Beating a benchmark: dynamic programming may not be the right numerical approach.

Pieter M. van Staden*, Peter A. Forsyth†, and Yuying Li‡

Abstract. We analyze dynamic investment strategies for benchmark outperformance using two widely-used objectives of practical interest to investors: (i) maximizing the information ratio (IR), and (ii) obtaining a favorable tracking difference (cumulative outperformance) relative to the benchmark. In the case of the tracking difference, we propose a simple and intuitive objective function based on the quadratic deviation (QD) from an elevated benchmark. In order to gain some intuition about these strategies, we provide closed form solutions for the controls under idealized assumptions. For more realistic cases, we represent the control using a Neural Network (NN) and directly solve a sampled optimization problem, which approximates the original optimal stochastic control formulation. Unlike the typical approach based on dynamic programming (DP), e.g. reinforcement learning, solving the sampled optimization with an NN as a control avoids computing conditional expectations and leads to an optimization problem with a small number of variables. In addition, our NN parameter size is independent of the number of portfolio rebalancing times. Under some assumptions, we prove that a traditional dynamic programming approach results in high dimensional problem, whereas directly solving for the control without using DP yields a low dimensional problem. Our analytical and numerical results illustrate that, compared with IR-optimal strategies with the same expected value of terminal wealth, the QD-optimal investment strategies result in comparatively more diversified asset allocations during certain periods of the investment time horizon.

Key words. Asset allocation, stochastic control, benchmark outperformance, neural network

MSC codes. 91G60, 93E20

1. Introduction.

Despite the considerable professional talent attracted to the field of active portfolio management, where a portfolio manager (or an investment institution) brings their expertise to bear on actively pursuing an investment strategy with the explicit goal of *outperforming* an appropriate pre-specified benchmark ([68, 76, 4, 120, 72]), it remains a disappointing fact that the promised outperformance hardly ever seems to materialize in practice. In fact, underperforming their benchmarks is something professional portfolio managers achieve with "surprising consistency" ([52]).

It is worth noting that many government pension plans also report performance relative to a benchmark of publicly traded financial assets ([21, 53]). Typically, the benchmark in these cases is a portfolio with a constant weight in a stock index and a bond index ([28, 90]).

*Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada, N2L 3G1, email: pieter.vanstaden@gmail.com

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

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

Given these observations, it therefore comes as no surprise that despite the large existing literature on the subject of deriving investment strategies designed to outperform a benchmark (see for example [41, 89, 83, 120, 119, 118, 11, 110, 18, 19, 20, 30, 2, 91]), this remains an active area of research (for recent examples, see [95, 15, 79, 3, 97, 55, 106, 1]). In addition, machine learning techniques are also increasingly used to address the problems associated with attempting to track or to outperform a given benchmark (see for example [88, 93, 75, 69, 8, 104, 70]).

However, in surveying the literature on deriving dynamic (multi-period) investment strategies for benchmark outperformance, we observe that the objective functions and assumptions that are very popular in the academic literature often do not appear to align very well with the performance metrics used and constraints applied by investors in practice.

The objective functions used in the literature often include the use of explicit or implied utility functions ([89, 83, 11, 110, 2, 3, 30, 91]), with the use of log utility (of outperformance) appearing to be especially popular.

Furthermore, the objective functions are often formulated in terms of the ratio of the active portfolio wealth to the benchmark wealth ([89, 83, 20, 18, 2, 3, 30, 91]). Using the wealth ratio, often in conjunction with log utility, means that contributions to or withdrawals from the portfolio cannot be included in the analysis due to analytical tractability considerations. This is undesirable in many contexts ([46, 45, 44]).

Another quantity enjoying significant popularity in the objective functions considered in the literature is the tracking *error*, which typically measures the standard deviation of the differences between the returns of the active portfolio and the returns of the benchmark (see for example [103, 67, 27, 25, 61]). However, many authors have criticized this metric ([65, 60, 23, 16]).

Finally, we note that the literature is typically concerned with obtaining closed-form solutions to the specified optimization problems, which necessarily makes idealized assumptions (e.g. continuous trading, unbounded leverage). Examples include [14, 89, 83, 120, 119, 118, 11, 110, 18, 19, 20, 30, 91, 9].

In this paper, we wish to address these considerations. In terms of objective functions, we limit our focus to two objectives for outperformance assessment, namely (i) the tracking difference and (ii) the information ratio.

(i) Tracking difference: In contrast to the tracking error (discussed above), the tracking *difference* is simply the difference between the *cumulative* returns of the active portfolio and that of the benchmark over a fixed time horizon([24]). For this reason, the tracking difference is recognized in the popular investment literature as a potentially more relevant and important metric than the tracking error for the investor (see for example [114, 17, 37, 60, 96]). Its importance is also recognized by regulators such as European Securities and Markets Authority, who requires its disclosure ([36]).

   We will focus on a tracking difference objective function which encapsulates both risk and reward in a natural manner.

(ii) Information ratio (IR): In a dynamic (or multi-period) context, the IR is typically defined ([9]) as the ratio of the expectation to the standard deviation of the difference between the terminal wealth of the active portfolio and the terminal wealth of the benchmark portfolio.
It is widely acknowledged that the IR is immensely popular in investment practice when measuring benchmark outperformance and for purposes of performance comparisons between funds ([57, 15, 64, 9]), despite concerns that it could be manipulated ([49, 50]). However, deriving *dynamic* investment strategies aimed at implicitly or explicitly maximizing the IR have not received significant attention in the academic literature, with the exception of [9, 120, 49].

Given these observations, our contributions in this paper are as follows:

- We formulate the investment benchmark outperformance problem as a stochastic optimal control problem and consider the IR and the tracking difference objective functions. While the IR objective is standard in the literature (see for example [9]), we propose a novel and straightforward tracking difference objective, which involves the minimization of the quadratic deviation (QD) of the wealth of the active portfolio compared to an elevated benchmark. Our treatment allows contributions/withdrawals from the portfolio, which is of interest to practitioners.
- In order to gain a theoretical understanding of the behavior of the resulting optimal investment strategies, we first solve the problems analytically under idealized assumptions. All closed-form results associated with the QD (tracking difference) objective are novel. We also present closed-form comparison results regarding certain critical aspects of the IR- and QD-optimal investment strategies.
- Under some assumptions, we prove that the traditional dynamic programming (DP) approach to these benchmark outperformance problems, assuming discrete portfolio rebalancing, requires the solution of a high-dimensional performance criterion (i.e. an approximation to a conditional expectation) in order to obtain the low dimensional optimal control.
- To compute optimal dynamic investment strategies under realistic constraints, we propose to use an NN representing control and directly solve a single sample optimization problem, which approximates the original stochastic optimal control problem. This direct approach exploits the lower dimensionality in optimal control and bypasses the problem of the approximation of conditional expectations associated with traditional DP methods. We note that this general idea was also used in [111]. However, in contrast with [111], we introduce time as a parameter directly in the NN, thus ensuring (under certain assumptions) that the (limiting) investment control is a continuous function of time, which is a desirable practical requirement. The idea of solving for the control directly, without using dynamic programming, has also been suggested in [102, 56]. Note that the approach in [56] uses the stacked NN technique as in [111]. In contrast, in our approach, the number of NN parameters does not increase with the

number of rebalancing times.

- Our numerical approach requires sample distributions to approximate the original stochastic optimal control problem. This is, of course, trivial if we restrict attention to parametric stochastic models. However, practitioners often prefer to test strategies by directly resampling the market data ([26, 33, 105, 22, 107, 5]). This is perhaps partly based on the belief that the empirical distribution is the least prejudiced estimate of the underlying distribution. Bootstrap resampling, first proposed by [35], is a simple but powerful technique to non-parametrically approximate sampling distributions, see, e.g., [98]. For illustrative purposes, here we use stationary block bootstrap resampling ([98]). Block bootstrap resampling is designed for weakly stationary series having serial dependence. We note that [100] and [99] suggest methods for resampling non-stationary time series, which we do not explore in this work. We emphasize that our method for solution of the optimal control is agnostic as to the particular technique used to augment the data. We only require a sufficiently large set of stochastic paths.
- Comparing the results using IR- and QD-optimal investment strategies obtained numerically using bootstrap resampling, we show how the closed-form comparison results apply qualitatively to in-sample investment results. In addition, the associated out-of-sample implications are often surprising. In particular, while the IR-optimal strategy retains a slightly higher probability of benchmark outperformance in-sample, the higher portfolio diversification associated with the QD-optimal strategy results in superior out-of-sample benchmark outperformance.

The remainder of the paper is organized as follows. Section 2 presents the problem formulation. Section 3 discusses analytical results under idealized assumptions. Section 4 discusses the inefficiencies of using DP-based techniques to solve benchmark outperformance problems such as the IR and QD problems in particular. Section 5 describes the preferred numerical solution approach based on approximating the optimal control by an NN, which allows the solution of the problems under more realistic constraints (bounded leverage, discrete rebalancing). Section 6 provides a comparison of the numerical method with the closed form solution, using simulated data. In addition, results obtained from resampling of historical data are presented. Finally, Section 7 concludes the paper and outlines possible future work.

2. Formulation.

We start by formulating the problem of outperforming a given benchmark investment strategy in general terms.

Let $T > 0$ denote the fixed investment time horizon/maturity of the active portfolio manager (henceforth simply referred to as the "investor"), and let time $t_0 \equiv 0$ denote the start of the investment period. The investor's controlled wealth process, with the control representing the investor's investment strategy, is denoted by $W(t)$, $t \in [t_0, T]$. Similarly, given some benchmark investment strategy, the benchmark portfolio's controlled wealth process is denoted by $\hat{W}(t)$, $t \in [t_0, T]$. For convenience, the time-t_0 wealth invested in both the benchmark and investor portfolio is assumed to be $w_0 = W(t_0) = \hat{W}(t_0) > 0$.

Assume that there are N_a candidate investment assets. Let $\hat{p}_i\left(t, \hat{\boldsymbol{X}}(t)\right)$ denote the proportion of the benchmark wealth $\hat{W}(t)$ invested in asset $i \in \{1, .., N_a\}$ at time $t \in [t_0, T]$, where $\hat{\boldsymbol{X}}(t)$ denotes the state of the system (or informally, the information) taken into account by the benchmark strategy for allocation decision $\hat{p}_i$. The vector $\hat{\boldsymbol{p}}\left(t, \hat{\boldsymbol{X}}(t)\right) = \left(\hat{p}_i\left(t, \hat{\boldsymbol{X}}(t)\right) : i = 1, .., N_a\right) \in \mathbb{R}^{N_a}$ denotes the asset allocation of the benchmark at time $t \in [t_0, T]$.

Similarly, let $p_i(t, \boldsymbol{X}(t))$ denote the proportion of the investor's wealth $W(t)$ invested in asset $i \in \{1, .., N_a\}$ at time $t \in [t_0, T]$, where $\boldsymbol{X}(t)$ denotes the information taken into account by the investor in making the asset allocation decision. As a concrete example, we consider the case where $\boldsymbol{X}(t) = \left(W(t), \hat{W}(t)\right)$ in Section 3,, but more general cases incorporating additional information in $\boldsymbol{X}(t)$ are also allowed in Section 5. The vector $\boldsymbol{p}(t, \boldsymbol{X}(t)) = (p_i(t, \boldsymbol{X}(t)) : i = 1, .., N_a) \in \mathbb{R}^{N_a}$ denotes the asset allocation of the investor at time $t \in [t_0, T]$.

Define the set of rebalancing events $\mathcal{T} \subseteq [t_0, T]$, where we have $\mathcal{T} = [t_0, T]$ in the case of continuous rebalancing, and a strict (discrete) subset $\mathcal{T} \subset [t_0, T]$ in the case of discrete rebalancing. The investor and benchmark investment strategies over the time horizon $[t_0, T]$, respectively, are then defined as the sets

$$\mathcal{P} = \{\boldsymbol{p}(t, \boldsymbol{X}(t)),\ t \in \mathcal{T}\}, \qquad \text{and} \qquad \hat{\mathcal{P}} = \left\{\hat{\boldsymbol{p}}\left(t, \hat{\boldsymbol{X}}(t)\right),\ t \in \mathcal{T}\right\}. \tag{2.1}$$

Here we implicitly assume that the investor and benchmark strategies invest in the same N_a underlying assets, which is relevant in the case of analytical solutions (Section 3). However, this requirement is also relaxed in the numerical solution approach discussed in Section 5.

We define $\mathcal{A}$ as the set of admissible controls, and $\mathcal{Z}$ as the set of admissible values of each vector $\boldsymbol{p}(t, \boldsymbol{X}(t))$, i.e., $\mathcal{P} \in \mathcal{A}$ if and only if $\mathcal{P} = \{\boldsymbol{p}(t, \boldsymbol{X}(t)) \in \mathcal{Z} : t \in \mathcal{T}\}$. Note that $\mathcal{Z}$, and therefore by extension $\mathcal{A}$, encode the investment constraints faced by the investor, such as leverage constraints or short-selling restrictions.

Since the investor wishes to *outperform* the benchmark according to a performance metric adopted in practice, we introduce two investment objectives to achieve this aim in the following subsections. In terms of notation, let $E_{\mathcal{P}}^{t_0, w_0}[\cdot]$ denote the expectation of some quantity taken with respect to a given initial wealth $w_0 = W(t_0) = \hat{W}(t_0)$ at time $t_0 = 0$, and using control $\mathcal{P} \in \mathcal{A}$ over $[t_0, T]$. The benchmark strategy $\hat{\mathcal{P}}$ that the investor wishes to outperform remains implicit in this notation. Similarly, we will use $Var_{\mathcal{P}}^{t_0, w_0}[\cdot]$ and $P_{\mathcal{P}}^{t_0, w_0}[\cdot]$ to denote the variance and probability, respectively, calculated under the control $\mathcal{P}$ and initial time and wealth given by (t_0, w_0).

2.1. Information ratio: Problem $IR(\gamma)$. The first investment objective involves maximizing the information ratio (IR), which in a dynamic setting is defined as ([49, 9])

$$\mathcal{IR}_{\mathcal{P}}^{t_0,w_0} = \frac{E_{\mathcal{P}}^{t_0,w_0}\left[W(T) - \hat{W}(T)\right]}{Stdev_{\mathcal{P}}^{t_0,w_0}\left[W(T) - \hat{W}(T)\right]}. \tag{2.2}$$

As discussed in [9], maximizing the IR (2.2) is achieved by solving the following mean-variance (MV) optimization problem with scalarization parameter ρ,

$$\sup_{\mathcal{P}\in\mathcal{A}} \left\{ E_{\mathcal{P}}^{t_0,w_0}\left[W(T) - \hat{W}(T)\right] - \rho \cdot Var_{\mathcal{P}}^{t_0,w_0}\left[W(T) - \hat{W}(T)\right] \right\}, \quad \rho > 0. \tag{2.3}$$

To solve (2.3), we use the embedding technique of [78, 121], which states that for any $\rho > 0$ and the associated control $\mathcal{P}_{ir}^{*} \in \mathcal{A}$ maximizing (2.3), there exists a value of an embedding parameter γ such that $\mathcal{P}_{ir}^{*} \in \mathcal{A}$ is also optimal for the following problem[1],

$$(IR(\gamma)) : \inf_{\mathcal{P}\in\mathcal{A}} E_{\mathcal{P}}^{t_0,w_0}\left[\left(W(T) - \left[\hat{W}(T) + \gamma\right]\right)^2\right], \qquad \gamma > 0. \tag{2.4}$$

Note that (2.4) is formulated here only for the range $\gamma > 0$ in order to ensure that economically meaningful strategies for benchmark outperformance are obtained.

As a result of the aforementioned equivalence of (2.3) and (2.4), we will subsequently refer to (2.4) simply as the IR (maximization) problem, abbreviated by $IR(\gamma)$. The exact relationship between γ in (2.4) and ρ in (2.3) is not important for the purposes of this paper, and it is indeed also of limited practical significance to the investor. For further clarification, the following remark highlights some practical aspects of our preference for formulation (2.4).

Remark 2.1. (Time-consistency of the $IR(\gamma)$-optimal control) As elaborated in [81, 45], there appears to be some controversy in the literature regarding the time-consistency (or lack thereof) of the optimal controls associated with problems of the form (2.4). By analogy with dynamic MV optimization (see [10, 13]), the IR-optimal control for the embedding problem (2.4) is typically time-inconsistent from the perspective of the MV formulation (2.3). This raises practical concerns as to whether the resulting IR-optimal control is in fact feasible to implement as a trading strategy. However, it should be emphasized that time-consistency is ultimately a matter of perspective, since for a fixed value of γ in (2.4), the resulting $IR(\gamma)$-optimal control is in fact a time-consistent control from the perspective of the quadratic objective (2.4), and is therefore clearly feasible as a trading strategy ([109]). As discussed in [115]

[1] Formally proving the equivalence of problems (2.3) and (2.4) proceeds along the same lines as the proof of the embedding result of [78, 121], and is therefore omitted. As shown in [29], the embedding result holds in great generality, in that it does not require restrictions on the admissible set $\mathcal{A}$ or on the underlying wealth dynamics.

and elaborated further below, a quadratic objective such as (2.4) also allows for a straightforward interpretation in terms of a "target" (in this case, $\hat{W}(T)+\gamma$). As a result, in this paper we always view the IR-optimal control as the time-consistent investment strategy that minimizes the induced objective function (2.4), and correspondingly formulate our results in terms of the embedding parameter γ.

The following additional observations regarding the IR objective (2.4) are relevant to the subsequent results:

(i) The investor wishing to maximize the IR effectively sets an *elevated* benchmark terminal wealth value, $\hat{W}(T)+\gamma$, and minimizes the (expected) quadratic deviation of the investor's wealth $W(T)$ from this elevated target.

(ii) In Section 3 below, we show that under some conditions, the IR problem (2.4) is equivalent to the more intuitive one-sided quadratic objective,

$$\inf_{\mathcal{P}\in\mathcal{A}} E_{\mathcal{P}}^{t_0,w_0}\left[\left(\min\left\{W(T)-\left[\hat{W}(T)+\gamma\right],0\right\}\right)^2\right], \qquad \gamma>0, \tag{2.5}$$

where only the *shortfall* of $W(T)$ below the elevated target $\hat{W}(T)+\gamma$ is penalized. While the equivalence between (2.4) and (2.5) can only be proven analytically under certain assumptions, numerical results nevertheless suggest that the results using (2.4) and (2.5) are indistinguishable even in more general cases where the conditions for analytical equivalence do not hold.

We now consider our second objective for outperforming the benchmark.

2.2. Tracking difference: Problem $QD(\beta)$. As discussed in the Introduction, the tracking difference measures the cumulative performance gap between the investor's portfolio and the benchmark portfolio over the time horizon $[t_0,T]$ ([24]).

In a dynamic setting, we propose the following straightforward objective function based on minimizing the quadratic deviation (QD) of the investor's terminal wealth from the terminal wealth of an elevated benchmark,

$$(QD(\beta)): \inf_{\mathcal{P}\in\mathcal{A}} E_{\mathcal{P}}^{t_0,w_0}\left[\left(W(T)-e^{\beta T}\hat{W}(T)\right)^2\right], \qquad \beta>0. \tag{2.6}$$

We will subsequently refer to problem (2.6) as the QD problem, and we make the following observations:

(i) The attractiveness of the formulation (2.6) lies in its simplicity, since the objective of obtaining a favorable tracking difference, widely publicized as a quantity of key interest to investors and regulators alike ([114, 17, 37, 60, 96, 24, 36, 66]) is the central object of consideration.

(ii) The parameter β in the QD problem (2.6) has a conveniently practical interpretation as the annual outperformance spread that the investor targets for the tracking difference of the active portfolio. Assuming that the active portfolio has access to at least the same set of assets as the benchmark, then as $\beta \to 0$, the optimal strategy is to simply invest in the benchmark. As β increases, we can expect that the optimal strategy will incur more risk, as measured by the value of the objective function, in order to achieve the desired outperformance.

(iii) By formulating (2.6) in terms of wealth, not only do we respect the cumulative aspect of the definition of the tracking difference, but the formulation also allows for the treatment of contributions to and withdrawals from the portfolio without difficulty (see Sections 3 and 5).

(iv) Like the IR problem (2.4), the QD problem (2.6) also formulates the outperformance objective in terms of an elevated benchmark terminal wealth value. However, in the case of the QD problem, the elevation is applied to $\hat{W}(T)$ by the multiplicative scaling factor $e^{\beta T}$, in contrast to the IR problem where the elevation is additive (i.e. by adding a constant γ to $\hat{W}(T)$ in (2.4)). The investor using the QD objective therefore wishes, where possible, to outperform the benchmark terminal wealth by a constant *factor*, and not by a constant *amount* as in the case of the IR problem.

(v) As in the case of the IR problem (see (2.5)), we show in Section 3 that under some conditions, the QD problem (2.6) also admits the equivalent, and perhaps more intuitive, one-sided quadratic formulation,

$$\inf_{\mathcal{P}\in\mathcal{A}} E_{\mathcal{P}}^{t_0,w_0}\left[\left(\min\left\{W(T) - e^{\beta T}\hat{W}(T), 0\right\}\right)^2\right], \qquad \beta > 0, \tag{2.7}$$

where only underperformance relative to the elevated benchmark $e^{\beta T}\hat{W}(T)$ is penalized.

In summary, two fundamentally different yet practical investment objectives for outperforming a given benchmark are considered. The following sections are devoted to explore the resulting investment outcomes of the IR and QD problems, using both closed-form solutions (where available) and numerical solutions.

3. Analytical (closed-form) solutions.

In order to gain a theoretical understanding of the behavior of the optimal investment strategies associated with the IR and QD objectives, we first solve the problems analytically under idealized assumptions. All closed-form results associated with the QD (tracking difference) objective, as well as selected results associated with the IR objective, are novel. We also present closed-form comparison results regarding certain critical aspects of the IR- and QD-optimal investment strategies. In our analytical solutions, we explicitly allow for contributions to the portfolio and jumps in the risky asset processes, both of which only receives limited treatment in the existing benchmark outperformance literature ([14, 89],[83, 120, 119, 118, 11, 110, 18, 19, 20, 30, 91, 2]).

Assumption 3.1 summarizes the assumptions required for deriving the subsequent closed-form results, which we emphasize are not required in the case of the numerical solutions discussed in Section 5.

Assumption 3.1. (No market frictions, continuous rebalancing) For the purposes of the closed-form results, we assume that trading continues in the event of insolvency. Specifically, trading continues even if $W(t) < 0$ for some $t \in [t_0, T]$. No transaction costs are applicable, and no investment constraints (such as leverage or short-selling restrictions) are in effect. In addition, the portfolios are rebalanced continuously, and cash is contributed to the investor and benchmark portfolios at a constant rate of $q \geq 0$ per year.

Note that the cash contributions are made to both the investor and the benchmark portfolios in order to ensure that a meaningful performance comparison is obtained.

In this section, the N_a underlying assets are assumed to consist of one risk-free asset and N_a^r risky assets evolving according to specified dynamics. Let $\boldsymbol{\varrho}(t, \boldsymbol{X}(t)) = \left(\varrho_1(t, \boldsymbol{X}(t)), ..., \varrho_{N_a^r}(t, \boldsymbol{X}(t))\right) \in \mathbb{R}^{N_a^r}$ and $\hat{\boldsymbol{\varrho}}\left(t, \hat{\boldsymbol{X}}(t)\right) = \left(\hat{\varrho}_1\left(t, \hat{\boldsymbol{X}}(t)\right), ..., \hat{\varrho}_{N_a^r}\left(t, \hat{\boldsymbol{X}}(t)\right)\right) \in \mathbb{R}^{N_a^r}$ denote the proportional allocations of the investor and benchmark wealth, respectively, to each of the *risky* assets at time $t \in [t_0, T]$. Specifically, $\varrho_i(t, \boldsymbol{X}(t))$ denotes the proportion of the investor's wealth $W(t)$ invested in risky asset i at time t given information $\boldsymbol{X}(t)$, while $\hat{\varrho}_i\left(t, \hat{\boldsymbol{X}}(t)\right)$ denotes the proportion of benchmark wealth $\hat{W}(t)$ invested in risky asset i at time t given information $\hat{\boldsymbol{X}}(t)$.

With regards to the benchmark strategy, we introduce the following assumption.

Assumption 3.2. (Information known about the benchmark strategy) For the closed-form solutions of this section, we assume that the benchmark's risky asset allocation strategy is an adapted feedback control of the form $\hat{\boldsymbol{\varrho}}\left(t, \hat{\boldsymbol{X}}(t)\right) = \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right), t \in [t_0, T]$, and that the investor is limited to investing in the same set of underlying assets as the benchmark. We also assume that the investor can instantaneously observe the vector $\hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right)$ at each $t \in [t_0, T]$, so that the investor wishes to derive $\boldsymbol{\varrho}(t, \boldsymbol{X}(t)) = \boldsymbol{\varrho}\left(t, W(t), \hat{W}(t), \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right)\right), t \in [t_0, T]$, the adapted feedback control representing the fraction of the investor's wealth $W(t)$ invested in each risky asset at time t according to the investor's strategy.

We observe that Assumption 3.2 is reasonable in the case of the investment benchmarks typically considered by government pension plans (see for example [21, 53]), as well as many of the popular benchmarks used in the literature and in practice ([14, 63, 120, 11, 4]).. Note that in the numerical solutions (Section 5), the requirement that the investor invests in the same assets as the benchmark is relaxed.

Combining definition (2.1) and Assumption 3.2, we therefore consider the following forms of the investor and benchmark strategies in this section,

$$\mathcal{P}=\left\{\boldsymbol{p}\left(t,\boldsymbol{X}\left(t\right)\right)=\left(1-\sum_{i=1}^{N_a^r}\varrho_i\left(t,\boldsymbol{X}\left(t\right)\right),\varrho_1\left(t,\boldsymbol{X}\left(t\right)\right),...,\varrho_{N_a^r}\left(t,\boldsymbol{X}\left(t\right)\right)\right):t\in\left[t_0,T\right]\right\},$$

$$\hat{\mathcal{P}}=\left\{\hat{\boldsymbol{p}}\left(t,\hat{W}\left(t\right)\right)=\left(1-\sum_{i=1}^{N_a^r}\hat{\varrho}_i\left(t,\hat{W}\left(t\right)\right),\hat{\varrho}_1\left(t,\hat{W}\left(t\right)\right),...,\hat{\varrho}_{N_a^r}\left(t,\hat{W}\left(t\right)\right)\right):t\in\left[t_0,T\right]\right\},$$

(3.1)

where $\boldsymbol{X}(t)=\left(W(t),\hat{W}(t),\hat{\boldsymbol{\varrho}}\left(t,\hat{W}(t)\right)\right)$. Due to the form of (3.1), we will informally refer to the *risky* asset allocations $\boldsymbol{\varrho}(t,\boldsymbol{X}(t))$ and $\hat{\boldsymbol{\varrho}}\left(t,\hat{W}(t)\right)$ as the investor and benchmark investment strategies, respectively, although the original definition (3.1) will be used in the numerical solutions of Section 5.

Given Assumption 3.1, Assumption 3.2 and the form of the controls (3.1), the investor's set of admissible controls can be written in terms of only the risky asset allocation vector $\boldsymbol{\varrho}$,

$$\mathcal{A}_0=\left\{\boldsymbol{\varrho}(t,x)=\boldsymbol{\varrho}(t,w,\hat{w},\hat{\boldsymbol{\varrho}}(t,w))\,|\,\boldsymbol{\varrho}:[t_0,T]\times\mathbb{R}^{N_a^r+2}\to\mathbb{R}^{N_a^r}\right\}, \tag{3.2}$$

so that the IR- and QD-problems analyzed in this section are of the following form,

$$(IR(\gamma)):\inf_{\boldsymbol{\varrho}\in\mathcal{A}_0}E_{\boldsymbol{\varrho}}^{t_0,w_0}\left[\left(W(T)-\left[\hat{W}(T)+\gamma\right]\right)^2\right],\qquad\gamma>0. \tag{3.3}$$

$$(QD(\beta)):\inf_{\boldsymbol{\varrho}\in\mathcal{A}_0}E_{\boldsymbol{\varrho}}^{t_0,w_0}\left[\left(W(T)-e^{\beta T}\hat{W}(T)\right)^2\right],\qquad\beta>0. \tag{3.4}$$

3.1. Asset and wealth dynamics. Since the closed-form solutions are based on specified underlying dynamics, let $S_0(t)$ denote the unit value of the risk-free asset at time $t\in[t_0,T]$, with dynamics given in terms of the risk-free rate $r>0$ as

$$dS_0(t)=rS_0(t)\,dt. \tag{3.5}$$

Define the risky asset value vector $\boldsymbol{S}(t)=\left(S_i(t):i=1,...,N_a^r\right)^\top$, where the ith component $S_i(t)$ denotes the unit value of the risky asset i at time $t\in[t_0,T]$. The superscript "$\top$" denotes the transpose. We allow for any of the typical finite-activity jump-diffusion models in finance (see for example [85, 73]) for the dynamics of $S_i(t)$. Let $\boldsymbol{Z}(t)=\left(Z_i(t):i=1,...,N_a^r\right)^\top$ denote a standard N_a^r-dimensional Brownian motion. Let $\boldsymbol{\xi}=\left(\xi_i:i=1,...,N_a^r\right)^\top$, where ξ_i denotes the random variable giving the jump multiplier associated with the ith risky asset with corresponding probability density function (pdf) $f_{\xi_i}(\xi_i)$. We also define

$$\kappa_i^{(1)}=\mathbb{E}\left[\xi_i-1\right],\qquad\kappa_i^{(2)}=\mathbb{E}\left[\left(\xi_i-1\right)^2\right],\qquad i=1,...,N_a^r, \tag{3.6}$$

as well as $\boldsymbol{\kappa}^{(1)} = \left(\kappa_i^{(1)} : i = 1, ..., N_a^r\right)^\top$ and $\boldsymbol{\kappa}^{(2)} = \left(\kappa_i^{(2)} : i = 1, ..., N_a^r\right)^\top$. If a jump occurs in the dynamics of risky asset i at time t, its value is assumed to jump from $S_i(t^-)$ to $S_i(t) = \xi_i \cdot S_i(t^-)$, where, given any functional $\psi(t), t \in [t_0, T]$, we use the notation $\psi(t^-)$ and $\psi(t^+)$ as shorthand for the one-sided limits $\psi(t^-) = \lim_{\epsilon \downarrow 0} \psi(t - \epsilon)$ and $\psi(t^+) = \lim_{\epsilon \downarrow 0} \psi(t + \epsilon)$, respectively. We assume that the jump components of the different risky asset processes are independent, so that $\boldsymbol{\xi}$ has independent components. However, while we can relax this assumption without technical difficulties (see for example [74]), this would be associated with the corresponding disadvantage of significantly increased notational complexity; the assumption of independence is therefore for ease of exposition.

Let $\boldsymbol{\pi}(t) = (\pi_i(t) : i = 1, ..., N_a^r)^\top$ denote a vector of N_a^r independent Poisson processes, with each $\pi_i(t)$ having the corresponding intensity $\lambda_i \geq 0$, and define $\boldsymbol{\lambda} = (\lambda_i : i = 1, ..., N_a^r)^\top$. We assume that ξ_i, $\pi_j(t)$ and $Z_k(t)$ are mutually independent for all $i, j, k \in \{1, ..., N_a^r\}$.

The vector of risky asset drift coefficients under the objective (or real-world) probability measure is denoted by $\boldsymbol{\mu} = (\mu_i : i = 1, ..., N_a^r)^\top$. Let $\boldsymbol{\sigma} = (\sigma_{i,j})_{i,j=1,...,N_a^r} \in \mathbb{R}^{N_a^r \times N_a^r}$ denote the volatility matrix, and define

$$\boldsymbol{\Sigma} = \boldsymbol{\sigma}\boldsymbol{\sigma}^\top, \qquad \boldsymbol{\Lambda} = \mathrm{diag}\left(\lambda_i \kappa_i^{(2)} : i = 1, ..., N_a^r\right). \tag{3.7}$$

We make the standard assumptions that $\mu_i > r$, for all i, and assume that the covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{\sigma}\boldsymbol{\sigma}^\top$ is positive definite (see for example [12, 121]).

The dynamics of $S_i(t)$ is therefore assumed to be of the form

$$\frac{dS_i(t)}{S_i(t^-)} = \left(\mu_i - \lambda_i \kappa_i^{(1)}\right) \cdot dt + \sum_{j=1}^{N_a^r} \sigma_{ij} \cdot dZ_j(t) + d\left(\sum_{k=1}^{\pi_i(t)} \left(\xi_i^{(k)} - 1\right)\right), \qquad i = 1, ..., N_a^r, \tag{3.8}$$

where $\xi_i^{(k)}$ are i.i.d. random variables with the same distribution as ξ_i. To lighten subsequent notation, define the vector $d\boldsymbol{\mathcal{N}}(t) = \left(\int_0^\infty (\xi_i - 1) N_i(dt, d\xi_i) : i = 1, ..., N_a^r\right)^\top$, where N_i is the Poisson random measure ([92]) corresponding to the dynamics of $S_i(t)$ in (3.8). We also define the following combinations of parameters,

$$\boldsymbol{\alpha} = \left(\mu_i - r - \lambda_i \kappa_i^{(1)} : i = 1, ..., N_a^r\right)^\top, \qquad \tilde{\boldsymbol{\mu}} = (\mu_i - r : i = 1, ..., N_a^r)^\top, \tag{3.9}$$

and

$$\eta = \tilde{\boldsymbol{\mu}}^\top \cdot (\boldsymbol{\Sigma} + \boldsymbol{\Lambda})^{-1} \cdot \tilde{\boldsymbol{\mu}}. \tag{3.10}$$

With strategies (3.1), and dynamics (3.5)-(3.8), the investor and benchmark controlled wealth processes therefore have the following dynamics for $t \in (t_0, T]$, respectively,

$$dW(t) = \left\{ W(t^-) \cdot \left[r + \boldsymbol{\alpha}^\top \boldsymbol{\varrho}(t, \boldsymbol{X}(t)) \right] + q \right\} \cdot dt + W(t^-) \left(\boldsymbol{\varrho}(t, \boldsymbol{X}(t)) \right)^\top \boldsymbol{\sigma} \cdot d\boldsymbol{Z}(t)$$
$$+ W(t^-) \left(\boldsymbol{\varrho}(t, \boldsymbol{X}(t)) \right)^\top \cdot d\boldsymbol{\mathcal{N}}(t), \tag{3.11}$$

$$d\hat{W}(t) = \left\{ \hat{W}(t^-) \cdot \left[r + \boldsymbol{\alpha}^\top \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) \right] + q \right\} \cdot dt + \hat{W}(t^-) \left(\hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) \right)^\top \boldsymbol{\sigma} \cdot d\boldsymbol{Z}(t)$$
$$+ \hat{W}(t^-) \left(\hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) \right)^\top \cdot d\boldsymbol{\mathcal{N}}(t), \tag{3.12}$$

where $W(t) = \hat{W}(t) = w_0$ and $\boldsymbol{X}(t) = \left(W(t), \hat{W}(t), \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) \right)$, while $q \geq 0$ denotes the constant rate per year of continuous cash injection into the portfolios (see Assumption 3.1).

In the following subsections, we derive and compare the closed-form solutions to the IR and QD problems subject to Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12).

3.2. Analytical solution: $IR(\gamma)$ problem.

We have the following verification theorem and corresponding Hamilton-Jacobi-Bellman (HJB) equation for the IR problem (3.3).

Theorem 3.3. *(IR problem: Verification theorem) Suppose that for all $(y, t) \in \mathbb{R} \times [t_0, T]$, there exist functions $V_{ir}(y, t) : \mathbb{R} \times [t_0, T] \to \mathbb{R}$ and $\boldsymbol{u}^*_{ir} : \mathbb{R} \times [t_0, T] \to \mathbb{R}^{N_a^r}$ with the following properties: (i) V_{ir} and $\boldsymbol{u}^*_{ir}$ are sufficiently smooth and solve the HJB partial integro-differential equation (PIDE) (3.13)-(3.14), and (ii) the function $\boldsymbol{u}^*_{ir}$ attains the pointwise supremum in (3.13).*

$$\frac{\partial V_{ir}}{\partial t} + \inf_{\boldsymbol{u} \in \mathbb{R}^{N_a^r}} \left\{ \left[ry + \boldsymbol{\alpha}^\top \boldsymbol{u} \right] \cdot \frac{\partial V_{ir}}{\partial y} + \frac{1}{2} \boldsymbol{u}^\top \boldsymbol{\Sigma} \boldsymbol{u} \cdot \frac{\partial^2 V_{ir}}{\partial y^2} - \left(\sum_{i=1}^{N_a^r} \lambda_i \right) \cdot V_{ir} \right.$$
$$\left. + \sum_{i=1}^{N_a^r} \lambda_i \int_0^\infty V_{ir}\left(y + u_i(\xi_i - 1), t\right) \cdot f_{\xi_i}(\xi_i)\, d\xi_i \right\} = 0, \tag{3.13}$$

$$V_{ir}(y, T) = (y - \gamma)^2. \tag{3.14}$$

Define the auxiliary process $Y_{ir}(t)$ by

$$Y_{ir}(t) := W(t) - \hat{W}(t), \quad \forall t \in (t_0, T], \quad \text{with} \quad Y_{ir}(t_0) = y_0 = 0. \tag{3.15}$$

Let the auxiliary control $\boldsymbol{u}(t) := \boldsymbol{u}(Y_{ir}(t), t) := \boldsymbol{u}(Y_{ir}(t), t; \boldsymbol{X}(t))$ be given by

$$\boldsymbol{u}(t) := W(t) \cdot \boldsymbol{\varrho}(t, \boldsymbol{X}(t)) - \hat{W}(t) \cdot \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right), \text{ where } \boldsymbol{X}(t) = \left(W(t), \hat{W}(t), \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) \right). \tag{3.16}$$

Let $\mathcal{A}_{\boldsymbol{u},0} = \left\{ \boldsymbol{u}(t) = \boldsymbol{u}(y, t; x) \middle| \boldsymbol{u} : \mathbb{R} \times [t_0, T] \to \mathbb{R}^{N_a^r} \right\}$. Then under Assumption 3.1, Assump-

tion 3.2 and wealth dynamics (3.11)-(3.12), V_{ir} is the value function and $\boldsymbol{u}_{ir}^*$ is the optimal control for the following control problem,

$$\inf_{\boldsymbol{u}\in\mathcal{A}_{\boldsymbol{u},0}} E_{\boldsymbol{u}}^{t_0,y_0}\left[\left(Y_{ir}\left(T\right)-\gamma\right)^2\right], \qquad \gamma>0. \tag{3.17}$$

Proof. See Appendix A.1. ■

By solving the HJB PIDE (3.13)-(3.14), the following lemma presents the IR-optimal investment strategy.

Lemma 3.4. *(IR-optimal investment strategy) Suppose that Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12) are applicable. Then the optimal fraction of the investor's wealth invested in risky asset $i\in\{1,...,N_a^r\}$ for problem $IR(\gamma)$ in (3.3) is given by the i^{th} component of the vector $\boldsymbol{\varrho}_{ir}^*\left(t,\boldsymbol{X}_{ir}^*\left(t\right)\right)$, where*

$$W_{ir}^*\left(t\right)\cdot\boldsymbol{\varrho}_{ir}^*\left(t,\boldsymbol{X}_{ir}^*\left(t\right)\right)=\left[\gamma e^{-r(T-t)}-\left(W_{ir}^*\left(t\right)-\hat{W}\left(t\right)\right)\right]\cdot\left(\boldsymbol{\Sigma}+\boldsymbol{\Lambda}\right)^{-1}\tilde{\boldsymbol{\mu}}+\hat{W}\left(t\right)\cdot\hat{\boldsymbol{\varrho}}\left(t,\hat{W}\left(t\right)\right), \tag{3.18}$$

with $W_{ir}^*\left(t\right)$ denoting the investor's wealth process (3.11) under the IR-optimal control $\boldsymbol{\varrho}_{ir}^*$, and $\boldsymbol{X}_{ir}^*\left(t\right)=\left(W_{ir}^*\left(t\right),\hat{W}\left(t\right),\hat{\boldsymbol{\varrho}}\left(t,\hat{W}\left(t\right)\right)\right)$. This control results in an optimal information ratio (2.2) of

$$\mathcal{IR}_{\boldsymbol{\varrho}_{ir}^*}^{t_0,w_0}=\left(e^{\eta T}-1\right)^{1/2}, \tag{3.19}$$

where η is given by (3.10).

Proof. See Appendix A.1. ■

It is noteworthy that the IR-optimal control $\boldsymbol{\varrho}_{ir}^*\left(t,\boldsymbol{X}_{ir}^*\left(t\right)\right)$ only depends on the instantaneous benchmark allocation $\hat{\boldsymbol{\varrho}}\left(t,\hat{W}\left(t\right)\right)$ at time t, and not on the future or the past of the benchmark investment strategy. The contribution rate q does not appear in the solution (3.18), which follows from the fortunate cancellation of terms in the auxiliary process $Y_{ir}\left(t\right)$. The optimal IR (3.19) extends the known IR results of [50] to the case of multiple risky assets containing jumps in their associated value dynamics. Specifically, if we consider the case of only a single risky asset with no jumps (i.e. setting $\lambda_1=0$), the expression for η in (3.10) reduces to $\eta=\left(\mu_1-r\right)^2/\sigma_1^2$, so that the optimal IR (3.19) reduces to the result reported in [49, 50].

The following lemma presents an important property of the IR-optimal strategy (3.18) when sufficient outperformance can be assured.

Lemma 3.5. *(IR: Matching the benchmark risky asset amounts) Given Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12), suppose that at some time $\bar{t}\in(t_0,T]$, the*

IR-optimal investor observes a wealth value $W_{ir}^{*}\left(\bar{t}\right)$ *of*

$$W_{ir}^{*}\left(\bar{t}\right)=\gamma e^{-r\left(T-\bar{t}\right)}+\hat{W}\left(\bar{t}\right). \tag{3.20}$$

Then for the remainder of the investment time horizon $t\in\left[\bar{t},T\right]$*, the IR-optimal investor (using strategy (3.18)) will simply match the benchmark strategy in terms of the amounts invested in the risky assets. In other words,*

$$W_{ir}^{*}\left(t\right)\cdot\boldsymbol{\varrho}_{ir}^{*}\left(t,\boldsymbol{X}_{ir}^{*}\left(t\right)\right)=\hat{W}\left(t\right)\cdot\hat{\boldsymbol{\varrho}}\left(t,\hat{W}\left(t\right)\right),\qquad\forall t\in\left[\bar{t},T\right]. \tag{3.21}$$

Proof. See Appendix A.1. ■

Note that Lemma 3.5 does *not* imply that the investor and benchmark strategies $\boldsymbol{\varrho}_{ir}^{*}$ and $\hat{\boldsymbol{\varrho}}$ are equal, since if (3.20) is satisfied at some $\bar{t}\in\left(t_{0},T\right]$, the results of Appendix A.1 (see (A.5)) imply that $W_{ir}^{*}\left(t\right)>\hat{W}\left(t\right)$, $\forall t\in\left[\bar{t},T\right]$.

Lemma 3.6 below reports that condition (3.20) is never satisfied in the special case when there are no jumps in the risky asset processes, with the implication that equivalence of problems (2.4) and (2.5) can be established analytically[2].

Lemma 3.6. *(IR: equivalence with only penalizing underperformance) If Assumption 3.1, Assumption 3.2, and wealth dynamics (3.11)-(3.12) apply with no jumps (i.e.* $\boldsymbol{\lambda}=\mathbf{0}\in\mathbb{R}^{N_{a}^{r}}$*) in the risky asset processes (3.8), then*

$$W_{ir}^{*}\left(t\right)<\gamma e^{-r\left(T-t\right)}+\hat{W}\left(t\right),\qquad\forall t\in\left[t_{0},T\right]. \tag{3.22}$$

As a result, in this case the IR optimization problem (2.4) is equivalent to the one-sided quadratic problem (2.5), where only the underperformance of the investor's portfolio (compared to the elevated benchmark) is penalized.

Proof. See Appendix A.1. ■

If the assumptions of this section are violated, both Lemma 3.5 and the more restrictive Lemma 3.6 provide valuable intuition for understanding the behavior of the IR-optimal investment strategies, which we will demonstrate in Section 6.

The following lemma shows that if we apply the assumption of no jumps as in Lemma 3.6, then the probability of the IR investor underperforming the benchmark admits a simple analytical expression. Note that we prefer formulating the result in the negative sense of *under*performance, since it directly expresses a key quantity of concern for the active investor.

Lemma 3.7. *(IR: probability of underperformance) If Assumption 3.1, Assumption 3.2, and wealth dynamics (3.11)-(3.12) apply with no jumps (i.e.* $\boldsymbol{\lambda}=\mathbf{0}\in\mathbb{R}^{N_{a}^{r}}$*) in the risky asset processes (3.8), the probability of the IR-optimal wealth falling below the benchmark wealth at*

[2] The proof of Lemma 3.6 uses the results of [31], which hold only when there are no jumps in a risky asset process. However, even in the case where there are jumps, the behavior of the optimal strategy typically satisfies (3.22), but this can only be verified numerically.

any $t \in (t_0, T]$ is given by

$$P^{t_0,w_0}_{\boldsymbol{\varrho}^*_{ir}}\left[W^*_{ir}(t) \leq \hat{W}(t)\right] = \Phi\left(-\frac{3}{2}\sqrt{\eta t}\right), \qquad \forall t \in (t_0, T], \tag{3.23}$$

where Φ denotes the standard normal cumulative distribution function (CDF), and η is as defined in (3.10).

Proof. See Appendix A.1. ∎

Remark 3.8 (γ independence of equation (3.23)). Note that the IR-optimal probability of underperformance (3.23) does not depend on the value of γ. We conjecture that this lack of dependence on γ is due to the assumption that trading continues if insolvent (this is commonly required in order to obtain closed form solutions). In the pure mean variance case, [80] prove the 80% rule, which states that given any expected value for final wealth, no matter how large, there is at least an 80% probability of reaching this target. However [116] show that this is entirely due to the allowance of trading if insolvent.

3.3. Analytical solution: $QD(\beta)$ problem.

The closed-form solutions associated with the novel objective function (2.6) are now discussed. The following verification theorem reports the HJB equation satisfied in the case of the QD problem (3.4).

Theorem 3.9. *(QD problem: Verification theorem) Suppose that for all $(y,t) \in \mathbb{R} \times [t_0, T]$, there exist functions $V_{qd}(y,t) : \mathbb{R} \times [t_0, T] \to \mathbb{R}$ and $\boldsymbol{v}^*_{qd}(y,t) : \mathbb{R} \times [t_0, T] \to \mathbb{R}^{N^r_a}$ with the following two properties. (i) V_{qd} and $\boldsymbol{v}^*_{qd}$ are sufficiently smooth and solve the HJB PIDE (3.24)-(3.25), and (ii) the function $\boldsymbol{v}^*_{qd}(y,t)$ attains the pointwise supremum in (3.24).*

$$\frac{\partial V_{qd}}{\partial t} + \inf_{\boldsymbol{v}\in\mathbb{R}^{N^r_a}}\left\{\left[ry + q\left(1 - e^{\beta T}\right) + \boldsymbol{\alpha}^\top \boldsymbol{v}\right] \cdot \frac{\partial V_{qd}}{\partial y} + \frac{1}{2}\boldsymbol{u}^\top \boldsymbol{\Sigma} \boldsymbol{u} \cdot \frac{\partial^2 V_{qd}}{\partial y^2} - \left(\sum_{i=1}^{N^r_a} \lambda_i\right) \cdot V_{qd} \right.$$

$$\left. + \sum_{i=1}^{N^r_a} \lambda_i \int_0^\infty V_{qd}\left(y + u_i(\xi_i - 1), t\right) \cdot f_{\xi_i}(\xi_i)\, d\xi_i \right\} = 0, \tag{3.24}$$

$$V_{qd}(y, T) = y^2. \tag{3.25}$$

Define the auxiliary process $Y_{qd}(t)$ by

$$Y_{qd}(t) := W(t) - e^{\beta T}\hat{W}(t), \quad \forall t \in (t_0, T], \quad \text{with} \quad Y_{qd}(t_0) = y_0 = w_0\left(1 - e^{\beta T}\right). \tag{3.26}$$

Let the auxiliary control $\boldsymbol{v}(t) := \boldsymbol{v}(Y_{qd}(t), t) := \boldsymbol{v}(Y_{qd}(t), t; \boldsymbol{X}(t))$ *be given by*

$$v(t) := W(t) \cdot \boldsymbol{\varrho}(t, \boldsymbol{X}(t)) - e^{\beta T} \hat{W}(t) \cdot \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right), \text{ where } \boldsymbol{X}(t) i = \left(W(t), \hat{W}(t), \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right)\right). \tag{3.27}$$

Let $\mathcal{A}_{\boldsymbol{u},0}$ *be as defined in Theorem 3.3. Then under Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12),* V_{qd} *is the value function and* $\boldsymbol{v}^*_{qd}$ *is the optimal control for the following control problem,*

$$\inf_{\boldsymbol{v} \in \mathcal{A}_{\boldsymbol{u},0}} E^{t_0, y_0}_{\boldsymbol{v}} \left[(Y_{qd}(T))^2 \right]. \tag{3.28}$$

Proof. See Appendix A.1. ■

Solving the HJB PIDE (3.24)-(3.25), we obtain the QD-optimal control as reported by the following lemma. As in the case of the IR-optimal control (see Lemma 3.4,) the QD-optimal control $\boldsymbol{\varrho}^*_{qd}\left(t, \boldsymbol{X}^*_{qd}(t)\right)$ also only depends on the instantaneous benchmark allocation $\hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right)$ and not on its past or future.

Lemma 3.10. *(QD-optimal control) Suppose that Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12) are applicable. Then the optimal fraction of the investor's wealth invested in risky asset* $i \in \{1, ..., N^r_a\}$ *for problem* $QD(\beta)$ *in (3.4) is given by the* i^{th} *component of the vector* $\boldsymbol{\varrho}^*_{qd}\left(t, \boldsymbol{X}^*_{qd}(t)\right)$*, where*

$$W^*_{qd}(t) \cdot \boldsymbol{\varrho}^*_{qd}\left(t, \boldsymbol{X}^*_{qd}(t)\right) = \left[h_\beta(t) - \left(W^*_{qd}(t) - e^{\beta T} \hat{W}(t)\right)\right] \cdot (\boldsymbol{\Sigma} + \boldsymbol{\Lambda})^{-1} \tilde{\boldsymbol{\mu}} + e^{\beta T} \hat{W}(t) \cdot \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right). \tag{3.29}$$

Here, $W^*_{qd}(t)$ *denotes the investor's wealth process (3.11) under the QD-optimal control* $\boldsymbol{\varrho}^*_{qd}$ *with* $\boldsymbol{X}^*_{qd}(t) = \left(W^*_{qd}(t), \hat{W}(t), \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right)\right)$*, and* $h_\beta(t) = \frac{q}{r}\left(e^{\beta T} - 1\right)\left(1 - e^{-r(T-t)}\right), t \in [t_0, T]$.

Proof. See Appendix A.1. ■

The following lemma shows that once sufficient outperformance can be assured, the QD-optimal amounts in the risky assets will agree with the corresponding benchmark amounts multiplied by the constant scaling factor $e^{\beta T}$.

Lemma 3.11. *(QD: Matching the elevated benchmark risky asset amount) Given Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12), suppose that at some time* $\bar{t} \in (t_0, T]$*, the QD-optimal investor observes a wealth value* $W^*_{qd}(\bar{t})$ *satisfying*

$$W^*_{qd}(\bar{t}) = e^{\beta T} \hat{W}(\bar{t}) + h_\beta(\bar{t}). \tag{3.30}$$

Then for the remainder of the investment time horizon $t \in \left[\bar{t}, T\right]$*, the QD-optimal investor* *(using strategy (3.29)) will invest the following amounts in the risky assets,*

$$W^*_{qd}(t) \cdot \boldsymbol{\varrho}^*_{qd}\left(t, \boldsymbol{X}^*_{qd}(t)\right) = e^{\beta T} \cdot \hat{W}(t) \cdot \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right), \qquad \forall t \in \left[\bar{t}, T\right]. \tag{3.31}$$

Proof. See Appendix A.1. ■

By analogy with Lemma 3.6, the following lemma establishes some conditions under which the equivalence of problems (2.6) and (2.7) can be established analytically.

Lemma 3.12. *(QD: equivalence with only penalizing underperformance) If Assumption 3.1, Assumption 3.2, and wealth dynamics (3.11)-(3.12) apply with no jumps (i.e.* $\boldsymbol{\lambda} = \mathbf{0} \in \mathbb{R}^{N_a^r}$*) in the risky asset processes (3.8), then*

$$W^*_{qd}(t) < h_\beta(t) + e^{\beta T}\hat{W}(t), \qquad \forall t \in [t_0, T]. \tag{3.32}$$

As a result, in this case the QD optimization problem (2.6) is equivalent to the one-sided quadratic problem (2.7), where only the underperformance of the investor's portfolio (compared to the elevated benchmark) is penalized.

Proof. See Appendix A.1. ■

As in the case of the IR problem, Lemma 3.11 and Lemma 3.12 provide intuition for the behavior of the QD-optimal investment strategies even if the assumptions of this section are relaxed.

For the QD problem, it appears unlikely that the probability of underperforming the benchmark can be established analytically for an *arbitrary* adapted feedback benchmark strategy (i.e. of the form $\hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right)$ as per Assumption 3.2) as in the case of the IR problem (see Lemma 3.7). However, when a constant proportion benchmark $\hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) \equiv \hat{\boldsymbol{\varrho}}$ for all t is used, the following lemma shows that the QD-optimal probability of underperforming the benchmark can be obtained analytically.

Lemma 3.13. *(QD: probability of underperformance) Suppose the following assumptions hold: (i) Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12) with no jumps (i.e.* $\boldsymbol{\lambda} = \mathbf{0} \in \mathbb{R}^{N_a^r}$*) in the risky asset processes (3.8); (ii) contributions are zero (*$q = 0$*), and (iii) the benchmark strategy is a constant proportion strategy with* $\hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) \equiv \hat{\boldsymbol{\varrho}}$ *for* $t \in [t_0, T]$*. Then the probability of the QD-optimal wealth underperforming the benchmark wealth at any* $t \in (t_0, T]$ *is given by*

$$P^{t_0, w_0}_{\boldsymbol{\varrho}^*_{qd}}\left[W^*_{qd}(t) \leq \hat{W}(t)\right] = \Phi\left(\frac{\left[\frac{1}{2}\hat{\boldsymbol{\varrho}}^\top \boldsymbol{\Sigma} \hat{\boldsymbol{\varrho}} - \tilde{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\varrho}} - \frac{3}{2}\eta\right]\sqrt{t}}{\left[\hat{\boldsymbol{\varrho}}^\top \boldsymbol{\Sigma} \hat{\boldsymbol{\varrho}} + 2\tilde{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\varrho}} + \eta\right]^{1/2}}\right), \qquad t \in (t_0, T]. \tag{3.33}$$

Proof. See Appendix A.1. ■

We emphasize that, in contrast to the IR-optimal probability of underperformance (see (3.23)), the closed-form expression (3.33) is obtained under the assumptions of a constant proportion benchmark strategy and zero contributions. Under these assumptions, we observe that (3.33) does not depend on the targeted outperformance spread β. As in Remark 3.8, we conjecture that this can be explained due to the assumption of allowing trading to continue if insolvent.

3.4. Analytical comparison results. We now present analytical comparison results that are general in the sense of holding regardless of the values of the parameters γ and β chosen for the $IR(\gamma)$ and $QD(\beta)$ problems, respectively. Supplementary comparison results, based on particular choices of γ and β such that equal expectations of terminal wealth is obtained, are presented in Appendix A.2.

The following lemma compares the wealth allocation to the risky asset *basket*. Specifically, let $\varrho^*_{ir,i}(t)$ and $\varrho^*_{qd,i}(t)$ denote the ith components (i.e. the proportional allocations to the ith risky asset) of the optimal controls $\boldsymbol{\varrho}^*_{ir}(t, \boldsymbol{X}^*_{ir}(t))$ and $\boldsymbol{\varrho}^*_{qd}\left(t, \boldsymbol{X}^*_{qd}(t)\right)$, respectively, where we drop the dependence on $\boldsymbol{X}^*_{ir}(t)$ and $\boldsymbol{X}^*_{qd}(t)$ to lighten notation. Similarly, $\hat{\varrho}_i(t)$ denotes the benchmark allocation to the ith risky asset. The total proportional wealth allocation to the risky asset basket according to each strategy is therefore

$$\mathcal{R}^*_{ir}(t) = \sum_{i=1}^{N_a^r} \varrho^*_{ir,i}(t), \qquad \mathcal{R}^*_{qd}(t) = \sum_{i=1}^{N_a^r} \varrho^*_{qd,i}(t), \qquad \hat{\mathcal{R}}(t) = \sum_{i=1}^{N_a^r} \hat{\varrho}_i(t). \tag{3.34}$$

In the case of the simple continuous-time mean-variance control reported in [121], the optimal risky basket composition is independent of the state. As the following corollary shows, in the case of the IR and QD objectives, the optimal risky asset basket compositions do depend on the state of the system, but rather weakly, in the sense that certain ratios remain independent of the state.

Corollary 3.14. *(Constant risky asset basket ratios) Suppose that Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12) hold. Since $W^*_{ir}(t), W^*_{qd}(t), \hat{W}(t)$ and $\hat{\mathcal{R}}(t)$ represent information known to the investor at time t, the total optimal risky asset basket allocations $\mathcal{R}^*_{ir}(t)$ and $\mathcal{R}^*_{qd}(t)$ can be determined from the following constant ratios,*

$$\frac{W^*_{ir}(t)\cdot\mathcal{R}^*_{ir}(t) - \hat{W}(t)\cdot\hat{\mathcal{R}}(t)}{\left[\gamma e^{-r(T-t)} - \left(W^*_{ir}(t) - \hat{W}(t)\right)\right]} = \frac{W^*_{qd}(t)\cdot\mathcal{R}^*_{qd}(t) - \hat{W}(t)\cdot\hat{\mathcal{R}}(t)}{\left[h_\beta(t) - \left(W^*_{qd}(t) - \hat{W}(t)\right)\right]} = \sum_{k=1}^{N_a^r}\left[(\boldsymbol{\Sigma}+\boldsymbol{\Lambda})^{-1}\tilde{\boldsymbol{\mu}}\right]_k. \tag{3.35}$$

Within each risky asset basket, the optimal allocations to risky asset $i \in \{1, ..., N_a^r\}$, $\varrho^*_{ir,i}(t)$ and $\varrho^*_{qd,i}(t)$, satisfy the following constant ratios,

$$\frac{W^*(t)\cdot\varrho^*_{ir,i}(t) - \hat{W}(t)\cdot\hat{\varrho}_i(t)}{W^*(t)\cdot\mathcal{R}^*_{ir}(t) - \hat{W}(t)\cdot\hat{\mathcal{R}}(t)} = \frac{W^*(t)\cdot\varrho^*_{qd,i}(t) - e^{\beta T}\hat{W}(t)\cdot\hat{\varrho}_i(t)}{W^*(t)\cdot\mathcal{R}^*_{qd}(t) - e^{\beta T}\hat{W}(t)\cdot\hat{\mathcal{R}}(t)} = \frac{\sum_{j=1}^{N_a^r}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda})^{-1}_{ij}(\mu_j - r)}{\sum_{k=1}^{N_a^r}\left[(\boldsymbol{\Sigma}+\boldsymbol{\Lambda})^{-1}\tilde{\boldsymbol{\mu}}\right]_k}. \tag{3.36}$$

Proof. Note that $[\boldsymbol{v}]_k$ denotes the kth component of any vector $\boldsymbol{v}$. The results follow from definition (3.34), Lemma 3.4 and Lemma 3.10. ■

As a result of Corollary 3.14, in our numerical experiments (Section 6) we analyze the behavior of the analytical solutions using only a "single" risky asset assumed to be a diversified stock index, since this focuses on the key aspect of the asset allocation (3.35). In particular, Corollary 3.14 encourages the interpretation of the optimal controls as primarily determining the overall risky asset basket allocations $\mathcal{R}^*_{ir}$ and $\mathcal{R}^*_{qd}$, since once this is known, determining individual allocations using (3.36) is trivial.

Lemma 3.15 below presents a simple but interesting comparison result for the probability of benchmark underperformance associated with the IR- and QD-optimal investment strategies.

Lemma 3.15. *(QD vs IR: Probability of underperformance) Suppose that the assumptions of Lemma 3.13 hold. In addition, we assume that the benchmark strategy, which is assumed to be a constant proportion strategy* $\hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) \equiv \hat{\boldsymbol{\varrho}} = \left(\hat{\varrho}_1, ..., \hat{\varrho}_{N_a^r}\right)$ *as per Lemma 3.13, satisfies the following: (i)* $\hat{\varrho}_i \geq 0$ *for all* $i \in \{1, ..., N_a^r\}$*, and (ii)* $\hat{\varrho}_i > 0$ *for at least one* $i \in \{1, ..., N_a^r\}$*. Then the probability that the QD-optimal strategy underperforms the benchmark always exceeds the corresponding probability associated with the IR-optimal strategy, in other words*

$$P^{t_0,w_0}_{\boldsymbol{\varrho}^*_{qd}}\left[W^*_{qd}(t) \leq \hat{W}(t)\right] \geq P^{t_0,w_0}_{\boldsymbol{\varrho}^*_{ir}}\left[W^*_{ir}(t) \leq \hat{W}(t)\right], \qquad \forall t \in [t_0, T]. \tag{3.37}$$

Proof. See Appendix A.1. ■

In numerical tests, we observe that (3.37) appears to remain true provided Assumption 3.1 holds, even if we allow for contributions ($q > 0$) and jumps in the risky asset processes.

While it is an interesting result, it should be emphasized that Lemma 3.15 only considers a single point of a cumulative distribution function, namely $P^{t_0,w_0}_{\boldsymbol{\varrho}^*_{qd}}\left[W^*_{qd}(t)/\hat{W}(t) \leq 1\right]$. As the results of Section 6 show, this is a very unreliable basis for the practical evaluation and comparison of investment strategies, especially since no mention is made of tail behavior (upside or downside) of the different strategies.

We conclude this section with some final remarks on the closed-form results. We observed in Section 2 that the objective functions suggest that the QD investor wishes (where possible) to outperform the benchmark terminal wealth by a constant *factor*, whereas the IR investor hopes to achieve the benchmark terminal wealth by a constant *amount* irrespective of the underlying market scenario. The results of Lemmas 3.4, 3.6, 3.10 and 3.12 confirm that this intuition not only holds at time T, but also for all $t < T$.

Specifically, at time $t < T$, the IR-optimal strategy can be interpreted as having an implicit wealth target of $\left[\gamma e^{-r(T-t)} + \hat{W}(t)\right]$; see (3.18), (3.20) and (3.22). Similarly, ignoring contributions, the QD-optimal strategy can be interpreted as having an implicit target of $e^{\beta T}\hat{W}(t)$ for $W^*_{qd}(t)$; see (3.29), (3.30) and (3.32). By "implicit target", we mean that in the case of both the IR and QD strategies, the risky asset basket exposure is increased in direct proportion with the extent to which the investor's wealth is underperforming the above-mentioned target values at time t. As a result, in adverse market scenarios (which of course also affects the benchmark), the IR strategy effectively aims to outperform the benchmark by a larger *factor* than in "typical" market scenarios due to the constant *amount* of specified outperformance, and is thus required to take on more extreme positions in the riskiest asset compared to the QD strategy. Similarly, early in the investment time horizon, when the investor's wealth is expected to be small relative to wealth at later stages, the IR strategy is therefore expected to take on significantly more risk (i.e. investing more in the riskiest asset) than the QD strategy due to its higher relative target implied by the constant amount of outperformance.

These statements can be made rigorous in the case of two assets under Assumption 3.1 (see Appendix A, in particular Theorem A.3), but the numerical results in Section 6 show that these observations remain true in more general cases where constraints are applicable.

4. Traditional dynamic programming: an unnecessarily high dimensional approximation problem.

If problems (2.4) or (2.6) cannot be solved analytically, for example when multiple investment constraints are applicable or the portfolio is rebalanced at discrete time intervals, then the standard numerical solution approach is to rely on dynamic programming (DP). For example, we could use the Q-learning algorithm, which is arguably the most popular data-driven Reinforcement Learning (RL) algorithm (see for example [34, 94, 84, 48]) that fundamentally relies on the DP principle to solve (2.4) or (2.6).

Many of the well-known concerns with using DP-based techniques, including in multi-asset portfolio optimization settings (see [111, 82]), follow from the fact that an approximation to a conditional expectation is required at each solution step. This is the essence of value iteration employed in RL and the Q-learning algorithm, which implies that an optimization problem has to be solved to determine the value function using the performance criterion ([92]) at each portfolio rebalancing event, recursively backwards from the terminal time T. This can cause significant challenges with regards to the stability and convergence associated with the estimated value function and estimated optimal control due to the amplification of the estimation errors over each iteration (see for example [111, 82, 117]).

While these challenges with DP do enjoy some recognition in the literature, in this section we present an *additional* motivation for avoiding the use of the DP principle to solve problems specifically of the form (2.4) or (2.6).

4.1. Formulation requiring a numerical solution approach.

We start by formulating a more realistic setting for the investment problems (2.4) and (2.6), which would necessitate the use of numerical solution techniques.

We assume that the investor only rebalances the portfolio at each of N_{rb} rebalancing times in the investment time horizon $[t_0 = 0, T]$, so that the set $\mathcal{T}$ of rebalancing times is given by

$$\mathcal{T} = \{ t_n = n\Delta t |\, n = 0, ..., N_{rb} - 1 \}, \qquad \Delta t = T/N_{rb}. \tag{4.1}$$

For convenience, we assume that the rebalancing times are equally spaced in (4.1), and that contributions to the portfolio are a priori specified and made only at rebalancing times. We therefore assume a given cash contribution schedule $\{q(t_n) : n = 0, ..., N_{rb} - 1\}$, where $q(t_n)$ denotes the amount of cash contributed to each portfolio (investor and benchmark portfolios) at $t_n \in \mathcal{T}$.

Note that the basic aspects of the formulation remains as in Section 2, including the use of N_a assets. In particular, the investor strategy and benchmark strategies are of the form (2.1) using $\mathcal{T}$ given by (4.1).

We do not make assumptions about underlying dynamics, but instead simply observe that if $R_i(t_n)$ denotes the return on asset $i \in \{1, .., N_a\}$ over the time interval $[t_n, t_{n+1}]$, then the investor and benchmark wealth dynamics are given by

$$W\left(t_{n+1}^-\right) = \left[W\left(t_n^-\right) + q(t_n)\right] \cdot \sum_{i=1}^{N_a} p_i\left(t_n, \boldsymbol{X}(t_n)\right) \cdot \left[1 + R_i(t_n)\right], \tag{4.2}$$

$$\hat{W}\left(t_{n+1}^-\right) = \left[\hat{W}\left(t_n^-\right) + q(t_n)\right] \cdot \sum_{i=1}^{N_a} \hat{p}_i\left(t_n, \hat{\boldsymbol{X}}(t_n)\right) \cdot \left[1 + R_i(t_n)\right], \tag{4.3}$$

where $n = 0, ..., N_{rb} - 1$ and $W\left(t_0^-\right) = \hat{W}\left(t_0^-\right) := w_0 > 0$. The minimal form of $\boldsymbol{X}$ is assumed to be $\boldsymbol{X}(t_n) = \left(W(t_n), \hat{W}(t_n)\right)$, which is suggested by the results presented in Subsection 4.2 below.

Finally, we assume that the investor is subject to the investment constraints of (i) no shorting and (ii) no leverage. In particular, this means that we consider the sets of admissibility (see Section 2) for the investor strategy given by

$$\mathcal{A} = \left\{\mathcal{P} = \left\{\boldsymbol{p}(t_n, \boldsymbol{X}(t_n)) : t_n \in \mathcal{T}\right\} \middle| \, \boldsymbol{p}(t_n, \boldsymbol{X}(t_n)) \in \mathcal{Z}, \quad \forall t_n \in \mathcal{T}\right\}, \tag{4.4}$$

$$\text{where} \qquad \mathcal{Z} = \left\{(y_1, ..., y_{N_a}) \in \mathbb{R}^{N_a} : \sum_{i=1}^{N_a} y_i = 1 \text{ and } y_i \geq 0 \text{ for all } i = 1, ..., N_a\right\}, \tag{4.5}$$

which also ensures that the investor's wealth (with dynamics (4.2)) remains non-negative.

We are therefore concerned with solving the IR and QD problems where $T = N_{rb} \cdot \Delta t$, the set of rebalancing times $\mathcal{T}$ is given by (4.1), wealth dynamics are given by (4.2) and (4.3), and the investor strategy $\mathcal{P}$ takes values in the admissible set $\mathcal{A}$ in (4.4).

4.2. High-dimensional performance criterion, low-dimensional control. We now present an additional challenge with DP-based solution techniques. Specifically, in the Proposition 4.1 we show that the DP approach is, in a sense, *unnecessarily* high-dimensional in the case of benchmark outperformance problems of the form (2.4) and (2.6). For concreteness and illustrative purposes, note that Proposition 4.1 incorporates some assumptions which are not required subsequently, since different DP approaches will treat the solution of the performance criterion (a conditional expectation) between rebalancing events in different ways. However, qualitatively similar observations regarding dimensionality will remain applicable.

Proposition 4.1. *(Discrete rebalancing: Dimensions of the dynamic programming solutions* *to the IR and QD problems) Suppose the $IR(\gamma)$ and $QD(\beta)$ problems in (2.4) and (2.6) are solved using dynamic programming in the case where the portfolio is only rebalanced at the set of discrete rebalancing times $\mathcal{T}$ in (4.1). For concreteness and illustrative purposes, we make the following additional simplifying assumptions: (i) The N_a underlying assets, representing the set of investable assets for both the investor and the benchmark, are risky assets with dynamics given by (3.8). (ii) The benchmark's asset allocation strategy is an adapted feedback control of the form $\hat{\boldsymbol{p}}\left(t_n, \hat{\boldsymbol{X}}(t_n)\right) = \hat{\boldsymbol{p}}\left(t_n, \hat{W}(t_n)\right), t_n \in \mathcal{T}$. (iii) At each rebalancing event, the investor can observe the benchmark asset allocation vector $\hat{\boldsymbol{p}}\left(t_n, \hat{W}(t_n)\right)$.*

Then at each fixed rebalancing time $t_n \in \mathcal{T}$, regardless of the number of underlying assets N_a, the optimal controls of problems $IR(\gamma)$ and $QD(\beta)$ in (2.4) and (2.6) are functions only of the investor's wealth and the benchmark wealth. In other words, at each rebalancing time t_n, the optimal investor control for each problem consists of the vectors $\boldsymbol{p}^*_{ir}(t_n, \boldsymbol{X}^*_{ir}(t_n))$ and $\boldsymbol{p}^*_{qd}(t_n, \boldsymbol{X}^*_{qd}(t_n)), t_n \in \mathcal{T}$, respectively, where $\boldsymbol{X}^*_{ir}(t_n) = \left(W^*_{ir}(t_n), \hat{W}(t_n)\right)$ and $\boldsymbol{X}^*_{qd}(t_n) = \left(W^*_{qd}(t_n), \hat{W}(t_n)\right)$.

However, in using dynamic programming to obtain the optimal controls $\boldsymbol{p}^*_k : \mathbb{R}^{(2+1)} \rightarrow \mathbb{R}^{N_a}, k \in \{ir, qd\}$, which are only two-dimensional controls at each fixed rebalancing time $t_n \in \mathcal{T}$, the investor requires the solution of a $(2N_a + 1)$-dimensional performance criterion $J : \mathbb{R}^{(2N_a+1)} \rightarrow \mathbb{R}$, for each problem, between each pair of adjacent rebalancing times $t_n, t_{n+1} \in \mathcal{T}$.

Proof. See Appendix A.3. ■

Therefore, given the stated assumptions, Proposition 4.1 shows that the case of discrete rebalancing[3], the investor needs to solve for a $(2N_a + 1)$-dimensional performance criterion during each value iteration (rebalancing time step), which can be expressed as a 2-dimensional function (corresponding to the value function if the optimal control is used) only at each rebalancing time $t_n \in \mathcal{T}$.

Proposition 4.1 demonstrates that it is inefficient to solve (2.4) and (2.6) by DP, in addition to the aforementioned challenges resulting from error amplification. We advocate solving the original stochastic optimal control problems, e.g., (2.4) and (2.6), directly without DP. In particular, we represent control by an NN, which explicitly exploits its lower dimensionality. As a result, significant computational advantages follow, since the optimal control is computed *without* the need to solve for the corresponding performance criterion.

5. Neural network (NN) solution approach.

We now discuss the numerical solution of problems (2.4) and (2.6) using a data-driven neural network (NN) approach that does not rely on the DP principle, but instead solves directly for the optimal control. This approach therefore avoids both the dimensionality and error amplification issues outlined in the previous section. While our approach is broadly inspired by some of our previous work (see [81, 88, 113]), it is specialized in this section to problems of the form (2.4) and (2.6). A brief summary of the

[3] In contrast, in the case of continuous rebalancing, the results of Section 3 show that that the investor only requires the solution of a 2-dimensional value function at every given $t \in [t_0, T]$.

approach is provided, with more information available in Appendix C.

Our basic task in solving problems (2.4) and (2.6) is to determine the control $\mathcal{P}$ (see (2.1)) in feedback form $\boldsymbol{p}\left(t, \boldsymbol{X}\left(t\right)\right)$. We assume that $\boldsymbol{p}(t, \boldsymbol{X}) \in \mathcal{Z}$ is a continuous function of $(t, \boldsymbol{X})$, which enforces the condition that, in the limit as $\Delta t \to 0$, the approximate control remains a *continuous* function of time. We believe that this is a necessary practical constraint to any investment policy, since investors would surely be reluctant to follow a strategy where the asset allocations exhibited non-smooth behavior as a function of time if the observed information $\boldsymbol{X}(t)$ is a smooth function of time. Since the portfolio is rebalanced only at discrete time intervals, the investment strategy can be found by evaluating this continuous function at discrete time intervals, i.e. $\left(t_n, \boldsymbol{X}\left(t_n\right)\right) \to \mathcal{P}\left(t_n, \boldsymbol{X}\left(t_n\right)\right) = \boldsymbol{p}\left(t_n, \boldsymbol{X}\left(t_n\right)\right), t_n \in \mathcal{T}$.

Appealing to the Universal Approximation Theorem (see [32, 47, 58, 59, 77, 108]), we approximate the continuous control function $\boldsymbol{p}(t, \boldsymbol{X})$ by a NN $\boldsymbol{F}(t, \boldsymbol{X}(t); \boldsymbol{\theta}) \equiv \boldsymbol{F}(\cdot, \boldsymbol{\theta})$, where $\boldsymbol{\theta} \in \mathbb{R}^{\eta_\theta}$ is the set of NN parameters (i.e. the NN weights and biases), so that

$$\boldsymbol{p}(t, \boldsymbol{X}(t)) \simeq \boldsymbol{F}(t, \boldsymbol{X}(t); \boldsymbol{\theta}) \equiv \boldsymbol{F}(\cdot, \boldsymbol{\theta}). \tag{5.1}$$

While we use a standard fully-connected feed-forward NN (see for example [51]), it has the following specific structural properties: (i) The minimal inputs (features) consists of time t, investor wealth $W\left(t\right)$ and benchmark wealth $\hat{W}\left(t\right)$ after incorporating contributions. (ii) The number of output nodes correspond to the number of assets. (iii) A softmax activation is used in the output layer to ensure the NN generates outputs in $\mathcal{Z} \subset \mathbb{R}^{N_a}$ as per (4.5). We place no fixed requirements on the number of hidden layers or activation functions, since these are typically tailored to a given portfolio optimization problem based on numerical experiments (see Appendix C). In the subsequent results, we use two hidden layers, each with $N_a + 2$ hidden nodes, and logistic sigmoid activations. The general NN structure is illustrated in Figure 5.1.

Since the NN $\boldsymbol{F}(\cdot, \boldsymbol{\theta})$ generates values in $\mathcal{Z}$, problems (2.4) and (2.6) are then approximated by the *unconstrained* optimization problems

(5.2)

$$\inf_{\boldsymbol{\theta} \in \mathbb{R}^{\eta_\theta}} E^{t_0, w_0}_{\boldsymbol{F}(\cdot; \boldsymbol{\theta})}\left[\left(W\left(T; \boldsymbol{\theta}\right) - \left[\hat{W}\left(T\right) + \gamma\right]\right)^2\right], \quad \text{and} \quad \inf_{\boldsymbol{\theta} \in \mathbb{R}^{\eta_\theta}} E^{t_0, w_0}_{\boldsymbol{F}(\cdot; \boldsymbol{\theta})}\left[\left(W\left(T; \boldsymbol{\theta}\right) - e^{\beta T}\hat{W}\left(T\right)\right)^2\right].$$

From a computational point of view, the expectations $E^{t_0, w_0}(\cdot)$ in (5.2) are approximated using a finite set of samples Y, which in the usual terminology (see [51]) serves as the training data set of the NN. Y is assumed to be of the form $Y = \left\{Y^{(j)} : j = 1, ..., N_d\right\}$, where each $Y^{(j)}$ represents a path of *joint* asset return observations $R_i, i \in \{1, .., N_a\}$ observed at each $t_n \in \mathcal{T}$.

Our solution approach is agnostic as to the particular technique used to generate the training data set Y. If we restrict attention to parametric stochastic models, then Y can be generated trivially from Monte Carlo simulation. However, it is more straightforward, and perhaps more convincing for practitioners, to use historical data directly, which (due to sparsity of data) necessarily requires some data augmentation or generation techniques. For

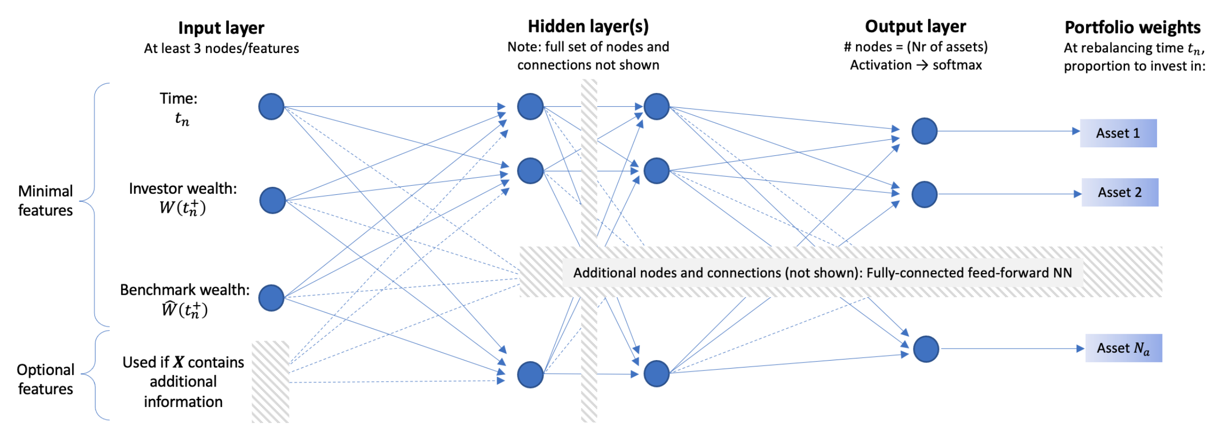


Figure 5.1: Illustration of the structure of the NN $\boldsymbol{F}$ used to model the control (investment strategy). The same NN is applied at all rebalancing times, with the asset allocations at a specific rebalancing time t_n obtained using the minimal features including time (t_n), investor wealth $W(t_n^+) = W(t_n^-) + q(t_n)$ and benchmark wealth $\hat{W}(t_n^+) = \hat{W}(t_n^-) + q(t_n)$.

illustrative purposes, we use stationary block bootstrap resampling ([98]) in the results of Section 6, which is popular with practitioners ([26, 33, 105, 22, 107, 5]) and designed for weakly stationary series having serial dependence. Note that [100] and [99] suggest methods for resampling non-stationary time series, which we do not explore in this paper.

Consider a given training dataset Y, regardless how it is obtained. For a given $\boldsymbol{\theta} \in \mathbb{R}^{\eta_\theta}$ in (5.2) and a given training sample path $Y^{(j)} \in Y$, we can obtain the corresponding wealth outcomes $W^{(j)}(T)$ and $\hat{W}^{(j)}(T)$ calculated using (4.2)-(4.3) and (5.1). Our final computational problems for (5.2) can therefore be expressed as

$$\min_{\boldsymbol{\theta}\in\mathbb{R}^{\eta_\theta}} \left\{ \frac{1}{N_d} \sum_{j=1}^{N_d} \left(W^{(j)}(T;\boldsymbol{\theta}) - \left[\hat{W}^{(j)}(T) + \gamma \right] \right)^2 \right\}, \quad \text{and} \quad \min_{\boldsymbol{\theta}\in\mathbb{R}^{\eta_\theta}} \left\{ \frac{1}{N_d} \sum_{j=1}^{N_d} \left(W^{(j)}(T;\boldsymbol{\theta}) - e^{\beta T} \hat{W}^{(j)}(T) \right)^2 \right\}. \tag{5.3}$$

The optimal NN parameter vectors for (5.3), denoted by $\boldsymbol{\theta}_k^*, k \in \{ir, qd\}$ respectively, can then be obtained using standard (unconstrained) optimization methods - see Appendix C. The resulting optimal investment strategies $\boldsymbol{p}_k^*(\cdot, \boldsymbol{X}(\cdot)) \simeq \boldsymbol{F}(\cdot, \boldsymbol{\theta}_k^*), k \in \{ir, qd\}$ can be implemented on a testing data set Y^{test} to assess the out-of-sample performance of the resulting strategies. While the contents of Y^{test} is expected to differ from that of the training dataset Y, for example it might be based on different data generation assumptions, it is assumed to have a similar structure to the training dataset.

We highlight the following important properties of this NN solution approach:

(i) We approximate the control directly using a NN, and do not rely on DP techniques. In particular, the problems of the approximation of (high-dimensional) conditional expectations and value iteration discussed in Section 4 are avoided entirely. Note that the idea of solving for the control directly, without using DP, has also been suggested in [102, 56].

(ii) Time is an input into the NN, which simultaneously implies that smooth behavior of the control as rebalancing time interval $\Delta t \rightarrow 0$ is automatically guaranteed, while also ensuring that the size of the NN parameter vector does not depend on the number of portfolio rebalancing events. These advantages contrasts our approach from that of for example [56, 111, 62].

For further details, including ground truth results, the reader is referred to Appendix C

6. Illustration of investment results.

In this section, we illustrate the results from investing according to the IR and QD optimal strategies, using both analytical solutions (Section 3), as well as numerical solutions using the NN approach (Section 5).

For illustrative purposes, we formulate a typical investment scenario where the investor wishes to outperform reasonable and popular benchmarks over the long term using both "standard assets" (a broad stock market index, Treasury bills and bonds) as well as two popular investment "factors" from the factor investing literature (see for example [6]). The investor is not necessarily limited to investing in the same assets that are used by the benchmark.

6.1. Investment scenario.

Table 6.1 summarizes the general investment scenario assumptions for the illustrative results. The time horizon of $T = 10$ years is chosen for an investor primarily concerned with long-run benchmark outperformance. The case of continuous rebalancing is approximated using 3,600 time steps in $[0, T]$, while the discrete rebalancing scenario assumes the annual or quarterly rebalancing of the portfolio.

Table 6.1: Key investment scenario assumptions

Parameter	Analytical solutions (no constraints)	Numerical solutions (with constraints)	
Investment constraints	None	No short-selling, no leverage allowed	
T	10 years	10 years	
w_0	120	120	
Rebalancing frequency	Continuous	Annual rebalancing	Quarterly rebalancing
N_{rb} (# rebalancing events)	3600	10	40
Contributions	$q = 12$ (rate per year)	$q(t_n) = 12, \forall n$ (annual contribution)	$q(t_n) = 3, \forall n$ (quarterly contribution)

Since there are many possibilities for the basis of comparison of the IR- and QD-optimal investment strategies, we assume that the investor aims to achieve an expected terminal wealth of $\mathcal{E}$ regardless of whether the IR or QD strategy is followed. Specifically, if the benchmark investment strategy results in $E_{\hat{\mathcal{P}}}^{t_0,w_0}\left[\hat{W}(T)\right] = \mathcal{K}$, we assume the investor chooses some value of $\hat{\beta} > 0$ in (6.1) to achieve an expected terminal wealth of $\mathcal{E}$:

$$E^{t_0,w_0}_{\mathcal{P}^{\mathcal{E}*}_{ir}}\left[W^{\mathcal{E}*}_{ir}(T)\right] \equiv E^{t_0,w_0}_{\mathcal{P}^{\mathcal{E}*}_{qd}}\left[W^{\mathcal{E}*}_{qd}(T)\right] := \mathcal{E} := e^{\hat{\beta}T}\cdot\mathcal{K} = e^{\hat{\beta}T}\cdot E^{t_0,w_0}_{\hat{\mathcal{P}}}\left[\hat{W}(T)\right]. \tag{6.1}$$

The desired target expectation (6.1) can be achieved by solving numerically (or in some cases, analytically - see Appendix A) for values of $\gamma = \gamma^{\mathcal{E}}_{ir}$ in the $IR(\gamma)$ problem and $\beta = \beta^{\mathcal{E}}_{qd}$ in the $QD(\beta)$ problem. Note that $\hat{\beta} > 0$ in (6.1) implies that we always have the strict inequality $\mathcal{E} > \mathcal{K}$, which is required since if $\mathcal{E} = \mathcal{K}$, then the IR- and QD-optimal strategies will be identical to the benchmark strategy[4].

Table 6.2 summarizes the underlying assets considered. Candidate assets for the investor portfolio are identified by the label "Px", $x \in \{0,1\}$, while benchmarks are identified by the label "BMx", $x \in \{0,1\}$. Both benchmarks portfolios are equally-weighted between stocks and bonds. We assume that the investor will construct portfolio P0 ($N_a = 2$) to outperform benchmark BM0 (also 2 assets), and portfolio P1 ($N_a = 5$) to outperform benchmark BM1 (3 assets with nonzero investment).

More information regarding the definition and historical returns data for the assets in Table 6.2 can be found in Appendix B.1. All data was obtained for the period from 1963:07 to 2020:12, which includes the period of significant market volatility experienced during 2020. Due to the reasonably long investment time horizon (Table 6.1), we assume as in for example [45, 44] that the investor is primarily interested in the real (or inflation-adjusted) performance of the portfolio. Therefore, prior to calculations or NN training/testing data set constructions, all time series of returns were inflation-adjusted using data from the US Bureau of Labor Statistics.

6.2. Illustration of analytical solutions. For the illustration of the analytical results of Section 3, we assume that investor portfolio P0 is constructed to outperform benchmark BM0 as per Table 6.2, while it is sufficient to consider only $N_a = 2$ assets (see Corollary 3.14). In the terminology of Section 3, T10 and Market (Table 6.2) are associated with the risk-free and risky assets, respectively. For the risky asset, we assume the [73] model, with more information on the parameters and calibration provided in Appendix B.1.

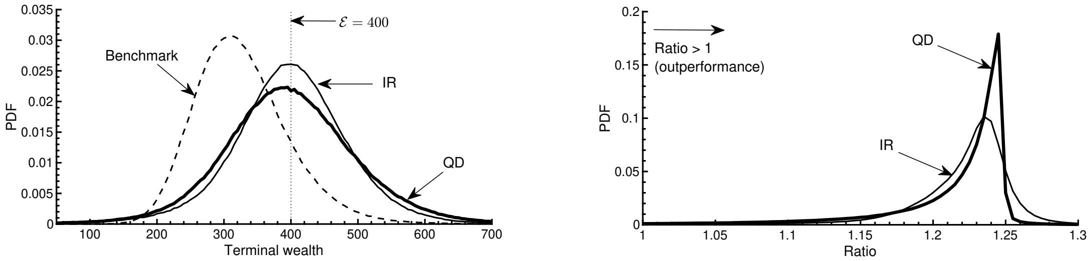
We now compare analytical investment results on the basis of (6.1), using 10^6 Monte Carlo simulations of asset dynamics (3.5) and (3.8) with parameters as in Table B.1. Figure 6.1 illustrates the simulated probability density functions (PDFs) associated with $\mathcal{E} = 400$ ($\hat{\beta} \simeq 2\%$), with results shown for both the terminal wealth (absolute performance) and the wealth ratio (relative performance). In Figure 6.1, the probability of benchmark underperformance is larger for the QD strategy (3.36%) than for the IR strategy (2.61%), which is expected as per

[4]While intuitive, the fact that $\mathcal{E} = \mathcal{K}$ implies $\mathcal{P}^{\mathcal{E}*}_{ir} = \mathcal{P}^{\mathcal{E}*}_{qd} = \hat{\mathcal{P}}$ can also be shown analytically by setting $\mathcal{E} = \mathcal{K}$ in the expressions for $\gamma^{\mathcal{E}}_{ir}$ and $\beta^{\mathcal{E}}_{qd}$ in Lemma A.2 in Appendix A, and then substituting the resulting values into the optimal controls (3.18) and (3.10).

Table 6.2: Portfolios of candidate assets considered by the investor "Px", $x \in \{0, 1\}$ and benchmarks "BMx", $x \in \{0, 1\}$. The tick mark "✓" indicates the inclusion of the asset in the portfolio optimization problem. The benchmark asset allocation is shown as a percentage of wealth.

Assets		Investor portfolios		Benchmarks	
Label	Asset description	P0	P1	BM0	BM1
T30	30-day Treasury bill	✓	✓	50%	25%
B10	10-year Treasury bond		✓		25%
Market	Market portfolio (broad equity market index)	✓	✓	50%	50%
Size	Portfolio of small stocks		✓		
Value	Portfolio of value stocks		✓		
	Number of candidate assets (N_a):	2	5	2	3

Lemma 3.15.

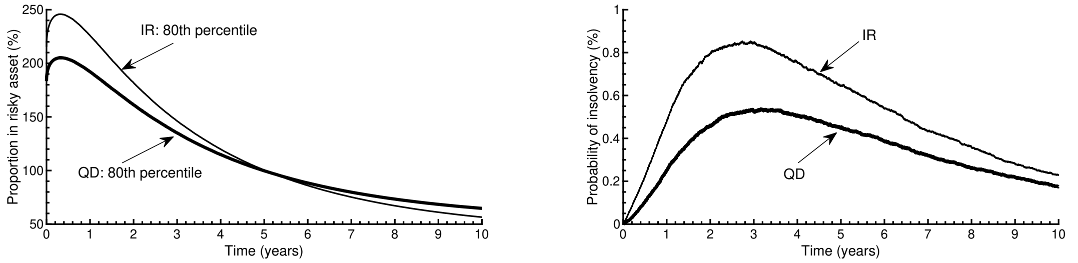


(a) PDFs of $\hat{W}(T)$, $W_j^{\mathcal{E}*}(T)$, $j \in \{ir, qd\}$ **(b)** PDFs of $W_j^{\mathcal{E}*}(T) / \hat{W}(T)$, $j \in \{ir, qd\}$

Figure 6.1: Analytical solutions, no constraints, investor portfolio P0, benchmark BM0: Simulated PDFs of benchmark and investor's target terminal wealth $\hat{W}(T)$ and $W_j^{\mathcal{E}*}(T)$, respectively, as well as the ratio $W_j^{\mathcal{E}*}(T) / \hat{W}(T)$, for $j \in \{ir, qd\}$. 10^6 Monte Carlo simulations, $\mathcal{E} = 400$ in (6.1). The corresponding CDFs are shown in Figure B.1 in Appendix B.

To illustrate the underlying analytical investment strategies, Figure 6.2(a) shows the relatively larger reliance placed by the IR strategy on the risky asset early in the investment time horizon, which has the effect (Figure 6.2(b)) that the IR strategy relies more heavily on trading in bankruptcy (allowed in this case as per Assumption 3.1) to achieve the desired benchmark outperformance. For both strategies, Figure 6.2(a) also illustrates that as time passes, the risky asset holdings of both the IR- and QD-optimal investment strategies trend closer to the benchmark holdings, which is (qualitatively) to be expected given the results of Lemma 3.5 and Lemma 3.11.

Note that the quantitative aspects of the relative behavior of the optimal investment strate-



(a) 80th percentiles of $p_j^{\mathcal{E}*}(t), j \in \{ir, qd\}$ **(b)** Probability of insolvency (expressed as %)

Figure 6.2: Analytical solutions, no constraints, investor portfolio P0, benchmark BM0: 80th percentiles of the investment in the single risky asset $p_j^{\mathcal{E}*}(t)$ and probability of insolvency as a function of time $t \rightarrow P_{p_j^{\mathcal{E}*}}^{t_0,w_0}\left[W_j^{\mathcal{E}*}(t) \leq 0\right]$, for $j \in \{ir, qd\}$. 10^6 Monte Carlo simulations, $\mathcal{E} = 400$ in (6.1).

gies observed in Figure 6.2(a) is analyzed in Appendix A (see Theorem A.3).

6.3. Illustration of numerical solutions. We now consider the scenario of multiple investment constraints and discrete rebalancing (see Subsection 4.1), so that the problems are solved using the NN approach outlined in Section 5 and Appendix C. Investment outcomes are still compared on the basis of (6.1), where the targeted expected value $\mathcal{E} = e^{\hat{\beta}T}\mathcal{K}$ (see (6.1)) is to be achieved on the neural network's *training* data set Y.

To construct both the training and testing data sets for the neural network, Y and Y^{test} respectively, we use stationary block bootstrap resampling for illustrative purposes (see discussion in Section 5). However, we emphasize that the NN approach is agnostic as to the particular technique used to obtain the data.

Table 6.3 outlines the key assumptions underlying Y and Y^{test}, with $\mathcal{K}$ reporting the mean benchmark terminal wealth on each training data set. For data sets DS1 and DS2, the relatively shorter expected blocksizes used for the testing data is due to the relatively shorter historical time period (11 years) of source data used for out-of-sample testing. Note that all subsequent results were also tested using various different assumptions for expected blocksizes, and since qualitatively similar results were obtained (as expected based on the robustness assessments presented in [88, 81]), only results for the data sets as outlined in Table 6.3 are presented.

Remark 6.1. (Rationale for training data period selections) While only for illustrative purposes, the data sets in Table 6.3 are constructed with specific goals. Data set DS0, obtained using simulation of specified asset dynamics, is included to illustrate the impact of discrete rebalancing and investment constraints on the results of Subsection 6.2. DS1 and DS2 incorporate data since 1963 due to data availability constraints for investable factors. In an ideal scenario, including data as far back as for example 1926 would be preferable, since it would include a wider range of economic and geopolitical events, such as the Great Depression and the second World War. A possible objection to using so much historical data (even if we limit our attention to data since 1963) might be that the historical data might not be relevant to current market conditions, and thus more recent data would be preferable. However, the last ~30 years exhibited a historical anomaly in that real interest rates have been declining almost monotonically, thus making investments in long-maturity low-risk government bonds particularly attractive, whereas it is exceedingly unlikely that this market regime would continue (see for example [42]). The training data of data sets DS1 and DS2 are specifically chosen to include periods of high inflation such as 1963-1985, including the 1970s where economic growth was stagnant in conjunction with high inflation, since this data might in fact be *more* relevant to current market conditions than more recent data. Regardless of these observations, we also include data set DS3, which incorporates training data only dating to 1995, since this might reflect the perspective of an investor considering the benchmark outperformance problems in 2010 (the start of the testing data set for DS3), and who wishes to use only the "most recent 15 years" (1995:01 - 2009:12) of training data for investable factors after Size and Value investments have been popularized with the publication of [39, 40]. DS3 involves more frequent rebalancing.

Table 6.3: Data set combinations, labelled DSx, $x \in \{0,1,2\}$, used for training and testing the neural network. "SBBR" refers to stationary block bootstrap resampling, with expected blocksize reported in brackets.

Label	Rebal. freq.	Training data set Y ($N_d = 10^6$)			Testing data set Y^{test} ($N_d^{test} = 5 \times 10^5$)	
		Source data	Data set generation	Benchmark exp. val.	Source data	Data set generation
DS0	Continuous	10^6 Monte Carlo simulations of asset dynamics (3.5), (3.8), (B.1) with parameters as in Table B.1		BM0: $\mathcal{K} = 334$	N/a	N/a
DS1	Annually	Historical data, 1963:07 - 2009:12	SBBR (6 months)	BM1: $\mathcal{K} = 338$	Historical data, 2010:01 - 2020:12	SBBR (3 months)
DS2	Annually	Historical data, 1963:07 - 1999:12	SBBR (6 months)	BM1: $\mathcal{K} = 364$	Historical data, 2000:01 - 2010:12	SBBR (3 months)
DS3	Quarterly	Historical data, 1995:01 - 2009:12	SBBR (3 months)	BM1: $\mathcal{K} = 352$	Historical data, 2010:01 - 2020:12	SBBR (3 months)

Table 6.4 provides the combinations of investor portfolios and benchmarks, as well as the targeted level of outperformance chosen for illustrative purposes. In the case of using portfolio P1 (5 assets) to outperform BM1 (3 assets), we use a slightly more ambitious value of $\hat{\beta} \simeq 1.7\%$ in (6.1), since the investor has more opportunities for outperformance given that factors are available for investment (see [113]). Note that the $\mathcal{E}$ values reported are different due to different values of $\mathcal{K}$ (see Table 6.3).

Table 6.4: Numerical solutions, with constraints: Target expected values for combinations of the investor portfolios, benchmarks and data set combinations. As per Table 6.3, both the investor portfolio and benchmark use continuous rebalancing in the case of DS0, annual rebalancing in the case of DS1 and DS2, and quarterly rebalancing in the case of DS3.

Investor portfolio	To outperform benchmark:	
	BM0 (2 assets)	BM1 (3 assets)
P0 (2 assets)	DS0: $\mathcal{E} = 370$ ($\hat{\beta} \simeq 1.0\%$)	N/a
P1 (5 assets)	N/a	DS1: $\mathcal{E} = 400$ ($\hat{\beta} \simeq 1.7\%$) DS2: $\mathcal{E} = 430$ ($\hat{\beta} \simeq 1.7\%$) DS3: $\mathcal{E} = 420$ ($\hat{\beta} \simeq 1.7\%$)

Table 6.5, based on using portfolio P0 to outperform benchmark BM0 on training data set DS0, shows the impact of applying investment constraints and discrete rebalancing to the results of Subsection 6.2: (i) with constraints, the QD-optimal probability of underperformance is now *lower* than the corresponding IR-optimal value, and thus the results of Lemma 3.15 no longer qualitatively hold; (ii) the QD-optimal strategy results in better downside performance than the IR strategy for both the wealth and the wealth ratio when constraints are applied. We note that while these results are obtained on the training data set of DS0, qualitatively similar training data ("in-sample") results hold for other data sets when investment constraints are applied - see for example the results for DS2 in Table B.2 (Appendix B). As a result, we will focus on the testing ("out-of-sample") outcomes in the subsequent results.

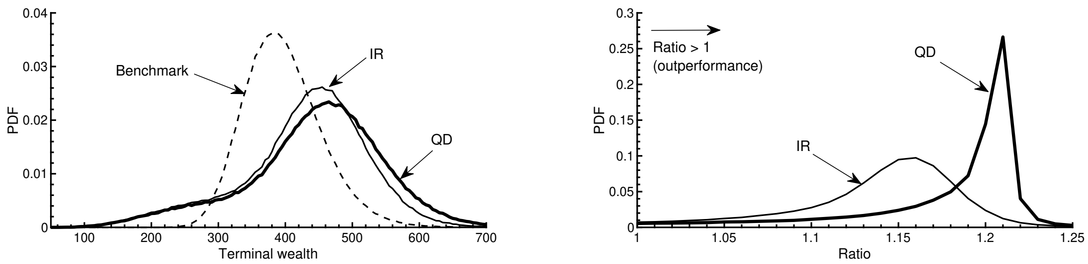
Table 6.5: Effect of constraints: analytical solutions vs. numerical solutions, investor portfolio P0, benchmark BM0. "No constraints" and "With constraints" columns are based on the assumptions for the analytical solutions and numerical solutions, respectively, as per Table 6.1. NN trained on data set DS0. Since no out-of-sample testing is conducted for DS0 (see Table 6.3), the "With constraints" results are obtained on the training data set.

Quantity	No constraints: P0, $\mathcal{E} = 370$					With constraints: P0, $\mathcal{E} = 370$				
	BM0	$W_j^{\mathcal{E}*}(T)$		$W_j^{\mathcal{E}*}(T)/\hat{W}(T)$		BM0	$W_j^{\mathcal{E}*}(T)$		$W_j^{\mathcal{E}*}(T)/\hat{W}(T)$	
	$\hat{W}(T)$	IR	QD	IR	QD	$\hat{W}(T)$	IR	QD	IR	QD
Mean	330	370	370	1.12	1.11	334	370	370	1.10	1.10
CExp 5%	208	193	191	0.90	0.90	207	172	176	0.82	0.84
5th pctile	228	244	236	1.07	1.03	227	210	214	0.91	0.94
Median	323	368	365	1.13	1.13	325	370	365	1.12	1.12
95th pctile	454	504	518	1.15	1.14	470	524	536	1.16	1.14
Prob. underp.				2.62%	3.35%				9.30%	8.55%

Figure 6.3 and Figure 6.4 illustrate the results for the out-of-sample (testing) data of DS1 (annual rebalancing) and DS3 (quarterly rebalancing). The corresponding CDFs are illustrated

in Appendix B. While the wealth distribution of the QD strategy is possibly preferable (Figures 6.3(a) and 6.4(a)), the wealth ratio distributions (Figures 6.3(b) and 6.4(b)) show that the QD strategy can result in a much more desirable outperformance profile than the IR strategy. Note that the potential risk of underperforming the benchmark is significantly larger out-of-sample than in-sample (for details, see Table B.2 where DS2 is used as an example), which is to be expected since with the true underlying data generating process is not known.

From a practical perspective, the CDF plots in Appendix B.2 show that the QD strategy has an 80% chance of outperforming the benchmark by about 100 bps per year. We remind the reader that this is an out-of-sample result, and makes use of standard index investments.



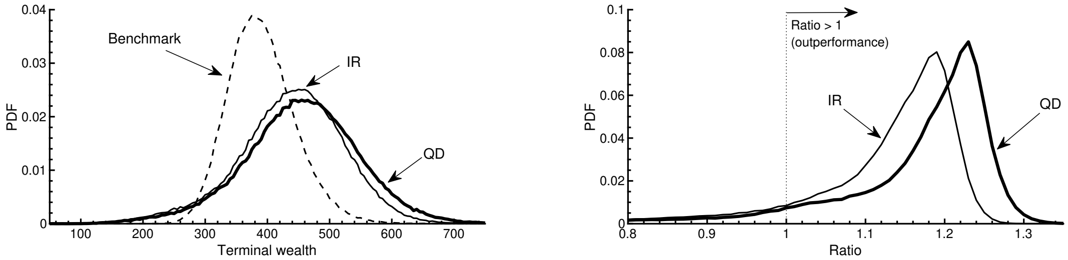
(a) PDFs of $\hat{W}(T)$, $W_j^{\mathcal{E}*}(T)$, $j \in \{ir, qd\}$ **(b)** PDFs of $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, $j \in \{ir, qd\}$

Figure 6.3: Out-of-sample (testing) results for DS1 using annual rebalancing, numerical solutions, with constraints, investor portfolio P1, benchmark BM1: Simulated probability density functions (PDFs) of benchmark and investor's target terminal wealth $\hat{W}(T)$ and $W_j^{\mathcal{E}*}(T)$, respectively, as well as the ratio $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, for $j \in \{ir, qd\}$. Note that both strategies result in $\mathcal{E} = 400$ on the *training* data of DS1, whereas figures show testing data results.

To explain the relative success of the QD strategy out-of-sample, Figure 6.5 illustrates the 80th percentiles of the proportion of wealth invested in each candidate asset[5] in P1 over time according to the IR- and QD-optimal investment strategies, on the training data set of DS1. We observe that the key qualitative observations regarding the analytical solutions discussed in Subsection 6.2 and Appendix A hold even if investment constraints are applied. Specifically, compared to the QD strategy, Figure 6.5 shows that the IR strategy maintains a larger stake in both the riskiest asset (Value) as well as the asset with the least risk (T30). In this sense, the IR strategy is less diversified than the QD strategy, in the sense that it takes more extreme positions in the assets with the most extreme risk/return trade-offs.

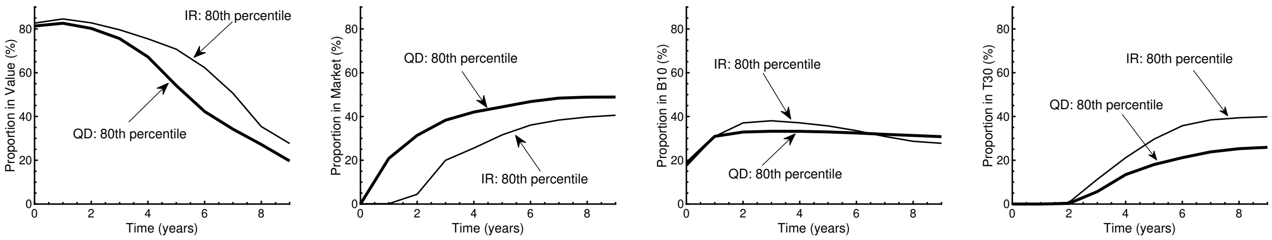
Finally, Table 6.6 presents the performance on the (single) historical path of the QD and IR strategies implemented starting the month indicated by the first column and continuing

[5]The zero investment in Size, as well as the large investment in Value, are to be expected given their historical performance (see [113]).



(a) PDFs of $\hat{W}(T)$, $W_j^{\mathcal{E}*}(T)$, $j \in \{ir, qd\}$ **(b)** PDFs of $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, $j \in \{ir, qd\}$

Figure 6.4: Out-of-sample (testing) results for DS3 using quarterly rebalancing, numerical solutions, with constraints, investor portfolio P1, benchmark BM1: Simulated probability density functions (PDFs) of benchmark and investor's target terminal wealth $\hat{W}(T)$ and $W_j^{\mathcal{E}*}(T)$, respectively, as well as the ratio $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, for $j \in \{ir, qd\}$. Note that both strategies result in $\mathcal{E} = 420$ on the *training* data of DS3, whereas figures show testing data results.



(a) Value **(b)** Market **(c)** B10 **(d)** T30

Figure 6.5: Numerical solutions, with constraints, investor portfolio P1, benchmark BM1, data set DS1, $\mathcal{E} = 400$: 80th percentile of the proportion of wealth invested in each asset over time on the training data set (DS1). Zero investment in Size, thus it is omitted. Note the same scale on the y-axis, and that the last rebalancing event is at $t = T - \Delta t = 9$ years.

until the maturity $T = 10 + t_0$ years is reached. Note that there is significant overlap (5 years) between the underlying data of each pair of adjacent rows. Table 6.6 presents *out-of-sample* results, since the probability that the actual historical path appears in the training data set constructed using block bootstrap resampling is vanishingly small ([88]). With the exception of single investment time period $[t_0, T + t_0]$ commencing in January 2000, the QD strategy consistently outperforms the IR strategy on the historical path.

Table 6.6 therefore illustrates the attractiveness in terms of historical performance of directly targeting the tracking difference using the proposed QD objective, and shows that the relatively lower reliance on the riskiest asset by the QD strategy early in the investment time horizon (Figures 6.2 and 6.5) improves its out-of-sample performance. In contrast, the IR strategy retains some resemblance to the results of MV optimization, and can be viewed as a

Table 6.6: Terminal wealth $W_j^{\mathcal{E}*}(T)$ for portfolio P1 obtained on the actual historical path by implementing the optimal strategies obtained numerically (with constraints) after training the NN on the training data sets DS1, DS2 and DS3 with benchmark BM1. The column "Best" indicates the strategy with the highest terminal wealth.

t_0 for $[t_0, T+t_0]$	Annual rebalancing							Quarterly rebalancing			
	BM1	NN trained on DS1 (1963:07 - 2009:12)			NN trained on DS2 (1963:07 - 1999:12)			BM1	NN trained on DS3 (1995:01 - 2009:12)		
		IR	QD	*Best*	IR	QD	*Best*		IR	QD	*Best*
1980:01	463	537	556	*QD*	534	548	*QD*	467	561	565	*QD*
1985:01	400	467	479	*QD*	467	475	*QD*	399	453	457	*QD*
1990:01	497	568	593	*QD*	567	588	*QD*	492	547	583	*QD*
1995:01	384	460	474	*QD*	454	459	*QD*	382	474	475	*QD*
2000:01	260	315	309	*IR*	321	310	*IR*	256	344	307	*IR*
2005:01	342	400	405	*QD*	383	406	*QD*	336	385	389	*QD*
2010:01	370	432	442	*QD*	435	438	*QD*	367	410	427	*QD*

"high conviction" strategy (see for example [76]), since it is comparatively less diversified near the start and near the end of the investment time horizon.

7. Conclusion. As noted in the Introduction, various objective functions have been formulated in the literature for benchmark outperformance. In this paper we have made the deliberate choice to target metrics which are valued by investors in practice (see the Introduction for a discussion).

We have considered two dynamic investment strategies for outperforming a benchmark, namely (i) maximizing information ratio (IR) and (ii) maximizing the tracking difference (cumulative outperformance). In the case of the tracking difference, we introduced a simple and intuitive objective function (the QD objective) for achieving this goal. Closed-form solutions under idealized assumptions are presented in order to gain intuition regarding the underlying investment strategies.

In particular, the closed form solutions show that the QD strategy is more diversified than the IR policy, and takes less risky positions. However, some properties of the closed form solutions are misleading, such as the results for probability of underperformance. We suspect that this due to allowing trading if insolvent (for the closed form solutions), similar to the pure mean-variance case [116]. This suggests that full numerical solutions with realistic constraints should be used to compare strategies.

Under certain assumptions, it can be shown that any dynamic programming approach for solving for the optimal control (which includes reinforcement learning) requires approximation of a high dimensional performance criterion, even if the control is low dimensional.

Abandoning traditional DP, we propose to directly solve the original optimal stochastic control problems, e.g., (2.4) and (2.6). In particular, we represent the control by a Neural Network (NN), which explicitly exploits its lower dimensionality. The proposed NN approach avoids inefficiency in approximating a high dimensional performance criterion (i.e. the conditional expectation), as well as avoiding potential instability from backward error propagation. Furthermore, the number of NN parameters does not depend on the number of portfolio rebalancing times.

Our approach requires sampling many stochastic paths in order to determine the optimal control. We are agnostic as to the method used to generate these paths. Our numerical examples generate these paths using parametric models calibrated to historical data, as well as block resampling of the historical data. Note that the resampling technique makes no assumptions about stochastic processes, and is popular amongst practitioners.

Both the analytical and numerical results illustrate that, compared with IR-optimal strategies with the same expected value of terminal wealth, the QD-optimal investment strategies result in comparatively more diversified asset allocations during certain periods of the investment time horizon.

Out-of-sample tests indicate that the QD-optimal strategy has an 80% chance of beating the benchmark by about about 100 bps per year. Note that this strategy does not require use of exotic instruments (e.g. alternative assets, private credit).

A. Additional analytical results and selected proofs.

In this appendix, additional analytical results are presented which relate to the various sections of the paper as indicated.

A.1. Proofs of the key results of Section 3.

Proof of Theorem 3.3. Let $\tilde{N}_i$ denote the compensated Poisson random measure ([92]) associated with the S_i-dynamics in (3.8), and define the vector

$$d\tilde{\boldsymbol{\mathcal{N}}}\left(t\right)=\left(\int_{0}^{\infty}\left(\xi_{i}-1\right)\tilde{N}_{i}\left(dt,d\xi_{i}\right):i=1,...,N_{a}^{r}\right)^{\top}. \tag{A.1}$$

It can be shown that the auxiliary process $Y_{ir}\left(t\right)$ in (3.15) has the following dynamics in terms of auxiliary control $\boldsymbol{u}\left(t\right)$ in (3.16),

$$dY_{ir}\left(t\right)=\left[rY_{ir}\left(t\right)+\left(\boldsymbol{u}\left(t\right)\right)^{\top}\tilde{\boldsymbol{\mu}}\right]dt+\left(\boldsymbol{u}\left(t\right)\right)^{\top}\boldsymbol{\sigma}\cdot d\boldsymbol{Z}\left(t\right)+\left(\boldsymbol{u}\left(t^{-}\right)\right)^{\top}\cdot d\tilde{\boldsymbol{\mathcal{N}}}\left(t\right). \tag{A.2}$$

Using the dynamics (A.2), the proof applies the techniques outlined in [92, 7] to the analysis of problem (3.17), with further details omitted.

Proof of Lemma 3.4. Considering the form of terminal condition (3.14), we make the ansatz that $V_{ir}\left(y,t\right)$ is of the form $V_{ir}\left(y,t\right)=A_{ir}\left(t\right)y^{2}+B_{ir}\left(t\right)y+C_{ir}\left(t\right)$ for unknown functions of time A_{ir},B_{ir} and C_{ir}. If this is indeed the case, then the pointwise supremum in

(3.13) is attained by the auxiliary control $\boldsymbol{u}_{ir}^{*}(t)$, where

$$\boldsymbol{u}_{ir}^{*}(t) = W_{ir}^{*}(t)\cdot \boldsymbol{\varrho}_{ir}^{*}(t, \boldsymbol{X}_{ir}^{*}(t)) - \hat{W}(t)\cdot \hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) = -\left[x + \frac{B_{ir}(t)}{2A_{ir}(t)}\right]\cdot (\boldsymbol{\Sigma}+\boldsymbol{\Lambda})^{-1}\tilde{\boldsymbol{\mu}}. \tag{A.3}$$

The substitution of V_{ir} and $\boldsymbol{u}_{ir}^{*}$ into (3.13)-(3.14) yields three ordinary differential equations (ODEs) for A_{ir}, B_{ir} and C_{ir}. Solving these equations to obtain $A_{ir}(t) = e^{(2r-\eta)(T-t)}$ and $B_{ir}(t) = -2\gamma e^{(r-\eta)(T-t)}$, where η is given by (3.10). Substitution into (A.3) and simplification results in (3.18).

After substituting the optimal control (3.18) into the dynamics of $Y_{ir}(t)$ in (A.2), we obtain the resulting auxiliary dynamics under the IR-optimal control, in other words $Y_{ir}^{*}(t) := W_{ir}^{*}(t) - \hat{W}(t)$. Techniques as in [92] give the following results

$$E_{\boldsymbol{\varrho}_{ir}^{*}}^{t_0,w_0}\left[W_{ir}^{*}(T) - \hat{W}(T)\right] = \gamma\left(1-e^{-\eta T}\right), \quad Var_{\boldsymbol{\varrho}_{ir}^{*}}^{t_0,w_0}\left[W_{ir}^{*}(T) - \hat{W}(T)\right] = \gamma^2 e^{-2\eta T}\left(e^{\eta T}-1\right), \tag{A.4}$$

so that the definition (2.2) gives the result (3.19) after some simplification.

Proof of Lemma 3.5. Given the form of (3.18), the assertion is obvious when $t = \bar{t}$. To show that (3.21) also holds for $t > \bar{t}$, we observe that combining (3.18) and (A.2) imply that the auxiliary process $Q_{ir}^{*}(t) := \gamma e^{-r(T-t)} - \left[W_{ir}^{*}(t) - \hat{W}(t)\right]$ has dynamics given by

$$\frac{dQ_{ir}^{*}(t)}{Q_{ir}^{*}(t^{-})} = (r-\eta)\cdot dt - \tilde{\boldsymbol{\mu}}^{\top}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda})^{-1}\boldsymbol{\sigma}\cdot d\boldsymbol{Z}(t) - \tilde{\boldsymbol{\mu}}^{\top}(\boldsymbol{\Sigma}+\boldsymbol{\Lambda})^{-1}\cdot d\tilde{\boldsymbol{\mathcal{N}}}(t), \tag{A.5}$$

with $Q_{ir}^{*}(\bar{t}) = 0$. Since $Q_{ir}^{*}(t) = 0$ for $t > \bar{t}$, (3.18) reduces to (3.21).

Proof of Lemma 3.6. The equivalence assertion follows from the results of [31], provided that (3.22) holds. Since in the case of no jumps, $Q_{ir}^{*}(t)$ in (A.5) reduces to a GBM with initial value $Q_{ir}^{*}(t_0) = \gamma e^{-r(T-t_0)} > 0$, we have $Q_{ir}^{*}(t) > 0$ for all $t \in [t_0, T]$, which is (3.22).

Proof of Lemma 3.7. Since it is assumed that there are no jumps in the risky asset dynamics, note that (3.10) reduces to $\eta = \tilde{\boldsymbol{\mu}}^{\top}\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{\mu}}$. Furthermore, as noted in the proof of Lemma 3.6, in the case of no jumps the dynamics of $Q_{ir}^{*}(t)$ in (A.5) is a GBM, with (3.23) following from the relationship $P_{\boldsymbol{\varrho}_{ir}^{*}}^{t_0,w_0}\left[W_{ir}^{*}(t) \leq \hat{W}(t)\right] = P_{\boldsymbol{\varrho}_{ir}^{*}}^{t_0,w_0}\left[Q_{ir}^{*}(t) \geq \gamma e^{-r(T-t)}\right]$.

Proof of Theorem 3.9. The dynamics of the auxiliary process $Y_{qd}(t)$ defined in (3.26) can be written in terms of the auxiliary control $\boldsymbol{v}(t)$, defined in (3.27), as

$$dY_{qd}(t) = \left[h_{\beta}'(t) - r\cdot(h_{\beta}(t) - Y_{qd}(t)) + (\boldsymbol{v}(t))^{\top}\tilde{\boldsymbol{\mu}}\right]dt + (\boldsymbol{v}(t))^{\top}\boldsymbol{\sigma}\cdot d\boldsymbol{Z}(t) + \left(\boldsymbol{v}(t^{-})\right)^{\top}\cdot d\tilde{\boldsymbol{\mathcal{N}}}(t). \tag{A.6}$$

Here, for a fixed value of the parameter β and the contribution rate q, we define $h_\beta(t)$ as the following function of time (this definition is also used in Lemma 3.10),

$$h_\beta(t) := \left(e^{\beta T}-1\right)\cdot\int_t^T qe^{-r(T-z)}dz = \frac{q}{r}\left(e^{\beta T}-1\right)\left(1-e^{-r(T-t)}\right), \qquad t\in[t_0,T], \tag{A.7}$$

with $h'_\beta(t) = \frac{d}{dt}h_\beta(t)$. The results of Theorem 3.9 then follows from the application of the techniques outlined in [92].

Proof of Lemma 3.10. The terminal condition (3.25) suggests an ansatz for V_{qd} that is quadratic in y, in other words $V_{qd}(y,t) = A_{qd}(t)y^2 + B_{qd}(t)y + C_{qd}(t)$. In this case, the pointwise supremum in (3.24) is attained by the auxiliary control $\boldsymbol{v}^*_{qd}(t)$ with a qualitatively similar form in terms of (y,t) as the result reported in (A.3). Substituting V_{qd} and $\boldsymbol{v}^*_{qd}$ into (3.24)-(3.25) yields ODEs for A_{qd}, B_{qd} and C_{qd}, which are solved to obtain $A_{qd}(t) = e^{(2r-\eta)(T-t)}$ and

$$B_{qd}(t) = \frac{2q}{r}\left(1-e^{\beta T}\right)\cdot\left[e^{(2r-\eta)(T-t)} - e^{(r-\eta)(T-t)}\right], \tag{A.8}$$

where η is given by (3.10). The necessary substitution and simplification yields (3.29).

Proof of Lemma 3.11. Substituting (A.7) into (3.30), note that condition (3.30) can equivalently be written as

$$W^*_{qd}(\bar{t}) + \int_{\bar{t}}^T qe^{-r(T-z)}dz = e^{\beta T}\cdot\left[\hat{W}(\bar{t}) + \int_{\bar{t}}^T qe^{-r(T-z)}dz\right], \tag{A.9}$$

which provides intuition as to why result (3.31) should hold. The proof proceeds along the same lines as in the case of Lemma 3.5, except that (3.31) can be established using the properties of the auxiliary process

$$Q^*_{qd}(t) := h_\beta(t) - \left[W^*_{qd}(t) - e^{\beta T}\hat{W}(t)\right], \tag{A.10}$$

which has dynamics that are formally the same as those of Q^*_{ir} in (A.5).

Proof of Lemma 3.12. The proof is structurally similar to that of Lemma 3.6, but follows from analyzing the properties of Q^*_{qd} in (A.10) after setting $\boldsymbol{\lambda} = \mathbf{0}$.

Proof of Lemma 3.13. Using the definition of Q^*_{qd} in (A.10), and observing that $q=0$ implies that $h_\beta(t) \equiv 0$ for all t, we have $P^{t_0,w_0}_{\boldsymbol{\varrho}^*_{qd}}\left[W^*_{qd}(t) \leq \hat{W}(t)\right] = P^{t_0,w_0}_{\boldsymbol{\varrho}^*_{qd}}\left[Q^*_{qd}(t) \geq \left(e^{\beta T}-1\right)\hat{W}(t)\right]$. Recalling that the dynamics of Q^*_{qd} are formally the same as the dynamics of Q^*_{ir} in (A.5), under the stated conditions of this lemma it can be shown that $Q^*_{qd}(t) \geq \left(e^{\beta T}-1\right)\hat{W}(t)$ if and only if

$$\left[\hat{\boldsymbol{\varrho}}^{\top}+\tilde{\boldsymbol{\mu}}^{\top}\boldsymbol{\Sigma}^{-1}\right]\cdot\boldsymbol{\sigma}\cdot\boldsymbol{Z}\left(t\right)\leq\left[\frac{1}{2}\hat{\boldsymbol{\varrho}}^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{\varrho}}-\tilde{\boldsymbol{\mu}}^{\top}\hat{\boldsymbol{\varrho}}-\frac{3}{2}\eta\right]t. \tag{A.11}$$

Observing that the left-hand side of (A.11) is a normally distributed random variable with zero mean and a variance of $\left[\hat{\boldsymbol{\varrho}}^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{\varrho}}+2\tilde{\boldsymbol{\mu}}^{\top}\hat{\boldsymbol{\varrho}}+\eta\right]\cdot t$, the result (3.33) follows.

Proof of Lemma 3.15. The assumptions of Lemma 3.13 are required to hold since the proof requires the analytical result (3.33) for the left-hand side of (3.37). Since this also implies that the assumptions of Lemma 3.7 are satisfied, the right-hand side of (3.37) is given by (3.23). Using the fact that the CDF $\Phi\left(\cdot\right)$ is non-decreasing, it then follows that (3.33) holds if and only if

$$-\frac{3}{2}\sqrt{\eta}\left[\hat{\boldsymbol{\varrho}}^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{\varrho}}+2\tilde{\boldsymbol{\mu}}^{\top}\hat{\boldsymbol{\varrho}}+\eta\right]^{1/2}\cdot\sqrt{t}\leq\left[\frac{1}{2}\hat{\boldsymbol{\varrho}}^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{\varrho}}-\tilde{\boldsymbol{\mu}}^{\top}\hat{\boldsymbol{\varrho}}-\frac{3}{2}\eta\right]\cdot\sqrt{t}, \tag{A.12}$$

where (since the assumptions of Lemma 3.13 including the absence of jumps in the risky asset processes hold), we have $\eta=\tilde{\boldsymbol{\mu}}^{\top}\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{\mu}}$. Since $\boldsymbol{\Sigma}$ is positive definite, so is $\boldsymbol{\Sigma}^{-1}$. Therefore, there exists matrices $\boldsymbol{\Sigma}^{1/2}$ and $\boldsymbol{\Sigma}^{-1/2}$ such that we have the (unique) decompositions $\boldsymbol{\Sigma}=\boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{1/2}$ and $\boldsymbol{\Sigma}^{-1}=\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Sigma}^{-1/2}$. As a result, recalling the conditions on the (constant proportion) benchmark strategy $\hat{\boldsymbol{\varrho}}$ and the assumption that the risky asset drift terms satisfy $\mu_i>r$ for all $i\in\{1,...,N_a^r\}$, the Cauchy-Schwarz inequality implies that

$$\begin{aligned}
-\frac{3}{2}\sqrt{\eta}\left[\hat{\boldsymbol{\varrho}}^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{\varrho}}+2\tilde{\boldsymbol{\mu}}^{\top}\hat{\boldsymbol{\varrho}}+\eta\right]^{1/2} &= -\frac{3}{2}\left\|\boldsymbol{\Sigma}^{-1/2}\cdot\tilde{\boldsymbol{\mu}}\right\|_2\left\|\boldsymbol{\Sigma}^{1/2}\hat{\boldsymbol{\varrho}}+\boldsymbol{\Sigma}^{-1/2}\tilde{\boldsymbol{\mu}}\right\|_2\\
&\leq -\frac{3}{2}\left(\tilde{\boldsymbol{\mu}}^{\top}\hat{\boldsymbol{\varrho}}+\eta\right)\\
&< \frac{1}{2}\hat{\boldsymbol{\varrho}}^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{\varrho}}-\tilde{\boldsymbol{\mu}}^{\top}\hat{\boldsymbol{\varrho}}-\frac{3}{2}\eta, \qquad \text{(A.13)}
\end{aligned}$$

thereby confirming that (3.37) holds for all $t\geq t_0=0$.

A.2. Additional analytical comparison results. As a supplement to Subsection 3.4, we present additional analytical comparison results which rely on specific choices of γ and β for the $IR\left(\gamma\right)$ and $QD\left(\beta\right)$ problems, respectively. Since the strategies are compared in Section 6 on the basis of equal expectation of terminal wealth (see (6.1)), we formally introduce Assumption A.1 outlining the basis of the comparison of the subsequent results. Note that these results are all derived within the setting of Section 3.

Assumption A.1. (Expected value target for terminal wealth) Assume that Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12) hold. Suppose that the benchmark investment strategy, given by the fractions of wealth in the risky assets $\hat{\boldsymbol{\varrho}}\left(t,\hat{W}\left(t\right)\right)$, results in an expected value of benchmark terminal wealth satisfying

$$E_{\hat{\varrho}}^{t_0,w_0}\left[\hat{W}(T)\right] := \mathcal{K}, \qquad \text{where } \mathcal{K} > w_0 e^{rT}. \tag{A.14}$$

We assume the investor chooses parameters $\gamma = \gamma_{ir}^{\mathcal{E}}$ in the $IR(\gamma)$ problem and $\beta = \beta_{qd}^{\mathcal{E}}$ in the $QD(\beta)$ problem such that the associated IR- and QD-optimal strategies $\boldsymbol{\varrho}_{ir}^{\mathcal{E}*}$ and $\boldsymbol{\varrho}_{qd}^{\mathcal{E}*}$, respectively, result in the same desired expected value of terminal wealth,

$$E_{\boldsymbol{\varrho}_{ir}^{\mathcal{E}*}}^{t_0,w_0}\left[W_{ir}^{\mathcal{E}*}(T)\right] = E_{\boldsymbol{\varrho}_{qd}^{\mathcal{E}*}}^{t_0,w_0}\left[W_{qd}^{\mathcal{E}*}(T)\right] = \mathcal{E} = e^{\hat{\beta}T}\mathcal{K}, \quad \text{for some } \hat{\beta} > 0. \tag{A.15}$$

The value of $\mathcal{E}$ (A.15) will be referred to as the expected value target for terminal wealth.

Subject to the assumptions of Section 3, the following lemma shows that the values of $\gamma = \gamma_{ir}^{\mathcal{E}}$ and $\beta = \beta_{qd}^{\mathcal{E}}$ achieving (A.15) can be derived analytically.

Lemma A.2. *(Analytical values γ and β achieving expected value target). Suppose that Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12) hold. The optimal controls of problems $IR\left(\gamma = \gamma_{ir}^{\mathcal{E}}\right)$ and $QD\left(\beta = \beta_{qd}^{\mathcal{E}}\right)$ achieve the required expected value target $E_{\boldsymbol{\varrho}_j^{\mathcal{E}*}}^{t_0,w_0}\left[W_j^{\mathcal{E}*}(T)\right] \equiv \mathcal{E}, j \in \{ir, qd\}$, provided $\gamma_{ir}^{\mathcal{E}}$ and $\beta_{qd}^{\mathcal{E}}$ are given respectively by*

$$\gamma_{ir}^{\mathcal{E}} = \frac{(\mathcal{E} - \mathcal{K})}{(1 - e^{-\eta T})}, \qquad \text{and} \qquad \beta_{qd}^{\mathcal{E}} = \frac{1}{T}\log\left[\frac{\mathcal{E} - \left[\frac{q}{r}\left(1 - e^{-rT}\right) + w_0\right]e^{(r-\eta)T}}{\mathcal{K} - \left[\frac{q}{r}\left(1 - e^{-rT}\right) + w_0\right]e^{(r-\eta)T}}\right], \tag{A.16}$$

where η is given by (3.10).

Proof. Using $\gamma_{ir}^{\mathcal{E}}$ as an example, we rearrange (A.4) and use definition (A.15). The value of $\beta_{qd}^{\mathcal{E}}$ is obtained similarly. ∎

To provide further analysis of the particular results observed in Subsection 6.2, we present the following closed-form result for the specific case of 2 assets (a single risky asset and a risk-free asset) in combination with a constant proportion benchmark strategy.

Theorem A.3. *(QD-optimal vs. IR-optimal strategies, $N_a = 2$: Risky asset exposure over time) Suppose the following assumptions hold: (i) Assumption 3.1, Assumption 3.2 and wealth dynamics (3.11)-(3.12) with a single risky asset ($N_a^r = 1$); (ii) the investor compares investment strategies on the basis of Assumption A.1; (iii) contributions are zero ($q = 0$); (iv) the benchmark strategy is a constant proportion strategy with $\hat{\boldsymbol{\varrho}}\left(t, \hat{W}(t)\right) \equiv \hat{\varrho} > 0$ for $t \in [t_0, T]$. Note that $\boldsymbol{X}_{ir}^{\mathcal{E}*}(t) := \left(W_{ir}^{\mathcal{E}*}(t), \hat{W}(t), \hat{\varrho}\right)$ and $\boldsymbol{X}_{qd}^{\mathcal{E}*}(t) := \left(W_{qd}^{\mathcal{E}*}(t), \hat{W}(t), \hat{\varrho}\right)$.*

Then, at inception $t = t_0 = 0$, the IR-optimal strategy $\boldsymbol{\varrho}_{ir}^{\mathcal{E}*} := \varrho_{ir}^{\mathcal{E}*}$ requires a larger investment in the single risky asset than the QD-optimal strategy $\boldsymbol{\varrho}_{qd}^{\mathcal{E}*} := \varrho_{qd}^{\mathcal{E}*}$,

$$\varrho_{ir}^{\mathcal{E}*}\left(t_0, \boldsymbol{X}_{ir}^{\mathcal{E}*}(t_0)\right) > \varrho_{qd}^{\mathcal{E}*}\left(t_0; \boldsymbol{X}_{qd}^{\mathcal{E}*}(t_0)\right). \tag{A.17}$$

At maturity $t = T$, the IR-optimal strategy is expected to invest less wealth in the risky asset than the QD-optimal strategy,

$$E^{t_0,w_0}_{\varrho^{\mathcal{E}*}_{ir}}\left[\varrho^{\mathcal{E}*}_{ir}\left(T,\boldsymbol{X}^{\mathcal{E}*}_{ir}(T)\right)\cdot W^{\mathcal{E}*}_{ir}(T)\right] < E^{t_0,w_0}_{\varrho^{\mathcal{E}*}_{qd}}\left[\varrho^{\mathcal{E}*}_{qd}\left(T,\boldsymbol{X}^{\mathcal{E}*}_{qd}(T)\right)\cdot W^{\mathcal{E}*}_{qd}(T)\right]. \tag{A.18}$$

Furthermore, if it is additionally assumed that η (see (3.10)) satisfies $\eta > r$, then the function

$$t \rightarrow f(t) := E^{t_0,w_0}_{\varrho^{\mathcal{E}*}_{ir}}\left[\varrho^{\mathcal{E}*}_{ir}\left(t,\boldsymbol{X}^{\mathcal{E}*}_{ir}(t)\right)\cdot W^{\mathcal{E}*}_{ir}(t)\right] - E^{t_0,w_0}_{\varrho^{\mathcal{E}*}_{qd}}\left[\varrho^{\mathcal{E}*}_{qd}\left(t,\boldsymbol{X}^{\mathcal{E}*}_{qd}(t)\right)\cdot W^{\mathcal{E}*}_{qd}(t)\right] \tag{A.19}$$

is monotonically decreasing on $t \in [t_0, T]$.

Proof. Considering benchmark wealth dynamics (3.12) after setting $\hat{\boldsymbol{\varrho}}\left(t,\hat{W}(t)\right) \equiv \hat{\varrho} > 0$ and $q = 0$, it can be shown that a given value of $E^{t_0,w_0}_{\hat{\mathcal{P}}}\left[\hat{W}(T)\right] \equiv \mathcal{K}$ can be achieved by choosing the constant $\hat{\varrho}$ according to

$$\hat{\varrho} = \frac{1}{(\mu - r)\,T}\log\left(\frac{\mathcal{K}}{w_0 e^{rT}}\right), \qquad (\text{if } q = 0). \tag{A.20}$$

where we recall that $\mathcal{K} > w_0 e^{rT}$ (see (A.15)). Combining, under the stated assumptions, the results (3.18), (3.29), (A.15), (A.16) and (A.20), tedious algebra results in the function $f(t)$ in (A.19) given by the following expression on $t \in [t_0 = 0, T]$,

$$f(t) = \frac{(\mathcal{E} - \mathcal{K})\,w_0}{(\mu - r)\left(\mathcal{K} - w_0 e^{(r-\eta)T}\right)}\cdot\left[\left(\frac{\eta}{1 - e^{-\eta T}}\right)\left(\frac{\mathcal{K}}{w_0 e^{rT}} - 1\right)\cdot e^{(r-\eta)t} - \frac{1}{T}\log\left(\frac{\mathcal{K}}{w_0 e^{rT}}\right)\cdot\left(\frac{\mathcal{K}}{w_0}\right)^{t/T}\right]. \tag{A.21}$$

The results (A.17), (A.18) and (A.19) follow from an analysis of the properties of the function f (A.21). ■

Note that the additional requirement $\eta > r$ leading to (A.19) is indeed satisfied in the case of typical process parameters, including by the parameters in Table B.1.

Theorem A.3 suggests that in order to achieve the same expected value of terminal wealth, the IR strategy relies on a larger investment in the riskiest asset early in the investment time horizon than the QD strategy. Once the desired outperformance become increasingly likely, the IR strategy's exposure to the riskiest asset is expected to be reduced to a level below that of the QD strategy. Note that the qualitative implications of Theorem A.3 hold even if the underlying assumptions are relaxed (see Section 6).

A.3. Proof of Proposition 4.1. In this proof, we consider only the QD problem (2.6), since the proof for the IR problem (2.4) proceeds along similar lines.

In the case of discrete rebalancing and cash injections into the portfolio at each $t_n \in \mathcal{T}$, we consider the *amounts* invested in each asset, since it is no longer sufficient to consider only the aggregate wealth processes for reasons that will become obvious when using the dynamic programming (DP) approach for solving the problems. To this end, let $\boldsymbol{U}(t) = (U_i(t) : i = 1, ..., N_a)^\top$ and $\hat{\boldsymbol{U}}(t) = \left(\hat{U}_i(t) : i = 1, ..., N_a\right)^\top$ denote the amounts invested at time t in each asset, according to the investor and benchmark strategy, respectively. The investor and benchmark wealth therefore satisfy $W(t) = \sum_{i=1}^{N_a} U_i(t)$ and $\hat{W}(t) = \sum_{i=1}^{N_a} \hat{U}_i(t)$, respectively.

For an arbitrary admissible investor strategy $\mathcal{P} \in \mathcal{A}$ with discrete rebalancing, define $\mathcal{P}_t = \{\boldsymbol{p}(t_m, \boldsymbol{X}(t_m)) \in \mathcal{P} | \, t_m \geq t, t_m \in \mathcal{T}\}$, where $\mathcal{T}$ is given by (4.1). To solve the QD problem (2.6) using DP, we define the performance criterion (see [92]), which at time $t \in [t_0, T]$ is given by the conditional expectation

$$J\left(t, \boldsymbol{u}^-, \hat{\boldsymbol{u}}^-, \mathcal{P}_t\right) = E_{\mathcal{P}_t}^{t^-, \boldsymbol{u}^-, \hat{\boldsymbol{u}}^-}\left[\left(W(T) - e^{\beta T}\hat{W}(T)\right)^2 \middle| \left(\boldsymbol{U}(t^-), \hat{\boldsymbol{U}}(t^-)\right) = \left(\boldsymbol{u}^-, \hat{\boldsymbol{u}}^-\right)\right], \tag{A.22}$$

where $\boldsymbol{u}^- = \left(u_1^-, ..., u_{N_a}^-\right)^\top$ and $\hat{\boldsymbol{u}}^- = \left(\hat{u}_1^-, ..., \hat{u}_{N_a}^-\right)^\top$. Note that (A.22) is not just defined at rebalancing times.

Fix a rebalancing time $t_n \in \mathcal{T}$ and given cash contribution $q(t_n)$, and introduce the notation $\boldsymbol{p}_n := \boldsymbol{p}(t_n, \boldsymbol{X}(t_n))$ and $\mathcal{P}_n = \{\boldsymbol{p}_m \in \mathcal{P} | \, t_m \geq t_n\}$,so that $\mathcal{P}_n = \boldsymbol{p}_n \cup \mathcal{P}_{n+1}$. We also define $\mathcal{A}_n = \{\mathcal{P}_n | \, \boldsymbol{p} \in \mathcal{Z}, \quad \forall \boldsymbol{p} \in \mathcal{P}_n\}$. The investor and benchmark wealth immediately prior to the cash contribution at t_n is therefore given by $W(t_n^-) := w^- = \sum_{i=1}^{N_a} u_i^-$ and $\hat{W}(t_n^-) := \hat{w}^- = \sum_{i=1}^{N_a} \hat{u}_i^-$, respectively. After incorporating the cash contribution $q(t_n)$, we therefore have $W(t_n^+) := w^+ = w^- + q(t_n)$ and $\hat{W}(t_n^-) := \hat{w}^+ = \hat{w}^- + q(t_n)$. As per the stated assumptions of Proposition 4.1, the investor can observe the benchmark allocation $\hat{\boldsymbol{p}}_n := \hat{\boldsymbol{p}}(t_n, \hat{w}^+)$, while we have amount dynamics between rebalancing events, i.e. for $t \in (t_n, t_{n+1})$, given by

$$\frac{d\boldsymbol{U}(t)}{\boldsymbol{U}(t^-)} = \left(\boldsymbol{\mu} - \boldsymbol{\lambda} \circ \boldsymbol{\kappa}^{(1)}\right) dt + \boldsymbol{\sigma} \cdot d\boldsymbol{Z}(t) + d\boldsymbol{\mathcal{N}}(t), \qquad \boldsymbol{U}\left(t_n^+\right) = \boldsymbol{u}^+ = w^+ \cdot \boldsymbol{p}_n, \tag{A.23}$$

$$\frac{d\hat{\boldsymbol{U}}(t)}{\hat{\boldsymbol{U}}(t^-)} = \left(\boldsymbol{\mu} - \boldsymbol{\lambda} \circ \boldsymbol{\kappa}^{(1)}\right) dt + \boldsymbol{\sigma} \cdot d\boldsymbol{Z}(t) + d\boldsymbol{\mathcal{N}}(t), \qquad \hat{\boldsymbol{U}}\left(t_n^+\right) = \hat{\boldsymbol{u}}^+ = \hat{w}^+ \cdot \hat{\boldsymbol{p}}_n. \tag{A.24}$$

By definition of the QD problem, at rebalancing time t_n we therefore have the auxiliary value function

$$V\left(t_n^-, w^-, \hat{w}^-\right) = \inf_{\mathcal{P}_n \in \mathcal{A}_n} E_{\mathcal{P}_n}^{t^-, w^-, \hat{w}^-}\left[\left(W(T) - e^{\beta T}\hat{W}(T)\right)^2 \middle| \left(W(t_n^-), \hat{W}(t_n^-)\right) = \left(w^-, \hat{w}^-\right)\right], \tag{A.25}$$

$$\equiv J\left(t_n^-, \boldsymbol{u}^-, \hat{\boldsymbol{u}}^-, \mathcal{P}_n^* = \boldsymbol{p}_n^* \cup \mathcal{P}_{n+1}^*\right), \tag{A.26}$$

where $\mathcal{P}_n^* \in \mathcal{A}_n$ denotes the control realizing the infimum in (A.25), whereas the dependence of (A.25) and (A.26) on $(w^-, \hat{w}^-)$ and $\left(\boldsymbol{u}^-, \hat{\boldsymbol{u}}^-\right)$, respectively, will be clarified below.

At the terminal time T, there are no rebalancing events (i.e. no control applied) or cash contributions, so in the case of the QD problem we simply have

$$V\left(T, w, \hat{w}\right) = V\left(T^-, w^- = \sum_{i=1}^{N_a} u_i^-, \hat{w}^- = \sum_{i=1}^{N_a} \hat{u}_i^-\right)$$

$$\equiv J\left(T^-, \boldsymbol{u}^-, \hat{\boldsymbol{u}}^-, \mathcal{P}_{N_{rb}}^* \equiv \emptyset\right) = \left[\left(\sum_{i=1}^{N_a} u_i^-\right) - e^{\beta T}\left(\sum_{i=1}^{N_a} \hat{u}_i^-\right)\right]^2, \tag{A.27}$$

From (A.27), it is obvious that the performance criterion J and value function V at time T can be expressed as a function of the investor wealth and benchmark wealth only.

Stepping backwards in time, consider the problem at a fixed rebalancing time $t_n \in \mathcal{T}$, and assume that the function $J\left(t_{n+1}^-, \boldsymbol{u}^-, \hat{\boldsymbol{u}}^-, \mathcal{P}_{n+1}\right)$ is given, along with the optimal control $\mathcal{P}_{n+1}^*$ which is applicable to the interval $[t_{n+1}, T]$. Despite the fact that by (A.26), we have $V\left(t_{n+1}^-, w^-, \hat{w}^-\right) = J\left(t_{n+1}^-, \boldsymbol{u}^-, \hat{\boldsymbol{u}}^-, \mathcal{P}_{n+1}^*\right)$, we do require the performance criterion $J\left(t_{n+1}^-, \cdot\right)$ as a function of the amounts $\left(\boldsymbol{u}^-, \hat{\boldsymbol{u}}^-\right)$, since $J\left(t_{n+1}^-, \boldsymbol{u}^-, \hat{\boldsymbol{u}}^-, \mathcal{P}_{n+1}^*\right)$ will serve as the terminal condition to be satisfied by the (at this point, unknown) performance criterion function $J\left(t, \boldsymbol{u}, \hat{\boldsymbol{u}}, \mathcal{P}_t\right), t \in (t_n, t_{n+1})$. Between rebalancing times, i.e. for $t \in (t_n, t_{n+1})$, there are no controls applied, cash flows or discounting. Considering the role of inflation, note that we can always make use of inflation-adjusted quantities, as is done in Section 6. The dynamic programming principle, definition (A.22) and dynamics (A.23)-(A.24) therefore imply that $J\left(t, \boldsymbol{u}, \hat{\boldsymbol{u}}, \mathcal{P}_t\right)$ satisfies the following $(2N_a + 1)$-dimensional PIDE on $t \in (t_n, t_{n+1})$ with given terminal condition $J\left(t_{n+1}^-, \boldsymbol{u}^-, \hat{\boldsymbol{u}}^-, \mathcal{P}_{n+1}^*\right)$:

$$\begin{aligned}
0 = {} & J_t + \left(\boldsymbol{u} \circ \left[\boldsymbol{\mu} - \left(\boldsymbol{\lambda} \circ \boldsymbol{\kappa}^{(1)}\right)\right]\right)^\top \cdot \boldsymbol{\nabla} J_{\boldsymbol{u}} + \left(\hat{\boldsymbol{u}} \circ \left[\boldsymbol{\mu} - \left(\boldsymbol{\lambda} \circ \boldsymbol{\kappa}^{(1)}\right)\right]\right)^\top \cdot \boldsymbol{\nabla} J_{\hat{\boldsymbol{u}}} \\
& + \frac{1}{2} \mathrm{tr}\left[\mathrm{diag}\left(\boldsymbol{u}\right) \cdot \boldsymbol{\Sigma} \cdot \mathrm{diag}\left(\boldsymbol{u}\right) \cdot \boldsymbol{\nabla}^2 J_{\boldsymbol{u}\boldsymbol{u}}\right] + \frac{1}{2} \mathrm{tr}\left[\mathrm{diag}\left(\hat{\boldsymbol{u}}\right) \cdot \boldsymbol{\Sigma} \cdot \mathrm{diag}\left(\hat{\boldsymbol{u}}\right) \cdot \boldsymbol{\nabla}^2 J_{\hat{\boldsymbol{u}}\hat{\boldsymbol{u}}}\right] \\
& + \mathrm{tr}\left[\mathrm{diag}\left(\boldsymbol{u}\right) \cdot \boldsymbol{\Sigma} \cdot \mathrm{diag}\left(\hat{\boldsymbol{u}}\right) \cdot \boldsymbol{\nabla}^2 J_{\boldsymbol{u}\hat{\boldsymbol{u}}}\right] - \left(\sum_{i=1}^{N_a} \lambda_i\right) \cdot J\left(t, \boldsymbol{u}, \hat{\boldsymbol{u}}\right) \\
& + \sum_{i=1}^{N_a} \lambda_i \int_0^\infty \left[J\left(t\ ,\ \boldsymbol{u} + u_i\left(\xi_i - 1\right) \cdot \boldsymbol{e}_i\ ,\ \hat{\boldsymbol{u}} + \hat{u}_i\left(\xi_i - 1\right) \cdot \boldsymbol{e}_i\right)\right] f_{\xi_i}\left(\xi_i\right) d\xi_i.
\end{aligned} \tag{A.28}$$

In (A.28), $\mathrm{tr}\left(\cdot\right)$ denotes the trace of a matrix, $\mathrm{diag}\left(\boldsymbol{v}\right)$ denotes the diagonal matrix with vector $\boldsymbol{v}$ on the main diagonal, $\boldsymbol{e}_i \in \mathbb{R}^{N_a}$ is the ith standard basis vector in $\mathbb{R}^{N_a}$, and we have gradients $\boldsymbol{\nabla} J_{\boldsymbol{u}} = \left[\frac{\partial J}{\partial u_i} : i = 1, ..., N_a\right]^\top$ and $\boldsymbol{\nabla} J_{\hat{\boldsymbol{u}}} = \left[\frac{\partial J}{\partial \hat{u}_i} : i = 1, ..., N_a\right]^\top$, as well as matrices of second derivatives $\boldsymbol{\nabla}^2 J_{\boldsymbol{u}\boldsymbol{u}} = \left(\frac{\partial^2 J}{\partial u_i \partial u_j}\right)_{i,j=1,...,N_a}$, $\boldsymbol{\nabla}^2 J_{\hat{\boldsymbol{u}}\hat{\boldsymbol{u}}} = \left(\frac{\partial^2 J}{\partial \hat{u}_i \partial \hat{u}_j}\right)_{i,j=1,...,N_a}$ as well as

$\nabla^2 J_{\boldsymbol{u}\hat{\boldsymbol{u}}} = \left(\frac{\partial^2 J}{\partial u_i \partial \hat{u}_j}\right)_{i,j=1,\ldots,N_a}$.

Let $\underline{J}$ denote the lower semi-continuous envelope of the function J obtained by solving (A.28). Under the stated assumptions, the QD-optimal control at time t_n is therefore a function of the investor wealth w^+ and benchmark wealth $\hat{w}^+$ (after the cash injection) only, since

$$\boldsymbol{p}_n^* = \boldsymbol{p}^*\left(t_n, w^+, \hat{w}^+\right) = \underset{\boldsymbol{p}_n \in \mathcal{Z}}{\operatorname{argmin}} \underline{J}\left(t_n^+, \boldsymbol{u}^+ = w^+ \cdot \boldsymbol{p}_n, \hat{\boldsymbol{u}}^+ = \hat{w}^+ \cdot \hat{\boldsymbol{p}}_n, \mathcal{P}_n = \boldsymbol{p}_n \cup \mathcal{P}_{n+1}^*\right), \tag{A.29}$$

with $w^+ = \sum_{i=1}^{N_a} u_i^- + q\left(t_n\right)$ and $\hat{w}^+ = \sum_{i=1}^{N_a} \hat{u}_i^- + q\left(t_n\right)$. Applying the DP principle at t_n, we advance J backwards across the rebalancing event at t_n, and also obtain the value function at time t_n, using

$$V\left(t_n^-, w^-, \hat{w}^-\right) = \underline{J}\left(t_n^+, \boldsymbol{u}^+ = w^+ \cdot \boldsymbol{p}_n^*, \hat{\boldsymbol{u}}^+ = \hat{w}^+ \cdot \hat{\boldsymbol{p}}_n, \mathcal{P}_n^* = \boldsymbol{p}_n^* \cup \mathcal{P}_{n+1}^*\right) \tag{A.30}$$

$$= J\left(t_n^-, \boldsymbol{u}^-, \hat{\boldsymbol{u}}^-, \mathcal{P}_n^*\right), \tag{A.31}$$

where $\boldsymbol{p}_n^*$ is given by (A.29).

The results (A.27) and (A.31) therefore show that it is only at each fixed rebalancing event $t_n \in \mathcal{T}$ and at the terminal time T can we express the performance criterion J as a function of investor and benchmark wealth. By definition, at each rebalancing time J also coincides with the value function if the optimal control is used, and therefore at each fixed $t_n \in \mathcal{T}$ the value function is also only a function of the investor and benchmark wealth. However, in general, the DP approach requires the solution of a $(2N_a+1)$-dimensional performance criterion $J : \mathbb{R}^{(2N_a+1)} \to \mathbb{R}$, obtained in this case by solving the PIDE (A.28).

B. Supplementary information for numerical results. This appendix provides supplementary information for the numerical results of Section 6.

B.1. Source data and parameters. The historical returns data for the basic assets such as the T-bills/bonds and the broad market index were obtained from the CRSP [6], whereas factor data for Size and Value (see [39, 38]) were obtained from Kenneth French's data library[7] (KFDL). The detailed time series sourced for each asset is as follows:

(i) T30 (30-day Treasury bill): CRSP, monthly returns for 30-day Treasury bill.
(ii) B10 (10-year Treasury bond): CRSP, monthly returns for 10-year Treasury bond.
(iii) Market (broad equity market index): CRSP, monthly returns, including dividends and distributions, for a capitalization-weighted index consisting of all domestic stocks trading on major US exchanges (the VWD index).

[6] Calculations were based on data from the Historical Indexes 2020©, Center for Research in Security Prices (CRSP), The University of Chicago Booth School of Business. Wharton Research Data Services 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.

[7] See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

(iv) Size (Portfolio of small stocks): KFDL, "Portfolios Formed on Size", which consists of monthly returns on a capitalization-weighted index consisting of the firms (listed on major US exchanges) with market value of equity, or market capitalization, at or below the 30th percentile (i.e. smallest 30%) of market capitalization values of NYSE-listed firms.

(v) Value (Portfolio of value stocks): KFDL, "Portfolios Formed on Book-to-Market", which consists of monthly returns on a capitalization-weighted index of the firms (listed on major US exchanges) consisting of the firms (listed on major US exchanges) with book-to-market value of equity ratios at or above the 70th percentile (i.e. highest 30%) of book-to-market ratios of NYSE-listed firms.

Data was obtained for the period from 1963:07 to 2020:12, and inflation-adjusted using inflation data from the US Bureau of Labor Statistics[8].

For the illustration of analytical solutions in Subsection 6.2, the parameters of (3.5) and (3.8) are to be determined. We use the same calibration methodology as outlined in [29, 43], and assume that the risky asset evolves according to the dynamics of the [73] model, with $\log \xi$ having an asymmetric double-exponential distribution,

$$f_{\xi}(\xi) = \nu\zeta_1\xi^{-\zeta_1-1}\mathbb{I}_{[\xi\geq 1]}(\xi) + (1-\nu)\,\zeta_2\xi^{\zeta_2-1}\mathbb{I}_{[0\leq\xi<1]}(\xi)\,, \upsilon\in[0,1] \text{ and } \zeta_1>1, \zeta_2>0, \tag{B.1}$$

where ν denotes the probability of an upward jump given that a jump occurs. Table B.1 summarizes the resulting parameters obtained using the filtering technique for the calibration of jump diffusion processes - see [29, 43] for the relevant methodological details.

Table B.1: Analytical solutions: Calibrated, inflation-adjusted parameters for asset dynamics (3.5) and (3.8), with $f_{\xi}(\xi)$ given by (B.1). For calibration purposes, a jump threshold equal to 3 has been used in the methodology of [29].

Parameter	r	μ	σ	λ	υ	ζ_1	ζ_2
Value	0.0074	0.0749	0.1392	0.2090	0.2500	7.7830	6.1074

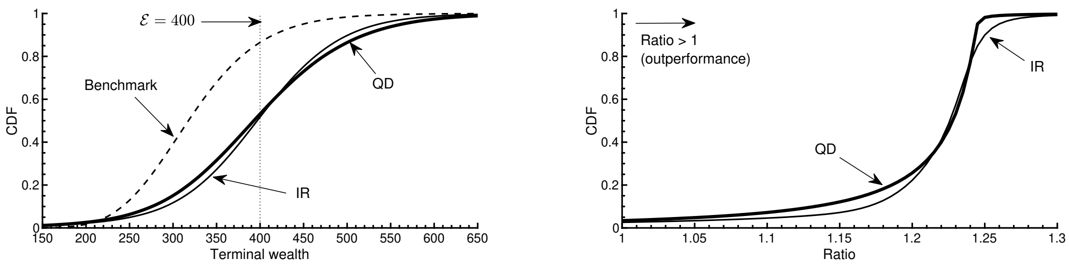
B.2. Additional numerical results.

As a supplement to the results in Subsection 6.2, Figure B.1 illustrates CDFs corresponding to the PDFs presented in Figure 6.1. Recall that Lemma 3.15 focused on just one point of the CDF, whereas Figure B.1(b) illustrates the complete CDFs. We observe that Figure B.1 appears to show a form of (partial) stochastic dominance of IR over QD for wealth outcomes below the mean $\mathcal{E}$ (see [112] for a definition and discussion).

[8] The annual average CPI-U index, which is based on inflation data for urban consumers, were used - see `http://www.bls.gov.cpi`

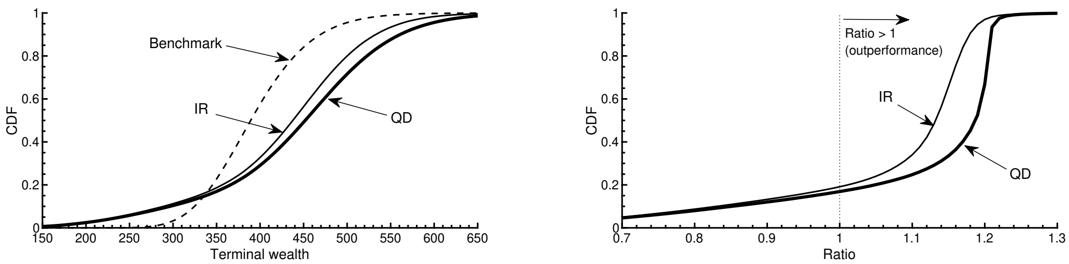
However, the situation changes when investment constraints are applied. This can be observed in Figures B.2 and B.3, which illustrate the corresponding CDFs to the PDFs presented in Figures 6.3 and 6.4 (Subsection 6.3). In this case, it appears that QD effectively achieves stochastic dominance over IR (and not just partial stochastic dominance for downside outcomes) regardless of whether wealth or the wealth ratio is considered.

From a practical perspective, Figures B.2 and B.3 show that the QD strategy has an 80% probability (out of sample) of outperforming the benchmark by about 100 bps per year. We remind the reader that this requires no stock picking ability, or use of exotic financial instruments, simply application of optimal control.



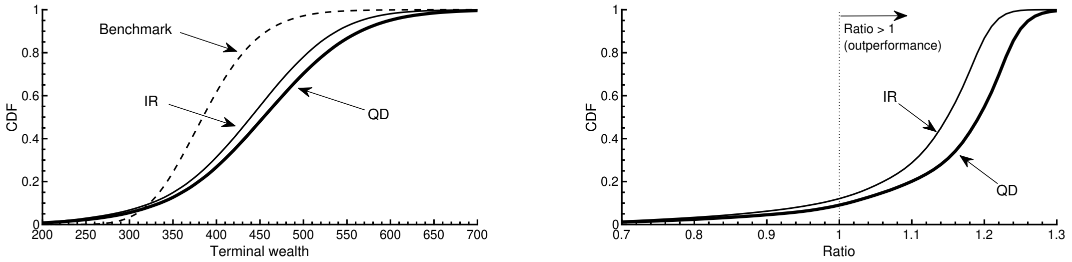
(a) CDFs of $\hat{W}(T)$, $W_j^{\mathcal{E}*}(T)$, $j \in \{ir, qd\}$ **(b)** CDFs of $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, $j \in \{ir, qd\}$

Figure B.1: Analytical solutions, no constraints, investor portfolio P0, benchmark BM0: Simulated CDFs of benchmark and investor's target terminal wealth $\hat{W}(T)$ and $W_j^{\mathcal{E}*}(T)$, respectively, as well as the ratio $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, for $j \in \{ir, qd\}$. 10^6 Monte Carlo simulations, $\mathcal{E} = 400$ in (6.1).



(a) CDFs of $\hat{W}(T)$, $W_j^{\mathcal{E}*}(T)$, $j \in \{ir, qd\}$ **(b)** CDFs of $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, $j \in \{ir, qd\}$

Figure B.2: Out-of-sample (testing) results for DS1 using annual rebalancing, numerical solutions, with constraints, investor portfolio P1, benchmark BM1: Simulated CDFs of benchmark and investor's target terminal wealth $\hat{W}(T)$ and $W_j^{\mathcal{E}*}(T)$, respectively, as well as the ratio $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, for $j \in \{ir, qd\}$.



(a) CDFs of $\hat{W}(T)$, $W_j^{\mathcal{E}*}(T)$, $j \in \{ir, qd\}$ **(b)** CDFs of $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, $j \in \{ir, qd\}$

Figure B.3: Out-of-sample (testing) results for DS3 using quarterly rebalancing, numerical solutions, with constraints, investor portfolio P1, benchmark BM1: Simulated CDFs of benchmark and investor's target terminal wealth $\hat{W}(T)$ and $W_j^{\mathcal{E}*}(T)$, respectively, as well as the ratio $W_j^{\mathcal{E}*}(T)/\hat{W}(T)$, for $j \in \{ir, qd\}$.

Table B.2 presents results for using investor portfolio P1 to outperform benchmark BM1 on data set DS, from which we conclude that the qualitative aspects of the comparative performance of the IR and QD-optimal strategies also hold on data set DS2.

Table B.2: Numerical solutions, with constraints, investor portfolio P1, benchmark B1, data set DS2, annual rebalancing: Training and testing results for mean terminal wealth $\mathcal{E} = 430$ ($\hat{\beta} \simeq 1.7\%$ in (6.1)) on the training data.

Quantity	Training data DS2 (1963:07 - 1999:12)					Testing data DS2 (2000:01 - 2010:12)				
	$\hat{W}(T)$	$W_j^{\mathcal{E}*}(T)$		$W_j^{\mathcal{E}*}(T)/\hat{W}(T)$		$\hat{W}(T)$	$W_j^{\mathcal{E}*}(T)$		$W_j^{\mathcal{E}*}(T)/\hat{W}(T)$	
	BM1	IR	QD	IR	QD	BM1	IR	QD	IR	QD
Mean	364	430	430	1.19	1.18	273	315	309	1.14	1.12
CExp 5%	212	249	235	1.04	1.05	172	104	110	0.53	0.55
5th pctile	235	286	268	1.11	1.12	187	135	143	0.64	0.67
Median	354	422	419	1.19	1.19	266	326	311	1.21	1.19
95th pctile	531	601	630	1.28	1.20	381	454	455	1.37	1.25
Prob. underp.				1.22%	1.05%				19.26%	15.64%

C. Neural network (NN) approach - additional details.

In this appendix, we discuss a number of additional details related to the neural network (NN) approach discussed in Section 5.

C.1. Implementation parameters and gradient descent algorithm.

The NN is trained with stochastic gradient descent using the Gadam algorithm of [54]. This combines the Adam algorithm ([71]) with tail iterate averaging for improved convergence properties and variance reduction ([101, 86, 87]). Numerical experiments showed that the default algorithm parameters of [71] performed well in our setting. Additionally, we used 64,000 stochastic gradient descent steps, together with a mini-batch size of 100 paths from the training data set Y on each gradient descent iteration. Numerical tests showed that results with this configuration were very stable and reliable; for example, essentially identical results are obtained each time the NN is trained independently on the same underlying data.

In terms of the structure of the NN, the minimal features were used (time, investor wealth, benchmark wealth) for illustrative purposes. As noted in Section 5, two hidden layers, each with $N_a + 2$ nodes, were found to capture sufficient complexity for both benchmark outperformance problems, while ensuring that stable results were obtained on the numerical solutions as well as the ground truth solutions (see Appendix C).

C.2. Ground truth results. To show that the numerical solutions obtained as described in Section 5 can converge under suitable conditions to the closed-form solutions as described in Section 3, we encounter the problem that the numerical solutions are explicitly constructed (via the NN output layer activation function) to enforce the desired investment constraints. While a different output layer activation function could be implemented, the treatment of trading in the case of insolvency (i.e when wealth crosses zero into the negative domain) needs to be carefully addressed in any numerical solution.

Instead of modifying the methodology used to obtain numerical solutions, we observe that if a relatively short time horizon (e.g. $T = 1$ year) is combined with a reasonable outperformance target (e.g. $\hat{\beta} \simeq 1.0\%$ in (6.1)), then the probability of insolvency is negligible, as is the need for leverage or short-selling in the closed-form solutions. This allows us to use the numerical solutions (with constraints) to approximate the closed-form solutions (no constraints), provided the underlying data is the same. We can therefore use a NN training data set based on simulated data with parameters as in Table B.1, and use the same data for the implementation of analytical solutions. The results, obtained using 10^6 Monte Carlo simulations, are illustrated in Table C.1. Investor portfolio P0 and benchmark BM0 are used, and we assume contributions are zero to avoid discrete approximation errors when comparing a continuous contribution rate to discrete contribution amounts made at rebalancing times. Table C.1 confirms that the numerical results using the NN approach recovers the analytical results as desired.

Table C.1: Ground truth comparison, investor portfolio P0, benchmark BM0, and data set DS0 used for NN training data: $w_0 = 100$, $q = q(t_n) = 0$, $T = 1$ year. Since BM0 results in an expected terminal wealth $\mathcal{K} = 104.20$, a value of $\mathcal{E} = 105.25$ implies $\hat{\beta} \simeq 1.0\%$. Analytical solutions based on 360 rebalancing events approximating continuous rebalancing. Numerical results are based on only 36 discrete rebalancing events to ensure that computation times remain reasonable.

Quantity	Analytical solutions: P0					Numerical solutions (using NN): P0				
	BM0	$W_j^{\mathcal{E}*}(T)$		$W_j^{\mathcal{E}*}(T)/\hat{W}(T)$		BM0	$W_j^{\mathcal{E}*}(T)$		$W_j^{\mathcal{E}*}(T)/\hat{W}(T)$	
	$\hat{W}(T)$	IR	QD	IR	QD	$\hat{W}(T)$	IR	QD	IR	QD
Mean	104.2	105.3	105.3	1.01	1.01	104.2	105.2	105.2	1.01	1.01
CExp 5%	85.6	80.1	80.2	0.93	0.93	85.6	80.2	80.2	0.93	0.93
5th pctile	90.7	87.4	87.4	0.96	0.96	90.7	87.2	87.2	0.96	0.96
Median	104.1	105.6	105.5	1.01	1.01	104.1	105.6	105.5	1.01	1.01
95th pctile	117.9	121.9	122.1	1.03	1.04	117.9	121.8	122.0	1.03	1.03

REFERENCES

[1] M. Aguilar and A. Custovic, *Portfolio optimization without optimization*, SSRN 3770361, (2021).

[2] A. Al-Aradi and S. Jaimungal, *Outperformance and tracking: dynamic asset allocation for active and passive portfolio management*, Applied Mathematical Finance, 25 (2018), pp. 268–294.

[3] A. Al-Aradi and S. Jaimungal, *Active and passive portfolio management with latent factors*, Quantitative Finance, 21 (2021), pp. 1437–1459.

[4] A. G. Alekseev and M. V. Sokolov, *Benchmark-based evaluation of portfolio performance: a characterization*, Annals of Finance, 12 (2016), pp. 409–440.

[5] A. Anarkulova, S. Cederburg, and M. S. O'Doherty, *Stocks for the long run? Evidence from a broad sample of developed markets*, Journal of Financial Economics, 143:1 (2022), pp. 409–433.

[6] A. Ang, *Asset management - A systematic approach to factor investing*, Oxford University Press, 2014.

[7] D. Applebaum, *Lévy processes and stochastic calculus*, Cambridge University Press, 2004.

[8] R. Badrania, D. P. Neghab, and M. Shafizadeh, *State-dependent stock selection in index tracing: a machine learning approach*, Financial Markets and Portfolio Management, 36 (2022), pp. 1–28.

[9] I. Bajeux-Besnainou, R. Portait, and G. Tergny, *Optimal portfolio allocations with tracking error volatility and stochastic hedging constraints*, Quantitative Finance, 13 (2013), pp. 1599–1612.

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

[11] S. Basak, A. Shapiro, and L. Tepla, *Risk management with benchmarking*, Management Science, 52 (2006), pp. 542–557.

[12] T. Bjork, *Arbitrage Theory in Continuous Time*, Oxford University Press, 2009.

[13] T. Bjork and A. Murgoci, *A theory of Markovian time-inconsistent stochastic control in discrete time*, Finance and Stochastics, 18 (2014), pp. 545–592.

[14] L. Bo, H. Liao, and X. Yu, *Optimal tracking portfolio with a ratcheting capital benchmark*, SIAM Journal on Control and Optimization, 59 (2021), pp. 2346–2380.

[15] A. Bolshakov and L. B. Chincarini, *Manager skill and portfolio size with respect to a benchmark*, European Financial Management, 26 (2020), pp. 176–197.

[16] E. Boyde, *ETFs: why tracking difference usually matters more than tracking error*, Financial Times, (2021). https://www.ft.com/content/44b08a6c-55d7-4841-8e8e-aad6451a4cc3.

[17] E. Boyde, *ETFs: why tracking difference usually matters more than tracking error*, Financial Times,

https://www.ft.com/content/44b08a6c-55d7-4841-8e8e-aad6451a4cc3 (2021).
[18] S. Browne, *Beating a moving target: optimal portfolio strategies for outperforming a stochastic benchmark*, Finance and Stochastics, 3 (1999), pp. 275–294.
[19] S. Browne, *Reaching goals by a deadline: Digital options and continuous-time active portfolio management*, Advances in Applied Probability, 31 (1999), pp. 551–577.
[20] S. Browne, *Risk-constrained dynamic active portfolio management*, Management Science, 46 (2000), pp. 1188–1199.
[21] Canadian Pension Plan, *Our investment strategy*. https://www.cppinvestments.com/the-fund/how-we-invest/our-investment-strategy, 2022.
[22] S. Cavaglia, L. Scott, K. Blay, and S. Hixon, *Multi-asset class factor premia: A strategic asset allocation perspective*, The Journal of Portfolio Management, 48:9 (2022), pp. 14–32.
[23] A. Charteris and K. McCullough, *Tracking error vs. tracking difference: does it matter?*, Investment Analysts Journal, 49:3 (2020), pp. 269–287.
[24] A. Charteris and K. McCullough, *Tracking error vs tracking difference: Does it matter?*, Investment Analysts Journal, 49 (2020), pp. 269–287.
[25] R. G. Clarke, S. Krase, and M. Statman, *Tracking errors, regret, and tactical asset allocation*, Journal of Portfolio Management, 20 (1994), pp. 16–24.
[26] P. Cogneau and V. Zakalmouline, *Block bootstrap methods and the choice of stocks for the long run*, Quantitative Finance, 13 (2013), pp. 1443–1457.
[27] T. F. Coleman, Y. Li, and J. Henniger, *Minimizing tracking error while restricting the number of assets*, Journal of Risk, 8 (2006), pp. 33–55.
[28] CPP Annual Report. https://cdn4.cppinvestments.com/wp-content/uploads/2022/06/CPP-Investments_F2022-Annual-Report-EN.pdf, 2022.
[29] D. Dang and P. Forsyth, *Better than pre-commitment mean-variance portfolio allocation strategies: A semi-self-financing Hamilton–Jacobi–Bellman equation approach*, European Journal of Operational Research, 250 (2016), pp. 827–841.
[30] M. Davis and S. Lleo, *Risk-sensitive benchmarked asset management*, Quantitative Finance, 8 (2008), pp. 415–426.
[31] M. Di Giacinto, S. Federico, F. Gozzi, and E. Vigna, *Constrained portfolio choices in the decumulation phase of a pension plan*, Carlo Alberto Notebooks 155, Collegio Carlo Alberto, (2010).
[32] P. Diamond and I. V. Fomenko, *Robustness of representations in multilayer feedforward neural networks*, Circuits Systems Signal Process, 12 (1993), pp. 211–221.
[33] H. Dichtl, W. Drobetz, and M. Wambach, *Testing rebalancing strategies for stock-bond portfolos across different asset allocations*, Applied Economics, 48 (2016), pp. 772–788.
[34] M. F. Dixon, I. Halperin, and P. Bilokon, *Machine learning in finance*, Springer International Publishing, 2020.
[35] B. Efron, *Bootstrap methods: Another look at the jacklife*, The Annals of Statistics, 7 (1979), pp. 1–26.
[36] ESMA, *Guidelines on ETFs and other UCITS issues: Guidelines for competent authorities and UCITS management companies*, European Securities and Markets Authority, ESMA/2014/937EN (2014).
[37] ETF.com, *Understanding tracking difference and tracking error*, ETF.com Education Center Article, https://www.etf.com/etf-education-center/etf-basics/understanding-tracking-difference-and-tracking-error (2021).
[38] E. Fama and K. French, *A five-factor asset pricing model*, Journal of Financial Economics, 116 (2015), pp. 1–22.
[39] E. F. Fama and K. R. French, *The cross-section of expected stock returns*, Journal of Finance, 47 (1992), pp. 427–465.
[40] E. F. Fama and K. R. French, *Common risk factors in the returns on stocks and bonds*, Journal of Financial Economics, 33 (1993), pp. 3–56.
[41] C. Filippi, G. Guastaroba, and M. Speranza, *A heuristic framework for the bi-objective enhanced*

index tracking problem, Omega, 65 (2016), pp. 122–137.

[42] P. Forsyth, *Bonds in a balanced portfolio: long term or short term?*, White paper, PWL capital, (2018), www.pwlcapital.com/bonds-balanced-portfolio-long-term-or-short-term/.

[43] P. Forsyth and K. Vetzal, *Dynamic mean variance asset allocation: Tests for robustness*, International Journal of Financial Engineering, 4:2 (2017). 1750021 (electronic).

[44] P. Forsyth and K. Vetzal, *Optimal asset allocation for retirement saving: Deterministic vs. time consistent adaptive strategies*, Applied Mathematical Finance, 26 (2019), pp. 1–37.

[45] P. Forsyth, K. Vetzal, and G. Westmacott, *Management of portfolio depletion risk through optimal life cycle asset allocation*, North American Actuarial Journal, 23 (2019), pp. 447–468.

[46] P. A. Forsyth, *A stochastic control approach to defined contribution plan decumulation: The nastiest, hardest problem in finance*, North American Actuarial Journal, 26:2 (2022), pp. 227–252.

[47] K.-I. Funahashi, *On the approximate realization of continuous mappings by neural networks*, Neural Networks, 2 (1989), pp. 183–189.

[48] Z. Gao, Y. Gao, Y. Hu, Z. Jiang, and J. Su, *Application of deep q-network in portfolio management*, in 2020 5th IEEE International Conference on Big Data Analytics (ICBDA), 2020, pp. 268–275, https://doi.org/10.1109/ICBDA49040.2020.9101333.

[49] W. Goetzmann, J. Ingersoll, M. Spiegel, and I. Welch, *Sharpening Sharpe ratios*, NBER Working Paper 9116, (2002).

[50] W. Goetzmann, J. Ingersoll, M. Spiegel, and I. Welch, *Manipulation and manipulation-proof performance measures*, Review of Financial Economics, 20 (2007), pp. 1503–1546.

[51] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*, MIT press, 2016.

[52] L. R. Gorman, S. G. Sapra, and R. A. Weigand, *The cross-sectional dispersion of stock returns, alpha and the information ratio*, Journal of Investing, 19 (2010), p. 113.

[53] Government Pension Fund Global, *Investment strategy*. https://www.nbim.no/en/the-fund/how-we-invest/investment-strategy/, 2022.

[54] D. Granziol, X. Wan, and S. Roberts, *Gadam: Combining adaptivity with iterate averaging gives greater generalisation*, arXiv 2003:01247, (2020).

[55] G. Guastaroba, R. Mansini, W. Ogryczak, and M. G. Speranza, *Enhanced index tracking with cvar-based ratio measures*, Annals of Operations Research, 292 (2020), pp. 883–931.

[56] J. Han and E. Weinan, *Deep learning approximation for stochastic control problems*, NIPS Deep Reinforcement Learning Workshop, (2016).

[57] M. Hassine and T. Roncalli, *Measuring performance of exchange-traded funds*, The Journal of Index Investing, 4 (2013), pp. 57–85.

[58] K. Hornik, *Approximation capabilities of multilayer feedforward networks*, Neural Networks, 4 (1991), pp. 251–257.

[59] K. Hornik, M. Stinchcombe, and H. White, *Multilayer feedforward networks are universal approximators*, Neural Networks, 2 (1989), pp. 359–366.

[60] M. Hougan, *Tracking difference: the perfect ETF metric*. https://www.etf.com/sections/blog/tracking-difference-perfect-etf-metric, 2015.

[61] J. Huang, Y. Li, and H. Yao, *Index tracking model, downside risk and non-parametric kernel estimation*, Journal of Economic Dynamics and Control, 92 (2018), pp. 103–128.

[62] C. Huré, H. Pham, A. Bachouch, and N. Langrené, *Deep neural networks algorithms for stochastic control problems on finite horizon: Convergence analysis*, SIAM Journal on Numerical Analysis, 59 (2021), pp. 525–557.

[63] R. Israelov and H. Tummala, *An alternative option to portfolio rebalancing*, The Journal of Derivatives, 25 (2018), pp. 1–27.

[64] C. K. Israelsen and G. F. Cogswell, *The error of tracking error*, Journal of Asset Management, 7 (2007), pp. 419–424.

[65] B. Johnson, H. Bioy, A. Kellet, and L. Davidson, *On the right track: measuring tracking efficiency*

in ETFs, The Journal of Index Investing, 4:3 (2013), pp. 35–41.
[66] B. Johnson, H. Bioy, A. Kellett, and L. Davidson, *On the right track: Measuring tracking efficiency in ETFs*, The Journal of Index Investing, 4 (2013), pp. 35–41.
[67] P. Jorion, *Portfolio optimization with tracking-error constraints*, Financial Analysts Journal, 59 (2003), pp. 70–82.
[68] A. K. Kashyap, N. Kovrijnykh, J. Li, and A. Pavlova, *The benchmark inclusion subsidy*, Journal of Financial Economics, 142 (2021), pp. 756–774.
[69] S. Kim, *Deep time series forecasting for enhanced index tracking*, Applied Economics, 53 (2021), pp. 1916–1934.
[70] S. Kim and S. Kim, *Index tracking through deep latent representation learning*, Quantitative Finance, 20 (2020), pp. 639–652.
[71] D. P. Kingma and J. L. Ba, *Adam: A method for stochastic optimization*, Published as a conference paper at ICLR 2015, (2015).
[72] R. Korn and C. Lindberg, *Portfolio optimization for an investor with a benchmark*, Decisions in Economics and Finance, 37 (2014), pp. 373–384.
[73] S. G. Kou, *A jump-diffusion model for option pricing*, Management Science, 48 (2002), pp. 1086–1101.
[74] S. G. Kou, *Jump-diffusion models for asset pricing in financial engineering*, in Handbooks in OR and MS, Vol. 15, J. Birge and V. Linetsky, eds., Elsevier B.V., 2008, ch. 2.
[75] Y. Kwak, J. Song, and H. Lee, *Neural network with fixed noise for index-tracking portfolio optimization*, Expert Systems With Applications, 183 (2021), pp. 1–11.
[76] C.-A. Lehalle and G. Simon, *Portfolio selection with active strategies: how long only constraints shape convictions*, Journal of Asset Management, 22 (2021), pp. 443–463.
[77] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken, *Multilayer feedforward networks with a nonpolynomial activation function can approximate any function*, Neural Networks, 6 (1993), pp. 861–867.
[78] D. Li and W.-L. Ng, *Optimal dynamic portfolio selection: multi period mean variance formulation*, Mathematical Finance, 10 (2000), pp. 387–406.
[79] H. Li, Q. Huang, and B. Wu, *Improving the naive diversification: an enhanced indexation approach*, Finance Research Letters, 39 (2021), pp. 1–6.
[80] X. Li and X. Y. Zhou, *Continuous time mean variance efficiency and the 80 per cent rule*, Annals of Applied Probability, 16 (2006), pp. 1751–1763.
[81] Y. Li and P. Forsyth, *A data-driven neural network approach to optimal asset allocation for target based defined contribution pension plans*, Insurance: Mathematics and Economics, 86 (2019), pp. 189–204.
[82] Z. Li, K. H. Tsang, and H. Y. Wong, *Lasso-based simulation for high-dimensional multi-period portfolio optimization*, IMA Journal of Management Mathematics, 31 (2020), pp. 257–280.
[83] A. Lim and B. Wong, *A benchmark approach to optimal asset allocation for insurers and pension funds*, Insurance: Mathematics and Economics, 46 (2010), pp. 317–327.
[84] G. Lucarelli and M. Borrotti, *A deep q-learning portfolio management framework for the cryptocurrency market*, Neural Computing and Applications, 32 (2020), pp. 17229–17244.
[85] R. Merton, *Option pricing when underlying stock returns are discontinuous*, Journal of Financial Economics, 3 (1976), pp. 125–144.
[86] N. Mucke, G. Neu, and L. Rosasco, *Beating SGD saturation with tail-averaging and minibatching*, 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), (2019).
[87] G. Neu and L. Rosasco, *Iterate averaging as regularization for stochastic gradient descent*, Proceedings of Machine Learning Research, 31st Annual Conference on Learning Theory, 75 (2018), pp. 1–21.
[88] C. Ni, Y. Li, P. Forsyth, and R. Carroll, *Optimal asset allocation for outperforming a stochastic benchmark target*, Quantitative Finance, 22:9 (2022), pp. 1595–1626.
[89] M. Nicolosi, F. Angelini, and S. Herzel, *Portfolio management with benchmark related incentives under mean reverting processes*, Annals of Operations Research, 266 (2018), pp. 373–394.

[90] Norges Bank, *Annual report.* https://www.nbim.no/en/publications/reports/2021/annual-report-2021/, 2021.
[91] G. Oderda, *Stochastic portfolio theory optimization and the origin of rule based investing*, Quantitative Finance, 15 (2015), pp. 1259–1266.
[92] B. Oksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Springer, 3rd ed., 2019.
[93] H. Ouyang, X. Zhang, and H. Yan, *Index tracking based on deep neural network*, Cognitive Systems Research, 57 (2019), pp. 107–114.
[94] H. Park, M. K. Sim, and D. G. Choi, *An intelligent financial portfolio trading strategy using deep q-learning*, Expert Systems with Applications, 158 (2020). Article 113573.
[95] S. Park, H. Song, and S. Lee, *Linear programming models for portfolio optimization using a benchmark*, The European Journal of Finance, 25 (2019), pp. 435–457.
[96] M.-S. Pastant, *Tracking difference - a must when comparing ETFs*, BNPParibas, https://investors-corner.bnpparibas-am.com/investing/tracking-difference-comparing-etfs (2018).
[97] S. M. Pesenti and S. Jaimungal, *Portfolio optimization within a Wasserstein ball*, (2021). SSRN 3744994.
[98] D. Politis and J. Romano, *The stationary bootstrap*, Journal of the American Statistical Association, 89 (1994), pp. 1303–1313.
[99] D. Politis, J. Romano, and M. Wolf, *Subsampling*, Springer-Verlag, 1999. Chapter 4: Subsampling for nonstationary time series.
[100] D. N. Politis, *The impact of bootstrap methods on time series analysis*, Statistical Science, 18:2 (2003), pp. 219–230.
[101] B. T. Polyak and A. B. Juditsky, *Acceleration of stochastic approximation by averaging*, SIAM Journal on Control and Optimization, 30 (1992), pp. 838–855.
[102] A. M. Reppen, H. M. Soner, and V. Tissot-Daguette, *Deep stochastic optimization in finance.* https://arxiv.org/abs/2205.04604, 2022.
[103] R. Roll, *A mean-variance analysis of tracking error*, Journal of Portfolio Management, Summer (1992), pp. 13–22.
[104] Y.-L. K. Samo and A. Vervuurt, *Stochastic portfolio theory: a machine learning perspective*, Working paper, ArXiv 1605.02654, (2016).
[105] L. Scott and S. Cavaglia, *A wealth management perspective on factor premia and the value of downside protection*, The Journal of Portfolio Management, 43:3 (2017), pp. 1–9.
[106] R. Sehgal and A. Mehra, *Enhanced indexing using weighted conditional value at risk*, Annals of Operations Research, 280 (2019), pp. 211–240.
[107] J. Simonian and A. Martirosyan, *Sharpe parity redux*, The Journal of Portfolio Management, 48:9 (2022), pp. 183–193.
[108] S. Sonoda and N. Murata, *Neural network with unbounded activation functions is universal approximator*, Applied and Computational Harmonic Analysis, 43 (2017), pp. 233–268.
[109] M. Strub, D. Li, and X. Cui, *An enhanced mean-variance framework for robo-advising applications.* SSRN 3302111, 2019.
[110] L. Tepla, *Optimal investment with minimum performance constraints*, Journal of Economic Dynamics and Control, 25 (2001), pp. 1629–1645.
[111] K. H. Tsang and H. Y. Wong, *Deep-learning solution to portfolio selection with serially dependent returns*, SIAM Journal on Financial Mathematics, 11 (2020), pp. 593–619.
[112] P. M. Van Staden, D. Dang, and P. Forsyth, *On the distribution of terminal wealth under dynamic mean-variance optimal investment strategies*, SIAM Journal on Financial Mathematics, 12 (2021), pp. 566–603.
[113] P. M. Van Staden, P. A. Forsyth, and Y. Li, *A data-driven neural network approach to dynamic factor investing with transaction costs*, Working paper, Cheriton School, University of Waterloo, (2021).

[114] Vanguard, *Understanding tracking difference and tracking error*, Vanguard White Paper, https://www.vanguard.com.hk/documents/understanding-td-and-te-en.pdf (2014).

[115] E. Vigna, *On efficiency of mean-variance based portfolio selection in defined contribution pension schemes*, Quantitative Finance, 14 (2014), pp. 237–258.

[116] J. Wang and P. Forsyth, *Numerical solution of the Hamilton-Jacobi-Bellman formulation for continuous time mean variance asset allocation*, Journal of Economic Dynamics and Control, 34 (2010), pp. 207–230.

[117] R. Wang and D. P. Foster, *What are the statistical limits of offline RL with linear function approximation?*, arXiv:2010.11895, (2020).

[118] D. D. Yao, S. Zhang, and X. Y. Zhou, *Tracking a financial benchmark using a few assets*, Operations Research, 54 (2006), pp. 232–246.

[119] Q. Zhang and Y. Gao, *Portfolio selection based on a benchmark process with dynamic value-at-risk constraints*, Journal of Computational and Applied Mathematics, 313 (2017), pp. 440–447.

[120] Y. Zhao, *A dynamic model of active portfolio management with benchmark orientation*, Journal of Banking and Finance, 31 (2007), pp. 3336–3356.

[121] X. Zhou and D. Li, *Continuous time mean variance portfolio selection: a stochastic LQ framework*, Applied Mathematics and Optimization, 42 (2000), pp. 19–33.