)} f_k'(\theta)\right\} \wedge 1 \wedge \frac{x}{G_k(y)}$.

(g) *Let* $b_k(y) := \sup\{x \geq 0 : \partial_x^+ V(T_k, x, y) > \alpha_k f_k'(0)/G_k(y)\} \vee 0$ *with the convention that* $\sup \emptyset = -\infty$. *Then* $\theta_k^*(x, y) = 0$ *if* $0 \leq x \leq b_k(y)$.

(h) *For any* $y \in \mathbb{R}^d$, *the consumption threshold* $b_k(y) \leq \mathfrak{s}(T_k, y)$. *Moreover,* $b_k(y) > 0$ *if* $\alpha_k f_k'(0)/G_k(y) \leq \max_{j>k} \mathbb{E}[\alpha_j f_j'(0)/G_j(Y_{T_j}^{T_k,y})]$ *and* $\mathbb{E}[\int_{T_k}^{T_M} I(Y_s^{T_k,y})ds]$ *is sufficiently small.*

Proof. See Appendix F. □

In the proof of part (d) of the above theorem, the main technical challenge is that the control constraint is factor-dependent. We refer the reader to Appendix E and F for the full proof, and here provide a brief overview of the main steps. We first provide an equivalent representation of the wealth dynamics whose coefficients are Lipschitz. We then prove the weak DPP for our specific control problem, leveraging results from Touzi [2013] and adapting them to deal with our portfolio volatility constrained setting. Using the weak DPP, we can then obtain the viscosity solution property of our value function.

Remark 4.2. *In Proposition 3.2, we discuss a methodology to reduce the complexity of evaluating the Hamiltonian. Our technique is designed to work with general preferences, possibly non-concave. A non-concave client's preference means the value function of the control problem may be convex in certain regions of the state space and concave in others, which leads to varying signs for the argument* A *in the function* $\mathcal{H}_c(a, g, A, y)$ *defined in* (3.12). *The case* $A \leq 0$ *has been treated in Proposition 3.2. We describe how to compute* $\mathcal{H}_c(a, g, A, y)$ *for* $A > 0$ *in Appendix D.*

5 Comparative Statics Analysis

We perform a sensitivity analysis to numerically assess. We specify the benchmark parameters used for our analysis in Section 5.1. We present and explain the results from our sensitivity analysis in Section 5.2.

5.1 Benchmark Parameter Specification and Client's Preferences

We consider the following benchmark parameter specification in our numerical analysis:

- Market parameters. We assume $d = 0$, i.e., no factor process. We estimate the drifts and covariance matrix of the assets using historical returns from the 20 year period between January 1998 to December 2017 for index funds representing U.S. bonds and U.S stocks. This yields the following table for the mean and covariance of returns (see also Das et al. [2020]):

$$\mu = \begin{pmatrix} 4.93\% \\ 8.86\% \end{pmatrix}, \quad C = \Sigma\Sigma^\top = \begin{pmatrix} 0.0017, -0.0021 \\ -0.0021, 0.0392 \end{pmatrix}.$$

We set the risk-free rate $r = 0.05\%$, which is of the same order of magnitude as the average 1M US Treasury rate for the period 1998-2017.

- Client parameters. The client has three goals with deadlines $(T_1, T_2, T_3) = (0.5, 2, 5)$ years and amounts $(G_1, G_2, G_3) =$ (\$30,000, \$10,000, \$250,000). We take the importance weights to be the same, i.e., $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$.[8] We set $I =$ \$1,000/month, and $x_0 =$ \$10,000. We also impose the constraint that the portfolio volatility does not exceed a given threshold, which we set to $c = 7.5\%$. This constraint reflects the attitude of robo-advisors to limit the risk exposure of a client's portfolio.

We consider two risk preference functions for the client. The first is the linear risk preference function, i.e., the client's utility is linear in the funding ratio of each goal. The second is the all-or-thing risk preference where the client's utility from a goal is one if such goal is fully funded, and zero otherwise. The all-or-nothing preference function is specified as $f_k(\theta) = 1_{\{\theta=1\}}$ which is discontinuous at $\theta = 1$ and thus violates the continuity assumption on the client's preference function in Theorem 4.1. Note that the purpose of this assumption is to guarantee the continuity of the terminal condition of the PDE. By regularizing the discontinuous terminal condition, the viscosity characterization can be applied. The smoothing is done by averaging the values of the terminal condition in a neighborhood of radius \$2,500 around the discontinuities. We set the upper boundary of the numerical domain for the finite difference method detailed in Appendix A to be $\ln(\tilde{x}_s + 2500)$, where $\tilde{x}_s$ is the flattened safe level defined in Appendix A.1.

Remark 5.1. *Das et al. [2022] restrict the investment strategy to a set of 15 portfolios on the efficient frontier. They also use a time grid with step size equal to one year. By contrast, we allow for 133 admissible investment portfolios, and the portfolio is rebalanced daily, i.e.,*

[8] These goals can be mapped, for instance, to a shorter-term goal of purchasing a car, a medium-term goal of financing a vacation, and a longer-term goal of making a downpayment for the purchase of a house.

each time step is $1/252$ *years. Hence, our numerical procedure sacrifices time complexity for the benefit of accuracy.*

We solve for value function and optimal strategies using the finite-difference PDE method, detailed in Appendix A, where we set the time step size $\Delta t = 1/252$ years and the space step size in the transformed domain $\Delta z = 0.01$. Additionally, we discretize Γ_c with a step size of $\Delta m = 0.05\%$, and consider 130 portfolios on Γ_c that are equally spaced in m, where the portfolio with the minimum m is the risk-free portfolio. Together with three corner points of the opportunity set, the total number of portfolios considered in the dynamic optimization problem is 133.

5.2 Results from Sensitivity Analysis

We conduct a comprehensive set of comparative statics, with the objective of measuring: (i) the sensitivity of the client's portfolio risk profile to changes in funding ratio preferences, goal importance weights, goal amounts, and time remaining till goal deadlines; (ii) the relation between optimal funding ratios and goal probabilities; (iii) the relationship between saving-consumption trade-off, goal characteristics, and the client's financial resources. In all plots, we measure wealth in thousand dollars, and time in years.

5.2.1 Value Functions

The all-or-nothing preferences are binary in nature: they offer full satisfaction to the investor when the goal amount is attained and zero satisfaction if not achieved by the deadline. This is clearly manifested in the form of the value function. As illustrated in Figure 1, the value function for an investor with all-or-nothing preferences presents a notably steeper staircase structure when compared to the case of linear preferences in funding ratios.

5.2.2 Expected Return and Volatility of Optimal Portfolio

We analyze how the optimal allocation strategy depends on the goal amounts near the goal deadlines. Figures 2 and 3 show that the client becomes more risk-seeking if her wealth is lower than the goal amount and the deadline is near. She then adjusts her portfolio to increase both its expected return and volatility, taking on more risk in order to achieve a higher funding ratio.

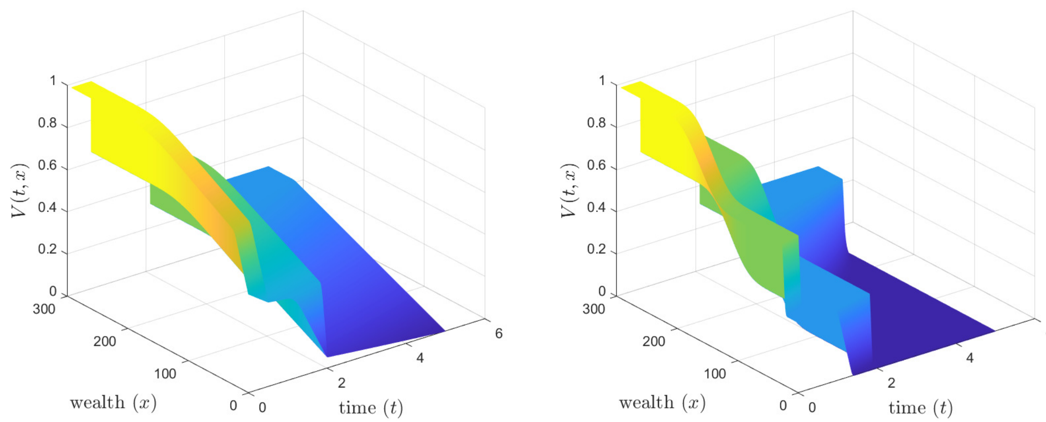

Figure 1: Left panel: Value function $V(t,x)$ for a client with linear funding preferences. Right panel: Value function $V(t,x)$ for all-or-nothing funding preferences.

The presence of "stripes" in the bottom left-hand corners of Figures 2 and 3 indicates that when the client's wealth is sufficiently close to or slightly higher than the goal amount near the deadline, she reduces the riskiness of her portfolio to ensure the goal is met with certainty. Figure 4 presents time slices of these stripes. At fixed points in time, as her wealth approaches levels close to the next goal's amount, the client transitions abruptly from a risky portfolio to a safer one to avoid the risk of falling short on the goal. This shift occurs earlier if the planning horizon is longer, i.e., the investment time is further from the goal deadline.

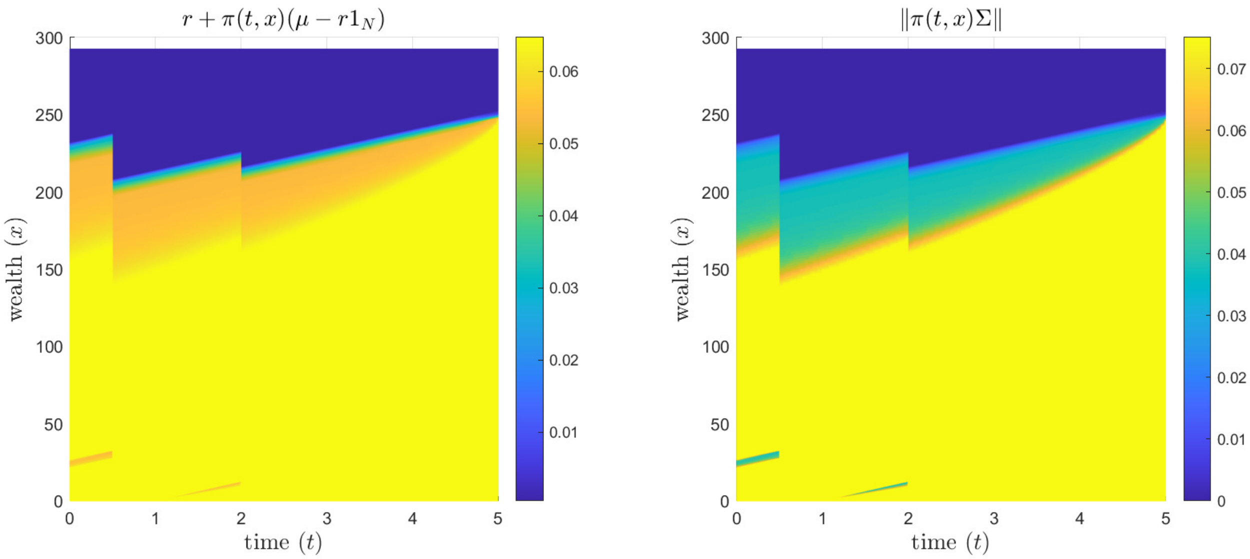

Figure 2: Left panel: Expected return of the optimal portfolio, given by $r+\pi(t,x)(\mu-r\mathbb{1}_N)$. Right panel: Volatility of the optimal portfolio, given by $\|\pi(t,x)\Sigma\| = \sqrt{\pi(t,x)^\top C\pi(t,x)}$. The client's preferences in the funding ratio are linear.

Additionally, the graphs in Figure 4 highlight that if the client's wealth is large enough to securely fund the first goal, she takes higher risks (the portfolio volatility constraint becomes binding) to accumulate enough expected returns to fund the remaining goals. If the client's wealth is sufficient to meet the two remaining goals (\$10,000+\$250,000) with high probability, she reduces the riskiness of her portfolio.

The client invests in a risk-free asset portfolio when her wealth reaches the safe level, which depends in a nonlinear fashion on various factors, including the investment time, goal times and deadlines, and current wealth amount. The shaded area defined by zig-zag lines at the top of each graph in Figures 2 and 3 illustrates this nonlinear relationship.

The investment patterns evidenced from Figure 5 highlight the fundamental differences between a client with linear preferences and one with all-or-nothing preferences. First, the dip in portfolio allocation is prompted at a comparatively lower wealth level for the investor with linear funding preferences. This is because the all-or-nothing investor is fully risk-seeking until her wealth is sufficiently close to the amount of the upcoming goal. Second, this dip is more pronounced for the all-or-nothing investor, who abruptly shifts to being highly risk-averse once she has accumulated enough wealth to meet the goal. This behavior resembles investors with S-shaped utility functions, as evidenced in Figure 5 with a higher

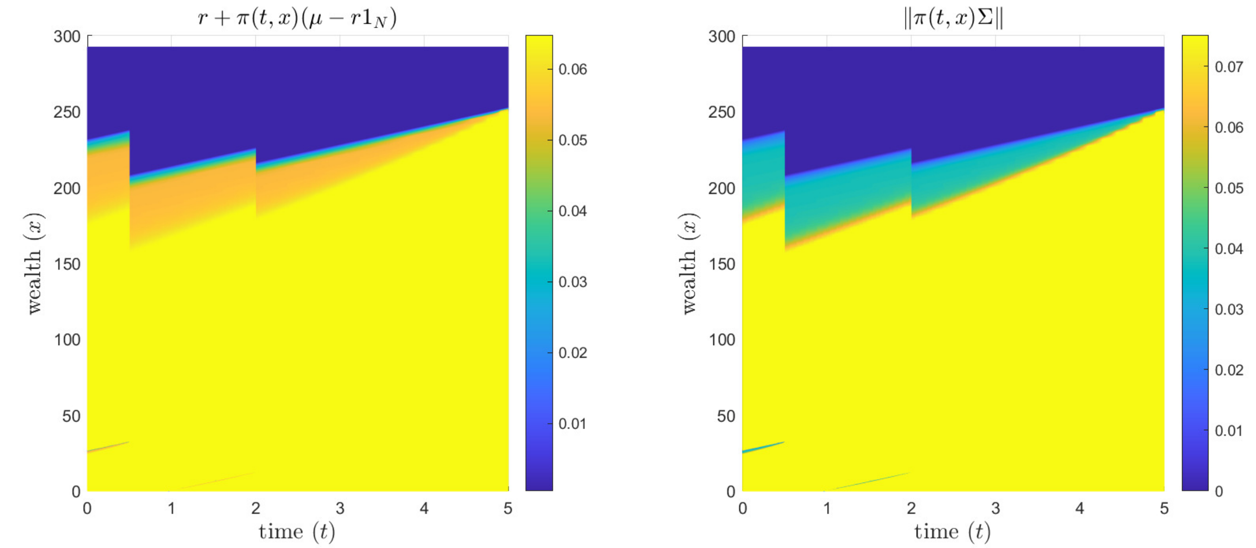

Figure 3: Left panel: Expected return of the optimal portfolio, given by $r+\pi(t,x)(\mu-r\mathbb{1}_N)$. Right panel: Volatility of the optimal portfolio, given by $\|\pi(t,x)\Sigma\| = \sqrt{\pi(t,x)^\top C\pi(t,x)}$. The client's preferences in the funding ratio are of the all-or-nothing type.

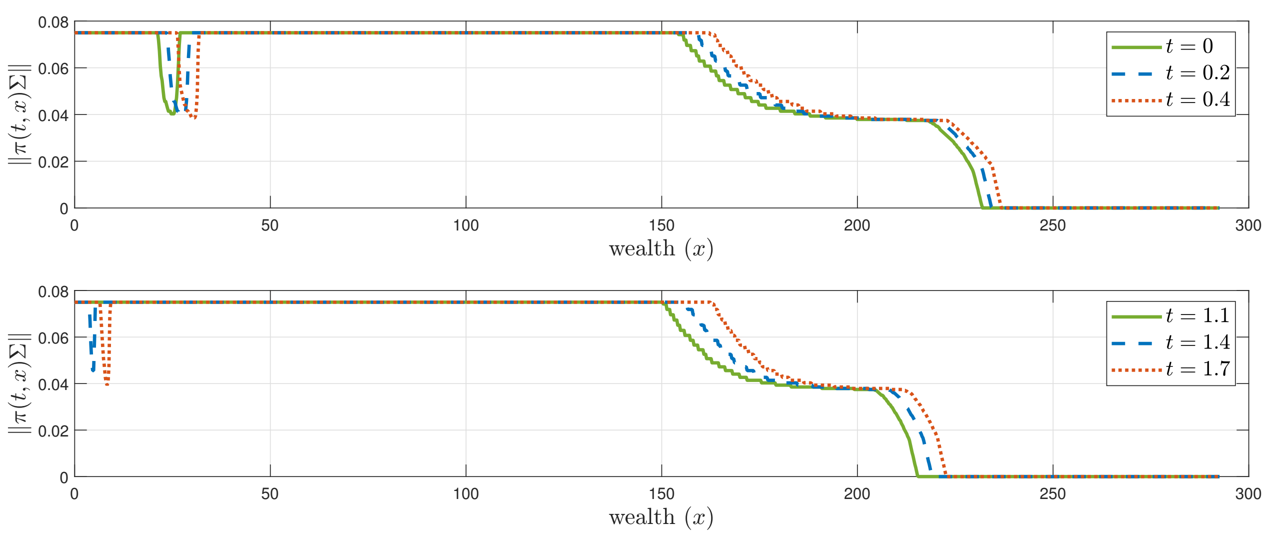

Figure 4: Top panel: Volatility of the optimal portfolio $\|\pi(t,x)\Sigma\| = \sqrt{\pi(t,x)^\top C\pi(t,x)}$ sliced at three different times prior to the first goal deadline. Bottom panel: Volatility of the optimal portfolio sliced at three different times between the first and second goal deadline. The client's preferences are linear in the funding ratio.

convexity of the value function around the kinks. These kinks correspond to dips in the portfolio volatility.

It is evident from the top and middle panels of Figure 5 that the investor takes high risk to ensure funding of remaining goals if her current wealth sufficiently covers the first goal but is inadequate for subsequent goals. Contrarily, the bottom panels reveal a lack of volatility dips in the investor's portfolio. Instead, we observe a sustained decrease in portfolio volatility once the safe wealth level is attained, capturing a guaranteed fulfillment of the third goal and the absence of further goals afterwards.

It is worth recalling that we have imposed a cap on the maximum portfolio volatility. Such a constraint becomes binding for both all-or-nothing and linear funding ratio preferences when the investor's wealth is either insufficient for the upcoming goal or adequately meets the first goal but falls short for the second. Absent this constraint, the all-or-nothing preferences would likely exhibit higher portfolio volatility in parameter regions where the wealth is either low or sufficient for the initial goal but not entirely for the subsequent one.

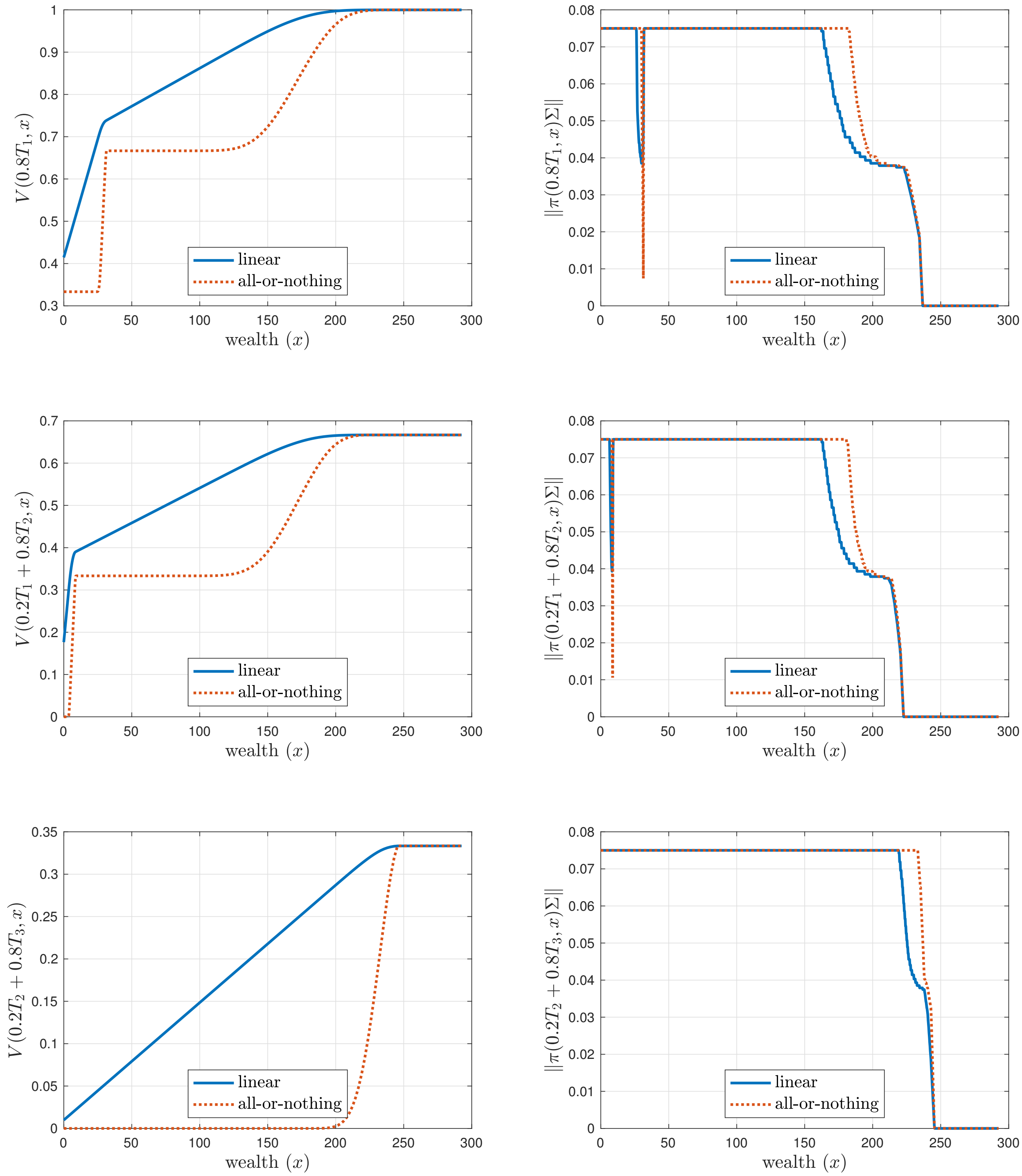

Figure 5: Left panels: Value functions $V(t,x)$ sliced at $t = 0.8T_1$ (top), $t = 0.2T_1 + 0.8T_2$ (middle), and $t = 0.2T_2 + 0.8T_3$ (bottom), for all-or-nothing and linear funding ratio preferences. Right panels: Portfolio volatility $\|\pi(t,x)\Sigma\| = \sqrt{\pi(t,x)^\top C \pi(t,x)}$ sliced at time $t = 0.8T_1$ (top), $t = 0.2T_1 + 0.8T_2$ (middle), and $t = 0.2T_2 + 0.8T_3$ (bottom), for all-or-nothing and linear funding ratio preferences.

5.2.3 Dependence of Portfolio Volatility on Importance Weights

In this section, we study how the risk profile of a client with linear preferences in funding ratio varies with changes in her goal importance weights.

Suppose, for instance, that the importance weight of goal "1" relative to goal "2" is increased. The benefit of funding goal "1" is then higher. As the deadline of goal "1" is approaching, we expect the client to reduce the risk profile of her portfolio. This investment behavior is well highlighted in the left panel of Figure 6.

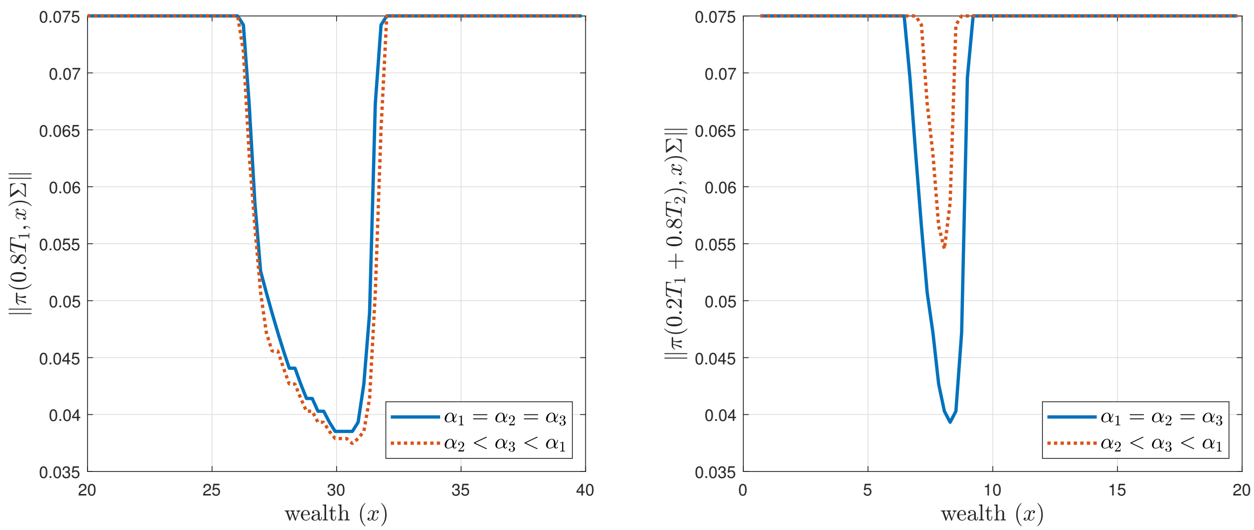

Figure 6: Volatility of the optimal portfolio for two goal importance weight profiles. The first profile is that of equal importance weights, i.e., $(\alpha_1, \alpha_2, \alpha_3) = (1/3, 1/3, 1/3)$. In the second profile, we increase the importance weight of goal "1" relative to goal "2", but leave the importance weight of goal "3" unchanged, and specifically set $(\alpha_1, \alpha_2, \alpha_3) = (7/12, 1/12, 1/3)$. Left panel: Portfolio volatility $\|\pi(t,x)\Sigma\| = \sqrt{\pi(t,x)^\top C\pi(t,x)}$ sliced at time $t = 0.8T_1$. Right panel: Portfolio volatility $\|\pi(t,x)\Sigma\|$ sliced at time $t = 0.2T_1 + 0.8T_2$ (i.e., close to the second goal deadline). We consider a client whose preferences in the funding ratio are linear.

Close to the deadline of goal "2", we expect the risk profile of her portfolio to be higher compared to the benchmark of equal importance weights. This investment pattern is visible from the right panel of Figure 6, which indicates that the client transfers wealth from safer to riskier assets as the importance weight of goal "2" is lower.

Taken together, these findings are consistent with the implications of the consumer behavior model developed by (Thaler [1985a] and Thaler [1985b]). As the importance weight of a goal increases, the optimal strategy is to reduce the portfolio risk profile to avoid falling short of fulfilling such goal.

5.2.4 Dependence of Funding Ratios on Goal Probabilities

In this section, we investigate how the distribution of goal funding ratios changes as we vary the financial resources (initial capital and income stream) of the client. Specifically, we vary both the client's initial capital, and the income stream added to the portfolio throughout the investment horizon.

Table 1 reports the mean and standard deviation of goal fundedness for a client whose preferences are linear in the funding ratios. If financial resources are very low (first row of the table), then the optimal strategy is to forego the first goal. Under these circumstances, even if the first goal has a higher cost-adjusted importance than the third goal, the deadline is too tight for the first goal to be sufficiently well funded.

x_0	I	$\mathbb{E}[\boldsymbol{\theta}_1]$	$\text{sd}(\boldsymbol{\theta}_1)$	$\mathbb{E}[\boldsymbol{\theta}_2]$	$\text{sd}(\boldsymbol{\theta}_2)$	$\mathbb{E}[\boldsymbol{\theta}_3]$	$\text{sd}(\boldsymbol{\theta}_3)$
\$2,000	\$100/mo.	0	0	0.4834	0.0385	0.0159	0.0012
\$2,000	\$500/mo.	0.1371	0.0065	0.9983	0.0077	0.0820	0.0066
\$2,000	\$1000/mo.	0.2720	0.0095	1	0	0.2022	0.0180
\$2,000	\$2000/mo.	0.4751	0.0155	1	0	0.4531	0.0418
\$10,000	\$100/mo.	0.0978	0.0186	0.9968	0.0121	0.0189	0.0026
\$10,000	\$500/mo.	0.4124	0.0208	0.9983	0.0077	0.0820	0.0066
\$10,000	\$1000/mo.	0.5473	0.0236	1	0	0.2022	0.0180
\$10,000	\$2000/mo.	0.7504	0.0294	1	0	0.4531	0.0418
\$50,000	\$100/mo.	1	0	1	0	0.0955	0.0217
\$50,000	\$500/mo.	1	0	1	0	0.2090	0.0302
\$50,000	\$1000/mo.	1	0	1	0	0.3508	0.0428
\$50,000	\$2000/mo.	1	0	1	0	0.6343	0.0700

Table 1: We simulate the optimally controlled wealth process, and solve for the optimal funding ratio on each sample path. We report estimated mean ($\mathbb{E}[\boldsymbol{\theta}_k]$) and standard deviation ($\text{sd}(\boldsymbol{\theta}_k)$) of the funding ratios ($\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3$) using 5,000 simulated sample paths. We consider a client whose preferences in the funding ratio are linear.

If financial resources are adequate, then the optimal strategy is to allocate the initial wealth to goals in decreasing order of marginal benefit. That means to first fund goal "2", then goal "1", and finally to increase the funding ratio of goal "3".

Consistent with intuition, for a fixed initial capital the funding ratio of each goal increases with I. A steady income stream over time can be directed towards goals as they draw closer to their deadlines.

Table 2 highlights the different behavior of an all-or-nothing client compared to a client with linear preferences. If financial resources are low, both types of clients would prioritize goal “2” because it has the highest marginal benefit. However, if financial resources are just enough to satisfy goal “1” with certainty, but not both goals “1” and “2”, then their allocation strategy would be different. A client with linear preferences would choose to fund goal “2”, and use the residual amount towards goal “1”. Because all goals have the same importance weight, such a client would attain 1/3 of the total benefit from fully funding the second goal, and extract partial benefit (between 0 and 1/3) from the first goal.

By contrast, a client with all-or-nothing preferences would prioritize goal “1”. If she were to forgo goal 1, she would permanently lose the 1/3 contribution to her total benefit coming from that goal. An all-or-nothing client would then gamble to raise enough wealth to satisfy both goals. In most sample paths, she would be able to satisfy either goal “1” or goal “2”, but by gambling she increases the total probability of satisfying both goal “1” and goal “2”. Even if the all-or-nothing client were to forego goal 1 because of an unfavorable realization of risk factors, the chances of funding goal 2, and thus of attaining 1/3 of the total benefit, would still be high. If she were not to take this risk, then she would be certain of not being able to satisfy both goals “1” and “2”, and thus only capture 1/3 of the total benefit. Under this risk-taking strategy, if the client has enough wealth to satisfy goal “1” by its deadline, she would fund this goal and be left with smaller financial resources to fund goal “2” with high probability. Clearly, the larger the amount of financial resources, and the higher the chance of satisfying both goals. These investment patterns can be seen by comparing the goal reaching probabilities corresponding to the rows of Tables 2 where $I = \$100$/mo. and $x_0 = \$30{,}000$, \$32,500, \$35,000.

If financial resources are sufficiently large to fully fund the first two goals, but still too low to meet the third goal with sufficient probability, the client invests to fund in full goals “1” and “2” because they are attainable. This investment profile is followed, for instance, by a client with monthly income \$1,000 and initial wealth equal to \$100,000.

5.2.5 Dependence of Consumption Threshold on Client Parameters

In this section, we analyze how the consumption threshold b_1 (defined in Corollary 3.1) of a client with linear funding ratio preferences changes with income stream, goal amounts, importance weights, and deadlines.

x_0	I	$\mathbb{P}(\theta_1 = 1)$	$\mathbb{P}(\theta_2 = 1)$	$\mathbb{P}(\theta_3 = 1)$
\$2,000	\$100/mo.	0	0	0
\$5,000	\$100/mo.	0	0.0174	0
\$20,000	\$100/mo.	0	1	0
\$30,000	\$100/mo.	0.1950	0.8064	0
\$32,500	\$100/mo.	0.7600	0.2894	0
\$35,000	\$100/mo.	0.9840	0.3898	0
\$200,000	\$100/mo.	1	1	0.3776
\$250,000	\$100/mo.	1	1	0.9690
\$2,000	\$500/mo.	0	1	0
\$20,000	\$500/mo.	0	1	0
\$22,500	\$500/mo.	0.0046	0.9976	0
\$25,000	\$500/mo.	0.2034	0.9204	0
\$27,500	\$500/mo.	0.8152	0.8862	0
\$30,000	\$500/mo.	1	0.9994	0
\$100,000	\$500/mo.	1	1	0
\$150,000	\$500/mo.	1	1	0.0574
\$200,000	\$500/mo.	1	1	0.6896
\$250,000	\$500/mo.	1	1	0.9994
\$2,000	\$1000/mo.	0	1	0
\$20,000	\$1000/mo.	0.001	1	0
\$22,500	\$1000/mo.	0.1432	1	0
\$25,000	\$1000/mo.	0.9780	1	0
\$30,000	\$1000/mo.	1	1	0
\$100,000	\$1000/mo.	1	1	0
\$150,000	\$1000/mo.	1	1	0.1832
\$175,000	\$1000/mo.	1	1	0.5742
\$200,000	\$1000/mo.	1	1	0.9510
\$2,000	\$2000/mo.	0	1	0
\$15,000	\$2000/mo.	0.0236	1	0
\$17,500	\$2000/mo.	0.5486	1	0
\$20,000	\$2000/mo.	0.9974	1	0
\$30,000	\$2000/mo.	1	1	0
\$100,000	\$2000/mo.	1	1	0.2780
\$150,000	\$2000/mo.	1	1	0.9876

Table 2: We simulate the optimally controlled wealth process, and solve for the optimal funding ratio on each sample path. We report the estimated goal reaching probability $\mathbb{P}(\theta_k = 1)$, $k = 1, 2, 3$ using 5,000 simulated sample paths.

The plots in Figure 7 suggest that the threshold beyond which it is optimal to consume towards the first goal decreases with the income stream. This is intuitive, because a larger income stream would make future goals easier to satisfy. However, whether or not the consumption threshold decreases as the goal becomes easier to satisfy depends on the level of attractiveness for the goal. Let us begin by considering goal "2", i.e., an easy to satisfy goal due to its low amount. If goal "2" becomes even easier to satisfy, because its amount goes down or its deadline is postponed, the consumption threshold for the first goal becomes even lower.

The situation is different if we consider goal "3", which is hard to satisfy under the benchmark setting due to its large amount of the $250,000 large relative to the client's initial capital and the guaranteed future income stream. Then, the optimal strategy is to consume towards the first two goals, rather than saving and accruing wealth towards the third goal, whose fundedness would be rather limited. However, if goal "3" becomes easier to satisfy, either because its amount is lowered or its deadline postponed, the marginal benefit of saving towards goal "3" increases. As a result, the consumption threshold for the first goal increases, because the client aims at attaining a higher fundedness for goal "3".

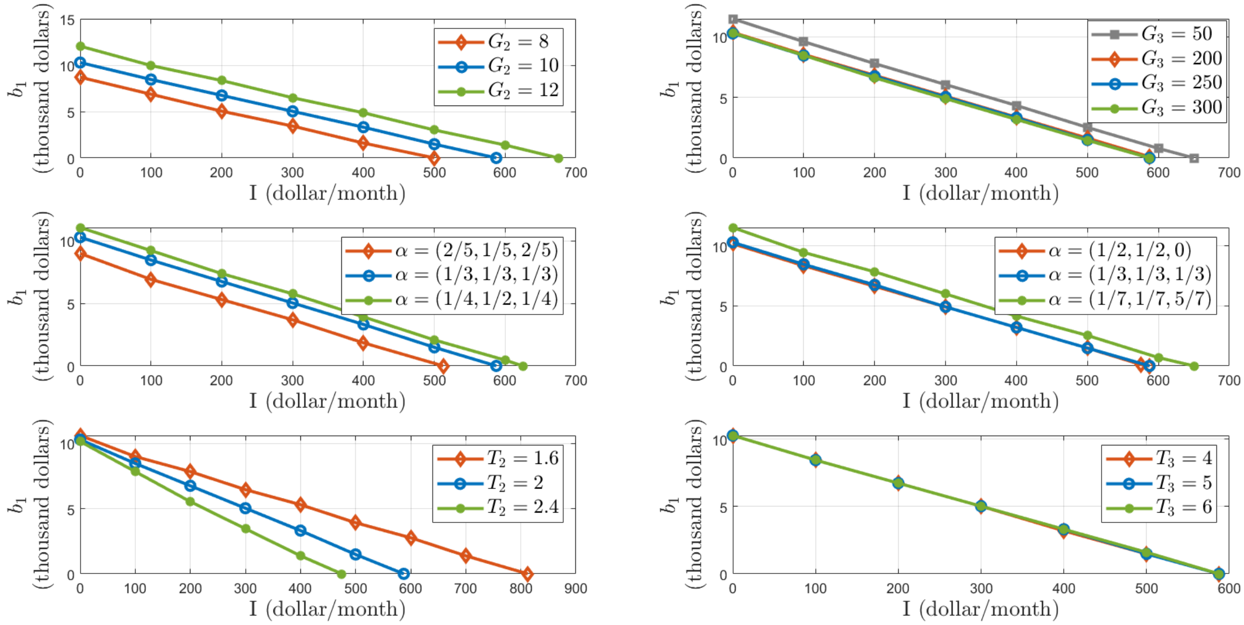

Figure 7: The consumption threshold b_1 as a function of the income stream I, for a client whose preferences are linear in the funding ratio. Top left panel: b_1 for different values of G_2. Top right panel: b_1 for different values of G_3. Middle left panel: b_1 for different values of α_2. We adjust α_1 and α_3, so that $\alpha_1 = \alpha_3$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Middle right panel: b_1 for different values of α_3. We adjust α_1 and α_2, so that $\alpha_1 = \alpha_2$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Bottom left panel: b_1 for different values of T_2. Bottom right panel: b_1 for different values of T_3.

As the importance weight of an upcoming goal increases, there is a higher aversion towards the risk of not fulfilling it. Unless the client has already accumulated a large enough amount of wealth, it becomes optimal for her to save for future goals. These investment patterns are evident in the middle panel of Figure 7, where it can be seen that the consumption threshold for goal "1" goes up. If the cost-adjusted importance of a goal is too low (e.g. its amount is too large to be attained with an investment strategy by the deadline), then the consumption threshold has little sensitivity to changes in its importance weight unless the change is significant (as seen from the middle right panel of Figure 7, the consumption threshold for goal "3" increases only if its importance weight is 5 times larger than that of other goals).

6 Computational Complexity and Run Times

In this section, we evaluate the computational performance of the procedure to solve the goal-based optimization problem described in Section 3.2 and Appendix D, comparing it to the computational complexity of solving the original HJB equation.

As discussed in Proposition 3.2 and Appendix D, the pointwise optimization in (3.12), which yields the unique optimizer $(\hat{s}(a,g,A,y), \hat{m}(a,g,A,y), \hat{\nu}(a,g,A,y))$ when $A < 0$, can be split into a "client-independent" optimization (i.e., the computation of $\Gamma_c(y)$ or extreme points of convex sets related to $\mathcal{P}_c(y)$) and a "client-dependent" optimization (i.e., the computation of $\mathcal{H}_c(V_x, V_{xy}, xV_{xx}, y)$). The "client-independent" problem remains unaffected by the individual client's goal characteristics and value function, thereby being the same across different clients. Consequently, our analysis narrows down to contrasting the complexity of the "client-dependent" optimization problem with that of solving the original HJB equation, which is inherently client-dependent.

Define C_k as the complexity of solving a general k-dimensional convex optimization problem, namely, minimizing a convex function of k variables within a convex domain. Further, let n_t, n_x, and n_y denote the number of points in the discretized grids for time, wealth, and factor respectively. Solving the original HJB equation requires solving the optimization problem in (3.8) for each (t,x,y). This amounts to a complexity of $C_N O(n_t n_x n_y)$. An alternative pathway is to use the transformed HJB equation, analyzed in Section 3.2 for the case of linear preferences and in Appendix D for the case of general preferences. For the case of linear preferences, the complexity of the numerical procedure for computing $\mathcal{H}_c(V_x, V_{xy}, xV_{xx}, y)$ is $C_{d+1} O(n_t n_x n_y)$. For general preferences, one needs to solve the optimization problem in Eq. (D.1) whose complexity is also bounded by $C_{d+1} O(n_t n_x n_y)$. If $d+1 < N$, the alternative pathway is computationally faster (see also the discussion right after Eq. (D.1) in the Appendix).

Table 3 presents the runtime for the "client-dependent" optimization, encompassing both the static optimization and the finite difference method utilized to solve the PDE across all successive time intervals $[T_{k-1}, T_k)$, the details of which are provided in Appendix A. The run times are measured using Colab, a hosted Jupyter notebook service offered by Google.[9]

[9] It is important to note that resource utilization can vary within the Colab environment. Therefore, the reported run times should be considered as rough estimates of the computational complexity involved in the PDE procedure.

Δt	Δz	Δm	n_t	n_z	n_π	linear	all-or-nothing
1/12	0.0259	0.25%	60	486	30	1.71 sec	1.45 sec
1/12	0.0259	0.05%	60	486	133	1.73 sec	1.48 sec
1/12	0.0259	0.01%	60	486	646	1.96 sec	1.67 sec
1/52	0.0114	0.25%	260	1103	30	14.45 sec	11.85 sec
1/52	0.0114	0.05%	260	1103	133	15.01 sec	12.49 sec
1/52	0.0114	0.01%	260	1103	646	16.14 sec	13.06 sec
1/252	0.0050	0.25%	1260	2538	30	150.07 sec	121.12 sec
1/252	0.0050	0.05%	1260	2538	133	155.60 sec	127.15 sec
1/252	0.0050	0.01%	1260	2538	646	167.79 sec	134.81 sec

Table 3: Run times of the "client-dependent" optimization for both linear and all-or-nothing preferences, averaged over 10 runs. The run times exclude the computation of minimum variance curve and corner points of the opportunity set. Observe that, in the absence of a factor process, the minimum variance curve Γ_c defined in Eq. (C.1) becomes a one-dimensional curve which can be parametrized by m. We vary Δm, Δt, and set Δz as specified in (A.6) with $\epsilon = 0.01187$. The columns n_t and n_z are the number of points in the time and log-wealth grids, respectively, which are inversely proportional to Δt and Δz. The column n_π gives the number of portfolios from the discretized minimum variance curve (which is inverse proportional to Δm) plus the number of isolated extreme points. We fix the market and client parameters to the benchmark values specified in Section 5.1.

As evident from the table, the numerical procedure is computationally efficient, with run times on the order of few minutes even with a high grid resolution, such as daily rebalancing and log-wealth changes of 0.005 (as shown in the last row of the table). Several noteworthy patterns emerge from the table. Firstly, the numerical procedure tends to be slightly faster when dealing with all-or-nothing preferences in the funding ratio. This is because searching over extreme points is typically faster than finding the interior solution of an optimization problem. Secondly, the run time of the procedure exhibits only modest sensitivity to the size of the m-grid. Furthermore, when we increase the rebalancing frequency from monthly to weekly (note that we need $dt = O((dz)^2)$ for the explicit Euler scheme to work, and thus dz is roughly reduced by half), we observe that the run time ratio between the finer grid and the coarser grid is approximately 8 (precisely, 14.45/1.71). If we proceed from weekly to daily rebalancing while further halving dz, the ratio between the run times expands to about 10 (precisely, 150.07/14.45). These numerical estimates align with theory, as it is well-established that the computational complexity of the explicit Euler method scales proportional to n_z^3.

7 Conclusion

Goals-based wealth management is evolving to a mainstream paradigm for retail investment management. Despite the increased adoption in the asset management industry, there is a limited number of analytical studies which quantify investment and goal consumption strategies, accounting for both the goal characteristics and intertemporal relations between them.

Our paper provides the first analytically tractable framework for sequential goal based wealth management in continuous time, and accommodates a large class of client's preferences towards goals. Client characteristics and asset price dynamics are coupled through an observable factor process, which impacts goal amounts and income well as expected return and volatility of investment assets. We have quantified analytically the optimal tradeoff between consuming to meet an expiring goal versus keeping wealth in the portfolio and investing it toward future goals. Our numerical analysis reveals distinct risk-taking behaviors and goal consumption patterns between a client with flexible goals and one with all-or-nothing goals.

Our framework offers numerous avenues for extension. One such extension involves allowing the clients to optimize their investment portfolio by choosing an optimal income rate I for deposits. Another interesting extension is to expand the model, currently able to deal with goals that have a fixed deadline, and incorporate an emergency fund which can be used to deal with unforeseen events. We defer these extensions to future research.

A Description of Finite Difference Method

In Sections A.1 and A.2, we perform two change of variables to improve the stability of the numerical algorithm. In Section A.3, we describe the explicit Euler scheme to solve the transformed PDE. In our implementation, we regularize the terminal condition by local averaging to ensure continuity or improve numerical stability.

In our benchmark parameter specification, we set $d = 0$, i.e., we do not consider the factor process. Hence, we decribe the numerical method with the variable y being omitted.

As discussed in the viscosity characterization, we allow for negative values of x. For $x > 0$, the HJB equation is equivalent to (3.11) with the y variable removed, i.e.,

$$V_t + IV_x + xV_x\hat{m}(V_x, xV_{xx}) + \frac{1}{2}x^2V_{xx}\hat{s}^2(V_x, xV_{xx}) = 0. \tag{A.1}$$

For $x \leq 0$, the HJB equation takes the simpler form $V_t + (I + rx)V_x = 0$. The terminal condition is determined by the static optimization problem (4.4). The safe level serves as a natural upper boundary for the numerical domain, with corresponding boundary condition

$$V\left(t, \mathfrak{s}(t)\right) = \sum_{j=k}^{M} \alpha_j f_j(1).$$

The lower boundary will become clear after introducing the first change of variable.

A.1 Flatten the Safe Level by a First Change of Variable

Discretization of Eq. (A.1) leads to a piecewise constant safe level $\mathfrak{s}(t)$. Due to the high sensitivity of the optimal portfolio near the safe level (since V_{xx} is close to a Dirac delta function there), the rounding error introduced by discretization causes instability. We remove the rounding error by a change of variable from (t, x) to $(t, \tilde{x})$, where

$$\tilde{x} = e^{r(T_k - t)} x + \frac{I}{r}\left(e^{r(T_k - t)} - 1\right).$$

Observe that $\tilde{x} = x$ if $t = T_k$. It is also easy to check that $\tilde{x} = \mathfrak{s}(T_k-) =: \tilde{x}_s$ when $x = \mathfrak{s}(t)$, i.e. the safe level in $\tilde{x}$ is constant. However, the bankruptcy level $\tilde{x}_\ell$ (corresponding to $x = 0$) is no longer flat:

$$\tilde{x}_\ell(t) = \frac{I}{r}(e^{r(T_k - t)} - 1).$$

Nevertheless, the optimal portfolio is not as sensitive near the bankruptcy level; so the rounding error here does not cause instability. In addition, we will see below that the transformed HJB equation for $\tilde{x} \leq \tilde{x}_\ell(t)$ gives rise to a Dirichlet boundary condition at $\tilde{x} = 0$. Let $\tilde{V}(t, \tilde{x}) = V(t, x)$ and write

$$(\tilde{\sigma}, \tilde{\mu})(t, \tilde{x}) := (\hat{s}, \hat{m})(\tilde{V}_{\tilde{x}}(t, \tilde{x})e^{-r(T_k - t)}, x\tilde{V}_{\tilde{x}\tilde{x}}(t, \tilde{x})),$$

where we recall that $(\hat{s}, \hat{m})$ is the unique optimizer of (3.12). Then $\tilde{V}$ satisfies a diffusion equation

$$\tilde{V}_t + e^{2r(T_k - t)} \left\{ x e^{-r(T_k - t)} \left(\tilde{\mu}(t, \tilde{x}) - r\right) \tilde{V}_{\tilde{x}} + \frac{1}{2} x^2 \tilde{\sigma}^2(t, \tilde{x}) \tilde{V}_{\tilde{x}\tilde{x}} \right\} = 0 \tag{A.2}$$

on $\{(t, \tilde{x}) : T_{k-1} \leq t < T_k, \tilde{x} > \tilde{x}_\ell(t)\}$, and a trivial transport equation $\tilde{V}_t = 0$ on $\{(t, \tilde{x}) : T_{k-1} \leq t < T_k, \tilde{x} \leq \tilde{x}_\ell(t)\}$. The terminal condition and the upper boundary condition are respectively given by

$$\tilde{V}(T_k-, \tilde{x}) = V(T_k-, x) = V(T_k-, \tilde{x}) \quad \text{and} \quad \tilde{V}\,(t, \tilde{x}_s) = \sum_{j=k}^{M} \alpha_j f_j(1).$$

Since $\tilde{x}_\ell(t) \geq 0$ and the direction of transport below $\tilde{x}_\ell(t)$ is horizontal, we can choose $\tilde{x} = 0$ as a constant lower boundary for our numerical domain along which $\tilde{V}(t, 0) = \tilde{V}(T_k-, 0)$. At $\tilde{x} > 0$, we have $\tilde{V}(t, \tilde{x}) = \tilde{V}(t', \tilde{x})$ where $t' := \inf\{s \geq t : \tilde{x}_\ell(s) \leq \tilde{x}\} < T_k$.

A.2 Second Change of Variable to Logarithmic Wealth

The explicit Euler method for (A.2) requires the time discretization Δt and the space discretization $\Delta\tilde{x}$ to satisfy the stability condition $\Delta t = O((\Delta\tilde{x})^2)$, where the constant is inverse proportional to x^2. For large goal amount, this constant is very small near the safe level, resulting in a large number of time steps. To reduce the time complexity while maintaining stability, let ϵ be a positive constant close to zero. For $\tilde{x} \geq \epsilon$, we perform a second change of variable from $(t, \tilde{x})$ to (t, z), where $z = \ln \tilde{x}$. Let $U(t, z) = \tilde{V}(t, \tilde{x}) = V(t, x)$. Also, define $z_s := \ln(\tilde{x}_s \vee \epsilon)$, $z_\ell(t) := \ln(\tilde{x}_\ell(t) \vee \epsilon)$ and

$$(\hat{\sigma}, \hat{\mu})(t, z) = (\tilde{\sigma}, \tilde{\mu})(t, \tilde{x}) = (\hat{s}, \hat{m})\left(U_z e^{-r(T_k - t)}, \frac{x}{\tilde{x}}(U_{zz} - U_z)\right). \tag{A.3}$$

We obtain that U satisfies

$$U_t + \left[\frac{x}{\tilde{x}} e^{r(T_k - t)}\left(\hat{\mu}(t, z) - r\right) - \frac{1}{2}\left(\frac{x e^{r(T_k - t)}}{\tilde{x}}\right)^2 \hat{\sigma}^2(t, z)\right] U_z + \frac{1}{2}\left(\frac{x e^{r(T_k - t)}}{\tilde{x}}\right)^2 \hat{\sigma}^2(t, z) U_{zz} = 0, \tag{A.4}$$

if $z > z_\ell(t)$. For $z \leq z_\ell(t)$, we approximately have $U_t(t, z) = 0$. The terminal condition is

$$U(T_k-, z) = \tilde{V}(T_k-, e^z) = V(T_k-, e^z)$$

and the boundary condition is

$$U\,(t, z_s) = \sum_{j=k}^{M} \alpha_j f_j(1), \quad U(t, \ln \epsilon) \approx U(T_k-, \ln \epsilon).$$

Note that the term $x^2e^{2r(T_k-t)}$ in the coefficient of the second derivative is now balanced out by $\tilde{x}^2$, thanks to the bound $|xe^{r(T_k-t)}/\tilde{x}| \leq 1$. This makes the stability condition much easier to satisfy; see (A.6).

A.3 Explicit Euler Method

Let Δt and Δz be the time and space discretizations in the (t, z)-coordinate. Write $z_j = z_{\min} + j\Delta z$, $\tilde{x}_j = e^{z_j}$, $U_j^n \approx U(n\Delta t, z_j)$, $\hat{\mu}_j^n \approx \hat{\mu}(n\Delta t, z_j)$ and $\hat{\sigma}_j^n \approx \hat{\sigma}(n\Delta t, z_j)$, where $\hat{\sigma}$ and $\hat{\mu}$ are defined in (A.3). Also let x_j^n be the inverse transform of $(n\Delta t, \tilde{x}_j)$ back to the original (t, x)-coordinate. The explicit Euler method for (A.4) reads

$$U_j^{n-1} = U_j^n + \Delta t H_j^n(\hat{\sigma}_j^n, \hat{\mu}_j^n), \quad j = L^{n-1}, L^{n-1}+1, \ldots, J$$

where $z_\ell(n\Delta t) \approx z_{\min} + L^n\Delta z$, $z_s \approx z_{\min} + J\Delta z$,

$$\begin{aligned}
\hat{\mu}_j^n &= \hat{m}\left(e^{-r(T_k-n\Delta t)}\mathcal{D}_j^n U^n, \frac{x_j^n}{\tilde{x}_j}\left(\frac{U_{j+1}^n + U_{j-1}^n - 2U_j^n}{(\Delta z)^2} - \mathcal{D}_j^n U^n\right)\right), \\
\hat{\sigma}_j^n &= \hat{s}\left(e^{-r(T_k-n\Delta t)}\mathcal{D}_j^n U^n, \frac{x_j^n}{\tilde{x}_j}\left(\frac{U_{j+1}^n + U_{j-1}^n - 2U_j^n}{(\Delta z)^2} - \mathcal{D}_j^n U^n\right)\right),
\end{aligned}$$

$$H_j^n(s, m) = \left(\frac{x_j^n e^{r(T_k-n\Delta t)}}{\tilde{x}_j}\right)^2 \left\{\left(\frac{\tilde{x}_j(m-r)}{x_j^n e^{r(T_k-n\Delta t)}} - \frac{s^2}{2}\right)\mathcal{D}_j^n U^n + \frac{s^2}{2} \cdot \frac{U_{j+1}^n + U_{j-1}^n - 2U_j^n}{(\Delta z)^2}\right\},$$

and $\mathcal{D}_j^n$ is a differencing operator in the z-variable. To get a simple positive coefficient method, we choose $\mathcal{D}_j^n U^n = (U_{j+1}^n - U_j^n)/\Delta z$. This leads to

$$H_j^n(s, m) = p_+(s, m)U_{j+1}^n + p_-(s, m)U_{j-1}^n - (p_+(s, m) + p_-(s, m))U_j^n.$$

where

$$p_+(s, m) = \left(\frac{x_j^n e^{r(T_k-n\Delta t)}}{\tilde{x}_j}\right)^2 \left\{\frac{s^2}{2(\Delta z)^2} + \frac{1}{\Delta z}\left[\frac{\tilde{x}_j}{x_j^n e^{r(T_k-n\Delta t)}}(m-r) - \frac{s^2}{2}\right]\right\}.$$

and

$$p_-(s, m) = \frac{s^2}{2(\Delta z)^2}\left(\frac{x_j^n e^{r(T_k-n\Delta t)}}{\tilde{x}_j}\right)^2 \geq 0.$$

For convergence of the numerical solution, it suffices to require the positive coefficient condition $p_\pm \geq 0$ and the stability condition $(p_+ + p_-)\Delta t \leq 1$ at all nodes (n, j); see e.g. Forsyth and Labahn [2007/08] and Barles and Souganidis [1991]. The positive coefficient condition is always satisfied for $\Delta z \leq 1$ when $\mu_i \geq r$ for all i; the stability condition requires

$$\frac{x_j^n e^{r(T_k - n\Delta t)}}{\tilde{x}_j} \left\{ \frac{s^2 \Delta t}{(\Delta z)^2} + \frac{\Delta t}{\Delta z} \left[(m - r) - \frac{s^2}{2} \frac{x_j^n e^{r(T_k - n\Delta t)}}{\tilde{x}_j} \right] \right\} \leq 1. \tag{A.5}$$

Set $M_{var} = c^2 \wedge \max_{i=1,\dots,N} C_{ii}$, and $M_{exc} = \max_{i=1,\dots,N}(\mu_i - r)^+$, where we recall that C_{ii} is the i-th diagonal entry of the covariance matrix $C = \Sigma\Sigma^\top$. Since $|xe^{r(T_k - t)}/\tilde{x}| \leq 1$, a sufficient condition for (A.5) is

$$\Delta z = \frac{M_{exc}\Delta t + \sqrt{M_{exc}^2 \Delta t^2 + 4M_{var}\Delta t}}{2}(1 + \epsilon) \tag{A.6}$$

for an arbitrarily small ϵ.

Under the benchmark parameters setting in the main body of the paper, both (A.6) and the positive coefficient condition are satisfied.

Remark A.1. *The numerical procedure described in this Appendix uses a finite wealth grid to solve the nonlinear PDE specified by Eq.* (3.6)*-*(3.7)*. The viscosity uniqueness characterized in Theorem 4.1 (d) also holds on bounded domains of the form* $\{(t, x, y) \in \mathbf{S}_k : (x, y) \in O(t) \subseteq \mathbb{R}^{1+d}\}$ *with prescribed Dirichlet boundary conditions. In this case, the polynomial growth condition in Theorem 4.1 (d) is not needed.*

B Proof of Corollary 3.1

Fix $y \in \mathbb{R}^d$. Setting $f_k(\theta) = \theta$ for all $k = 1, \dots, M$ in Theorem 4.1(f) yields $\theta_k^*(x, y) = 0$ for $0 \leq x \leq b_k(y)$. If $x > b_k(y)$, let $z := 1 \wedge \frac{x - b_k(y)}{G_k(y)} > 0$. For any $\theta \in [0, z)$, we have $x - \theta G_k(y) > b_k(y)$ and $\partial_x^+ V(T_k, x - \theta G_k(y), y) \leq \alpha_k / G_k(y)$. By Proposition 3.3, this can only happen if $\theta \leq \theta_k^*(x, y)$. Since $\theta \in [0, z)$ is arbitrary, we conclude that $\theta_k^*(x, y) \geq z$. If $z = 1$, then we trivially have $\theta_k^*(x, y) \leq 1 = z$. If $z < 1$, then for any $\theta > z$, we have $x - \theta G_k(y) < b_k(y)$ and $\partial_x^+ V(T_k, x - \theta G_k(y), y) > \alpha_k / G_k(y)$. Proposition 3.3 then implies that $\theta_k^*(x, y) \leq \theta$. Letting $\theta \searrow z$ yields $\theta_k^*(x, y) \leq z$. Combining the two inequalities, we obtain $\theta_k^*(x, y) = z = 1 \wedge \frac{x - b_k(y)}{G_k(y)}$ for $x > b_k(y)$, which is piecewise linear in x.

C Proof of Proposition 3.2 and Technical Details for Computing $\Gamma_c(y)$ and $\mathcal{E}(dom(\mathcal{S}_c(y)))$

Lemma C.1. *For $c : \mathbb{R}^d \to \mathbb{R}_+ \cup \{\infty\}$ and $y \in \mathbb{R}^d$, the set-valued map $\mathcal{S}_c(y)$ is compact-valued, convex-valued and continuous. Moreover, $s_{\min}(m, \nu; y, c) : \mathbb{R}^{1+d} \to \mathbb{R} \cup \{\infty\}$ is a convex function. Consequently, $dom(\mathcal{S}_c(y)) = \{(m, \nu) \in \mathbb{R}^{1+d} : s_{\min}(m, \nu; y, c) < \infty\}$ is a convex set.*

Proof. First observe that $\mathcal{P}_c(y)$ is compact because it is the continuous image of the compact set $\Delta_c(y)$. Let $(m, \nu) \in dom(\mathcal{S}_c(y))$. Being the (m, ν)-section of the compact set $\mathcal{P}_c(y)$, $\mathcal{S}_c(y)(m, \nu)$ is also compact. To see convexity (equivalent to connectedness in $\mathbb{R}$), we observe that $\mathcal{S}_c(y)(m, \nu)$ is the image of the convex set $\Xi(m, \nu) := \{\pi \in \Delta_c(y) : r + \pi(\mu(y) - r\mathbb{1}_N) = m, \pi\Sigma\Sigma_Y^\top(y) = \nu\}$ under the continuous function $\pi \mapsto \|\pi\Sigma(y)\|$, and thus, is connected.

Next, we show the continuity of $(m, \nu) \mapsto \mathcal{S}_c(y)(m, \nu)$. Upper-hemicontinuity follows from the closedness of $\mathcal{P}_c(y)$ (see [Aliprantis and Border, 2006, Theorem 17.11]). Lower-hemicontinuity can be deduced from the lower-hemicontinuity of Ξ (see [Aubin and Frankowska, 2009, Proposition 1.5.1]) and [Aliprantis and Border, 2006, Theorem 17.23].

Finally, the convexity of $(m, \nu) \mapsto s_{\min}(m, \nu; y, c) = \inf\{\|\pi\Sigma(y)\| : \pi \in \Xi(m, \nu)\}$ is a consequence of the convexity of $\pi \mapsto \|\pi\Sigma(y)\|$ and the linearity of $\pi \mapsto (r + \pi(\mu(y) - r\mathbb{1}_N), \pi\Sigma\Sigma_Y^\top(y))$. Indeed, let $\pi_i \in \Xi(m_i, \nu_i)$ be the minimizer for $s_{\min}(m_i, \nu_i; y, c)$, $i = 1, 2$ (the minimizer exists because $\Xi(m, \nu)$ is compact), and let $\lambda \in [0, 1]$. It is easy to check that $\pi := \lambda\pi_1 + (1 - \lambda)\pi_2 \in \Xi(m, \nu)$ where $(m, \nu) := \lambda(m_1, \nu_1) + (1 - \lambda)(m_2, \nu_2)$. We then have $s_{\min}(m, \nu; y, c) \le \|\pi\Sigma(y)\| \le \lambda\|\pi_1\Sigma(y)\| + (1 - \lambda)\|\pi_2\Sigma(y)\| = \lambda s_{\min}(m_1, \nu_1; y, c) + (1 - \lambda)s_{\min}(m_2, \nu_2; y, c)$. □

Proof of Proposition 3.2. Equations (3.11)-(3.12) follow directly from the definition of $\mathcal{P}_c(y)$. The optimizer $(\hat{s}, \hat{m}, \hat{\nu})$ always exists because $\mathcal{P}_c(y)$ is compact and nonempty (it always contains the risk-free asset).

For $(m, \nu) \in dom(\mathcal{S}_c(y))$, Lemma C.1 implies that $s_{\min}(m, \nu; y, c) = \inf \mathcal{S}_c(y)(m, \nu)$ is convex and thus continuous in (m, ν). It follows that $\Gamma_c(y)$ is a continuous hypersurface. Since $\mathcal{P}_c(y)$ lives in a subspace of dimension $2 + d'(y)$, both $\Gamma_c(y)$ and $dom(\mathcal{S}_c(y))$ have dimension $1 + d'(y)$. It remains to show (3.13)-(3.16).

When $A < 0$, we have

$$\begin{aligned}\mathcal{H}_c(a, g, A, y) &= A \inf_{(m,\nu)\in dom(\mathcal{S}_c(y))} \left\{ \frac{a}{A}m + \frac{\nu g}{A} + \frac{1}{2} \inf_{s\in\mathcal{S}_c(y)(m,\nu)} s^2 \right\} \\ &= A \inf_{(m,\nu)\in dom(\mathcal{S}_c(y))} \left\{ \frac{a}{A}m + \frac{\nu g}{A} + \frac{1}{2} s^2_{\min}(m, \nu; y, c) \right\} \\ &= A \inf_{(s,m,\nu)\in\Gamma_c(y)} \left\{ \frac{a}{A}m + \frac{\nu g}{A} + \frac{1}{2}s^2 \right\} = Ah_c \left(\frac{a}{A}, \frac{g}{A}, y \right),\end{aligned}$$

where h_c is defined in (3.15). Because $(s, m, \nu) \mapsto \gamma_1 m + \nu\gamma_2 + s^2/2$ is strictly convex, the optimizer for $h_c(\gamma_1, \gamma_2, y)$ is unique. If $A = 0$, we have

$$\mathcal{H}_c(a, g, A, y) = \sup_{(m,\nu)\in dom(\mathcal{S}_c(y))} \sup_{s\in\mathcal{S}_c(y)(m,\nu)} \{am + \nu g\} = \sup_{(m,\nu)\in dom(\mathcal{S}_c(y))} \{am + \nu g\}.$$

Since the objective function $(m, \nu) \mapsto am + \nu g$ is linear, maximizing over the convex set $dom(\mathcal{S}_c(y))$ reduces to maximizing over its extreme points, which yields (3.16). □

We conclude this section with a discussion of the methods used to compute $\Gamma_c(y)$ and $\mathcal{E}(dom(\mathcal{S}_c(y)))$. Define $\mathcal{S}(y)$ and $s_{\min}(\cdot; y)$ by setting $c(y) = \infty$ for all y in the definitions of $\mathcal{S}_c(y)$ and $s_{\min}(\cdot; y, c)$. Since $\mathcal{S}(y)$ is compact-valued, $s_{\min}(m, \nu; y)$ is attained for any $(m, \nu) \in \mathcal{S}(y)$. Using the attainability of $s_{\min}(\cdot; y)$, it can be easily shown that

$$\Gamma_c(y) = \{(s_{\min}(m, \nu; y), m, \nu) : (m, \nu) \in dom(\mathcal{S}(y))\} \cap ([0, c(y)] \times \mathbb{R}^{1+d}). \tag{C.1}$$

Hence the computation of $\Gamma_c(y)$ depends on the evaluation of $dom(\mathcal{S}(y))$ and $s_{\min}(m, \nu; y)$.

- Computation of $dom(\mathcal{S}(y))$. Using the fact that a polytope is the convex hull of its vertices, and that a linear transformation commutes with the convex hull operation, we have that $dom(\mathcal{S}(y))$ is a polytope whose vertex set is the image of the vertices of Δ under the linear map $\pi \mapsto (r + \pi(\mu(y) - r\mathbb{1}_N), \pi\Sigma\Sigma_Y^\top(y))$. That is,

$$dom(\mathcal{S}(y)) = co\left(\{(r, \mathbb{0}_d), (\mu_1(y), \sigma_1(y)\Sigma_Y^\top(y)), \dots, (\mu_N(y), \sigma_N(y)\Sigma_Y^\top(y))\}\right),$$

  where $co(\cdot)$ denotes the convex hull and $\mathbb{0}_d$ is the zero vector in $\mathbb{R}^d$. If $d'(y) = 0$, $dom(\mathcal{S}(y))$ reduces to an interval $[\min_{i=0,\dots,N} \mu_i(y), \max_{i=0,\dots,N} \mu_i(y)]$, where we set $\mu_0(y) := r(y)$.

- Computation of $s_{\min}(m,\nu;y)$. The minimum variance problem

$$s^2_{\min}(m,\nu;y)=\inf\{\|\pi\Sigma(y)\|^2:\pi\in\Delta, r+\pi(\mu(y)-r\mathbb{1}_N)=m,\pi\Sigma\Sigma_Y^\top(y)=\nu\} \quad (C.2)$$

  is a quadratic programming problem in $\mathbb{R}^N$, which can be efficiently solved using state-of-art packages such as CPLEX. Because the optimization problem in (C.2) is convex, the optimal π can be uniquely identified, giving us a bijection between $\Gamma_c(y)$ and $\Phi^{-1}(\Gamma_c(y);y)$, where we recall that Φ has been defined in Eq. (3.9). The map to recover π from $(s,m,\nu)\in\Gamma_c(y)$ can be *pre-computed* and *stored*, and is independent of client information such as her wealth and value function.

D Computation of $\mathcal{H}_c(a,g,A,y)$ for Positive A

In this section, we describe how to compute $\mathcal{H}_c(a,g,A,y)$ for $A>0$. Recall that

$$\begin{aligned}\mathcal{H}_c(a,g,A,y)&=\sup_{(s,m,\nu)\in\mathcal{P}_c(y)}\left\{am+\nu g+\frac{1}{2}As^2\right\}\\ &=\sup_{\pi\in\Delta_c(y)}\left\{a[r+\pi(\mu(y)-r\mathbb{1}_N)]+\pi\Sigma\Sigma_Y^\top(y)g+\frac{1}{2}A\|\pi\Sigma(y)\|^2\right\}\end{aligned}$$

We start with the original convex domain $\Delta_c(y)$ whose extreme points are easier to characterize. In the original domain, the objective function is convex in π. Hence, for each y, it suffices to search over the set of extreme points of $\Delta_c(y)$ which are of the form: (i) any single asset (including risk-free asset) whose volatility is no larger than $c(y)$, or (ii) any convex combination of assets whose portfolio volatility is exactly equal to $c(y)$. Denote the set of extreme points by E_y. We have

$$E_y=(\{\mathbb{0}_N\}\cup\{e_i:\|\sigma_i(y)\|\le c(y)\}_{i=1,\dots,N})\cup\{\pi\in\Delta:\|\pi\Sigma(y)\|=c(y)\}=:E_y^1\cup E_y^2,$$

where $\mathbb{0}_N$ is the zero vector in $\mathbb{R}^N$ and e_i is the unit vector in $\mathbb{R}^N$ with one in its i-th coordinate and zero elsewhere. We then map $\mathcal{E}_y$ to the (s,m,ν)-space. Its image is given by

$$\Phi(E_y;y)=\Phi(E_y^1;y)\cup\Phi(E_y^2;y)=\Phi(E_y^1;y)\cup\{(c(y),m,\nu):(m,\nu)\in\mathcal{D}_y\},$$

where $\mathcal{D}_y:=\{(m,\nu)\in dom(\mathcal{S}_c(y)):(c(y),m,\nu)\in\mathcal{P}_c(y)\}$. Observe that $\Phi(E_y^1;y)$ is a finite set containing at most $N+1$ points, and $\Phi(E_y^2; y)$ lives in the hyperplane $s = c(y)$ which has dimension $1 + d'(y)$.

Next, we turn to the transformed domain $\mathcal{P}_c(y)$ in which the objective function $(s, m, \nu) \mapsto am + \nu g + \frac{1}{2}As^2$ remains convex. This means we only need to search in a subset of $\Phi(\mathcal{E}_y; y)$ whose points are also extreme points of $co(\mathcal{P}_c(y))$. Here, we take the convex hull because $\mathcal{P}_c(y)$ need not be convex (see Figure A1) despite having convex (m, ν)-slices by Lemma C.1. Also observe that the objective function restricted to the hyperplane $s = c(y)$ is linear. Hence the optimization problem reduces to

$$\max\left\{\sup_{(s,m,\nu)\in\Phi(E_y^1;y)\cap\mathcal{E}(co(\mathcal{P}_c(y)))} am + \nu g + \frac{1}{2}As^2, \sup_{(m,\nu)\in co(\mathcal{D}_y)} am + \nu g + \frac{1}{2}Ac^2(y)\right\}, \tag{D.1}$$

where $co(\cdot)$ and $\mathcal{E}(\cdot)$ denote the convex hull and extreme points operator, respectively. The left subproblem amounts to a search over the finite set $\Phi(E_y^1; y) \cap \mathcal{E}(co(\mathcal{P}_c(y)))$. The right subproblem is a $(1 + d'(y))$-dimensional convex optimization problem where the convex set $co(\mathcal{D}_y)$ can further be replaced by its extreme points $\mathcal{E}(co(\mathcal{D}_y))$. The benefit of analyzing $\Delta_c(y)$ first is that it enables us to eliminate many points on the minimum variance hypersurface $\Gamma_c(y)$ that are extreme points of $co(\mathcal{P}_c(y))$ but do not lie in $\Phi(E_y; y)$.

Remark D.1. *For each y satisfying $d'(y) = 0$, the number of extreme points which need to be stored is at most $N+3$. This is because $\Phi(\mathcal{E}_y^1; y)$ contains at most $N+1$ points, and if $d'(y) = 0$ the set $co(\mathcal{D}_y)$ degenerates to an interval (since the only feasible ν is zero). We only need to track its two end points which correspond to the maximum mean and minimum mean when the volatility is fixed to be $c(y)$. The total number of points in $\Phi(\mathcal{E}_y; y)$ that are also extreme points of $co(\mathcal{P}_c(y))$ is therefore bounded by $N+3$.*

E Auxiliary Results for the Proof of Theorem 4.1

Section E.1 requires no assumption; Section E.2 requires f_k to be continuous; Section E.3 requires f_k to be differentiable, concave and increasing.

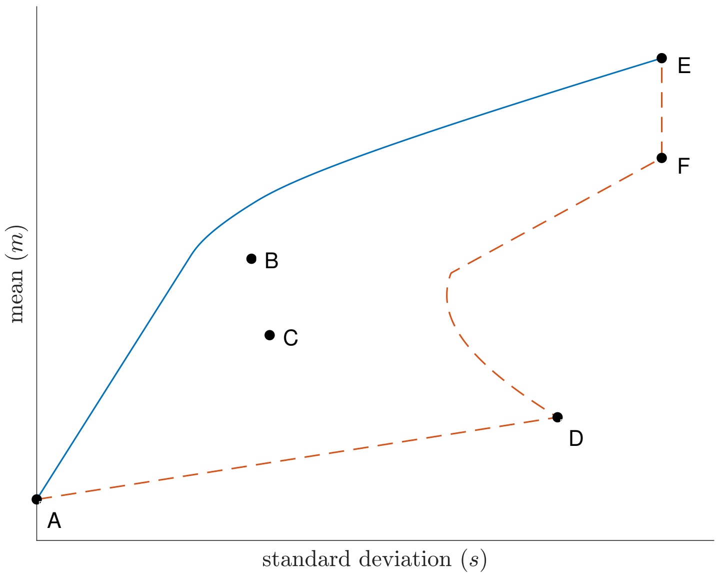

Figure A1: An example of a non-convex opportunity set formed by $N = 4$ risky assets and a risk-free asset, without a factor process. The solid blue line is the minimum variance curve Γ_c; the dashed red line is the maximum variance curve. The region enclosed by the two curves is the opportunity set $\mathcal{P}_c$ which has convex horizontal slices, but is not a convex set in the (s, m)-plane. The set $\Phi(\mathcal{E})$ is given by four isolated points A, B, C, D and the vertical line connecting E and F. Among these points, only the four corner points A, D, E, F are extreme points of $co(\mathcal{P}_c)$.

E.1 Value Function Equivalence

Lemma E.1. *Let $\tilde{V}$ be the value function defined by the right hand side of* (4.2)*, together with factor and wealth dynamics specified, respectively, by* (2.1) *and*

$$dX_t = [I(Y_t) + rX_t + \rho(X_t)\boldsymbol{\pi}_t(\mu(Y_t) - r\mathbb{1}_N)]\,dt + \rho(X_t)\boldsymbol{\pi}_t\Sigma(Y_t)dW_t, \tag{E.1}$$

where $\rho(\cdot)$ is defined in Proposition 3.1. Then $\tilde{V} = V$.

Proof. The equality clearly holds at the last goal deadline where $\tilde{V}(T_M, x, y) = V(T_M, x, y) = 0$. If $t < T_M$, let $(\boldsymbol{\pi}, \boldsymbol{\theta})$ be an ε-optimal control for $V(t, x, y)$. Observe that if $X_s \geq K = \sup_{(t,y)\in[0,T_M]\times\mathbb{R}^d} |\mathfrak{s}(t, y)|$, the investor already has sufficient wealth to fully fund all remaining goals, so that any further investment in the risky asset cannot improve the expected payoff. In other words, $\tilde{\boldsymbol{\pi}}_s = \boldsymbol{\pi}_s 1_{\{s<\tau_K\}}$ with $\tau_K := \inf\{s \geq t : X_s^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}} \geq K\}$ is at least as good as $\boldsymbol{\pi}$. By replacing $\boldsymbol{\pi}$ by $\tilde{\boldsymbol{\pi}}$ if needed, we may assume without loss of generality that the ε-optimal control satisfies $\boldsymbol{\pi}_s = 0$ for $s \geq \tau_K$.

Denote by X the wealth process in (3.5) and by $\tilde{X}$ the wealth process in (E.1). Since $\rho(x) = x^+$ if $x \leq K$, $X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}$ and $\tilde{X}^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}$ evolve in exactly the same way up to time τ_K. After time τ_K, the two controlled SDEs continue to coincide thanks to $\boldsymbol{\pi} = 0$. Since $X_s^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}} = \tilde{X}_s^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}$ for all $s \in [t, T_M]$, it is clear that $(\boldsymbol{\pi}, \boldsymbol{\theta})$ is admissible for $\tilde{V}(t,x,y)$ and consequently, $V(t,x,y) - \varepsilon \leq \tilde{V}(t,x,y)$. Sending $\varepsilon \to 0$ yields $V(t,x,y) \leq \tilde{V}(t,x,y)$. The reverse inequality $\tilde{V} \leq V$ can be proved similarly. $\square$

Remark E.1. *The benefit of replacing* (3.5) *by* (E.1) *is that the SDE coefficients of the latter is Lipschitz continuous in* $(x,y) \in \mathbb{R} \times \mathbb{R}^d$*, as the product of two bounded Lipschitz continuous functions is Lipschitz continuous. This is crucial for the application of the comparison principle in the proof of Theorem 4.1 (d), and also simplifies the SDE estimates in the proof of Lemma E.2 and Theorem 4.1 (c).*

E.2 Weak dynamic programming principle

The weak dynamic programing principle (DPP) is taken from [Touzi, 2013, Theorem 3.3] whose proof is tweaked to accommodate our volatility constraint. Recall that g^* and g_* are the upper and lower semicontinuous envelopes of g defined in (4.3).

Lemma E.2. *Suppose* $f_k : [0,1] \to [0,1]$ *is continuous for all* $k = 1, \ldots, M$*. Fix* $(t,x,y) \in \mathbf{S} := [0, T_M) \times \mathbb{R} \times \mathbb{R}^d$*. Let* $\{\tau^{\boldsymbol{\pi},\boldsymbol{\theta}} : (\boldsymbol{\pi},\boldsymbol{\theta}) \in \mathcal{A}_t(x,y)\}$ *be a family of* $[t, T_M]$*-valued stopping times. Then*

$$V(t,x,y) \leq \sup_{(\boldsymbol{\pi},\boldsymbol{\theta}) \in \mathcal{A}_t(x,y)} \mathbb{E}\left[V^*(\tau^{\boldsymbol{\pi},\boldsymbol{\theta}}, X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}_{\tau^{\boldsymbol{\pi},\boldsymbol{\theta}}}, Y^{t,y}_{\tau^{\boldsymbol{\pi},\boldsymbol{\theta}}}) \right], \tag{E.2}$$

and

$$V(t,x,y) \geq \sup_{(\boldsymbol{\pi},\boldsymbol{\theta}) \in \mathcal{A}_t(x,y)} \mathbb{E}\left[V_*(\tau^{\boldsymbol{\pi},\boldsymbol{\theta}}, X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}_{\tau^{\boldsymbol{\pi},\boldsymbol{\theta}}}, Y^{t,y}_{\tau^{\boldsymbol{\pi},\boldsymbol{\theta}}}) \right]. \tag{E.3}$$

Proof. We only mention the difference from [Touzi, 2013, Theorem 3.3]. The nontrivial direction (E.3) of the weak DPP is typically proved using a covering argument where one partitions the domain into countable disjoint regions, and constructs a global control by pasting together finitely many representative controls, one from each region where $(\tau^{\boldsymbol{\pi},\boldsymbol{\theta}}, X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}_{\tau^{\boldsymbol{\pi},\boldsymbol{\theta}}}, Y^{t,y}_{\tau^{\boldsymbol{\pi},\boldsymbol{\theta}}})$ may lie in. A key step is to show that each representative is approximately optimal for all starting points within its region (cf. Eq. (3.1) in Touzi [2013]). The subtlety in our problem is that the control $\boldsymbol{\pi}$ is subject to a factor-dependent volatility constraint; so a control that is near optimal for the starting point (t,x,y) may fail to be admissible for a nearby starting point (t',x',y'). However, we show below that it is possible to ensure admissibility by a suitable truncation, without losing much performance.

Let us first remove the funding ratio constraint $\boldsymbol{\theta}_k G_k(Y_{T_k}) \le X^+_{T_k-}$ by a reparametrization. Express each $\boldsymbol{\theta}$ by $\boldsymbol{\theta}_k = [\boldsymbol{\omega}_k X^+_{T_k-}/G_k(Y_{T_k})] \wedge 1$, where $\boldsymbol{\omega}_k$ can be interpreted as the consumption fraction, and can take any values in $[0,1]$. In terms of $\boldsymbol{\omega}$, the control problem can be rewritten as $V(t,x,y) = \sup_{(\boldsymbol{\pi},\boldsymbol{\omega})} J(t,x,y;\boldsymbol{\pi},\boldsymbol{\omega})$ where

$$J(t,x,y;\boldsymbol{\pi},\boldsymbol{\omega}) := \mathbb{E}\left[\sum_{k=1}^{M} \alpha_k f_k \left(\frac{\boldsymbol{\omega}_k (X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\omega}}_{T_k-})^+}{G_k(Y^{t,y}_{T_k})} \wedge 1 \right) 1_{\{T_k>t\}} \right],$$

and the supremum is taken over all progressively measurable, $\Delta_c(Y^{t,y})$-valued processes $\boldsymbol{\pi}$, and random vectors $\boldsymbol{\omega} = (\boldsymbol{\omega}_1,\dots,\boldsymbol{\omega}_M)$ such that $\boldsymbol{\omega}_k$ is $\mathcal{F}_{T_k}$-measurable, $[0,1]$-valued for all k for which $T_k > t$. With a little abuse of notation, we still use $\mathcal{A}_t(x,y)$ to denote the set of admissible $(\boldsymbol{\pi},\boldsymbol{\omega})$.

Let $(\boldsymbol{\pi},\boldsymbol{\omega})$ be ε-optimal for $V(t,x,y)$. Define the cylinder $B_\delta(t,x,y) := \{(t',x',y') \in \mathbf{S} : t-\delta < t' \le t, \|(x',y')-(x,y)\| < \delta\}$. To follow the proof in Touzi [2013], we need to show there exists $\delta = \delta_{t,x,y} > 0$ such that for any $(t',x',y') \in B_\delta(t,x,y)$, we can find $(\boldsymbol{\pi}',\boldsymbol{\omega}) \in \mathcal{A}_{t'}(x',y')$ satisfying $J(t,x,y;\boldsymbol{\pi},\boldsymbol{\omega}) - J(t',x',y';\boldsymbol{\pi}',\boldsymbol{\omega}) \le \varepsilon$. A natural candidate is

$$\boldsymbol{\pi}' := \left(1 \wedge \frac{c(Y^{t',y'})}{\|\boldsymbol{\pi}\Sigma(Y^{t',y'})\|}\right)\boldsymbol{\pi}. \tag{E.4}$$

Clearly, $(\boldsymbol{\pi}',\boldsymbol{\omega}) \in \mathcal{A}_{t'}(x',y')$ by construction. It suffices to show

$$\lim_{(t',x',y')\to(t,x,y)} J(t',x',y';\boldsymbol{\pi}',\boldsymbol{\omega}) \ge J(t,x,y;\boldsymbol{\pi},\boldsymbol{\omega}).$$

For a fixed $(\boldsymbol{\pi},\boldsymbol{\omega})$, the lower semicontinuity $\lim_{(t',x',y')\to(t,x,y)} J(t',x',y';\boldsymbol{\pi},\boldsymbol{\omega}) \ge J(t,x,y;\boldsymbol{\pi},\boldsymbol{\omega})$ follows from standard SDE estimates and is mentioned in [Touzi, 2013, Theorem 3.3]. So we only need to show

$$\lim_{(t',x',y')\to(t,x,y)} |J(t',x',y';\boldsymbol{\pi}',\boldsymbol{\omega}) - J(t',x',y';\boldsymbol{\pi},\boldsymbol{\omega})| = 0. \tag{E.5}$$

By Lemma E.1, we may assume the wealth dynamics is given by (E.1) so that the SDE coefficients of (X,Y) satisfy the Lipschitz and linear growth condition. Using standard SDE estimates (see e.g. [Pham, 2009, Theorem 1.3.16]), we obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{t\le s\le T_M}\left|Y_s^{t',y'}-Y_s^{t,y}\right|^2\right]&=\mathbb{E}\left[\sup_{t\le s\le T_M}\left|Y_s^{t,Y_t^{t',y'}}-Y_s^{t,y}\right|^2\right]\\&\le Ce^{C(T_M-t)}\mathbb{E}\left[|Y_t^{t',y'}-y'|^2+|y'-y|^2\right]\\&\le Ce^{C(T_M-t)}\mathbb{E}\left[Ce^{C(t-t')}(1+|y'|^2)(t-t')+|y'-y|^2\right].\end{aligned}$$

Here and in the sequel, C is a constant which only depends on the exogenous model input $(r, I, \mu_Y, \Sigma_Y, \mu, \Sigma, c, \rho)$ and which may change from line to line. We see that

$$\mathbb{E}\left[\sup_{t\le s\le T_M}\left|Y_s^{t',y'}-Y_s^{t,y}\right|^2\right]\to 0\quad\text{as}\quad(t',x',y')\to(t,x,y).$$

The above convergence also implies $Y_s^{t',y'}\to Y_s^{t,y}$ almost surely (abbreviated as a.s.) for all $s\in[t,T_M]$. By the continuity of $c(\cdot),\Sigma(\cdot)$ and the admissibility of $\boldsymbol{\pi}$, we further obtain $\boldsymbol{\pi}_s'\to\left(1\wedge\frac{c(Y_s^{t,y})}{\|\boldsymbol{\pi}\Sigma(Y_s^{t,y})\|}\right)\boldsymbol{\pi}_s=\boldsymbol{\pi}_s$ a.s. for all $s\in[t,T_M]$, and by bounded convergence theorem,

$$\mathbb{E}\int_t^{T_M}|\boldsymbol{\pi}_s'-\boldsymbol{\pi}_s|^2ds\to 0. \tag{E.6}$$

Next, we turn to the wealth process and suppose $t\in[T_{k-1},T_k)$ with $T_0:=0$. A similar Gronwall-type SDE estimate shows that for any square integrable $\mathcal{F}_u$-measurable random variables (ξ_1,ξ_2,ζ), where ξ_1,ξ_2 are $\mathbb{R}$-valued, and ζ is $\mathbb{R}^d$-valued,

$$\mathbb{E}\left[\sup_{u\le s<T_j}\left|X_s^{u,\xi_1,\zeta,\boldsymbol{\pi}',\boldsymbol{\omega}}-X_s^{u,\xi_2,\zeta,\boldsymbol{\pi},\boldsymbol{\omega}}\right|^2\right]\le Ce^{C(T_j-u)}\mathbb{E}\left[|\xi_1-\xi_2|^2+\int_u^{T_j}|\boldsymbol{\pi}_s'-\boldsymbol{\pi}_s|^2ds\right], \tag{E.7}$$

where T_j is the first goal deadline after time u. Take $u=t'$, $\xi_1=\xi_2=x'$ and $\zeta=y'$. (E.6) and (E.7) imply $\mathbb{E}[|X_{T_k-}^{t',x',y',\boldsymbol{\pi}',\boldsymbol{\omega}}-X_{T_k-}^{t',x',y',\boldsymbol{\pi},\boldsymbol{\omega}}|^2]\to 0$. As a result, $X_{T_k-}^{t',x',y',\boldsymbol{\pi}',\boldsymbol{\omega}}\to X_{T_k-}^{t',x',y',\boldsymbol{\pi},\boldsymbol{\omega}}$ a.s. and

$$\mathbb{E}\left[\left|X_{T_k}^{t',x',y',\boldsymbol{\pi}',\boldsymbol{\omega}}-X_{T_k}^{t',x',y',\boldsymbol{\pi},\boldsymbol{\omega}}\right|^2\right]\le\mathbb{E}\left[(1+\boldsymbol{\omega}_k)^2\left|X_{T_k-}^{t',x',y',\boldsymbol{\pi}',\boldsymbol{\omega}}-X_{T_k-}^{t',x',y',\boldsymbol{\pi},\boldsymbol{\omega}}\right|^2\right]\to 0.$$

Repeat the SDE estimate (E.7) on each successive interval $[T_k,T_{k+1})$ with $u=T_k$, and $\xi_1=X_{T_k}^{t',x',y',\boldsymbol{\pi}',\boldsymbol{\omega}}$, $\xi_2=X_{T_k}^{t',x',y',\boldsymbol{\pi},\boldsymbol{\omega}}$, $\zeta=Y_{T_k}^{t',y'}$, until we reach the last goal deadline. We obtain $X_{T_k-}^{t',x',y',\boldsymbol{\pi}',\boldsymbol{\omega}}\to X_{T_k-}^{t',x',y',\boldsymbol{\pi},\boldsymbol{\omega}}$ a.s. for all k for which $T_k>t$. The desired limit (E.5) follows from the continuity of f_k and bounded convergence theorem.

For the pasting of controls to work, we also need the truncation in (E.4) to be progressively measurable in the sense that if we replace (t', y') by a random point $(\tau, Y_\tau^{t_0,y_0})$, where τ is any $[t_0, T_M]$-valued stopping time, then the modulating factor multiplied with $\boldsymbol{\pi}$ needs to be progressively measurable. But this is automatic by the progressive measurability of $Y_s^{\tau, Y_\tau^{t_0,y_0}} = Y_s^{t_0,y_0}$ for $s \geq \tau$.

Finally, we remark that [Touzi, 2013, Theorem 3.3] requires $(X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\omega}}, Y^{t,y})$ to be a.s. bounded on $[t, \tau^{\boldsymbol{\pi},\boldsymbol{\omega}}]$ for inequality (E.3) to hold. This is not needed in our setting as our value function is bounded; we can take the upper semicontinuous minorant φ of V in their proof to be bounded. $\square$

E.3 Auxiliary results for the Proof of Theorem 4.1 (h)

All results in this section are stated and proven under the assumption of Theorem 4.1 parts (e)-(h), which we do not explicitly state.

Proposition E.1. *Fix $y \in \mathbb{R}^d$. If $x > 0$ and $\mathbb{E}[\int_{T_k}^{T_M} I(Y_s^{T_k,y})ds]$ are sufficiently small, then*

$$\partial_x^+ V\left(T_k, x, y\right) > \max_{j>k} \mathbb{E}\left[\frac{\alpha_j f_j'(0)}{G_j(Y_{T_j}^{T_k,y})}\right].$$

The proof uses probabilistic arguments and consists of a series of lemmas. Before proceeding, let us comment on the need of obtaining an optimal funding ratio $\boldsymbol{\theta}^*$, which is crucial in the proof of Proposition E.1. If we were to work with an ε-optimal funding ratio, the choice of small x and $\mathbb{E}[\int_{T_k}^{T_M} I(Y_s^{T_k,y})ds]$ would be ε-dependent, and it is not clear if this dependence is uniform. Although we know $\theta^*(x,y)$ is well-defined in feedback form, this is not equivalent to the existence of an optimal funding ratio $\boldsymbol{\theta}^*$ as a random variable in the sense that we do not yet know whether an optimal $\boldsymbol{\pi}^*$ and hence an optimally controlled wealth process exists. One could try to investigate the regularity of the value function V and express $\boldsymbol{\pi}^*$ in feedback form using the derivatives of V, but since our main focus is on the optimal funding ratio and an approximate optimal investment strategy can always be found by numerically solving the HJB equation, we choose not to pursue this direction. A much simpler solution is to relax the set of admissible funding ratios so that an optimizer is guaranteed to exist.

To this end, let us rewrite the value function as

$$V(t,x,y) = \sup_{\boldsymbol{\theta}\in\mathcal{R}_t(x,y)} \mathbb{E}_{t,x,y}\left[\sum_{k=1}^{M}\alpha_k f_k(\boldsymbol{\theta}_k)1_{\{T_k>t\}}\right], \tag{E.8}$$

where $\mathcal{R}_t(x,y)$ is the projection of $\mathcal{A}_t(x,y)$ onto the $\boldsymbol{\theta}$ space. By bounded convergence theorem, the supremum stays unchanged if we replace $\mathcal{R}_t(x,y)$ by its closure $\overline{\mathcal{R}_t(x,y)}$ under a.s. convergence.

Lemma E.3. *Problem* (E.8) *admits an optimizer in* $\overline{\mathcal{R}_t(x,y)}$.

Proof. We first show that $\mathcal{R}_t(x,y)$ is convex. Let $(\boldsymbol{\pi}^i,\boldsymbol{\theta}^i)\in\mathcal{A}_t(x,y)$, $i=1,2$ and let $\boldsymbol{\theta}=\lambda\boldsymbol{\theta}^1+(1-\lambda)\boldsymbol{\theta}^2$ with $\lambda\in[0,1]$. We want to show that $\boldsymbol{\theta}\in\mathcal{R}_t(x,y)$. Let $X^i=(X^i)^{t,x,y,\boldsymbol{\pi}^i,\boldsymbol{\theta}^i}$ be the wealth process controlled by $(\boldsymbol{\pi}^i,\boldsymbol{\theta}^i)$. Define $X=\lambda X^1+(1-\lambda)X^2$ and $\boldsymbol{\pi}=1_{\{X>0\}}(\lambda\boldsymbol{\pi}^1X^1+(1-\lambda)\boldsymbol{\pi}^2X^2)/X$. Then it is easy to see that $X\geq 0$, $\boldsymbol{\pi}\in\Delta_c(Y^{t,y})$ and that X is precisely the wealth process controlled by $(\boldsymbol{\pi},\boldsymbol{\theta})$. This implies $(\boldsymbol{\pi},\boldsymbol{\theta})\in\mathcal{A}_t(x,y)$, and thus $\boldsymbol{\theta}\in\mathcal{R}_t(x,y)$.

Next, let $\boldsymbol{\theta}^n\in\mathcal{R}_t(x,y)$ be an optimizing sequence for $V(t,x,y)$, say, each $\boldsymbol{\theta}^n$ is $(1/n)$-optimal. Since $\mathcal{R}_t(x,y)$ is bounded in L^0_+ (the set of non-negative random variables that is a.s. finite-valued), by [Kardaras and Zitkovic, 2013, Lemma 2.1], there exists a sequence h^n of forward convex combinations of $\boldsymbol{\theta}^n$ that converges in probability to $h\in L^0_+$. By passing to a subsequence, we may assume the convergence is almost sure. The convexity of $\mathcal{R}_t(x,y)$ implies that $h^n\in\mathcal{R}_t(x,y)$ and thus, $h\in\overline{\mathcal{R}_t(x,y)}$. It remains to check that h is optimal. Suppose $h^n=\sum_{j\geq n}\lambda_j^n\boldsymbol{\theta}^j$ with λ^n being the finite-dimensional weight vector. We then have

$$\begin{aligned}
\mathbb{E}_{t,x,y}\left[\sum_{k=1}^{M}\alpha_k f_k(h_k)1_{\{T_k>t\}}\right] &= \lim_n \mathbb{E}_{t,x,y}\left[\sum_{k=1}^{M}\alpha_k f_k(h_k^n)1_{\{T_k>t\}}\right]\\
&\geq \lim_n \sum_{j\geq n}\lambda_j^n\mathbb{E}_{t,x,y}\left[\sum_{k=1}^{M}\alpha_k f_k(\boldsymbol{\theta}_k^j)1_{\{T_k>t\}}\right]\\
&\geq \lim_n \sum_{j\geq n}\lambda_j^n(V(t,x,y)-1/j)\\
&\geq \lim_n \sum_{j\geq n}\lambda_j^n(V(t,x,y)-1/n)\\
&= \lim_n(V(t,x,y)-1/n) = V(t,x,y),
\end{aligned}$$

where the first inequality follows from the concavity of the function f_k, and the second inequality from the fact that $\boldsymbol{\theta}^j$ is $(1/j)$-optimal. The last equality is due to $\sum_{j\geq n}\lambda_j^n = 1$. □

Lemma E.4. *Let $\boldsymbol{\theta}^* \in \overline{\mathcal{R}(t,x,y)}$ be optimal for $V(t,x,y)$. Then $\sup_{k:T_k>t}\boldsymbol{\theta}_k^* \to 0$ a.s. as $x \to 0$ and $\mathbb{E}[\int_t^{T_M} I(Y_s^{t,y})ds] \to 0$.*

Proof. Let $\boldsymbol{\theta}^n \to \boldsymbol{\theta}^*$ be an optimizing sequence in $\mathcal{R}_t(x,y)$ with associated controlled wealth process $X^n := X^{t,x,y,\boldsymbol{\pi}^n,\boldsymbol{\theta}^n}$ and factor process $Y = Y^{t,y}$. Let E_s be the solution to

$$\frac{dE_s}{ds} = rE_s + I(Y_s), \qquad E_t = 0, \tag{E.9}$$

which is the time-s value of the cumulative income in the time period $[t,s]$. Recall the market price of risk process $\Lambda(Y_t)$ defined in Eq. (2.3). Let $\mathbb{Q}$ be the associated equivalent martingale measure, under which $W_s^{\mathbb{Q}} = W_s + \int_0^s \Lambda(Y_u)du$ is a Brownian motion by the Girsanov theorem. Applying Ito's formula to the process $e^{-rs}(X_s^n - E_s)$ between consecutive goal deadlines, we obtain

$$d(e^{-rs}(X_s^n - E_s)) = e^{-rs}X_s^n\pi_s^n\Sigma(Y_s)dW_s^{\mathbb{Q}}.$$

Since $\boldsymbol{\pi}^n, \Sigma(Y)$ are bounded, and X^n is square integrable, the process $(e^{-rs}(X_s^n - E_s))$ is a $\mathbb{Q}$-martingale between goal deadlines, and a $\mathbb{Q}$-supermartingale if the nonpositive jumps at goal deadlines are included. It follows from the admissibility of $\boldsymbol{\theta}^n$ and the $\mathbb{Q}$-supermartingale property of $e^{-rs}(X_s^n - E_s)$ that for any $T_k > t$,

$$\mathbb{E}^{\mathbb{Q}}[e^{-rT_k}(\boldsymbol{\theta}_k^n G_k(Y_{T_k}) - E_{T_k})] \leq \mathbb{E}^{\mathbb{Q}}[e^{-rT_k}(X_{T_k-}^n - E_{T_k})] \leq e^{-rt}x.$$

Letting $n \to \infty$, we have by Fatou's lemma that

$$\mathbb{E}^{\mathbb{Q}}[\boldsymbol{\theta}_k^* G_k(Y_{T_k})] \leq e^{r(T_k-t)}x + \mathbb{E}^{\mathbb{Q}}[E_{T_k}].$$

Plugging in the solution to equation (E.9), given by $E_{T_k} = \int_t^{T_k} e^{r(T_k-s)}I(Y_s)ds$, we obtain that the right hand side of the above inequality converges to zero as $x \to 0$ and $\mathbb{E}[\int_t^{T_M} I(Y_s)ds] \to 0$, which implies the almost sure convergence of $\boldsymbol{\theta}_k^*$ to zero. □

Proof of Proposition E.1. Let $\boldsymbol{\theta}^* \in \overline{\mathcal{R}_{T_k}(x,y)}$ be optimal for $V(T_k,x,y)$. There exists $(\boldsymbol{\pi}^n, \boldsymbol{\theta}^n) \in \mathcal{A}_{T_k}(x,y)$ such that $\boldsymbol{\theta}^n \to \boldsymbol{\theta}^*$ a.s. Let $X^n = X^{T_k,x,y,\boldsymbol{\pi}^n,\boldsymbol{\theta}^n}$ be the corresponding wealth process. Let us perturb the initial capital by some $\Delta x > 0$, i.e. suppose $X_{T_k} = x + \Delta x$.

Let $j_0 > k$ be arbitrary. Consider the following investment strategy: from initial capital x, construct a portfolio on (T_k, T_{j_0}) by collecting all the income stream $I(Y^{T_k,y})$ and following $(\boldsymbol{\pi}^n, \boldsymbol{\theta}^n)$. In a separate portfolio, the remaining capital Δx is fully invested in the risk-free asset with terminal value $\Delta x \cdot e^{r(T_{j_0}-T_k)}$ Due to the linearity of the wealth dynamics, the two portfolios can be combined into a single admissible one with $X_{T_{j_0}-} = X^n_{T_{j_0}-} + \Delta x \cdot e^{r(T_{j_0}-T_k)}$. This yields

$$\begin{aligned}
&V(T_k, x+\Delta x, y)\\
&\geq \mathbb{E}\left[\sum_{j=1}^{M} \alpha_j f_j(\boldsymbol{\theta}^n_j) 1_{\{T_j > T_k\}} + \alpha_{j_0} f_{j_0}\left(\boldsymbol{\theta}^n_{j_0} + \frac{\Delta x \cdot e^{r(T_{j_0}-T_k)}}{G_{j_0}(Y^{T_k,y}_{T_{j_0}})} \wedge (1-\boldsymbol{\theta}^n_{j_0})\right) - \alpha_{j_0} f_{j_0}(\boldsymbol{\theta}^n_{j_0})\right].
\end{aligned}$$

Letting $n \to \infty$ and using the optimality of $\boldsymbol{\theta}^*$, we obtain

$$V(T_k, x+\Delta x, y) \geq V(T_k, x, y) + \alpha_{j_0}\mathbb{E}\left[f_{j_0}\left(\boldsymbol{\theta}^*_{j_0} + \frac{\Delta x \cdot e^{r(T_{j_0}-T_k)}}{G_{j_0}(Y^{T_k,y}_{T_{j_0}})} \wedge (1-\boldsymbol{\theta}^*_{j_0})\right) - f_{j_0}(\boldsymbol{\theta}^*_{j_0})\right].$$

It then follows from the concavity of V in x (see Theorem 4.1 (e)) that

$$\begin{aligned}
\partial_x^+ V(T_k, x, y) &\geq \frac{V(T_k, x+\Delta x, y) - V(T_k, x, y)}{\Delta x}\\
&\geq \frac{\alpha_{j_0}}{\Delta x}\mathbb{E}\left[f_{j_0}\left(\boldsymbol{\theta}^*_{j_0} + \frac{\Delta x \cdot e^{r(T_{j_0}-T_k)}}{G_{j_0}(Y^{T_k,y}_{T_{j_0}})} \wedge (1-\boldsymbol{\theta}^*_{j_0})\right) - f_{j_0}(\boldsymbol{\theta}^*_{j_0})\right].
\end{aligned}$$

Applying Lemma E.4 with $t = T_k$, we get that $\boldsymbol{\theta}^*_{j_0} \to 0$ a.s. as $x \to 0$ and $\mathbb{E}[\int_{T_k}^{T_M} I(Y^{T_k,y}_s)ds] \to 0$. Let $\gamma \in (e^{-r(T_{j_0}-T_k)}, 1)$. Dominated convergence theorem then implies that if x and $\mathbb{E}[\int_{T_k}^{T_M} I(Y^{T_k,y}_s)ds]$ are sufficiently small, then

$$\partial_x^+ V(T_k, x, y) > \frac{\alpha_{j_0}}{\Delta x}\gamma\mathbb{E}\left[f_{j_0}\left(\frac{\Delta x \cdot e^{r(T_{j_0}-T_k)}}{G_{j_0}(Y^{T_k,y}_{T_{j_0}})} \wedge 1\right) - f_{j_0}(0)\right].$$

For such x and $\mathbb{E}[\int_{T_k}^{T_M} I(Y_s^{T_k,y})ds]$, we have

$$\partial_x^+ V(T_k,x,y) \geq \lim_{\Delta x\to 0} \frac{\alpha_{j_0}}{\Delta x}\gamma \mathbb{E}\left[f_{j_0}\left(\frac{\Delta x \cdot e^{r(T_{j_0}-T_k)}}{G_{j_0}(Y_{T_{j_0}}^{T_k,y})} \wedge 1 \right) - f_{j_0}(0) \right]$$
$$= \alpha_{j_0}\gamma \mathbb{E}\left[f'_{j_0}(0) \frac{e^{r(T_{j_0}-T_k)}}{G_{j_0}(Y_{T_{j_0}}^{T_k,y})} \right] > \alpha_{j_0} f'_{j_0}(0) \mathbb{E}\left[\frac{1}{G_{j_0}(Y_{T_{j_0}}^{T_k,y})} \right],$$

where in the equality above we have used the dominated convergence theorem to exchange the limit and the expectation. □

F Proof of Theorem 4.1

Proof of Parts (a)-(d). Clearly, $V(T_M,\cdot,\cdot)\equiv 0$ is continuous. Suppose we have already obtained the continuity of $V(T_k,\cdot,\cdot)$ if $k=M$ and the continuity of V in $\mathbf{S}_{k+1}$ if $k<M$. Write (4.4) as

$$g_k(x,y) = \max_{\theta\in\varphi(x,y)} h(x,y,\theta)$$

where $h(x,y,\theta) = \alpha_k f_k(\theta) + V\left(T_k, x - G_k(y)\theta, y\right)$ and $\varphi(x,y) = \left\{\theta\in[0,1] : G_k(y)\theta \leq x^+\right\}$. Clearly, $h : \mathbb{R}\times\mathbb{R}^d\times[0,1]\to\mathbb{R}$ is continuous, thanks to the continuity of $f_k(\cdot)$, $V(T_k,\cdot,\cdot)$ and $G_k(\cdot)$. Moreover, φ is a continuous set-valued map with non-empty compact values by Theorems 17.20 and 17.21 in Aliprantis and Border [2006]. Thus, (a) and (b) are direct consequences of Berge's maximum theorem. We next prove (c), and then (d) which includes the continuity of V in $\mathbf{S}_k$.

First claim that

$$V(t,x,y) = \sup_{\boldsymbol{\pi}} \mathbb{E}\left[g_k(X_{T_k-}^{t,x,y,\boldsymbol{\pi}}, Y_{T_k}^{t,y}) \right], \quad (t,x,y)\in\mathbf{S}_k. \tag{F.1}$$

If $k=M$, this follows directly from optimizing over $\boldsymbol{\theta}_M$ in the definition of V. If $k<M$, for some small $\varepsilon>0$, we apply the weak DPP (Lemma E.2) to get

$$V(t,x,y) \leq \sup_{(\boldsymbol{\pi},\boldsymbol{\theta})\in\mathcal{A}_t(x,y)} \mathbb{E}\left[V^*(T_k+\varepsilon, X_{T_k+\varepsilon}^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}, Y_{T_k+\varepsilon}^{t,y}) \right]$$

and

$$V(t,x,y) \geq \sup_{(\boldsymbol{\pi},\boldsymbol{\theta})\in\mathcal{A}_t(x,y)} \mathbb{E}\left[V_*(T_k+\varepsilon, X_{T_k+\varepsilon}^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}, Y_{T_k+\varepsilon}^{t,y}) \right].$$

Since V is continuous in $\mathbf{S}_{k+1}$, $V^* = V_* = V$ at time $T_k + \varepsilon$, So the weak DPP simplifies to the classical DPP: $V(t,x,y) = \sup_{(\boldsymbol{\pi},\boldsymbol{\theta})\in\mathcal{A}_t(x,y)} \mathbb{E}[V(T_k+\varepsilon, X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}_{T_k+\varepsilon}, Y^{t,y}_{T_k+\varepsilon})]$. We then let $\varepsilon \to 0$ and use the continuity and boundedness of V in $\mathbf{S}_{k+1}$ to deduce

$$V(t,x,y) = \sup_{(\boldsymbol{\pi},\boldsymbol{\theta})\in\mathcal{A}_t(x,y)} \mathbb{E}\left[V(T_k, X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}}_{T_k}, Y^{t,y}_{T_k})\right].$$

Further optimizing over $\boldsymbol{\theta}_k$ verifies the claim (F.1).

In the proof of parts (c) and (d) below, we focus on the subproblem (F.1). In view of Lemma E.1, we take (E.1) as the defining equation for the wealth process on $[T_{k-1}, T_k)$.

Part (c). By the definition of upper (resp. lower) semicontinuous envelope, there exists a sequence $(t_n, x_n, y_n) \in \mathbf{S}_k$ such that $(t_n, x_n, y_n) \to (T_k, x, y)$ and $V(t_n, x_n, y_n) \to V^*(T_k, x, y)$ (resp. $V(t_n, x_n, y_n) \to V_*(T_k, x, y)$). We want to show $\lim_n V(t_n, x_n, y_n) = g_k(x, y)$.

Let $(\boldsymbol{\pi}^n, \boldsymbol{\theta}^n)$ be an $(1/n)$-optimal control for $V(t_n, x_n, y_n)$. It follows from (F.1) and the definition of an $(1/n)$-optimal control that

$$\mathbb{E}\left[g_k(X^{t_n,x_n,y_n,\boldsymbol{\pi}^n}_{T_k-}, Y^{t,y_n}_{T_k})\right] \le V(t_n, x_n, y_n) \le \mathbb{E}\left[g_k(X^{t_n,x_n,y_n,\boldsymbol{\pi}^n}_{T_k-}, Y^{t,y_n}_{T_k})\right] + \frac{1}{n}. \tag{F.2}$$

By standard SDE estimates,

$$\begin{aligned}
&\mathbb{E}\left[|(X^{t_n,x_n,y_n,\boldsymbol{\pi}^n}_{T_k-}, Y^{t_n,y_n}_{T_k}) - (x,y)|^2\right] \\
&\le 2\mathbb{E}\left[|(X^{t_n,x_n,y_n,\boldsymbol{\pi}^n}_{T_k-}, Y^{t_n,y_n}_{T_k}) - (x_n,y_n)|^2\right] + 2(x_n - x)^2 + 2(y_n - y)^2 \\
&\le 2Ce^{C(T_k - t_n)}(|x_n|^2 + |y_n|^2 + C)(T_k - t_n) + 2(x_n - x)^2 + 2(y_n - y)^2,
\end{aligned}$$

for some constant C independent of n, where we have used the uniform boundedness of $\boldsymbol{\pi}^n$ (recall that $\boldsymbol{\pi}^n \in \Delta$). The above implies that $(X^{t_n,x_n,y_n,\boldsymbol{\pi}^n}_{T_k-}, Y^{t_n,y_n}_{T_k}) \to (x,y)$ a.s. as $n \to \infty$. Together with the continuity and boundedness of g_k, we obtain the desired limit by letting $n \to \infty$ in (F.2).

Part (d). Observe that F is continuous, by the continuity of the set-valued map $\Delta_c(\cdot)$ (see [Aliprantis and Border, 2006, Theorems 17.20 and 17.21]) and Berge's Maximum Theorem. By [Touzi, 2013, Theorem 7.4], the bounded function V is a viscosity solution to (4.5)-(4.6); that is, V^* is a viscosity subsolution and V_* is a viscosity supersolution. As in Lemma E.2, their proof needs to be slightly modified to accommodate the volatility constraint. Specifically, in the supersolution proof (see [Touzi, 2013, Proposition 7.2]), the constant control process needs to be replaced by the feedback control $\boldsymbol{\pi} = (1 \wedge \frac{c(Y)}{\|\pi\Sigma(Y)\|})\pi$ where $\pi \in \Delta$ is any constant vector satisfying $\|\pi\Sigma(y)\| \le c(y)$.

It remains to show uniqueness and continuity by the comparison principle. To simplify notation, let $\mathcal{S}_n$ be the set of $n \times n$ symmetric real matrices, and write $F(x, y, p, q, g, A, B) = F(x, y, p, q, Z)$ with

$$Z = \begin{pmatrix} A & g^\top \\ g & B \end{pmatrix} \in \mathcal{S}_{1+d}.$$

We first show that there exists a function $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $\omega(0+) = 0$ such that

$$\begin{aligned} &F(x', y', \beta(x - x'), \beta(y - y'), Z') - F(x, y, \beta(x - x'), \beta(y - y'), Z) \\ &\le \omega(\beta|x - x'|^2 + \beta|y - y'|^2 + |x - x'| + |y - y'|) \end{aligned} \tag{F.3}$$

whenever $(x, y), (x', y') \in \mathbb{R}^{1+d}$, $Z, Z' \in \mathcal{S}_{1+d}$, and

$$-3\beta \begin{pmatrix} I_{1+d} & 0 \\ 0 & I_{1+d} \end{pmatrix} \le \begin{pmatrix} Z & 0 \\ 0 & -Z' \end{pmatrix} \le 3\beta \begin{pmatrix} I_{1+d} & -I_{1+d} \\ I_{1+d} & I_{1+d} \end{pmatrix}.[10] \tag{F.4}$$

As before, C denotes a constant that only depends on the model input and that may vary from line to line. By the Lipschitz continuity of $I(\cdot)$ and $\mu_Y(\cdot)$, we have

$$\begin{aligned} &F(x', y', \beta(x - x'), \beta(y - y'), Z') - F(x, y, \beta(x - x'), \beta(y - y'), Z) \\ &= \beta(x - x')\left[I(y) - I(y') + r(x - x')\right] + \beta(y - y')(\mu_Y^\top(y) - \mu_Y^\top(y')) + H_1 - H_2 \\ &\le C\beta(|x - x'|^2 + |y - y'|^2) + H_1 - H_2 \end{aligned}$$

where

$$\begin{aligned} H_1 &= \sup_{\pi \in \Delta_c(y)} \left\{ \beta(x - x')\rho(x)\pi(\mu(y) - r\mathbb{1}_N) + \frac{1}{2}\operatorname{Tr}\left(\begin{pmatrix} \rho(x)\pi\Sigma(y) \\ \Sigma_Y(y) \end{pmatrix} \begin{pmatrix} \rho(x)\pi\Sigma(y) \\ \Sigma_Y(y) \end{pmatrix}^\top Z \right) \right\}, \\ H_2 &= \sup_{\pi \in \Delta_c(y')} \left\{ \beta(x - x')\rho(x')\pi(\mu(y') - r\mathbb{1}_N) + \frac{1}{2}\operatorname{Tr}\left(\begin{pmatrix} \rho(x')\pi\Sigma(y') \\ \Sigma_Y(y') \end{pmatrix} \begin{pmatrix} \rho(x')\pi\Sigma(y') \\ \Sigma_Y(y') \end{pmatrix}^\top Z' \right) \right\}. \end{aligned}$$

[10] For $Z, Z' \in \mathcal{S}_n$, the relation $Z \ge Z'$ means $Z - Z'$ is positive semidefinite.

Let $\pi^* \in \Delta_c(y)$ be optimal for H_1. Then $\tilde{\pi} := (1 \wedge \frac{c(y')}{\|\pi^*\Sigma(y')\|})\pi^* \in \Delta_c(y')$. We thus have

$$H_1 - H_2 \leq \beta(x-x')\rho(x)\pi^*(\mu(y) - r\mathbb{1}_N) + \frac{1}{2}\operatorname{Tr}\left(\begin{pmatrix}\rho(x)\pi^*\Sigma(y)\\ \Sigma_Y(y)\end{pmatrix}\begin{pmatrix}\rho(x)\pi^*\Sigma(y)\\ \Sigma_Y(y)\end{pmatrix}^\top Z\right)$$
$$- \beta(x-x')\rho(x')\tilde{\pi}(\mu(y') - r\mathbb{1}_N) - \frac{1}{2}\operatorname{Tr}\left(\begin{pmatrix}\rho(x')\tilde{\pi}\Sigma(y')\\ \Sigma_Y(y')\end{pmatrix}\begin{pmatrix}\rho(x')\tilde{\pi}\Sigma(y')\\ \Sigma_Y(y')\end{pmatrix}^\top Z'\right).$$

Using (F.4), the boundedness of π^*, the bounded Lipschitz property of $\rho(\cdot)$, $\mu(\cdot)$, $\Sigma(\cdot)$, and the Lipschitz continuity of $\Sigma_Y(\cdot)$, we get

$$\begin{aligned} H_1 - H_2 &\leq \beta(x-x')C(|x-x'| + |y-y'| + |\pi^* - \tilde{\pi}|) \\ &\quad + \frac{3\beta}{2}\left\|\begin{pmatrix}\rho(x)\pi^*\Sigma(y)\\ \Sigma_Y(y)\end{pmatrix} - \begin{pmatrix}\rho(x')\tilde{\pi}\Sigma(y')\\ \Sigma_Y(y')\end{pmatrix}\right\|^2 \\ &\leq C\beta(|x-x'|^2 + |y-y'|^2 + |\pi^* - \tilde{\pi}|^2). \end{aligned}$$

Observe that $\pi^* \in \Delta_c(y)$ implies $\|\pi^*\Sigma(y)\| \leq c(y)$, from which we deduce

$$\|\pi^*\Sigma(y')\| - c(y') = \|\pi^*\Sigma(y')\| - \|\pi^*\Sigma(y)\| + \|\pi^*\Sigma(y)\| - c(y) + c(y) - c(y') \leq C|y-y'|.$$

Moreover, it follows directly from the definition of $\tilde{\pi}$ that

$$|\pi^* - \tilde{\pi}| = |\pi^*|\frac{(\|\pi^*\Sigma(y')\| - c(y'))^+}{\|\pi^*\Sigma(y')\| \vee c(y')} \leq \frac{C|y-y'|}{c(y')} \leq C|y-y'|.$$

Putting everything together, we have verified (F.3)-(F.4). By the viscosity comparison principle (see e.g. [Pham, 2009, Theorem 4.4.5][11]), $V^* = V_* = g_k$ at time T_k implies $V^* \leq V_*$ in $\mathbf{S}_k$, from which the continuity of V follows (since $V^* = V_* = V$). The comparison principle additionally implies the uniqueness of viscosity solution with polynomial growth.

Proof of Part (e). Fix $(t,y) \in [0,T_M]\times\mathbb{R}^d$. Let $x = \lambda x_1 + (1-\lambda)x_2$ for some $x_1, x_2 \geq 0$ and $\lambda \in [0,1]$. For $i = 1,2$, let $(\boldsymbol{\pi}^i, \boldsymbol{\theta}^i)$ be an ε-optimal control for $V(t,x_i,y)$, and $X^i = X^{t,x_i,y,\boldsymbol{\pi}^i,\boldsymbol{\theta}^i}$

[11] [Pham, 2009, Theorem 4.4.5] is proven for problems where the control takes values in a constant set, but the proof also works for factor-dependent control space, as long as the Hamiltonian F satisfies (F.3)-(F.4).

be the corresponding controlled wealth process. We have

$$V(t,x_i,y) \le \mathbb{E}\left[\sum_{k=1}^{M} \alpha_k f_k(\boldsymbol{\theta}_k^i, y) 1_{\{T_k>t\}}\right] + \varepsilon.$$

Let $X := \lambda X^1 + (1-\lambda)X^2$, $\boldsymbol{\pi} := 1_{\{X>0\}}(\lambda \boldsymbol{\pi}^1 X^1 + (1-\lambda)\boldsymbol{\pi}^2 X^2)/X$ and $\boldsymbol{\theta} := \lambda\boldsymbol{\theta}^1 + (1-\lambda)\boldsymbol{\theta}^2$. It is easy to see that $X = X^{t,x,y,\boldsymbol{\pi},\boldsymbol{\theta}} \ge 0$, $\boldsymbol{\pi} \in \Delta_c(Y^{t,y})$ and $\boldsymbol{\theta} \in [0,1]$. So $(\boldsymbol{\pi},\boldsymbol{\theta}) \in \mathcal{A}_t(x,y)$. It follows that

$$\begin{aligned}
V(t,x,y) &\ge \mathbb{E}\left[\sum_{k=1}^{M} \alpha_k f_k(\boldsymbol{\theta}_k) 1_{\{T_k>t\}}\right] \\
&\ge \lambda \mathbb{E}\left[\sum_{k=1}^{M} \alpha_k f_k(\boldsymbol{\theta}_k^1) 1_{\{T_k>t\}}\right] + (1-\lambda)\mathbb{E}\left[\sum_{k=1}^{M} \alpha_k f_k(\boldsymbol{\theta}_k^2) 1_{\{T_k>t\}}\right] \\
&\ge \lambda V(t,x_1,y) + (1-\lambda)V(t,x_2,y) - \varepsilon,
\end{aligned}$$

where we used the concavity of f_k in the second step. Since $\varepsilon > 0$ is arbitrary, we conclude that V is concave in x.

Proof of Part (f). Any optimizer θ_k of (4.4) is characterized by the generalized version of the Karush-Kuhn-Tucker (KKT) conditions:

$$\begin{aligned}
&\text{(Stationarity)} \quad 0 = -\alpha_k f_k'(\theta_k) + vG_k(y) + s - t + \lambda G_k(y) \\
&\qquad\qquad\qquad \text{for some } v \in [\partial_x^+ V, \partial_x^- V](T_k, x - G_k(y)\theta_k, y); && \text{(F.5)}\\
&\text{(Primal feasibility)} \quad 0 \le \theta_k \le 1, \quad G_k(y)\theta_k \le x; && \text{(F.6)}\\
&\text{(Dual feasibility)} \quad s,t,\lambda \ge 0; && \text{(F.7)}\\
&\text{(Complementary Slackness)} \quad s(1-\theta_k) = 0, \quad t\theta_k = 0, \quad \lambda(x - G_k(y)\theta_k) = 0. && \text{(F.8)}
\end{aligned}$$

Let $\theta_0 := (\inf A) \wedge 1 \wedge (x/G_k(y))$ with $A := \{\theta \in [0, x/G_k(y)] : \partial_x^+ V(T_k, x - G_k(y)\theta, y) > \alpha_k f_k'(\theta)/G_k(y)\}$, and let v, s, t, λ be the subdifferential and dual variables in the generalized KKT condition for $\theta_k^*(x,y)$. We want to show $\theta_k^*(x,y) = \theta_0$.

1) "$\theta_k^*(x,y) \le \theta_0$": Suppose on the contrary, $\theta_k^*(x,y) > \theta_0$. By the primal feasibility condition $\theta_k^*(x,y) \le 1 \wedge (x/G_k(y))$, we must have that $\theta_0 = \inf A$. By the definition of infimum, $\theta_k^*(x,y) \ge \theta_1$ for some $\theta_1 \in A$. It then follows from the monotonicity of $\partial_x^+ V$ and

f_k' (implied by concavity) that

$$\partial_x^+ V(T_k, x-G_k(y)\theta_k^*(x,y), y) \geq \partial_x^+ V(T_k, x-G_k(y)\theta_1, y) > \frac{\alpha_k}{G_k(y)} f_k'(\theta_1) \geq \frac{\alpha_k}{G_k(y)} f_k'(\theta_k^*(x,y)).$$

Hence, $-\alpha_k f_k'(\theta_k^*(x,y))+vG_k(y) \geq -\alpha_k f_k'(\theta_k^*(x,y))+\partial_x^+ V(T_k, x-G_k(y)\theta_k^*(x,y), y)G_k(y) > 0$. By the stationarity condition (F.5) and dual feasibility, we then have $t = -\alpha_k f'(\theta_k^*(x,y)) + vG_k(y) + s + \lambda G_k(y) > 0$. By the complementary slackness condition (F.8), it holds that $\theta_k^*(x,y) = 0$, contradicting the assumption that $\theta_k^*(x,y) > \theta_0 \geq 0$.

2) "$\theta_k^*(x,y) \geq \theta_0$": If $\theta_k^*(x,y) = 1 \wedge (x/G_k(y))$, then $\theta_k^*(x) \geq \theta_0$ holds trivially. Hence, we assume $\theta_k^*(x,y) < 1 \wedge (x/G_k(y))$. By complementary slackness, this implies that $s = \lambda = 0$.

We claim that $\alpha_k f_k'(\theta_k^*(x,y))/G_k(y) \leq \partial_x^- V(T_k, x-G_k(y)\theta_k^*(x,y), y)$. Suppose on the contrary, $\alpha_k f_k'(\theta_k^*(x,y))/G_k(y) > \partial_x^- V(T_k, x-G_k(y)\theta_k^*(x,y), y)$, then $\alpha_k f_k'(\theta_k^*(x,y)) - vG_k(y) \geq \alpha_k f_k'(\theta_k^*(x,y)) - \partial_x^- V(T_k, x-G_k(y)\theta_k^*(x,y), y)G_k(y) > 0$. The stationarity and dual feasibility conditions then imply $s + \lambda G_k(y) = \alpha_k f'(\theta_k^*(x,y)) - vG_k(y) + t > 0$, which contradicts $s = \lambda = 0$.

Next, let $\tilde\theta \in (\theta_k^*(x,y), 1 \wedge (x/G_k(y)))$ be arbitrary. Then $\tilde\theta$ is primal feasible and trivially satisfies the complimentary slackness conditions $s(1-\tilde\theta) = \lambda(x - G_k(y)\tilde\theta) = 0$ because $s = \lambda = 0$. By the concavity of V in x and the concavity of f_k in θ together with the previous claim, we have $\alpha_k f_k'(\tilde\theta)/G_k(y) \leq \alpha_k f_k'(\theta_k^*(x,y))/G_k(y) \leq \partial_x^- V(T_k, x-G_k(y)\theta_k^*(x,y), y) \leq \partial_x^+ V(T_k, x-G_k(y)\tilde\theta, y)$. If $\alpha_k f_k'(\tilde\theta)/G_k(y) = \partial_x^+ V(T_k, x-G_k(y)\tilde\theta, y)$, then by setting $\tilde v := \alpha_k f_k'(\tilde\theta)/G_k(y)$ and $\tilde t := 0$, we have $0 = -\alpha_k f_k'(\tilde\theta) + \tilde v G_k(y) + s - \tilde t + \lambda G_k(y)$ and $\tilde t\tilde\theta = 0$. This shows that $(\tilde\theta, \tilde v, s, \tilde t, \lambda)$ satisfies the generalized KKT condition. Hence $\tilde\theta$ is also optimal, contradicting our assumption that $\theta_k^*(x,y)$ is the largest optimizer. Consequently, we must have $\alpha_k f_k'(\tilde\theta)/G_k(y) < \partial_x^+ V(T_k, x-G_k(y)\tilde\theta, y)$ and thus, $\tilde\theta \geq \theta_0$. But $\tilde\theta \in (\theta_k^*(x,y), 1 \wedge (x/G_k(y)))$ is arbitrary. Letting $\tilde\theta \searrow \theta_k^*(x,y)$ gives us $\theta_k^*(x,y) \geq \theta_0$.

Proof of Part (g). Fix $y \in \mathbb{R}^d$ and let $b_k(y)$ be defined as in the theorem statement. We have $\theta_k^*(0,y) = 0$ by the primal feasibility condition. For $x \in (0, b_k(y)]$, by the definition of supremum and the positivity of $G_k(\cdot)$, for any $\epsilon > 0$ we have that $x - \epsilon G_k(y) < z$ for some z satisfying $\partial_x^+ V(T_k, z, y) > \alpha_k f_k'(0)/G_k(y)$. The concavity of $V(T_k, \cdot, y)$ and f_k then implies $\partial_x^+ V(T_k, x-\epsilon G_k(y), y) \geq \partial_x^+ V(T_k, z, y) > \alpha_k f_k'(0)/G_k(y) \geq \alpha_k f_k'(\epsilon)/G_k(y)$. By Theorem 4.1 (e), we deduce that $\theta_k^*(x,y) \leq \epsilon$. Sending $\epsilon \searrow 0$ conclude the proof that $\theta_k^*(x,y) = 0$.

Proof of Part (h) Fix $y \in \mathbb{R}^d$. For any $x > \mathfrak{s}(T_k, y)$, $V(T_k, x, y) = \sum_{j>k} \alpha_j f_j(1)$ and $\partial_x V(T_k, x, y) = 0 \leq \alpha_k f_k'(0)/G_k(y)$. Hence, $b_k(y) \leq \mathfrak{s}(T_k, y)$. Suppose $\alpha_k f_k'(0)/G_k(y) \leq \max_{j>k} \mathbb{E}[\alpha_j f_j'(0)/G_j(Y_{T_j}^{T_k,y})]$. By Proposition E.1, if $x> 0$ and $\mathbb{E}[\int_{T_k}^{T_M} I(Y_s^{T_k,y})ds]$ are sufficiently small, we have

$$\partial_x^+ V\left(T_k, x, y\right) > \max_{j>k} \mathbb{E}\left[\frac{\alpha_j f_j'(0)}{G_j(Y_{T_j}^{T_k,y})}\right] \geq \frac{\alpha_k f_k'(0)}{G_k(y)}.$$

It follows that $b_k(y) \geq x > 0$. □

References

C. D. Aliprantis and K. C. Border. *Infinite dimensional analysis : a hitchhiker's guide.* Springer-Verlag Berlin Heidelberg, 3rd edition, 2006.

J.-P. Aubin and H. Frankowska. *Set-valued analysis.* Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2009. ISBN 978-0-8176-4847-3. doi: 10.1007/978-0-8176-4848-0. Reprint of the 1990 edition [MR1048347].

G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991. ISSN 0921-7134.

J. Brunel. *Goals-Based Wealth Management: An Integrated and Practical Approach to Changing the Structure of Wealth Advisory Practices.* Wiley, New York, 2015.

A. Chhabra. Beyond markowitz: A comprehensive wealth allocation framework for individual investors. *Journal of Wealth Management*, 7(4):8–34, 2005.

G. Consigli and M. Dempster. Planning for retirement: Asset liability management for individuals. *Annals of Operations Research*, 81:1116–1132, 1998.

M. G. Crandall, H. Ishii, and P. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of American Mathematical Society*, 27(1):1–67, 1992.

J. Cvitanic, S. Kou, X. Wan, and K. Williams. Pi portfolio management: Reaching goals while avoiding drawdowns. *Working Paper*, 2020.

M. Dai, S. Kou, S. Qian, and X. Wane. Nonconcave utility maximization with portfolio bounds. *Management Science*, 75(3):1327–1370, 2020.

S. Das, H. Markowitz, J. Scheid, and M. Statman. Portfolio optimization with mental accounts. *Journal of Financial and Quantitative Analysis*, 45(2):331–334, 2010.

S. Das, D. Ostrov, A. Radhakrishnan, and D. Srivastav. A new approach to goals-based wealth management. *Journal of Investment Management*, 16(3):1—27, 2018.

S. Das, D. Ostrov, A. Radhakrishnan, and D. Srivastav. Dynamic portfolio allocation in goals-based wealth management. *Computational Management Science*, 17:613–640, 2020.

S. Das, D. Ostrov, A. Radhakrishnan, and D. Srivastav. Dynamic optimization for multi-goals wealth management. *Journal of Banking and Finance*, 140:106192, 2022.

R. Deguest and L. Martellini. *Goal-based Investing: Theory and Practice*. World Scientific, 2021.

M. Dempster and E. Medova. Planning for retirement: Asset liability management for individuals. In *Asset and Liability Management Handbook*, pages 409–432.

M. Dempster, E. Germano, E. Medova, and M. Villaverde. Global asset liability management. *British Actuarial Journal*, 9(1):1116–1132, 2003.

M. Denault and J. Simonato. A note on a dynamic goal-based wealth management problem. *Finance Research Letters*, 46:102404, 2022.

M. Dixon and I. Alperin. Goal-based wealth management with generative reinforcement learning. *Risk*, July 2021.

E. Fama and G. Schwert. Asset returns and inflation. *Journal of Financial Economics*, 5 (5):115–146, 1977.

P. A. Forsyth and G. Labahn. Numerical methods for controlled hamilton-jacobi-bellman pdes in finance. *Journal of Computational Finance*, 11(2):1–43, 2007/08.

A. Gargano and A. Rossi. Goal setting and saving in the Fintech era. *Journal of Finance, Forthcoming*, 2023.

R. Jensen. The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations. *Arch. Rational Mech. Anal.*, 101(1):1–27, 1988. ISSN 0003-9527. doi: 10.1007/BF00281780.

C. Kardaras and G. Zitkovic. Forward-convex convergence of sequences of nonnegative random variables. *Proc. Amer. Math. Soc.*, pages 919–929, 2013.

M. Leibowitz, A. Lo, R. Merton, S. Ross, and J. Siegel. Q group panel discussion: Looking to the future, 2016.

E. A. Locke and G. P. Latham. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9):705, 2002.

J. Mulvey and H. Vladimirou. Stochastic network programming for financial planning problems. *Management Science*, 38(11):1642–1664, 2012. ISSN 0363-0129. doi: 10.1137/110852942.

D. Nevins. Goals-based investing: Integrating traditional behavioral finance. *Journal of Wealth Management*, 6(4):8–23, 2004.

A. Pagan and G. Schwert. Alternative models for conditional stock volatility. *Journal of Econometrics*, 45(45):267–290, 1990.

H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*, volume 61. Springer Verlag, Berlin, 2009.

A. Ruszczynski. *Nonlinear Optimization*. Princeton University Press, 2006.

H. Shefrin and M. Statman. Behavioral portfolio theory. *Journal of Financial and Quantitative Analysis*, 35(2):127–151, 2000.

R. Thaler. Mental accounting and consumer choice. *Marketing Science*, 4:199–214, 1985a.

R. Thaler. Mental accounting matters. *Journal of Behavioral Decision Making*, 12:183–206, 1985b.

N. Touzi. *Optimal stochastic control, stochastic target problems, and backward SDE*, volume 29 of *Fields Institute Monographs*. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2013. ISBN 978-1-4614-4285-1; 978-1-4614-4286-8. doi: 10.1007/978-1-4614-4286-8. With Chapter 13 by Angès Tourin.