

Neural Network Approach to Portfolio Optimization with Leverage Constraints: a Case Study on High Inflation Investment

Chendi Ni¹, Yuying Li², and Peter Forsyth³

¹Cheriton School of Computer Science, University of Waterloo, Waterloo, N2L 3G1,
Canada, chendi.ni@uwaterloo.ca

²Cheriton School of Computer Science, University of Waterloo, Waterloo, N2L 3G1,
Canada, yuying@uwaterloo.ca

³Cheriton School of Computer Science, University of Waterloo, Waterloo, N2L 3G1,
Canada, paforsyt@uwaterloo.ca

May 26, 2023

Abstract

Motivated by the current global high inflation scenario, we aim to discover a dynamic multi-period allocation strategy to optimally outperform a passive benchmark while adhering to a bounded leverage limit. To this end, we formulate an optimal control problem to outperform a benchmark portfolio throughout the investment horizon. Assuming the asset prices follow the jump-diffusion model during high inflation periods, we first establish a closed-form solution for the optimal strategy that outperforms a passive strategy under the cumulative quadratic tracking difference (CD) objective, assuming continuous trading and no bankruptcy. To obtain strategies under the bounded leverage constraint among other realistic constraints, we then propose a novel leverage-feasible neural network (LFNN) to represent control, which converts the original constrained optimization problem into an unconstrained optimization problem that is computationally feasible with standard optimization methods. We establish mathematically that the LFNN approximation can yield a solution that is arbitrarily close to the solution of the original optimal control problem with bounded leverage. We further apply the LFNN approach to a four-asset investment scenario with bootstrap resampled asset returns from the filtered high inflation regime data. The LFNN strategy is shown to consistently outperform the passive benchmark strategy by about 200 bps (median annualized return), with a greater than 90% probability of outperforming the benchmark at the end of the investment horizon.

Keywords: cumulative tracking difference, leveraged portfolio, benchmark outperformance, asset allocation, machine learning

JEL codes: G11, G22

AMS codes: 91G, 35Q93, 68T07

1 Introduction

Since the global outbreak of COVID-19 in March 2020, there has been a significant increase in worldwide inflation. Specifically, from May 2021 to February 2023, the 12-month change in the CPI index in the U.S. has not dropped below 5% (Bureau of Labor Statistics, 2023). Prior to the pandemic, the U.S. economy experienced nearly four decades of low inflation. The abrupt shift from a long-term low-inflation environment

to a high-inflation environment has created substantial uncertainty and volatility in the financial markets. In 2022, the technology-heavy NASDAQ stock index recorded a yearly return of -33.10% (NASDAQ, 2023).

Equally concerning is the uncertainty around the duration of this round of high inflation. Some believe that the geopolitical tensions and the COVID-19 pandemic will overturn the trend of globalization and lead to global supply chain restructuring (Javorcik, 2020) which may result in a higher cost of production in the foreseeable future. Moreover, Ball et al. (2022) suggests that the future inflation rate may remain high if the unemployment rate remains low.

In this article, we aim to answer the following question: with the goal of outperforming a passive benchmark, how should an active investor optimize the portfolio during high inflation? It is important to note that we do not attempt to make predictions about future inflation conditions. Instead, we approach the problem by formulating a multi-period optimal control problem that considers bounded leverage constraints and specific investment criteria. This optimal control problem requires the specification of an appropriate objective function, realistic constraints, as well as stochastic models for returns of traded assets during high inflation regimes. Given the complexities and challenges associated, it is crucial to develop an efficient method capable of computing optimal solutions, accommodating flexible data sources, handling high-dimensional cases, and dealing with complex constraints. In this paper, we propose a framework to address these challenges.

In Section 2, we assume that the real (inflation-adjusted) asset returns during a high-inflation regime follow stochastic processes and treat allocation decisions as the control of a dynamic system. Specifically, we formulate an optimal control problem to outperform a fixed-mix benchmark portfolio consistently throughout the investment horizon by minimizing a cumulative quadratic tracking difference (CD) objective. There is a large amount of extant literature on closed-form solutions for beating a stochastic benchmark under synthetic market assumptions (Browne, 1999, 2000; Tepla, 2001; Basak et al., 2006; Davis and Lleo, 2008; Lim and Wong, 2010a; Oderda, 2015; Alekseev and Sokolov, 2016; Al-Aradi and Jaimungal, 2018a). In these articles, the common objective function involves a log-utility function, e.g. the log wealth ratio. Under the log wealth ratio formulation, it is often hard to accommodate a fixed stream of cash injections, which is a common characteristic of open-ended funds. Forsyth et al. (2022) consider a scenario where a fixed amount of cash injections is allowed and provides a closed-form solution under a cumulative quadratic tracking difference (CD) objective given the assumption that the stock price follows a double exponential jump-diffusion model and the bond price is deterministic. Since the assumption that the bond index price is stochastic and has jumps is more reasonable under a high-inflation scenario, we develop a closed-form solution under the case that both the stock index and the bond index follow jump-diffusion models.

The closed-form solution is derived, unfortunately, under unrealistic assumptions such as continuous rebalancing, infinite leverage, and continued trading when insolvent. A discrete-time multi-period asset allocation problem is generally solved using a dynamic programming (DP) based approach, which converts a multi-step optimization problem into multiple single-step optimization problems. However, van Staden et al. (2023) point out that dynamic programming-based approaches require the evaluation of a high-dimensional performance criterion to obtain the optimal control which is comparatively low-dimensional. This means that solving the discrete-time problem numerically using dynamic programming-based techniques (for example numerical solutions to the corresponding PIDE (Wang and Forsyth, 2010), or reinforcement learning (RL) techniques (Dixon et al., 2020; Park et al., 2020; Lucarelli and Borrotti, 2020; Gao et al., 2020)) are inefficient and are computationally prone to known issues such as error amplification over recursions (Wang et al., 2020).

Acknowledging these limitations, in Section 2.7, we propose to use a single neural network model to approximate the optimal control and solve the original optimal control problem directly via a single standard finite-dimensional optimization. This direct approximation of the control exploits the lower dimensionality of the optimal control and bypasses the problem of solving high-dimensional conditional expectations associated with DP methods. We note that the idea of using a neural network to directly approximate the control process is also used in Han et al. (2016); Buehler et al. (2019); Tsang and Wong (2020); Reppen et al. (2022), in which they propose a stacked neural network approach that includes individual sub-networks for each rebalancing step. In contrast, we propose a single shallow neural network that includes time as an input feature, and thus avoids the need to have multiple sub-networks for each rebalancing step and greatly reduces the computational and modeling complexity. Furthermore, using time as a feature in the neural

network approximation function is consistent with the observation that (under assumptions) the optimal control is naturally a continuous function of time, which we discuss in detail in Section 2.4.

The idea of using a single neural network to approximate controls has also been explored in previous studies such as Li and Forsyth (2019) and Ni et al. (2022). These studies focus on portfolio optimization problems with long-only constraints. The neural network architecture proposed in these studies transforms the constrained portfolio optimization problems into unconstrained optimization problems, making them computationally easier to solve. However, these existing neural network architectures do not address the bounded leverage constraint, which limits the total long exposure in the portfolio. The limited literature on portfolio optimization with bounded leverage is likely due to the added complexity arising from the combination of long and short positions in the portfolio.

A significant contribution of this article is the introduction of a novel leverage-feasible neural network (LFNN) model, which converts the leverage-constrained optimization problem into an unconstrained optimization problem. This model enables the incorporation of the bounded leverage constraint into the portfolio optimization framework. Additionally, in Section 2.8, we provide a mathematical proof that, under reasonable assumptions, the solution of the unconstrained optimization problem obtained using the LFNN model can approximate the optimal control of the original problem arbitrarily well. This mathematical justification validates the effectiveness and validity of the LFNN approach.

In Section 3, we present a case study on active portfolio optimization in a high-inflation regime. To identify historical high-inflation periods, we employ a simple filtering method. Subsequently, we use bootstrap resampling to generate training and testing data sets, which consist of price paths for four assets: the equal-weighted and cap-weighted stock indexes, as well as the 30-day and 10-year U.S. treasury indexes.

Using the leverage-feasible neural network (LFNN) model and the cumulative quadratic tracking shortfall (CS) objective, we derive a leverage-constrained strategy for portfolio optimization. Our results demonstrate that the LFNN model produces a strategy that consistently outperforms the fixed-mix benchmark. Specifically, the strategy achieves a median (annualized) internal rate of return (IRR) that is more than 2% higher than the benchmark. Moreover, there is a probability of over 90% that the strategy will yield a higher terminal wealth compared to the benchmark.

These findings highlight the efficacy of the LFNN model in optimizing portfolios under high-inflation conditions. By incorporating the bounded leverage constraint and utilizing the CS objective, our approach enables investors to achieve superior performance and mitigate risks in a high-inflation environment.

Our contributions are summarized below:

- (i) To gain intuition about the behavior of the optimal controls, we derive the closed-form solution under a jump-diffusion asset price model and other typical assumptions (such as continuous rebalancing) for a two-asset case. The closed-form solution provides important insights into the properties of the optimal control as well as meaningful interpretations of the neural network models that approximate the controls.
- (ii) We propose to represent the control directly by a neural network representation so that the stochastic optimal control problem can be solved numerically under realistic constraints such as discrete rebalancing and limited leverage. Particularly, we propose the novel leverage-feasible neural network (LFNN) model to convert the original complex leverage-constrained optimization problem into an unconstrained optimization problem that can be solved easily by standard optimization methods.
- (iii) We prove that, with a suitable choice of the hyperparameter of the LFNN model, the solution of the parameterized unconstrained optimization problem can approximate the optimal control arbitrarily well. This provides a mathematical justification for the validity of the LFNN approach. This is further supported by the numerical results that the performance of the LFNN model matches the clipped form of the closed-form solution on simulated data.
- (iv) In the case study on active portfolio optimization in high-inflation, we apply the neural network method to bootstrap resampled asset returns with four underlying assets, including the equal-weighted/cap-weighted stock indexes, and the 30-day/10-year treasury bond indexes. The dynamic strategy from

the learned LFNN model outperforms the fixed-mix benchmark strategy consistently throughout the investment horizon, with a 2% higher median (annualized) internal rate of return (IRR), and more than 90% probability of achieving a higher terminal wealth. Furthermore, the learned allocation strategy suggests that the equal-weighted stock index and short-term bonds are preferable investment assets during high-inflation regimes.

2 Outperform dynamic benchmark under bounded leverage

2.1 Sovereign wealth funds and benchmark targets

Instead of taking a passive approach, some of the largest sovereign wealth funds often adopt an active management philosophy and use passive portfolios as the benchmark to evaluate the efficiency of active management. For example, the Canadian Pension Plan (CPP) uses a base reference portfolio of 85% global equity and 15% Canadian government bonds (CPP Investments, 2022). Another example is the Government Pension Fund Global of Norway (also known as the oil fund) managed by Norges Bank Investment Management (NBIM), which uses a benchmark index consisting of 70% equity index and 30% bond index.¹ The benchmark equity index is constructed based on the market capitalization for equities in the countries included in the benchmark. The benchmark index for bonds specifies a defined allocation between government bonds and corporate bonds, with a weight of 70 percent to government bonds and 30 percent to corporate bonds (Norges Bank, 2022).

However, the excess return that these well-known sovereign wealth funds have achieved over their respective passive benchmark portfolios cannot be described as impressive. In the 2022 fiscal year report, CPP claims to have beaten the base reference portfolio by an annualized 80 bps after fees over the past 5 years (CPP Investments, 2022). On the other hand, NBIM reports a mere average of 27 bps of annual excess return over the benchmark over the last decade (see Table 2.1). It is worth noting that these behemoth funds achieve seemingly meager results by hiring thousands of highly paid investment professionals and spending billions of dollars on day-to-day operations. For example, the CPP 2021 annual report (CPP Investments, 2021) lists personnel costs as CAD 938 million, for 1,936 employees, which translates to average costs of about CAD 500,000 per employee-year.

Year	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Average
Excess return (%)	0.21	0.99	-0.77	0.45	0.15	0.70	-0.30	0.23	0.27	0.74	0.27

Table 2.1: Norges Bank Investment Management, relative return to benchmark portfolio

The stark contrast between the enormous spending of sovereign wealth funds and the meager outperformance of the funds relative to the passive benchmark portfolios is probably provocative to taxpayers and pensioners who invest their hard-earned money in the funds. Equally concerning is the potential of a long, persistent inflation regime and the funds' ability to consistently beat the benchmark portfolio in such times. After all, both the CPP Investments and NBIM were established in the late 1990s, a decade after the last long inflation period ended in the mid-1980s.

These concerns prompt us to ask the following question: in a presumed persistent high-inflation environment, can a fund manager find a simple asset allocation strategy that consistently beats the benchmark passive portfolios by a reasonable margin (preferably without spending billions of dollars in personnel costs)?

2.2 Mathematical formulation

In this section, we mathematically formulate the problem of outperforming a benchmark. Let $[t_0(= 0), T]$ denote the investment horizon, and let $W(t)$ denote the wealth (value) of the portfolio actively managed

¹The Ministry of Finance of Norway sets the allocation fraction between the equity index and the bond index. It gradually raised the weight for equities from 60% to 70% from 2015-2018.

by the manager at time $t \in [t_0, T]$. We refer to the actively managed portfolio as the “active portfolio”. Furthermore, let $\hat{W}(t)$ denote the wealth of the benchmark portfolio at time $t \in [t_0, T]$. To ensure a fair assessment of the relative performance of the two portfolios, we assume both portfolios start with an equal initial wealth amount $w_0 > 0$, i.e., $W(t_0) = \hat{W}(t_0) = w_0 > 0$.

Technically, the admissible sets of underlying assets for the active and passive portfolio need not be identical. However, for simplicity, we assume that both the active portfolio and the benchmark portfolio can allocate wealth to the same set of N_a assets.

Let vector $\mathbf{S}(t) = (S_i(t) : i = 1, \dots, N_a)^\top \in \mathbb{R}^{N_a}$ denote the asset prices of the N_a underlying assets at time $t \in [t_0, T]$. In addition, let vectors $\mathbf{p}(t) = (p_i(t) : i = 1, \dots, N_a)^\top \in \mathbb{R}^{N_a}$ and $\hat{\mathbf{p}}(t) = (\hat{p}_i(t) : i = 1, \dots, N_a)^\top \in \mathbb{R}^{N_a}$ denote the fraction of wealth allocated to the N_a underlying assets at time $t \in [t_0, T]$, respectively, for the active portfolio and the benchmark portfolio. From a control theory perspective, the allocation vector \mathbf{p} can be regarded as the control of the system, as it determines how the wealth of the active portfolio evolves over time. We will seek to find the optimal feedback control. In other words, the closed-loop controls (allocation decisions) are assumed to also depend on the value of the state variables (e.g. portfolio wealth). Therefore, we consider the control \mathbf{p} to be a function of time as well as the relevant state variables. In addition, the benchmark portfolio allocation $\hat{\mathbf{p}}$ can be regarded as a known function of its state variables and time as well.

Mathematically, $\mathbf{p}(\mathbf{X}(t)) = (p_i(\mathbf{X}(t)) : i = 1, \dots, N_a)^\top \in \mathbb{R}^{N_a}$ and $\hat{\mathbf{p}}(\hat{\mathbf{X}}(t)) = (\hat{p}_i(\hat{\mathbf{X}}(t)) : i = 1, \dots, N_a)^\top \in \mathbb{R}^{N_a}$, where $\mathbf{X}(t) \in \mathcal{X} \subseteq \mathbb{R}^{N_x}$ and $\hat{\mathbf{X}}(t) \in \hat{\mathcal{X}} \subseteq \mathbb{R}^{N_{\hat{x}}}$ are the state variables taken into account by the active portfolio and the benchmark portfolio respectively. Here we include t in $\mathbf{X}(t)$ and $\hat{\mathbf{X}}(t)$ for notational simplicity. In this article, we consider the particular problem of outperforming a passive portfolio, in which $\mathbf{X}(t) = (t, W(t), \hat{W}(t))^\top$.

We assume that the active portfolio and the benchmark portfolio follow the same rebalancing schedule denoted by $\mathcal{T} \subseteq [t_0, T]$. In the case of discrete rebalancing, $\mathcal{T} \subset [t_0, T]$ is a discrete set. In the case of continuous rebalancing, $\mathcal{T} = [t_0, T]$, i.e., rebalancing happens continuously throughout the entire investment horizon.

Additionally, we assume both portfolios follow the same deterministic sequence of cash injections, defined by the set $\mathcal{C} = \{c(t), t \in \mathcal{T}_c\}$, where $\mathcal{T}_c \subseteq [t_0, T]$ is the schedule of the cash injections. When \mathcal{T}_c is a discrete injection schedule, $c(t)$ is the amount of cash injection at t . In the case of continuous cash injections, i.e., $\mathcal{T} = [t_0, T]$, $c(t)$ is the rate of cash injection at t , i.e., the total cash injection amount during $[t, t + dt]$ is $c(t)dt$, where dt is an infinitesimal time interval. For simplicity, we assume that $\mathcal{T}_c = \mathcal{T}$, so that the cash injections schedule is the same as the rebalancing schedule. At $t \in \mathcal{T}$, $W(t)$ and $\hat{W}(t)$ always denote the wealth after the cash injection (assuming there is a cash injection event happening at t).

The active and benchmark strategies, respectively, are defined as the sequence of the allocation fractions following the rebalancing schedule. Mathematically, the active and benchmark strategies are defined by sets

$$\mathcal{P} = \{\mathbf{p}(\mathbf{X}(t)), t \in \mathcal{T}\}, \quad \text{and} \quad \hat{\mathcal{P}} = \{\hat{\mathbf{p}}(\hat{\mathbf{X}}(t)), t \in \mathcal{T}\}. \quad (2.1)$$

Denote \mathcal{A} as the set of admissible strategies, which reflects the investment constraints on the controls. We assume that admissibility can vary with state and let $\{\mathcal{X}_i : i = 1, \dots, k\}$ be a partition of \mathcal{X} (the state variable space), i.e.

$$\begin{cases} \bigcup_{i=1}^k \mathcal{X}_i = \mathcal{X}, \\ \mathcal{X}_i \cap \mathcal{X}_j = \emptyset, \forall 1 \leq i < j \leq k, \end{cases} \quad (2.2)$$

and $\{\mathcal{Z}_i \subseteq \mathbb{R}^{N_a} : i = 1, \dots, k\}$ be the corresponding value sets of feasible controls such that any feasible control \mathbf{p} satisfies

$$\mathbf{p}(\mathbf{x}) \in \mathcal{Z}_i, \forall \mathbf{x} \in \mathcal{X}_i, \forall i \in \{1, \dots, k\}. \quad (2.3)$$

We say that strategy \mathcal{P} is an admissible strategy, i.e., $\mathcal{P} \in \mathcal{A}$, if and only if

$$\mathcal{P} = \left\{ \mathbf{p}(\mathbf{X}(t)), t \in \mathcal{T} \mid \mathbf{p}(\mathbf{X}(t)) \in \mathcal{Z}_i, \text{ if } \mathbf{X}(t) \in \mathcal{X}_i \right\} \quad (2.4)$$

Consider a discrete rebalancing schedule $\mathcal{T} = \{t_j, j = 0, \dots, N\}$ with N rebalancing events, where $t_0 < t_1 < \dots < t_N = T$.² Then, the wealth evolution of the active portfolio and the benchmark portfolio can be described by the equations

$$\begin{cases} W(t_{j+1}) = \left(\sum_{i=1}^{N_a} p_i(\mathbf{X}(t_j)) \cdot \frac{S_i(t_{j+1}) - S_i(t_j)}{S_i(t_j)} \right) W(t_j) + c(t_{j+1}), & j = 0, \dots, N-1, \\ \hat{W}(t_{j+1}) = \left(\sum_{i=1}^{N_a} \hat{p}_i(\hat{\mathbf{X}}(t_j)) \cdot \frac{S_i(t_{j+1}) - S_i(t_j)}{S_i(t_j)} \right) \hat{W}(t_j) + c(t_{j+1}), & j = 0, \dots, N-1. \end{cases} \quad (2.5)$$

In the continuous rebalancing case, $\mathcal{T} = [t_0, T]$. Let $dS_i(t)$ denote the instantaneous change in price for asset i , $i \in [1, \dots, N_a]$.³ Then, at $t \in \mathcal{T} = [t_0, T]$, the wealth dynamics of the active portfolio and the benchmark portfolio, following their respective strategies \mathcal{P} and $\hat{\mathcal{P}}$, can be described by the equations

$$\begin{cases} dW(t) = \left(\sum_{i=1}^{N_a} p_i(\mathbf{X}(t)) \cdot \frac{dS_i(t)}{S_i(t)} \right) W(t) + c(t)dt, \\ d\hat{W}(t) = \left(\sum_{i=1}^{N_a} \hat{p}_i(\hat{\mathbf{X}}(t)) \cdot \frac{dS_i(t)}{S_i(t)} \right) \hat{W}(t) + c(t)dt. \end{cases} \quad (2.6)$$

Let sets $\mathcal{W}_{\mathcal{P}} = \{W(t), t \in \mathcal{T}\}$ and $\hat{\mathcal{W}}_{\hat{\mathcal{P}}} = \{\hat{W}(t), t \in \mathcal{T}\}$ denote the wealth trajectories of the active portfolio and the benchmark portfolio following their respective investment strategies \mathcal{P} and $\hat{\mathcal{P}}$. Let $F(\mathcal{W}_{\mathcal{P}}, \hat{\mathcal{W}}_{\hat{\mathcal{P}}}) \in \mathbb{R}$ denote an investment metric that measures the performances of the active and benchmark strategies, based on their respective wealth trajectories.

In this article, we assume that the asset prices $\mathbf{S}(t) \in \mathbb{R}^{N_a}$ are stochastic. Then, the wealth trajectories $\mathcal{W}_{\mathcal{P}}$ and $\hat{\mathcal{W}}_{\hat{\mathcal{P}}}$ are also stochastic, as well as the performance metric $F(\mathcal{W}_{\mathcal{P}}, \hat{\mathcal{W}}_{\hat{\mathcal{P}}})$, which measures the relative performance of the active strategy with respect to the benchmark strategy. Therefore, when investment managers target to optimize an investment metric, the evaluation is often on the expectation of the random metric.

Let $\mathbb{E}_{\mathcal{P}}^{(t_0, w_0)}[F(\mathcal{W}_{\mathcal{P}}, \hat{\mathcal{W}}_{\hat{\mathcal{P}}})]$ denote the expectation of the value of the performance metric F , with respect to a given initial wealth $w_0 = W(0) = \hat{W}(0)$ at time $t_0 = 0$, following an admissible investment strategies $\mathcal{P} \in \mathcal{A}$, and the benchmark investment strategy $\hat{\mathcal{P}}$. Since the benchmark strategy is often pre-determined and known, we keep the benchmark strategy $\hat{\mathcal{P}}$ implicit in this notation for simplicity. Subsequently we use $\mathbb{E}_{\mathcal{P}}^{(t_0, w_0)}[F(\mathcal{W}_{\mathcal{P}}, \hat{\mathcal{W}}_{\hat{\mathcal{P}}})]$, the expectation of a desired performance metric, as the (*investment*) *objective function* and solve

$$\text{(Optimization problem):} \quad \inf_{\mathcal{P} \in \mathcal{A}} \mathbb{E}_{\mathcal{P}}^{(t_0, w_0)}[F(\mathcal{W}_{\mathcal{P}}, \hat{\mathcal{W}}_{\hat{\mathcal{P}}})]. \quad (2.7)$$

2.3 Choice of investment objective

The first step to designing a proper outperforming investment objective is to clarify the definition of *beating the benchmark*. In the context of measuring the performance of the portfolio against the benchmark, a common metric is the tracking error, which measures the volatility of the difference in returns, i.e.,

$$\text{Tracking error} = \text{stdev}(R - \hat{R}), \quad (2.8)$$

where R denotes the return of the active portfolio, and \hat{R} denotes the return of the benchmark portfolio. Note that the returns of the active portfolio and the benchmark portfolio are determined from their respective wealth trajectories ($\mathcal{W}_{\mathcal{P}}$ and $\hat{\mathcal{W}}_{\hat{\mathcal{P}}}$) that are evaluated under the same investment horizon and same market

²Technically, at $t = t_0$, the manager makes the initial asset allocation, rather than a “rebalancing” of the portfolio. However, despite the different purposes, a rebalancing of the portfolio is simply a new allocation of the portfolio wealth. Therefore, for notational simplicity, we include t_0 in the rebalancing schedule.

³For illustration purposes, here we assume $S_i(t)$, $i \in [1, \dots, N_a]$ follow standard diffusion processes, i.e., no jumps. We will discuss the case with jumps in detail in Section 2.4.

conditions. The tracking error measures the volatility of the difference in returns over the investment horizon. A criticism of the tracking error is that it only measures the variability in the difference in returns, but does not reflect the magnitude of the return difference itself. For example, an active strategy with a constant negative return difference over the investment horizon would yield a better tracking error than an active strategy with a positive but volatile return difference. For this reason, many prefer the tracking difference (Johnson et al., 2013; Hougan, 2015; Charteris and McCullough, 2020; Boyde, 2021), which is defined as the annualized difference between the active portfolio’s cumulative return and the benchmark portfolio’s cumulative return over a specific period. Note that both tracking error and tracking difference metrics measure the return difference of the active portfolio over the benchmark portfolio. In other words, these metrics measure how closely the return of the active portfolio tracks the return of the benchmark portfolio. In practice, if an investment manager aims to achieve a certain annualized relative return target, e.g., β , then the tracking difference metric may not be appropriate. To address this, van Staden et al. (2022) suggests the investment objective

$$\inf_{\mathcal{P} \in \mathcal{A}} \mathbb{E}_{\mathcal{P}}^{(t_0, w_0)} \left[\left(W(T) - e^{\beta T} \hat{W}(T) \right)^2 \right], \quad (2.9)$$

where $W(T)$ and $\hat{W}(T)$ are the respective terminal wealth of the active portfolio and the benchmark portfolio at terminal time T , and β is the annualized relative return target.

The optimal control problem (2.9) aims to produce an active strategy that minimizes the quadratic difference between $W(T)$ and the terminal portfolio value target of $e^{\beta T} \hat{W}(T)$. In other words, the optimal control policy tries to outperform the benchmark portfolio by a total factor of $e^{\beta T}$ over the time horizon $[0, T]$, which is equivalent to an annualized relative return of β . The quadratic term of the difference incentivizes the terminal wealth of the active portfolio $W(T)$ to closely track the *elevated target* $e^{\beta T}$.

It is worth noting that the relative return target β can be intuitively interpreted as the manager’s willingness to take more risks. As $\beta \downarrow 0$, the optimal solution to problem (2.9) is simply to mimic the benchmark strategy. However, as β grows larger, the manager needs to take on more risk (for more return) in order to beat the benchmark portfolio by the relative return target rate.

A criticism of the investment objective (2.9) is that it is symmetrical in terms of the outperformance and underperformance of $W(T)$ relative to the elevated target $e^{\beta T} \hat{W}(T)$. This is a common issue for volatility-based measures, such as the Sharpe ratio (Ziemba, 2005). In practice, investors may favor outperformance more than underperformance, while still aiming to track the elevated target closely. Acknowledging this, instead of (2.9), Ni et al. (2022) propose the following asymmetrical objective function,

$$\inf_{\mathcal{P} \in \mathcal{A}} \mathbb{E}_{\mathcal{P}}^{(t_0, w_0)} \left[\left(\min(W(T) - e^{\beta T} \hat{W}(T), 0) \right)^2 + \max(W(T) - e^{\beta T} \hat{W}(T), 0) \right]. \quad (2.10)$$

The investment objective (2.10) penalizes the outperformance (of $W(T)$ relative to the elevated target $e^{\beta T} \hat{W}(T)$) linearly but the underperformance quadratically, thus encouraging the optimal policy to favor outperformance more than underperformance when necessary. Note that the use of objective function (2.10) does not permit closed-form solutions and machine learning techniques are used (Ni et al., 2022) to compute the desired optimal strategy numerically.

Another criticism of the investment objective (2.9) and (2.10) is that both are only concerned with the relative performance at terminal time T . In reality, investment managers are often required to report intermediate portfolio performance internally or externally at regular time intervals. Instead of achieving the annualized relative return target when reviewing the portfolio performance at the end of the investment horizon, managers may want to consistently achieve the relative return target throughout the entire investment horizon. In this case, managers may need to set an investment objective function to control the deviation of the wealth of the portfolio from the target along a market scenario within the investment horizon. Consequently, van Staden et al. (2022) propose the following cumulative quadratic tracking difference (CD) objectives

$$(CD(\beta)) : \begin{cases} \inf_{\mathcal{P} \in \mathcal{A}} \mathbb{E}_{\mathcal{P}}^{(t_0, w_0)} \left[\int_{t_0}^T (W(t) - e^{\beta t} \hat{W}(t))^2 dt \right], & \text{if } \mathcal{T} = [t_0, T], \\ \inf_{\mathcal{P} \in \mathcal{A}} \mathbb{E}_{\mathcal{P}}^{(t_0, w_0)} \left[\sum_{t \in \mathcal{T} \cup \{T\}} (W(t) - e^{\beta t} \hat{W}(t))^2 \right], & \text{if } \mathcal{T} \subseteq [t_0, T], \mathcal{T} \text{ discrete.} \end{cases} \quad (2.11)$$

Here, note that objective (2.11) is for the continuous rebalancing case, and (2.12) for discrete rebalancing. Both (2.11) and (2.12) measure the cumulative deviation of the wealth of the active portfolio relative to the target, along a market scenario within the entire investment horizon. Therefore, they measure the intermediate performance deviations effectively. However, similar to (2.9), (2.11) and (2.12) penalize outperformance and underperformance symmetrically. Therefore, we also consider following cumulative quadratic shortfall (CS) objectives that only penalize the shortfall (underperformance with respect to the target)

$$(CS(\beta)) : \begin{cases} \inf_{\mathcal{P} \in \mathcal{A}} \mathbb{E}_{\mathcal{P}}^{(t_0, w_0)} \left[\int_{t_0}^T (\min(W(t) - e^{\beta t} \hat{W}(t), 0))^2 dt + \epsilon W(T) \right], & \text{if } \mathcal{T} = [t_0, T], \\ \inf_{\mathcal{P} \in \mathcal{A}} \mathbb{E}_{\mathcal{P}}^{(t_0, w_0)} \left[\sum_{t \in \mathcal{T} \cup \{T\}} (\min(W(t) - e^{\beta t} \hat{W}(t), 0))^2 + \epsilon W(T) \right], & \text{if } \mathcal{T} \subseteq [t_0, T], \mathcal{T} \text{ discrete.} \end{cases} \quad (2.13)$$

Here (2.13) and (2.14) are the investment objectives for the continuous rebalancing and discrete rebalancing cases respectively. ϵ is a small regularization parameter to ensure that problem (2.13) and (2.14) are well-posed. A more detailed comparison of the CD and CS objective functions can be found in Appendix H in Ni et al. (2023).

2.4 Closed-form solution for CD problem

In this section, we present the closed-form solution to the CD problem (2.11) under several assumptions. The closed-form solution not only provides us with insights for understanding the CD-optimal controls for problem (2.11), but also serves as the baseline for understanding the numerical results derived from the neural network method (discussed in later sections). Specifically, in this section, we consider the case that all asset prices follow jump-diffusion processes and portfolios with cash injections, which are aspects not frequently considered in benchmark outperformance literature Browne (1999, 2000); Tepla (2001); Basak et al. (2006); Yao et al. (2006); Zhao (2007); Davis and Lleo (2008); Lim and Wong (2010b); Oderda (2015); Zhang and Gao (2017); Al-Arabi and Jaimungal (2018b); Nicolosi et al. (2018); Bo et al. (2021).

We first summarize the assumptions for obtaining the closed-form solution to CD problem (2.11).

Assumption 2.1. (*Two assets, no friction, unlimited leverage, trading in insolvency, constant rate of cash injection*) The active portfolio and the benchmark portfolio have access to two underlying assets, a stock index, and a constant-maturity bond index. Both portfolios are rebalanced continuously, i.e., $\mathcal{T} = [t_0, T]$. There is no transaction cost and no leverage limit. Furthermore, we assume that trading continues in the event of insolvency, i.e., when $W(t) < 0$ for some $t \in [t_0, T]$. Finally, we assume both portfolios receive constant cash injections with an injection rate of c , which means during any time interval $[t, t + \Delta t] \subseteq [t_0, T], \forall \Delta t > 0$, both portfolios receive cash injection amount of $c\Delta t$.

Remark 2.1. (Remark on Assumption 2.1) For illustration purposes, we assume only two underlying assets. However, the technique for deriving the closed-form solution can be extended to multiple assets. We remark that unlimited leverage is unrealistic, and is only assumed for deriving the closed-form solution. In Section C.1, we will discuss the technique for handling the leverage constraint in more detail. We also acknowledge that it is not realistic to assume that the manager can continue to trade and borrow when insolvent. However,

this assumption is typically required for obtaining closed-form solutions, see Zhou and Li (2000); Li and Ng (2000) for the case of a multi-period mean-variance asset allocation problem. In Appendix C.1, we have more discussion on the impact of insolvency and its handling in experiments.

Assumption 2.2. (*Fixed-mix benchmark strategy*) We assume that the benchmark strategy is a fixed-mix strategy (also known as the constant weight strategy). We assume the benchmark always allocates a constant fraction of $\hat{\varrho}$ ($\in \mathbb{R}$) of the portfolio wealth to the stock index, and a constant fraction of $1 - \hat{\varrho}$ to the bond index. Let $\hat{\varrho} = (\hat{\varrho}, 1 - \hat{\varrho})^\top \in \mathbb{R}^2$ denote the vector of allocation fractions to the stock index and the bond index, the benchmark strategy is the fixed-mix strategy defined by $\hat{\mathcal{P}} = \{\hat{\mathbf{p}}(\hat{\mathbf{X}}(t)) = (\hat{p}_1(\hat{\mathbf{X}}(t)), \hat{p}_2(\hat{\mathbf{X}}(t)))^\top \equiv \hat{\varrho}, \forall t \in \mathcal{T}\}$.

Finally, we assume the stock index price and bond index price follow the jump-diffusion processes described below.

Assumption 2.3. (*Jump-diffusion processes*) Let $S_1(t)$ and $S_2(t)$ denote the deflated (adjusted by inflation) price of the stock index and the bond index at time $t \in [t_0, T]$. We assume $S_i(t)$, $i \in \{1, 2\}$ follow the jump-diffusion processes

$$\frac{dS_i(t)}{S_i(t^-)} = (\mu_i - \lambda_i \kappa_i + r_i \cdot \mathbf{1}_{S_i(t^-) < 0})dt + \sigma_i dZ_i(t) + d\left(\sum_{k=1}^{\pi_i(t)} (\xi_i^{(k)} - 1)\right), \quad i = 1, 2. \quad (2.15)$$

Here μ_i are the (uncompensated) drift rate, σ_i is the diffusive volatility, $Z_1(t), Z_2(t)$ are correlated Brownian motions, where $\mathbb{E}[dZ_1(t) \cdot dZ_2(t)] = \rho dt$. r_i are the borrowing premiums when $S_i(t^-)$ is negative.⁴ $\pi_i(t)$ is a Poisson process with positive intensity parameter λ_i . $\{\xi_i^{(k)}, k = 1, \dots, \pi_i(t)\}$ are i.i.d. positive random variables that describe jump multipliers associated with the assets. If a jump occurs for asset i at time $t \in (t_0, T]$, its underlying price jumps from $S_i(t^-)$ to $S_i(t) = \xi_i \cdot S_i(t^-)$.⁵ $\kappa_i = \mathbb{E}[\xi_i - 1]$. ξ_i and $\pi_i(t)$ are independent of each other. Moreover, $\pi_1(t)$ and $\pi_2(t)$ are assumed to be independent.⁶

Remark 2.2. (Motivation for jump-diffusion model) The assumption of stock index price following a jump-diffusion model is common in the financial mathematics literature (Merton, 1976; Kou, 2002). In addition, we follow the practitioner approach and directly model the returns of the constant maturity bond index as a stochastic process, see for example Lin et al. (2015); MacMinn et al. (2014). As in MacMinn et al. (2014), we also assume that the constant maturity bond index follows a jump-diffusion process. During high-inflation regimes, central banks often make rate hikes to curb inflation, which causes sudden jumps in bond prices (Lahaye et al., 2011). We believe this is an appropriate assumption for bonds in high-inflation regimes.

Under the jump-diffusion model (2.15), the wealth processes for the active portfolio and benchmark portfolio are

$$\begin{cases} dW(t) = \left(\sum_{i=1}^{N_a} p_i(\mathbf{X}(t^-)) \cdot \frac{dS_i(t)}{S_i(t^-)}\right)W(t_j^-) + cdt, \\ d\hat{W}(t) = \left(\sum_{i=1}^{N_a} \hat{p}_i(\hat{\mathbf{X}}(t^-)) \cdot \frac{dS_i(t)}{S_i(t^-)}\right)\hat{W}(t_j^-) + cdt, \end{cases} \quad (2.16)$$

where $t \in (t_0, T]$, $W(t_0) = \hat{W}(t_0) = w_0$ and $X(t^-) = (t, W(t^-), \hat{W}(t^-))^\top \in \mathbb{R}^3$ is the state variable vector.

We now derive the closed-form solution of the CD problem (2.11) under Assumption 2.1, 2.2 and 2.3. We first present the verification theorem for the HJB integro-differential equation (PIDE) satisfied by the value function and the optimal control of the CD problem (2.11).

Theorem 2.1. (*Verification theorem for CD problem (2.11)*) For a fixed $\beta > 0$, assume that for all $(t, w, \hat{w}, \hat{\varrho}) \in [t_0, T] \times \mathbb{R}^3$, there exists a function $V(t, w, \hat{w}, \hat{\varrho}) : [t_0, T] \times \mathbb{R}^3 \mapsto \mathbb{R}$ and $p^*(t, w, \hat{w}, \hat{\varrho}) : [t_0, T] \times \mathbb{R}^3 \mapsto \mathbb{R}^2$ that

⁴Intuitively, there is a premium for shorting an asset. In the closed-form solution derivation, we assume $r_i = 0$.

⁵For any functional $\psi(t)$, we use the notation $\psi(t^-)$ as shorthand for the left-sided limit $\psi(t^-) = \lim_{\Delta t \downarrow 0} \psi(t - \Delta t)$.

⁶See Forsyth (2020) for the discussion on the empirical evidence for stock-bond jump independence. Also note that the assumption of independent jumps can be relaxed without technical difficulty if needed (Kou, 2002), but will significantly increase the complexity of notations.

satisfy the following two properties. (i) V and \mathbf{p}^* are sufficiently smooth and solve the HJB PIDE (2.17), and (ii) the function $\mathbf{p}^*(t, w, \hat{w}, \hat{\varrho})$ attains the pointwise infimum in (2.17) below

$$\begin{cases} \frac{\partial V}{\partial t} + (w - e^{\beta t} \hat{w})^2 + \inf_{\mathbf{p} \in \mathbb{R}^2} H(\mathbf{p}; t, w, \hat{w}, \hat{\varrho}) = 0, \\ V(T, w, \hat{w}, \hat{\varrho}) = 0, \end{cases} \quad (2.17)$$

where

$$\begin{aligned} H(\mathbf{p}; t, w, \hat{w}, \hat{\varrho}) &= (w \cdot \boldsymbol{\alpha}^\top \mathbf{p} + c) \cdot \frac{\partial V}{\partial w} + (\hat{w} \cdot \boldsymbol{\alpha}^\top \hat{\varrho} + c) \cdot \frac{\partial V}{\partial \hat{w}} - \left(\sum_i \lambda_i \right) \cdot V(t, w, \hat{w}, \hat{\varrho}) \\ &+ \frac{w^2}{2} \cdot (\mathbf{p}^\top \boldsymbol{\Sigma} \mathbf{p}) \cdot \frac{\partial^2 V}{\partial w^2} + \frac{\hat{w}^2}{2} \cdot (\hat{\varrho}^\top \boldsymbol{\Sigma} \hat{\varrho}) \cdot \frac{\partial^2 V}{\partial \hat{w}^2} + w \hat{w} \cdot (\mathbf{p}^\top \boldsymbol{\Sigma} \hat{\varrho}) \cdot \frac{\partial^2 V}{\partial w \partial \hat{w}} \\ &+ \sum_i \lambda_i \int_0^\infty V(w + p_i w(\xi - 1), \hat{w} + \hat{p}_i \hat{w}(\xi - 1), t, \hat{\varrho}) f_{\xi_i}(\xi) d\xi. \end{aligned} \quad (2.18)$$

Here $\boldsymbol{\alpha} = (\mu_1 - \lambda_1 \kappa_1, \mu_2 - \lambda_2 \kappa_2)^\top$ is the vector of (compensated) drift rates, $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix}$ is the covariance matrix, and f_{ξ_i} is the density function for ξ_i .

Then, under Assumption 2.1, 2.2 and 2.3, V is the value function and \mathbf{p}^* is the optimal control for the CD problem (2.11).

Proof. See Appendix A.1 □

Define several auxiliary variables

$$\begin{cases} \kappa_i^{(2)} = \mathbb{E}[(\xi_i - 1)^2], & (\sigma_i^{(2)})^2 = (\sigma_i)^2 + \lambda_i \kappa_i^{(2)}, \quad i \in \{1, 2\}, \\ \vartheta = \sigma_1 \sigma_2 \rho - (\sigma_2^{(2)})^2, & \gamma = (\sigma_1^{(2)})^2 + (\sigma_2^{(2)})^2 - 2\sigma_1 \sigma_2 \rho, \\ \phi = \frac{(\mu_1 - \mu_2)(\mu_1 - \mu_2 + \vartheta)}{\gamma}, & \eta = \frac{(\mu_1 - \mu_2 + \vartheta)^2}{\gamma} - (\sigma_2^{(2)})^2, \end{cases} \quad (2.19)$$

then we have the following proposition regarding the optimal control of problem (2.11).

Proposition 2.1. (CD-optimal control) Suppose Assumption 2.1, 2.2 and 2.3 are applicable, then the optimal control fraction of the wealth of the active portfolio to be invested in the stock index for the CD(β) problem (2.11) is given by $p^*(t, w, \hat{w}, \hat{\varrho}) \in \mathbb{R}$, where

$$p^*(t, w, \hat{w}, \hat{\varrho}) = \frac{1}{W^*(t)} \left[\frac{(\mu_1 - \mu_2)}{\gamma} h(t; \beta, c) + \frac{(\mu_1 - \mu_2 + \vartheta)}{\gamma} \left(g(t; \beta) \hat{W}(t) - W^*(t) \right) + g(t; \beta) \hat{W}(t) \cdot \hat{\varrho} \right]. \quad (2.20)$$

Here $W^*(t)$ denotes the wealth process of the active portfolio from (2.6) following control $\mathbf{p}^*(t, W^*(t), \hat{W}(t), \hat{\varrho}) = \left(p^*(t, W^*(t), \hat{W}(t), \hat{\varrho}), 1 - p^*(t, W^*(t), \hat{W}(t), \hat{\varrho}) \right)^\top$, where p^* is the optimal stock allocation described in (2.20), and $\hat{W}(t)$ is the wealth process of the benchmark portfolio following the fixed-mixed strategy described in Assumption 2.2. Here, h and g are deterministic functions of time,

$$g(t; \beta) = -\frac{D(t; \beta)}{2A(t)}, \quad h(t; \beta, c) = -\frac{B(t; \beta, c)}{2A(t)}, \quad (2.21)$$

where A, D and B are deterministic functions defined as

$$A(t) = \frac{e^{(2\mu_2 - \eta)(T-t)} - 1}{(2\mu_2 - \eta)}, \quad D(t; \beta) = 2e^{\beta T} \left(\frac{e^{-\beta(T-t)} - e^{(2\mu_2 - \eta)(T-t)}}{2\mu_2 - \eta + \beta} \right), \quad (2.22)$$

and

$$B(t; \beta, c) = \frac{2c}{2\mu_2 - \eta} \left(\frac{e^{(2\mu_2 - \eta)(T-t)} - e^{(\mu_2 - \phi)(T-t)}}{\mu_2 + \phi - \eta} - \frac{e^{(\mu_2 - \phi)(T-t)} - 1}{\mu_2 - \phi} \right) + \frac{2ce^{\beta T}}{2\mu_2 - \eta + \beta} \left(\frac{e^{(\mu_2 - \phi)(T-t)} - e^{-\beta(T-t)}}{\mu_2 - \phi + \beta} - \frac{e^{(2\mu_2 - \eta)(T-t)} - e^{(\mu_2 - \phi)(T-t)}}{\mu_2 + \phi - \eta} \right). \quad (2.23)$$

Proof. See Appendix A.2. □

2.4.1 Insights from CD-optimal control

The CD-optimal control (2.20) provides insights into the behaviour of the optimal allocation policy. For ease of exposition, we first establish the following properties of $g(t; \beta)$ and $h(t; \beta, c)$.

Corollary 2.1. (*Properties of $g(t; \beta)$*) The function $g(t; \beta)$ defined in (2.21) has the following properties for $t \in [t_0, T]$ and $\beta > 0$:

(i) For fixed $t \in [t_0, T]$, $g(t; \beta)$ is strictly increasing on $\beta \in (0, \infty)$.

(ii) For fixed $\beta > 0$, $g(t; \beta)$ is strictly increasing on $t \in [t_0, T]$.

(iii) $g(t; \beta)$ admits the following bounds:

$$e^{\beta t} \leq g(t; \beta) \leq e^{\beta T}. \quad (2.24)$$

Proof. See Appendix A.3. □

Corollary 2.2. (*Properties of $h(t; \beta, c)$*) The function $h(t; \beta, c)$ defined in (2.21) has the following properties for $t \in [t_0, T]$, $\beta > 0$ and $c \geq 0$:

(i) For fixed $t \in [t_0, T]$ and $c > 0$, $h(t; \beta, c)$ is strictly increasing on $\beta \in (0, \infty)$.

(ii) $h(t; \beta, c) \geq 0$, $\forall (t, \beta, c) \in [t_0, T] \times (0, \infty) \times [0, \infty)$.

(iii) For fixed $t \in [t_0, T]$ and $\beta > 0$, $h(t; \beta, c)$ is strictly increasing on $c \in [0, \infty)$. $h(t; \beta, 0) \equiv 0$. Moreover, $h(t; \beta, c) \propto c$, i.e. $h(t; \beta, c)$ is proportional to c .

Proof. See Appendix A.3. □

In order to analyze the closed-form solution, we make the following assumptions.

Assumption 2.4. (*Drift rates of the two assets*) We assume that the drift rates of the stock and the bond index μ_1 and μ_2 satisfy the following properties,

$$\mu_1 - \mu_2 > 0, \quad \mu_1 - \mu_2 + \vartheta > 0, \quad (2.25)$$

where ϑ is defined in (2.19).

Remark 2.3. (Remark on drift rate assumptions) The first inequality $\mu_1 - \mu_2 > 0$ indicates that the stock index has a higher drift rate than the bond index, which is a standard assumption.⁷ The second inequality $\mu_1 - \mu_2 + \vartheta > 0$ is also practically reasonable. ϑ is a variance term that is usually on a smaller scale compared to the drift rates. In reality, it is unlikely that $\mu_1 - \mu_2 > 0$ but $\mu_1 - \mu_2 + \vartheta \leq 0$.⁸

⁷In fact, in this two-asset case, this assumption does not cause loss of generality.

⁸For reference, based on the calibrated jump-diffusion model (2.15) on historical high-inflation regimes, $\mu_1 = 0.051$, $\mu_2 = -0.014$, $\vartheta = -0.00024$, and thus both inequalities are satisfied.

Now we proceed to summarize the insights from the CD-optimal control (2.20). The first obvious observation is that the CD-optimal control is a contrarian strategy. This can be seen from the fact that fixing time and the wealth of the benchmark portfolio $\hat{W}(t)$, the allocation to the more risky stock index decreases when the wealth of the active portfolio $W^*(t)$ increases.

If we take a deeper look at (2.20), we can see that the CD-optimal control consists of two components: a cash injection component p_{cash}^* and a tracking component p_{track}^* . Mathematically,

$$p^*(t, w, \hat{w}, \hat{\varrho}) = p_{cash}^*(t, w, \hat{w}) + p_{track}^*(t, w, \hat{w}, \hat{\varrho}), \quad (2.26)$$

where

$$\begin{cases} p_{cash}^*(t, w, \hat{w}) = \frac{1}{W^*(t)} \left[\frac{(\mu_1 - \mu_2)}{\gamma} h(t; \beta, c) \right], \\ p_{track}^*(t, w, \hat{w}, \hat{\varrho}) = \frac{1}{W^*(t)} \left[\frac{(\mu_1 - \mu_2 + \vartheta)}{\gamma} (g(t; \beta) \hat{W}(t) - W^*(t)) + g(t; \beta) \hat{W}(t) \cdot \hat{\varrho} \right]. \end{cases} \quad (2.27)$$

Based on Assumption 2.4 and Corollary 2.2, the cash injection component p_{cash} is always non-negative. Furthermore, from Corollary 2.2, we know that the stock allocation from the cash injection component is proportional to the cash injection rate c . In addition, as $t \uparrow T$, $h(t; \beta, c)$ increases, and thus the stock allocation from the cash injection component also increases with time.

On the other hand, the tracking component p_{track} does not depend on the cash injection rate c , but only concerns the tracking performance of the active portfolio. One key finding is that

$$\begin{cases} p_{track}^*(t, w, \hat{w}, \hat{\varrho}) \geq \hat{\varrho}, & \text{if } W^*(t) \leq g(t; \beta) \hat{W}(t), \\ p_{track}^*(t, w, \hat{w}, \hat{\varrho}) < \hat{\varrho}, & \text{if } W^*(t) > g(t; \beta) \hat{W}(t). \end{cases} \quad (2.28)$$

This means that the CD-optimal control uses $g(t; \beta) \hat{W}(t)$ as the true target for the active portfolio to decide if the active portfolio should take more or less risk than the benchmark portfolio. This is a key observation, since the CD objective function (2.11) measures the difference between $W(t)$ and $e^{\beta t} \hat{W}(t)$. One would naively think that the optimal strategy would be based on the deviation from $e^{\beta t} \hat{W}(t)$. In contrast, from Corollary 2.1, we know that the true target $g(t; \beta) \hat{W}(t)$ used for decision making is greater than $e^{\beta t} \hat{W}(t)$. The insight from this observation is that if the manager wants to track an elevated target $e^{\beta t} \hat{W}(t)$, she should aim higher than the target itself.

2.5 Leverage constraints

In practice, large pension funds such as the Canadian Pension Plan often have exposures to alternative assets, such as private equity (CPP Investments, 2022). Unfortunately, due to practical limitations, we only have access to long-term historical returns of publicly traded stock indexes and treasury bond indexes. Although controversial, some literature suggests that returns on private equity can be replicated using a leveraged small-cap stock index (Phalippou, 2014; L'Her et al., 2016). Following this line of argument, we allow managers to take leverage to invest in public stock index funds to roughly mimic the pension fund portfolios with some exposure to private equities.

Essentially, taking leverage to invest in stocks requires borrowing additional capital, which incurs borrowing costs. For simplicity, we assume the borrowing activity is represented by shorting some bond assets within the portfolio, and thus the manager is required to pay the cost of shorting these shortable assets. We assume that the cost consists of two parts: the returns of the shorted assets, and an additional borrowing premium (rate depends on specific investment scenarios) so that the total borrowing cost reflects both the interest rate environment (the return of shorted bond assets) and is reasonably estimated (with the added borrowing premium).

Following the notation from Section 2.2, we assume that the total N_a underlying assets are divided into two groups. The first group of N_l assets are long-only assets, which we index by the set $\{1, \dots, N_l\}$. The

second group of $N_a - N_l$ assets are shortable assets that can be shorted to create leverage and are indexed by the set $\{N_l + 1, \dots, N_a\}$. Recall the notation of $p_i(\mathbf{X}(t))$ for the allocation fraction for asset i at time t .

For long-only assets, the wealth fraction needs to be non-negative, hence we have

$$\text{(Long-only constraint): } p_i(\mathbf{X}(t)) \geq 0, \quad i \in \{1, \dots, N_l\}, \quad t \in \mathcal{T}. \quad (2.29)$$

Furthermore, the total allocation fraction for all assets should be one. Therefore, the following summation constraint needs to be satisfied

$$\text{(Summation constraint): } \sum_{i=1}^{N_a} p_i(\mathbf{X}(t)) = 1, \quad t \in \mathcal{T}. \quad (2.30)$$

In practice, due to borrowing costs (from taking leverage) and risk management mandates, the use of leverage is often constrained. For this reason, we cap the maximum leverage by introducing a constant p_{max} , which represents the total allocation fraction for long-only assets. Therefore,

$$\text{(Maximum leverage constraint): } \sum_{i=1}^{N_l} p_i(\mathbf{X}(t)) \leq p_{max}, \quad t \in \mathcal{T}. \quad (2.31)$$

Note that no leverage is permitted if $p_{max} = 1$.

Finally, we make the following assumption on the scenario of shorting multiple shortable assets.

Assumption 2.5. (*Simultaneous shorting*) *If one shortable asset has a negative weight, other shortable assets must have nonpositive weights. Mathematically, this assumption can be expressed as*

$$\text{(Simultaneous shorting constraint): } \begin{cases} p_i(\mathbf{X}(t)) \leq 0, \quad \forall i \in \{N_l + 1, \dots, N_a\}, \text{ if } \sum_{i=1}^{N_l} p_i(\mathbf{X}(t)) > 1, \quad t \in \mathcal{T} \\ p_i(\mathbf{X}(t)) \geq 0, \quad \forall i \in \{N_l + 1, \dots, N_a\}, \text{ if } \sum_{i=1}^{N_l} p_i(\mathbf{X}(t)) \leq 1, \quad t \in \mathcal{T} \end{cases}. \quad (2.32)$$

Remark 2.4. (Remark on Assumption 2.5) This assumption avoids the ambiguity between the long-only assets and shortable assets in scenarios that involve leverage. When leveraging occurs, all shortable assets are treated as one group to provide the needed liquidity to achieve the desired leverage level.

The above constraints consider scenarios with non-negative portfolio wealth. Before we proceed to the handling of the negative portfolio wealth scenarios, we first define the following partition of the state space \mathcal{X} ,

Definition 2.1. (*Partition of state space*) *We define $\{\mathcal{X}_1, \mathcal{X}_2\}$ to be a partition of the state space \mathcal{X} , such that*

$$\begin{cases} \mathcal{X}_1 = \left\{ x = (t, W, \hat{W})^\top \in \mathcal{X} \mid W \geq 0 \right\}, \\ \mathcal{X}_2 = \left\{ x = (t, W, \hat{W})^\top \in \mathcal{X} \mid W < 0 \right\}. \end{cases} \quad (2.33)$$

Intuitively, we separate the state space \mathcal{X} into two regions by the wealth of the active portfolio, one with non-negative wealth and the other with negative wealth. Then, we present the following assumption concerning the negative wealth (insolvency) scenarios.

Assumption 2.6. (*No trading in insolvency*) *If the wealth of the active portfolio is negative, then all long-only asset positions should be liquidated, and all the debt (i.e. the negative wealth) is allocated to the least-risky shortable asset (in terms of volatility). Particularly, without loss of generality, we assume all debt is allocated to asset $N_l + 1$. Let $\mathbf{e}_i \in \mathbb{R}^{N_a} = (0, \dots, 0, 1, 0, \dots, 0)^\top$ denote the standard basis vector of which the i -th entry is 1 and all other entries are 0. Then, we can formulate this assumption as follows.*

$$\text{(No trading in insolvency): } p(\mathbf{X}(t)) = \mathbf{e}_{N_l+1}, \quad \text{if } \mathbf{X}(t) \in \mathcal{X}_2. \quad (2.34)$$

Remark 2.5. (Remark on Assumption 2.6) Essentially, when the portfolio wealth is negative, we assume the debt is allocated to a short-term bond asset and accumulates over time.

Summarizing the constraints, we can define two sets $\mathcal{Z}_1, \mathcal{Z}_2$:

$$\mathcal{Z}_1 = \left\{ \mathbf{z} \in \mathbb{R}^{N_a} \left| \begin{cases} z_i \geq 0, \forall i \in \{1, \dots, N_l\}, \\ \sum_{i=1}^{N_a} z_i = 1, \\ \sum_{i=1}^{N_l} z_i \leq p_{max}, \\ z_i \leq 0, \forall i \in \{N_l + 1, \dots, N_a\}, \text{ if } \sum_{i=1}^{N_l} z_i > 1, \\ z_i \geq 0, \forall i \in \{N_l + 1, \dots, N_a\}, \text{ if } \sum_{i=1}^{N_l} z_i \leq 1 \end{cases} \right. \right\}, \quad (2.35)$$

$$\mathcal{Z}_2 = \{\mathbf{e}_{N_l+1}\}, \quad (2.36)$$

Then, the corresponding space of feasible control vector values \mathcal{Z} and the admissible strategy set \mathcal{A} are

$$\begin{aligned} & \left\{ \begin{array}{l} \mathcal{Z} = \mathcal{Z}_1 \cup \mathcal{Z}_2, \\ \mathcal{A} = \left\{ \mathcal{P} = \left\{ \mathbf{p}(\mathbf{X}(t)), t \in \mathcal{T} \left| \begin{cases} \mathbf{p}(\mathbf{X}(t)) \in \mathcal{Z}_1, \text{ if } \mathbf{X}(t) \in \mathcal{X}_1, \\ \mathbf{p}(\mathbf{X}(t)) \in \mathcal{Z}_2, \text{ if } \mathbf{X}(t) \in \mathcal{X}_2, \end{cases} \right. \right\} \right\}. \end{array} \right. \end{aligned} \quad (2.37)$$

$$(2.38)$$

It is not obvious how the conditional constraints in (2.37) and (2.38) can be formulated into a standard constrained optimization problem.

2.6 Neural network method

In Section 2.4, we derive the closed-form solution under the jump-diffusion model, which requires several unrealistic assumptions such as continuous rebalancing, unlimited leverage, and trading in insolvency. Furthermore, the closed-form solution is specific to the investment objective defined in the CD problem (2.11). To discover optimal strategies for high inflation regimes, capability in solving general investment problem (2.7) for different objectives and under realistic constraints, such as discrete rebalancing and limited leverage (i.e., leverage constraints discussed in Section 2.5), is critically beneficial. Therefore, we need computationally efficient methods to solve these problems numerically, particularly in high-dimensional cases.

Solving a discrete-time multi-period optimal asset allocation problem often utilizes dynamic programming (DP). For example, Dixon et al. (2020); Park et al. (2020); Lucarelli and Borrotti (2020); Gao et al. (2020) use Q-learning algorithms to solve the discrete-time multi-period optimal allocation problem. In general, if there are N_a assets to invest in, then the use of Q-learning involves approximation of an action-value function (“Q” function) which is a $(2N_a + 1)$ -dimensional function (van Staden et al., 2023) which represents the conditional expectation of the cumulative rewards at an intermediate state.⁹ Meanwhile, the optimal control is a mapping from the state space to the allocation fractions to the assets. If the state space is relatively low-dimensional,¹⁰ then the DP-based approaches are potentially unnecessarily high-dimensional.

Instead of using dynamic programming methods, Han et al. (2016); Buehler et al. (2019); Tsang and Wong (2020); Reppen et al. (2022) propose to approximate the optimal control function by neural network functions directly. In particular, they propose a stacked neural network approach that essentially uses a sub-network to approximate the control at every rebalancing step. Therefore, the number of neural networks required grows linearly with the number of rebalancing periods. Note that, in the taxonomy of Powell (2023), this method is termed as Policy Function Approximation (PFA).

In this article, we follow the lines of Li and Forsyth (2019); Ni et al. (2022) and propose a single neural network to approximate the optimal control function. The direct representation of the control function avoids the high-dimensional approximation required in DP-based methods. In addition, we consider time t as an

⁹Intuitively, the dimensionality comes from tracking the allocation in the N_a assets for both the active portfolio and benchmark portfolio when evaluating the changes in wealth of both portfolios over one period in the action-value function.

¹⁰For example, the state space of problem (2.11) with assumptions of a fixed-mix strategy is a vector in \mathbb{R}^3 .

input feature (along with the wealth of the active portfolio and benchmark portfolio), therefore avoiding the need for multiple sub-networks in the stacked neural network approach.

The numerical solution to the general problem (2.7) requires solving for the feedback control \mathbf{p} . We approximate the control function $\boldsymbol{\theta}$ by a neural network function $f(\mathbf{X}(t); \boldsymbol{\theta}) : \mathcal{X} \mapsto \mathbb{R}^{N_a}$, where $\boldsymbol{\theta} \in \mathbb{R}^{N_\theta}$ represents the parameters of the neural network (i.e., weights and biases). In other words,

$$\mathbf{p}(\mathbf{X}(t)) \simeq f(\mathbf{X}(t); \boldsymbol{\theta}) \equiv f(\cdot; \boldsymbol{\theta}). \quad (2.39)$$

Then, the optimization problem (2.7) can be converted to solving the following optimization problem.

$$\text{(Parameterized optimization problem): } \inf_{\boldsymbol{\theta} \in \mathcal{Z}_\theta} \mathbb{E}_{f(\cdot; \boldsymbol{\theta})}^{(t_0, w_0)} [F(\mathcal{W}_\theta, \hat{\mathcal{W}}_{\hat{p}})]. \quad (2.40)$$

Here \mathcal{W}_θ is the wealth trajectory of the active portfolio following the neural network approximation function parameterized by θ . $\mathcal{Z}_\theta \subseteq \mathbb{R}^{N_\theta}$ is the feasibility domain of the parameter $\boldsymbol{\theta}$, which is translated from the constraints of the original problem, e.g., (2.37) and (2.38). Mathematically,

$$\mathcal{Z}_\theta = \left\{ \boldsymbol{\theta} : \begin{cases} f(\mathbf{X}; \boldsymbol{\theta}) \in \mathcal{Z}_1, & \text{if } \mathbf{X} \in \mathcal{X}_1, \\ f(\mathbf{X}; \boldsymbol{\theta}) \in \mathcal{Z}_2, & \text{if } \mathbf{X} \in \mathcal{X}_2. \end{cases} \right\}. \quad (2.41)$$

Here $\mathcal{Z}_1, \mathcal{Z}_2$ are defined in (2.35), (2.36) and $\mathcal{X}_1, \mathcal{X}_2$ are partitions of the state space \mathcal{X} defined in Definition 2.1.

Note here that \mathcal{Z}_θ depends on the structure of the neural network function $f(\cdot; \boldsymbol{\theta})$. Intuitively, \mathcal{Z}_θ is the preimage of \mathcal{Z} , i.e., any $\theta \in \mathcal{Z}_\theta$, $f(\cdot; \boldsymbol{\theta}) \in \mathcal{Z}$. Specific neural network model design may result in $\mathcal{Z}_\theta = \mathbb{R}^{N_\theta}$, which means (2.40) becomes an unconstrained optimization problem. For long-only investment problems, the only constraints are the long-only constraint (2.29) and the summation constraint (2.30). Previous work has proposed a neural network architecture with a softmax activation function at the last layer so that the output (vector of allocation fractions) automatically satisfies the two constraints, and thus $\mathcal{Z}_\theta = \mathbb{R}^{N_\theta}$ and problem (2.40) becomes an unconstrained optimization problem (see, e.g., Li and Forsyth (2019); Ni et al. (2022)). However, as discussed in Section 2.5, we consider the more complicated case where leverage and shorting are allowed. The problem thus involves more constraints than the long-only case and therefore we would like to design a new model architecture to convert the constrained optimization problem to an unconstrained problem. We will discuss the design of the *leverage-feasible neural network* (LFNN) model in the next section, and how the LFNN model achieves this goal.

It is worth noting that for the particular CD problem (2.12) and CS problem (2.14), our technique may be formulated to appear similar to policy gradient methods in RL literature (Silver et al., 2014) on a high level. Examples of policy gradient methods in financial problems include Coache and Jaimungal (2021), in which the authors develop an actor-critic algorithm for portfolio optimization problems with convex risk measures. However, there are two main differences between our proposed methodology and policy gradient algorithms. Firstly, we assume that the randomness of the environment (i.e., asset returns) over the entire investment horizon is readily available upfront (e.g., through calibration of parametric models or resampling of historical data), which is a common assumption adopted by practitioners when backtesting investment strategies. On the other hand, RL literature often considers an unknown environment, and the algorithms focus on the exploration of the agent to learn from the unknown environment and thus may be unnecessarily complicated for our use case. Secondly, our proposed methodology is not limited to the cumulative reward framework in RL and thus is more universal and suitable for problems in which the investment objective cannot be easily expressed in the form of a cumulative reward.

2.7 Leverage-feasible neural network (LFNN)

In this section, we propose the leverage-feasible neural network (LFNN) model, which yields $\mathcal{Z}_\theta = \mathbb{R}^{N_\theta}$ for leverage constraints defined in equation (2.37), and converts a constrained optimization problem (2.40) to an unconstrained problem.

Let vector $\mathbf{x} = (t, W(t), \hat{W}(t))^\top \in \mathcal{X}$ be the feature (input) vector. We first define a standard fully-connected feedforward neural network (FNN) function $\tilde{f} : \mathcal{X} \mapsto \mathbb{R}^{N_a+1}$ as follows:

$$(\text{FNN}) : \begin{cases} h_j^{(1)} = \text{Sigmoid}\left(\sum_{i=1}^{N_x} x_i \theta_{ij}^{(1)} + b_j^{(1)}\right), j = 1, \dots, N_h^{(1)}, \\ h_j^{(k)} = \text{Sigmoid}\left(\sum_{i=1}^{N_h^{(k-1)}} h_i^{(k-1)} \theta_{ij}^{(k)} + b_j^{(k)}\right), j = 1, \dots, N_h^{(k)}, \forall k \in \{2, \dots, K\}, \\ o_j = \sum_{i=1}^{N_h^{(K)}} h_i \theta_{ij}^{(K+1)}, j = 1, \dots, N_a + 1, \\ \tilde{f}(\mathbf{x}; \boldsymbol{\theta}) := (o_1, \dots, o_{N_a+1})^\top. \end{cases} \quad (2.42)$$

Here $\text{Sigmoid}(\cdot)$ denotes the sigmoid activation function, K denotes the number of hidden layers, h_j^k denotes the value of the j -th node in the k -th hidden layer, and $N_h^{(k)}$ is the number of nodes in the k -th hidden layer. Additionally, $\boldsymbol{\theta}^{(k)} = (\theta_{ij}^{(k)}) \in \mathbb{R}^{N_h^{(k)} \times N_h^{(k-1)}}$ and $\mathbf{b}^{(k)} = (b_j^{(k)}) \in \mathbb{R}^{N_h^{(k)}}$ are the (vectorized) weight matrix and bias vector for the k -th layer,¹¹ and the parameter vector of the entire neural network is $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(1)}, \mathbf{b}^{(1)}, \dots, \boldsymbol{\theta}^{(K)}, \mathbf{b}^{(K)}, \boldsymbol{\theta}^{(K+1)})^\top \in \mathbb{R}^{N_\theta}$, where $N_\theta = \sum_{k=1}^{K+1} N_h^{(k)} \cdot N_h^{(k-1)} + \sum_{k=1}^K N_h^{(k)}$.

Building on \tilde{f} , we propose the following *leverage-feasible neural network* (LFNN) model $f : \mathcal{X} \mapsto \mathcal{Z}$:

$$(\text{LFNN}) : f(\mathbf{x}; \boldsymbol{\theta}) := \psi\left(\tilde{f}(\mathbf{x}; \boldsymbol{\theta}), \mathbf{x}\right) \in \mathcal{Z}. \quad (2.43)$$

Here, $\psi(\cdot)$ is the *leverage-feasible activation function*. For $\mathbf{o} = (o_1, \dots, o_{N_a+1})^\top \in \mathbb{R}^{N_a+1}$, and $\mathbf{p} = \psi(\mathbf{o}, \mathbf{x})$, $\psi(\cdot) : (\mathbf{o}, \mathbf{x}) \in \mathbb{R}^{N_a+1} \times \mathcal{X} \mapsto \mathcal{Z}$ is defined by

$$\mathbf{p} = \psi(\mathbf{o}, \mathbf{x}) = \begin{cases} \begin{cases} l = p_{max} \cdot \text{Sigmoid}(o_{N_a+1}), \\ p_i = l \cdot \frac{e^{o_i}}{\sum_{k=1}^{N_l} e^{o_k}}, i \in \{1, \dots, N_l\}, \\ p_i = (1-l) \cdot \frac{e^{o_i}}{\sum_{k=N_l+1}^{N_a} e^{o_k}}, i \in \{N_l+1, \dots, N_a\}, \end{cases} & \text{if } \mathbf{x} \in \mathcal{X}_1, \\ e_{N_l+1}, & \text{if } \mathbf{x} \in \mathcal{X}_2. \end{cases} \quad (2.44)$$

Recall that N_l is the number of long-only assets and p_{max} is the maximum leverage allowed. We show that the leverage-feasible activation function ψ has the following property.

Lemma 2.1. (*Decomposition of ψ*) The leverage-feasible function ψ defined in (2.44) has the function decomposition that

$$\psi(\mathbf{o}, \mathbf{x}) = \varphi(\zeta(\mathbf{o}), \mathbf{x}), \quad (2.45)$$

where

$$\begin{cases} \zeta : \mathbb{R}^{N_a+1} \mapsto \tilde{\mathcal{Z}}, \zeta(\mathbf{o}) = \left(\text{Softmax}\left((o_1, \dots, o_{N_l})\right), \text{Softmax}\left((o_{N_l+1}, \dots, o_{N_a})\right), p_{max} \cdot \text{Sigmoid}(o_{N_a+1}) \right)^\top, \\ \varphi : \tilde{\mathcal{Z}} \times \mathcal{X} \mapsto \mathcal{Z}, \varphi(z) = \left(z_{N_a+1} \cdot (z_1, \dots, z_{N_l}), (1 - z_{N_a+1}) \cdot (z_{N_l+1}, \dots, z_{N_a}) \right)^\top \cdot \mathbf{1}_{\mathbf{x} \in \mathcal{X}_1} + e_{N_l+1} \cdot \mathbf{1}_{\mathbf{x} \in \mathcal{X}_2}, \end{cases} \quad (2.46)$$

and

$$\tilde{\mathcal{Z}} = \left\{ z \in \mathbb{R}^{N_a+1}, \sum_{i=1}^{N_l} z_i = 1, \sum_{i=N_l+1}^{N_a} z_i = 1, z_{N_a+1} \leq p_{max}, z_i \geq 0, \forall i \right\}. \quad (2.47)$$

Proof. This is easily verifiable by definition of ψ in (2.44). \square

Remark 2.6. (Remark on Lemma 2.1) The leverage-feasible activation function ψ corresponds to a two-step decision process described by ζ and φ . Intuitively, ζ first determines the internal allocations within long-only assets and shortable assets, as well as the total leverage. Then, φ converts the internal allocations and total leverage into final allocation fractions, which depend on the wealth of the active portfolio.

¹¹For $k = K + 1$, define $N_h^{(K+1)} = N_a + 1$ so $\boldsymbol{\theta}^{(K+1)}$ is well-defined.

With the LFNN model outlined above, the parameterized optimization problem (2.40) becomes an unconstrained optimization problem. Specifically, we present the following theorem regarding the feasibility domain \mathcal{Z}_θ associated with the LFNN model (2.43).

Theorem 2.2. (*Unconstrained feasibility domain*) *The feasibility domain \mathcal{Z}_θ defined in (2.41) associated with the LFNN model (2.43) is \mathbb{R}^{N_θ} .*

Proof. See Appendix B.1. □

Following Theorem 2.2, the constrained optimization problem (2.7) can be transformed into the following unconstrained optimization problem

$$\text{(Unconstrained parameterized problem): } \inf_{\theta \in \mathbb{R}^{N_\theta}} \mathbb{E}_{f(\cdot; \theta)}^{(t_0, w_0)} [F(\mathcal{W}_\theta, \hat{\mathcal{W}}_{\hat{\mathcal{P}}})]. \quad (2.48)$$

2.8 Mathematical justification for LFNN approach

By approximating the feasible control with a parameterized LFNN model, we have shown that the original constrained optimization problem is transformed into an unconstrained optimization problem, which is computationally more implementable.

However, an important question remains: is the solution to the parameterized unconstrained optimization problem (2.48) capable of yielding the optimal control of the original problem (2.7)? In other words, suppose θ^* is the solution to (2.48), can $f(\cdot; \theta^*)$ approximate solution to (2.7) with desired accuracy?

In this section, we prove that under benign assumptions and appropriate choices of the hyperparameter of the LFNN model (2.43), solving the unconstrained problem (2.48) provides an arbitrarily close approximation the original problem (2.7). We start by establishing the following lemma.

Lemma 2.2. (*Structure of feasible control*) *Any feasible control function $p : \mathcal{X} \mapsto \mathcal{Z}$, where \mathcal{Z} is defined in (2.38), has the function decomposition*

$$p(x) = \varphi(\omega(x), x), \quad (2.49)$$

where $\varphi : \tilde{\mathcal{Z}} \times \mathcal{X} \mapsto \mathcal{Z}$ is defined in (2.46) and $\omega : \mathcal{X} \mapsto \tilde{\mathcal{Z}}$.

Proof. See Appendix B.2. □

Next, we propose the following benign assumptions on the state space and the optimal control.

Assumption 2.7. (*Assumption on state space and optimal control*)

- (i) *The space \mathcal{X} of state variables is a compact set.*
- (ii) *Following Lemma 2.2, the optimal control $p^* : \mathcal{X} \mapsto \mathcal{Z}$ has the decomposition $p^*(x) = \varphi(\omega^*(x), x)$ for some $\omega^* : \mathcal{X} \mapsto \tilde{\mathcal{Z}}$. We assume $\omega^* \in C(\mathcal{X}, \tilde{\mathcal{Z}})$, where $C(\mathcal{X}, \tilde{\mathcal{Z}})$ denotes the set of continuous mappings from \mathcal{X} to $\tilde{\mathcal{Z}}$.*

Remark 2.7. (Remark on Assumption 2.7) In our particular problem of outperforming a benchmark portfolio, the state variable vector is $X(t) = (t, W(t), \hat{W}(t))^\top \in \mathcal{X}$ where $t \in [0, T]$. In this case, assumption (i) is equivalent to the assumption that the wealth of the active portfolio and benchmark portfolio is bounded, i.e. $\mathcal{X} = [0, T] \times [w_{min}, w_{max}] \times [w_{min}, \hat{w}_{max}]$, where w_{min}, w_{max} and $\hat{w}_{min}, \hat{w}_{max}$ are the respective wealth bounds for the portfolios. Intuitively, assumption (ii) states that the decision process for the optimal control to obtain the internal allocation fractions within the long-only assets, shortable assets, and the total leverage is a continuous function. This is a natural extension of the long-only case, in which it is commonly assumed that the allocation within long-only assets is a continuous function of state variables.

Finally, we present the following theorem.

Theorem 2.3. (*Approximation of optimal control*) Following Assumption 2.7, $\forall \epsilon > 0$, there exists $N_h \in \mathbb{N}$, and $\boldsymbol{\theta} \in \mathbb{R}^{N_\theta}$ such that the corresponding LFNN model $f(\cdot; \boldsymbol{\theta})$ described in (2.43) satisfies the following:

$$\sup_{x \in \mathcal{X}} \|f(x; \boldsymbol{\theta}) - p^*(x)\| < \epsilon. \quad (2.50)$$

Proof. See Appendix B.2. □

Theorem 2.3 shows that given any arbitrarily small tolerance $\epsilon > 0$, there exists a suitable choice of the hyperparameter of the LFNN model (e.g. the number of hidden layers and nodes), and a parameter vector $\boldsymbol{\theta}$, such that the corresponding parameterized LFNN function is within this tolerance of the optimal control function.¹² In other words, with a large enough LFNN model (in terms of the number of hidden nodes), solving the unconstrained parameterized problem (2.48) approximately solves the original optimization problem (2.7) with any required precision.

Remark 2.8. (*Empirical evidence of approximation*) In practice, we find that a small neural network structure with one single hidden layer and only 10 hidden nodes achieves excellent approximation performance. In particular, in a numerical experiment with simulated data, we compare the LFNN model with the approximate form of the closed-form solution derived in Section 2.4, and find that the LFNN model mimics the closed-form solution very well. This provides further empirical evidence that supports Theorem 2.3. Additional details can be found in Appendix C.

2.9 Training LFNN

Since the numerical experiments involve the solution and evaluation of the optimal parameters $\boldsymbol{\theta}^*$ of the LFNN model (2.43) in problem (2.48), we briefly review how the parameters are computed in experiments.

In numerical experiments, the expectation in (2.48) is approximated by using a finite set of samples of the set $\mathbf{Y} = \{Y^{(j)} : j = 1, \dots, N_d\}$, where N_d is the number of samples, and $Y^{(j)}$ represents a time series sample of *joint* asset return observations $R_i(t)$, $i \in \{1, \dots, N_a\}$, observed at $t \in \mathcal{T}$.¹³ Mathematically, problem (2.48) is approximated by

$$\inf_{\boldsymbol{\theta} \in \mathbb{R}^{N_\theta}} \left\{ \frac{1}{N_d} \sum_{j=1}^{N_d} F \left(\mathcal{W}_{\boldsymbol{\theta}}^{(j)}, \hat{\mathcal{W}}_{\hat{\mathcal{P}}}^{(j)} \right) \right\}. \quad (2.51)$$

Here $\mathcal{W}_{\boldsymbol{\theta}}^{(j)}$ is the wealth trajectory of the active portfolio following the LFNN parameterized by $\boldsymbol{\theta}$, and $\hat{\mathcal{W}}^{(j)}$ is the wealth trajectory of the benchmark portfolio following the benchmark strategy $\hat{\mathcal{P}}$, both evaluated on $Y^{(j)}$, the j -th time series sample.

We use a shallow neural network model, specifically, an LFNN model with one single hidden layer with 10 hidden nodes, i.e., $K = 1$ and $N_h^{(1)} = 10$. We use the 3-tuple vector $(t, W_{\boldsymbol{\theta}}(t), \hat{W}(t))^{\top}$ as the input (feature) to the LFNN network. At $t \in [t_0, T]$, $W_{\boldsymbol{\theta}}(t)$ is the wealth of the active portfolio of the strategy that follows the LFNN model parameterized by $\boldsymbol{\theta}$, and $\hat{W}(t)$ is the wealth of the benchmark portfolio.

Then, the optimal parameter $\boldsymbol{\theta}^*$ can be numerically obtained by solving problem (2.51) using standard optimization algorithms such as ADAM (Kingma and Ba, 2014). This process is commonly referred to as “training” of the neural network model, and \mathbf{Y} is often referred to as the training data set (Goodfellow et al., 2016).

Once $\boldsymbol{\theta}^*$ is numerically obtained, the resulting optimal strategy $f(\cdot; \boldsymbol{\theta}^*)$ is evaluated on a separate “testing” data set \mathbf{Y}^{test} , which contains a different set of samples generated from either the same distribution of the training process or a different process (depending on experiment purposes) so that the “out-of-sample” performance of $f(\cdot; \boldsymbol{\theta}^*)$ is assessed.

¹²The distance is defined in (2.50), i.e. the supremum of the pointwise distance over the extended state space \mathcal{X} .

¹³Note that the corresponding set of asset prices can be easily inferred from the set of asset returns, or vice versa.

3 Numerical experiments

In this section, we present a case study that explores optimal asset allocation during high-inflation periods using the LFNN model through numerical experiments. To conduct our analysis, we need data specifically from high inflation periods. Such data can be acquired using parametric modeling or non-parametric sample generation methods. It is important to note that our LFNN approach is agnostic to the choice of data modeling methods. While there is no universally accepted method for identifying or modeling high inflation regimes, for the purpose of this demonstration, we employ a simple filtering technique to identify inflation regime data and generate the required samples for training the LFNN.

3.1 Filtering historical inflation regimes

We use the U.S. CPI index and monthly data from the Center for Research in Security Prices (CRSP) over the 1926:1-2022:1 period.¹⁴¹⁵ We select high-inflation periods as determined by the CPI index using the following filtering procedure. Using a moving window of k months, we determine the cumulative CPI index log return (annualized) in this window. If the cumulative annualized CPI index log return is greater than a cutoff, then all the months in the window are flagged as part of a high-inflation regime. Note that some months may appear in more than one moving window. Any months which do not meet this criterion are considered to be in low-inflation regimes. Pseudo-code can be found in Appendix D in Ni et al. (2023).

Since the average annual inflation over the period 1926:1-2022:1 was 2.9%, and Federal Reserve policymakers have been targeting the inflation rate of 2% over the long run to achieve maximum employment and price stability (The Federal Reserve, 2011), we use a cutoff of 5% as the threshold for high inflation. In addition, we use the moving window size of 5 years (see Appendix D.2 in Ni et al. (2023) for more discussion). This uncovers two inflation regimes: 1940:8-1951:7 and 1968:9-1985:10, which correspond to well-known market shocks (i.e. the second world war, and price controls; the oil price shocks and stagflation of the seventies).

Table 3.1 shows the average annual inflation over the two regimes identified from our filter.

Time Period	Average Annualized Inflation
1940:8-1951:7	.0564
1968:9-1985:10	.0661

Table 3.1: Inflation regimes determined using a five-year moving window with a cutoff inflation rate of 0.05.

For possible investment assets, we consider the 30-day U.S. T-bill index (CRSP designation “t30ind”), a constant maturity 10-year U.S. treasury index,¹⁶ and the cap-weighted stock index (CapWt) and the equal-weighted stock index (EqWt), also from CRSP.¹⁷ All of these various indexes are adjusted for inflation by using the U.S. CPI index.

We find that the equal-weighted stock index has a higher average return and higher volatility than the cap-weighted stock index. In addition, we find that the 30-day T-bill index has a similar average return as the

¹⁴The date convention is that, for example, 1926:1 refers to January 1, 1926.

¹⁵More specifically, results presented here were calculated based on data from Historical Indexes, ©2022 Center for Research in Security Prices (CRSP), The University of Chicago Booth School of Business. Wharton Research Data Services (WRDS) was used in preparing this article. This service and the data available thereon constitute valuable intellectual property and trade secrets of WRDS and/or its third-party suppliers.

¹⁶The 10-year treasury index was generated from monthly returns from CRSP back to 1941 (CRSP designation “b10ind”). The data for 1926-1941 are interpolated from annual returns in Homer and Sylla (1996). The 10-year treasury index is constructed by (a) buying a 10-year treasury at the start of each month, (b) collecting interest during the month, and then (c) selling the treasury at the end of the month. We repeat the process at the start of the next month. The gains in the index then reflect both interest and capital gains and losses.

¹⁷The capitalization-weighted total returns have the CRSP designation “vwretd”, and the equal-weighted total returns have the CRSP designation “ewretd”.

10-year T-bond index, but much lower volatility, see Appendix D.3 in Ni et al. (2023) for more details. This indicates that the T-bill index is the better choice of a defensive asset during high inflation. Subsequently, we consider the equal-weighted stock index, the cap-weighted stock index, and the 30-day T-bill index.

3.2 Bootstrap resampling

Once we have obtained the filtered historical high-inflation data series from Section 3.1, it becomes necessary to generate training and testing data sets from the original time series data. While one common approach is to assume and fit a parametric model to the underlying data, it is important to acknowledge the limitations associated with this choice.

Parametric models have several drawbacks, including the difficulty of accurately estimating their parameters (Black, 1993). Even for a simple geometric Brownian motion (GBM) model, accurately estimating the drift rate can be challenging and prone to errors, requiring a long historical period of data coverage (Brigo et al., 2008). More complex models, such as the jump-diffusion model (2.15), introduce additional components to the stochastic model, which necessitates the estimation of extra parameters. Furthermore, parametric models inherently make assumptions about the true stochastic model for asset prices, which can be subject to debate.

Acknowledging the above limitations of parametric market data models, we turn to the alternative non-parametric method of bootstrap resampling as a data-generating process for numerical experiments. Unlike parametric models, non-parametric models such as bootstrap resampling do not make assumptions about the parametric form of the asset price dynamics. Intuitively speaking, the bootstrap resampling method randomly chooses data points from the historical time series data and reassembles them into new paths of time series data. The bootstrap was initially proposed as a statistical method for estimating the sampling distribution of statistics (Efron, 1992). We use it as a data-generating procedure, as the philosophy behind bootstrap resampling is consistent with the idea that *“history does not repeat, but it rhymes.”* The bootstrap resampling provides an empirical distribution, which seems to be the least prejudiced estimate possible of the underlying distribution of the data-generating process. We also note that bootstrap resampling is widely adopted by practitioners (Alizadeh and Nomikos, 2007; Cogneau and Zakamouline, 2013; Dichtl et al., 2016; Scott and Cavaglia, 2017; Shahzad et al., 2019; Cavaglia et al., 2022; Simonian and Martirosyan, 2022) as well as academics (Anarkulova et al., 2022).

Specifically, we choose to use the stationary block bootstrap resampling method (Politis and Romano, 1994). See Appendix E.1 in Ni et al. (2023) for the detailed pseudo-code for bootstrap resampling. Compared to the traditional bootstrap method, the block bootstrap technique preserves the local dependency of data within blocks. Furthermore, the stationary block bootstrap uses random block sizes which preserves the stationarity of the original time series data. An important parameter is the expected block size, which, informally, is a measure of serial correlation in the return data. A challenge in using block bootstrap resampling is the need to choose a single block size for multiple underlying time series data so that the bootstrapped data entries for different assets are synchronized in time. Subsequently, we use the expected block size of 6 months for all time series data. However, we have compared different numerical experiments using a range of block sizes, including i.i.d. assumptions (i.e. expected block size equal to one month), and find that the results are relatively insensitive to block size, as discussed in more detail in Appendix E.2 in Ni et al. (2023).

Typically, the bootstrap technique resamples from data sourced from one contiguous segment of historical periods. However, the moving-window filtering algorithm has identified two non-contiguous historical inflation regimes. To apply the bootstrap method, there are two intuitive possibilities: 1) concatenate the two historical inflation regimes first, then bootstrap from the concatenated combined series, or 2) bootstrap within each regime (i.e., using circular block bootstrap resampling within each regime), then combine the bootstrapped resampled data points. We have experimented with both methods and find that the difference is minimal (see Appendix E.3 in Ni et al. (2023)). In this article, we adopt the first method, i.e., we concatenate the historical regimes first, then bootstrap from the combined series. This method is also adopted by Anarkulova et al. (2022), where stock returns from different countries are concatenated and the bootstrap is applied to the combined data.

3.3 A case study on high inflation investment: a 4-asset scenario

3.3.1 Experiment setup

In this section, we conduct a case study on optimal asset allocation during a consistent high-inflation regime. The details of the investment specification are given in Table 3.2. Briefly, the active portfolio and benchmark portfolio begin with the same initial wealth of 100 at $t_0 = 0$. Both portfolios are rebalanced monthly. The investment horizon is 10 years, and there is an annual cash injection of 10 for both portfolios, evenly divided over 12 months.

We consider an empirical case in which we allow the manager to allocate between four investment assets: the equal-weighted stock index, the cap-weighted stock index, the 30-day U.S. T-bill index, and the 10-year U.S. T-bond index. We assume that the stock indexes and the 10-year T-bond index are long-only assets. The manager can short the T-bill index to take leverage and invest in the long-only assets (with maximum total leverage of 1.3). In this experiment, we assume the borrowing premium rate is zero. Essentially, we assume that the manager can borrow short-term funding to take leverage at the same cost as the treasury bill. This may be a reasonable assumption for sovereign wealth funds, as they are state-owned and enjoy a high credit rating. We remark that the borrowing premium does not really affect the results significantly.¹⁸ The annual outperformance target β is set to be 2% (i.e. 200 bps per year).

It is worth noting that we choose the benchmark portfolio to be a fixed-mix portfolio that maintains a 70% weight in the equal-weighted stock index and 30% in the 30-day U.S. T-bill index. We select this fixed-mix portfolio as the benchmark based on our observation that the equal-weighted stock index shows superior performance compared to the cap-weighted stock index during high-inflation environments. Surprisingly, when analyzing bootstrap resampled data from the historical inflation regimes, we find that the fixed-mix portfolio consisting of 70% in the equal-weighted stock index and 30% in the 30-day U.S. T-bill index partially stochastically dominates the fixed-mix portfolio consisting of 70% in the cap-weighted stock index and 30% in the 30-day U.S. T-bill index. For more detailed information, interested readers can refer to Appendix F in Ni et al. (2023).

Investment horizon T (years)	10
Equity market indexes	CRSP cap-weighted/equal-weighted index (real)
Bond index	CRSP 30-day/10-year U.S. treasury index (real)
Index samples for bootstrap	Concatenated 1940:8-1951:7, 1968:9-1985:10
Initial portfolio wealth/annual cash injection	100/10
Rebalancing frequency	Monthly
Maximum leverage	1.3
Outperformance target rate β	2%

Table 3.2: Investment scenario.

As discussed in the previous section, we use the stationary bootstrap resampling algorithm to generate a training data set \mathbf{Y} and a testing data set \mathbf{Y}^{test} (both with 10,000 resampled paths) from the concatenated index samples from two historical inflation regimes: 1940:8-1951:7 and 1968:9-1985:10, using an expected blocksize of 6 months. The testing data set \mathbf{Y}^{test} is generated using a different random seed as the training data set \mathbf{Y} , and thus the probability of seeing the same sample in \mathbf{Y} and \mathbf{Y}^{test} is near zero (see Ni et al. (2022) for proof).

We remark that in this experiment, we train the LFNN model (2.43) on \mathbf{Y} under the discrete-time CS objective (2.14), instead of the CD objective (2.12). As discussed in Section 2.3, the CS objective function only penalizes underperformance relative to the elevated target. Numerical comparisons of the two objective functions suggest that the CS objective function indeed yields more favorable investment results than the CD objective (see Appendix H in Ni et al. (2023)).

¹⁸See Appendix I in Ni et al. (2023) for a more detailed discussion.

In this section, unless stated otherwise, all the results presented are testing results.

3.4 Experiment results

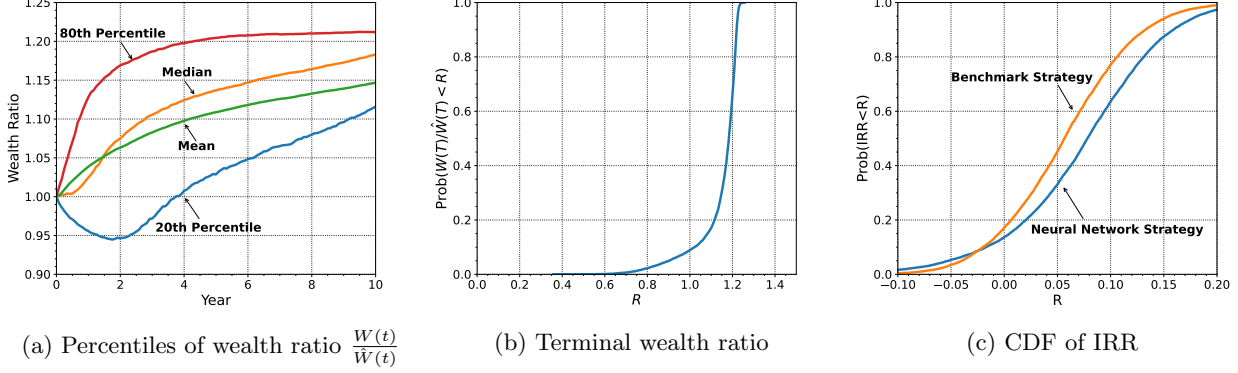


Figure 3.1: Percentiles of wealth ratio $\frac{W(t)}{\hat{W}(t)}$ and CDF of terminal wealth ratio $\frac{W(T)}{\hat{W}(T)}$ and internal rate of return (IRR). Results are based on the evaluation of the learned neural network model on \mathbf{Y}^{test} .

Strategy	Median[W_T]	E[W_T]	std[W_T]	5th Percentile	Median IRR (annual)
Neural network	364.2	403.4	211.8	136.3	0.078
Benchmark	308.5	342.9	165.0	149.0	0.056

Table 3.3: Statistics of strategies. Results are based on the evaluation results on the testing data set.

The analysis of Figure 3.1a reveals that the neural network strategy (the strategy following the training LFNN model) consistently outperforms the benchmark strategy in terms of the wealth ratio $W(t)/\hat{W}(t)$. Over time, both the mean and median wealth ratios demonstrate a smooth and consistent increase. Regarding tail performance (20th percentile), the neural network strategy initially falls behind the benchmark but gradually recovers and ultimately achieves 10% greater wealth at the terminal time. This observation indicates that the neural network strategy effectively manages tail risk.

An additional metric that holds significant interest for managers is the distribution of the terminal wealth ratio $\frac{W(T)}{\hat{W}(T)}$. This metric examines the relative performance of the strategies at the end of the investment period. Figure 3.1b illustrates that there is a greater than 90% chance that the neural network strategy outperforms the benchmark strategy in terms of terminal wealth. This outcome is particularly noteworthy as the objective function (2.14) does not directly target the terminal wealth ratio.

Given the constant cash injections in the portfolios, it is appropriate to employ the internal rate of return (IRR) as a measure of the portfolio’s annualized performance. Figure 3.1c demonstrates that the neural network strategy has a more than 90% chance of producing a higher IRR. Furthermore, the median IRR of the neural network strategy exceeds that of the benchmark strategy by slightly over 2%, aligning with the chosen target outperformance rate of $\beta = 0.02$. This indicates that the neural network model consistently achieves the desired target performance across most outcomes.

The results from Table 3.3 indicate that the 5th percentile of the terminal wealth for the neural network strategy is lower than that of the benchmark strategy. This suggests that in some scenarios, particularly during persistent bear markets when stocks perform poorly, the neural network strategy may experience lower terminal wealth compared to the benchmark strategy. The neural network strategy takes on more

risk by allocating a higher fraction of wealth to the equal-weighted stock index, which is considered a riskier asset, in comparison to the benchmark portfolio.

It’s important to note, however, that these scenarios occur with low probability. As depicted in Figure 3.1b, the neural network strategy exhibits a significantly high probability of outperforming the benchmark in terms of terminal wealth, exceeding 90%. This implies that while there might be instances where the neural network strategy suffers relative to the benchmark, the overall performance is consistently strong, resulting in a high likelihood of achieving superior terminal value.

To gain insight into the strong performance of the neural network strategy, we further examine its allocation profile.

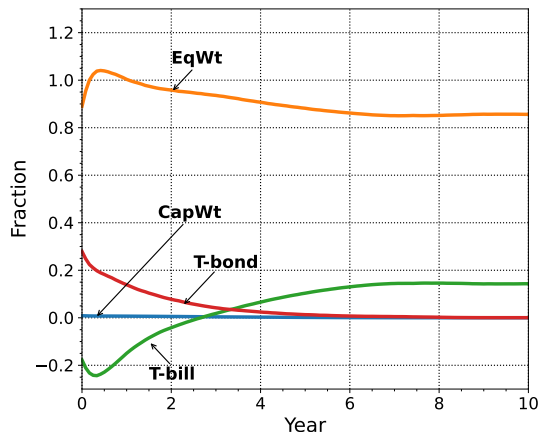


Figure 3.2: Mean allocation fraction over time, evaluated on \mathbf{Y}^{test}

We begin by examining the mean allocation fraction for the four assets over time, as depicted in Figure 3.2. The first noteworthy observation from Figure 3.2 is that, on average, the neural network strategy does not allocate wealth to the cap-weighted stock index. Initially, this might appear surprising; however, it aligns with historical data indicating significantly higher real returns for the equal-weighted stock index during periods of high inflation (refer to Appendix D.3 in Ni et al. (2023)). Given that the objective is to outperform a benchmark heavily invested in the equal-weighted index (70%), it is logical to avoid allocating wealth to a comparatively weaker index in the active portfolio.

The second observation derived from Figure 3.2 pertains to the evolution of mean bond allocation fractions. Initially, the neural network strategy shorts the 30-day T-bill index and assumes some leverage while heavily investing in the equal-weighted stock index during the first two years. This indicates a deliberate risk-taking approach early on to establish an advantage over the benchmark strategy. Subsequently, the allocation to the 10-year T-bond decreases, coinciding with the reduction in the allocation to the equal-weighted index. This suggests that the initial allocation to the T-bond was primarily for leveraging purposes, with the 10-year bond being the only defensive asset available. As leverage is no longer used in later years, the neural network strategy favors the T-bill over the 10-year bond.

Overall, despite the gradual decrease in stock allocation over time, the neural network strategy maintains an average allocation of more than 80% to the equal-weighted stock index. This is expected, as outperforming an aggressive benchmark with a 70% allocation to the equal-weighted stock index necessitates assuming higher levels of risk. Despite the higher allocation to riskier assets, the neural network strategy consistently delivers strong results compared to the benchmark strategy, as illustrated in Figure 3.1.

Lastly, it is worth noting that the neural network strategy, trained under high-inflation regimes, exhibits remarkable performance on low-inflation testing datasets. This unexpected outcome highlights the robustness of the strategy. For further discussion on this topic, interested readers can refer to Appendix J in Ni et al.

(2023).

4 Conclusion

In this paper, our primary objective is to propose a framework that generates optimal dynamic allocation strategies under leverage constraints in order to outperform a benchmark during high inflation regimes. Imposing leverage-constraint in multi-period asset allocation is consistent with the practice in large sovereign wealth funds, which often have exposures to alternative assets. Our proposed framework efficiently solves high-dimensional optimal control problems, accommodating diverse objective functions, constraints, and data sources.

We begin by assuming that both asset prices follow jump-diffusion models. Under this assumption, we derive a closed-form solution for a two-asset case using the cumulative tracking difference (CD) objective function. However, to obtain this closed-form solution, we need to make additional unrealistic assumptions such as continuous rebalancing, unlimited leverage, and continued trading in insolvency. Despite these assumptions, the closed-form solution provides valuable insights into the optimal control behavior. Notably, to track the elevated target, the optimal control needs to aim higher than the target when making allocation decisions.

To overcome the limitations of unrealistic assumptions and derive a more practical solution, we introduce a novel leverage-feasible neural network (LFNN) model. The LFNN model approximates the optimal control directly, eliminating the need for high-dimensional approximations of conditional expectations required in dynamic programming approaches. Additionally, the LFNN model converts the leverage-constrained optimization problem into an unconstrained optimization problem. Importantly, we justify the validity of the LFNN approach by mathematically proving that the solution to the parameterized unconstrained optimization problem can approximate the solution to the original constrained optimization problem with arbitrary precision.

To illustrate the effectiveness of our proposed approach, we conduct a case study on optimal asset allocation during high-inflation regimes. We apply the LFNN model to bootstrap resampled data from filtered historical high-inflation data. In our numerical experiment, we consider an investment case with four assets in high inflation regimes. The results consistently demonstrate that the neural network strategy outperforms the benchmark strategy throughout the investment period. Specifically, the neural network strategy achieves a 2% higher median Internal Rate of Return (IRR) compared to the benchmark strategy and yields a higher terminal wealth with more than a 90% probability. The allocation strategy derived from the LFNN model suggests that managers should favor the equal-weighted stock index over the cap-weighted stock index and short-term bonds over long-term bonds during high-inflation periods.

5 Acknowledgements

Forsyth’s work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) grant RGPIN-2017-03760. Li’s work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) grant RGPIN-2020-04331.

6 Conflicts of interest

The authors have no conflicts of interest to report.

A Technical details of closed-form solution

A.1 Proof of Theorem (2.17)

At any state $(t, w, \hat{w}) \in [t_0, T] \times \mathbb{R}^2$, define the value function $V(w, \hat{w}, t)$ to the CD problem (2.11) as

$$V(t, w, \hat{w}, \hat{\rho}) = \inf_{\mathbf{p}} \left\{ \mathbb{E}_{\mathbf{p}} \left[\int_t^T (W(s) - e^{\beta s} \hat{W}(s))^2 ds \mid W(t) = w, \hat{W}(t) = \hat{w} \right] \right\}. \quad (\text{A.1})$$

By the dynamic programming principle, we have

$$V(t, w, \hat{w}, \hat{\rho}) = \inf_{\mathbf{p}} \left\{ \mathbb{E}_{\mathbf{p}} \left[\left(V(t+\Delta t, W(t+\Delta t), \hat{W}(t+\Delta t), \hat{\rho}) + \int_t^{t+\Delta t} (W(s) - e^{\beta s} \hat{W}(s))^2 ds \right) \mid W(t) = w, \hat{W}(t) = \hat{w} \right] \right\} \quad (\text{A.2})$$

Rearrange equation (A.2) to obtain

$$\inf_{\mathbf{p}} \left\{ \mathbb{E}_{\mathbf{p}} \left[\left(dV(t, w, \hat{w}, \hat{\rho}) + \int_t^{t+\Delta t} (W(s) - e^{\beta s} \hat{W}(s))^2 ds \right) \mid W(t) = w, \hat{W}(t) = \hat{w} \right] \right\} = 0 \quad (\text{A.3})$$

Then, apply Itô's lemma with jumps (Cont et al., 2011), substitute dW and $d\hat{W}$ terms with (2.16), and take limits as $\Delta t \downarrow 0$, we obtain (2.17).

The above results merely serve as an intuitive guide to obtain (2.17). The formal proof of (2.17) proceeds by using a suitably smooth test function, see for example (Øksendal and Sulem, 2007).

A.2 Proof of results for CD-optimal control

In Section 2.4, we emphasized the dependence of B and D (defined in (2.23) and (2.22)) on parameters β and c for understanding the optimal control function. As β and c are fixed parameters, in this proof, we omit the dependence of B and D on them for notational simplicity.

The quadratic source term $(w - e^{\beta t} \hat{w})^2$ in Theorem (2.17) suggests the following *ansatz* for the value function V in Theorem 2.17 of the form

$$V(t, w, \hat{w}) = A(t)w^2 + B(t)w + C(t) + \hat{A}(t)\hat{w}^2 + \hat{B}(t)\hat{w} + D(t)w\hat{w}, \quad (\text{A.4})$$

where $A, B, C, \hat{A}, \hat{B}, D$ are unknown deterministic functions of time t . If (A.4) is correct, then the pointwise infimum in (2.17) is attained by p^* satisfying the relationship

$$\left(w \cdot \frac{\partial^2 V}{\partial w^2} \right) \cdot p^* = -\frac{1}{\gamma} \left((\mu_1 - \mu_2) \cdot \frac{\partial V}{\partial w} + (\hat{\rho}\gamma + \theta) \cdot \hat{w} \cdot \frac{\partial^2 V}{\partial w \hat{w}} + \theta \cdot w \cdot \frac{\partial^2 V}{\partial w^2} \right), \quad (\text{A.5})$$

assuming $A(t) > 0$. Here γ and θ are defined in (2.19). (A.4) implies that the relevant partial derivatives of V are of the form

$$\frac{\partial^2 V}{\partial w^2} = 2A(t), \quad \frac{\partial V}{\partial w} = 2A(t)w + B(t) + D(t)\hat{w}, \quad \frac{\partial^2 V}{\partial w \hat{w}} = D(t). \quad (\text{A.6})$$

Substituting (A.6) into (A.5), the optimal control p^* obtained is in the form of (2.20), where h and g are given by (2.21). Then, it only remains to determine the functions A, B, D . Substituting (A.5) into PIDE (2.17), we can obtain the following ordinary differential equations (ODE) for A, B, D ,

$$\begin{cases} \frac{dA(t)}{dt} = -\left(2\mu_2 - \eta\right)A(t) - 1, & A(T) = 0, \\ \frac{dD(t)}{dt} = -\left(2\mu_2 - \eta\right)D(t) + 2e^{\beta t}, & D(T) = 0, \\ \frac{dB(t)}{dt} = -(\mu_2 - \phi)B(t) - 2cA(t) - cD(t), & B(T) = 0, \end{cases} \quad (\text{A.7})$$

Solving the ODE system gives us the A, B, D defined in (2.22) and (2.23). We also note that $A(t) > 0$, thus completing the proof.

A.3 Proof of Corollary (2.1) and (2.2)

van Staden et al. (2022) derive the CD-optimal control under the assumption that the stock price follows the double-exponential jump-diffusion model and the bond is risk-free with the bond price $B(t)$ following

$$\frac{dB(t)}{B(t)} = r. \quad (\text{A.8})$$

Under such, assumptions, van Staden et al. (2022) shows that the CD-optimal control can be expressed in a similar form as in (2.20) with g and h functions. The g and h functions satisfy the same properties as in Corollary (2.1) and (2.2). Despite the fact that we assume the bond price follows a jump-diffusion model, the proof of Corollary (2.1) and (2.2) follows similar steps as the proof in van Staden et al. (2022).

B Technical details of LFNN model

B.1 Proof of Theorem 2.2

Theorem 2.2. (Unconstrained feasibility domain) The feasibility domain \mathcal{Z}_θ defined in (2.41) associated with the LFNN model (2.43) is \mathbb{R}^{N_θ} .

Proof. First, it is obvious that $\mathcal{Z}_\theta \subseteq \mathbb{R}^{N_\theta}$ by definition of (2.41). Next, we show that $\mathbb{R}^{N_\theta} \subseteq \mathcal{Z}_\theta$. To prove this, we need to show that for any $\theta \in \mathbb{R}^{N_\theta}$,

$$f(x; \theta) = p \in \begin{cases} \mathcal{Z}_1, & \text{if } x \in \mathcal{X}_1, \\ \mathcal{Z}_2, & \text{if } x \in \mathcal{X}_2, \end{cases} \quad \forall x \in \mathcal{X}. \quad (\text{B.1})$$

Here f is the LFNN function defined in (2.43), $p = (p_1, \dots, p_{N_a})^\top \in \mathbb{R}^{N_a}$ is the output of the LFNN model that represents the wealth allocation to the assets, \mathcal{Z} is the feasibility domain defined in (2.37), and $x = (t, W(t), \hat{W}(t))^\top \in \mathcal{X}$ is a feature vector. To prove (B.1), we verify the two scenarios ($x \in \mathcal{X}_1$ and $x \in \mathcal{X}_2$) separately.

When $x \in \mathcal{X}_2$, it is easily verifiable that $p = e_{N_l+1}$ via the definition of the leverage-feasible activation function (2.44).

Next, we verify that when $x \in \mathcal{X}_1$, $p \in \mathcal{Z}_1$. To prove this, we need to show that constraints of (2.29)-(2.32) are satisfied when $x \in \mathcal{X}_1$.

By definition of (2.44), it is obvious that the long-only constraint (2.29) holds for long-only assets.

It is also easy to verify that the summation constraint (2.30) is satisfied. This can be observed after the fact that

$$\sum_{i=1}^{N_l} p_i = l, \quad \text{and} \quad \sum_{i=N_l+1}^{N_a} p_i = 1 - l. \quad (\text{B.2})$$

The maximum leverage constraint (2.31) is also satisfied, as

$$\sum_{i=1}^{N_l} p_i = l = p_{max} \cdot \text{Sigmoid}(-o_{N_a+1}) \leq p_{max}. \quad (\text{B.3})$$

Finally, the simultaneous shorting constraint (2.5) is satisfied. To see this, we examine the scenario when leverage occurs, i.e., $\sum_{i=1}^{N_l} p_i = l > 1$. Then, by definition from (2.44), we know

$$p_i = (1 - l) \cdot \frac{e^{o_i}}{\sum_{k=N_l+1}^{N_a} e^{o_k}} \leq 0, \quad \forall i \in \{N_l + 1, \dots, N_a\} \quad (\text{B.4})$$

From (B.4) it is clear that if $l \leq 1$, then $p_i \geq 0, \forall i$.

Therefore, for any $\theta \in \mathbb{R}^{N_\theta}$, (B.1) is satisfied. This implies $\mathbb{R}^{N_\theta} \subseteq \mathcal{Z}_\theta$. □

B.2 Proof of Lemma 2.2 and Theorem 2.3

Lemma 2.2. (Structure of feasible control) Any feasible control function $p : \mathcal{X} \mapsto \mathcal{Z}$, where \mathcal{Z} is defined in (2.38), has the function decomposition

$$p(x) = \overline{\varphi(\omega(x), x)}, \quad (\text{B.5})$$

where $\varphi : \tilde{\mathcal{Z}} \times \mathcal{X} \mapsto \mathcal{Z}$ is defined in (2.46), i.e.

$$\varphi(z) = \left(z_{N_a+1} \cdot (z_1, \dots, z_{N_l}), (1 - z_{N_a+1}) \cdot (z_{N_l+1}, \dots, z_{N_a}) \right)^\top \cdot \mathbf{1}_{x \in \mathcal{X}_1} + \mathbf{e}_{N_l+1} \cdot \mathbf{1}_{x \in \mathcal{X}_2}, \quad (\text{B.6})$$

and $\omega : \mathcal{X} \mapsto \tilde{\mathcal{Z}}$. Here

$$\tilde{\mathcal{Z}} = \left\{ z \in \mathbb{R}^{N_a+1}, \sum_{i=1}^{N_l} z_i = 1, \sum_{i=N_l+1}^{N_a} z_i = 1, z_{N_a+1} \leq p_{max}, z_i \geq 0, \forall i \right\}. \quad (\text{B.7})$$

Proof. We prove the lemma by existence.

Define ω as

$$\omega(x) = \begin{cases} \phi(p(x)), & \text{if } x \in \mathcal{X}_1, \\ \left(\frac{1}{N_l}, \dots, \frac{1}{N_l}, \frac{1}{N_a - N_l}, \dots, \frac{1}{N_a - N_l}, 0 \right)^\top, & \text{if } x \in \mathcal{X}_2, \end{cases} \quad (\text{B.8})$$

where for $\forall z = (z_1, \dots, z_{N_a})^\top \in \mathcal{Z}_1$, $y = \phi(z) \in \mathbb{R}^{N_a+1}$ is as defined below

$$\phi(z) \equiv y = \begin{cases} \begin{cases} y_i = \frac{z_i}{\sum_{j=1}^{N_l} z_j}, & i \in \{1, \dots, N_l\}, \\ y_i = \frac{z_i}{1 - \sum_{j=1}^{N_l} z_j}, & i \in \{N_l, \dots, N_a\}, \\ y_{N_a+1} = \sum_{j=1}^{N_l} z_j, \end{cases} & \text{if } \sum_{i=1}^{N_l} z_i \in (0, 1) \cup (1, p_{max}], \\ \begin{cases} y_i = z_i, & i \in \{1, \dots, N_l\}, \\ y_i = 1/(N_a - N_l), & i \in \{N_l, \dots, N_a\}, \\ y_{N_a+1} = 1, \end{cases} & \text{if } \sum_{i=1}^{N_l} z_i = 1, \\ \begin{cases} y_i = 0, & i \in \{1, \dots, N_l\}, \\ y_i = z_i, & i \in \{N_l, \dots, N_a\}, \\ y_{N_a+1} = 0, \end{cases} & \text{if } \sum_{i=1}^{N_l} z_i = 0, \end{cases} \quad (\text{B.9})$$

It can then be easily verified that $\omega : \mathcal{X} \mapsto \tilde{\mathcal{Z}}$, and that $p(x) = \overline{\varphi(\omega(x), x)}$. □

Lemma B.1. (Approximation of controls with a specific structure) Assume a control function $p : \mathcal{X} \mapsto \mathcal{Z}$ has the structure

$$p(x) = \Phi(\Omega(x), x), x \in \mathcal{X}, \quad (\text{B.10})$$

where \mathcal{X} is compact, $\Omega \in C(\mathcal{X}, \mathcal{Y})$, i.e. Ω is a continuous mapping from \mathcal{X} to \mathcal{Y} , and $\Phi : \mathcal{Y} \times \mathcal{X} \mapsto \mathcal{Z}$ is Lipschitz continuous on $\mathcal{Y} \times \mathcal{X}_i$, $\forall i = 1, \dots, n$, where $\{\mathcal{X}_i, i = 1, \dots, n\}$ is a partition of \mathcal{X} , i.e.

$$\begin{cases} \bigcup_{i=1}^m \mathcal{X}_i = \mathcal{X}, \\ \mathcal{X}_i \cap \mathcal{X}_j = \emptyset, \forall 1 \leq i, j \leq n. \end{cases} \quad (\text{B.11})$$

If $\exists m \in \mathbb{N}$ and $\Upsilon : \mathbb{R}^m \mapsto \mathcal{Y}$ such that

(i) Υ has a continuous right inverse on $\text{Im}(\Upsilon)$.

(ii) $\text{Im}(\Upsilon)$ is dense in \mathcal{Y} , then $\forall \epsilon > 0$.

(iii) $\partial \text{Im}(\Upsilon)$ is collared.

Then there exists a choice of N_h and $\boldsymbol{\theta}$ such that the fully connected feedforward neural network function $\tilde{f}(\cdot; \boldsymbol{\theta})$ defined in (2.42) satisfies

$$\sup_{x \in \mathcal{X}} \|\Phi(\Upsilon(\tilde{f}(x; \boldsymbol{\theta})), x) - p(x)\| < \epsilon. \quad (\text{B.12})$$

Proof. Let

$$L_\Phi = \max_{1 \leq i \leq n} L_i, \quad (\text{B.13})$$

where L_i is the Lipschitz constant for Φ on $\mathcal{Y} \times \mathcal{X}_i$.

Since $\Omega \in C(\mathcal{X})$ in compact \mathcal{X} , following Kratsios and Bilokopytov (2020), we know that $\forall \epsilon$, there exists $N_h \in \mathbb{N}$ and $\boldsymbol{\theta} \in \mathbb{R}^{N_\theta}$ such that the corresponding FNN $\tilde{f}(\cdot; \boldsymbol{\theta}) : \mathcal{X} \mapsto \mathbb{R}^m$ defined in (2.42) satisfies

$$\sup_{x \in \mathcal{X}} \|\Upsilon(\tilde{f}(x; \boldsymbol{\theta})) - \Omega(x)\| < \epsilon/L_\Phi, \quad (\text{B.14})$$

Then

$$\sup_{x \in \mathcal{X}} \|\Phi(\Upsilon(\tilde{f}(x; \boldsymbol{\theta})), x) - p(x)\| = \sup_{1 \leq i \leq n} \sup_{x \in \mathcal{X}_i} \|\Phi(\Upsilon(\tilde{f}(x; \boldsymbol{\theta})), x) - \Phi(\Omega(x), x)\| \quad (\text{B.15})$$

$$\leq \sup_{1 \leq i \leq n} \sup_{x \in \mathcal{X}_i} L_i \cdot (\|\Upsilon(\tilde{f}(x; \boldsymbol{\theta})) - \Omega(x)\|) \quad (\text{B.16})$$

$$< \sup_{1 \leq i \leq n} \frac{L_i}{L_\Phi} \epsilon \quad (\text{B.17})$$

$$\leq \epsilon. \quad (\text{B.18})$$

□

Remark B.1. (Remark on Lemma B.1) Normally, the universal approximation theorem only applies to the approximation of continuous functions defined on a compact set (Hornik, 1991). Lemma B.1 extends the universal approximation theorem to a broader class of functions that have the structure of (B.10). Furthermore, Lemma B.1 provides guidance on constructing neural network functions that handle stochastic constraints on controls which are usually difficult to address in stochastic optimal control problems. Consider the following example: the control $p : \mathcal{X} \mapsto \mathbb{R}^{N_a}$ has stochastic constraints such that $p(\boldsymbol{x}) \in [a(\boldsymbol{x}), b(\boldsymbol{x})]$ where $a, b : \mathcal{X} \mapsto \mathbb{R}^{N_a}$ are deterministic functions. This is a common setting in portfolio optimization problems in which allocation fractions to specific assets are subject to thresholds tied to the performance of the portfolio. With Lemma B.1, with a bit of engineering, one can easily construct a Φ so that the corresponding neural network satisfies the constraints naturally and be guaranteed that such a neural network can approximate the control well.

We then proceed to prove Theorem 2.3.

Theorem 2.3. (Approximation of optimal control) Following Assumption 2.7, $\forall \epsilon > 0$, there exists $N_h \in \mathbb{N}$, and $\boldsymbol{\theta} \in \mathbb{R}^{N_\theta}$ such that the corresponding LFNN model $f(\cdot; \boldsymbol{\theta})$ described in (2.43) satisfies the following:

$$\sup_{x \in \mathcal{X}} \|f(x; \boldsymbol{\theta}) - p^*(x)\| < \epsilon. \quad (\text{B.19})$$

Proof. From (2.43) and Lemma 2.1, we know that

$$f(x; \boldsymbol{\theta}) = \psi(\tilde{f}(x; \boldsymbol{\theta}), x) = \varphi(\zeta(\tilde{f}(x; \boldsymbol{\theta})), x), \quad (\text{B.20})$$

where \tilde{f} is the FNN defined in (2.42) and $\varphi : \tilde{\mathcal{Z}} \times \mathcal{X} \mapsto \mathbb{R}^{N_a}$, $\zeta : \mathbb{R}^{N_a+1} \mapsto \tilde{\mathcal{Z}}$ are defined in (2.46).

It can be easily verified that ζ satisfies the following:

(i) ζ has a continuous right inverse, e.g.

$$\zeta^{-1}(z) : Im(\zeta) \mapsto \mathbb{R}^{N_a+1}, \zeta^{-1}(z) = \left(\log(z_1), \dots, \log(z_{N_a}), \sigma^{-1}(z_{N_a+1}/p_{max}) \right)^\top, \quad (\text{B.21})$$

where σ^{-1} is the inverse function of the sigmoid function.

(ii) $Im(\zeta)$ is dense in $\tilde{\mathcal{Z}}$. This is because $\overline{Im(\zeta)}$, the closure of $Im(\zeta)$, is $\tilde{\mathcal{Z}}$.

(iii) $\partial Im(\zeta)$ is collared (Brown, 1962; Connelly, 1971; Baillif, 2022).

Furthermore, consider the partition of \mathcal{X} , $\{\mathcal{X}_1, \mathcal{X}_2\}$, which is defined in Definition 2.1. It is easily verifiable that φ is Lipschitz continuous on $\tilde{\mathcal{Z}} \times \mathcal{X}_1$ and $\tilde{\mathcal{Z}} \times \mathcal{X}_2$ respectively.

Finally, according to Assumption 2.7, $p^*(x) = \varphi(\omega^*(x), x)$, where $\omega^* \in C(\mathcal{X}, \tilde{\mathcal{Z}})$.

Applying Lemma B.1 with $\mathcal{Y} = \tilde{\mathcal{Z}}$, $\Omega(\cdot) = \omega^*(\cdot)$, $\Upsilon(\cdot) = \zeta(\cdot)$, and $\Phi(\cdot, \cdot) = \varphi(\cdot, \cdot)$, we know that there exists $N_h \in \mathbb{N}$, and $\theta \in \mathbb{R}^{N_\theta}$ such that the corresponding LFNN model $f(x; \theta) = \varphi(\zeta(\tilde{f}(x; \theta)), x)$ satisfies the following:

$$\sup_{x \in \mathcal{X}} \|f(x; \theta) - p^*(x)\| < \epsilon. \quad (\text{B.22})$$

□

C Comparing LFNN with closed-form solution

In this section, we compare the performance of the strategy following the learned shallow LFNN model (which we refer to as the “neural network strategy” from now on) with the closed-form solution (2.20), and provide empirical validation of the LFNN approach.

C.1 Approximate form under realistic assumptions

We first note that the closed-form solution p^* defined in (2.20) is obtained under several unrealistic assumptions, namely continuous rebalancing, unlimited leverage, and continuing trading in insolvency.¹⁹ In practice, investors have constraints such as discrete rebalancing, limited leverage, and no trading when insolvent. For a meaningful comparison, instead of comparing the neural network strategy with the closed-form solution p^* directly, we compare the neural network strategy with an easily obtainable approximation to the closed-form solution which satisfies realistic constraints.

In particular, we consider an equally-spaced discrete rebalancing schedule $\mathcal{T}_{\Delta t}$ defined as

$$\mathcal{T}_{\Delta t} = \{t_i : i = 0, \dots, N\}, \quad (\text{C.1})$$

where $t_i = i\Delta t$, and $\Delta t = T/N$. Then, the *clipped form* $\bar{p}_{\Delta t} : \mathcal{T}_{\Delta t} \times \mathbb{R}^3 \mapsto \mathbb{R}$ is defined as

$$(\text{Clipped form}) : \quad \bar{p}_{\Delta t}(t_i, \bar{W}_{\Delta t}(t_i), \hat{W}_{\Delta t}(t_i), \hat{\rho}) = \min \left(\max \left(p^*(t_i, \bar{W}_{\Delta t}(t_i), \hat{W}_{\Delta t}(t_i), \hat{\rho}), p_{min} \right), p_{max} \right). \quad (\text{C.2})$$

Here $[p_{min}, p_{max}]$, where $p_{min} = 0$ and $p_{max} \geq 1$, is the allowed range, $\bar{W}_{\Delta t}(t_i)$ is the wealth of the active portfolio at t_i following $\bar{p}_{\Delta t}$ from t_0 to t_i , $\hat{W}_{\Delta t}(t_i)$ is the wealth of the benchmark portfolio at t_i following the fixed-mix strategy described by constant allocation fraction $\hat{\rho}$, but only rebalanced discretely according to $\mathcal{T}_{\Delta t}$. Clearly, the allocation strategy from $\bar{p}_{\Delta t}$ follows the discrete schedule of $\mathcal{T}_{\Delta t}$, and satisfies the leverage constraint that $\bar{p}_{\Delta t} \in [p_{min}, p_{max}]$. $\bar{p}_{\Delta t}$ approaches the closed-form solution p^* as $\Delta t \downarrow 0, p_{min} \downarrow -\infty$ and

¹⁹Note that we consider a two-asset scenario here, thus the scalar $p^* \in \mathbb{R}$ (allocation fraction for the stock index) fully describes the allocation strategy \mathbf{p}^* , since $\mathbf{p}^* = (p^*, 1 - p^*)^\top$.

$p_{max} \uparrow \infty$. We note that a similar clipping idea is explored in Vigna (2014) in the context of closed-form solutions for multi-period mean-variance asset allocation. However, it should be emphasized that the clipped form $\bar{p}_{\Delta t}$ with finite (p_{min}, p_{max}) is a feasible, but in general sub-optimal, control of the leverage-constrained CD problem (2.12).

We then address the assumption that trading continues when insolvent, i.e., when the wealth of the portfolio reaches zero. While necessary for the mathematical derivation of the closed-form solution, we acknowledge that this is by no means reasonable for practitioners. Under the continuous rebalancing case (no jumps), if the control (allocation) is bounded, it is shown that the wealth of the portfolio can never be negative (Wang and Forsyth, 2012). However, with discrete rebalancing, even with a bounded control, as long as the upper bound $p_{max} > 1$, it is theoretically possible that the portfolio value becomes negative. We address this assumption by applying an overlay on strategies so that in the case of insolvency, we assume the manager liquidates the long-only positions and allocates the debt (negative wealth) to a shortable (bond) asset (consistent with Assumption 2.6) to allow outstanding debt to accumulate until the end of the investment horizon. Going forward, when we refer to any strategy (e.g. neural network strategy, clipped form), we mean the strategy with this overlay applied. We remark that in practice, this overlay has little effect. In numerical experiments with 10,000 samples of observed wealth trajectories (based on calibrated jump-diffusion model or bootstrap resampled data paths), we do not observe any single wealth trajectory that ever hits negative wealth for any strategy (e.g., neural network strategy, clipped form, etc).

In summary, the clipped form satisfies the realistic constraints and is a comparable benchmark for the neural network strategy. In the following section, we will numerically compare the performance of the clipped form, the neural network strategy, and the closed-form solution.

C.2 Comparison: LFNN strategy vs clipped-form solution

We assess and compare the performance of the neural network strategy and the clipped form, we assume the following investment scenario described in Table C.1.

Investment horizon T (years)	10
Assets	CRSP cap-weighted index (real)/30-day T-bill (U.S.) (real)
Index Samples	Concatenated 1940:8-1951:7, 1968:9-1985:10
Initial portfolio wealth/annual cash injection	100/10
Rebalancing frequency	Monthly, quarterly, semi-annually, annually
Maximum leverage	1.3
Benchmark equity percentage	0.7
Outperformance target rate β	1% (100 bps)

Table C.1: Investment scenario.

We assume the stock index and the bond index prices follow a double exponential jump model (2.15), see e.g., (Kou, 2002; Kou and Wang, 2004), i.e., for the jump variable ξ_i , $y_i = \log(\xi_i)$ follows the double exponential distribution with density functions $g_i(y_i)$ defined as follows

$$g_i(y_i) = \nu_i \iota_i e^{-\iota_i y_i} \mathbf{1}_{y_i \geq 0} + (1 - \nu_i) \varsigma_i e^{-\varsigma_i y_i} \mathbf{1}_{y_i < 0}, \quad i = 1, 2. \quad (\text{C.3})$$

where ν_i is the probability for an upward jump, and ι_i and ς_i are parameters that describe the upward jump and downward jump respectively. The double exponential jump-diffusion model allows the flexibility of modeling asymmetric upward and downward jumps in asset prices, which seems an appropriate assumption for inflation regimes.²⁰

Using the threshold technique (Mancini, 2009; Cont et al., 2011; Dang and Forsyth, 2016), we calibrate the double exponential jump-diffusion models to the historical high-inflation periods described in Section 3.1.

²⁰We remind the reader that the closed-form solution is derived under the jump-diffusion model.

The calibrated parameters can be found in Appendix G in Ni et al. (2023). Then, we construct a training data set \mathbf{Y} and a testing data set \mathbf{Y}^{test} by sampling the calibrated model, each with 10,000 samples.

The neural network strategy follows the LFNN model learned from \mathbf{Y} . We then evaluate the performance of the neural network strategy and the approximate form (C.2) on the testing data set \mathbf{Y}^{test} . Specifically, we compare the value of the CD objective function (2.12) for the neural network strategy and the clipped form on \mathbf{Y}^{test} . In particular, this training/testing process is repeated for various rebalancing frequencies from monthly to annually, as described in Table C.1.

Closed-form solution objective function value: 418					
Strategy	$\Delta t = 1$	$\Delta t = 1/2$	$\Delta t = 1/4$	$\Delta t = 1/12$	$\Delta t = 0$
Clipped form	545	504	479	467	461 (extrapolated)
Neural network	537	498	476	464	458 (extrapolated)

Table C.2: CD objective function values. Results shown are evaluated on \mathbf{Y}^{test} , the lower the better.

In Table C.2, we can see that the neural network strategy consistently outperforms the clipped form in terms of the objective function value for all rebalancing frequencies. From Table 4.2 we can see that the objective function values of both the neural network strategy and the clipped form converge at roughly a first-order rate as $\Delta t \downarrow 0$. Assuming this to be true, we extrapolate the solution to $\Delta t = 0$ using Richardson extrapolation. These extrapolated values are estimates of the exact value of the continuous-time CD objective function (2.11) for the clipped form and the neural network strategy. We can see that the neural network strategy still outperforms the clipped form in terms of the extrapolated objective function value. We can also see that the extrapolated neural network objective function value is lower than the (suboptimal) clipped form extrapolated value, but, of course, larger than the unconstrained closed-form solution.

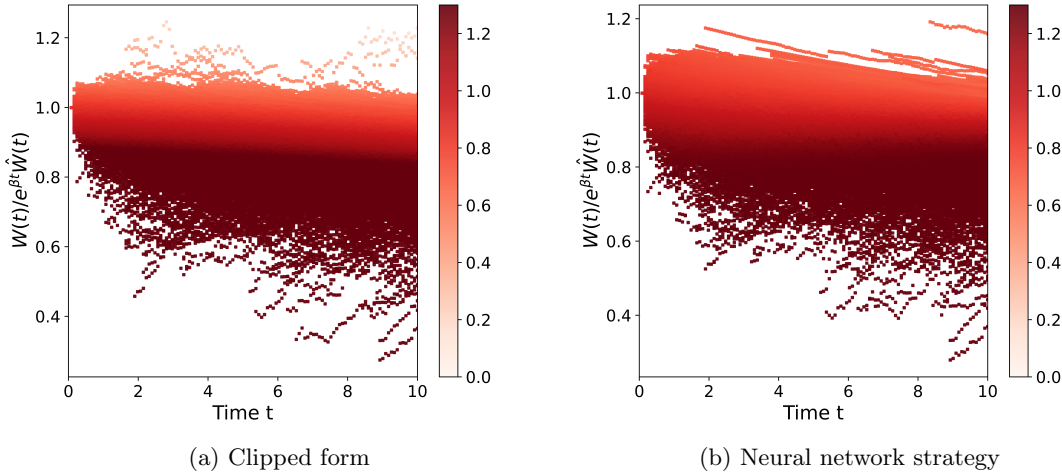


Figure C.1: Stock allocation fraction w.r.t. tracking ratio $W(t)/\left(e^{\beta t}\hat{W}(t)\right)$ and time t . Results are based on the evaluation on the testing data set \mathbf{Y}^{test} , and monthly rebalancing (i.e. $\Delta t = 1/12$).

Finally, we compare the neural network allocation strategy with the clipped form strategy. Specifically, in Figure C.1, we consider the case of monthly rebalancing and present the scatter plots of the allocation fraction in the stock index with respect to time t and the ratio between the wealth of the active portfolio $W(t)$ and the elevated target $e^{\beta t}\hat{W}(t)$. For simplicity, we call this ratio the “tracking ratio”. We plot the 3-tuple $\left(\frac{W(t)}{e^{\beta t}\hat{W}(t)}, t, p_1(W(t), \hat{W}(t), t)\right)$ (obtained from the evaluation of the strategies on samples from \mathbf{Y}^{test})

by using time t as the x-axis, the tracking ratio $\frac{W(t)}{e^{\beta t} \hat{W}(t)}$ as the y-axis, and the values of the corresponding allocation fraction to the cap-weighted index $p_1(W(t), \hat{W}(t), t)$ to color the scattered dots on the plot. A darker shade of the color indicates a higher allocation fraction.

As we can see from Figure C.1, the stock allocation fraction of the neural network strategy behaves similarly to the stock allocation fraction from the clipped form. Both strategies invest more wealth in the stock when the tracking ratio is lower, which is consistent with the insights we obtained in Section 2.4.1. In addition, the transition patterns of the allocation fraction of the two strategies are also highly similar. One can almost draw an imaginary horizontal dividing line around $\frac{W(t)}{e^{\beta t} \hat{W}(t)} = 0.9$ that separates high stock allocation and low stock allocation for both strategies.

We remark that a common criticism towards the use of neural networks is about the lack of interpretability compared to more interpretable counterparts such as the regression models (Rudin, 2019). In this section, we see that the neural network strategy closely resembles the closed-form solution for the CD objective. The closed-form solution, in turn, complements the neural network model and offers an alternative way of interpreting results obtained from the neural network.

References

- Al-Aradi, A. and S. Jaimungal (2018a). Outperformance and tracking: dynamic asset allocation for active and passive portfolio management. *Applied Mathematical Finance* 25(3), 268–294.
- Al-Aradi, A. and S. Jaimungal (2018b). Outperformance and tracking: Dynamic asset allocation for active and passive portfolio management. *Applied Mathematical Finance* 25(3), 268–294.
- Alekseev, A. G. and M. V. Sokolov (2016). Benchmark-based evaluation of portfolio performance: a characterization. *Annals of Finance* 12, 409–440.
- Alizadeh, A. H. and N. K. Nomikos (2007). Investment timing and trading strategies in the sale and purchase market for ships. *Transportation Research Part B: Methodological* 41(1), 126–143.
- Anarkulova, A., S. Cederburg, and M. S. O’Doherty (2022). Stocks for the long run? evidence from a broad sample of developed markets. *Journal of Financial Economics* 143(1), 409–433.
- Baillif, M. (2022). Collared and non-collared manifold boundaries. *L’Enseignement Mathématique* 68(1), 161–180.
- Ball, L. M., D. Leigh, and P. Mishra (2022). Understanding us inflation during the covid era. Technical report, National Bureau of Economic Research.
- Basak, S., A. Shapiro, and L. Tepla (2006). Risk management with benchmarking. *Management Science* 52(4), 542–557.
- Black, F. (1993). Estimating expected return. *Financial Analysts Journal* 49(5), 36–38.
- Bo, L., H. Liao, and X. Yu (2021). Optimal tracking portfolio with a ratcheting capital benchmark. *SIAM Journal on Control and Optimization* 59(3), 2346–2380.
- Boyde, E. (2021). ETFs: why tracking difference usually matters more than tracking error. <https://www.ft.com/content/44b08a6c-55d7-4841-8e8e-aad6451a4cc3>. [Online; accessed 6-Dec-2022].
- Brigo, D., A. Dalessandro, M. Neugebauer, and F. Triki (2008). A stochastic processes toolkit for risk management. *arXiv preprint arXiv:0812.4210*.
- Brown, M. (1962). Locally flat imbeddings of topological manifolds. *Annals of Mathematics*, 331–341.
- Browne, S. (1999). Beating a moving target: optimal portfolio strategies for outperforming a stochastic benchmark. *Finance and Stochastics* 3, 275–294.
- Browne, S. (2000). Risk-constrained dynamic active portfolio management. *Management Science* 46(9), 1188–1199.
- Buehler, H., L. Gonon, J. Teichmann, and B. Wood (2019). Deep hedging. *Quantitative Finance* 19:8, 1271–1291.
- Bureau of Labor Statistics (2023). Consumer price index summary.
- Cavaglia, S., L. Scott, K. Blay, and S. Hixon (2022). Multi-asset class factor premia: A strategic asset allocation perspective. *The Journal of Portfolio Management* 48(4), 14–32.
- Charteris, A. and K. McCullough (2020). Tracking error vs tracking difference: Does it matter? *Investment Analysts Journal* 49(3), 269–287.
- Coache, A. and S. Jaimungal (2021). Reinforcement learning with dynamic convex risk measures. *arXiv preprint arXiv:2112.13414*.

- Cogneau, P. and V. Zakamouline (2013). Block bootstrap methods and the choice of stocks for the long run. *Quantitative Finance* 13(9), 1443–1457.
- Connelly, R. (1971). A new proof of brown’s collaring theorem. *Proceedings of the American Mathematical Society* 27(1), 180–182.
- Cont, R., C. Mancini, et al. (2011). Nonparametric tests for pathwise properties of semimartingales. *Bernoulli* 17(2), 781–813.
- CPP Investments (2021). Annual Report 2021. <https://www.cppinvestments.com/the-fund/our-performance/financial-results/>. [Online; accessed 5-Dec-2022].
- CPP Investments (2022). Annual Report 2022. https://www.cppinvestments.com/wp-content/uploads/2022/06/ CPP-Investments_F2022-Annual-Report-EN.pdf. [Online; accessed 5-Dec-2022].
- Dang, D.-M. and P. A. Forsyth (2016). Better than pre-commitment mean-variance portfolio allocation strategies: A semi-self-financing Hamilton–Jacobi–Bellman equation approach. *European Journal of Operational Research* 250(3), 827–841.
- Davis, M. and S. Lleo (2008). Risk-sensitive benchmarked asset management. *Quantitative Finance* 8(4), 415–426.
- Dichtl, H., W. Drobetz, and L. Kryzanowski (2016). Timing the stock market: Does it really make no sense? *Journal of Behavioral and Experimental Finance* 10, 88–104.
- Dixon, M. F., I. Halperin, and P. Bilokon (2020). *Machine learning in Finance*, Volume 1406. Springer.
- Efron, B. (1992). Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pp. 569–593. Springer.
- Forsyth, P. A. (2020). Optimal dynamic asset allocation for dc plan accumulation/decumulation: Ambition-cvar. *Insurance: Mathematics and Economics* 93, 230–245.
- Forsyth, P. A., P. van Staden, and Y. Li (2022). Beating a constant weight benchmark: easier done than said. Working paper, Cheriton School of Computer Science, University of Waterloo.
- Gao, Z., Y. Gao, Y. Hu, Z. Jiang, and J. Su (2020). Application of deep q-network in portfolio management. In *2020 5th IEEE International Conference on Big Data Analytics (ICBDA)*, pp. 268–275. IEEE.
- Goodfellow, I., Y. Bengio, and A. Courville (2016). *Deep learning*. MIT press.
- Han, J. et al. (2016). Deep learning approximation for stochastic control problems. *arXiv preprint arXiv:1611.07422*.
- Homer, S. and R. E. Sylla (1996). *A history of interest rates*. Rutgers University Press.
- Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural networks* 4(2), 251–257.
- Hougan, M. (2015). Tracking difference, the perfect etf metric. *Retrieved* 8(1), 2020.
- Javorcik, B. (2020). Global supply chains will not be the same in the post-covid-19 world. *COVID-19 and trade policy: Why turning inward won’t work* 111.
- Johnson, B., H. Bioy, A. Kellett, and L. Davidson (2013). On the right track: Measuring tracking efficiency in etfs. *The Journal of Index Investing* 4(3), 35–41.
- Kingma, D. P. and J. Ba (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

- Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management science* 48(8), 1086–1101.
- Kou, S. G. and H. Wang (2004). Option pricing under a double exponential jump diffusion model. *Management science* 50(9), 1178–1192.
- Kratsios, A. and I. Bilokopytov (2020). Non-euclidean universal approximation. *Advances in Neural Information Processing Systems* 33, 10635–10646.
- Lahaye, J., S. Laurent, and C. J. Neely (2011). Jumps, cojumps and macro announcements. *Journal of Applied Econometrics* 26(6), 893–921.
- Li, D. and W.-L. Ng (2000). Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. *Mathematical finance* 10(3), 387–406.
- Li, Y. and P. A. Forsyth (2019). A data-driven neural network approach to optimal asset allocation for target based defined contribution pension plans. *Insurance: Mathematics and Economics* 86, 189–204.
- Lim, A. and B. Wong (2010a). A benchmark approach to optimal asset allocation for insurers and pension funds. *Insurance: Mathematics and Economics* 46(2), 317–327.
- Lim, A. E. and B. Wong (2010b). A benchmarking approach to optimal asset allocation for insurers and pension funds. *Insurance: Mathematics and Economics* 46(2), 317–327.
- Lin, Y., R. D. MacMinn, and R. Tian (2015). De-risking defined benefit plans. *Insurance: Mathematics and Economics* 63, 52–65.
- Lucarelli, G. and M. Borrotti (2020). A deep q-learning portfolio management framework for the cryptocurrency market. *Neural Computing and Applications* 32(23), 17229–17244.
- L’Her, J.-F., R. Stoyanova, K. Shaw, W. Scott, and C. Lai (2016). A bottom-up approach to the risk-adjusted performance of the buyout fund market. *Financial Analysts Journal* 72(4), 36–48.
- MacMinn, R., P. Brockett, J. Wang, Y. Lin, and R. Tian (2014). The securitization of longevity risk and its implications for retirement security. *Recreating sustainable retirement*, 134–160.
- Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics* 36(2), 270–296.
- Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of financial economics* 3(1-2), 125–144.
- NASDAQ (2023). NASDAQ Composite Index. <https://www.nasdaq.com/market-activity/index/comp/historical>. [Online; accessed 15-Feb-2023].
- Ni, C., Y. Li, P. Forsyth, and R. Carroll (2022). Optimal asset allocation for outperforming a stochastic benchmark target. *Quantitative Finance* 22(9), 1595–1626.
- Ni, C., Y. Li, and P. A. Forsyth (2023). Neural network approach to portfolio optimization with leverage constraints: a case study on high inflation investment. *arXiv preprint arXiv:2304.05297v2*.
- Nicolosi, M., F. Angelini, and S. Herzel (2018). Portfolio management with benchmark related incentives under mean reverting processes. *Annals of Operations Research* 266, 373–394.
- Norges Bank (2022). Investment Strategy. <https://www.nbim.no/en/the-fund/how-we-invest/investment-strategy/>. [Online; accessed 5-Dec-2022].
- Oderda, G. (2015). Stochastic portfolio theory optimization and the origin of rule based investing. *Quantitative Finance* 15(8), 1259–1266.

- Øksendal, B. K. and A. Sulem (2007). *Applied stochastic control of jump diffusions*, Volume 498. Springer.
- Park, H., M. K. Sim, and D. G. Choi (2020). An intelligent financial portfolio trading strategy using deep q-learning. *Expert Systems with Applications* 158, 113573.
- Phalippou, L. (2014). Performance of buyout funds revisited? *Review of Finance* 18(1), 189–218.
- Politis, D. N. and J. P. Romano (1994). The stationary bootstrap. *Journal of the American Statistical Association* 89(428), 1303–1313.
- Powell, W. (2023). A universal framework for sequential decision problems. *OR/MS Today February*. <https://tinyurl.com/PowellORMSfeature/>.
- Reppen, A. M., H. M. Soner, and V. Tissot-Daguette (2022). Deep stochastic optimization in finance. *arXiv preprint arXiv:2205.04604*.
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature machine intelligence* 1(5), 206–215.
- Scott, L. and S. Cavaglia (2017). A wealth management perspective on factor premia and the value of downside protection. *The Journal of Portfolio Management* 43(3), 33–41.
- Shahzad, S. J. H., E. Bouri, D. Roubaud, L. Kristoufek, and B. Lucey (2019). Is bitcoin a better safe-haven investment than gold and commodities? *International Review of Financial Analysis* 63, 322–330.
- Silver, D., G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller (2014). Deterministic policy gradient algorithms. In *International conference on machine learning*, pp. 387–395. Pmlr.
- Simonian, J. and A. Martirosyan (2022). Sharpe parity redux. *The Journal of Portfolio Management* 48(4), 183–193.
- Tepla, L. (2001). Optimal investment with minimum performance constraints. *Journal of Economic Dynamics and Control* 25, 1629–1645.
- The Federal Reserve (2011). What is an acceptable level of inflation? <https://www.federalreserve.gov/faqs/5D58E72F066A4DBDA80BBA659C55F774.htm>. [Online; accessed 5-Dec-2022].
- Tsang, K. H. and H. Y. Wong (2020). Deep-learning solution to portfolio selection with serially dependent returns. *SIAM Journal on Financial Mathematics* 11(2), 593–619.
- van Staden, P. M., P. A. Forsyth, and Y. Li (2022). Across-time risk-aware strategies for outperforming a benchmark. Working paper.
- van Staden, P. M., P. A. Forsyth, and Y. Li (2023). Beating a benchmark: dynamic programming may not be the right numerical approach. *SIAM Journal on Financial Mathematics* 14:2, 407–451.
- Vigna, E. (2014). On efficiency of mean–variance based portfolio selection in defined contribution pension schemes. *Quantitative finance* 14(2), 237–258.
- Wang, J. and P. A. Forsyth (2010). Numerical solution of the hamilton–jacobi–bellman formulation for continuous time mean variance asset allocation. *Journal of Economic Dynamics and control* 34(2), 207–230.
- Wang, J. and P. A. Forsyth (2012). Comparison of mean variance like strategies for optimal asset allocation problems. *International journal of theoretical and applied finance* 15(02), 1250014.
- Wang, R., D. P. Foster, and S. M. Kakade (2020). What are the statistical limits of offline rl with linear function approximation? *arXiv preprint arXiv:2010.11895*.

- Yao, D. D., S. Zhang, and X. Y. Zhou (2006). Tracking a financial benchmark using a few assets. *Operations Research* 54(2), 232–246.
- Zhang, Q. and Y. Gao (2017). Portfolio selection based on a benchmark process with dynamic value-at-risk constraints. *Journal of Computational and Applied Mathematics* 313, 440–447.
- Zhao, Y. (2007). A dynamic model of active portfolio management with benchmark orientation. *Journal of Banking & Finance* 31(11), 3336–3356.
- Zhou, X. Y. and D. Li (2000). Continuous-time mean-variance portfolio selection: A stochastic lq framework. *Applied Mathematics and Optimization* 42, 19–33.
- Ziembka, W. T. (2005). The symmetric downside-risk sharpe ratio. *The Journal of Portfolio Management* 32(1), 108–122.