Optimal Asset Allocation in a High Inflation Regime: 
a Leverage-feasible Neural Network Approach

Chendi Ni\textsuperscript{1}, Yuying Li\textsuperscript{2}, and Peter Forsyth\textsuperscript{3}

\textsuperscript{1}Cheriton School of Computer Science, University of Waterloo, Waterloo, N2L 3G1, Canada, chendi.ni@uwaterloo.ca
\textsuperscript{2}Cheriton School of Computer Science, University of Waterloo, Waterloo, N2L 3G1, Canada, yuying@uwaterloo.ca
\textsuperscript{3}Cheriton School of Computer Science, University of Waterloo, Waterloo, N2L 3G1, Canada, paforsyt@uwaterloo.ca

April 11, 2023

Abstract

We study the optimal multi-period asset allocation problem with leverage constraints in a persistent, high-inflation environment. Based on filtered high-inflation regimes, we discover that a portfolio containing an equal-weighted stock index partially stochastically dominates a portfolio containing a capitalization-weighted stock index. Assuming the asset prices follow the jump diffusion model during high inflation periods, we establish a closed-form solution for the optimal strategy that outperforms a passive strategy under the cumulative quadratic tracking difference (CD) objective. The closed-form solution provides insights but requires unrealistic constraints. To obtain strategies under more practical considerations, we consider a constrained optimal control problem with bounded leverage. To solve this optimal control problem, we propose a novel leverage-feasible neural network (LFNN) model that approximates the optimal control directly. The LFNN model avoids high-dimensional evaluation of the conditional expectation (common in dynamic programming (DP) approaches). 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. Numerical experiments show that the LFNN model achieves comparable performance to the closed-form solution on simulated data. We apply the LFNN approach to a four-asset investment scenario with bootstrap resampled asset returns. The LFNN strategy consistently outperforms the passive benchmark strategy by about 200 bps (median annualized return), with a greater than 90% probability of outperforming the benchmark at the terminal date. These results suggest that during persistent inflation regimes, investors should favor short-term bonds over long-term bonds, and the equal-weighted stock index over the cap-weighted stock index.

1 Introduction

Since the global outbreak of COVID-19 in March 2020, we have observed significant world-wide inflation. Particularly, from May 2021 to February 2023 (the latest information as of the time of writing), the 12-month change in the CPI index in the U.S. has not been lower than 5% \cite{Bureau_of_Labor_Statistics_2023}. Before the pandemic, the U.S. economy has seen almost four decades of low inflation. The sudden shift from a long-term low-inflation environment to a high-inflation environment causes large uncertainty and volatility in the financial market. In 2022, the technology-heavy NASDAQ stock index recorded a yearly return of $-33.10\%$ \cite{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 cause 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 do not attempt to make predictions about the future inflation situation. Instead, we look into historical high-inflation regimes, and try to answer the following question: how should investors invest in a persistent high-inflation environment? More specifically, we formulate the investment problem as a multi-period asset allocation problem, in which an investment manager rebalances an active portfolio periodically, with the goal of finding asset allocation strategies that perform well under specific investment criteria during high-inflation regimes.

To gain some initial insights into investing during high-inflation regimes, in Section 2 we study the performance of fixed-mix (constant weight) strategies on bootstrap resampled data from historical high-inflation periods. Fixed-mix strategies allocate a fixed fraction of the portfolio wealth to the underlying assets and are common choices of benchmark strategies for large sovereign wealth funds (CPP Investments, 2022; Norges Bank, 2022). Our preliminary findings show that the fixed-mix strategies with the equal-weighted stock index achieves partial first-order stochastic dominance over fixed-mix strategies which use the cap-weighted stock index during high-inflation regimes.

In Section 3, we seek an active investment strategy (i.e. a dynamic asset allocation), so that the active portfolio would outperform a fixed-mix benchmark portfolio consistently throughout the investment horizon. We assume that the real (inflation adjusted) asset returns during a high-inflation regime follow stochastic processes and treat the asset allocation decisions as the control of a dynamic system. The multi-period asset allocation problem then becomes a stochastic optimal control problem.

To solve the stochastic optimal control problem, in Section 3.4, we study a two-asset case under a synthetic market assumption. The two assets are a stock index, and a constant maturity bond index. There is a large amount of extant literature on closed-form solution methods for beating a stochastic benchmark under synthetic market assumptions (Browne, 1999, 2000; Teplá, 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 many of these articles, the 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 as well as the assumption that the stock price follows a double exponential jump-diffusion model and the bond price is deterministic. We extend the results in Forsyth et al. (2022) by considering the case that both the stock index and the bond index follow jump-diffusion models. While this introduces additional complexity to the problem, we believe the assumption that the bond index price is stochastic and has jumps is appropriate under a high-inflation scenario. In Section 3.4, we present the closed-form solution under the CD objective derived under the synthetic model. The closed-form solution provides insights for understanding the behaviour of the optimal control (allocation strategy).

However, to derive the closed-form solution, some unrealistic assumptions such as continuous rebalancing, infinite leverage, and continued trading when insolvent, are required. Ideally, we would like to solve the stochastic optimal control problem under realistic constraints, i.e., discrete rebalancing, limited leverage, and no trading when insolvent. To solve a discrete-time multi-period asset allocation problem, it is common to use the dynamic programming (DP) approach to convert 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 3.7, we propose to use a single neural network model to approximate the optimal control, which converts the multi-period optimization problem into a single finite-
dimensional optimization problem, which is solved by learning the parameters of the neural network model on data sets of asset returns using standard optimization methods. 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 which 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 known to be a continuous function of time, which we discuss in detail in Section 3.4.

In terms of the modeling framework, the closest work is Li and Forsyth (2019); Ni et al. (2022), which also use a single neural network to approximate the controls under an optimal defined-contribution plan problem. However, the neural network models proposed in Li and Forsyth (2019); Ni et al. (2022) are limited to solving long-only investment problems, in which leverage is not permitted. In contrast, we consider the more practical investment scenario that allows a certain amount of leveraging and shorting in the investment portfolio to mimic the investment behavior of large sovereign wealth funds. The leverage assumption introduces additional constraints to the optimization problem, and thus requires a new model architecture. Particularly, in Section 3.7, we propose the novel leverage-feasible neural network (LFNN) model that converts the leverage-constrained optimization problem into an unconstrained optimization problem. In Section 3.8, we prove that, under benign assumptions, the solution of the unconstrained optimization problem (parameterized by the LFNN model) can approximate the optimal control arbitrarily well, which mathematically justifies the validity of the LFNN approach.

In addition, compared to Li and Forsyth (2019); Ni et al. (2022), we consider the more practical cumulative tracking difference-based investment objective that evaluates the relative performance of the active portfolio versus the benchmark portfolio throughout the investment horizon, instead of only considering the performance at the terminal time of the investment horizon.

In Section 4, we discuss various numerical experiments and results. In Section 4.1, we first compare the leverage-constrained investment strategy computed using the LFNN model with the clipped form of the closed-form solution on synthetic market data. We show that the LFNN model produces a strategy that achieves comparable performance against the closed-form solution, thus substantiating the LFNN approach. In Section 4.2, we further apply the neural network model to a more realistic case with four underlying assets, which include the equal-weighted stock index, the cap-weighted stock index, the 30-day U.S. treasury bill index, and the 10-year U.S. treasury bond index. We conduct numerical experiments on bootstrap resampled data under the cumulative quadratic tracking shortfall (CS) objective and obtain a leverage-constrained strategy that outperforms the benchmark with more 2% higher median (annualized) internal rate of return (IRR), and more than 90% probability of achieving a higher terminal wealth.

We summarize our contributions below:

(i) We assess the performance of the passive investment strategies on resampled data from filtered historical inflation periods. We find that the fixed-mix strategy with the equal-weighted stock index partially stochastically dominates the fixed-mix strategy with the cap-weighted stock index in resampled high-inflation regimes. This suggests that the equal-weighted stock index is preferred over the cap-weighted stock index during high inflation periods.

(ii) To discover dynamic allocation strategies that outperform the passive fixed-mix strategies in high-inflation regimes, we formulate the problem as a stochastic optimal control problem with suitable objective functions under leverage constraints. To gain intuition about the behavior of the optimal controls, we derive the closed-form solution under a jump-diffusion asset price model and 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.
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 that allows us to convert the original leverage-constrained optimization problem into an unconstrained optimization problem.

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.

We benchmark the performance of the LFNN model with the clipped form of the closed-form solution and show that the LFNN model achieves comparable objective function values, and the allocation strategy from the LFNN model achieves highly comparable performance as the clipped form and the closed-form solution. The results provide empirical validation for the neural network approach.

In a more realistic stochastic market scenario, we apply the neural network method to bootstrap resampled asset returns with four underlying assets, including the equal-weighted and cap-weighted stock index, and the 30-day/10-year treasury bond indexes. The dynamic strategy that follows 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 Passive investing in historical high-inflation periods

Inflation has been a topic of much recent concern. Before the recent COVID-19 pandemic, we have just observed a long period of benign inflation, and low (real) short-term interest rates. Some would argue that this has led to a bubble in asset prices. Now, after the coronavirus crisis, we are seeing high inflation statistics in most of the world. Going forward, we have to be cognizant of the risk of inflation. If we enter into a long period of even moderate inflation, will the traditional passive approach using a mix of stock indexes and moderate-term bonds still work?

In this section, we first filter the historical time series (1926-2022) to uncover periods of high, sustained inflation. We concatenate these high-inflation regimes to produce asset return series which have been observed in inflationary times. We then examine the performance of various passive investment strategies on bootstrapped returns from the high-inflation series, aiming to gain a crude understanding of high-inflation investing. Our preliminary finding is that an investor should use a mix of short-term bonds and an equal-weighted stock index, to mitigate inflationary effects.

2.1 Filtering historical inflation regimes

We use monthly data from the Center for Research in Security Prices (CRSP) over the 1926:1-2022:1 period. We also use the U.S. CPI index, also supplied by CRSP.

Our objective is to select high-inflation periods as determined by the CPI index. Monthly data is quite volatile, so we use the following filtering procedure. We use a moving window of $k$ months, and 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.

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

More 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.
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. See Algorithm A.1 in Appendix A.1 for the pseudo-code.

This approach requires the specification of the cutoff and the window size. The average annual inflation over the period 1926:1-2022:1 was 2.9%. In other words, inflation of about 3% was normal historically. 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). After some experimentation, we use a cutoff of 5% as the threshold for high inflation, which is more than double the Fed’s target inflation rate. In addition, we use the moving window size of 5 years (see Appendix A.2 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 2.1 shows the average annual inflation over the two regimes identified from the moving-window filter.

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Average Annualized Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940:8-1951:7</td>
<td>0.0564</td>
</tr>
<tr>
<td>1968:9-1985:10</td>
<td>0.0661</td>
</tr>
</tbody>
</table>

Table 2.1: 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”). In addition, we construct a constant maturity 10-year U.S. treasury index. We also study the cap-weighted stock index (CapWt) and the equal-weighted stock index (EqWt), also from CRSP. The CRSP indexes are total return indexes, which include all distributions for all domestic stocks trading on major U.S. exchanges. All of these various indexes are in nominal terms, so we adjust them for inflation by using the U.S. CPI index, also supplied by CRSP.

We calculate the average historical returns of the assets during the two inflation periods (see details in Appendix A.3), and 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 similar average return as the 10-year T-bond index, but much lower volatility. This seems to indicate that the T-bill index is the better choice of defensive asset during high inflation. Therefore, we consider the equal-weighted stock index, the cap-weighted stock index and the 30-day T-bill index for the following experiment.

2.2 Bootstrap resampling

We choose to test allocation strategies using stationary block bootstrap resampling (Politis and Romano, 1994). See Appendix B.1 for detailed pseudo-code for bootstrap resampling. Compared to the traditional bootstrap method, the block bootstrap technique preserves the local dependency of data within blocks. The stationary block bootstrap uses random block sizes which preserves the stationarity of the original time series data. Briefly, each bootstrap resample consists of (i) selecting a random starting date in the historical return series, (ii) then selecting a block (of random size) of consecutive returns from this start date, and (iii) repeating this process until a sample of the total desired length is obtained.

An important parameter is the expected blocksize, 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 blocksize for

---

3 The 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.

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

5
multiple underlying time series data so that the bootstrapped data entries for different assets are synchronized in time. In the following experiments, we use the expected blocksize of 6 months for all time series data. However, we have compared different numerical experiments using a range of blocksizes, including i.i.d. assumptions (i.e. expected blocksize equal to one month), and find that the results are relatively insensitive to blocksize, as discussed in more detail in Appendix B.2.

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 B.3). 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.

### 2.3 Comparison of passive strategies

In this section, we compare the performances of two fixed-mix strategies. The first strategy, the “EqWt” strategy, maintains a 70% allocation to the equal-weighted index, and 30% allocation to the 30-day T-bill index. The second strategy, the “CapWt” strategy, maintains a 70% allocation to the cap-weighted index, and 30% allocation to the 30-day T-bill index.

<table>
<thead>
<tr>
<th>Investment horizon $T$ (years)</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity market indexes</td>
<td>CRSP cap-weighted/equal-weighted index (real)</td>
</tr>
<tr>
<td>Bond index</td>
<td>30-day T-bill (U.S.) (real)</td>
</tr>
<tr>
<td>Index Samples</td>
<td>Concatenated 1940:8-1951:7, 1968:9-1985:10</td>
</tr>
<tr>
<td>Initial portfolio wealth</td>
<td>100</td>
</tr>
<tr>
<td>Rebalancing frequency</td>
<td>Monthly</td>
</tr>
</tbody>
</table>

Table 2.2: Investment scenario.

In the numerical experiment, we consider the investment scenario described in Table 2.2. Briefly, we begin with an initial wealth of 100 for both strategies, with no further cash injections and withdrawals. The investment horizon is 10 years, with monthly rebalancing to maintain the constant weights in the portfolio. We evaluate the investment results by examining the distribution of the final wealth $W(T)$ at $T = 10$ years. We evaluate the two strategies on $10,000$ block bootstrap resampled data samples (Politis and Romano 1991; Dichtl et al. 2016; Anarkulova et al. 2022) from the concatenated CRSP combined time series from 1940:8-1951:7 and 1968:9-1985:10, with an expected blocksize of six months. We then compare the CDF (cumulative distribution functions) of the terminal wealth of the EqWt strategy and the CapWt strategy.

Before we discuss the experiment results, we quickly review the concept of partial stochastic dominance. Suppose two investment strategies $A$ and $B$ are evaluated on a set of data samples under the same investment scenario. We consider the CDFs of terminal wealth $W$ associated with both strategies. Specifically, we denote the CDF of strategy A by $CDF_A(W)$ and that of strategy B by $CDF_B(W)$. Let $W_T$ be the random wealth at time $T$ and $W$ be a possible wealth realization, then we can interpret $CDF_A(W)$ as

$$CDF_A(W) = Prob(W_T \leq W). \quad (2.1)$$

Following Atkinson (1987); van Staden et al. (2021), we define partial first-order stochastic dominance.
Definition 2.1 (Partial first order stochastic dominance). Given an investment strategy A which generates a CDF of terminal wealth $W$ given by $F_A(W)$, and a strategy B with CDF $F_B(W)$, then strategy A partially stochastically dominates strategy B (to first order) in the interval $(W_{lo}, W_{hi})$ if

$$F_A(W) \leq F_B(W), \forall W \in (W_{lo}, W_{hi})$$

(2.2)

with strict inequality for at least one point in $(W_{lo}, W_{hi})$.

The arguments for relaxing the usual definition of stochastic dominance are given in (Atkinson, 1987; van Staden et al., 2021). Consider first the case of the upper bound $W_{hi}$. Given some initial wealth $W_0$, if $W_{hi} \gg W_0$, then an investor may not be concerned that strategy A underperforms strategy B at these very high wealth values. In this case, the investor is fabulously wealthy.

Turning attention to the lower bound $W_{lo}$, we consider two cases. Suppose that $W_{lo} \ll W_0$, Assume $F_A(W_{lo}) = F_B(W_{lo})$. As an extreme example, suppose $W_{lo} =$ two cents. The fact that strategy B has a higher probability of ending up with one cent, compared with strategy A is cold comfort, and not particularly interesting. On the other hand, suppose $F(W_{lo}) \ll 1$. Again, an investor may not be interested in events with exceptionally low probabilities.

![Figure 2.1: Cumulative distribution function of final real wealth $W$ at $T = 10$ years, bootstrap resampling expected blocksize six months, 10,000 resamples (Appendix A.1). $T = 10$ years. Data: concatenated returns, 1940:8-1951:7, 1968:9-1985:10. Scenario described in Table 2.2](image)

This seems to be the case for our experiment here. In Figure 2.1 we compare the CDFs of the terminal wealth of the EqWt strategy and the CapWt strategy. Remarkably, the EqWt strategy appears to partially stochastically dominate the CapWt strategy, since the CDF curve of the EqWt strategy almost appears to be entirely on the right side of the CDF curve of the CapWt strategy, except at very low probability values. In fact, close examination shows that the curves cross at the point $F_{EqWt}(W_{lo}) = F_{CapWt}(W_{lo}) \simeq .02$, with a slight underperformance of the EqWt strategy compared to the CapWt strategy in this extreme left tail.

The fact that the EqWt strategy partially stochastically dominates the CapWt strategy seems to suggest that the equal-weighted stock index is the better choice for the stock index than the cap-weight stock index during high inflation times. In a later numerical experiment in Section 4.2, where the admissible stock indexes include both cap-weighted and equal-weighted indexes, we show that the computed optimal strategy almost always ignores the cap-weighted stock index and picks the equal-weighted stock index, which further confirms the argument that the equal-weighted stock index is the preferred asset for investors during high inflation scenarios.
However, if we examine more recent data\(^5\) the situation is not so clear (Taljaard and Mare, 2021), since the equal-weighted index appears to underperform. However, Taljaard and Mare (2021) suggests that this is due to the recent market concentration in tech stocks. In fact, a plausible explanation for the outperformance (historically) of an equal-weighted index is that this is simply due to the small-cap effect, which was not widely known until about 1981 (Banz, 1981). Plyakha et al. (2021) acknowledge that the equal-weighted index has significant exposure to the size factor. However, Plyakha et al. (2021) argue that the equal-weighted index also has a larger exposure to the value factor. In addition, there is a significant \textit{alpha} effect due to the contrarian strategy of frequent rebalancing to equal weights. It would appear to be simplistic to dismiss an equal weight strategy on the grounds that this is simply a small cap effect that has become less effective.

## 3 Dynamic outperforming benchmark under bounded leverage

### 3.1 Sovereign wealth funds and benchmark targets

So far, we have examined the performance of passive portfolios in bootstrap resampled sustained inflation regimes. In practice, 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.\(^6\) 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 3.1). 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.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Excess return (%)</td>
<td>0.21</td>
<td>0.99</td>
<td>-0.77</td>
<td>0.45</td>
<td>0.15</td>
<td>0.70</td>
<td>-0.30</td>
<td>0.23</td>
<td>0.27</td>
<td>0.74</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Table 3.1: 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.

\(^5\)Since about 2010. Of course, this is outside a period of sustained high inflation.

\(^6\)As of February, 2023, Apple, Microsoft, Amazon and Alphabet (A and C) in total comprised 17% of the market capitalization of the S&P 500.

\(^7\)The 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.
3.2 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 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 \(S(t) = (S_i(t) : i = 1, \cdots, N_a)^\top \in \mathbb{R}^{N_a}\) denote the asset prices of the \(N_a\) underlying assets at time \(t \in [t_0, T]\). In addition, let vectors \(p(t) = (p_i(t) : i = 1, \cdots, N_a)^\top \in \mathbb{R}^{N_a}\) and \(\hat{p}(t) = (\hat{p}_i(t) : i = 1, \cdots, N_a)^\top \in \mathbb{R}^{N_a}\) 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 vectors \(p\) and \(\hat{p}\) can be regarded as the controls of the system, as they determine how the wealth of the portfolios 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 \(p\) and \(\hat{p}\) as functions of time as well as the state variables.

Mathematically, \(p(X(t)) = (p_i(X(t)) : i = 1, \cdots, N_a)^\top \in \mathbb{R}^{N_a}\) and \(\hat{p}(\hat{X}(t)) = (\hat{p}_i(\hat{X}(t)) : i = 1, \cdots, N_a)^\top \in \mathbb{R}^{N_a}\), where \(X(t) \in \mathcal{X} \subseteq \mathbb{R}^{N_s}\) and \(\hat{X}(t) \in \hat{\mathcal{X}} \subseteq \mathbb{R}^{N_s}\) are the state variables taken into account by the active portfolio and the benchmark portfolio respectively. Since the controls depend on time \(t\), here we include \(t\) in \(X(t)\) and \(\hat{X}(t)\) for notational simplicity. In this article, we consider the particular problem of outperforming a passive portfolio, in which \(X(t) = (t, W(t), \hat{W}(t))^\top\). Without further specification, we use \(X(t)\) to represent \((t, W(t), \hat{W}(t))^\top\) in the rest of the article.

We assume that the active portfolio and the benchmark portfolio follow the same rebalancing schedule denoted by \(T \subseteq [t_0, T]\). In the case of discrete rebalancing, \(T \subset [t_0, T]\) is a discrete set. In the case of continuous rebalancing, \(T = [t_0, T]\), 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 \(C = \{c(t), t \in T_c\}\), where \(T_c \subseteq [t_0, T]\) is the schedule of the cash injections. When \(T_c\) is a discrete injection schedule, \(c(t)\) is the amount of cash injection at \(t\). In the case of continuous cash injection, i.e., \(T = [t_0, T]\), \(c(t)\) 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 \(T_c = T\), so that the cash injections schedule is the same as the rebalancing schedule. At \(t \in T\), \(W(t)\) 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 strategy are defined by sets \(\mathcal{P} = \{p(X(t)), t \in T\}\), and \(\hat{\mathcal{P}} = \{\hat{p}(\hat{X}(t)), t \in T\}\).

Denote \(\mathcal{A}\) as the set of admissible strategies, which reflects the investment constraints on the controls. Mathematically, let \(\{\mathcal{X}_i : i = 1, \cdots, k\}\) be a partition of \(\mathcal{X}\) (the state variable space), i.e.

\[
\begin{align*}
\bigcup_{i=1}^k \mathcal{X}_i &= \mathcal{X}, \\
\mathcal{X}_i \cap \mathcal{X}_j &= \emptyset, \forall 1 \leq i < j \leq k,
\end{align*}
\]
and \( \{ Z_i \subseteq \mathbb{R}^N : i = 1, \cdots, k \} \) be the corresponding range of feasible controls such that any feasible control \( p \) satisfies
\[
p(x) \in Z_i, \forall x \in X_i, \forall i \in \{ 1, \cdots, k \}. \tag{3.3}
\]
We say that strategy \( \mathcal{P} \) is an admissible strategy, i.e., \( \mathcal{P} \in \mathcal{A} \), if and only if
\[
\mathcal{P} = \left\{ p(X(t)), t \in T \mid p(X(t)) \in Z_i, \text{ if } X(t) \in X_i \right\} \tag{3.4}
\]
Consider a discrete rebalancing schedule \( T = \{ t_j, j = 0, \cdots, N \} \) with \( N \) rebalancing events, where \( t_0 < t_1 < \cdots < t_N = T \). Then, the wealth evolution of the active portfolio and the benchmark portfolio can be described by the equations
\[
\begin{align*}
W(t_{j+1}) &= \left( \sum_{i=1}^{N} p_i(X(t_j)) \cdot \frac{S_i(t_{j+1})-S_i(t_j)}{S_i(t_j)} \right) W(t_j) + c(t_{j+1}), \quad j = 0, \cdots, N - 1, \\
\hat{W}(t_{j+1}) &= \left( \sum_{i=1}^{N} \hat{p}_i(X(t_j)) \cdot \frac{S_i(t_{j+1})-S_i(t_j)}{S_i(t_j)} \right) \hat{W}(t_j) + c(t_{j+1}), \quad j = 0, \cdots, N - 1.
\end{align*}
\tag{3.5}
\]
In the continuous rebalancing case, \( T = [t_0, T] \). Let \( dS_i(t) \) denote the instantaneous change in price for asset \( i, i \in \{ 1, \cdots, N \} \). Then, at \( t \in T = [t_0, T] \), 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
\[
\begin{align*}
dW(t) &= \left( \sum_{i=1}^{N} p_i(X(t)) \cdot \frac{dS_i(t)}{S_i(t)} \right) W(t_j) + c(t)dt, \\
d\hat{W}(t) &= \left( \sum_{i=1}^{N} \hat{p}_i(X(t)) \cdot \frac{dS_i(t)}{S_i(t)} \right) \hat{W}(t_j) + c(t)dt.
\end{align*}
\tag{3.6}
\]
Let sets \( \mathcal{W}_\mathcal{P} = \{ W(t), t \in T \} \) and \( \hat{\mathcal{W}}_\hat{\mathcal{P}} = \{ \hat{W}(t), t \in T \} \) 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 \( S(t) \in \mathbb{R}^N \) 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 via taking the expectation of the random metric.

Let \( \mathbb{E}^{(t_0, w_0)}_\mathcal{P}[F(\mathcal{W}_\mathcal{P}, \hat{\mathcal{W}}_\hat{\mathcal{P}})] \) 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 \( \hat{\mathcal{P}} \). The benchmark strategy is often pre-determined and known, hence we keep the benchmark strategy \( \hat{\mathcal{P}} \) implicit in this notation for simplicity. Going forward, we refer to \( \mathbb{E}^{(t_0, w_0)}_\mathcal{P}[F(\mathcal{W}_\mathcal{P}, \hat{\mathcal{W}}_\hat{\mathcal{P}})] \), the expectation of a desired performance metric, as the (investment) objective function.

Finally, we can mathematically define the investment objective (the optimization problem) associated with a general performance metric \( F \) as
\[
\text{(Optimization problem)} \quad \inf_{\mathcal{P} \in \mathcal{A}} \mathbb{E}^{(t_0, w_0)}_\mathcal{P}[F(\mathcal{W}_\mathcal{P}, \hat{\mathcal{W}}_\hat{\mathcal{P}})]. \tag{3.7}
\]

---

8Technically, 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.

9For illustration purposes, here we assume \( S_i(t), i \in \{ 1, \cdots, \} \) follow standard diffusion processes, i.e., no jumps. We will discuss the case with jumps in detail in Section 3.4.
### 3.3 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.,

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

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 ($W_P$ and $\hat{W}_P$) that are evaluated under the same investment horizon and same market 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., $\beta$, then the tracking difference metric may not be appropriate. To address this, van Staden et al. (2022) suggests the investment objective

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

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 $\beta$ is the annualized relative return target.

The optimal control problem (3.9) 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 $\beta$. 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 $\beta$ can be intuitively interpreted as the manager’s willingness to take more risks. As $\beta \downarrow 0$, the optimal solution to problem (3.9) is simply to mimic the benchmark strategy. However, as $\beta$ 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 (3.9) 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-type 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 (3.9), Ni et al. (2022) propose the following asymmetrical objective function,

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

The investment objective (3.10) 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 (3.10) does not permit closed-form solutions and requires machine learning techniques (Ni et al., 2022) to derive the desired optimal strategy numerically.
Another criticism of the investment objective \((3.9)\) and \((3.10)\) is that both are only concerned with the relative performance at terminal time \(T\). However, 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, it is more desirable for managers to consistently achieve the relative return target throughout the entire investment horizon. Therefore, it is often important for investment managers 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 entire investment horizon. For this reason, van Staden et al. (2022) propose the following cumulative quadratic tracking difference (CD) objectives

\[
(CD(\beta)) : \left\{ \begin{array}{ll}
\inf_{P \in A} E_P^{(t_0,w_0)} \left[ \int_{t_0}^{T} \left( W(t) - e^{\beta t} \hat{W}(t) \right)^2 dt \right], & \text{if } T = [t_0,T], \\
\inf_{P \in A} E_P^{(t_0,w_0)} \left[ \sum_{t \in T \cup \{T\}} \left( W(t) - e^{\beta t} \hat{W}(t) \right)^2 \right], & \text{if } T \subseteq [t_0,T], T \text{ discrete.}
\end{array} \right.
\tag{3.11}
\]

Here, note that objective \((3.11)\) is for the continuous rebalancing case, and \((3.12)\) for discrete rebalancing. Both \((3.11)\) and \((3.12)\) 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 \((3.9)\), \((3.11)\) and \((3.12)\) 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)

\[
(CS(\beta)) : \left\{ \begin{array}{ll}
\inf_{P \in A} E_P^{(t_0,w_0)} \left[ \int_{t_0}^{T} \left( \min (W(t) - e^{\beta t} \hat{W}(t), 0) \right)^2 dt + \epsilon W(T) \right], & \text{if } T = [t_0,T], \\
\inf_{P \in A} E_P^{(t_0,w_0)} \left[ \sum_{t \in T \cup \{T\}} \left( \min (W(t) - e^{\beta t} \hat{W}(t), 0) \right)^2 + \epsilon W(T) \right], & \text{if } T \subseteq [t_0,T], T \text{ discrete.}
\end{array} \right.
\tag{3.13}
\]

Here \((3.13)\) and \((3.14)\) are the investment objectives for the continuous rebalancing and discrete rebalancing cases respectively. \(\epsilon\) is a small regularization parameter to ensure that problem \((3.13)\) and \((3.14)\) are well-posed. A more detailed comparison of the CD and CS objective functions can be found in Appendix F.

### 3.4 Closed-form solution for CD problem

In this section, we present the closed-form solution to the CD problem \((3.11)\) under several assumptions. The closed-form solution not only provides us with insights for understanding the CD-optimal controls for problem \((3.11)\), 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 assets 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-Aradi 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 \((3.11)\).

**Assumption 3.1.** (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., \(T = [t_0,T]\). 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 3.1.** (Remark on Assumption 3.1) 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 4.1.1, 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 Section 4.1.1 we will discuss the impact of insolvency and its handling in experiments.

**Assumption 3.2.** (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 \( \hat{\varrho} \) 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 \( \mathcal{P} = \{ \hat{\varrho}X(t) = (\hat{\varrho}X(t), \hat{\varrho}2X(t)) \}^\top \equiv \hat{\varrho}, \forall t \in T \).

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

**Assumption 3.3.** (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

\[
dS_i(t) = \left( \mu_i - \lambda_i \kappa_i + r_i \cdot \mathbf{1}_{S_i(t^-) < 0} \right) dt + \sigma_i dZ_i(t) + d\left( \sum_{k=1}^{\pi_i(t)} \xi_i^{(k)} - 1 \right), \quad i = 1, 2. \tag{3.15}
\]

Here \( \mu_i \) are the (uncompensated) drift rate, \( \sigma_i \) 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. \( \pi_i(t) \) is a Poisson process with positive intensity parameter \( \lambda_i \). \( \{\xi_i^{(k)}, k = 1, \ldots, \pi_i(t)\} \) 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 \( S_i(t) = \xi_i \cdot S_i(t^-) \). \( \xi_i \) and \( \pi_i(t) \) are independent of each other. Moreover, \( \pi_1(t) \) and \( \pi_2(t) \) are assumed to be independent.

**Remark 3.2.** (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 (3.15), the wealth processes for the active portfolio and benchmark portfolio are

\[
\begin{align*}
\{dW(t) &= \left( \sum_{i=1}^{N_i} \pi_i(X(t^-)) \cdot \frac{dS_i(t)}{S_i(t^-)} \right) W(t^-) + \rho dt, \\
\{d\hat{W}(t) &= \left( \sum_{i=1}^{N_i} \hat{\pi}_i(X(t^-)) \cdot \frac{d\hat{S}_i(t)}{\hat{S}_i(t^-)} \right) \hat{W}(t^-) + \rho dt,
\end{align*}
\tag{3.16}
\]

\( ^{10} \)Intuitively, there is a premium for shorting an asset. In the closed-form solution derivation, we assume \( r_1 = 0 \).

\( ^{11} \)For any functional \( \psi(t) \), we use the notation \( \psi(t^-) \) as shorthand for the left-sided limit \( \lim_{\Delta t \downarrow 0} \psi(t - \Delta t) \).

\( ^{12} \)See 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.
where \( t \in (t_0, T] \), \( W(t_0) = \hat{W}(t_0) = w_0 \) and \( X(t^-) = (t, W(t^-), \hat{W}(t^-)) \in \mathbb{R}^3 \) is the state variable vector.

We now derive the closed-form solution of the CD problem (3.11) under Assumption 3.1, 3.2 and 3.3. 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 (3.11).

**Theorem 3.1.** (Verification theorem for CD problem (3.11)) 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 \( V(t, w, \hat{w}, \hat{\varrho}) : [t_0, T] \times \mathbb{R}^3 \to \mathbb{R} \) and \( p^*(t, w, \hat{w}, \hat{\varrho}) : [t_0, T] \times \mathbb{R}^3 \to \mathbb{R}^2 \) that satisfy the following two properties. (i) \( V \) and \( p^* \) are sufficiently smooth and solve the HJB PIDE (3.17), and (ii) the function \( p^*(t, w, \hat{w}, \hat{\varrho}) \) attains the pointwise infimum in (3.17) below

\[
\begin{align*}
\frac{\partial V}{\partial t} &+ (w - e^{\beta t} \hat{w})^2 + \inf_{p^* \in \mathbb{R}^2} H(p; t, w, \hat{w}, \hat{\varrho}) = 0, \\
V(T, w, \hat{w}, \hat{\varrho}) &= 0,
\end{align*}
\]  

(3.17)

where

\[
H(p; t, w, \hat{w}, \hat{\varrho}) = (w \cdot \alpha^\top p + c) \cdot \frac{\partial V}{\partial w} + (\hat{w} \cdot \alpha^\top \hat{\varrho} + c) \cdot \frac{\partial V}{\partial \hat{w}} - \left( \sum_i \lambda_i \right) \cdot V(t, w, \hat{w}, \hat{\varrho})
\]

\[+ \frac{w^2}{2} \cdot \left( p^\top \sigma \right) \cdot \frac{\partial^2 V}{\partial w^2} + \frac{\hat{w}^2}{2} \cdot \left( \hat{\varrho}^\top \sigma \hat{\varrho} \right) \cdot \frac{\partial^2 V}{\partial \hat{w}^2} + w \hat{w} \cdot (p^\top \sigma \hat{\varrho}) \cdot \frac{\partial^2 V}{\partial w \partial \hat{w}}
\]

\[+ \sum_i \lambda_i \int_0^\infty V(w + p_i(w) (\xi - 1), \hat{w} + \hat{p}_i \hat{w} (\xi - 1), t, \hat{\varrho}) f_\xi(\xi) d\xi.
\]  

(3.18)

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

Then, under Assumption 3.1, 3.2 and 3.3 \( V \) is the value function and \( p^* \) is the optimal control for the CD problem (3.11).

**Proof.** See Appendix C.1

---

Define several auxiliary variables

\[
\begin{align*}
\kappa_i^{(2)} &= \mathbb{E}[(\xi_i - 1)^2], \\
(\sigma_i^{(2)})^2 &= (\sigma_i)^2 + \lambda_i \kappa_i^{(2)}, i \in \{1, 2\}, \\
\varrho &= \sigma_1 \sigma_2 \rho - (\sigma_2^{(2)})^2, \\
\gamma &= (\sigma_1^{(2)})^2 + (\sigma_2^{(2)})^2 - 2 \sigma_1 \sigma_2 \rho, \\
\phi &= \frac{(\mu_1 - \mu_2) (\mu_1 - \mu_2 + \theta)}{\gamma}, \\
\eta &= \frac{(\mu_1 - \mu_2 + \theta)^2}{\gamma} - (\sigma_2^{(2)})^2,
\end{align*}
\]  

(3.19)

then we have the following proposition regarding the optimal control to problem (3.11).

**Proposition 3.1.** (CD-optimal control) Suppose Assumption 3.1, 3.2 and 3.3 are applicable, then the optimal control fraction of the wealth of the active portfolio to be invested in the stock index for the CD(\( \beta \)) problem (3.11) is given by \( p^*(t, w, \hat{w}, \hat{\varrho}) \in \mathbb{R} \), where

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

(3.20)

Here \( W^*(t) \) denotes the wealth process of the active portfolio from (3.6) following control \( p^*(t, W^*(t), \hat{W}(t), \hat{\varrho}) = (p^*(t, W^*(t), \hat{W}(t), \hat{\varrho}), 1 - p^*(t, W^*(t), \hat{W}(t), \hat{\varrho})) \), where \( p^* \) is the optimal stock allocation described in (3.20), and \( \hat{W} \) is the wealth process of the benchmark portfolio following the fixed-mixed strategy described in Assumption 3.2. Here, \( h \) and \( g \) are deterministic functions of time,

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

(3.21)

---
where $A, D$ and $B$ are deterministic functions defined as

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

and

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

(3.23)

Proof. See Appendix C.2.

3.4.1 Insights from CD-optimal control

The CD-optimal control (3.20) 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 3.1. (Properties of $g(t; \beta)$) The function $g(t; \beta)$ defined in (3.21) 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:

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

(3.24)

Proof. See Appendix C.3.

Corollary 3.2. (Properties of $h(t; \beta, c)$) The function $h(t; \beta, c)$ defined in (3.21) 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 C.3.

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

Assumption 3.4. (Drift rates of the two assets) We assume that the drift rates of the stock and the bond index $\mu_1$ and $\mu_2$ satisfy the following properties,

$$\mu_1 - \mu_2 > 0, \quad \mu_1 - \mu_2 + \vartheta > 0,$$

where $\vartheta$ is defined in (3.19).

Remark 3.3. (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. The second inequality $\mu_1 - \mu_2 + \vartheta > 0$ is also practically reasonable. $\vartheta$ 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 $\mu_1 - \mu_2 + \vartheta \leq 0$.  

\[13\] In fact, in this two-asset case, this assumption does not cause loss of generality.

\[14\] For reference, based on the calibrated jump-diffusion model (3.15) 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 \((3.20)\). 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 \((3.20)\), we can see that the CD-optimal control consists of two components: a cash injection component \(p_{\text{cash}}^*\) and a tracking component \(p_{\text{track}}^*\). Mathematically,

\[
P^*(t, w, \hat{w}; \hat{\varrho}) = p_{\text{cash}}^*(t, w, \hat{w}) + p_{\text{track}}^*(t, w, \hat{w}, \hat{\varrho}),
\]

where

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

Based on Assumption \(3.4\) and Corollary \(3.2\), the cash injection component \(p_{\text{cash}}^*\) is always non-negative. Furthermore, from Corollary \(3.2\) 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 \(p_{\text{track}}^*\) does not depend on the cash injection rate \(c\), but only concerns the tracking performance of the active portfolio. One key finding is that

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

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 \((3.11)\) 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 \(3.1\) 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.

### 3.5 Leverage constraints

In practice, large pension funds such as the Canadian Pension Plan often have exposures to alternative assets, such as private equity. 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 \(3.2\), 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 \(\{1, \cdots, N_l\}\). 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 \( \{ N_l + 1, \cdots, N_a \} \). 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

\[
\text{(Long-only constraint):} \quad p_i(\mathbf{X}(t)) \geq 0, \; i \in \{1, \cdots, N_l\}, \; t \in \mathcal{T}.. \tag{3.29}
\]

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

\[
\text{(Summation constraint):} \quad \sum_{i=1}^{N_a} p_i(\mathbf{X}(t)) = 1, \; t \in \mathcal{T}. \tag{3.30}
\]

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,

\[
\text{(Maximum leverage constraint):} \quad \sum_{i=1}^{N_l} p_i(\mathbf{X}(t)) \leq p_{\max}, \; t \in \mathcal{T}. \tag{3.31}
\]

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

**Assumption 3.5.** (Simultaneous shorting) If one shortable asset has a negative weight, other shortable assets must have nonpositive weights. Mathematically, this assumption can be expressed as

\[
\text{(Simultaneous shorting constraint):} \quad \begin{cases} 
p_i(\mathbf{X}(t)) \leq 0, & \forall i \in \{ N_l + 1, \cdots, N_a\}, \text{ if } \sum_{i=1}^{N_l} p_i(\mathbf{X}(t)) > 1, \; t \in \mathcal{T} \\
p_i(\mathbf{X}(t)) \geq 0, & \forall i \in \{ N_l + 1, \cdots, N_a\}, \text{ if } \sum_{i=1}^{N_l} p_i(\mathbf{X}(t)) \leq 1, \; t \in \mathcal{T}.
\end{cases} \tag{3.32}
\]

**Remark 3.4.** (Remark on Assumption 3.5) 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 3.1.** (Partition of state space) We define \( \{ \mathcal{X}_1, \mathcal{X}_2 \} \) to be a partition of the state space \( \mathcal{X} \), such that

\[
\mathcal{X}_1 = \left\{ x = (t, \mathbf{W}(t), \dot{\mathbf{W}}(t))^\top \in \mathcal{X} | \mathbf{W}(t) \geq 0 \right\},
\]

\[
\mathcal{X}_2 = \left\{ x = (t, \mathbf{W}(t), \dot{\mathbf{W}}(t))^\top \in \mathcal{X} | \mathbf{W}(t) < 0 \right\}. \tag{3.33}
\]

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 3.6.** (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 \( \mathbf{e}_i \in \mathbb{R}^{N_a} = (0, \cdots, 0, 1, 0, \cdots, 0)^\top \) 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.

\[
\text{(No trading in insolvency):} \quad p(\mathbf{X}(t)) = \mathbf{e}_{N_l+1}, \; \text{if } \mathbf{X}(t) \in \mathcal{X}_2. \tag{3.34}
\]
Remark 3.5. (Remark on Assumption 3.6) 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, the corresponding space of feasible control vector values $Z$ and the admissible strategy set $A$ are

\[
(\text{Admissible set}) : \left\{ \begin{array}{c}
Z = \{ z \in \mathbb{R}^{N_a} \} \\
A = \{ P(X(t)), t \in T \} \end{array} \right. \\
\begin{align}
&z_i \geq 0, \forall i \in \{1, \cdots, N\}, \\
&\sum_{i=1}^{N_a} z_i = 1, \\
&\sum_{i=1}^{N_a} z_i \leq p_{max}, \\
&z_i \leq 0, \forall i \in \{N_i + 1, \cdots, N_a\}, \text{ if } \sum_{i=1}^{N_i} z_i > 1, \\
&z_i \geq 0, \forall i \in \{N_i + 1, \cdots, N_a\}, \text{ if } \sum_{i=1}^{N_i} z_i \leq 1
\end{align}\]

Note that the conditional constraints in (3.35) and (3.36) cannot be formulated into a standard constrained optimization problem.

### 3.6 Neural network method

#### 3.6.1 Motivation for neural network method

In Section 3.4, we derive the closed-form solution under the jump-diffusion model, by obtaining the first-order optimality of the corresponding PIDE of the problem. The closed-form solution 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 (3.11). Ideally, we would like to solve the general investment problem (3.7) under realistic constraints, such as discrete rebalancing and limited leverage (i.e., leverage constraints discussed in Section 3.5). Unfortunately, for such cases, closed-form solutions rarely exist. Therefore, we often need to solve the problem numerically.

One method of solving a discrete-time multi-period optimal asset allocation problem utilizes dynamic programming (DP) based reinforcement learning (RL) algorithms. For example, Dixon et al. (2020); Park et al. (2020); Lucarelli and Borroto (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 $(2^{N_a} + 1)$-dimensional function which represents the conditional expectation of the cumulative rewards at an intermediate state. 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, 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 has been termed 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
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.

Mathematically, the numerical solution to the general problem (3.7) requires solving for the feedback control $p$. We approximate the control function $\theta$ by a neural network function $f(X(t); \theta) : X \mapsto \mathbb{R}^N$, where $\theta \in \mathbb{R}^N$ represents the parameters of the neural network (i.e., weights and biases). In other words,

$$p(X(t)) \approx f(X(t); \theta) \equiv f(\cdot; \theta).$$

Then, the optimization problem (3.7) can be converted to solving the following optimization problem.

$$(\text{Parameterized optimization problem}): \inf_{\theta \in \mathcal{Z}_\theta} \mathbb{E}_{f(\cdot; \theta)}[F(W_\theta, \hat{W}_P)].$$

Here $W_\theta$ is the wealth trajectory of the active portfolio following the neural network approximation function parameterized by $\theta$. $\mathcal{Z}_\theta \subseteq \mathbb{R}^N$ is the feasibility domain of the parameter $\theta$, which is translated from the constraints of the original problem, e.g., (3.35) and (3.36). Mathematically,

$$\mathcal{Z}_\theta = \left\{ \theta : \begin{array}{ll}
  f(X(t); \theta) \in \mathcal{Z}, & \text{if } X(t) \in \mathcal{X}_1, \\
  f(X(t); \theta) \in \{e_{N_i+1}\}, & \text{if } X(t) \in \mathcal{X}_2.
\end{array} \right\}$$ (3.39)

Here $\mathcal{Z}$ is defined in (3.35) and $\{\mathcal{X}_1, \mathcal{X}_2\}$ is the partition of the state space $X$ defined in Definition 3.1.

Note here that $\mathcal{Z}_\theta$ is dependent on the structure of the neural network function $f(\cdot; \theta)$. Intuitively, $\mathcal{Z}_\theta$ is the preimage of $\mathcal{Z}$ and $\{e_{N_i+1}\}$ under $f(\cdot; \theta)$. Specific neural network model design may result in $\mathcal{Z}_\theta = \mathbb{R}^N$, which means (3.38) becomes an unconstrained optimization problem. For long-only investment problems, the only constraints are the long-only constraint (3.29) and the summation constraint (3.30). 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 $\mathcal{Z}_\theta = \mathbb{R}^N$ and problem (3.38) becomes an unconstrained optimization problem (see, e.g., Li and Forsyth 2019; Ni et al. 2022). However, as discussed in Section 3.5 we consider the more complicated case where leverage and shorting is allowed. The problem thus involves more constraints than the long-only case does and therefore requires 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 (3.12) and CS problem (3.14), 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.

### 3.7 Leverage-feasible neural network (LFNN)

In this section, we demonstrate the leverage-feasible neural network (LFNN) model design, which yields $\mathcal{Z}_\theta = \mathbb{R}^N$ for leverage constraints defined in equation (3.35), and converts a constrained optimization problem (3.38) to an unconstrained problem.
Let vector $x = (t, W(t), \dot{W}(t))^T \in \mathcal{X}$ be the feature (input) vector. We first define a standard fully-connected feedforward neural network (FNN) function $f : \mathcal{X} \mapsto \mathbb{R}^{N_a+1}$ as follows:

$$(\text{FNN}) : \begin{cases} h_j^{(1)} = \text{Sigmoid} \left( \sum_{i=1}^{N_t} x_i \theta_{ij}^{(1)} + b_j^{(1)} \right), & j = 1, \ldots, N_h^{(1)}, \\
    h_j^{(k)} = \text{Sigmoid} \left( \sum_{i=1}^{N_h^{(k-1)}} h_i^{(k-1)} \theta_{ij}^{(k)} + b_j^{(k)} \right), & j = 1, \ldots, N_h^{(k)}, \quad \forall k \in \{2, \ldots, K\}, \\
    o_j = \sum_{i=1}^{N_h^{(K)}} h_i \theta_{ij}^{(K+1)}, & j = 1, \ldots, N_a + 1, \\
    \hat{f}(x; \theta) := (o_1, \ldots, o_{N_a+1})^T. \end{cases}$$

(3.40)

Here Sigmoid(·) 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. $\theta^{(k)} = (\theta_{ij}^{(k)}) \in \mathbb{R}^{N_h^{(k)} \times N_h^{(k-1)}}$ and $b^{(k)} = (b_j^{(k)}) \in \mathbb{R}^{N_h^{(k)}}$ are the (vectorized) weight matrix and bias vector for the $k$-th layer and the final parameter vector of the neural network is $\theta = (\theta^{(1)}, b^{(1)}, \ldots, \theta^{(K)}, b^{(K)}, \theta^{(K+1)})^T \in \mathbb{R}^{N_{\theta}}$, where $N_{\theta} = \sum_{k=1}^{K+1} N_h^{(k)} \cdot N_h^{(k-1)} + \sum_{k=1}^{K} N_h^{(k)}$.

Building upon $\hat{f}$, we propose the following **leverage-feasible neural network** (LFNN) model $f : \mathcal{X} \mapsto \mathbb{R}^{N_a}$:

$$(\text{LFNN}) : \ f(x; \theta) := \psi \left( \hat{f}(x; \theta), x \right) \in \mathbb{R}^{N_a}.$$ (3.41)

Here, $\psi(\cdot)$ is the **leverage-feasible activation function**. For $o = (o_1, \ldots, o_{N_a+1})^T \in \mathbb{R}^{N_a+1}$ and $p = (p_1, \ldots, p_{N_a})^T \in \mathbb{R}^{N_a}$, $\psi : (o, x) \in \mathbb{R}^{N_a+1} \times \mathcal{X} \mapsto p \in \mathbb{R}^{N_a}$ is defined by

$$
\psi(o, x) = p = \begin{cases} \ l = p_{\max} \cdot \text{Sigmoid}(o_{N_a+1}), \\
    p_i = l \cdot \frac{e^{o_i}}{\sum_{k=1}^{N_a+1} e^{o_k}}, & i \in \{1, \ldots, N_l\}, \\
    p_i = (1 - l) \cdot \frac{e^{o_i}}{\sum_{k=1}^{N_a+1} e^{o_k}}, & i \in \{N_l+1, \ldots, N_a\}, \\
    e_{N_l+1}, & \text{if } x \in \mathcal{X}_2. \end{cases}$$ (3.42)

Note 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 $\psi$ has the following property.

**Lemma 3.1.** (Decomposition of $\psi$) The leverage-feasible function $\psi$ defined in (3.42) has the function decomposition that

$$\psi(o, x) = \varphi(\zeta(o), x),$$ (3.43)

where

$$
\zeta : \mathbb{R}^{N_a+1} \mapsto \hat{Z}, \zeta(o) = \left( \text{Softmax} \left( (o_1, \ldots, o_{N_l}) \right), \text{Softmax} \left( (o_{N_l+1}, \ldots, o_{N_a}) \right), p_{\max} \cdot \text{Sigmoid}(o_{N_a+1}) \right)^T,
\varphi : \hat{Z} \times \mathcal{X} \mapsto \mathbb{R}^{N_a}, \varphi(z) = \left( z_{N_a+1} \cdot (z_1, \ldots, z_{N_l}), (1 - z_{N_a+1}) \cdot (z_{N_l+1}, \ldots, z_{N_a}) \right)^T \cdot I_{x \in \mathcal{X}_1} + e_{N_l+1} \cdot I_{x \in \mathcal{X}_2},
$$

and

$$\hat{Z} = \left\{ z \in \mathbb{R}^{N_a+1} : \sum_{i=1}^{N_l} z_i = 1, \sum_{i=N_l+1}^{N_a} z_i = 1, z_{N_a+1} \leq p_{\max}, z_i \geq 0, \forall i \right\}.$$ (3.44)

**Proof.** This is easily verifiable by definition of $\psi$ in (3.42). \hfill $\Box$

**Remark 3.6.** (Remark on Lemma 3.1) The leverage-feasible activation function $\psi$ corresponds to a two-step decision process described by $\zeta$ and $\varphi$. Intuitively, $\zeta$ first determines the internal allocations within long-only assets and shortable assets, as well as the total leverage. Then, $\varphi$ converts the internal allocations and total leverage into final allocation fractions, which depend on the wealth of the active portfolio.

\footnote{For $k = K+1$, define $N_h^{(K+1)} = N_a + 1$ so $\theta^{(K+1)}$ is well-defined.}
With the LFNN model outlined above, the parameterized optimization problem (3.38) becomes an unconstrained optimization problem. Specifically, we present the following theorem regarding the feasibility domain $\mathcal{Z}_\theta$ associated with the LFNN model (3.41).

**Theorem 3.2.** (Unconstrained feasibility domain) The feasibility domain $\mathcal{Z}_\theta$ defined in (3.39) associated with the LFNN model (3.41) is $\mathbb{R}^{N_\theta}$.

Proof. See Appendix D.1.

Following Theorem 3.2, the constrained optimization problem (3.7) can be transformed into the following unconstrained optimization problem

(Unconstrained parameterized problem) $\inf_{\theta \in \mathbb{R}^{N_\theta}} \mathbb{E}^{(t_0, w_0)} \left[ F(W_\theta, \hat{W}) \right]$. (3.46)

### 3.8 Mathematical justification for LFNN approach

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

However, an important question remains: does solving the parameterized unconstrained optimization problem (3.46) yield the optimal control of the original problem (3.7)? In other words, suppose $\theta^*$ is the solution to (3.46), is $f(\cdot; \theta^*)$ the solution to (3.7)?

In this section, we prove that under benign assumptions and appropriate choices of the hyperparameter of the LFNN model (3.41), solving the unconstrained problem (3.46) provides an arbitrarily close approximation to the original problem (3.7). We start by establishing the following lemma.

**Lemma 3.2.** (Structure of feasible control) Any feasible control function $p : \mathcal{X} \mapsto \mathbb{R}^{N_\alpha}$ has the function decomposition

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

where $\varphi : \hat{\mathcal{Z}} \times \mathcal{X} \mapsto \mathbb{R}^{N_\alpha}$ is defined in (3.44) and $\omega : \mathcal{X} \mapsto \hat{\mathcal{Z}}$.

Proof. See Appendix D.2.

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

**Assumption 3.7.** (Assumption on state space and optimal control)

(i) The space of state variables $\mathcal{X}$ is a compact set.

(ii) Following Lemma 3.2, the optimal control $p^* : \mathcal{X} \mapsto \mathbb{R}^{N_\alpha}$ has the decomposition $p^*(x) = \varphi(\omega^*(x), x)$ for some $\omega^* : \mathcal{X} \mapsto \hat{\mathcal{Z}}$. We assume $\omega^* \in C(\mathcal{X}, \hat{\mathcal{Z}})$.

**Remark 3.7.** (Remark on Assumption 3.7) In our particular problem of outperforming a benchmark portfolio, the state variable vector is $X(t) = (t, W(t), \hat{W}(t))^T \in \mathcal{X}$ 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 are bounded, i.e. $\mathcal{X} = [0, T] \times [w_{\text{min}}, w_{\text{max}}] \times [\hat{w}_{\text{min}}, \hat{w}_{\text{max}}]$, where $w_{\text{min}}, w_{\text{max}}$ and $\hat{w}_{\text{min}}, \hat{w}_{\text{max}}$ 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 are a continuous function of state variables.

Finally, we present the following theorem.
Theorem 3.3. (Approximation of optimal control) Following Assumption 3.7, \( \forall \epsilon > 0 \), there exists \( N_h \in \mathbb{N} \), and \( \theta \in \mathbb{R}^{N_\theta} \) such that the corresponding LFNN model \( f(\cdot; \theta) \) described in (3.41) satisfies the following:

\[
\sup_{x \in X} \| f(x; \theta) - p^*(x) \| < \epsilon.
\]  

(3.48)

Proof. See Appendix D.2.

Theorem 3.3 shows that given any arbitrarily small distance, there exists a suitable choice of the hyper-parameter of the LFNN model (e.g. number of hidden layers and nodes), and a parameter vector \( \theta \), such that the corresponding parameterized LFNN function is within this distance of the optimal control function. In other words, with a large enough LFNN model (in terms of number of hidden nodes), solving the unconstrained parameterized problem (3.46) approximately solves the original optimization problem (3.7) with any required precision.

4 Numerical experiments

In this section, we discuss several numerical experiments and their results. Since the numerical experiments involve the solution and evaluation of the optimal parameters \( \theta^* \) of the LFNN model (3.41) in problem (3.46), we briefly review how the parameters are computed in experiments.

In numerical experiments, the expectation in (3.46) is approximated by using a finite set of samples of the set \( Y = \{ Y^{(j)} : j = 1, \cdots, N_d \} \), where \( N_d \) is the number of samples, and \( Y^{(j)} \) represents a time series sample of joint asset return observations \( R_i(t), i \in \{1, \cdots, N_a\} \), observed at \( t \in T \) \(^{19}\). Mathematically, problem (3.46) is approximated by

\[
\inf_{\theta \in \mathbb{R}^{N_\theta}} \left\{ \frac{1}{N_d} \sum_{j=1}^{N_d} F(W^{(j)}_{\theta}, \hat{W}^{(j)}_p) \right\}.
\]  

(4.1)

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

Based on our experiments, we find that a shallow neural network model is capable of achieving good results. Specifically, we use 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 \( (t, W_\theta(t), \hat{W}(t))^T \) as the input (feature) to the LFNN network. At \( t \in [t_0, T] \), \( W_\theta(t) \) is the wealth of the active portfolio of the strategy that follows the LFNN model parameterized by \( \theta \), and \( \hat{W}(t) \) is the wealth of the benchmark portfolio.

Then, the optimal parameter \( \theta^* \) can be numerically obtained by solving problem (4.1) using standard optimization algorithms such as ADAM \(^{18}\). This process is commonly referred to as “training” of the neural network model, and \( Y \) is often referred to as the training data set \(^{20}\).

Once \( \theta^* \) is numerically obtained, the resulting optimal strategy \( f(\cdot; \theta^*) \) is evaluated on a separate “testing” data set \( Y^{test} \), 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 \( f(\cdot; \theta^*) \) is assessed.

4.1 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 (3.20), and provide empirical validation of the LFNN approach.

\(^{18}\)The distance is defined in (3.48), i.e. the supremum of the pointwise distance over the extended state space \( X \).

\(^{19}\)Note that the corresponding set of asset prices can be easily inferred from the set of asset returns, or vice versa.
4.1.1 Approximate form under realistic assumptions

We first note that the closed-form solution $p^*$ defined in (3.20) is obtained under several unrealistic assumptions, namely continuous rebalancing, unlimited leverage, and continuing trading in insolvency. 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 $T_{\Delta t}$ defined as

$$T_{\Delta t} = \left\{ t_i : i = 0, \ldots, N \right\},$$

(4.2)

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

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

(4.3)

Here $[\min p, \max p > 1]$ is the allowed leverage 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 $t_i$, $\hat{\bar{W}}_{\Delta t}(t_i)$ is the wealth of the benchmark portfolio at $t_i$ following the fixed-mix strategy described by constant allocation fraction $\hat{\varrho}$, but only rebalanced discretely according to $T_{\Delta t}$. Clearly, the allocation strategy from $\bar{p}_{\Delta t}$ follows the discrete schedule of $T_{\Delta t}$, and satisfies the leverage constraint that $\bar{p}_{\Delta t} \in [\min p, \max p]$. $p_{\Delta t}$ approaches the closed-form solution $p^*$ as $\Delta t \downarrow 0$, $\min p \downarrow -\infty$ and $\max p \uparrow \infty$. 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 $(\min p, \max p)$ is a feasible, but in general sub-optimal, control of the leverage-constrained CD problem (3.12).

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 become negative [Wang and Forsyth (2012)]. However, with discrete rebalancing, even with a bounded control, as long as the upper bound $\max p > 1$, 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 3.6) 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.

4.1.2 Performance of neural network strategy

We assess and compare the performance of the neural network strategy and the clipped form, we assume the following investment scenario described in Table 4.1.

We assume the stock index and the bond index prices follow the jump-diffusion model (3.15). Furthermore, we assume that the prices follow a double exponential jump model [Kou (2002), Kou and Wang (2004)].

20 Note 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 $p^*$, since $p^* = (p^*, 1 - p^*)'$.

21 We remind the reader that the closed-form solution is derived under the jump-diffusion model.
Investment horizon $T$ (years) 10

Assets
CRSP cap-weighted index (real)/30-day T-bill (U.S.) (real)

Index Samples

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 $\beta$ 1% (100 bps)

Table 4.1: Investment scenario.

i.e., for the jump variable $\xi_t$, $y_i = \log(\xi_i)$ follows the double exponential distribution with density functions $g_i(y_i)$ defined as follows

$$g_i(y_i) = \nu_i \iota_i e^{-\iota_i y_i} 1_{y_i \geq 0} + (1 - \nu_i) \varsigma_i e^{-\varsigma_i y_i} 1_{y_i < 0}, \ i = 1, 2. \ (4.4)$$

where $\nu_i$ is the probability for an upward jump, and $\iota_i$ and $\varsigma_i$ 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 is an appropriate assumption for inflation regimes.

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 2. The calibrated parameters can be found in Appendix E. Then, we sample a training data set $Y$ and a testing data set $Y_{\text{test}}$ from the calibrated model, each with 10,000 samples.

The neural network strategy follows the LFNN model learned on $Y$. We then evaluate the performance of the neural network strategy and the approximate form (4.3) on the testing data set $Y_{\text{test}}$. Specifically, we compare the value of the CD objective function (3.12) for the neural network strategy and the clipped form on $Y_{\text{test}}$. In particular, this training/testing process is repeated for various rebalancing frequencies from monthly to annually, as described in Table 4.1.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>$\Delta t = 1$</th>
<th>$\Delta t = 1/2$</th>
<th>$\Delta t = 1/4$</th>
<th>$\Delta t = 1/12$</th>
<th>$\Delta t = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clipped form</td>
<td>545</td>
<td>504</td>
<td>479</td>
<td>467</td>
<td>461 (extrapolated)</td>
</tr>
<tr>
<td>Neural network</td>
<td>537</td>
<td>498</td>
<td>476</td>
<td>464</td>
<td>458 (extrapolated)</td>
</tr>
<tr>
<td>Closed-form solution</td>
<td>418</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: CD objective function values. Results shown are evaluated on $Y_{\text{test}}$, the lower the better.

In Table 4.2, 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 (3.11) 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 the unconstrained closed-form solution.

Finally, we compare the allocation strategies from the neural network strategy with the clipped form. Specifically, in Figure 4.1, 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 $e^{\beta t} W(t)$. For simplicity, we call this ratio the “tracking ratio”. We
plot the 3-tuple \( \left( \frac{W(t)}{e^{\beta t} \hat{W}(t)}, t, \tilde{p}_1(W(t), \hat{W}(t), t) \right) \) (obtained from the evaluation of the strategies on samples from \( Y_{test} \)) 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 \( \tilde{p}_1(W(t), \hat{W}(t), t) \) 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 4.1, 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 3.4.1. 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.

4.2 Case study under realistic market scenario

4.2.1 Learning from bootstrap resampled data

So far, our discussion has focused on scenarios in which asset prices follow a parametric model (specifically, the jump-diffusion model (3.15)). While the assumption of parametric models brings many benefits, most notably allowing us to derive the closed-form solution, we also acknowledge that the assumption of the parametric model has several limitations.

However, the parameter estimation for a parametric model is very difficult (Black, 1993). Even for a simple geometric Brownian motion (GBM) model, estimating the drift rate accurately is error prone and requires the data to cover a long historical period (Brigo et al., 2008). More complicated models, such as the double exponential jump-diffusion model we have discussed in detail, introduce more components to the stochastic model and thus require estimation of the extra parameters. Furthermore, parametric models
implicitly introduce many assumptions concerning the true stochastic model for asset prices, which are sometimes debatable.

Acknowledging the above limitations of parametric market data models, we turn to the non-parametric method of bootstrap resampling as a data-generating process for numerical experiments. Unlike parametric models, non-parametric models such as bootstrap resampling aim to 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 Cavaglià, 2017; Shahzad et al., 2019; Cavaglià et al., 2022; Simonian and Martirosyan, 2022) as well as academics (Anarkulova et al., 2022).

As we have discussed in Section 2.2, we bootstrap from inflation-adjusted returns from the concatenated series consisting of two identified high-inflation regimes: 1940:8-1951:7 and 1968:9-1985:10. In the subsequent numerical experiments, we use the stationary block bootstrap resampling (see Appendix B.1) to generate the training and testing data sets. The stationary block bootstrap resampling method is a special form of block bootstrap resampling that preserves the weak stationarity of the original time series data (Politis and Romano, 1994), while taking into account serial correlation effects in the time series.

4.2.2 Experiment results

The details of the investment specification are given in Table 4.3. 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 benchmark portfolio is rebalanced to maintain the weight of 70% in the equal-weighted stock index and 30% in the 30-day U.S. T-bill index. We select this fixed-mix portfolio based on our observation that the equal-weighted stock index shows superior performance to the cap-weighted stock index during high-inflation environments, as discussed in Section 2.2.

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 (see Appendix C for a more detailed discussion). The annual outperformance target $β$ is set to be 2% (i.e., 200 bps per year).

We use the bootstrap resampling algorithm (see Appendix B.1) to generate a training data set $Y$ and a testing data set $Y^{test}$ (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 $Y^{test}$ is generated using a different random seed as the training data set $Y$, and thus the probability of seeing the same sample in $Y$ and $Y^{test}$ is near zero (see Ni et al. (2022) for proof).

It is worth noting that in this experiment, we train the LFNN model (3.41) on $Y$ under the discrete-time CS objective (F.1), instead of the CD objective (3.12). As discussed in Section 3.3, the CS objective function only penalizes underperformance relative to the elevated target. Numerical comparisons of the two objective

---

22 After all, it is not interesting to try to beat a weak benchmark, such as a 70% cap-weighted stock index/30% bond benchmark.
functions suggest that the CS objective function indeed yields more favorable investment results than the CD objective (see Appendix F).

<table>
<thead>
<tr>
<th>Investment horizon $T$ (years)</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equity market indexes</td>
<td>CRSP cap-weighted/equal-weighted index (real)</td>
</tr>
<tr>
<td>Bond index</td>
<td>CRSP 30-day/10-year U.S. treasury index (real)</td>
</tr>
<tr>
<td>Initial portfolio wealth/annual cash injection</td>
<td>100/10</td>
</tr>
<tr>
<td>Rebalancing frequency</td>
<td>Monthly</td>
</tr>
<tr>
<td>Maximum leverage</td>
<td>1.3</td>
</tr>
<tr>
<td>Outperformance target rate $\beta$</td>
<td>2%</td>
</tr>
</tbody>
</table>

Table 4.3: Investment scenario.

Figure 4.2: 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 $Y^{test}$.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural network</td>
<td>364.2</td>
<td>403.4</td>
<td>211.8</td>
<td>136.3</td>
<td>0.078</td>
</tr>
<tr>
<td>Benchmark</td>
<td>308.5</td>
<td>342.9</td>
<td>165.0</td>
<td>149.0</td>
<td>0.056</td>
</tr>
</tbody>
</table>

Table 4.4: Statistics of strategies. Results are based on the evaluation results on the testing data set.

From Figure 4.2a we observe that, on average, the neural network strategy consistently outperforms the benchmark in terms of the wealth ratio $W(t)/\hat{W}(t)$. The mean and median wealth ratio increases smoothly and consistently over time. In terms of tail performance (20th percentile), the neural network strategy initially trails behind the benchmark, but eventually recovers over time, and ends up with 10% greater wealth at the terminal time. This shows that the neural network strategy manages tail risk well.

Another metric that might be of great interest to managers is the distribution of the terminal wealth ratio $\frac{W(T)}{\hat{W}(T)}$. This metric is concerned with the relative performance of the strategy at the end of the investment period. We can see from Figure 4.2b that there is a more than 90% chance that the neural network strategy outperforms the benchmark strategy in terms of terminal wealth. This is a surprisingly good result, considering that the terminal wealth ratio is not directly targeted in the objective function (F.1).
Since there are constant cash injections in the portfolios, it may be suitable to use the internal rate of return (IRR) to measure the annualized performance of the portfolio. As we can see from Figure 4.2c, the neural network strategy has a more than 90% chance of producing a higher IRR. The median IRR of the neural network strategy is slightly more than 2% higher than the median IRR of the benchmark strategy, which is in line with the choice of the target outperformance rate $\beta = 0.02$. This indicates that the neural network model achieves the desired target performance in most of the outcomes.

We remark that the neural network has a worse extreme left tail compared to the benchmark strategy. As listed in Table 4.4, the 5th percentile of the terminal wealth of the neural network strategy is lower than that of the benchmark strategy. Indeed, as we will see shortly, the neural network strategy takes more risk than the benchmark strategy, as it allocates a higher fraction of wealth to the equal-weighted stock index, the riskier asset, than the benchmark portfolio. Therefore, in a persistent bear market in which the stocks perform badly, the neural network strategy suffers more than the benchmark strategy. However, such scenarios happen with a low probability. As shown in Figure 4.2b, the neural network strategy outperforms the benchmark in terms of terminal value with an extremely high probability of over 90%.

At this point, readers may wonder how the neural network strategy performs asset allocation to achieve such good results. Below we present the allocation strategies for the neural network strategy.

![Figure 4.3: Mean allocation fraction over time, evaluated on $Y^{test}$](image)

We first plot the mean allocation fraction for the four assets over time in Figure 4.3. From Figure 4.3, the first interesting observation is that the neural network strategy does not allocate wealth to the cap-weighted stock index on average. While seemingly surprising at first, this is actually in line with the fact that the equal-weighted stock index has significantly higher real returns than the cap-weighted stock index during historical high-inflation periods (see Appendix A.3). Since the objective is to outperform a benchmark strategy heavily invested in the equal-weighted index (70%), it does not make sense to allocate wealth to a weaker index in the active portfolio.

The second observation from Figure 4.3 concerns the evolution of the mean bond allocation fractions. As we can see from Figure 4.3, on average, the neural network strategy shorts the 30-day T-bill index in the first two years. During the same period, the strategy takes on some leverage and invests heavily in the equal-weighted stock index. Intuitively speaking, such behavior indicates that the neural network strategy takes risks early on in order to establish an advantage over the benchmark strategy. After the first two years, the allocation to the 10-year T-bond decreases, and the allocation to the 30-day T-bill ramps up. The decrease in the 10-year bond allocation coincides with the decrease in the allocation to the equal-weighted index, indicating that the initial allocation to the T-bond was simply due to the need for leveraging - we assume that the manager can only short the T-bill for leveraging, thus making the 10-year bond the only
defensive asset available during the first two years. As we can see from Figure 4.3, once leverage is not needed in later years, the neural network strategy picks the T-bill over the 10-year bond.

Overall, while the stock allocation eventually decreases over time, the neural network strategy maintains more than 80% allocation to the equal-weighted stock index on average. This is not surprising, since beating an aggressive benchmark with 70% allocation to the equal-weighted stock index requires taking risk. Despite having a higher allocation to riskier assets, the neural network strategy achieves good results compared to the benchmark strategy, as shown in Figure 4.2.

Finally, we note that the neural network strategy learned under high-inflation regimes performs very well, somewhat surprisingly, on low-inflation testing data sets, demonstrating the robustness of the strategy. Interested readers may find more discussion in Appendix H.

5 Conclusion

In this paper, we studied a leverage-constrained multi-period asset allocation problem during high inflation periods.

We first identified the high inflation periods from historical data using a moving-window filtering algorithm. To gain intuition about high-inflation investing, we tested the passive fixed-mix strategies on the bootstrap resampled asset returns from the identified inflation regimes. We discovered empirically that passive strategies with the equal-weighted stock index partially stochastically dominate passive strategies with the cap-weighted stock index.

To mimic the behavior of large sovereign wealth funds, we then studied leverage-constrained multi-period asset allocation problem with the objective of outperforming a passive benchmark strategy.

Under the assumption that both asset prices follow jump-diffusion models, we derived a closed-form solution for a two-asset case under the cumulative tracking difference (CD) objective function. In order to obtain a closed-form solution, additional assumptions such as continuous rebalancing, unlimited leverage, and continued trading in insolvency are required. The closed-form solution provides useful insight for understanding the behavior of the optimal control. Notably, in order to track the elevated target, the optimal control needs to aim higher than the target when making allocation decisions.

To address the unrealistic assumptions (e.g., continuous rebalancing, unlimited leverage, trading if insolvent) required by the closed-form solution, we then proposed a novel leverage-feasible neural network (LFNN) model that approximates the optimal control directly. The direct approximation of a low-dimensional control bypasses the need for high-dimensional approximation of conditional expectations (which is required in dynamic programming approaches). In addition, the LFNN model converts the leverage-constrained optimization problem into an unconstrained optimization problem. Furthermore, we justify the validity of the LFNN approach by proving that solution to the parameterized unconstrained optimization problem can approximate the solution to the original constrained optimization problem with arbitrary precision.

In numerical experiments, we showed that the LFNN model achieves comparable results to an approximate form of the closed-form solution on synthetic data under the CD objective. This validates the capability of the LFNN (even with a shallow structure) to represent the optimal control accurately. In turn, the closed-form solution improves the interpretability of the neural network model.

Then, we studied an empirical investment case with four assets. We evaluated the LFNN model on bootstrap resampled data and showed that the neural network strategy achieves consistent outperformance over the benchmark strategy throughout the investment period. Specifically, the neural network strategy achieves a 2% higher median IRR over the benchmark strategy, and a higher terminal wealth with more than 90% probability. The allocation strategy suggests that managers should favor equal-weighted stock index over cap-weighted stock index and short-term bonds over long-term bonds during a high-inflation period.
A Moving-window inflation filter

A.1 Filtering algorithm

Algorithm A.1 presents the pseudocode for the moving-window filtering algorithm.

Algorithm A.1: Pseudocode window inflation filter

Data:
- CPI[i]; i = 1, ..., N /* CPI Index */
- Cutoff /* High inflation cutoff: annualized */
- ∆t /* CPI index time interval */
- K /* smoothing window size */

Result: Flag[i]; i = 1, ..., N /* = 1 high-inflation month; = 0 otherwise */

/* initialization */
Flag[i] = 0; i = 1, ..., N;

for ( i = 1, ..., N - K ) {
    if \( \log(CPI[i + K]/CPI[i])/(K \times \Delta t) > \text{Cutoff} \) then
        for ( j = 0, ..., K ) {
            Flag[i+j] = 1;
        }
    end
}

A.2 Effect of moving window size

Figure A.1 shows the filtering results for windows of size 12, 60, and 120 months. We can see that the five-year window produces two obvious 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). Increasing the window size to 10 years results in similar-looking plots as the five-year window size, but the number of months in each window increases, and the average inflation rate is lower. Since our objective is to determine the effect of high-inflation periods on allocation strategies, we choose the five-year window size.

Figure A.1: high-inflation regimes, using the moving-window method, with the window size shown. The cutoff for high-inflation regimes was 0.05. High-inflation months have a label value of one, and low-inflation months have a label value of zero. CPI data identified from the historical period 1926:1-2022:1.
A.3 Asset performance during high inflation

To gain some intuition on the behavior of asset returns during the inflation periods, we assume that each real (adjusted by CPI index) index follows geometric Brownian motion (GBM). For example, given an index with value $S$, then

$$\text{d}S = \mu S \text{d}t + \sigma S \text{d}Z$$  \hspace{1cm} (A.1)

where $\text{d}Z$ is the increment of a Wiener process. We use maximum likelihood estimation to fit the drift rate $\mu$ (expected arithmetic return) and volatility $\sigma$ in each regime, for each index, as shown in Table A.1. We also show a series constructed by: converting the indexes in each regime to returns, concatenating the two return series, and converting the concatenated return series back to an index. This concatenated index does not, of course, correspond to an actual historical index, but is a pseudo-index constructed from high-inflation regimes. This amounts to a worst-case sequence of returns in terms of the duration of historical inflation periods, that could plausibly be expected during a long period of high inflation.

<table>
<thead>
<tr>
<th>Index</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$\mu - \sigma^2/2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1940:8-1951:7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CapWt</td>
<td>0.079</td>
<td>0.140</td>
<td>0.069</td>
</tr>
<tr>
<td>EqWt</td>
<td>0.145</td>
<td>0.190</td>
<td>0.127</td>
</tr>
<tr>
<td>10 Year Treasury</td>
<td>-0.035</td>
<td>0.036</td>
<td>-0.036</td>
</tr>
<tr>
<td>30-day T-bill</td>
<td>-0.050</td>
<td>0.029</td>
<td>-0.050</td>
</tr>
<tr>
<td>1968:9-1985:10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CapWt</td>
<td>0.026</td>
<td>0.164</td>
<td>0.013</td>
</tr>
<tr>
<td>EqWt</td>
<td>0.065</td>
<td>0.220</td>
<td>0.041</td>
</tr>
<tr>
<td>10 Year Treasury</td>
<td>0.011</td>
<td>0.093</td>
<td>0.007</td>
</tr>
<tr>
<td>30-day T-bill</td>
<td>0.009</td>
<td>0.012</td>
<td>0.009</td>
</tr>
<tr>
<td>CapWt</td>
<td>0.049</td>
<td>0.156</td>
<td>0.038</td>
</tr>
<tr>
<td>EqWt</td>
<td>0.098</td>
<td>0.209</td>
<td>0.076</td>
</tr>
<tr>
<td>10 Year Treasury</td>
<td>-0.008</td>
<td>0.076</td>
<td>-0.011</td>
</tr>
<tr>
<td>30-day T-bill</td>
<td>-0.014</td>
<td>0.022</td>
<td>-0.014</td>
</tr>
</tbody>
</table>

Table A.1: GBM parameters for the indexes shown. All indexes are real (deflated). $\mu$ is the expected annualized arithmetic return. $\sigma$ is the annualized volatility. $(\mu - \sigma^2/2)$ is the annualized mean geometric return.

It is striking that in each historical inflation regime (i.e., 1940:8-1951:7 and 1968:9-1985:10) in Table A.1 the drift rate $\mu$ for the equal-weighted index is much larger than the drift rate for the cap-weighted index. We can observe that the mean geometric return for the cap-weighted index, in the period 1968:9-1985:10, was only about one percent per year.

It is also noticeable that bonds performed very poorly in the period 1940:8-1951:7. As well, during the period 1968:9-1985:10, there was essentially no term premium for 10-year treasuries, compared with 30-day T-bills. In addition, the 10-year treasury index had much higher volatility compared to the 30-day T-bill index. Looking at the concatenated series, it appears that 30-day T-bills are arguably the better defensive asset here since the volatility of this index is quite low (but with a negative (real) drift rate).

Consequently, in the following, we will focus attention on 30-day T-bills, the cap-weighted index, and the equal-weighted index.
B Bootstrap resampling

B.1 Stationary block bootstrap algorithm

Algorithm B.1 presents the pseudocode for the stationary block bootstrap. See Ni et al. (2022) for more discussion.

---

Algorithm B.1: Pseudocode for stationary block bootstrap

/* initialization */
bootstrap_samples = [ ];
/* loop until the total number of required samples are reached */
while True
    /* choose random starting index in [1,...,N], N is the index of the last historical sample */
    index = UniformRandom( 1, N );
    /* actual blocksize follows a shifted geometric distribution with the expected value of exp(blocksize) */
    blocksize = GeometricRandom( exp(blocksize) );
    for ( i = 0; i < blocksize; i = i + 1 ) {
        /* if the chosen block exceeds the range of the historical data array, do a circular bootstrap */
        if index + i > N then
            bootstrap_samples.append( historical_data[ index + i - N ] );
        else
            bootstrap_samples.append( historical_data[ index + i ] );
        end
        if bootstrap_samples.len() == number_required then
            return bootstrap_samples;
        end
    }
end

---

B.2 Effect of blocksize

As discussed, we will use bootstrap resampling (Politis and Romano, 1994; Politis and White, 2004; Patton et al., 2009; Dichtl et al., 2016; Anarkulova et al., 2022), to analyze the performance of using the equal-weighted index compared to the cap-weighted index, during periods of high inflation (our concatenated series: 1940:8-1951:7, 1968:9-1985:10).

First, we examine the effect of the expected blocksize parameter in the bootstrap resampling algorithm. We will use a paired sampling approach, where we simultaneously draw returns from the bond and stock indexes. The algorithm in Politis and White (2004) was developed for single asset time series. It is therefore important to assess the effect of the blocksize on numerical results. In Table B.1 we examine the effect of different blocksizes on the statistics of stationary block bootstrap resampling.

Perhaps a more visual way of analyzing the effect of the expected blocksize is shown in Figure B.1, where we show the cumulative distribution function (CDF) of the final wealth after 10 years, for different blocksizes. We show the CDF since this gives us a visualization of the entire final wealth distribution, not just a few summary statistics.

---

This preserves correlation effects.
Since the data frequency is at one-month intervals, specifying a geometric mean expected blocksize of one month means that the blocksize is always a constant one month. This effectively means that we are assuming that the data is i.i.d. However, the one-month results are an outlier, compared to the other choices of expected blocksize. There is hardly any difference between the CDFs for any choice of expected blocksize in the range of 3-24 months. In this article, we use an expected blocksize of 6 months.


| Expected blocksize (months) | Median $|W_T|$ | $E[W_T]$ | std $|W_T|$ | 5th Percentile |
|-----------------------------|-----------|----------|-----------|---------------|
| 1                           | 170.9     | 191.6    | 97.6      | 78.6          |
| 3                           | 174.6     | 202.9    | 120.4     | 69.3          |
| 6                           | 174.2     | 204.2    | 125.9     | 66.8          |
| 12                          | 175.5     | 204.4    | 124.2     | 67.9          |
| 24                          | 179.2     | 205.1    | 118.4     | 68.7          |

Figure B.1: Cumulative distribution function (CDF), final wealth $W(T)$ at $T = 10$ years. The effect of expected blocksize. Constant weight, scenario in Table 2.2. Equity weight: 0.7, rebalanced monthly. Bond index: 30-day T-bill. Equity index: equal-weighted. Concatenated series: 1940:8-1951:7, 1968:9-1985:10 (high-inflation regimes). All quantities are real (inflation-adjusted). Initial wealth 100. Bootstrap resampling, expected blocksize one year, 10,000 resamples.

B.3 Bootstrapping from non-contiguous data segments

As discussed in Section 2.2 we have identified two historical inflation regimes: 1940:8-1951:7 and 1968:9-1985:10. As traditional bootstrap methods assume one segment of the underlying data segment, it naturally becomes a question of how to bootstrap from two non-contiguous segments of data appropriately. In the main sections of the article, we first concatenate the two data segments, then treat the concatenated data samples as a complete segment and apply bootstrap methods to them. This method is in line with the work of Anarkulova et al. (2022), in which the authors concatenate stock returns from different countries and bootstrap from the concatenated series.
A second intuitive bootstrap method would be to bootstrap randomly from each of the two segments. Briefly, each bootstrap resample consists of (i) selecting a random segment (probability proportional to the length of the segment), (ii) selecting a random starting date in the selected segment, (iii) then selecting a block (of random size) of consecutive returns from this start date, (iv) in the event that the end of the data set in a segment is reached, we use circular block bootstrap resampling within that segment, and (v) repeating this process until a sample of the total desired length is obtained.

We compare the bootstrapped data from concatenated segments and separate segments, by evaluating the performance of the 70%/30% equal-weighted index/T-bill fixed-mix portfolio, using the investment scenario described in Table 2.2.

<table>
<thead>
<tr>
<th></th>
<th>Median[W_T]</th>
<th>E[WT]</th>
<th>std[W_T]</th>
<th>5th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bootstrap from concatenated segments</td>
<td>174.2</td>
<td>204.2</td>
<td>125.9</td>
<td>66.8</td>
</tr>
<tr>
<td>Bootstrap from separate segments</td>
<td>176.9</td>
<td>208.0</td>
<td>132.4</td>
<td>65.4</td>
</tr>
</tbody>
</table>

Table B.2: Effect of bootstrap method - bootstrap from concatenated segments vs bootstrap from separate segments, on the statistics of the final wealth \(W(T)\) at \(T = 10\) years. Constant weight, scenario in Table 2.2. Equity weight: 0.7, rebalanced monthly. Bond index: 30-day T-bill. Equity index: equal-weighted. Concatenated series: 1940:8-1951:7, 1968:9-1985:10 (high-inflation regimes). All quantities are real (inflation-adjusted). Initial wealth 100. Bootstrap resampling, 10,000 resamples.

We can observe from Table B.2 that the strategy performance on bootstrap resampled data using two methods only varies slightly. This indicates that the two methods do not yield much difference for practical purposes. This is indeed expected - after all, the difference between the two methods only occurs when a random block crosses the edge of each of the segments. However, such a situation only occurs with a very low probability. Except for this low-probability situation, the two bootstrap methods are identical.

C Technical details of closed-form solution

C.1 Proof of Theorem (3.17)

At any state \((t, w, \hat{w}, \hat{\varrho}) \in [t_0, T] \times \mathbb{R}^2\), define the value function \(V(w, \hat{w}, t)\) to the CD problem (3.11) as

\[
V(t, w, \hat{w}, \hat{\varrho}) = \inf_p \left\{ \mathbb{E}_p \left[ \int_t^T \left( W(s) - e^{\beta s} \hat{W}(s) \right)^2 ds \right] | W(t) = w, \hat{W}(t) = \hat{w} \right\}. \tag{C.1}
\]

By the dynamic programming principle, we have

\[
V(t, w, \hat{w}, \hat{\varrho}) = \inf_p \left\{ \mathbb{E}_p \left[ \left( V(t+\Delta t, W(t+\Delta t), \hat{W}(t+\Delta t), \hat{\varrho}) + \int_t^{t+\Delta t} \left( W(s) - e^{\beta s} \hat{W}(s) \right)^2 ds \right) | W(t) = w, \hat{W}(t) = \hat{w} \right] \right\}. \tag{C.2}
\]

Rearrange equation (C.2) to obtain

\[
\inf_p \left\{ \mathbb{E}_p \left[ \left( dV(t, w, \hat{w}, \hat{\varrho}) + \int_t^{t+\Delta t} \left( W(s) - e^{\beta s} \hat{W}(s) \right)^2 ds \right) | W(t) = w, \hat{W}(t) = \hat{w} \right] \right\} = 0 \tag{C.3}
\]

Then, apply Itô’s lemma with jumps (Cont et al., 2011), substitute \(dW\) and \(d\hat{W}\) terms with (3.16), and take limits as \(\Delta t \downarrow 0\), we obtain (3.17).

The above results merely serve as an intuitive guide to obtain (3.17). The formal proof of (3.17) proceeds by using a suitably smooth test function, see for example (Øksendal and Sulem, 2007).
C.2 Proof of results for CD-optimal control

In Section 3.4, we emphasized the dependence of $B$ and $D$ (defined in (3.23) and (3.22)) on parameters $\beta$ and $c$ for understanding the optimal control function. As $\beta$ 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 (3.17) suggests the following ansatz for the value function $V$ in Theorem 3.17 of the form

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

where $A, B, C, \hat{A}, \hat{B}, D$ are unknown deterministic functions of time $t$. If (C.4) is correct, then the pointwise infimum in (3.17) is attained by $p^*$ satisfying the relationship

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

assuming $A(t) > 0$. Here $\gamma$ and $\theta$ are defined in (3.19). (C.4) implies that the relevant partial derivatives of $V$ are of the form

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

Substituting (C.6) into (C.5), the optimal control $p^*$ obtained is in the form of (3.20), where $h$ and $g$ are given by (3.21). Then, it only remains to determine the functions $A, B, D$. Substituting (C.5) into PIDE (3.17), we can obtain the following ordinary differential equations (ODE) for $A, B, D$,

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

Solving the ODE system gives us the $A, B, D$ defined in (3.22) and (3.23). We also note that $A(t) > 0$, thus completing the proof.

C.3 Proof of Corollary (3.1) and (3.2)

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

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

Under such assumptions, van Staden et al. (2022) shows that the CD-optimal control can be expressed in a similar form as in (3.20) with $g$ and $h$ functions. The $g$ and $h$ functions satisfy the same properties as in Corollary (3.1) and (3.2). In spite of the fact that we assume the bond price follows a jump-diffusion model, the proof of Corollary (3.1) and (3.2) follows similar steps as the proof in van Staden et al. (2022).

D Technical details of LFNN model

D.1 Proof of Theorem 3.2

Theorem 3.2 (Unconstrained feasibility domain) The feasibility domain $Z^\theta$ defined in (3.39) associated with the LFNN model (3.41) is $\mathbb{R}^{N^\theta}$. 

35
Proof. First, it is obvious that \( Z_\theta \subseteq \mathbb{R}^{N_a} \) by definition of (3.39). Next, we show that \( \mathbb{R}^{N_a} \subseteq Z_\theta \). To prove this, we need to show that for any \( \theta \in \mathbb{R}^{N_a} \),

\[
f(x; \theta) = p \in \begin{cases} \mathcal{Z}, & \text{if } x \in \mathcal{X}_1, \\ \{e_{N_l+1}\}, & \text{if } x \in \mathcal{X}_2. \end{cases} \quad \forall x \in \mathcal{X}. \quad (D.1)
\]

Here \( f \) is the LFNN function defined in (3.41), \( p = (p_1, \ldots, p_{N_a})^\top \in \mathbb{R}^{N_a} \) is the output of the LFNN model that represents the wealth allocation to the assets, \( \mathcal{Z} \) is the feasibility domain defined in (3.35), and \( x = (t, W(t), \hat{W}(t))^\top \in \mathcal{X} \) is a feature vector. To prove (D.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 (3.42).

Next, we verify that when \( x \in \mathcal{X}_1 \), \( p \in Z_\theta \). To prove this, we need to show that constraints of (3.29)-(3.32) are satisfied when \( x \in \mathcal{X}_1 \).

By definition of (3.42), it is obvious that the long-only constraint (3.29) holds for long-only assets. It is also easy to verify that the summation constraint (3.30) is satisfied. This can be observed after the fact that

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

The maximum leverage constraint (3.31) is also satisfied, as

\[
\sum_{i=1}^{N_l} p_i = l = p_{\max} \cdot \text{Sigmoid}(-a_{N_l+1}) \leq p_{\max}. \quad (D.3)
\]

Finally, the simultaneous shorting constraint (3.5) 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 (3.42), we know

\[
p_i = (1 - l) \cdot \frac{e^{o_i}}{\sum_{k=N_l+1}^{N_a} e^{o_k}} \leq 0, \quad \forall i \in \{N_l + 1, \ldots, N_a\} \quad (D.4)
\]

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

Therefore, for any \( \theta \in \mathbb{R}^{N_a} \), (D.1) is satisfied. This implies \( \mathbb{R}^{N_a} \subseteq Z_\theta \).

\[\square\]

**D.2 Proof of Lemma 3.2 and Theorem 3.3**

**Lemma 3.2.** (Structure of feasible control) Any feasible control function \( p : \mathcal{X} \mapsto \mathbb{R}^{N_a} \) has the function decomposition

\[
p(x) = \varphi(\omega(x), x), \quad (D.5)
\]

where \( \varphi : \hat{\mathcal{Z}} \times \mathcal{X} \mapsto \mathbb{R}^{N_a} \) is defined in (3.44), i.e.

\[
\varphi(z) = \left(z_{N_l+1} \cdot z_1, \ldots, z_{N_l}, (1 - z_{N_l+1}) \cdot z_1, \ldots, z_{N_a}\right)^\top \cdot 1_{x \in \mathcal{X}_1} + e_{N_l+1} \cdot 1_{x \in \mathcal{X}_2}, \quad (D.6)
\]

and \( \omega : \mathcal{X} \mapsto \hat{\mathcal{Z}} \). Here

\[
\hat{\mathcal{Z}} = \left\{ z \in \mathbb{R}^{N_a+1}; \sum_{i=1}^{N_l} z_i = 1, \sum_{i=N_l+1}^{N_a} z_i = 1, z_{N_l+1} \leq p_{\max}, z_i \geq 0, \forall i \right\}. \quad (D.7)
\]
Proof. We prove the lemma by existence. \( \forall z = (z_1, \cdots, z_{N_a})^T \in \mathcal{Z} \), define \( y = \phi(z) \in \mathbb{R}^{N_a+1} \) as

\[
y : \begin{cases}
y_i = \frac{z_i}{\sum_{j=1}^{N_i} z_j}, & i \in \{1, \cdots, N_i\}, \\
y_i = \frac{z_i}{1 - \sum_{j=1}^{N_i} z_j}, & i \in \{N_i, \cdots, N_a\}, \quad \text{if } \sum_{i=1}^{N_i} z_i \in (0, 1) \cup (1, p_{\text{max}}], \\
y_{N_a+1} = \sum_{j=1}^{N_1} z_j, \\
y_i = z_i, & i \in \{1, \cdots, N_1\}, \\
y_i = 1/(N_a - N_1), & i \in \{N_i, \cdots, N_a\}, \quad \text{if } \sum_{i=1}^{N_i} z_i = 1, \\
y_{N_a+1} = 1, \\
y_i = 0, & i \in \{1, \cdots, N_i\}, \\
y_i = z_i, & i \in \{N_i, \cdots, N_a\}, \quad \text{if } \sum_{i=1}^{N_i} z_i = 0, \\
y_{N_a+1} = 0,
\end{cases}
\]

Next, define \( \omega \) as

\[
\omega(x) = \begin{cases}
\phi(p(x)), & \text{if } x \in \mathcal{X}_1, \\
\left( \frac{1}{N_1}, \cdots, \frac{1}{N_1 - N_i}, \cdots, \frac{1}{N_a - N_i}, 0 \right)^T, & \text{if } x \in \mathcal{X}_2,
\end{cases}
\]

Then, it can be easily verified that \( \omega : \mathcal{X} \mapsto \hat{\mathcal{Z}} \), and that \( p(x) = \varphi(\omega(x), x) \). \( \square \)

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

\[
p(x) = \Phi(\Omega(x), x), \quad x \in \mathcal{X},
\]

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

\[
\left\{ \bigcup_{i=1}^{n} \mathcal{X}_i = \mathcal{X}, \quad \mathcal{X}_i \cap \mathcal{X}_j = \emptyset, \forall 1 \leq i, j \leq n. \right\}
\]

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

(i) \( \Upsilon \) 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 \( \theta \) such that the fully connected feedforward neural network function \( \hat{f}(\cdot; \theta) \) defined in \( (3.40) \) satisfies

\[
\sup_{x \in \mathcal{X}} \| \Phi(\Upsilon(\hat{f}(x; \theta)), x) - p(x) \| < \epsilon.
\]

**Proof.** Let

\[
L_{\phi} = \max_{1 \leq i \leq n} L_i,
\]

where \( L_i \) is the Lipschitz constant for \( \Phi \) on \( \mathcal{Y} \times \mathcal{X}_i \).

Then, following Kratsios and Bilokopytov (2020), we know that \( \forall \epsilon, L > 0 \), there exists \( N_h \in \mathbb{N} \) and \( \theta \in \mathbb{R}^{N_\theta} \) such that the corresponding FNN \( \hat{f}(\cdot; \theta) : \mathcal{X} \mapsto \mathbb{R}^m \) defined in \( (3.40) \) satisfies

\[
\sup_{x \in \mathcal{X}} \| \Upsilon(\hat{f}(x; \theta)) - \Omega(x) \| < \epsilon/L_{\phi},
\]

37
Theorem 3.3. (Approximation of optimal control) Following Assumption 3.7, the network satisfies the constraints naturally, and be guaranteed that such a neural network can approximate allocations to specific assets subject to thresholds tied to the performance of the portfolio.\(\theta\) and \(a, b\) 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 D.1, with a bit of engineering, one can easily construct a \(\Phi\) 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 3.3.

Theorem 3.3 (Approximation of optimal control) Following Assumption 3.7 \(\forall \varepsilon > 0\), there exists \(N_h \in \mathbb{N}\), and \(\theta \in \mathbb{R}^{n_{\theta}}\) such that the corresponding LFNN model \(f(\cdot; \theta)\) described in (3.41) satisfies the following:

\[
\sup_{x \in \mathcal{X}} \|f(x; \theta) - p^*(x)\| < \varepsilon.
\]

Proof. From (3.41) and Lemma 3.1 we know that

\[
f(x; \theta) = \psi(\tilde{f}(x; \theta), x) = \varphi(\zeta(\tilde{f}(x; \theta)), x),
\]

where \(\tilde{f}\) is the FNN defined in (3.40) and \(\varphi : \hat{Z} \times \mathcal{X} \mapsto \mathbb{R}^{N_a}, \zeta : \mathbb{R}^{N_a+1} \mapsto \hat{Z}\) are defined in (3.44).

It can be easily verified that \(\zeta\) satisfies the following:

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

\[
\zeta^{-1}(z) : Im(\zeta) \mapsto \mathbb{R}^{N_a+1}, \quad \zeta^{-1}(z) = \left(\log(z_1), \cdots, \log(z_{N_a}), \sigma^{-1}(z_{N_a+1}/p_{max})\right)^\top.
\]

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

(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 3.1. It is easily verifiable that \(\varphi\) is Lipschitz continuous on \(\mathcal{X}_1\) and \(\mathcal{X}_2\) respectively.

Finally, according to Assumption 3.7 \(p^*(x) = \varphi(\omega^*(x), x)\), where \(\omega^* \in C(\mathcal{X}, \hat{Z})\).

Applying Lemma D.1, 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:

\[
\sup_{x \in \mathcal{X}} \|f(x; \theta) - p^*(x)\| < \varepsilon.
\]

Remark D.1. (Remark on Lemma D.1) Normally, the universal approximation theorem only applies to approximation of continuous functions defined on a compact set (Hornik 1991). Lemma D.1 extends the universal approximation theorem to a broader class of functions which have the structure of (D.10). Furthermore, Lemma D.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 \(p(x) \in [a(x), b(x)]\) 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 D.1, a bit of engineering, one can easily construct a \(\Phi\) so that the corresponding neural network satisfies the constraints naturally, and be guaranteed that such a neural network can approximate the control well.

\[24\] \(\sigma^{-1}\) is the inverse function of the sigmoid function.
E Calibrated synthetic model parameters

<table>
<thead>
<tr>
<th>μ₁</th>
<th>σ₁</th>
<th>λ₁</th>
<th>ν₁</th>
<th>κ₁</th>
<th>μ₂</th>
<th>σ₂</th>
<th>λ₂</th>
<th>ν₂</th>
<th>κ₂</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.051</td>
<td>0.146</td>
<td>0.178</td>
<td>0.2</td>
<td>7.13</td>
<td>7.33</td>
<td>-0.014</td>
<td>0.017</td>
<td>0.321</td>
<td>0</td>
<td>N/A</td>
</tr>
</tbody>
</table>


F Comparison of CD and CS objectives

In this section, we numerically compare the CS objective function \((3.13)\) with the CD objective function \((3.11)\).

As we briefly discussed in Section 3.3, one caveat of the CD objective function is that it not only penalizes the underperformance relative to the elevated target but also penalizes the outperformance over the elevated target. In practice, the outperformance of the elevated target is favorable, and managers may not want to penalize the strategy when it happens. Therefore, in such cases, the cumulative quadratic shortfall (CS) objective \((3.13)\) and \((3.14)\) may be more appropriate. For the remainder of the paper, we focus on the discrete-time CS problem with the LFNN parameterization and the equally-spaced rebalancing schedule \(T_{\Delta t}\) defined in Section 4.1.1 i.e.,

\[
\begin{equation}
\text{(Parameterized CS}(\beta)\text{)}: \inf_{\theta \in \mathbb{R}^N} \mathbb{E}_{f(\cdot; \theta)}^{(t_0, w_0)} \left[ \sum_{t \in T_{\Delta t}} \left( \min (W_{\theta}(t) - e^{\beta t} \hat{W}(t), 0) \right)^2 + \epsilon W_{\theta}(T) \right], \quad \text{(F.1)}
\end{equation}
\]

The CS objective function in \((F.1)\) only penalizes the underperformance against the elevated target. Here \(\epsilon W(T)\) is a regularization term. We remark that problem \((F.1)\) without the regularization term can be ill-posed. To see this, consider a case where \(W_{\theta}(t) \gg e^{\beta t} \hat{W}(t)\), for some \(t \in [t_0, T]\). In this case, the future cumulative quadratic shortfall \((|t, T|)\) will almost surely be zero without the regularization term, so the control from thereon has no effect on the objective function under that scenario. We choose \(\epsilon\) to be a small positive scalar. As William Bernstein once said, “if you have won the game stop playing.” If one has accumulated as much wealth as Warren Buffet, then it does not matter what assets she invests in. The positive regularization factor of \(\epsilon\) forces the strategy to put all wealth into less risky assets when the portfolio has already performed extremely well.

We design a numerical experiment to compare the CS objective with the symmetric CD objective in the following problem \((F.2)\) with the same LFNN parameterization and equally-spaced rebalancing schedule \(T_{\Delta t}\)

\[
\begin{equation}
\text{(Parameterized CD}(\beta)\text{)}: \inf_{\theta \in \mathbb{R}^N} \mathbb{E}_{f(\cdot; \theta)}^{(t_0, w_0)} \left[ \sum_{t \in T_{\Delta t}} \left( W_{\theta}(t) - e^{\beta t} \hat{W}(t) \right)^2 \right]. \quad \text{(F.2)}
\end{equation}
\]

Specifically, we adopt the investment scenario in Table 4.1 with \(\beta = 0.02\) and reuse the training and testing data sets simulated from the calibrated double exponential jump-diffusion model. The neural network strategies follow the trained LFNN models on the training data set \(Y\) for both the CS and CD objective. We then evaluated both strategies on the same testing data set \(Y_{\text{test}}\).

We compare the wealth ratio, i.e., the wealth of the managed portfolio divided by the wealth of the benchmark portfolio, over time, for both strategies. The wealth ratio metric reflects how well the active strategy performs against the benchmark strategy along the investment horizon; a higher wealth ratio metric is better. Below we show the percentiles of the wealth ratio for both strategies evaluated on \(Y_{\text{test}}\).
Figure F.1: Percentiles of wealth ratio of the neural network strategy (i.e., the neural network model) learned under the cumulative quadratic tracking difference (CD) objective and the neural network strategy learned under the cumulative quadratic shortfall (CS) objective. The results shown are based on evaluations on the testing data set $Y_{\text{test}}$.

We can see from Figure F.1 that the CS strategy (neural network strategy trained under the CS objective) yields a more favorable wealth ratio than the CD strategy (neural network strategy trained under the CD objective). On average, the CS strategy achieves a consistently higher terminal wealth ratio than the CD strategy. Even in the 20th percentile case, the CS strategy lags initially but recovers over time. The result indicates that the CS objective might be a wiser choice for managers in practice. In the following numerical experiments with bootstrap resampled data, we use the CS objective (F.1) instead of the CD objective (F.2).

25 The CS strategy starts with a higher allocation to the stock, and thus encounters more volatility early on.

26 Unfortunately, we cannot derive a closed-form solution under the CS objective (3.13) under continuous rebalancing using the PIDE approach, due to the non-smoothness introduced by the $\text{min}(\cdot)$ function.
High borrowing premium

In Section 4.2, we conducted the numerical experiments, assuming the borrowing premium is zero. This assumption is based on the fact that large sovereign wealth funds are often considered to have almost risk-free credit ratings, due to their state-backed nature. In other words, we assume that sovereign wealth funds can borrow funding at the same rate as risk-free treasury bills.

This assumption may be too benign for general public funds. In general, it is unlikely that a non-sovereign wealth fund can borrow at a risk-free rate. However, the actual borrowing cost within large public funds is often unavailable. For this reason, we use the corporate bond yields issued by corporations with similar credit ratings as these large public funds as an approximation to the borrowing cost. Currently, large public funds such as the Blackstone Group or Apollo Global Management are rated between Aaa and Baa rating by Moody’s. We obtain the nominal corporate bond yields with Moody’s Aaa (Moody’s, 2023a) and Baa (Moody’s, 2023b) ratings and adjust them with CPI returns. During the two high-inflation regimes we have identified, Aaa-rated corporate bonds have an average real yield of 0.7%, while Baa-rated bonds have 1.8%. Taking an average of the two, we use 1.25% as an estimate for the real yield of corporate bonds as well as the borrowing cost of large public funds.27 As discussed in Section 2, the average real return for T-bill index is -1.6%. This gives us an average borrowing premium rate of 2.85%. In this section, as a stress test, we conduct the same experiment as in Section 4.2 except that we use a fixed borrowing premium of 3% instead of 0. We note that the historical corporate bond yields are based on bonds with a long maturity. Typically, long-term yields are higher than short-term yields, which accounts for the term risk. Therefore, the assumption of a 3% borrowing premium should be a quite aggressive stress test for the use of leverage.

![Figure G.1: Percentiles of wealth ratio over the investment horizon, and CDF of terminal wealth ratio. Annualized borrowing premium is 3%. Results are based on the evaluation of the learned neural network model (from high-inflation data) on the testing data set (low-inflation data).](image)

27Note that the corporate bonds from Moody's yield data have maturities of more than 20 years. Usually, long-term bonds have higher yields than short-term bonds. Thus, using corporate yields likely overestimates the borrowing cost, since we assume the manager is only borrowing short-term funding.
Table G.1: Statistic of strategies. Annualized borrowing premium is 3%. Results are based on the evaluation of the learned neural network model (from high-inflation data) on the testing data set (low-inflation data).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural network</td>
<td>362.7</td>
<td>401.3</td>
<td>212.6</td>
<td>133.9</td>
<td>0.078</td>
</tr>
<tr>
<td>Benchmark</td>
<td>308.5</td>
<td>342.9</td>
<td>165.0</td>
<td>149.0</td>
<td>0.056</td>
</tr>
<tr>
<td>Neural network (zero premium)</td>
<td>364.2</td>
<td>403.4</td>
<td>211.8</td>
<td>136.3</td>
<td>0.078</td>
</tr>
</tbody>
</table>

H Strategy performance in low inflation regimes

Until now, the entire discussion has been centered around the imaginary scenario of a long, persistent inflation environment. We have shown that the neural network strategy consistently outperforms the benchmark strategy under a high-inflation regime. However, what if we are wrong? More dramatically, if the high-inflation situation ends immediately and we enter a low-inflation environment, how will the neural network strategy learned under high-inflation regimes perform?

To answer these questions, we evaluate the neural network strategy learned under high-inflation regimes on a testing data set bootstrapped from low-inflation historical regimes. Specifically, we exclude the two
inflation regimes (1940:8-1951:7 and 1968:9-1985:10) from the full historical data of 1926:1-2022:1 and obtain several low-inflation data segments. We concatenate the low-inflation data segments and use the stationary bootstrap (Appendix B.1) to generate a testing data set. We adopt the investment scenario described in Table 4.3 and evaluate the performance of the neural network strategy obtained in Section 4.2 on this low-inflation data set.

![Figure H.1: Cumulative distribution functions (CDFs) for cap-weighted and equal-weighted indexes, as a function of final real wealth $W$ at $T = 10$ years. Initial stake $W_0 = 100$, no cash injections or withdrawals. Block bootstrap resampling, expected blocksize 6 months. 70% stocks, 30% bonds, rebalanced monthly. Bond index: 30-day U.S. T-bills. Stock index: CRSP capitalization-weighted or CRSP equal-weighted index. Underlying data excludes high-inflation regimes. All indexes are deflated by the CPI. 10,000 resamples. Data set 1926:1-2022:1, excluding high inflation regimes (1940:8-1951:7 and 1968:9-1985:10).](image)

Note that we continue to use the equal-weighted stock index/30-day T-bill fixed-mix portfolio as the benchmark. This is validated by Figure H.1 which plots the CDF of the terminal wealth of the fixed-mix portfolios using 70% equal-weighted stock index vs 70% cap-weighted stock index (both with 30% 30-day U.S. T-bill as the bond component). As we can see from Figure H.1, the fixed-mix portfolio with an equal-weighted stock index clearly has a more right-skewed distribution than the portfolio with a cap-weighted stock index. This seems to suggest that the equal-weighted index is the superior choice to use in the benchmark portfolio, even in low inflation regimes.

| Strategy          | Median $|W_T|$ | E($|W_T|$) | std($|W_T|$) | 5th Percentile | Median IRR (annual) |
|-------------------|---------|-------|---------|-----------|----------------|---------------------|
| Neural network    | 429.7   | 489.6 | 301.9   | 151.6     | 0.100          |                     |
| Benchmark         | 368.3   | 420.8 | 238.2   | 175.7     | 0.079          |                     |

Table H.1: Statistic of strategies. Results are based on the evaluation of the learned neural network model (from high-inflation data) on the low-inflation testing data set.

We then present the performance of the neural network strategy learned on high-inflation data on the testing data set bootstrapped from low-inflation historical returns. Surprisingly, as we can see from Figure H.2a, the performance of the neural network strategy learned under high-inflation regimes performs quite well in low-inflation environments. Compared to the testing results on the high-inflation data set, there is a noticeable performance degradation; for example, the probability of outperforming the benchmark strategy in terminal wealth is now slightly less than 90%. However, the degradation is quite minimal. The neural network strategy still has more than 85% chance of outperforming the benchmark strategy at the end of
the investment horizon. As shown in Table H.1 the median IRR of the neural network strategy is still 2% higher than the median IRR of the benchmark strategy, meeting the investment target.

The above results indicate that the neural network strategy is surprisingly robust. Despite being specifically trained under a high-inflation scenario, the strategy performs admirably well in a low-inflation environment.
References


Boyde, E. (2021). ETFs: why tracking difference usually matters more than tracking error. [https://www.ft.com/content/44b08a6c-55d7-4841-8e8e-aad6451a4cc3](https://www.ft.com/content/44b08a6c-55d7-4841-8e8e-aad6451a4cc3) [Online; accessed 6-Dec-2022].


Forsyth, P. A., P. van Staden, and Y. Li (2022). Is it too easy for an asset manager to beat a constant weight benchmark?


 Moody’s (2023a). Moody’s Seasoned Aaa Corporate Bond Yield [AAA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/AAA. [Online; accessed 18-Jan-2023].

 Moody’s (2023b). Moody’s Seasoned Baa Corporate Bond Yield [BAA], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/BAA. [Online; accessed 18-Jan-2023].


Taljaard, B. H. and E. Mare (2021). Why has the equal weight portfolio underperformed and what can we do about it? *Quantitative Finance* 21(11), 1855–1868.


