Naïve Markowitz Policies*

Lin Chen† Xun Yu Zhou‡

January 3, 2024

Abstract

We study a continuous-time Markowitz mean–variance portfolio selection model in which a naïve agent, unaware of the underlying time-inconsistency, continuously reoptimizes over time. We define the resulting naïve policies through the limit of discretely naïve policies that are committed only in very small time intervals, and derive them analytically and explicitly. We compare naïve policies with pre-committed optimal policies and with consistent planners' equilibrium policies in a Black–Scholes market, and find that the former achieve higher expected terminal returns than originally planned yet are mean–variance inefficient when the risk aversion level is sufficiently small, and always take strictly riskier exposure than equilibrium policies. We finally define an efficiency ratio for comparing return–risk tradeoff with the same original level of risk aversion, and show that naïve policies are always strictly less efficient than pre-committed and equilibrium policies.

Key Words. Continuous time, mean–variance model, time inconsistency, naïve agent, pre-committed agent, consistent planner, equilibrium policies.

1 Introduction

The Markowitz mean–variance (MV) portfolio selection model (Markowitz, 1952 and Markowitz, 1959) is a monumental work in quantitative finance. The model formulates the investment problem as striving to achieve the best balance between return and risk, represented respectively by the mean and variance of the final portfolio worth. Its variants, extensions and implications have been passionately studied in theory and applied in practice to this day.

The original MV model is formulated for a static single period and solved by quadratic program. It is natural and necessary to extend it to the dynamic setting, both in discrete time and continuous time. However, a dynamic MV model is inherently *time inconsistent*; namely, any "optimal" policy for the present moment will generally not be optimal for the next

*The first version of the paper was completed in 2017, and part of the results was included in the first author's PhD thesis defended in 2020. The paper was finalized when the second author was on vacation in Las Vegas, a place arguably ideal for observing the "naïve" behaviors studied in the paper.

†Department of Industrial Engineering and Operations Research, Columbia University, New York, New York 10027, USA, lc3110@columbia.edu

‡Department of Industrial Engineering and Operations Research, Columbia University, New York, New York 10027, USA, xz2574@columbia.edu

moment.[1] This inconsistency comes from the variance term that does not satisfy the tower rule: unlike the mean, there is no consistency over time in evaluating the same variance of the final wealth. As a result, in sharp contrast to the classical time-consistent models, there is no such notion as a *dynamically optimal policy* for a time-inconsistent model because any such policy, once planned for this moment, may need to be given up quickly (and *instantly* in a continuous-time setting) in favor of a different plan at the next moment. Technically, time-inconsistency poses fundamental challenges in "solving" – whatever "solving" means – the problem because the Bellman optimality principle, which is the very foundation of the classical dynamic programming for studying dynamic optimization problems, is no longer valid.

Economists have recognized and studied time-inconstancy since as early as the 1950s. The foundational paper Strotz (1956) *describes* three types of agents when facing time inconsistency. Type 1, a "naïveté" (or naïf), is unaware of the time inconsistency and at any given time and state of affairs seeks an "optimal" policy for that moment only, without knowing that he will not uphold that policy for long. As a result, his policies change all the times, and the eventual policy that is being actually carried out *ex post* can be vastly and characteristically different from any of his short-lived "optimal" policies he originally planned to execute.[2] The next two types realize the issue of time inconsistency but act differently. Type 2 is a "pre-committer" who solves the optimization problem only at time 0 and sticks to the resulting policy throughout (via some "commitment device" if necessary and available), recognizing that the original policy may no longer be optimal at later times. Type 3 is a "consistent planner" who is unable to precommit and understands that her future selves may abandon whatever plans she makes now. Her resolution is to optimize taking the future deviations from the current plan as *constraints*, effectively leading to a game among selves at different times. The resulting policies are called equilibrium ones.

It is important to note that it is not meaningful to determine which type is superior than the others, simply because superiority depends on what criteria one uses – and now there are no uniform criteria as in the time-consistent paradigm.[3] In this sense, the Strotzian approach to time-inconsistency is both *normative* (i.e. to advise people about the best course of actions, especially in Types 2 and 3) and *descriptive* (i.e. to describe what people are actually doing, as more with Type 1).

Mathematically, model formulations and solutions for deriving the three types of agent policies call for different treatments as they are very different from each other. The problems are also challenging due to the invalidity of the dynamic programming approach. In the last decade, there have been significant developments in studying time-inconsistent models analyt-

[1] Here a "policy" is a *plan* that maps any given time and state to an action (a portfolio in the MV model). It is also called a *feedback control law* in control theory.

[2] For instance, Barberis (2012) shows, in a casino gambling model (which is time-inconsistent in discrete time due to probability weighting), a naïve gambler's initial plan was to gamble as long as possible when winning but to stop if he started accumulating losses, he actually ends up doing the *opposite*: he gambles as long as possible when losing and stops once he accumulates some gains. Similar behaviors are also observed, and indeed prevalent, in stock investment especially with retail investors.

[3] In a recent note Zhang (2023), it is shown via a simple example that an equilibrium policy may not be Pareto optimal, i.e. there exists another policy that is no worse for any self but strictly better for at least one self, compared with the equilibrium policy.

ically, mainly in three different settings: MV portfolio selection, and optimization problems involving non-exponential discounting or probability weighting; see He and Zhou (2022) for a recent survey on the related works. For the MV models, earlier works focused on Type 2, pre-committed agents; see, e.g., Richardson (1989); Hakansson (1971); Li and Ng (2000); Zhou and Li (2000); Lim and Zhou (2002); Bielecki et al. (2005); Xia (2005); Li and Zhou (2006), although most of these works did not spell out that their solutions were pre-committed ones. Later research gradually shifted to Type 3, consistent planners; see, e.g. Basak and Chabakauri (2010); Hu et al. (2012); Björk and Murgoci (2014); Björk et al. (2014); He and Jiang (2022).

In contrast to the hitherto rich literature on pre-committed agents and consistent planners, there are far fewer works on the general behaviors of naïve agents, and almost none in continuous time (not necessarily limited to MV models).[4] Barberis (2012) and Hu et al. (2022) study naïve strategies in casino gambling models which are inherently discrete time. As shown in these papers, finding naïve policies in discrete time is rather straightforward technically if the pre-committed polices are already available: at each discrete time point one solves and obtains the corresponding pre-committed policy, holds it until the next time point when one re-solves the pre-committed problem, and repeats these steps until the terminal time. The *eventual* naïve policy is then just to "paste" these piece-wise pre-committed policies together.

This pasting approach, however, becomes problematic in the continuous-time setting. Indeed, assume that at each given time and state, say (s, y), there is a pre-committed policy $(t, x) \mapsto \pi^*(t, x; s, y)$. This policy is executed at (s, y) only and *instantaneously* discarded thereafter, while a policy applied for just one single time–state initial point (s, y) has no impact on the dynamics in continuous time. In other words, there would be no difference if one applied a different policy, say $\pi^{**}(s, y; s, y)$, at just the initial point (s, y). Therefore, while it is seemingly intuitive to define the function $(t, x) \mapsto \pi^*(t, x; t, x)$ as the overall naïve policy, the definition is nothing else than a *heuristic*, exactly because for each fixed (t, x) the action $\pi^*(t, x; t, x)$ does not impact the system and hence it is not clear how to *interpret* the resulting "continuously pasted" policy $(t, x) \mapsto \pi^*(t, x; t, x)$. In the setting of optimal stopping, Huang and Nguyen-Huu (2018) and Huang et al. (2020) define "naïve stopping policies" following this continuous pasting approach, which is, *in retrospect*, still a heuristic for the same reason.

In sum, the notion of naïve policies, which is defined originally and naturally for discrete time, *collapses* in continuous time. In this paper, we address this problem and make two main contributions. First, we take a different approach to define the naïve policies that is faithful to the original spirit of Strotz (1956) but adapted to the continuous-time setting, premised upon the notion that, *in general*, any continuous-time behavior is the limit of discrete-time behaviors when the time-step approaches zero.[5] We fix a set of discrete time points and consider a fictitious agent who only optimizes at each of these points and holds the resulting pre-committed policy until the next point. It is then natural to use the "limit" – in a certain sense – of these

[4]Huang and Nguyen-Huu (2018) and Huang et al. (2020) define naïve stopping policies for continuous-time optimal stopping problems. However, optimal stopping can be regarded as a special optimal stochastic control problem, but not vice versa. Moreover, there is some delicate issue with the definition of naïve stopping policies in Huang and Nguyen-Huu (2018) and Huang et al. (2020), which will be explained in the next paragraph.

[5]An analogy here is that the Brownian motion is just the limit of a simple random walk when the step size diminishes to zero.

discretely naïve agents when the step size becomes asymptotically small to describe the naïve behavior in the original continuous-time model. One technical subtlety here is that policies are generally only measurable functions whose limit is difficult to analyze. We consider instead the limiting process of the wealth processes – which are analytically better behaved – of those discrete agents, and find the policy that generates this limiting process as the wealth process. It is important to note that the resulting policy must be time-consistent to be qualified as our naïve policy, namely it should not depend on any specific time and state the discretely naïve agent starts. There are two advantages in our approach. On one hand, the approach is both general and constructive. It is general because the definition of a naïve policy applies readily to any time-inconsistent problems beyond MV, and it is constructive because the definition itself points to the direction of *deriving* a naïve policy. On the other hand, our approach is advantageous from the implementation perspective (relevant to what is necessary in e.g. practical applications and computer simulations): discrete-time processes are the only implementable ones and policies defined directly in continuous time make implicit use of quantities that are not necessarily accessible in a discrete-time setting, while our method naturally links to the discrete-time approximation for implementation.

The second contribution is to compare the naïve policies with the other Strotzian types of policies in a Black–Scholes market. Be mindful that it does not make much sense to use either mean or variance of the terminal wealth alone for comparison, as the essence of the MV model is to achieve a best trade-off between the two criteria. Instead, MV efficiency ought to be the primary criterion. We show that, between a naïveté and a pre-committer and starting from any given point of time and state, the former when sufficiently risk-seeking ends up with a higher expected terminal wealth than he originally planned but is still MV inefficient (while the latter is always MV efficient by definition). To compare naïve and equilibrium policies which are both MV inefficient, we use an objective metric which is the risky weight defined as the fraction of dollar amount invested in stocks. We show that a naïveté *always* allocate strictly higher risky weight than the two types of consistent planners considered by Björk et al. (2014) and He and Jiang (2022) respectively. This in turn suggests that the naïve policies tend to be more risk-taking than their consistent planning counterparts.[6] We finally define an efficiency ratio for comparing return–risk tradeoff with the *same* original level of risk aversion, and prove that naïve policies are always strictly less efficient than pre-committed and equilibrium policies.

Pendersen and Peskir (2017) introduce the notion of "dynamic optimality" in a continuous-time MV model, which bears some relevance to naïve policies (although the paper stops short of commenting on it). Definition 2 therein defines a dynamically optimal policy as there being no other policy applied at present time could produce a more favourable value at the terminal time. However, as discussed earlier, in a time-inconsistent problem there is no such thing as "dynamic optimality": as much as a naïveté attempts to reoptimize continuously over time, the resulting actual policy at any given time may significantly deviate from the pre-committed optimal one (and therefore is MV inefficient with respect to that given time, and indeed *not* optimal in any sense). On the other hand, Pendersen and Peskir (2017) conjecture the analytical formula

[6] An analogous result is proved in Hu et al. (2022) for a casino gambling model: a naïve gambler stops gambling no earlier than a gambler doing consistent planning.

of such a "dynamically optimal" policy for a single stock Black–Scholes market, which turns out to be exactly the "continuously pasted policy" discussed earlier, taking advantage of the availability of the explicitly expressed pre-committed policies. We have already explained why our definition of naïve policies is general and our derivation of these policies is constructive.

The rest of the paper is organized as follows. In Section 2 we formulate the continuous-time MV portfolio selection model. In Section 3 we introduce the so-called 2^{-n}-committed policies, which are commited only during a small interval of length 2^{-n}, before reoptimization. We consider the limit of the wealth processes under these policies as $n \to \infty$, and define the policy that generates this limiting wealth process as a naïve policy. We then state the main result that expresses naïve policies analytically. In Section 4 we compare naïve policies with other types of policies in a Black–Scholes market. Section 5 concludes the paper. Proofs related to the main result are placed in Appendices.

2 A Continuous-Time Markowitz Model

In this section we review the continuous-time Markowitz MV model. We first introduce notations.

Throughout this paper, $M^\top$ denotes the transpose of any vector or matrix M, while all vectors are *column* vectors unless otherwise specified. A fixed filtered complete probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0})$ is given along with a standard $\{\mathcal{F}_t\}_{t\geq 0}$-adapted, m-dimensional Brownian motion $W(t) \equiv (W^1(t), ..., W^m(t))^\top$. We use f or $f(\cdot)$ to denote the *function* f, and $f(x)$ to denote the *function value* of f at x. Likewise, we use X or $X(\cdot)$ to denote a stochastic process $X = \{X_s, s \geq 0\}$. Given a Hilbert space H and $b > a \geq 0$, we denote by $L^2([a, b]; H)$ the Hilbert space of H-valued, square-integrable functions f on $[a, b]$ endowed with the norm $(\int_a^b ||f(t)||_H^2 dt)^{1/2}$. Moreover, we denote by $L^2_{\mathcal{F}}([a, b]; \mathbb{R}^m)$ the Hilbert space of $\mathbb{R}^m$-valued, square-integrable and $\{\mathcal{F}_t\}_{t\geq 0}$-adapted stochastic processes g endowed with the norm $\left[\mathbb{E}\int_a^b ||g(t)||^2 dt\right]^{1/2}$, where $||\cdot||$ is the L^2 norm in a Euclidean space.

A financial market has $m+1$ assets being traded continuously. One of the assets is a bank account whose price process S_0 is subject to the following equation:

$$dS_0(t) = r(t)S_0(t)dt, \ t \geq 0; \ S_0(0) = s_0 > 0, \tag{1}$$

where the interest rate function $r(\cdot)$ is deterministic. The other m assets are stocks whose price processes $S_i, i = 1, ..., m$, satisfy the following stochastic differential equations (SDEs):

$$dS_i(t) = S_i(t)\left[b_i(t)dt + \sum_{j=1}^m \sigma_{ij}(t)dW^j(t)\right], \ t \geq 0; \ \ S_i(0) = s_i > 0, \tag{2}$$

where $b(\cdot)$ and $\sigma_{ij}(\cdot)$, the appreciation and volatility rates functions respectively, are scalar-valued and deterministic. Set the excess rate of return vector function and the volatility matrix function respectively as

$$B(t) := (b_1(t) - r(t), ..., b_m(t) - r(t))^\top, \ \ \sigma(t) := (\sigma_{ij}(t))_{m\times m}.$$

An agent has total *wealth* $X(t)$ at time $t \in [0, T]$, where T is a given terminal time of the investment horizon. Assuming that the trading of shares takes place in a self-financing fashion and that there are no transaction costs, the process X satisfies the *canonical wealth equation*

$$dX(t) = \Big[r(t)X(t) + B(t)^\top \pi(t)\Big]\,dt + \pi(t)^\top \sigma(t)dW(t),\ t \in [0, T], \tag{3}$$

where each $\pi_i(t), i = 1, 2, ..., m$, denotes the total market value of the agent's wealth in the i-th asset, resulting in a *portfolio* $(\pi_1(t), ..., \pi_m(t))^\top$, at time t. The agent considers portfolio choice at time s when her wealth is y, where $(s, y) \in [0, T)\times\mathbb{R}$ is given. The process $\pi \equiv (\pi_1, ..., \pi_m)^\top = \{\pi(t) : s \le t \le T\}$ is called an *admissible portfolio* (process) for (s, y) if $\pi \in L^2_{\mathcal{F}}([s, T]; \mathbb{R}^m)$ and the wealth equation (3) with initial condition $X(s) = y$ admits a unique strong solution. Denote by $\mathcal{U}_{s,y}$ the set of admissible portfolio processes for (s, y).

We focus on a portfolio *policy* $\boldsymbol{\pi} = \boldsymbol{\pi}(\cdot, \cdot)$ which is a *deterministic* map from $[0, T] \times \mathbb{R}$ to $\mathbb{R}^m$. Such a policy specifies a portfolio $\boldsymbol{\pi}(t, x)$ when time is t and wealth is x.[7] In the classical, time-consistent setting, a policy $\boldsymbol{\pi}(\cdot, \cdot)$ is independent of the initial time–state pair (s, y), meaning that it is implemented no matter when and where one starts. Such policies are called *time-consistent* ones. A time-consistent policy $\boldsymbol{\pi} = \boldsymbol{\pi}(\cdot, \cdot)$ is called admissible if for *any* $(s, y) \in [0, T) \times \mathbb{R}$, the following SDE obtained by substituting $\boldsymbol{\pi}$ into the wealth equation (3)

$$dX(t) = \Big[r(t)X(t) + B(t)^\top \boldsymbol{\pi}(t, X(t))\Big]\,dt + \boldsymbol{\pi}(t, X(t))^\top \sigma(t)dW(t),\ t \in [0, T];\ \ X(s) = y, \tag{4}$$

admits a unique strong solution X and, moreover, the resulting portfolio process $\pi \in \mathcal{U}_{s,y}$ where $\pi(t) := \boldsymbol{\pi}(t, X(t)), t \in [s, T]$. Note that the wealth–portfolio *process* pair (X, π) *depends on* the initial (s, y), and we say (X, π) is *generated* from the policy $\boldsymbol{\pi}$ with respect to (s, y).

The classical verification theorem for time-consistent problems (e.g. Yong and Zhou, 1999) dictates that, under standard assumptions, there exists a time-consistent policy that generates optimal wealth–portfolio process pair (X, π) for any given initial (s, y).

The following assumptions are in force throughout this paper.

(A1) $r(t), B(t)$ and $\sigma(t)$ are uniformly bounded on $[0, T]$.

(A2) $B(t) \neq 0$ a.e.$t \in [0, T]$ and $\sigma(t)\sigma(t)^\top \ge \delta I, \forall t \in [0, T]$ for some $\delta > 0$.

Given $(s, y) \in [0, T) \times \mathbb{R}$, the Markowitz mean–variance portfolio selection problem over $[s, T]$ is

$$\min_{\pi(\cdot)\in\mathcal{U}_{s,y}} \ \mathrm{Var}_{s,y}(X(T)) \tag{5}$$

$$\text{subject to } \begin{cases} \mathbb{E}_{s,y}[X(T)] = yf(s, T), \\ (X(\cdot), \pi(\cdot)) \text{ satisfy (3) with } X(s) = y \end{cases} \tag{6}$$

[7]In control theory, the policy here is also called the *feedback* control law, whereas the portfolio process corresponds to the *open-loop* control.

where $\mathrm{Var}_{s,y}$ and $\mathbb{E}_{s,y}$ denote respectively the variance and expectation conditional on $\mathcal{F}_s$ and $X(s) = y$, and $f(u, v), 0 \leq u \leq v \leq T$, is a given deterministic real-valued function satisfying $f(u, u) = 1, \forall u \in [0, T]$. The number $f(u, v)$ represents the desired growth factor over the time horizon $[u, v]$. It is economically sensible to consider the expected mean target to be dependent of the initial (s, y), which is equivalent to the state-dependend risk aversion considered in Björk et al. (2014). He and Jiang (2022) consider a more general target $L(s, y)$ instead of $yf(s, T)$; see also Section 4 of this paper.

We add an assumption on f throughout this paper:

(A3) $f \in C^1([0, T] \times [0, T])$, $f(u, v) \geq e^{\int_u^v r(t)dt}$, $\forall\, 0 \leq u \leq v \leq T$, and $-\infty < \frac{\partial f}{\partial t}(t, T)|_{t=T} < \infty$.

The second part of this assumption is natural, demanding the target return to be at least as great as the risk-free return.

Given (s, y), the relation between $\mathrm{Var}_{s,y}(X_*(T))$ (or equivalently $\sqrt{\mathrm{Var}_{s,y}(X_*(T))}$) and $\mathbb{E}_{s,y}[X_*(T)]$, where $X_*(T)$ is the optimal terminal wealth of the problem (5) – (6), is called an *efficient frontier* with respect to (s, y), which gives the best risk–return tradeoff for future investment when standing at (s, y).

The problem (5) – (6) has been solved explicitly in literature; see e.g. (Li and Zhou, 2006, Theorem 2.1),[8] with the following *unique* optimal policy (conditional on $\mathcal{F}_s$ and $X(s) = y$)

$$\boldsymbol{\pi}_*(t, x; s, y) = -[\sigma(t)\sigma(t)^\top]^{-1}B(t)^\top \left[x - \gamma(s, T)e^{-\int_t^T r(v)dv}y\right], \quad (t, x) \in [s, T) \times \mathbb{R}, \tag{7}$$

where

$$\gamma(s, T) := \frac{f(s, T) - e^{\int_s^T [r(v)-\rho(v)]dv}}{1 - e^{-\int_s^T \rho(v)dv}}, \quad s \in [0, T), \tag{8}$$

with

$$\rho(t) := B(t)[\sigma(t)\sigma(t)^\top]^{-1}B(t)^\top > 0.$$

Note that l'Hôspital's rule along with Assumptions **(A2)-(A3)** yield that $\gamma(\cdot, T)$ is continuous at T; hence is uniformly bounded on $[0, T]$.

Substituting the policy (7) into the wealth equation (3) we obtain that the corresponding optimal wealth process is determined by the following SDE:

$$\begin{cases} dX_*(t) = \left[(r(t) - \rho(t))X_*(t) + \gamma(s, T)\rho(t)e^{-\int_t^T r(v)dv}y\right] dt \\ \quad -B(t)(\sigma(t)\sigma(t)^\top)^{-1}\sigma(t)\left[X_*(t) - \gamma(s, T)e^{-\int_t^T r(v)dv}y\right] dW(t), \quad t \in [s, T], \\ X(s) = y. \end{cases} \tag{9}$$

Finally, the efficient frontier at (s, y) is

$$\mathrm{Var}_{s,y}(X_*(T)) = \frac{1}{e^{\int_s^T \rho(v)dv} - 1}\left(\mathbb{E}_{s,y}[X_*(T)] - ye^{\int_s^T r(v)dv}\right)^2. \tag{10}$$

[8] The previous results such as (Li and Zhou, 2006, Theorem 2.1) are for the case when $(s, y) = (0, x_0)$, but they extend readily to arbitrary initial (s, y) because the underlying mathematical problem of the latter is the same.

In sharp contrast to the time-consistent setting, the policy $\boldsymbol{\pi}_*(\cdot,\cdot;s,y)$ given by (7) now depends on the initial pair (s,y) *explicitly.* If the agent sticks to this policy during the entire future time period $[s,T]$ without subsequently altering it, then it is the so-called optimal *pre-committed* policy. If the agent is naïve *à la* Strotz who reoptimizes at every subsequent time moment, then the policy (7) will be abandoned immediately (indeed instantaneously) at any $\tilde{s} > s$. More precisely, suppose the agent carries out (7) for a (little) while and reaches the state $X(\tilde{s})$ at time $\tilde{s} > s$. Now the *current* initial time becomes $\tilde{s}$ and the *current* initial state is $X(\tilde{s})$. If the agent reoptimizes the problem for the remaining duration $[\tilde{s},T]$, then the corresponding policy at $(\tilde{s}, X(\tilde{s}))$ is (conditional on $\mathcal{F}_{\tilde{s}}$)

$$\boldsymbol{\pi}_*(t,x;\tilde{s},X(\tilde{s})) = -[\sigma(t)\sigma(t)^\top]^{-1}B(t)^\top \left[x - \gamma(\tilde{s},T)e^{-\int_t^T r(v)dv}X(\tilde{s})\right], \quad (t,x) \in [\tilde{s},T) \times \mathbb{R}. \tag{11}$$

Clearly, the two policies (7) and (11) are generally different as two *functions* on $[\tilde{s},T] \times \mathbb{R}$.

So, problem (5) – (6) admits a policy (7) that is optimal for the current (s,y) *only.* In other words, the pre-committed optimal policy depends inherently on (s,y), which in turn causes the time-inconsistency of the policy and hence that of the problem, as discussed above. A time-inconsistent policy of the type (7) is defined only for the given (s,y).

3 Naïve Policies

A naïvetè ("he") always "reoptimizes" under current information; as a result he devises policies and then instantly abandons them in the continuous-time setting. Although at each given time he tries to follow the pre-committed optimal policy (7) but his *eventual* policy due to the constant changes could be completely different from (7). In this section, we define naïve policies rigorously, and then derive them in analytical form for the MV problem (5)–(6).

3.1 A 2^{-n}-committed agent

As discussed earlier, the difficulty of defining and analyzing naïve policies lies in the continuous-time setting of the problem. We overcome this difficulty by introducing an auxiliary agent, named the 2^{-n}-*committed agent*, to approximate the behavior of the naïvetè.

A 2^{-n}-committed agent ("she") is one who behaves "in between" a pre-committer and a naïvetè. Specifically, she partitions the time horizon $[0,T]$ into 2^n equal-length intervals, with the partitioning points being $\{t_k\}_{k=0}^{2^n}$ where $t_k = \frac{kT}{2^n}$. She first solves problem (5)-(6) with $(s,y) = (0,x_0)$ to obtain the pre-committed optimal policy $\pi(\cdot,\cdot;0,x_0)$ defined by (7). She implements and commits to this policy until time t_1 when her wealth becomes $X(t_1)$, at which she resolves problem (5)-(6) with $(s,y) = (t_1, X(t_1))$ and switches to the policy $\pi(\cdot,\cdot;t_1,X(t_1))$. She commits to this new policy until t_2 before changing it to $\pi(\cdot,\cdot;t_2,X(t_2))$. She then repeats these steps until time T. Figure 1 illustrates the resulting wealth process under this construction.

Denote by $\{X_*(t;t_k) : t \in [t_k,t_{k+1}]\}$ the above wealth process in the time interval $[t_k,t_{k+1}]$, $k = 0,1,\cdots,2^{n-1}$, with $X_*(0;0) = x_0$. By (9), these processes $X_*(t;t_k), t \in [t_k,t_{k+1}], k = 0,1,\cdots,2^{n-1}$,

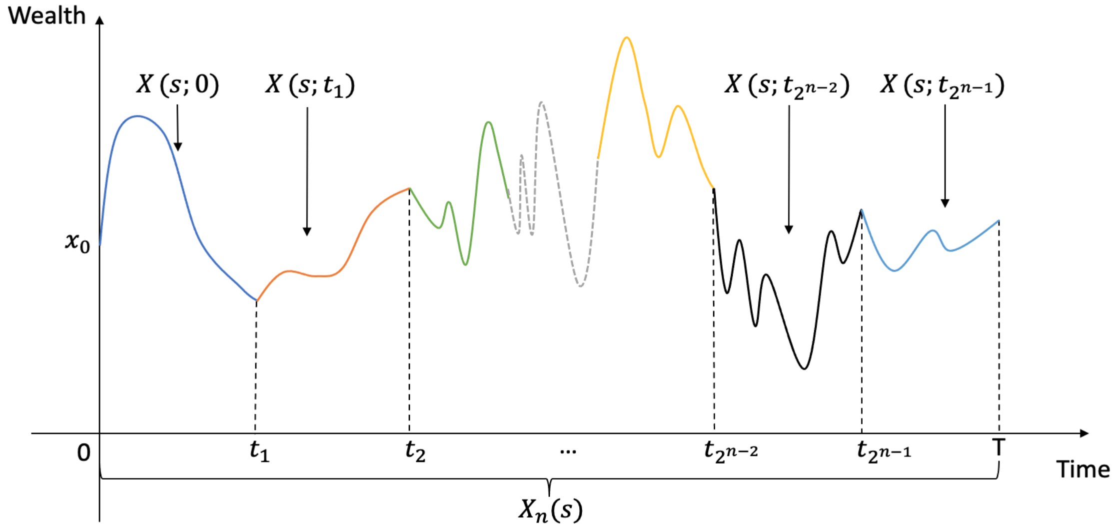

Figure 1: This figure shows a sample path of the wealth process $X_n(\cdot)$ of the 2^{-n}-committer. Each segment of the process, represented by a different color, follows the pre-committed optimal policy devised at the beginning of the corresponding time interval. The wealth process is continuous.

can be determined by the following SDEs recursively:

$$\begin{cases} dX_*(t;t_k) = \left[(r(t)-\rho(t))X_*(t;t_k) + \gamma(t_k,T)\rho(t)e^{-\int_t^T r(v)dv}X_*(t_k;t_{k-1})\right]dt \\ \quad -B(t)(\sigma(t)\sigma(t)^\top)^{-1}\sigma(t)\left[X_*(t;t_k) - \gamma(t_k,T)e^{-\int_t^T r(v)dv}X_*(t_k;t_{k-1})\right]dW(t), \quad t\in[t_k,t_{k+1}], \\ X_*(t_k;t_k) = X_*(t_k;t_{k-1}), \end{cases} \tag{12}$$

where $X_*(t_0;t_{-1})$ is defined as x_0.

Now, by "pasting" $X_*(\cdot;t_k)$, $k=0,1,...,2^n-1$, we obtain the following process:

$$X_n(s) := \begin{cases} X_*(s;0), & 0\le s< t_1, \\ X_*(s;t_1), & t_1\le s<t_2, \\ \dots \\ X_*(s;t_{2^n-1}), & t_{2^n-1}\le s\le T, \end{cases} \tag{13}$$

which is the wealth process of the 2^{-n}-committed agent, visualized by Figure 1. Obviously, this process is adapted and continuous on $[0,T]$.

3.2 Naïve policies

While this 2^{-n}-committed agent behaves somewhere between a pre-committed agent and a naïve one, she is closer to the latter when n becomes larger. Therefore, we define a naïve policy through the limit (in certain sense) of the 2^{-n}-committed wealth process as $n\to\infty$.

Definition 1 If the 2^{-n}-committed wealth process X_n converge to an adapted process X in some sense, and the limiting process X can be generated by a *time-consistent* admissible policy $\boldsymbol{\pi}^* = \boldsymbol{\pi}^*(\cdot,\cdot)$, then $\boldsymbol{\pi}^*$ is called a *naïve policy* of the problem (5)-(6).

Some remarks on this definition are in order. First, this definition applies to more general time-consistent problems instead of just the current Markowitz problem. As such, we intentionally leave vague the precise sense in which X_n converge to X in order to make the definition general and applicable to other problems. For the present problem, we will see momentarily that the convergence is in the weak-L^2 sense. Second, a naïve policy in itself must be time-consistent, meaning that it can no longer depend on any initial (s, y) and, in particular, on $(0, x_0)$, even though each X_n is indeed constructed starting from a specific pair $(0, x_0)$. Third, we do not define a naïve policy as simply the limit of 2^{-n}-committed policies, because policies are in general only measurable and they may not converge and are hard to analyze. Instead, we consider the limit of wealth processes that are much better behaved, and try to derive the equation satisfied by that limit. Then, in the current case of the mean–variance problem, we verify whether the limiting equation is of the canonical form (4). If yes, then we can immediately identify the corresponding naïve policy. If not, then there exists no naïve policy per our definition. For a more general time-consistent problem, we can compare the drift and diffusion coefficients of the limiting equation with those of the system equation respectively to determine the naïve policy or negate its existence. Finally, the choice of the dyadic sequence of partitions (leading to the 2^{-n}-committed policies) is *not* essential. Any other choice of partitions will lead to the same naïve policy – so long as the largest interval converges to 0 – due to the uniqueness of the limit of the corresponding wealth processes.

The following proposition, whose proof is deferred to Appendix A, indicates that the 2^{-n}-committed wealth processes X_n, $n = 1, 2, \cdots$, are uniformly bounded in $L^2_{\mathcal{F}}([0,T];\mathbb{R})$.

Proposition 1 *It holds that*

$$||X_n||^2 := \mathbb{E}\int_0^T |X_n(s)|^2 ds < \infty, \ \forall n.$$

Moreover, $||X_n||^2$ *is uniformly bounded in* n.

Due to Proposition 1, the sequence $\{X_n\}_{n=1}^{\infty}$ is uniformly bounded in the Hilbert space $L^2_{\mathcal{F}}([0,T];\mathbb{R})$, and hence is weakly compact. So there exists a weakly convergent subsequence (still denoted as $\{X_n\}_{n=1}^{\infty}$ without loss of generality) and a process $X \in L^2_{\mathcal{F}}([0,T];\mathbb{R})$ such that

$$X_n \to X \text{ weakly in } L^2_{\mathcal{F}}([0,T];\mathbb{R}).$$

The following theorem is the main result of the paper, which characterizes this limiting process and, consequently, the naïve policy.

Theorem 1 *The weakly limiting process X satisfies the following SDE:*

$$\begin{cases} dX(t) = \left[(r(t)-\rho(t)) + \gamma(t,T)\rho(t)e^{-\int_t^T r(s)ds}\right] X(t)dt \\ \quad -B(t)(\sigma(t)\sigma(t)^\top)^{-1}\sigma(t)\left[1-\gamma(t,T)e^{-\int_t^T r(s)ds}\right] X(t)dW(t), \quad t\in[0,T], \\ X(0)=x_0. \end{cases} \tag{14}$$

Moreover, the following is the naïve policy:

$$\boldsymbol{\pi}^*(t,x) = -[\sigma(t)\sigma(t)^\top]^{-1}B(t)^\top[1-\gamma(t,T)e^{-\int_t^T r(s)ds}]x, \quad (t,x)\in[0,T]\times\mathbb{R}. \tag{15}$$

A proof of Theorem 1 is delayed to Appendices B.

Note that the explicitly presented policy (15) indeed does not depend on any initial pair (s,y) and, in particular, on $(0,x_0)$. This means that even if the wealth process of the 2^{-n}-committed was to be alternatively constructed from a different initial pair (s,y), it would lead to the same naïve policy (15). On the other hand, it generates X as its wealth process for the given initial $(0,x_0)$.

To conclude this section, we illustrate our definition of naïve policies by presenting an example beyond the mean–variance setting.

Example 1 Given $(s,y)\in[0,T)\times\mathbb{R}$, consider the following deterministic control problem

$$\min_{u(\cdot)} \int_s^T |u(t)-2(t-s)|dt$$

$$\text{subject to } \dot{x}(t)=u(t), \quad t\in[s,T]; \quad x(s)=y.$$

This problem is taken from (Zhang, 2023, Section 4) except that there is no state process therein. To derive the naïve policy, consider a 2^{-n}-committed agent who starting from $(s,y)=(0,x_0)$ partitions $[0,T]$ with $t_k=\frac{kT}{2^n}$ where $k=0,1,\cdots,2^n$. Her control is

$$u(t)=2(t-t_k), \quad \text{for } t\in[t_k,t_{k+1}), \quad k=0,1,\cdots,2^n-1.$$

Let the corresponding state process be $x_n(\cdot)$. Then, for $t\in[t_k,t_{k+1})$,

$$\begin{aligned} x_n(t) =& x_n(t_k) + \int_{t_k}^t 2(t'-t_k)dt' = x_n(t_k)+(t-t_k)^2 \\ =& x_n(t_{k-1}) + (t_k-t_{k-1})^2 + (t-t_k)^2 \\ =& \cdots = x_0 + k\left(\frac{T}{2^n}\right)^2 + (t-t_k)^2. \end{aligned}$$

For any $t\in[0,T]$, there is k such that $t\in[t_k,t_{k+1})$. Hence

$$|x_n(t)-x_0| = k\left(\frac{T}{2^n}\right)^2 + (t-t_k)^2 \le 2^n\left(\frac{T}{2^n}\right)^2 + \left(\frac{T}{2^n}\right)^2 \to 0$$

as $n\to\infty$. So $x_n(t)$ converges to $x(t)\equiv x_0$ uniformly in $t\in[0,T]$. The policy that generates the limiting state process $x(t) \equiv x_0$ is then $u(t) = \dot{x}(t) = 0$. It coincides with the "naïve strategy" in (Zhang, 2023, Section 4) which is however based on the heuristic discussed earlier.

4 Comparison between Naïve and Other Types of Policies

In the continuous-time MV literature, two types of equilibrium policies by consistent planners have been introduced and studied: the *weak* equilibrium policies by Björk et al. (2014) and the *regular* equilibrium policies by He and Jiang (2022). In this section, we compare the naïve policies with these two types of equilibrium policies as well as the pre-committed ones, in a Black–Scholes market.

4.1 Weak and regular equilibrium policies

We first review the two types of equilibrium strategies, whose definitions can be found, in slight variants of the MV formulation, in Björk et al. (2014) and He and Jiang (2022) respectively.

Given $(s, y) \in [0, T] \times \mathbb{R}$, Björk et al. (2014) consider the following problem:

$$\max_{\pi(\cdot)\in\mathcal{U}_{s,y}} J(s, y; \pi(\cdot)) := \mathbb{E}_{s,y}[X(T)] - \frac{\alpha(s, y)}{2}\mathrm{Var}_{s,y}(X(T)) \tag{16}$$

$$\text{subject to } (X(\cdot), \pi(\cdot)) \text{ satisfy (3) with } X(s) = y. \tag{17}$$

In the objective function of this problem, there is a risk-aversion term $\alpha(s, y) > 0$ that depends on the initial time s and initial state y; see Björk et al. (2014) for the many discussions on the motivation of such a varying risk-aversion term.[9] The problem is again time-inconsistent. Björk et al. (2014) study the behavior of a consistent planner by considering the equilibrium policies defined as follows. Given an admissible (time-consistent) policy $\hat{\boldsymbol{\pi}}(\cdot, \cdot)$, construct a new policy $\boldsymbol{\pi}_h$ by

$$\boldsymbol{\pi}_h(t, x) := \begin{cases} \pi, & t \in [s, s+h),\ x \in \mathbb{R}, \\ \hat{\boldsymbol{\pi}}(t, x), & t \in [0, T] \setminus [s, s+h),\ x \in \mathbb{R}, \end{cases} \tag{18}$$

where $\pi \in \mathbb{R}^m$, $h > 0$ and $s \in [0, T)$ are aribitrarily given. Let $\pi(\cdot)$ and $\pi_h(\cdot)$ be respectively the portfolio processes generated by $\hat{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}_h$ starting from (s, y). We say that $\hat{\boldsymbol{\pi}}$ is a *weak equilibrium policy* if the pertubed policy $\boldsymbol{\pi}_h$ is admissible and

$$\liminf_{h\to 0} \frac{J(s, y; \hat{\pi}) - J(s, y; \pi_h)}{h} \geq 0, \tag{19}$$

for all $\pi \in \mathbb{R}^m$ and $(s, y) \in [0, T) \times \mathbb{R}$.

On the other hand, He and Jiang (2022) formulate the following problem:

$$\min_{\pi(\cdot)\in\mathcal{U}_{s,y}} \mathrm{Var}_{s,y}(X(T)) \tag{20}$$

[9] Björk et al. (2014) consider only the state-dependent risk aversion $\alpha(s, y) = \alpha(y)$, but the method and results therein readily extend to the time–state dependent case presetned here.

$$\text{subject to} \begin{cases} \mathbb{E}_{s,y}[X(T)] = L(s,y), \\ (X(\cdot), \pi(\cdot)) \text{ satisfy (3) with } X(s) = y, \end{cases} \tag{21}$$

where $L(s,y)$ indicates the expected terminal wealth target when the initial pair is (s,y).[10] When $L(s,y) = yf(s,T)$, the problem (20)–(21) reduces to the problem (5)–(6). He and Jiang (2022) also study a consistent planner, except that they use the notion of *regular* equilibrium policies which is very different from that of the weak equilibrium policies. Specifically, an admissible, time-consistent policy $\hat{\boldsymbol{\pi}}$ is called a *regular equilibrium policy* if for any $(s,y) \in [0,T) \times \mathbb{R}$, any $\pi \in \mathbb{R}^m$ such that $\boldsymbol{\pi}_h$ constructed by (18) is admissible for sufficeintly small $h > 0$, we have[11]

$$\mathrm{Var}_{s,y}(X^{\boldsymbol{\pi}_h}(T)) - \mathrm{Var}_{s,y}(X^{\hat{\boldsymbol{\pi}}}(T)) \geq 0 \tag{22}$$

for sufficiently small $h > 0$, where $X^{\hat{\boldsymbol{\pi}}}(T)$ and $X^{\boldsymbol{\pi}_h}(T)$ are the terminal wealth values, both starting from (s,y) and under $\hat{\boldsymbol{\pi}}$ and $\boldsymbol{\pi}_h$ respectively.

The difference between the problems (16)–(17) and (20)–(21) is that the former uses a weighting coefficient $\alpha(s,y)/2$ in its objective function while the latter takes $L(s,y)$ in its constraint. The two problems are related via the Lagrange multiplier method. As a result, if we choose $\alpha(s,y)$ and $L(s,y)$ in a certain way, then the respective pre-committed optimal polices for the two problems coincide, as stipulated in the following proposition.

Proposition 2 *If*

$$\frac{1}{\alpha(s,y)} e^{\int_s^T \rho(t)dt} + y e^{\int_s^T r(t)dt} = \frac{L(s,y) - e^{\int_s^T [r(t)-\rho(t)]dt} y}{1 - e^{-\int_s^T \rho(t)dt}}, \quad \forall (s,y) \in [0,T] \times \mathbb{R} \tag{23}$$

holds, then the pre-committed optimal policies for (16)–(17) and (20)–(21) are the same for any $(s,y) \in [0,T] \times \mathbb{R}$ in the following form

$$\bar{\boldsymbol{\pi}}_*(t,x;s,y) = -[\sigma(t)\sigma(t)^\top]^{-1} B(t)^\top \left[x - \bar{\gamma}(s,T,y) e^{-\int_t^T r(v)dv} \right], \quad (t,x) \in [s,T] \times \mathbb{R}, \tag{24}$$

where

$$\bar{\gamma}(s,T,y) := \frac{1}{\alpha(s,y)} e^{\int_s^T \rho(v)dv} + e^{\int_s^T r(v)dv} y.$$

Proof It follows from the equations (5.12), (5.1) and (6.7) in Zhou and Li (2000) that the pre-committed optimal policy of (16)–(17) is (24). On the other hand, it follows from (Li and Zhou, 2006, Theorem 2.1) that the precommitted strategy of (20)–(21) is

$$\tilde{\boldsymbol{\pi}}_*(t,x;s,y) = -[\sigma(t)\sigma(t)^\top]^{-1} B(t)^\top \left[x - \tilde{\gamma}(s,T,y) e^{-\int_t^T r(v)dv} \right], \quad (t,x) \in [s,T] \times \mathbb{R}, \tag{25}$$

where

$$\tilde{\gamma}(s,T,y) := \frac{L(s,y) - e^{\int_s^T [r(v)-\rho(v)]dv} y}{1 - e^{-\int_s^T \rho(v)dv}}.$$

[10] In the original formulation of He and Jiang (2022), the expected terminal wealth constraint is $\mathbb{E}_{s,y}[X(T)] \geq L(s,y)$, which is equivalent to the equality constraint formulated here.

[11] Here, the term "admissible" requires the corresponding portfolio processes generated by the relevant policies for (s,y) to also satisfy the expectation constraint in (21).

It is now evident that if (23) is satisfied, then $\bar{\gamma}(s,T,y) \equiv \tilde{\gamma}(s,T,y)$ leading to $\tilde{\boldsymbol{\pi}}_*(t,x;s,y) \equiv \bar{\boldsymbol{\pi}}_*(t,x;s,y)$. ∎

The condition (23) ensures that the pre-committed solutions of the two problems coincide. As a result, the naïve policies of the two problems are also identical because they are obtained via the limit of pre-committed policies. However, (23) does not necessarily lead to the same weak/regular equilibrium policies of the two problems, because equilibrium policies are not based on pre-committed ones.

4.2 Comparisons

We now compare the naïve policies with the weak/regular equilibrium policies and the pre-committed polices, in a Black–Scholes market for simplicity. Specifically, there is a risk-free asset and only one risky asset (i.e. $m = 1$) with $r(t) \equiv r > 0$, $B(t) \equiv b - r > 0$, $\sigma(t) \equiv \sigma > 0$. As a result, $\rho(t) \equiv \rho = (\frac{b-r}{\sigma})^2 > 0$.

We carry out the comparison for two cases. In Subsection 4.2.1, we choose $\alpha(s,y) = \frac{\alpha}{y}$ for some constant $\alpha > 0$ in the problem (16)–(17), which is also a case examined closely in Björk et al. (2014). The parameter α characterizes the risk preference of the agent, and a greater α implies a higher degree of risk aversion. Subsection 4.2.2 studies the case when $L(s,y) = ye^{k(T-s)}$ for some constant $k > r$ in the problem (20)–(21). Clearly, a greater k indicates a smaller level of risk aversion. In each case, we choose $f(\cdot,\cdot)$, $L(\cdot,\cdot)$ and $\alpha(\cdot,\cdot)$ in such a way (e.g. to satisfy (23)) that the different formulations of the MV problem are consistent in their respective pre-committed optimal policies.

4.2.1 The case $\alpha(s,y) = \frac{\alpha}{y}$

When $\alpha(s,y) = \frac{\alpha}{y}$, the corresponding L according to (23) is

$$L(s,y) = y\left[\frac{1}{\alpha}\left(e^{(T-s)\rho} - 1 + \alpha e^{(T-s)r}\right)\right], \tag{26}$$

whereas the corresponding f is

$$f(s,T) = \frac{1}{\alpha}\left[e^{(T-s)\rho} - 1 + \alpha e^{(T-s)r}\right]. \tag{27}$$

It is easy to check that this f satisfies Assumption **(A3)**. By Theorem 1, the naïve policy is

$$\boldsymbol{\pi}^*(t,x) = -\frac{b-r}{\sigma^2}\left[1 - \frac{f(t,T) - e^{(r-\rho)(T-t)}}{1 - e^{-\rho(T-t)}}e^{-r(T-t)}\right]x, \quad (t,x) \in [0,T] \times \mathbb{R}. \tag{28}$$

Substituting the expression of f in (27) into the above and going through some simple computation, we finally get

$$\boldsymbol{\pi}^*(t,x) = \frac{b-r}{\alpha\sigma^2}e^{(\rho-r)(T-t)}x, \quad (t,x) \in [0,T] \times \mathbb{R}. \tag{29}$$

The *risky weight* function of this policy, defined as the ratio between the dollar amount in the stock and the total wealth and denoted by c_{na}, is thereby

$$c_{na}(t) := \frac{\boldsymbol{\pi}^*(t,x)}{x} = \frac{b-r}{\alpha\sigma^2} e^{(\rho-r)(T-t)}, \quad t \in [0,T], \tag{30}$$

which turns out to be a function of t only.

On the other hand, when $\alpha(s,y) = \frac{\alpha}{y}$, Theorem 4.6 in Björk et al. (2014) gives the weak equilibrium policy of the problem (16)–(17) as

$$\boldsymbol{\pi}_{we}(t,x) = c_{we}(t)x, \tag{31}$$

where $c(t) \equiv c_{we}(t)$ is the unique solution to the following integral equation

$$c(t) = \frac{b-r}{\alpha\sigma^2}\left[e^{-\int_t^T [r+(b-r)c(s)+\sigma^2 c(s)^2]ds} + \alpha e^{-\int_t^T \sigma^2 c(s)^2 ds} - \alpha\right]. \tag{32}$$

Similarly, c_{we} is the risky weight function of the weak equilibrium policy.

Finally, we can rewrite (26) as

$$L(s,y) = y e^{\int_s^T [r+\psi(t)]dt} \tag{33}$$

where

$$\psi(t) := \frac{r + (\rho - r)e^{\rho(T-t)}}{\alpha e^{(T-t)r} + e^{\rho(T-t)} - 1}. \tag{34}$$

Applying Theorem 1-i in He and Jiang (2022) and noting that the solution to the problem (2.10) therein is $v^*(t) = \frac{\psi(t)}{b-r}$, we obtain the regular equilibrium policy for (20)–(21) to be

$$\boldsymbol{\pi}_{re}(t,x) = c_{re}(t)x, \tag{35}$$

where

$$c_{re}(t) := \frac{\psi(t)}{b-r} = \frac{1}{b-r}\frac{r + (\rho - r)e^{\rho(T-t)}}{\alpha e^{(T-t)r} + e^{\rho(T-t)} - 1}, \quad t \in [0,T] \tag{36}$$

is the risky weight of this equilibrium policy at $t \in [0,T]$.

The following proposition shows that the naïve policy allocates *strictly* more weight to the risky asset than the two equilibrium policies at *any* time before T.

Proposition 3 *In the Black–Scholes market, if $\alpha(s,y) = \frac{\alpha}{y}$, then we have*

$$c_{we}(t) < c_{na}(t), \quad c_{re}(t) < c_{na}(t), \quad \forall t \in [0,T),$$

for any $\alpha > 0$.

Proof Let us first prove $c(t) \equiv c_{we}(t) < c_{na}(t) \ \ \forall t \in [0,T)$. We have the obvious inequality

$$\rho + (b-r)c(s) + \sigma^2 c(s)^2 > 0, \quad \forall s \in [0,T) \tag{37}$$

because $\Delta := (b-r)^2 - 4\rho\sigma^2 = -3(b-r)^2 < 0$. Recalling that $c(\cdot)$ satisfies (32), we deduce

$$\begin{aligned} c_{we}(t) &= \frac{b-r}{\alpha\sigma^2}\left[e^{-\int_t^T [r+(b-r)c(s)+\sigma^2 c(s)^2]ds} + \alpha e^{-\int_t^T \sigma^2 c(s)^2 ds} - \alpha\right] \\ &\leq \frac{b-r}{\alpha\sigma^2} e^{-\int_t^T [r+(b-r)c(s)+\sigma^2 c(s)^2]ds} \\ &< \frac{b-r}{\alpha\sigma^2} e^{-\int_t^T (r-\rho)ds} \\ &= \frac{b-r}{\alpha\sigma^2} e^{(\rho-r)(T-t)} = c_{na}(t), \ \forall t \in [0,T). \end{aligned} \tag{38}$$

Next, we prove $c_{re}(t) < c_{na}(t) \ \ \forall t \in [0,T)$. Indeed

$$\begin{aligned} c_{re}(t) &= \frac{1}{b-r} \frac{r + (\rho - r)e^{\rho(T-t)}}{\alpha e^{(T-t)r} + e^{\rho(T-t)} - 1} \\ &< \frac{1}{b-r} \frac{\rho e^{\rho(T-t)}}{\alpha e^{(T-t)r}} \\ &= \frac{b-r}{\alpha\sigma^2} e^{(\rho-r)(T-t)} = c_{na}(t), \ \forall t \in [0,T). \end{aligned}$$

The proof is complete.

■

So naïve policies take more risky exposure than the two types of equilibrium policies. It is interesting to compare the naïvetè also with a pre-committer, realizing that the former strives to follow the latter at *every* initial pair (s,y). Take $(s,y) = (0,x_0)$ for example. The pre-committer's expected terminal wealth is

$$\mathbb{E}_{0,x_0}[X_*(T)] = x_0 f(0,T) = x_0 e^{rT} \frac{1}{\alpha}\left[e^{(\rho-r)T} - e^{-rT} + \alpha\right], \tag{39}$$

noting (27). Although the naïvetè's original expected target return was also $x_0 f(0,T)$ at $(0,x_0)$, he changes mind all the time subsequently so his *actual* target return at $(0,x_0)$ can be significantly deviated from the original one. To see this, plugging in the naïve policy (29) to the wealth equation (3) to obtain

$$dX^*(t) = \left[rX^*(t) + \frac{1}{\alpha}\rho e^{(\rho-r)(T-t)}X^*(t)\right]dt + \frac{b-r}{\alpha\sigma}e^{(\rho-r)(T-t)}X^*(t)dW(t), \ t\in[0,T]; \ \ X^*(0) = x_0. \tag{40}$$

Taking the integral form of this SDE and applying expectation on both sides, we get an ODE in terms of $\mathbb{E}_{0,x_0}[X^*(\cdot)]$. Solving this ODE we arrive at

$$\mathbb{E}_{0,x_0}[X^*(T)] = x_0 e^{rT} e^{\frac{1}{\alpha}\frac{\rho}{\rho-r}[e^{(\rho-r)T}-1]}. \tag{41}$$

Recall that $\alpha > 0$ is the risk aversion coefficient, and the smaller α the less risk averse the agent is. Comparing (41) with (39) and noting that $\frac{\rho}{\rho-r}[e^{(\rho-r)T} - 1] > 0$ always holds, the naïvetè's expected terminal wealth is larger than the pre-committer's when α is small, and the former grows exponentially fast while the latter does only linearly in α^{-1} as $\alpha \to 0$. So a naïve policy ends up achieving a much higher expected terminal wealth than a pre-committed one which is also his *originally* planned target.[12] However, this by no means implies that the former is superior to the latter because in an MV model there are two criteria and the variance is as important as the return. Instead, we ought to analyze the risk–return tradeoff of a naïve policy and compare it with the pre-committer's which is known to attain the best such tradeoff.

To do this, we first have the following result.

Proposition 4 *In the Black–Scholes market, the terminal wealth $X(T)$ of a policy $\boldsymbol{\pi}$ where $\boldsymbol{\pi}(t) = c(t)X(t)$ starting from $X_0 = x_0$ satisfies*

$$\sqrt{\mathrm{Var}_{0,x_0}[X(T)]} = \sqrt{e^{\int_0^T \sigma^2 c(t)^2 dt} - 1}\, \mathbb{E}_{0,x_0}[X(T)]. \tag{42}$$

Proof Under $\boldsymbol{\pi}(t) = c(t)X(t)$, the wealth process follows

$$dX(t) = [r + (b-r)c(t)]X(t)dt + \sigma c(t)X(t)dW(t), \quad X(0) = x_0.$$

Hence

$$\mathbb{E}_{0,x_0}[X(T)] = x_0 e^{\int_0^T [r+(b-r)c(t)]dt}. \tag{43}$$

By Itô's formula

$$dX(t)^2 = 2[r + (b-r)c(t) + \frac{1}{2}\sigma^2 c(t)^2]X(t)^2 dt + 2\sigma c(t)X(t)^2 dW(t),$$

leading to

$$\mathbb{E}_{0,x_0}[X(T)^2] = x_0^2 e^{\int_0^T [2r+2(b-r)c(t)+\sigma^2 c(t)^2]dt}. \tag{44}$$

It follows that

$$\begin{aligned} \mathrm{Var}_{0,x_0}[X(T)] &= \mathbb{E}_{0,x_0}[X(T)^2] - (\mathbb{E}_{0,x_0}[X(T)])^2 \\ &= x_0^2 e^{\int_0^T [2r+2(b-r)c(t)+\sigma^2 c(t)^2]dt} - x_0^2 e^{2\int_0^T [r+(b-r)c(t)]dt} \\ &= x_0^2 e^{2\int_0^T [r+(b-r)c(t)]dt}(e^{\int_0^T \sigma^2 c(t)^2 dt} - 1) \\ &= (e^{\int_0^T \sigma^2 c(t)^2 dt} - 1)(\mathbb{E}_{0,x_0}[X(T)])^2; \end{aligned} \tag{45}$$

and hence the desired result. ∎

A naïve agent with $\alpha(s,y) = \alpha/y$ applies the policy $\boldsymbol{\pi}^*(t) = c_{na}(t)X^*(t)$ where c_{na} is given by (30). Applying Proposition 4 we deduce

$$\sqrt{\mathrm{Var}_{0,x_0}[X^*(T)]} = \sqrt{e^{\frac{\rho}{2\alpha^2(\rho-r)}[e^{2(\rho-r)T}-1]} - 1}\, \mathbb{E}_{0,x_0}[X^*(T)], \quad \text{for } \alpha > 0. \tag{46}$$

Note the above is well defined when $\rho = r$ because the limit of the expression exists when $\rho \to r$. For the pre-committed agent, it follows from (10) and (39) that the corresponding relationship is

[12] This also reconciles with the previously proved fact that naïve policies are more exposed to the stock than equilibrium ones.

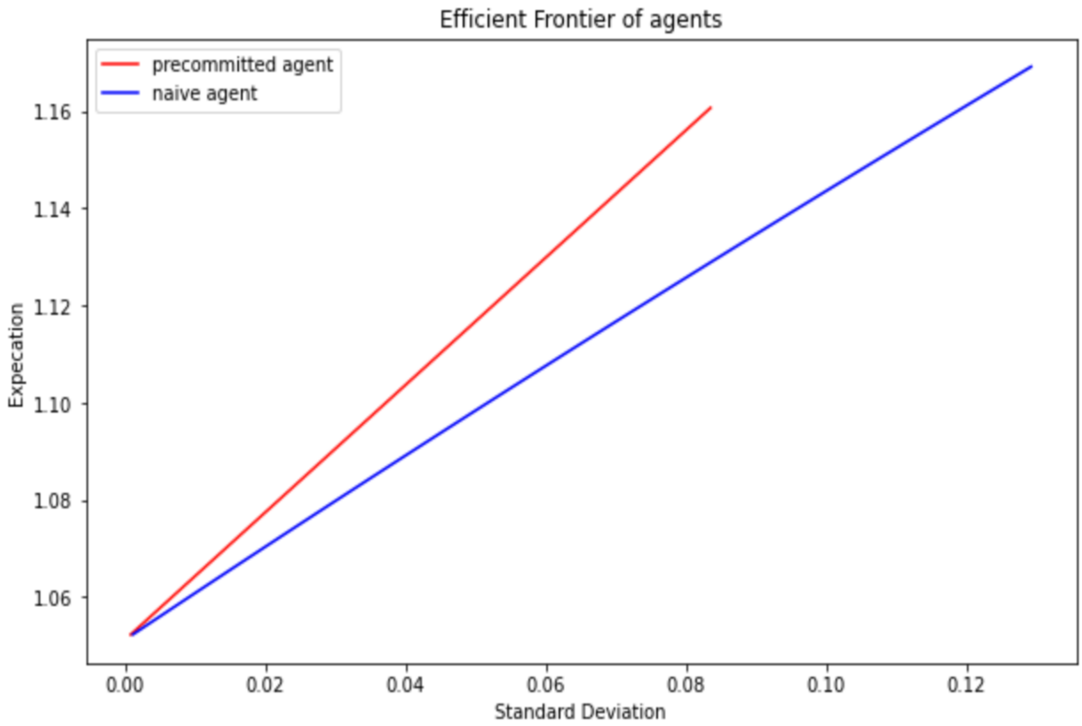

Figure 2: This figure shows the risk–return tradeoffs of pre-committed agents and naïve agents.

$$\sqrt{\mathrm{Var}_{0,x_0}[X_*(T)]} = \sqrt{\frac{1}{e^{\rho T}-1}}\left(\mathbb{E}_{0,x_0}[X_*(T)] - x_0 e^{rT}\right), \quad \text{for } \mathbb{E}_{0,x_0}[X_*(T)] \geq x_0 e^{rT}. \tag{47}$$

Because (47) is mean–variance efficient, the line represented by it must lie *above* that by (46) on the expectation–standard deviation plane; see Figure 2 for a visual demonstration in which $x_0 = 1, r = 0.05, \sigma = 0.1, b = 0.15$ and $T = 1$. Note both lines start at $(0, x_0 \exp(rT)) = (0, \exp(0.1))$, corresponding to $\alpha = \infty$ for the naïvetè and to $\mathbb{E}_{0,x_0}[X_*(T)] = x_0 e^{rT}$ for the pre-committer, and diverges as the expectation grows high. In other words, the naïve policy (28) takes more risk than it needs to - as dictated by the efficient frontier – in order to achieve a higher expected terminal wealth (41).

To sum, in the current MV setting, a naïve policy is more risk-loving than the other types of polices while expecting higher terminal wealth. Although at every (s, y) it tries to follow the pre-committed optimal policy, the actual policy turns out to be very different. It is MV inefficient and certainly not "dynamically optimal" in any sense at any given (s, y).

4.2.2 The case $L(t, x) = xe^{k(T-t)}$

We now consider the case when $L(t, x) = xe^{k(T-t)}$, where $k > r$ (otherwise the problem (20)–(21) is trivial). The corresponding f is $f(t, T) = e^{k(T-t)}$, which satisfies Assumption

(A3). Substituting this into (28) we obtain the naïve policy

$$\boldsymbol{\pi}^*(t,x) = c_{na}(t)x, \quad (t,x) \in [0,T] \times \mathbb{R}$$

where the risky weight is

$$c_{na}(t) = \frac{\boldsymbol{\pi}^*(t,x)}{x} = \frac{b-r}{\sigma^2} \frac{e^{(k-r)(T-t)} - 1}{1 - e^{-\rho(T-t)}}, \quad t \in [0,T]. \tag{48}$$

Next, it follows from (23) that the corresponding

$$\alpha(s,y) = \frac{\phi(s)}{y} \tag{49}$$

where $\phi(s) := \frac{e^{\rho(T-s)}-1}{e^{k(T-s)}-e^{r(T-s)}} > 0$. Again, by Theorem 4.6 in Björk et al. (2014) we get the weak equilibrium policy of the problem (16)–(17) to be

$$\boldsymbol{\pi}_{we}(t,x) = c_{we}(t)x,$$

where $c(t) \equiv c_{we}(t)$ uniquely solves

$$c(t) = \frac{b-r}{\phi(t)\sigma^2} \left[e^{-\int_t^T [r+(b-r)c(s)+\sigma^2 c(s)^2] ds} + \phi(t) e^{-\int_t^T \sigma^2 c(s)^2 ds} - \phi(t) \right]. \tag{50}$$

Finally, by Theorem 1-i in He and Jiang (2022), the regular equilibrium policy for (20)–(21) is

$$\boldsymbol{\pi}_{re}(t,x) = c_{re}(t)x,$$

where

$$c_{re}(t) := \frac{k-r}{b-r}, \quad t \in [0,T]. \tag{51}$$

Proposition 5 *In the Black-Scholes market, if $L(t,x) = xe^{k(T-t)}$, then we have*

$$c_{we}(t) < c_{na}(t), \quad c_{re}(t) < c_{na}(t), \quad \forall t \in [0,T),$$

for any $k > r$.

Proof It follows from (50) that

$$
\begin{aligned}
c_{we}(t) \equiv c(t) &= \frac{b-r}{\phi(t)\sigma^2}\left[e^{-\int_t^T [r+(b-r)c(s)+\sigma^2 c(s)^2]ds} + \phi(t)e^{-\int_t^T \sigma^2 c(s)^2 ds} - \phi(t)\right] \\
&\leq \frac{b-r}{\phi(t)\sigma^2} e^{-\int_t^T [r+(b-r)c(s)+\sigma^2 c(s)^2]ds} \\
&< \frac{b-r}{\phi(t)\sigma^2} e^{-\int_t^T (r-\rho)ds} \\
&= \frac{b-r}{\phi(t)\sigma^2} e^{(\rho-r)(T-t)} \\
&= \frac{b-r}{\sigma^2}\frac{e^{(k-r)(T-t)}-1}{1-e^{-\rho(T-t)}} = c_{na}(t), \ \forall t \in [0,T),
\end{aligned}
\tag{52}
$$

where we have utilized (37) to get the second inquality and noted the definition of $\phi(\cdot)$ to obtain the second to the last equality.

Next, applying the general inequality

$$
\frac{e^x - 1}{1-e^{-y}} > \frac{x}{y}, \quad \forall x > 0, \ y > 0, \tag{53}
$$

we deduce

$$
c_{na}(t) = \frac{b-r}{\sigma^2}\frac{e^{(k-r)(T-t)}-1}{1-e^{-\rho(T-t)}} > \frac{b-r}{\sigma^2}\frac{k-r}{\rho} = \frac{k-r}{b-r} = c_{re}(t).
$$

The proof is complete.of ∎

We can also compare the naïve policy with the pre-committed one with respect to any initial (s, y) in the current case. Because the analysis is similar to that in the previous subsection, we omit the details here.

4.3 Efficiency ratio

The naïve policy is not mean–variance efficient as discussed in Subsection 4.2.1 and illustrated by Figure 2. However, there is some subtlety in this comparison. The expected terminal wealth target of the pre-committer is exogenous and fixed over time, while that of the naïvetè is changing, as shown in Subsection 4.2.1, and hence is *endogenous.* Figure 2 shows that the red line is above the blue one, which means that for the *same* expected terminal wealth, the naïve policy results in larger standard deviation than the pre-committed one. The same terminal expected terminal wealth, however, does not correspond to the same risk aversion level represented by α (or by the compatible f or L) for the naïvetè and the pre-committer, because the former ends up with a much higher expected return than the latter even with the same original α, so long as α is sufficiently small.

Therefore, it will give a more complete picture of comparison if we compare the points on the red and blue lines with the same primitive α, instead of just comparing the points with the same expectation or standard deviation. To this end, we define the *efficiency ratio* $R = \frac{\mathbb{E}_{0,x_0}[X(T)]}{\sqrt{\mathrm{Var}_{0,x_0}[X(T)]}}$, which represents the tangent of any point on the expectation–standard deviation plane. Note R is a function of α. Clearly, the higher R is the better risk–return tradeoff an investment strategy achieves.

Proposition 6 *In the Black–Scholes market, the efficiency ratio of a policy $\boldsymbol{\pi}$ where $\boldsymbol{\pi}(t) = c(t)X(t)$ starting from $X_0 = x_0$ is*

$$R = \frac{1}{\sqrt{e^{\int_0^T \sigma^2 c(t)^2 dt} - 1}}.$$

Proof This is straightforward from Proposition 4. ∎

Let us compare a naïvetè's efficiency ratio, denoted by $R_{na}(\alpha)$, with a pre-committer's, denoted by $R_*(\alpha)$, for the same α in the setting of Subsection 4.2.1. The following result shows that the former is *always* strictly less efficient than the latter.

Proposition 7 *Under the assumption of Proposition 3, we have*

$$R_*(\alpha) > R_{na}(\alpha)$$

for any $\alpha > 0$.

Proof It follows from (10) and (39) that the ratio for a pre-committed policy is

$$\begin{aligned} R_*(\alpha) =& \frac{\mathbb{E}_{0,x_0}[X_*(T)]}{\sqrt{\mathrm{Var}_{0,x_0}(X_*(T))}} = \sqrt{e^{\rho T} - 1}\left(\frac{1}{1 - \frac{x_0 e^{rT}}{\mathbb{E}_{0,x_0}[X_*(T)]}}\right) \\ =& \sqrt{e^{\rho T} - 1}\left(\frac{1}{1 - \frac{\alpha}{e^{(\rho-r)T} - e^{-rT} + \alpha}}\right) = \frac{(e^{\rho T} - 1 + \alpha e^{rT})}{\sqrt{e^{\rho T} - 1}}. \end{aligned} \tag{54}$$

For a naïve policy, it follows from (46) that

$$R_{na}(\alpha) = \frac{1}{\sqrt{e^{\frac{\rho}{2\alpha^2(\rho-r)}[e^{2(\rho-r)T}-1]} - 1}} \leq \frac{1}{\sqrt{\frac{\rho}{2\alpha^2(\rho-r)}[e^{2(\rho-r)T} - 1]}} = \frac{\alpha}{\sqrt{\frac{\rho}{2(\rho-r)}[e^{2(\rho-r)T} - 1]}}. \tag{55}$$

The above results imply that as $\alpha \to 0$, $R_*(\alpha) \to e^{\rho T} - 1 > 0$ while $R_{na}(\alpha) \to 0$. This proves $R_*(\alpha) > R_{na}(\alpha)$ for sufficiently small $\alpha > 0$. The desired result will then follow if we can show that $R_*'(\alpha) \geq R_{na}'(\alpha)$ for any $\alpha > 0$.

To this end, we first prove a general inequality

$$\frac{e^{2x} - e^{2y}}{2(x-y)} \geq \frac{e^{2x} - 1}{2x}, \quad x \neq y, \quad x \neq 0. \tag{56}$$

Indeed, fix $x \neq 0$ and define a function $f(y) := \frac{e^{2x} - e^{2y}}{2(x-y)}$, $y \neq x$. Its first-order derivative is

$$f'(y) = \frac{e^{2x} - e^{2y} - 2xe^{2y} + 2ye^{2y}}{2(x-y)^2} =: \frac{g(y)}{2(x-y)^2}, \quad y \neq x. \tag{57}$$

However $g'(y) = 4(y-x)e^{2y}$. So $g(y)$ is decreasing when $y < x$ and increasing when $y > x$, leading to $\min_y g(y) = g(x) = 0$. Consequently, $f'(y) \geq 0$, $\forall y \neq x$, implying $\frac{e^{2x}-e^{2y}}{2(x-y)} = f(y) \geq f(0) = \frac{e^{2x}-1}{2x}$ $\forall y \neq x$ and $x \neq 0$ and proving (56).

It now follows from (56) that $\frac{e^{2\rho T}-e^{2rT}}{2(\rho-r)T} \geq \frac{e^{2\rho T}-1}{2\rho T}$, leading to

$$\frac{\frac{e^{\rho T}-1}{\rho T}}{\frac{e^{2\rho T}-e^{2rT}}{2(\rho-r)T}} \leq \frac{\frac{e^{\rho T}-1}{\rho T}}{\frac{e^{2\rho T}-1}{2\rho T}} \leq 1$$

where the second inequality is due to the monotonicity of $\frac{e^x-1}{x}$ in $x \geq 0$. The above is equivalent to

$$\sqrt{\frac{\frac{e^{\rho T}-1}{\rho T}}{\frac{e^{2(\rho-r)T}-1}{2(\rho-r)T}}} \leq e^{rT},$$

which is further equivalent to

$$R'_{na}(\alpha) \equiv \frac{1}{\sqrt{\frac{\rho}{2(\rho-r)}[e^{2(\rho-r)T}-1]}} \leq \frac{e^{rT}}{\sqrt{e^{\rho T}-1}} \equiv R'_*(\alpha), \quad \alpha > 0.$$

The proof is complete.

∎

Finally, let us compare the efficiency of the naïve policy with those of the weak and regular equilibrium policies, denoted by $R_{we}(\alpha)$ and $R_{re}(\alpha)$ respectively.

Proposition 8 *Under the assumptions of either Proposition 3 or Proposition 5, we have*

$$R_{we}(\alpha) > R_{na}(\alpha), \;\; R_{re}(\alpha) > R_{na}(\alpha)$$

for any $\alpha > 0$.

Proof It is immediate from Propositions 6, 3 and 5. ∎

So a naïve policy always has a smaller efficiency ratio than those of the corresponding two equilibrium policies. Once again, the comparison here is under the *same* level of the original risk aversion reflected by the compatible f, α or L.

5 Conclusions

In this paper we define precisely and derive rigorously the policies implemented by a naïve agent, a notion originally put forth by Strotz (1956), for a continuous-time Markowitz model that is intrinsically time inconsistent. Such an agent attempts to optimize at any given time but, since optimal policies depend on when and where one makes them in a time-inconsistent problem, in effect constantly changes his policies. Ironically, the policy a naïveté actually executes may be anything but he originally desired. At any given time and state he sets an expected investment target and wants to achieve mean–variance efficiency but we show that his final policy ends up with a (much) higher target return and an even higher variance that overall becomes mean–variance *in*efficient if he is sufficiently risk loving. Moreover, naïve policies are universally riskier than their consistent planning counterparts.

Studying naïve behaviors in continuous-time problems is a nearly uncharted research area where open questions abound. An outstanding problem is to derive sufficient conditions for a naïve policy, analogous to the verification theorems for (time-consistent) optimality. The essential difficulty of the problem is that a naïve policy ends up not optimal in any sense as we have emphasized in the paper; hence solving the problem likely calls for a very different approach compared with the classical verification theorems. From a behavioral economics perspective, on the other hand, it is fascinating to inquire and understand how an originally well-intended policy may go wrong or even go opposite when one insists on optimizing *all the time*. The definition of naïve policies and the approach to derive them in this paper are generalizable to other types of problems such as those with non-exponential discounting and probability weighting. As such, we hope the paper has also set a stage for further study of these problems.

Acknowledgements. Zhou is supported by a start-up grant and the Nie Center for Intelligent Asset Management at Columbia University. The authors thank the Editor-in-Chief and two anonymous referees for their constructive comments that have led to an improved version of the paper.

Appendices

A Proof of Proposition 1

The main idea of the proof is to find a *deterministic* function Y to bound X_n^2, which is stated in the following lemma.

Lemma 9 *Let Y satisfying the following ODE*

$$dY(s) = \left[R^* + (\gamma^*)^2 e^{-2\int_s^T r(v)dv}\rho(s)\right] Y(s)ds,\ s \in [0,T];\ \ Y(0) = x_0, \tag{58}$$

where

$$R^* := \max_{0\le s\le T} |2r(s) - \rho(s)|,\ \gamma^* := \max_{0\le s\le T} \gamma(s,T).$$

Then, we have, for every $k = 0, 1, ..., 2^n - 1$,

$$\mathbb{E}[X_*(s;t_k)^2] \le Y(s),\ \ s \in [t_k, t_{k+1}].$$

Proof By Assumptions **(A1)–(A3)**, it is clear that $R^* < \infty$ and $\gamma^* < \infty$.

Recall $X_*(\cdot;t_k)$ satisfies the SDE (12) on $[t_k, t_{k+1}]$ for $k = 0, 1, ..., 2^n - 1$. Applying Itô's formula to $X_*(t;t_k)^2$ and then taking conditional expectation on $\mathcal{F}_{t_k}$ we obtain the (ω-wise) ODE

$$\begin{cases} d\mathbb{E}[X_*(t;t_k)^2|\mathcal{F}_{t_k}] = \left\{(2r(t) - \rho(t))\mathbb{E}[X_*(t;t_k)^2|\mathcal{F}_{t_k}] + \gamma(t_k,T)^2\rho(t)e^{-2\int_t^T r(v)dv}X_*(t_k;t_k)^2\right\}dt,\ \ t \in [t_k, t_{k+1}], \\ \mathbb{E}[X_*(t_k;t_k)^2|\mathcal{F}_{t_k}] = X_*(t_k;t_k)^2. \end{cases} \tag{59}$$

Consider a new stochastic process $Z(\cdot;t_k)$ which satisfies the ODE on $[t_k, t_{k+1}]$ for $k = 0, 1, ..., 2^n - 1$:

$$\begin{cases} dZ(t;t_k) = \left[R^* Z(t;t_k) + \gamma(t_k,T)^2\rho(t)e^{-2\int_t^T r(v)dv}X_*(t_k;t_k)^2\right]dt,\ \ t \in [t_k, t_{k+1}], \\ Z(t_k;t_k) = X_*(t_k;t_k)^2. \end{cases} \tag{60}$$

Because $|2r(t) - \rho(t)| \le R^*$, $t \in [0,T]$, a comparison theorem of ODEs yields

$$\mathbb{E}[X_*(t;t_k)^2|\mathcal{F}_{t_k}] \le Z(t;t_k),\ \text{a.s.},\ \ k = 0, 1, ..., 2^n - 1. \tag{61}$$

Now, we construct another stochastic process $\bar{Z}(\cdot;t_k)$ on $[t_k, t_{k+1}]$ for $k = 0, 1, ..., 2^n - 1$:

$$\begin{cases} d\bar{Z}(t;t_k) = \left[R^* + (\gamma^*)^2\rho(t)e^{-2\int_t^T r(v)dv}\right]\bar{Z}(t;t_k)dt,\ \ t \in [t_k, t_{k+1}], \\ \bar{Z}(t_k;t_k) = X_*(t_k;t_k)^2. \end{cases} \tag{62}$$

It follows from (60) that $Z(t;t_k)$ increases in $t \in [t_k, t_{k+1}]$; hence $Z(t;t_k) \ge X_*(t_k;t_k)^2$ for

$t \in [t_k, t_{k+1}]$. Then, we get

$$\begin{aligned}\frac{dZ(t;t_k)}{dt} &= R^* Z(t;t_k) + \gamma(t_k,T)^2 \rho(t) e^{-2\int_t^T r(v)dv} X_*(t_k;t_k)^2 \\ &\le \left[R^* + \gamma(t_k,T)^2 \rho(t) e^{-2\int_t^T r(v)dv}\right] Z(t;t_k) \\ &\le \left[R^* + (\gamma^*)^2 \rho(t) e^{-2\int_t^T r(v)dv}\right] Z(t;t_k).\end{aligned} \tag{63}$$

Comparing (62) and (63), we conclude from the Grownwall inequality that

$$Z(t;t_k) \le \bar{Z}(t;t_k), \text{ a.s.}, \ t \in [t_k, t_{k+1}], \ k = 0, 1, ..., 2^n - 1. \tag{64}$$

To finish the proof we use mathematical induction on k. When $k = 0$, $t \in [0, t_1]$, it follows from (61) and (64) that

$$\mathbb{E}[X_*(t;0)^2] = \mathbb{E}[\mathbb{E}[X_*(t;0)^2|\mathcal{F}_0]] \le \mathbb{E}[Z(t;0)] \le \mathbb{E}[\bar{Z}(t;0)] = Y(t). \tag{65}$$

Now, assume that when $k = m - 1$, the following holds:

$$\mathbb{E}[X_*(t;t_{m-1})^2] \le Y(t), \quad t \in [t_{m-1}, t_m]. \tag{66}$$

By (61) and (64) we obtain

$$\begin{aligned}\mathbb{E}[X_*(t;t_m)^2] &= \mathbb{E}[\mathbb{E}[X_*(t;t_m)^2|\mathcal{F}_{t_m}]] \\ &\le \mathbb{E}[Z(t;t_m)] \\ &\le \mathbb{E}[\bar{Z}(t;t_m)], \quad t \in [t_m, t_{m+1}]\end{aligned} \tag{67}$$

where the initial value of $\mathbb{E}[\bar{Z}(\cdot;t_m)]$ on $[t_m, t_{m+1}]$ is $\mathbb{E}[X_*(t_m;t_m)^2] \equiv \mathbb{E}[X_*(t_m;t_{m-1})^2]$. However, (66) gives $\mathbb{E}[X_*(t_m;t_{m-1})^2] \le Y(t_m)$, whereas $\mathbb{E}[\bar{Z}(\cdot;t_m)]$ and $Y(\cdot)$ satisfy the same ODE on $[t_m, t_{m+1}]$. Thus $\mathbb{E}[\bar{Z}(t;t_m)] \le Y(t)$ on $[t_m, t_{m+1}]$. Combining with (67), we get the desired result. ∎

We are now ready to prove Proposition 1. By Lemma 9, we have

$$\begin{aligned}||X_n||^2 = \mathbb{E}\int_0^T X_n(s)^2 ds &= \sum_{k=1}^{2^n} \int_{t_{k-1}}^{t_k} \mathbb{E}[X_*(s;t_{k-1})^2] ds \\ &\le \sum_{k=1}^{2^n} \int_{t_{k-1}}^{t_k} Y(s) ds = \int_0^T Y(s) ds < \infty.\end{aligned} \tag{68}$$

B Proof of Theorem 1

To ease notation we use the following

$$\begin{cases}\gamma(t) := \gamma(t,T), \ A(t) := r(t) - \rho(t), \ C(t) := e^{-\int_t^T r(v)dv} \rho(t), \\ D(t) := B(t)(\sigma(t)\sigma(t)^\top)^{-1}\sigma(t) e^{-\int_t^T r(v)dv}, \ F(t) := B(t)(\sigma(t)\sigma(t)^\top)^{-1}\sigma(t),\end{cases} \tag{69}$$

with which we rewrite the SDE (12) as

$$\begin{cases} dX_*(t;t_k) = [A(t)X_*(t;t_k) + \gamma(t_k)C(t)X_*(t_k;t_{k-1})]\,dt \\ \quad + [-F(t)X_*(t;t_k) + \gamma(t_k)D(t)X_*(t_k;t_{k-1})]\,dW(t), \quad t \in [t_k, t_{k+1}], \\ X_*(t_k;t_k) = X_*(t_k;t_{k-1}). \end{cases} \tag{70}$$

Denote

$$A^* := \max_{t\in[0,T]} |A(t)|^2,\ C^* := \max_{t\in[0,T]} |C(t)|^2,\ D^* := \max_{t\in[0,T]} ||D(t)||^2,\ F^* := \max_{t\in[0,T]} ||F(t)||^2,$$

which are all finite due to the boundedness assumptions in **(A1)** and **(A2)**.

In order to prove Theorem 1, we need the following lemma.

Lemma 10 *The process X_n defined by (13) satisfies*

$$\lim_{n\to\infty} \max_{k\in\{0,\ldots,2^n-1\},s\in[t_k,t_{k+1}]} \mathbb{E}|X_n(s) - X_n(t_k)|^2 = 0.$$

Proof For $s \in [t_k, t_{k+1}]$, we bound the term $\mathbb{E}|X_n(s) - X_n(t_k)|^2$ as follows:

$$\begin{aligned} &\mathbb{E}|X_n(s) - X_n(t_k)|^2 = \mathbb{E}|X_*(s;t_k) - X_*(t_k, t_{k-1})|^2 \\ &\leq 2\mathbb{E}\left[\int_{t_k}^s (A(t)X_*(t;t_k) + \gamma(t_k)C(t)X_*(t_k;t_{k-1}))\,dt\right]^2 \\ &+ 2\mathbb{E}\left[\int_{t_k}^s (-F(t)X_*(t;t_k) + \gamma(t_k)D(t)X_*(t_k;t_{k-1}))\,dW(t)\right]^2. \end{aligned} \tag{71}$$

For bounding the first term on the right side of the above, we have by the Cauchy–Schwartz inequality

$$\begin{aligned} &\mathbb{E}\left[\int_{t_k}^s (A(t)X_*(t;t_k) + \gamma(t_k)C(t)X_*(t_k;t_{k-1}))\,dt\right]^2 \\ &\leq (s-t_k)\int_{t_k}^s \mathbb{E}\left|A(t)X_*(t;t_k) + \gamma(t_k)C(t)X_*(t_k;t_{k-1})\right|^2 dt \\ &\leq (s-t_k)\int_{t_k}^s 2\mathbb{E}|A(t)X_*(t;t_k)|^2 + 2\mathbb{E}|\gamma(t_k)C(t)X_*(t_k;t_{k-1})|^2 dt \\ &\leq (s-t_k)\int_{t_k}^s \left(2A^*\mathbb{E}|X_*(t;t_k)|^2 + 2\gamma^*C^*\mathbb{E}|X_*(t_k;t_{k-1})|^2\right) dt \\ &\leq (s-t_k)\int_{t_k}^s (2A^* + 2\gamma^*C^*)Y(T)dt = (2A^* + 2\gamma^*C^*)(s-t_k)^2 Y(T), \end{aligned} \tag{72}$$

where the last inequality follows from Lemma 9 and the fact that $Y(s)$ is increasing in $s \in [0,T]$.

For the second term, by virtue of Itô's isometry, we similarly have

$$\mathbb{E}\left[\int_{t_k}^s (-F(t)X_*(t;t_k) + \gamma(t_k)D(t)X_*(t_k;t_{k-1}))\,dW(t)\right]^2 \leq (2\gamma^*D^* + 2F^*)(s-t_k)Y(T). \tag{73}$$

Combining the above, we obtain

$$\mathbb{E}|X_n(s) - X_n(t_k)|^2 \leq 4(s - t_k)(A^* + \gamma^* C^* + \gamma^* D^* + F^*)Y(T),\ s \in [t_k, t_{k+1}]. \tag{74}$$

Thus,

$$\max_{k\in\{0,\ldots,2^n-1\}, s\in[t_k,t_{k+1}]} \mathbb{E}[X_n(s) - X_n(t_k)]^2 \leq \frac{4T}{2^n}(A^* + \gamma^* C^* + \gamma^* D^* + F^*)Y(T) \to 0$$

as $n \to \infty$. ∎

Because $X_n \to X$ weakly in $L^2_{\mathcal{F}}([0,T];\mathbb{R})$, it follows from Mazur's lemma that for each integer $n \geq 1$, there exists a positive integer $N(n)$ and a convex combination $V_n := \sum_{k=n}^{N(n)} a_{k,n} X_k$, where $a_{k,n} \geq 0$ and $\sum_{k=n}^{N(n)} a_{k,n} = 1$, such that

$$V_n \to X \text{ strongly in } L^2_{\mathcal{F}}([0,T];\mathbb{R}). \tag{75}$$

By the definition of V_n, it satisfies the SDE

$$\begin{cases} dV_n(t) = [A(t)V_n(t) + C(t)U_n(t)]dt + [-F(t)V_n(t) + D(t)U_n(t)]dW(t), \\ V_n(0) = x_0, \end{cases} \tag{76}$$

where

$$U_n(t) := \sum_{k=n}^{N(n)} a_{k,n}[\gamma(m_{t,k})X_k(m_{t,k})],\ m_{t,k} := \frac{N}{2^k}T \text{ when } \frac{N}{2^k}T \leq t < \frac{N+1}{2^k}T \text{ for some } N \in \mathbb{N}.$$

Consider the linear SDE

$$\begin{cases} dZ(t) = [A(t)X(t) + C(t)\gamma(t)X(t)]dt + [-F(t)X(t) + D(t)\gamma(t)X(t)]dW(t), \\ Z(0) = x_0. \end{cases} \tag{77}$$

We now prove that

$$\lim_{n\to\infty} \int_0^T \mathbb{E}|V_n(t) - Z(t)|^2 dt = 0. \tag{78}$$

To this end, we first analyze $V_n(t) - Z(t)$. We have

$$\begin{aligned} V_n(t) - Z(t) &= \int_0^t [A(u)(V_n(u) - X(u)) + C(u)(U_n(u) - \gamma(u)X(u))]\, du \\ &+ \int_0^t [-F(u)(V_n(u) - X(u)) + D(u)(U_n(u) - \gamma(u)X(u))]\, dW(u) \\ &=: Q_{1,n}(t) + Q_{2,n}(t). \end{aligned} \tag{79}$$

As a result,

$$\int_0^T \mathbb{E}|V_n(t) - Z(t)|^2 dt \leq 2\int_0^T \mathbb{E}|Q_{1,n}(t)|^2 dt + 2\int_0^T \mathbb{E}|Q_{2,n}(t)|^2 dt. \tag{80}$$

We proceed to analyze $\mathbb{E}[Q_{1,n}^2(t)]$ and $\mathbb{E}[Q_{2,n}^2(t)]$, respectively. First

$$\begin{aligned}\mathbb{E}|Q_{1,n}(t)|^2 &\leq T\mathbb{E}\int_0^t |A(u)(V_n(u)-X(u))+C(u)(U_n(u)-\gamma(u)X(u))|^2\,du \\ &\leq 2TA^*\int_0^t \mathbb{E}|V_n(u)-X(u)|^2du+2TC^*\int_0^t \mathbb{E}|U_n(u)-\gamma(u)X(u)|^2du.\end{aligned}\tag{81}$$

By the strong convergence of V_n to X, the first term above converges to 0 as $n\to\infty$. For the second term,

$$\begin{aligned}&\int_0^t \mathbb{E}|U_n(u)-\gamma(u)X(u)|^2du \\ &=\int_0^t \mathbb{E}|U_n(u)-\gamma(u)V_n(u)+\gamma(u)V_n(u)-\gamma(u)X(u)|^2du \\ &\leq 2(\gamma^*)^2\int_0^t \mathbb{E}|V_n(u)-X(u)|^2du+2\int_0^t\mathbb{E}\left|\sum_{k=n}^{N(n)}a_{k,n}\left[\gamma(u)X_k(u)-\gamma(m_{u,k})X_k(m_{u,k})\right]\right|^2du.\end{aligned}\tag{82}$$

Now,

$$\begin{aligned}&\int_0^t\mathbb{E}\left|\sum_{k=n}^{N(n)}a_{k,n}\left[\gamma(u)X_k(u)-\gamma(m_{u,k})X_k(m_{u,k})\right]\right|^2du \\ &=\int_0^t\mathbb{E}|\sum_{k=n}^{N(n)}a_{k,n}(\gamma(u)-\gamma(m_{u,k}))X_k(u)+a_{k,n}\gamma(m_{u,k})(X_k(u)-X_k(m_{u,k}))|^2du \\ &\leq 2\int_0^t\left[\mathbb{E}|\sum_{k=n}^{N(n)}a_{k,n}(\gamma(u)-\gamma(m_{u,k}))X_k(u)|^2+\mathbb{E}|\sum_{k=n}^{N(n)}a_{k,n}\gamma(m_{u,k})(X_k(u)-X_k(m_{u,k}))|^2\right]du \\ &\leq 2\int_0^t\left[\sum_{k=n}^{N(n)}a_{k,n}\mathbb{E}|(\gamma(u)-\gamma(m_{u,k}))X_k(u)|^2+\sum_{k=n}^{N(n)}a_{k,n}\mathbb{E}|\gamma(m_{u,k})(X_k(u)-X_k(m_{u,k}))|^2\right]du,\end{aligned}\tag{83}$$

where the last inequality follows from the convexity of the function $f(x)=x^2$. Because $\gamma(\cdot)$ is continuous on $[0,T]$, it is uniformly continuous. Hence, for any $\varepsilon>0$ there is $n_0\in\mathbb{N}$ such that $|\gamma(t)-\gamma(s)|\leq\varepsilon$ whenever $t,s\in[0,T]$ with $|t-s|\leq\frac{1}{2^{n_0}}T$. For $n\geq n_0$, we then have

$$\begin{aligned}&2\int_0^t\left[\sum_{k=n}^{N(n)}a_{k,n}\mathbb{E}|(\gamma(u)-\gamma(m_{u,k}))X_k(u)|^2+\sum_{k=n}^{N(n)}a_{k,n}\mathbb{E}|\gamma(m_{u,k})(X_k(u)-X_k(m_{u,k}))|^2\right]du \\ &\leq 2\int_0^t\left[\varepsilon^2\max_{n\leq k\leq N(n)}\mathbb{E}|X_k(u)|^2+(\gamma^*)^2\max_{n\leq k\leq N(n)}\mathbb{E}|X_k(u)-X_k(m_{u,k})|^2\right]du \\ &\leq 2\int_0^t\left[\varepsilon^2Y(u)+(\gamma^*)^2\frac{4T}{2^n}(A^*+\gamma^*C^*+\gamma^*D^*+F^*)Y(T)\right]du \\ &\leq 2\left[\varepsilon^2+(\gamma^*)^2\frac{4T}{2^n}(A^*+\gamma^*C^*+\gamma^*D^*+F^*)\right]TY(T),\end{aligned}\tag{84}$$

where the second inequality is by Lemma 9 and the proof of Lemma 10. Taking $n\to\infty$ and then $\varepsilon \to 0$, we obtain

$$\lim_{n\to\infty} \int_0^t \mathbb{E}| \sum_{k=n}^{N(n)} a_k^n \left(\gamma(u)X_k(u) - \gamma(m_{u,k})X_k(m_{u,k})\right) |^2 du = 0. \tag{85}$$

Combining (81), (82) and (85) yields

$$\lim_{n\to\infty} \mathbb{E}|Q_{1,n}(t)|^2 = 0. \tag{86}$$

Moreover, according to the above analysis the bound of $\mathbb{E}|Q_{1,n}(t)|^2$ does not depend on t; thus the dominated convergence theorem gives

$$\lim_{n\to\infty} \int_0^T \mathbb{E}|Q_{1,n}(t)|^2 dt = \int_0^T \lim_{n\to\infty} \mathbb{E}|Q_{1,n}(t)|^2 dt = 0. \tag{87}$$

Employing Itô's isometry we can derive similarly

$$\lim_{n\to\infty} \int_0^T \mathbb{E}|Q_{2,n}(t)|^2 dt = \int_0^T \lim_{n\to\infty} \mathbb{E}|Q_{2,n}(t)|^2 dt = 0. \tag{88}$$

By plugging (87) and (88) into (80) we establish (78), namely, $V_n \to Z$ strongly in $L^2_{\mathcal{F}}([0,T];\mathbb{R})$. Thus, $Z(t,\omega) = X(t,\omega)$ except on a zero measure set in the space of $[0,T] \times \Omega$. It follows that X satisfies the same SDE as Z or, equivalently, X satisfies (14). Moreover, it is immediate that this wealth equation is generated by the feedback policy (15). The proof is complete.

References

Barberis, N. A model of casino gambling. *Management Science*, 58:35–51, 2012.

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

Bielecki, T., Jin, H., Pliska, S., and Zhou, X. Continuous-time mean-variance portfolio selection with bankruptcy prohibition. *Mathematical Finance*, 15:213–244, 2005.

Björk, T. and Murgoci, A. A theory of markovian time-inconsistent stochastic control in discrete time. *Finance and Stochastics*, 18:545592, 2014.

Björk, T., Murgoci, A., and Zhou, X. Mean-variance portfolio optimization with state-dependent risk aversion. *Mathematics Finance*, 24:1–24, 2014.

Hakansson, N. Multi-period mean-variance analysis: Toward a general theory of portfolio choice. *The Journal of Finance*, 26:857–884, 1971.

He, X. and Jiang, Z. Mean–variance portfolio selection with dynamic targets for expected terminal wealth. *Mathematics of Operations Research*, pages 587–615, 2022.

He, X. and Zhou, X. Who are I: Time inconsistency and intrapersonal conflict and reconciliation. *Stochastic Analysis, Filtering, and Stochastic Optimization, Yin, G., Zariphopoulou, T. (eds)*, pages 177–208, 2022.

Hu, S., Obloj, J., and Zhou, X. A casino gambling model under cumulative prospect theory: Analysis and algorithm. *Management Science*, page to appear, 2022.

Hu, Y., Jin, H., and Zhou, X. Time-inconsistent stochastic linear-quadratic control. *SIAM Journal on Control and Optimization*, 50:1548–1572, 2012.

Huang, Y.-J. and Nguyen-Huu, A. Time-consistent stopping under decreasing impatience. *Finance and Stochastics*, 22:69–95, 2018.

Huang, Y.-J., Nguyen-Huu, A., and Zhou, X. General stopping behaviors of naïve and non-committed sophisticated agents, with application to probability distortion. *Mathematical Finance*, 30:310–340, 2020.

Li, D. and Ng, W. Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical Finance,*, 10:387–406, 2000.

Li, X. and Zhou, X. Continuous-time mean–variance efficiency: The 80% rule. *Annals of Applied Probability*, 16:1751–1763, 2006.

Lim, A. E. B. and Zhou, X. Quadratic hedging and mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operations Research*, 27:101–120, 2002.

Markowitz, H. Portfolio selection. *Journal of Finance*, 7:77–91, 1952.

Markowitz, H. *Portfolio Selection: Efficient Diversification of Investment.* Wiley, New York, 1959.

Pendersen, J. L. and Peskir, G. Optimal mean-variance portfolio selection. *Mathematics and Financial Economics*, 11:137–160, 2017.

Richardson, H. R. A minimum variance result in continuous trading portfolio optimization. *Management Science*, 35:1045–1055, 1989.

Strotz, R. H. Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23:165–180, 1956.

Xia, J. Mean-variance portfolio choice: Quadratic partial hedging. *Mathematical Finance*, 15: 533–538, 2005.

Yong, J. and Zhou, X. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer, New York, 1999.

Zhang, J. Is a sophisticated agent always a wise one? *SIAM Journal on Financial Mathematics*, 14:42–48, 2023.

Zhou, X. and Li, D. Continous-time mean-variance portfolio selection:a stochastic lq framework. *Applied Mathematics*, 42:19–33, 2000.