Dynamic optimal investment strategies for benchmark outperformance with widely-used performance metrics

Pieter M. van Staden∗ Peter A. Forsyth† Yuying Li‡

April 28, 2022

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. First, by employing standard assumptions such as no market frictions, we solve stochastic control problems for the optimal analytical (or closed-form) investment strategies, extending known results for the IR problem to the case where jumps are allowed in the underlying process dynamics and where contributions to the portfolio are allowed. The analytical solutions to the QD problem are novel, as are the analytical results comparing the resulting IR- and QD-optimal investment strategies. Second, employing a data-driven neural network approach to solve the control problems numerically without any assumptions regarding the underlying process dynamics, we compare the resulting investment strategies under more realistic assumptions, including the discrete rebalancing of the portfolio together with multiple investment constraints, such as leverage and short-selling restrictions. 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. This ultimately results in delivering superior out-of-sample benchmark outperformance for the QD-optimal investor, and demonstrates the effectiveness of targeting benchmark outperformance via the proposed QD objective.

Keywords: Asset allocation, portfolio optimization, benchmark outperformance, neural network

JEL classification: G11, C61

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 considerable expertise to bear on actively pursuing an investment strategy with the explicit goal of outperforming an appropriate pre-specified benchmark (Alekseev and Sokolov (2016); Kashyap et al. (2021); Korn and Lindberg (2014); Lehalle and Simon (2021); Zhao (2007)), 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” (Gorman et al. (2010)).

For example, S&P Global’s SPIVA 2020 research report (SPIVA (2020) and analysis by Coleman (2021)), which provides an assessment of the performance of active funds against their appropriate index benchmarks until December 2020, shows that an astonishing 86% of active domestic (US) equity funds underperformed their benchmarks over the last 20 years. In the case of active bond funds, between 82% and 97% of funds (depending on maturity and issuer type) underperformed their respective benchmarks over the last 15 years.

While the underperformance of benchmarks by active funds is of course not a recent phenomenon (see for example Ikenberry et al. (1998)), it seems to be enjoying substantial recent publicity in the popular investment literature and news articles (see for example Business Insider (2020); CNBC (2020b); New York Times (2021); Swedroe and Berk (2020)).

∗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
Participation in an active fund also typically costs the average investor about five times more than investing in an index fund with a passive investment strategy (CNBC, 2020a), where the aim of passive investment strategies is simply the replication of the performance of some benchmark. For investors, implementation of passive investment strategies are easy and affordable due to the widespread availability of low-cost exchange-traded funds (ETFs) aimed simply at benchmark replication. While single broad market indices like the S&P 500 are often used as benchmarks (Kashyap et al., 2021), plausible benchmarks also include (the returns of) a portfolio rebalanced according to some straightforward rebalancing rule known to yield satisfying performance historically, such as a constant proportion strategy like the 1/N rule (see for example Brightman et al., 2017; Israelov and Tumala, 2018; Li et al., 2021; Ni et al., 2022).

Unsurprisingly, investors appear to be taking note of these developments, since the total assets under management by passive US domestic equity funds have exceeded that of the corresponding active funds since around August 2018 (Bloomberg, 2021). As a result, the pressure has arguably never been greater on active portfolio managers to deliver on the promised outperformance.

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 Al-Aradi and Jaimungal, 2018; Basak et al., 2006; Browne, 1999a,b, 2000; Davis and Lleo, 2008; Filippi et al., 2016; Lim and Wong, 2010; Nicolosi et al., 2018; Oderda, 2015; Tepla, 2001; Yao et al., 2006; Zhang and Gao, 2017; Zhao, 2007), the question has taken on a new sense of urgency and remains an active area of research (for recent examples, see Aguilar and Custovic, 2021; Al-Aradi and Jaimungal, 2021; Bolshakov and Chincarini, 2020; Guastaroba et al., 2020; Li et al., 2021; Park et al., 2019; Pesenti and Jaimungal, 2021; Sehgal and Mehra, 2019). 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 Badrania et al., 2021; Kim, 2021; Kim and Kim, 2020; Kwak et al., 2021; Ni et al., 2022; Ouyang et al., 2019; Samo and Vervuurt, 2016).

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. To clarify, we briefly discuss the treatment of the benchmark outperformance problem in the relevant literature, before contrasting this with investment practice.

The objective functions used in the literature often include the use of explicit or implied utility functions (Al-Aradi and Jaimungal, 2018, 2021; Basak et al., 2006; Davis and Lleo, 2008; Lim and Wong, 2010; Nicolosi et al., 2018; Oderda, 2015; Tepla, 2001), with the use of log utility (of outperformance) appearing to be especially popular. However, many other considerations are often included (possibly as constraints or as objective functions in their own right), for example setting maximum and minimum limits on out-/under-performance (Basak et al., 2006), minimizing the expected time until outperformance subject to a constraint on the probability of underperformance (and some variations of this theme, see Browne, 1999a,b, 2000), and imposing penalties on underperforming the benchmark (Gaivoronski et al., 2005). In addition, fairly elaborate definitions of outperformance have been proposed for the sophisticated investor, involving for example the possibility of multiple benchmarks (Al-Aradi and Jaimungal, 2018, 2021) or constraints on the correlation or more generally the dependence structure with the benchmark (Bernard and Vanduffel, 2014; Pesenti and Jaimungal, 2021).

Furthermore, an important quantity of interest in many of the objective functions in the literature is the ratio of the active portfolio wealth to the benchmark wealth (Al-Aradi and Jaimungal, 2018, 2021; Browne, 1999a, 2000; Davis and Lleo, 2008; Lim and Wong, 2010; Nicolosi et al., 2018; Oderda, 2015). 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 implies that the resulting investment strategies are potentially of limited use to many institutional investors like defined contribution pension funds, where the modelling of contributions and withdrawals are critical aspects of the problem (Forsyth and Vetzal, 2019; Forsyth et al., 2019; Forsyth, 2021). As discussed below, we explicitly

---

1 It is also worth noting that some ETFs aim not just to replicate the performance of an index but to offer so-called “enhanced indexation” (see for example Filippi et al., 2016; Li et al., 2021), according to which the aims are a combination of some level of index tracking with index outperformance. For the purposes of this paper, we consider this as simply another example of an attempt at active portfolio management to outperform a given benchmark.

2 Note that we only consider dynamic investment strategies, arguably the only reasonable approach given the fairly long time horizon considered in this paper. Nevertheless, there is a large literature on applying Markowitz (1952)-inspired portfolio optimization models to derive portfolio weights for benchmark outperformance in single-period settings - see for example Filippi et al., 2016; Gaivoronski et al., 2005; Guastaroba et al., 2020; Huang et al., 2018; Jorion, 2003; Li et al., 2021; Reid, 1992.
consider contributions to the portfolio in this paper, leaving withdrawals for our future work.

Another quantity enjoying significant popularity in the objective functions considered in the literature is
the tracking error, which typically measures the standard deviation or average absolute deviation of the daily,
monthly or annual differences between the returns of the active portfolio and the returns of the benchmark (see
for example Clarke et al. (1994); Coleman et al. (2006); Huang et al. (2018); Jorion (2003); Koll (1992). It should
be emphasized that by measuring some statistic associated with a set of individual observations of return dif-
fferences, the tracking error does not include any cumulative information regarding the over/under-performance
of the portfolio vs. the benchmark over some time horizon. Therefore, an active portfolio underperforming the
benchmark consistently in each period (for example by similar ratios or amounts) might in fact have a very low
tracking error according the definition.

Despite this observation, tracking error minimization is not only employed as expected in benchmark replica-
tion settings, but also used in combination with additional objectives with the aim of obtaining outperformance
and/or other desirable properties such as improved downside risk (Ammann and Zimmermann (2001); Barro
and Canestrelli (2009); Beasley et al. (2003); Gaivoronski et al. (2005); Giusio (2017); Goel et al. (2018); Guas-
taroba et al. (2020); Huang et al. (2018); Jorion (2003); Kim and Kim (2020); Kwon and Wu (2017); Rossbach
and Karlow (2019); Sant’Anna et al. (2017); Sehgal and Mehra (2019); Strub and Baumann (2018); Vieira et al.
(2021); Wang et al. (2018); Xu et al. (2016)). The tracking error is also used by some investors for perfor-
mane measurement purposes, which is reasonable provided they are primarily concerned with performance
consistency rather than outperformance (Du and Overway (2021); Israelsen and Cogswell (2007); Johnson et al.
(2013; Vanguard (2014)).

Finally, we note that the literature is typically concerned with obtaining closed-form solutions to the specified
optimization problems, which necessarily assumes for example continuous trading, the possibility of unlimited
short-selling and leverage, and usually also the possibility of continuing trading in the event of insolvency.
Examples include Bajeux-Besnainou et al. (2013); Basak et al. (2006); Bo et al. (2021); Browne (1999a,b, 2000);
Davis and Lleo (2008); Lim and Wong (2010); Nicolosi et al. (2018); Oderda (2015); Tepla (2001); Yao et al.
(2006); Zhang and Gao (2017); Zhao (2007).

Of course, while the existing literature may be very valuable for some investors, active portfolio managers
arguably have goals that are simultaneously more mundane and more ambitious than the treatment offered in
the literature. On one hand, practical goals are more mundane, in the sense that outperformance is typically
measured by investors and institutions alike on a far more straightforward basis (discussed below) than the
aforementioned objective functions typically encountered in the literature. On the other hand, active portfolio
managers simultaneously also place more ambitious demands on the resulting investment strategies compared to
most of the literature, since they necessarily require dynamic investment strategies derived subject to realistic
investment constraints, such as discrete rebalancing, short-selling and leverage restrictions.

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

(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, and therefore provides a straightforward and intuitive measure of the “performance
gap” (Charteris and McCullough (2020)). 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, especially in the case of long-term investors (see for example Boyd (2021); ETF.com
2021); Hougan (2015); Pastant (2018); Vanguard (2014)). Its importance is also recognized by regulators
such as European Securities and Markets Authority, who requires its disclosure (ESMA (2014)).

Since the tracking error does not actually convey any information regarding the actual over-/under-
performance of the portfolio, it is perhaps not surprising that the relationship between the tracking
difference and tracking error has been shown to be very weak, resulting in significantly different rankings
of active portfolio managers (Charteris and McCullough (2020); Johnson et al. (2013)). This implies that
institutional investors concerned with minimizing the tracking error, such as many pension funds, may in
fact be sacrificing performance (Israelsen and Cogswell (2007)).

Despite these considerations, the tracking difference has received relatively scant attention in the academic
literature (see Bonelli (2015); Charteris and McCullough (2020); Johnson et al. (2013)). By explicitly
focusing on tracking difference in this paper using an intuitive objective function, we aim to address this
shortcoming.
(ii) Information ratio: In one-period portfolio optimization settings, the information ratio (IR) is simply defined as a Sharpe ratio calculated using the difference in returns of the active portfolio over the benchmark instead of just the returns of the active portfolio (see Treynor and Black (1973)). However, in dynamic (or multi-period) settings where the portfolio is rebalanced either continuously or at specific discrete time intervals, the IR is typically defined (see for example Bajeux-Besnainou et al. (2013)) 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. Since this paper is concerned with dynamic investment strategies for benchmark outperformance, we also make use of this latter definition.

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 (Bajeux-Besnainou et al. (2013); Bolshakov and Chincarini (2020); Hassine and Roncalli (2013); Israelsen and Cogswell (2007)), despite concerns that it could be manipulated (Goetzmann et al. (2002, 2007)). Its popularity is underscored by the fact that it is included in standard (single-period) portfolio optimization routines used in the industry (Kopman and Liu (2009)).

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 Bajeux-Besnainou et al. (2013); Goetzmann et al. (2002); Zhao (2007). Instead, IR maximization is often simply employed in single-period optimization settings (see for example Bolshakov and Chincarini (2020); Gaivoronski et al. (2005); Lai et al. (2011); Park et al. (2019); Roll (1992)).

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 two above-mentioned objectives, namely the IR and the tracking difference. While the IR objective is standard in the literature (see for example Bajeux-Besnainou et al. (2013)), 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 from the wealth of an elevated benchmark portfolio over a targeted time horizon. By formulating these objectives in terms of the wealth rather than for example a wealth ratio, our treatment explicitly considers the possibility of contributions to the portfolio, making the formulation potentially more attractive from the perspective of institutional investors and active portfolio managers.

- In order to gain a theoretical understanding of the behavior of the resulting optimal investment strategies, we first solve the problems analytically using standard assumptions, including parametric underlying dynamics, continuous portfolio rebalancing and no market frictions. We allow for any of the commonly-used jump diffusion models in finance to be used for the risky assets (Kou (2002); Merton (1976)), thereby extending the known results of Bajeux-Besnainou et al. (2013); Goetzmann et al. (2007); Zhao (2007) in the case of the IR objective to the setting where jumps are allowed in the risky asset processes. 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.

- We also present analytical results showing that a dynamic programming-based solution to these benchmark outperformance problems in the more realistic setting of discrete portfolio rebalancing would be unnecessarily high-dimensional. By this, we mean that any dynamic programming-based approach would require the solution of a high-dimensional performance criterion (i.e. an approximation to a conditional expectation) in order to obtain the object of key interest, namely the optimal control (investment strategy) which is comparatively low-dimensional. This implies that that solving these problems numerically using standard techniques (for example using numerical solutions to the corresponding PIDEs, or machine learning techniques relying on the dynamic programming principle to obtain the solution) would be very challenging, as the structure of these problems would make them exceptionally prone to well-known issues with dynamic programming in high dimensions such as error amplification over successive iterations.

- Next, in order to gain practical insights into the behavior of the resulting optimal dynamic investment strategies in a more realistic setting, we solve the problems numerically using a data-driven neural network (NN) approach. The key step in our numerical algorithm is to approximate the control directly using a NN, bypassing the problem of the approximation of conditional expectations. We note that this general idea was also used in Tsang and Wong (2020). However, in contrast with Tsang and Wong (2020), we
introduce time as a parameter directly in the NN, this ensuring (under certain assumptions) that the (limiting) investment control is a continuous function of time, which is a desirable practical requirement. The formulation of both problems are now expanded to include the following: (i) rebalancing the portfolio at discrete time intervals; (ii) periodic contributions to the portfolio; (iii) waiving the requirement to specify underlying parametric process dynamics, instead relying solely on empirical market data from 1963 until the end of 2020; (iv) imposing multiple realistic investment constraints such as short-selling and leverage restrictions. We also consider a 10-year investment time horizon, ensuring that the conclusions are relevant to the investor concerned with long-run benchmark outperformance.

• Comparing the results using IR- and QD-optimal investment strategies obtained numerically in our more realistic setting, we show not only how the closed-form comparison results do apply qualitatively to in-sample investment results, but also that 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 at certain points during the investment time period associated with the QD-optimal strategy ultimately results in superior out-of-sample benchmark outperformance. Our results therefore highlight the potential practical limitations of using analytical solutions available in the literature to reach qualitative conclusions regarding the behavior of optimal strategies in practice. Furthermore, the results demonstrate the effectiveness of targeting benchmark outperformance via the proposed QD objective.

The remainder of the paper is organized as follows. Section 2 presents the problem formulation. Section 3 and Section 4 discuss the analytical and numerical solutions of the problems, respectively. Finally, Section 5 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. No reference will be made to the portfolio rebalancing frequency or investment constraints, which will be addressed in subsequent sections.

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 assets available for investment. Let $\hat{p}_i(t, \hat{X}(t))$ 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 $\hat{X}(t)$ denotes the state of the system (or informally, the information) taken into account by the benchmark entity/investor in order to reach allocation decision $\hat{p}_i$. The vector $\hat{p}(t, \hat{X}(t)) = (\hat{p}_i(t, \hat{X}(t)) : i = 1, ..., N_a) \in \mathbb{R}^{N_a}$ denotes the asset allocation of the benchmark at time $t \in [t_0, T]$.

While the benchmark strategy could of course be deterministic, in which case we would simply have the function of time $t \rightarrow \hat{p}_i(t)$, we do not generally assume this to be the case when formulating and solving the problem. For example, in Section 3 where we consider feedback controls in a Markovian setting, we have $(t, \hat{W}(t)) \rightarrow \hat{p}_i(t, \hat{W}(t)), \text{ so that } \hat{X}(t) \text{ is simply the scalar } \hat{X}(t) = \hat{W}(t) \text{ at each } t \in [t_0, T]$. However, our numerical solution approach discussed in Section 4 allows for $\hat{X}(t)$ to be more general.

Similarly, let $p_i(t, X(t))$ 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 $X(t)$ denotes the information taken into account by the investor in making the asset allocation decision. As a concrete example, in Section 4 we consider the case where $X(t) = (W(t), \hat{W}(t), \hat{p}(t, \hat{W}(t)))$, but more general cases incorporating additional information in $X(t)$ are also allowed in Section 4. The vector $p(t, X(t)) = (p_i(t, X(t)) : i = 1, ..., N_a) \in \mathbb{R}^{N_a}$ denotes the asset allocation of the investor at time $t \in [t_0, T]$.

Define the set of rebalancing events $T \subseteq [t_0, T]$, where we have $T = [t_0, T]$ in the case of continuous rebalancing, and a strict (discrete) subset $T \subset [t_0, T]$ in the case of discrete rebalancing. The investor and
The standard deviation, variance, covariance and probability, respectively, calculated under the control initial time and wealth given by \( P \in A \) and using control implicit in this notation. Similarly, we will use expectation of some quantity taken with respect to a given initial wealth

We observe that definitions (2.1)-(2.2) implicitly assume that the investor and benchmark strategies invest in the same \( N_0 \) underlying assets. This is purely for convenience as this is relevant in the case of analytical solutions (Section 3), but it is not required in the case of the numerical solution approach discussed in Section 4.

At this point, we make no specific assumptions regarding the investment constraints underlying \( P \) or \( \hat{P} \).

However, in general terms, for the investor’s strategy \( P \), we define \( A \) as the set of admissible controls, and \( Z \) as the set of admissible values of each vector \( p(t, X(t)) \), so that \( P \in A \) if and only if \( P = \{ p(t, X(t)) : i = 1, \ldots, N_0 \} : t \in T \} \), \( \hat{P} \in A \) if and only if \( \hat{P} = \{ \hat{p}(t, \hat{X}(t)) : i = 1, \ldots, N_0 \} : t \in T \} \). Note that, \( Z \), and therefore by extension \( A \), encode the investment constraints faced by the investor, such as leverage constraints or short-selling restrictions. Throughout this paper, we make the realistic assumption that the investor may apply different investment constraints than the benchmark.

The following remark clarifies the information regarding the benchmark strategy that is assumed to be known by the investor.

**Remark 2.1.** (Information known about the benchmark strategy) Since we consider both the analytical and numerical solutions of the benchmark outperformance problems presented in this paper, it should be emphasized that in each of these settings, different assumptions are made regarding the information that is known by the investor regarding the benchmark strategy:

(i) In the case of the analytical solutions (Section 3), we assume that at each rebalancing time \( t \in T \), the investor can observe the corresponding instantaneous benchmark asset allocation vector, \( \hat{p}(t, \hat{X}(t)) \).

Note that the assumption that the investor has knowledge of the benchmark strategy is a fairly common assumption in the benchmark outperformance literature - see for example Bessler et al. (2017); Browne (2000); Davis and Lleo (2005); Kim and Kim (2020); Kwon and Wu (2017); Li et al. (2021); Yao et al. (2006). Considering concrete examples of many benchmarks actually used in the literature and in practice (Alekseev and Sokolov (2016); Basak et al. (2006); Bo et al. (2021); Israelov and Tummala (2018); Zhao (2007)), many popular benchmarks easily satisfy this requirement. Specific examples include constant proportion benchmark strategies combining stocks and bonds in some fixed proportions, a special case of which is the investor simply trying to “beat the market” (i.e. the benchmark strategy has all wealth invested in a broad market index at all times \( t \in T \)). Other popular investment strategies such as glide path strategies (discussed in for example Forsyth and Vetzal (2019)) also meet this requirement.

(ii) By contrast, in the case of the numerical solutions (Section 4), the neural network solution approach can learn the appropriate benchmark outperformance strategy based purely on the historical performance (returns) of the benchmark strategy \( \hat{P} \), so that no details regarding the benchmark asset allocation vector \( \hat{p}(t, \hat{X}(t)) \) is required. As a result, when we consider the benchmark outperformance problems subject to realistic investment constraints and solve these problems numerically, we also do not require the full specification of the complete benchmark asset allocation strategy \( P \), in contrast to the usual treatment in the literature where \( \hat{P} \) is assumed to be known (as noted above, see for example Bessler et al. (2017); Browne (2000); Davis and Lleo (2008); Kim and Kim (2020); Kwon and Wu (2017); Li et al. (2021); Yao et al. (2006)). Our methodology might therefore be of particular interest to portfolio managers or institutional investors who often wish to outperform competing institutions (Simões et al. (2018)), where the benchmark strategy itself may be proprietary while the benchmark returns are nevertheless observable.

Since the investor wishes to outperform the benchmark in some sense, we introduce two practical investment objectives to achieve this aim in the following subsections. In terms of notation, let \( E_{P}^{\tau_{0}, w_{0}}[\cdot] \) denote the expectation of some quantity taken with respect to a given initial wealth \( w_{0} = W(t_{0}) = W(t_{0}) \) at time \( t_{0} = 0 \), and using control \( P \in A \) over \( [t_{0}, T] \). The benchmark strategy \( \hat{P} \) that the investor wishes to outperform remains implicit in this notation. Similarly, we will use \( Stdev_{P}^{\tau_{0}, w_{0}}[\cdot] \), \( Var_{P}^{\tau_{0}, w_{0}}[\cdot] \), \( Cov_{P}^{\tau_{0}, w_{0}}[\cdot] \) and \( P_{P}^{\tau_{0}, w_{0}}[\cdot] \) to denote the standard deviation, variance, covariance and probability, respectively, calculated under the control \( P \) and initial time and wealth given by \( (t_{0}, w_{0}) \).
2.1 Information ratio: Problem \( IR(\gamma) \)

The first investment objective we consider involves maximizing the information ratio (IR), which in a dynamic setting is defined as (Bajeux-Besnainou et al. (2013); Goetzmann et al. (2002))

\[
IR_{t_0}^{\rho} = \frac{E_{P_\rho}^{t_0,w_0} \left[ W(T) - \hat{W}(T) \right]}{StdE_{P_\rho}^{t_0,w_0} \left[ W(T) - \hat{W}(T) \right]}.
\]  

(2.3)

As discussed in Bajeux-Besnainou et al. (2013), maximizing the IR (2.3) is achieved by solving the following mean-variance (MV) optimization problem with scalarization parameter \( \rho \),

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

(2.4)

where \( \rho \) effectively encodes the “risk appetite” of the investor for outperforming the benchmark.

In order to solve problem (2.4), we use the embedding technique of Li and Ng (2000); Zhou and Li (2000), which states that for any \( \rho > 0 \) and the associated control \( P_\rho^* \in A \) maximizing (2.4), there exists a value of an embedding parameter \( \gamma \) such that \( P_\rho^* \in A \) is also optimal for the following problem,

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

(2.5)

Note that (2.5) is formulated here only for the range \( \gamma > 0 \) in order to ensure that economically meaningful strategies for benchmark outperformance are obtained. For more information regarding the relationship between (2.4) and (2.5), please refer to Appendix A.1.

Associated with each value of \( \gamma \) in (2.5), we therefore obtain an investment strategy that maximizes the IR (2.3) for some value of the risk appetite for benchmark outperformance that is now encoded by \( \gamma \). As a result, we will subsequently refer to problem (2.5) simply as the IR (maximization) problem, abbreviated by \( IR(\gamma) \).

The exact relationship between \( \gamma \) in (2.5) and \( \rho \) in (2.4) 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 (2.5).

Remark 2.2. (Time-consistency of the IR(\( \gamma \))-optimal control) As elaborated in Forsyth et al. (2019); Li and Forsyth (2019), there appears to be some confusion in the literature regarding the time-consistency (or lack thereof) of the optimal controls associated with problems of the form (2.5). By analogy with dynamic MV optimization (see Basak and Chabakauri (2010); Bjork and Murgoci (2014)), the IR-optimal control for the embedding problem (2.5) is typically time-inconsistent from the perspective of the MV formulation (2.4). 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 \( \gamma \) in (2.5), the resulting IR(\( \gamma \))-optimal control is in fact a time-consistent control from the perspective of the quadratic objective (2.5), and is therefore clearly feasible as a trading strategy (Strub et al. (2019)). As discussed in Vigna (2014) and elaborated further below, a quadratic objective such as (2.5) also allows for a straightforward interpretation in terms of a “target” (in this case, \( \hat{W}(T) + \gamma \)), whereas very little guidance is offered in the literature regarding the selection of \( \rho \) in the case of a MV-type objectives of the form (2.4). 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 (2.5), and correspondingly formulate our results in terms of the embedding parameter \( \gamma \).

In contrast to the objective introduced in the next section, we emphasize that the preceding IR-related objective functions are not novel. In fact, the MV formulation (2.4) and its associated embedding formulation (2.5) clearly present reasonable objectives for benchmark outperformance in their own right, and are often used in both single-period and dynamic settings without necessarily recognizing the connections with IR maximization (see for example Gaivoronski et al. (2005); Jorion (2003); Perrin and Roncalli (2020); Roll (1992); Zhao (2007)). Here, we explicitly highlight the link to IR maximization due to the widespread popularity of using the IR for outperformance measurement (Bajeux-Besnainou et al. (2013); Bolshakov and Chincarini (2020); Hassine and Roncalli (2013); Israelsson and Cogswell (2007)) as noted in the Introduction.

The following additional observations regarding the IR objective (2.5) are relevant to the subsequent results:
(i) Since the embedding parameter $\gamma$ is a constant in problem $IR(\gamma)$, one way of interpreting (2.5) is that the investor wishes, where possible, to outperform the benchmark terminal wealth $\hat{W}(T)$ by a constant amount $\gamma$, regardless of the market conditions prevailing over $[t_0, T]$. The investor wishing to maximize the IR therefore 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) Re-arranging the IR problem (2.5), we observe that (2.5) is equivalent to
\[
\inf_{P \in A} \left\{ E_P^{t_0,w_0} \left[ \left( W(T) - \hat{W}(T) \right)^2 \right] - 2\gamma \cdot E_P^{t_0,w_0} [W(T)] \right\}, \quad \gamma > 0. \tag{2.6}
\]
In other words, the IR problem involves a trade-off between maximizing the performance of the investor’s portfolio as measured by $E_P^{t_0,w_0} [W(T)]$, while simultaneously minimizing the quadratic deviation of the investor’s wealth from the terminal wealth of the benchmark, $E_P^{t_0,w_0} \left[ \left( W(T) - \hat{W}(T) \right)^2 \right]$, with the embedding parameter $\gamma$ encoding the relative weighting of these objectives.

(iii) In Section 3 below, we show that under some conditions, the IR problem (2.5) is equivalent to the more intuitive one-sided quadratic objective,
\[
\inf_{P \in A} E_P^{t_0,w_0} \left[ \left( \min \left\{ W(T) - \left[ \hat{W}(T) + \gamma \right], 0 \right\} \right)^2 \right], \quad \gamma > 0, \tag{2.7}
\]
where only the shortfall of $W(T)$ below the elevated target $\hat{W}(T) + \gamma$ is penalized. While the equivalence between (2.5) and (2.7) can only be proven analytically under certain assumptions, numerical results nevertheless suggest that the results using (2.5) and (2.7) are indistinguishable even in more general cases where the conditions for analytical equivalence do not hold.

With these observations in mind, 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 $[t_0, T]$ (Charteris and McCullough (2020)).

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,

\[
(QD(\beta)) : \inf_{P \in A} E_P^{t_0,w_0} \left[ \left( W(T) - e^{\beta T} \hat{W}(T) \right)^2 \right], \quad \beta > 0. \tag{2.8}
\]

We will subsequently refer to problem (2.8) as the QD problem, and we make the following observations:

(i) The attractiveness of the formulation (2.8) 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 (Boyde 2021; Charteris and McCullough 2020; ESMA 2014; ETF.com 2021; Hougan 2015; Johnson et al. 2013; Pastant 2018; Vanguard 2014) is the central object of consideration. The simplicity of the objective function, however, does not imply a lack of sophistication of the resulting QD-optimal investment strategies, as will be evident from the results in the subsequent sections.

(ii) The parameter $\beta$ in the QD problem (2.8) has a conveniently practical interpretation as the annual outperformance spread that the investor targets for the tracking difference of the active portfolio. In particular, if the investor’s strategy achieves a terminal wealth of $W(T) = e^{\beta T} \hat{W}(T)$ under some ideal scenario, the investor has achieved an (average) annual outperformance return spread $\beta$ over the benchmark.

(iii) By formulating (2.8) 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 4).
(iv) Like the IR problem \(2.5\), the QD problem \(2.8\) 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 \(\gamma\) to \(\hat{W}(T)\) in \(2.5\)). 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. The implications of this observation on the resulting optimal investment strategies will be discussed in more detail in the subsequent sections.

(v) With the alternative formulation \(2.6\) of the IR problem in mind, we observe that the QD problem is equivalent to the following problem,

\[
\inf_{P \in \mathcal{A}} \left\{ E^{\mathcal{Q}, u_0}_{P, W_0} \left[ (W(T) - \hat{W}(T))^2 \right] - 2\hat{\gamma} \cdot E^{\mathcal{Q}, u_0}_{P, W_0} \left[ W(T) \cdot \hat{W}(T) \right] \right\}, \; \hat{\gamma} \equiv (e^{\beta T} - 1), \; \beta > 0. \tag{2.9}
\]

We therefore observe that the QD objective \(2.9\), like the IR objective (see \(2.6\)), involves minimizing the quadratic deviation of the investor’s wealth from the terminal wealth of the benchmark, \(E^{\mathcal{Q}, u_0}_{P, W_0} \left[ (W(T) - \hat{W}(T))^2 \right]\).

However, unlike in the case of the IR objective where this term is only combined with the simultaneous goal of maximizing the investor’s expected wealth \(E^{\mathcal{Q}, u_0}_{P, W_0} [W(T)]\), the QD objective combines this term with maximizing \(E^{\mathcal{Q}, u_0}_{P, W_0} [W(T) \cdot \hat{W}(T)]\). Of course, all else being equal, this still implies that the QD investor retains an incentive to obtain higher values of \(W(T)\). However, through the emphasis on covariance implied by \(2.9\), the role of maximizing the investor’s expected wealth is more nuanced than in the case of the IR problem \(2.6\).

(vi) As in the case of the IR problem (see \(2.7\)), we show in Section 3 that under some conditions, the QD problem \(2.8\) also admits the equivalent, and perhaps more intuitive, one-sided quadratic formulation,

\[
\inf_{P \in \mathcal{A}} E^{\mathcal{Q}, u_0}_{P, W_0} \left[ \left( \min \left\{ W(T) - e^{\beta T} \hat{W}(T) , 0 \right\} \right)^2 \right], \; \beta > 0, \tag{2.10}
\]

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

In summary, we have therefore discussed two fundamentally different, yet practical and easily-understood investment objectives for outperforming a given benchmark. The following sections are devoted to explore the practical implications of the preceding observations for the investor, using both closed-form solutions (where available) and numerical solutions of the IR and QD problems.

3 Analytical (closed-form) solutions

In this section, we investigate the analytical or closed-form solutions of the IR and QD problems introduced in Section 2. Of course, analytical solutions require the standard assumptions used in the literature, such as no market frictions and no investment constraints. However, we emphasize that these (unrealistic) assumptions, including any assumptions regarding the form of the underlying dynamics, are \(not\) required in Section 4 where we consider the problems subject to realistic investment constraints.

In this section, we consider \(N_a = N_r^* + 1\) assets evolving according to specified dynamics, consisting of one risk-free asset and \(N_r^*\) risky assets. In particular, we also allow for the modelling of jumps in the risky asset processes, in contrast to the existing literature on benchmark outperformance (see for example Al-Aradi and Jaimungal (2018); Basak et al. (2006); Browne (1999a,b, 2000); Davis and Lleo (2008); Lim and Wong (2010); Oderda (2015); Tepla (2001); Yao et al. (2006); Zhang and Gao (2017); Zhao (2007)).

In our analytical solutions, we explicitly allow for contributions to the portfolio, which only receives very limited treatment in the existing benchmark outperformance literature (Bo et al. (2021); Nicolosi et al. (2018)), while leaving further generalization of the results to the inclusion of withdrawals from the portfolio for our future work.

Assumption 3.1 summarizes the conditions under which the closed-form results are derived in this section. Note that the cash contribution is made to both the investor and the benchmark portfolios in order to ensure that a meaningful and fair performance comparison is obtained.
Assumption 3.1. (No market frictions, continuous rebalancing) We assume that trading continues in the event of insolvency; in other words, trading continues even if \( W(t) < 0 \) for some \( t \in [0, T] \). 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.

Recalling that we consider one risk-free asset and \( N_a^r \) risky assets, let \( \varrho(t, \mathbf{x}(t)) = (\varrho_1(t, \mathbf{x}(t)), \ldots, \varrho_{N_a^r}(t, \mathbf{x}(t))) \in \mathbb{R}^{N_a^r} \) and \( \hat{\varrho}(t, \hat{\mathbf{x}}(t)) = (\hat{\varrho}_1(t, \hat{\mathbf{x}}(t)), \ldots, \hat{\varrho}_{N_a^r}(t, \hat{\mathbf{x}}(t))) \in \mathbb{R}^{N_a^r} \) denote the proportional allocations of the investor and benchmark wealth, respectively, to each of the risky assets at time \( t \in [0, T] \). In more detail, for risky asset \( i \in \{1, \ldots, N_a^r\} \), \( \varrho_i(t, \mathbf{x}(t)) \) denotes the proportion of the investor’s wealth \( W(t) \) invested in asset \( i \) at time \( t \) given information \( \mathbf{x}(t) \), while \( \hat{\varrho}_i(t, \hat{\mathbf{x}}(t)) \) denotes the proportion of benchmark wealth \( \hat{W}(t) \) invested in asset \( i \) at time \( t \) given information \( \hat{\mathbf{x}}(t) \).

With reference to Remark 2.1, we make the following assumption regarding the benchmark strategy for the purposes of deriving the closed-form results of this section.

Assumption 3.2. (Analytical solutions: 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 \( \hat{\varrho}(t, \hat{\mathbf{x}}(t)) \) is the investor’s strategy at time \( t \in [0, T] \), and 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 \( \hat{\varrho}(t, \hat{W}(t)) \) at each \( t \in [0, T] \), so that the investor wishes to derive \( \varrho(t, \mathbf{x}(t)) = \varrho\left(t, W(t), \hat{W}(t), \hat{\varrho}(t, \hat{W}(t))\right) \), \( t \in [0, T] \), 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 emphasize that this assumption is not required in the case of the numerical solutions considered in Section 4 (see Assumption 4.1), but it should be noted that many benchmarks used in practice do in fact satisfy Assumption 3.2 (see Remark 2.1).

Combining definitions (2.1)–(2.2), Assumption 3.2 and incorporating the allocation to the risk-free asset, we therefore consider the following forms of the investor and benchmark strategies, respectively, in this section:

\[
P = \left\{ p(t, \mathbf{x}(t)) = \left( 1 - \sum_{i=1}^{N_a^r} \varrho_i(t, \mathbf{x}(t)), \varrho_1(t, \mathbf{x}(t)), \ldots, \varrho_{N_a^r}(t, \mathbf{x}(t)) \right) : t \in [0, T] \right\},
\]

\[
\hat{P} = \left\{ \hat{p}(t, \hat{W}(t)) = \left( 1 - \sum_{i=1}^{N_a^r} \hat{\varrho}_i(t, \hat{W}(t)), \hat{\varrho}_1(t, \hat{W}(t)), \ldots, \hat{\varrho}_{N_a^r}(t, \hat{W}(t)) \right) : t \in [0, T] \right\}, \tag{3.1}
\]

where \( \mathbf{x}(t) = (W(t), \hat{W}(t), \hat{\varrho}(t, \hat{W}(t))) \). Due to the forms of (3.1), in this section we will informally refer to the risky asset allocations \( \varrho(t, \mathbf{x}(t)) \) and \( \hat{\varrho}(t, \hat{W}(t)) \) as the investor and benchmark investment strategies, respectively, although the proper definition (3.1) should be kept in mind when extending the analysis to the multi-asset case presented in Section 4.

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 \( \varrho \).

\[
A_0 = \left\{ \varrho(t, \omega) = \varrho(t, w, \hat{\varrho}(t, w)) : [t_0, T] \times [0, T] \rightarrow \mathbb{R}^{N_a^r} \right\}, \tag{3.2}
\]

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

\[
\text{IR}(\gamma) : \inf_{\varrho \in A_0} E^{\omega,w}_{\varrho} \left[ (W(T) - \hat{W}(T) + \gamma)^2 \right], \quad \gamma > 0. \tag{3.3}
\]

\[
\text{QD}(\beta) : \inf_{\varrho \in A_0} E^{\omega,w}_{\varrho} \left[ (W(T) - e^{\beta T} \hat{W}(T))^2 \right], \quad \beta > 0. \tag{3.4}
\]

3.1 Asset and wealth dynamics

We now give a description of the underlying asset dynamics. First, with \( r > 0 \) denoting the continuously compounded risk-free rate, let \( S_0(t) \) denote the unit value of the risk-free asset at time \( t \in [0, T] \), with...
dS_0(t) = rS_0(t) \, dt. \tag{3.5}

Next, we consider the dynamics of the risky asset value vector \( S(t) = (S_i(t) : i = 1, \ldots, N^r_a) \), where the \( i \)-th component of \( S_i(t) \) denotes the unit value of the risky asset \( i \) at time \( t \in [t_0, T] \). Throughout this paper, we use the superscript \(^\top\) to denote the transpose.

Realistic modelling of \( S_i(t) \) requires the consideration of jumps in the risky asset process dynamics (see for example [Bjork, 2009]; [Zhou and Li, 2000]), and therefore so is the matrix \( \Lambda \) where \( S_i(t) \) satisfies
dynamics given by

where \( \xi \) is a random variable denoting the jump multiplier associated with the \( i \)-th risky asset with corresponding probability density function (pdf) \( f_{\xi_i}(\xi_i) \). The following expectations are associated with each \( \xi_i \).

For subsequent reference, we also define the random vector \( \xi = (\xi_i : i = 1, \ldots, N^r_a) \), as well as (non-random) vectors \( \kappa^{(1)} = (\kappa^{(1)}_i : i = 1, \ldots, N^r_a) \) and \( \kappa^{(2)} = (\kappa^{(2)}_i : i = 1, \ldots, N^r_a) \).

If a jump occurs in the dynamics of risky asset \( i \), 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 \( \lim_{t \to t^-} \psi(t) = \xi_i \cdot \lim_{t \to t^-} \psi(t) \) and \( \lim_{t \to t^+} \psi(t) = \xi_i \cdot \lim_{t \to t^+} \psi(t) \), respectively.

Next, we consider the dynamics of the risky asset value vector \( S(t) \), having the corresponding intensity \( \kappa^{(2)}_i \) for all \( i \). Let \( \pi = (\pi_i(t) : i = 1, \ldots, N^r_a) \) denote a vector of \( N^r_a \) independent Poisson processes, with each \( \pi_i(t) \) having the corresponding intensity \( \lambda_i \geq 0 \), and define \( \Lambda = (\lambda_i : i = 1, \ldots, N^r_a) \).

The vector of risky asset drift coefficients under the objective (or real-world) probability measure is denoted \( \mu = (\mu_i : i = 1, \ldots, N^r_a) \), where we make the economically reasonable assumption that the drift of risky asset \( i \) satisfies \( \mu_i > r \), for all \( i \). Let \( \sigma = (\sigma_{i,j})_{i,j=1,\ldots,N^r_a} \in \mathbb{R}^{N^r_a \times N^r_a} \) denote the volatility matrix, and for subsequent use we define two further matrices, \( \Sigma, \Lambda \in \mathbb{R}^{N^r_a \times N^r_a} \), as follows,

where \( \Lambda \) is the diagonal matrix with the Hadamard product \( \lambda \circ \kappa^{(2)} \) on the main diagonal. We make the standard assumption in this setting that the covariance matrix \( \Sigma = \sigma \sigma^\top \) is positive definite (see for example [Bjork, 2009]; [Zhou and Li, 2006]), and therefore so is the matrix \( \Sigma + \Lambda \).

Finally, let \( \mathbf{Z}(t) = (Z_i(t) : i = 1, \ldots, N^r_a) \) denote a standard \( N^r_a \)-dimensional Brownian motion. We assume that \( \xi_i, \pi_j(t) \) and \( Z_k(t) \) are mutually independent for all \( i,j,k \in \{1, \ldots, N^r_a\} \), with the covariance aspects of the risky asset dynamics being captured by the matrix \( \Sigma \) in (3.7).

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

where \( \xi_i^{(k)} \) are i.i.d. random variables with the same distribution as \( \xi_i \). Note that we can recover geometric Brownian motion (GBM) dynamics for \( S_i(t) \) by setting the intensity parameter \( \lambda_i \) to zero in (3.8).

To write the risky asset dynamics (3.8) in a more compact form for subsequent use, we define the vectors

where \( N_i \) is the Poisson random measure corresponding to the dynamics of \( S_i(t) \) in (3.8) - see for example
Using \( (3.9) \), and with \( \circ \) again denoting the Hadamard product, we have
\[
\frac{dS(t)}{S(t^-)} = \left( \mu - \lambda \circ \kappa^{(1)} \right) dt + \sigma \cdot dZ(t) + d\mathcal{N}(t).
\]  
(3.10)

For subsequent reference, we also define the following combinations of parameters of the underlying asset dynamics,
\[
\alpha = \left( \mu_i - r - \lambda_i \kappa_i^{(1)} : i = 1, \ldots, N^*_\alpha \right)^\top, \quad \hat{\mu} = \left( \mu_i - r : i = 1, \ldots, N^*_\alpha \right)^\top,
\]
(3.11)
as well as
\[
\eta = \hat{\mu}^\top \cdot (\Sigma + \Lambda)^{-1} \cdot \hat{\mu}.
\]
(3.12)

For investor and benchmark strategies \( \mathcal{P} \) and \( \hat{\mathcal{P}} \) of the form \( (3.1) \), together with the dynamics of the underlying assets \( (3.5) \) and \( (3.8) \), the investor and benchmark controlled wealth processes therefore have the following dynamics for \( t \in (t_0, T] \), respectively,
\[
dW(t) = \left\{ W(t^-) \cdot \left[ r + \alpha^\top \varrho(t, X(t)) \right] + q \right\} \cdot dt + W(t^-) \cdot \left( \varrho(t, X(t))^\top \right) \cdot \sigma \cdot dZ(t)
\]
\[
\quad + W(t^-) \cdot \left( \varrho(t, X(t))^\top \right) \cdot d\mathcal{N}(t),
\]
(3.13)
\[
d\tilde{W}(t) = \left\{ \tilde{W}(t^-) \cdot \left[ r + \alpha^\top \tilde{\varrho}(t, \tilde{W}(t)) \right] + q \right\} \cdot dt + \tilde{W}(t^-) \cdot \left( \tilde{\varrho}(t, \tilde{W}(t))^\top \right) \cdot \sigma \cdot dZ(t)
\]
\[
\quad + \tilde{W}(t^-) \cdot \left( \tilde{\varrho}(t, \tilde{W}(t))^\top \right) \cdot d\mathcal{N}(t),
\]
(3.14)
with \( W(t) = \tilde{W}(t) = w_0 \) and \( X(t) = \left( W(t), \tilde{W}(t), \varrho(t, \tilde{W}(t)) \right) \), 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 \( (3.3)-(3.4) \), subject to Assumption \( 3.1 \) Assumption \( 3.2 \) and wealth dynamics \( (3.13)-(3.14) \).

### 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 \( V_{ir}(y, t) : \mathbb{R} \times [t_0, T] \rightarrow \mathbb{R} \) and \( u^*_{ir} : \mathbb{R} \times [t_0, T] \rightarrow \mathbb{R}^{N^*_\alpha} \) with the following properties: (i) \( V_{ir} \) and \( u^*_{ir} \) are sufficiently smooth and solve the HJB partial integro-differential equation (PIDE) \( (3.15)-(3.16) \), and (ii) the function \( u^*_{ir} \) attains the pointwise supremum in \( (3.15) \).

\[
\frac{\partial V_{ir}}{\partial t} + \inf_{u \in \mathbb{R}^{N^*_\alpha}} \left\{ ry + \alpha^\top u \right\} = \frac{\partial V_{ir}}{\partial y} + \frac{1}{2} u^\top \Sigma u \cdot \frac{\partial^2 V_{ir}}{\partial y^2} - \left( \sum_{i=1}^{N^*_\alpha} \lambda_i \right) \cdot V_{ir}
\]
\[
\quad + \sum_{i=1}^{N^*_\alpha} \lambda_i \int_0^\infty V_{ir}(y + u_i (\xi_i - 1), t) \cdot f_\xi(\xi_i) \, d\xi_i \bigg\} = 0,
\]
(3.15)
\[
V_{ir}(y, T) = (y - \gamma)^2.
\]
(3.16)

Define the auxiliary process \( Y_{ir}(t) \) by
\[
Y_{ir}(t) := W(t) - \tilde{W}(t), \quad \forall t \in [t_0, T], \quad \text{with} \quad Y_{ir}(t_0) = y_0 = 0.
\]
(3.17)

Let the auxiliary control \( u(t) := u(Y_{ir}(t), t) \) be given by
\[
u(t) := \left( W(t) \cdot \varrho(t, X(t)) - \tilde{W}(t) \cdot \varrho(t, \tilde{W}(t)) \right), \quad \text{where} \quad X(t) = \left( W(t), \tilde{W}(t), \varrho(t, \tilde{W}(t)) \right).
\]
(3.18)

Let \( \mathcal{A}_{u,0} = \{ u(t) = u(y, t, x) : u \in \mathbb{R} \times [t_0, T] \rightarrow \mathbb{R}^{N^*_\alpha} \} \). Then under Assumption \( 3.1 \) Assumption \( 3.2 \) and wealth dynamics...
dynamics (3.13)-(3.14), $V_{ir}$ is the value function and $u^*_{ir}$ is the optimal control for the following control problem,

\[
\inf_{u \in \mathcal{A}_{u,0}} E^{t_0,y_0}_{\mathcal{F}} \left[ (Y_{ir} (T) - \gamma)^2 \right], \quad \gamma > 0.
\]  

(3.19)

**Proof.** Let $\tilde{N}$ denote the compensated Poisson random measure (see Oksendal and Sulem (2019)) associated with the $S_t$-dynamics in (3.8), and define the vector

\[
d\tilde{\mathbf{N}}(t) = \left( \int_0^\infty (\xi_i - 1) \tilde{N}_i (dt, d\xi_i) : i = 1, \ldots, N_t^* \right)^T.
\]  

(3.20)

It can then be shown that the auxiliary process $Y_{ir} (t)$ in (3.17) has the following dynamics in terms of auxiliary control $u (t)$ in (3.18),

\[
dY_{ir} (t) = \left[ rY_{ir} (t) + (u (t))^T \mu \right] dt + (u (t))^T \sigma \cdot dZ (t) + (u (t))^T \cdot d\tilde{\mathbf{N}} (t).
\]  

(3.21)

Using the dynamics (3.21), the proof applies standard techniques (see for example Applebaum 2004; Oksendal and Sulem 2019) to the analysis of problem (3.19), so that we omit further details. \hfill \Box

By solving the HJB PIDE (3.15)-(3.16), the following lemma reports the IR-optimal investment strategy.

**Lemma 3.4.** *(IR-optimal investment strategy)* Suppose that Assumption 3.2, Assumption 3.3 and wealth dynamics (3.13)-(3.14) are applicable. Then the optimal fraction of the investor’s wealth invested in risky asset $i \in \{1, \ldots, N_t^*\}$ for problem IR ($\gamma$) in (3.9) is given by the $i$th component of the vector $v^*_{ir} (t, X^*_{ir} (t))$, where

\[
W^*_{ir} (t) \cdot g^*_{ir} (t, X^*_{ir} (t)) = \left[ \gamma e^{-r(T-t)} - \left( W^*_{ir} (t) - \hat{W} (t) \right) \right] \cdot (\Sigma + \Lambda)^{-1} \hat{\mu} + \hat{W} (t) \cdot \hat{g} \left( t, \hat{W} (t) \right), \tag{3.22}
\]

with $W^*_{ir} (t)$ denoting the investor’s wealth process (3.13) under the IR-optimal control $g^*_{ir}$, and $X^*_{ir} (t) = (W^*_{ir} (t), W (t), \hat{g} \left( t, \hat{W} (t) \right))$.

**Proof.** Considering the form of terminal condition (3.16), we make the ansatz that $V_{ir} (y, t)$ is of the form $V_{ir} (y, t) = A_{ir} (t) y^2 + B_{ir} (t) y + C_{ir} (t)$ for unknown functions of time $A_{ir}, B_{ir}$ and $C_{ir}$. If this is indeed the case, then the pointwise supremum in (3.15) is attained by the auxiliary control $u^*_{ir} (t)$, where

\[
u^*_{ir} (t) = W^*_{ir} (t) \cdot g^*_{ir} (t, X^*_{ir} (t)) - \hat{W} (t) \cdot \hat{g} \left( t, \hat{W} (t) \right) = - \left[ x + \frac{B_{ir} (t)}{2A_{ir} (t)} \right] \cdot (\Sigma + \Lambda)^{-1} \hat{\mu}. \tag{3.23}
\]

The substitution of $V_{ir}$ and $u^*_{ir}$ into (3.15)-(3.16) 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 $\eta$ is given by (3.12). Substitution into (3.22) and simplification results in (3.22). \hfill \Box

It is noteworthy that the IR-optimal control $g^*_{ir} (t, X^*_{ir} (t))$ as per Lemma 3.4 only depends on the instantaneous benchmark allocation $\hat{g} \left( t, \hat{W} (t) \right)$ at time $t$, and not on the future or the past of the benchmark investment strategy.

We also highlight that the contribution rate $q$ does not appear in the solution (3.22), which follows from the fortunate cancellation of terms in the auxiliary process $Y_{ir} (t)$. Specifically, we recall that to ensure a fair comparison of strategies, equal contributions are made to both the investor and benchmark portfolios (see wealth dynamics (3.13)-(3.14)). From the definition of $Y_{ir} (t)$ in (3.17) as the difference in wealth values, the contribution terms cancel, so that we obtain dynamics (3.21) which contains no reference to $q$.

The following lemma reports the optimal IR that the investor can attain by implementing the IR-optimal control $g^*_{ir}$.

**Lemma 3.5.** *(IR: moments of $W^*_{ir} (T) - \hat{W} (T)$ and optimal IR)* Assume that Assumption 3.2, Assumption 3.3 and wealth dynamics (3.13)-(3.14) hold, with $t_0 = 0$. Implementing the IR-optimal control (3.22) gives

\[
E^{t_0,y_0}_{\mathcal{F}} \left[ W^*_{ir} (T) - \hat{W} (T) \right] = \gamma (1 - e^{-\eta T}), \quad V a r^{t_0,y_0}_{\mathcal{F}} \left[ W^*_{ir} (T) - \hat{W} (T) \right] = \gamma^2 e^{-2\eta T} (e^{\eta T} - 1), \tag{3.24}
\]

so that the IR investor obtains an optimal information ratio (2.3) of

\[
IR^{t_0,y_0}_{\mathcal{F}} = (e^{\eta T} - 1)^{1/2}, \tag{3.25}
\]
where η is given by (3.12).

Proof. After substituting the optimal control (3.22) into the dynamics of \( Y_{ir} (t) \) in (3.21), we obtain the resulting auxiliary dynamics under the IR-optimal control, in other words \( Y_{ir}^{\star} (t) := W_{ir}^{\star} (t) - \tilde{W} (t) \). Standard techniques (Oksendal and Sulem (2019)) give results (3.24), so that (3.25) follows from the definition (2.3).

Lemma 3.5 extends the known information ratio results of [Goetzmann et al. (2007)] 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 \( \eta \) in (3.12) reduces to \( \eta = (\mu_1 - r)^2 / \sigma_1^2 \), so that the optimal IR (3.25) reduces to the result reported in [Goetzmann et al. (2002, 2007)]. Note that the optimal IR (3.25) does not depend on the initial wealth \( w_0 \), since the IR is defined in terms of the wealth difference (see (2.3)).

The following lemma presents an important property of the IR-optimal strategy (3.22), whereby the IR-optimal investor will simply match the benchmark in terms of the amounts invested in the risky assets once sufficient outperformance can be assured.

**Lemma 3.6.** (IR: Matching the benchmark risky asset amounts) Given Assumption 3.1, Assumption 3.2 and wealth dynamics (3.13)-(3.14), suppose that at some time \( t \in (t_0, T) \), the IR-optimal investor observes a wealth value \( W_{ir}^{\star} (t) \) of

\[
W_{ir}^{\star} (t) = \gamma e^{-r(T-t)} + \tilde{W} (t).
\]

Then for the remainder of the investment time horizon \( t \in [t, T] \), the IR-optimal investor (using strategy (3.22)) will simply match the benchmark strategy in terms of the amounts invested in the risky assets. In other words,

\[
W_{ir}^{\star} (t) \cdot \tilde{g}_{ir}^{\star} (t, X_{ir}^{\star} (t)) = \tilde{W} (t) \cdot \tilde{\varphi} (t, \tilde{W} (t)), \quad \forall t \in [t, T].
\]

Proof. Given the form of (3.22), the assertion is obvious when \( t = \bar{t} \). To show that (3.27) also holds for \( t > \bar{t} \), we observe that combining (3.22) and (3.21) imply that the auxiliary process \( Q_{ir}^{\star} (t) := \gamma e^{-r(T-t)} - [W_{ir}^{\star} (t) - \tilde{W} (t)] \)

has dynamics given by

\[
\frac{dQ_{ir}^{\star} (t)}{Q_{ir}^{\star} (t^-)} = (r - \eta) \cdot dt - \mu^T (\Sigma + \Lambda)^{-1} \sigma \cdot dZ (t) - \mu^T (\Sigma + \Lambda)^{-1} \cdot d\mathcal{N} (t),
\]

with \( Q_{ir}^{\star} (\bar{t}) = 0 \). Since \( Q_{ir}^{\star} (t) \equiv 0 \) for \( t > \bar{t} \), (3.22) reduces to (3.27).

Note that Lemma 3.6 does not imply that the investor and benchmark strategies are equal (in the sense of definition (3.1)). Specifically, if the condition (3.26) is satisfied at some \( t \in (t_0, T) \), then dynamics (3.28) imply that \( W_{ir}^{\star} (t) > \tilde{W} (t) \), \( \forall t \in [t, T] \), and therefore (3.27) implies that \( g_{ir}^{\star} (t, X_{ir}^{\star} (t)) \) and \( \tilde{\varphi} (t, \tilde{W} (t)) \) will not be identical at any \( t \in [t, T] \).

While Lemma 3.6 allows for jumps in the risky asset processes, Lemma 3.7 below reports that condition (3.26) is in fact never satisfied in the special case when there are no jumps in the risky asset processes, with the implication that equivalence of problems (2.5) and (2.7) can be established analytically.

**Lemma 3.7.** (IR: equivalence with only penalizing underperformance) If Assumption 3.1, Assumption 3.2 and wealth dynamics (3.13)-(3.14) apply with no jumps (i.e. \( \lambda \in \mathbb{R}^{N_\lambda} \)) in the risky asset processes (3.10), then

\[
W_{ir}^{\star} (t) < \gamma e^{-r(T-t)} + \tilde{W} (t), \quad \forall t \in [t_0, T].
\]

As a result, in this case the IR optimization problem (2.5) 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. The equivalence assertion follows from the results of [Di Giacinto et al. (2010)], provided that (3.29) holds. Since in the case of no jumps, \( Q_{ir}^{\star} (t) \) in (3.28) reduces to a GBM with initial value \( Q_{ir}^{\star} (t_0) = \gamma e^{-r(T-t_0)} > 0 \), we have \( Q_{ir}^{\star} (t) > 0 \) for all \( t \in [t_0, T] \), which is (3.29).
An intuitive explanation as to why \((3.29)\) does not hold when the risky asset dynamics involve jumps, is that the resulting wealth dynamics with jumps \((3.13)\) can jump over the “barrier” (upper bound) as defined by \((3.29)\), whereas this cannot happen in the case of diffusion processes (no jumps).

In general, even if the assumptions of this section are violated, it should be emphasized that both Lemma 3.6 and the more restrictive Lemma 3.7 provide very valuable intuition for understanding the behavior of the IR-optimal investment strategies, which we will demonstrate in Section 5.

The following lemma shows that if we apply the assumption of no jumps as in Lemma 3.7, 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 underperformance, since it directly expresses a key quantity of concern for the active investor.

**Lemma 3.8.** \((IR: \text{probability of underperformance})\) If Assumptions 3.1 and 3.2 hold, and wealth dynamics \((3.13)-(3.14)\) apply with no jumps (i.e., \(\lambda = 0 \in \mathbb{R}^{N_x}\)) in the risky asset processes \((3.10)\), the probability of the IR-optimal wealth falling below the benchmark wealth at any \(t \in (t_0, T]\) is given by

\[
P_{\theta_{ir}}^{\nu_{ir}} \left[ W_{ir}^* (t) \leq \hat{W} (t) \right] = \Phi \left( -\frac{3}{2} \sqrt{\eta t} \right), \quad \forall t \in (t_0, T],
\]

where \(\Phi\) denotes the standard normal cumulative distribution function (CDF), and \(\eta\) is as defined in \((3.12)\).

**Proof.** Since it is assumed that there are no jumps in the risky asset dynamics, note that \((3.12)\) reduces to \(\eta = \bar{\mu}^T \Sigma^{-1} \mu\). Furthermore, as noted in the proof of Lemma 3.7 in the case of no jumps the dynamics of \(Q_{ir}^* (t)\) in \((3.28)\) is a GBM, with \((3.30)\) following from the relationship \(P_{\theta_{ir}}^{\nu_{ir}} \left[ W_{ir}^* (t) \leq \hat{W} (t) \right] = P_{\theta_{ir}}^{\nu_{ir}} \left[ Q_{ir}^* (t) \geq \gamma e^{-r(T-t)} \right]\).

While the result of Lemma 3.8 does not depend on contributions to the portfolio (which is to be expected, given the preceding results such as \((3.22)\) and the associated discussion), it is nevertheless remarkable that under the stated assumptions, the IR-optimal probability of underperformance \((3.30)\) does not depend on the value of \(\gamma\) nor on the specific form of the benchmark strategy other than Assumption 3.2. Even if the assumptions of Lemma 3.8 are violated, for example when multiple realistic investment constraints are applicable, the numerical results in Section 5 show that the dependence of the IR-optimal probability of underperformance on the value of \(\gamma\) remains very weak.

Although it is a quantity of relevance to the subsequent results, we do not report any analytical results for the IR-optimal probability of insolvency, defined for our purposes (see Assumption 3.1) simply as \(P_{\theta_{ir}}^{\nu_{ir}} \left[ W_{ir}^* (t) \leq 0 \right]\). The reason for the absence of a closed-form expression is that this probability remains analytically intractable even if we introduce restrictive assumptions such as those of Lemma 3.8. However, it is straightforward to assess the IR-optimal probability of insolvency numerically using the (joint) Monte Carlo simulation of the \(Q_{ir}^* (t)\) and \(\hat{W} (t)\) dynamics \((3.28)\) and \((3.14)\), since

\[
P_{\theta_{ir}}^{\nu_{ir}} \left[ W_{ir}^* (t) \leq 0 \right] = P_{\theta_{ir}}^{\nu_{ir}} \left[ Q_{ir}^* (t) - \hat{W} (t) \geq \gamma e^{-r(T-t)} \right], \quad t \in (t_0,T].
\]

### 3.3 Analytical solution: \(QD(\beta)\) problem

We now discuss the closed-form solution of the tracking difference problem using our proposed objective function \((3.28)\). The following verification theorem reports the HJB equation satisfied in the case of the QD problem \((3.34)\).

**Theorem 3.9.** \((QD \text{ problem: Verification theorem})\) 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 \(v_{qd}(y,t) : \mathbb{R} \times [t_0,T] \to \mathbb{R}^{N_x}\) with the following two properties. (i) \(V_{qd}\) and \(v_{qd}\) are sufficiently smooth and solve the HJB PIDE \((3.32)-(3.33)\), and (ii) the function \(v_{qd}(y,t)\) attains the pointwise supremum in \((3.32)\).

\(^{4}\)We do not offer a conjecture as to why \((3.29)\) does not depend on \(\gamma\). However, it is worth noting that this quantity corresponds to simply one point of the CDF \(P_{\theta_{ir}}^{\nu_{ir}} \left[ W_{ir}^* (t)/\hat{W} (t) \leq k \right]\), and the moments and tail behavior of this CDF is indeed significantly affected by the choice of \(\gamma\).
Define the auxiliary process $V_{qd}(t)$ by

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

Let the auxiliary control $v(t) := v(Y_q(t), t; \xi(t))$ be given by

$$
v(t) := W(t) \cdot \phi(t, \xi(t)) - e^{\gamma T} \hat{W}(t) \cdot \tilde{\phi} \left(t, \hat{W}(t)\right), \quad \text{where} \quad \xi(t) = \left(W(t), \hat{W}(t), \tilde{\phi} \left(t, \hat{W}(t)\right)\right).
$$

Proof. It can be shown that the dynamics of the auxiliary process $Y_{qd}(t)$ defined in (3.34) can be written in terms of the auxiliary control $v(t)$, defined in (3.35), as

$$
dY_{qd}(t) = \left[ h_{\beta}(t) - r \cdot (h_{\beta}(t) - Y_{qd}(t)) + (v(t))^T \bar{\mu} \right] dt + (v(t))^T \sigma \cdot dZ(t) + (v(t^-))^T \cdot d\bar{N}(t).
$$

Here, for a fixed value of the parameter $\beta$ and the contribution rate $q$, we define $h_{\beta}(t)$ as the following function of time,

$$
h_{\beta}(t) := \left(e^{\gamma T} - 1\right) \int_0^T q e^{-r(T-t)} dt = \frac{q}{r} \left(e^{\gamma T} - 1\right) \left(1 - e^{-r(T-t)}\right), \quad t \in [t_0, T],
$$

with $h_{\beta}'(t) = \frac{q}{r} h_{\beta}(t)$. As in the case of Theorem 3.3, the results of Theorem 3.9 then follows from the application of standard techniques (Oksendal and Sulem (2019)).

Solving the HJB PIDE (3.32)-(3.33), we obtain the QD-optimal control as reported by the following lemma.

Lemma 3.10. (QD-optimal control) Suppose that Assumption 3.1, Assumption 3.2 and wealth dynamics (3.13)-(3.14) are applicable. Then the optimal fraction of the investor’s wealth invested in risky asset $i \in \{1, ..., N_a\}$ for problem QD ($\beta$) in (3.4) is given by the $i$th component of the vector $\mathbf{g}_{qd}(t, X_{qd}(t))$, where

$$
W^*_q(t) \cdot \mathbf{g}_{qd}(t, X_{qd}(t)) = \left[ h_{\beta}(t) - \left( W^*_q(t) - e^{\gamma T} \hat{W}(t) \right) \right] \cdot \left( \Sigma + \Lambda \right)^{-1} \bar{\mu} + e^{\gamma T} \hat{W}(t) \cdot \tilde{\phi} \left(t, \hat{W}(t)\right),
$$

with $W^*_q(t)$ denoting the investor’s wealth process (3.13) under the QD-optimal control $\mathbf{g}_{qd}$, and $X_{qd}(t) = \left(W^*_q(t), \hat{W}(t), \tilde{\phi} \left(t, \hat{W}(t)\right)\right)$.

Proof. As in the case of Lemma 3.4, the terminal condition (3.33) 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 (3.32) is attained by the auxiliary control $v_{qd}(t)$ with a qualitatively similar form in terms of $(y, t)$ as the result reported in (3.23). Substituting $V_{qd}$ and $v_{qd}$ into (3.32)-(3.33) yields ODEs for $A_{qd}, B_{qd}$ and $C_{qd}$, which are solved to obtain $A_{qd}(t) = e^{(2r-\eta)(T-t)}$ and

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

where $\eta$ is given by (3.12). The necessary substitution and simplification yields (3.39).
Lemma 3.10 shows that as in the case of the IR-optimal control (see Lemma 3.4) the QD-optimal control \( \tilde{q}_d(t, X^*_d(t)) \) also only depends on the instantaneous benchmark allocation \( \tilde{q}(t, W(t)) \) and not on its past or future. However, in contrast to the IR-optimal control (3.22), the contribution rate \( q \) does affect the QD-optimal control (3.39) through the term \( h_\beta(t) \) (3.38).

The following lemma reports the first two moments of the difference between the investor’s wealth and that of the elevated benchmark, quantities which are subsequently useful when comparing investment outcomes.

**Lemma 3.11.** (QD: moments of \( W^*_d(T) - e^{\beta T} \hat{W}(T) \)) Assume that Assumption 3.1, Assumption 3.2 and wealth dynamics (3.13)-(3.14) hold, with \( t_0 = 0 \). Implementing the QD-optimal control (3.39) gives

\[
\begin{align*}
E_{W_0} & [ \hat{W}_d(T) - e^{\beta T} \hat{W}(T) ] = - [ h_\beta(0) + (e^{\beta T} - 1) w_0 ] e^{(r - \eta)T}, \\
Var_{W_0} & [ \hat{W}_d(T) - e^{\beta T} \hat{W}(T) ] = [ h_\beta(0) + (e^{\beta T} - 1) w_0 ]^2 e^{2(r - \eta)T} (e^{\beta T} - 1).
\end{align*}
\]

Proof. Substituting (3.39) in (3.37) gives the dynamics of \( Y_q(T) \). Standard techniques (Oksendal and Sulem (2019)) give results (3.41)-(3.42).

The following lemma shows that once sufficient outperformance can be assured, the QD-optimal amount in each risky asset will agree with the corresponding amount invested in the same asset by benchmark multiplied by the constant scaling factor \( e^{\beta T} \).

**Lemma 3.12.** (QD: Matching the elevated benchmark risky asset amount) Given Assumption 3.1, Assumption 3.2 and wealth dynamics (3.13)-(3.14), suppose that at some time \( T \in (t_0, T] \), the QD-optimal investor observes a wealth value \( W^*_d(T) \) satisfying

\[
W^*_d(T) = e^{\beta T} \hat{W}(T) + h_\beta(T).
\]

Then for the remainder of the investment time horizon \( t \in [T, T] \), the QD-optimal investor (using strategy (3.39)) will invest the following amounts in the risky assets,

\[
W^*_d(t) \cdot \tilde{q}_d(t, X^*_d(t)) = e^{\beta T} \cdot \hat{W}(t) \cdot \tilde{q}(t, W(t)), \quad \forall t \in [T, T].
\]

Proof. Substituting (3.38) into (3.43), note that condition (3.43) can equivalently be written as

\[
W^*_d(T) + \int_T^T q e^{-r(T-z)} dz = e^{\beta T} \cdot \hat{W}(T) + \int_T^T q e^{-(r - \eta)T} dz,
\]

which provides intuition as to why result (3.44) should hold. The proof proceeds along the same lines as in the case of Lemma 3.6 except that (3.44) can be established using the properties of the auxiliary process

\[
Q^*_d(t) = h_\beta(t) - [ W^*_d(t) - e^{\beta T} \hat{W}(t) ],
\]

which has dynamics that are formally the same as the dynamics of \( Q^*_r \) in (3.28).

By analogy with Lemma 3.7, the following lemma establishes some conditions under which the equivalence of problems (2.8) and (2.10) can be established analytically.

**Lemma 3.13.** (QD: equivalence with only penalizing underperformance) If Assumption 3.1, Assumption 3.2 and wealth dynamics (3.13)-(3.14) apply with no jumps (i.e. \( \lambda = 0 \in \mathbb{R}^{N^2} \)) in the risky asset processes (3.10), then

\[
W^*_d(t) < h_\beta(t) + e^{\beta T} \hat{W}(t), \quad \forall t \in [t_0, T].
\]

As a result, in this case the QD optimization problem (2.8) is equivalent to the one-sided quadratic problem (2.10), where only the underperformance of the investor’s portfolio (compared to the elevated benchmark) is penalized.

Proof. The proof proceeds along the same lines as for Lemma 3.7 but follows from analyzing the properties of \( Q^*_d \) in (3.46) with \( \lambda = 0 \).
As in the case of the IR problem, Lemma 3.12 and Lemma 3.13 provides 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 \( \hat{\varrho}(t, \tilde{W}(t)) \)) as per Assumption 3.2, as in the case of the IR problem (see Lemma 3.8). However, when a constant proportion benchmark \( \hat{\varrho}(t, \tilde{W}(t)) \equiv \hat{\varrho} \) for all \( t \) is used, both the QD-optimal probability of underperforming the benchmark and the QD-optimal probability of insolvency (i.e. probability of \( W^*_Q(t) \leq 0 \)) can be obtained analytically under some conditions, as the following lemma shows.

**Lemma 3.14.** (QD: probability of underperformance and insolvency) Suppose the following assumptions hold:

(i) Assumption 3.4: Assumption 3.3 and wealth dynamics (3.13)-(3.14) with no jumps (i.e. \( \lambda = 0 \in \mathbb{R}^{N_a} \)) in the risky asset processes (3.10); (ii) contributions are zero \((q = 0)\), and (iii) the benchmark strategy is a constant proportion strategy with \( \hat{\varrho}(t, \tilde{W}(t)) \equiv \hat{\varrho} \) 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

\[
P^{Q_d, w_d}_{\hat{\varrho}}[W^*_{Q_d}(t) \leq \tilde{W}(t)] = \Phi \left( \frac{\left[ \frac{1}{2} \hat{\varrho}^\top \Sigma \hat{\varrho} - \mu^\top \hat{\varrho} - \frac{3}{2} \eta \right]}{\sqrt{\frac{1}{2} \hat{\varrho}^\top \Sigma \hat{\varrho} + 2 \mu^\top \hat{\varrho} + \eta}} \right), \quad t \in (t_0, T],
\]

while the QD-optimal probability of insolvency at any time is given by

\[
P^{Q_d, w_d}_{\hat{\varrho}}[W^*_{Q_d}(t) \leq 0] = \Phi \left( \frac{\left[ \log \left( 1 - e^{-\beta T} \right) + \left[ \frac{1}{2} \hat{\varrho}^\top \Sigma \hat{\varrho} - \mu^\top \hat{\varrho} - \frac{3}{2} \eta \right] t \rho^\top \Sigma \hat{\varrho} + 2 \mu^\top \hat{\varrho} + \eta \right]}{\sqrt{\frac{\left[ \frac{1}{2} \hat{\varrho}^\top \Sigma \hat{\varrho} + 2 \mu^\top \hat{\varrho} + \eta \right]}} \right), \quad t \in (t_0, T].
\]

**Proof.** We provide a brief summary of the proof of (3.48), since the proof of (3.49) proceeds along similar lines. Using the definition of \( Q^*_d \) in (3.46), and observing that \( \eta = 0 \) implies that \( h_\beta(t) \equiv 0 \) for all \( t \), we have

\[
P^{Q_d, w_d}_{\hat{\varrho}}[W^*_{Q_d}(t) \leq \tilde{W}(t)] = P^{Q_d, w_d}_{\hat{\varrho}}[Q^*_d(t) \geq \left( e^{\beta T} - 1 \right) \tilde{W}(t)].
\]

Recalling that the dynamics of \( Q^*_d \) are formally the same as the dynamics of \( Q^*_r \) in (3.28), under the stated conditions of this lemma it can be shown that \( Q^*_d(t) \geq \left( e^{\beta T} - 1 \right) \tilde{W}(t) \) if and only if

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

(5.30)

Observing that the left-hand side of (5.30) is a normally distributed random variable with zero mean and a variance of \( \left[ \hat{\varrho}^\top \bar{\varrho} + 2 \mu^\top \hat{\varrho} + \eta \right] \cdot t \), the result (3.48) follows. \( \square \)

We emphasize that, in contrast to the IR-optimal probability of underperformance (see (3.30)), the closed-form expression (3.48) is obtained under the assumptions of a constant proportion benchmark strategy and zero contributions. Under these assumptions, we observe that (3.48) does not depend on the targeted outperformance spread \( \beta \). In addition, (3.49) shows that we can obtain a closed-form expression for the probability of insolvency in the case of the QD problem (under the stated assumptions), unlike in the case of the IR problem (see (3.31) and associated discussion).

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, in the sense that it holds regardless of the values of the parameters \( \gamma \) and \( \beta \) in the IR (\( \gamma \)) and QD (\( \beta \)) problems, respectively.

**Lemma 3.15.** (QD vs IR: Probability of underperformance) Suppose that the assumptions of Lemma 3.14 hold. In addition, we assume that the benchmark strategy, which is assumed to be a constant proportion strategy \( \hat{\varrho}(t, \tilde{W}(t)) \equiv \hat{\varrho} = (\hat{\varrho}_1, \ldots, \hat{\varrho}_{N_a}) \) as per Lemma 3.14, satisfies the following: (i) \( \hat{\varrho}_i \geq 0 \) for all \( i \in \{1, \ldots, N_a \} \), and (ii) \( \hat{\varrho}_i > 0 \) for at least one \( i \in \{1, \ldots, N_a \} \). 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

\[
P^{Q_d, w_d}_{\hat{\varrho}}[W^*_{Q_d}(t) \leq \tilde{W}(t)] \geq P^{Q_d, w_d}_{\hat{\varrho}_r}[W^*_{r}(t) \leq \tilde{W}(t)], \quad \forall t \in [t_0, T].
\]

**Proof.** Note that the assumptions of Lemma 3.14 are required to hold since the proof requires the analytical result (3.48) for the left-hand side of (3.51). Since these (more restrictive) assumptions also imply that the
assumptions of Lemma 3.8 are satisfied, the right-hand side of (3.51) is given by (3.30). Using the fact that the CDF $\Phi(\cdot)$ is non-decreasing, it then follows that (3.48) holds if and only if

$$
- \frac{3}{2} \sqrt{\eta} \left[ \tilde{\varrho}^\top \Sigma \tilde{\varrho} + 2 \tilde{\mu}^\top \tilde{\varrho} + \eta \right]^{1/2} \cdot \sqrt{\tau} \leq \left[ \frac{1}{2} \tilde{\varrho}^\top \Sigma \tilde{\varrho} - \tilde{\mu}^\top \tilde{\varrho} - \frac{3}{2} \cdot \sqrt{\eta} \right] \cdot \sqrt{\tau},
$$

(3.52)

where (since the assumptions of Lemma 3.14 including the absence of jumps in the risky asset processes hold), we have $\eta = \tilde{\mu}^\top \Sigma^{-1} \tilde{\mu}$. Since $\Sigma$ is positive definite, so is $\Sigma^{-1}$. Therefore, there exists matrices $\Sigma^{1/2}$ and $\Sigma^{-1/2}$ such that we have the (unique) decompositions $\Sigma = \Sigma^{1/2} \Sigma^{1/2}$ and $\Sigma^{-1} = \Sigma^{-1/2} \Sigma^{-1/2}$. As a result, recalling the conditions on the (constant proportion) benchmark strategy $\tilde{\varrho}$ and the assumption that the risky asset drift terms satisfy $\mu_i > r$ for all $i \in \{1, ..., N_\pi\}$, the Cauchy-Schwarz inequality implies that

$$
- \frac{3}{2} \sqrt{\eta} \left[ \tilde{\varrho}^\top \Sigma \tilde{\varrho} + 2 \tilde{\mu}^\top \tilde{\varrho} + \eta \right]^{1/2} = - \frac{3}{2} \left\| \Sigma^{-1/2} \cdot \tilde{\mu} \right\|_2 \left\| \Sigma^{-1/2} \tilde{\varrho} + \Sigma^{-1/2} \tilde{\mu} \right\|_2 \\
\leq - \frac{3}{2} \left( \tilde{\mu}^\top \tilde{\varrho} + \eta \right) \\
< \frac{1}{2} \tilde{\varrho}^\top \Sigma \tilde{\varrho} - \tilde{\mu}^\top \tilde{\varrho} - \frac{3}{2} \eta,
$$

(3.53)

thereby confirming that (3.51) holds for all $t \geq t_0 = 0$.

In our numerical tests, we observe that (3.51) 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 $P_{\tilde{\varrho}_{q,q}}^{t_0,\omega} \left[ W_{qd}(t) / \tilde{W}(t) \leq 1 \right]$. As the results of Section 5 and Appendix B 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.

Before proceeding to the numerical solutions of the problems under more realistic assumptions (Section 4), we conclude this section by briefly highlighting three qualitative observations regarding the analytical results.

(i) Two-stage asset allocation process: Implementing the optimal controls (3.22) and (3.39) can be interpreted as giving the outlines of a two-stage asset allocation process. The first stage consists of determining the total proportion of wealth in the risky asset basket according to each strategy. To illustrate this, let $g^{\ast, t}_{\pi, i}(t)$ and $g_{\pi, j}(t)$ denote the ith components (i.e. the allocations to the ith risky asset) of the optimal controls $g^*_{\pi}(t, X^{\ast}_{\pi}(t))$ and $g^*_{\pi}(t, X^{\ast}_{\pi}(t))$, respectively, where we drop the dependence on $X^{\ast}_{\pi}(t)$ and $X^{\ast}_{\pi}(t)$ to lighten notation. The total proportional wealth allocation to the risky asset basket according to each strategy is therefore given by

$$
\mathcal{R}^\ast_{\pi}(t) = \sum_{k=1}^{N_{\pi}} g^*_{\pi, k}(t), \quad \mathcal{R}^\ast_{qd}(t) = \sum_{k=1}^{N_{\pi}} g^*_{qd, k}(t).
$$

(3.54)

Since the investor observes $\tilde{W}(t)$ and intends to allocate their observable wealth ($W^\ast_{\pi}(t)$, or $W^\ast_{qd}(t)$) among the $N_{\pi}$ assets, the primary goal of the optimal controls (3.22) and (3.39) can then be interpreted as first determining $\mathcal{R}^\ast_{\pi}$ and $\mathcal{R}^\ast_{qd}$, respectively, using

$$
W^\ast_{\pi}(t) \cdot \mathcal{R}^\ast_{\pi}(t) = \left[ \gamma e^{-r(T-t)} - \left( W^\ast_{\pi}(t) - \tilde{W}(t) \right) \right] \cdot \sum_{k=1}^{N_{\pi}} \left[ (\Sigma + \Lambda)^{-1} \tilde{\mu} \right]_k + \tilde{W}(t) \cdot \bar{\mathcal{R}}(t),
$$

(3.55)

$$
W^\ast_{qd}(t) \cdot \mathcal{R}^\ast_{qd}(t) = \left[ h_{\pi} (t) - \left( W^\ast_{qd}(t) - \tilde{W}(t) \right) \right] \sum_{k=1}^{N_{\pi}} \left[ (\Sigma + \Lambda)^{-1} \tilde{\mu} \right]_k + \tilde{W}(t) \cdot \bar{\mathcal{R}}(t),
$$

(3.56)

where $\bar{\mathcal{R}}(t) = \sum_{k=1}^{N_{\pi}} \tilde{\varrho}_k(t)$ denotes the total allocation to the risky asset basket for the benchmark. In (3.55)-(3.56), we remind the reader that $W^\ast_{\pi}(t), W^\ast_{qd}(t), \tilde{W}(t)$ and $\bar{\mathcal{R}}(t)$ are the information known to the investor. Once $\mathcal{R}^\ast_{\pi}$ and $\mathcal{R}^\ast_{qd}$ are determined from (3.55)-(3.56), the second stage of the asset allocation
process is determining the allocation to the \( i \)th risky asset \((q_{ir,i}^*\text{ and } q_{qd,i}^*)\), which follows from

\[
\frac{W^*(t) \cdot q_{ir,i}^*(t) - \hat{W}(t) \cdot \hat{q}_i(t)}{W^*(t) \cdot R_{ir}^*(t) - W(t) \cdot \mathcal{R}(t)} = \frac{W^*(t) \cdot q_{qd,i}^*(t) - e^{rT} \hat{W}(t) \cdot \hat{q}_i(t)}{W^*(t) \cdot R_{qd}^*(t) - e^{rT}W(t) \cdot \mathcal{R}(t)} = \sum_{j=1}^{N^r} (\Sigma + \Lambda)^{-1} (\mu_j - r) \sum_{k=1}^{N^r} (\Sigma + \Lambda)^{-1} \hat{\mu}_k.
\]

The result (3.57) shows that, regardless of strategy, the amount each strategy allocates to the \( i \)th risky asset in excess of the corresponding (possibly elevated) benchmark amount, divided by the amount each strategy allocates to the entire risky asset basket in excess of that of the (possibly elevated) benchmark, is a deterministic constant. In this framework, the optimal controls (3.22) and (3.39) were therefore used primarily to determine \( R_{ir}^* \) and \( R_{qd}^* \), noting that the value of the constant in (3.57) uses knowledge of the form of the optimal controls.

Note that setting \( \hat{W} \equiv 0 \) in the case of the IR strategy (3.55), we obtain the simple continuous-time mean-variance control reported in Zhou and Li (2000), where the optimal risky asset composition is independent of the state. However, in our more general setting, we can see that the optimal risky asset composition depends on the state as shown in (3.57), but in a rather weak manner, compared to the risky asset basket vs. risk-free asset split, as shown in (3.55)-(3.56).

As a result of these observations, in our numerical experiments (Section 5) we analyze the behavior of the analytical solutions using only a “single” risky asset, which is assumed to be a diversified stock index, since this focuses on the key stage of the asset allocation problem ((3.55) and (3.56)). We will also observe that this is sufficient to provide the relevant intuition for the behavior of the optimal strategies when implemented in a setting of multiple risky assets and realistic investment constraints.

(ii) Implications for the asset allocation profile: We recall from Section 4 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, 3.7, 3.10 and 3.13 rigorously confirm that this intuition not only holds at time \( T \), but also for all \( t < T \). Specifically, we see that at time \( t < T \), the IR-optimal strategy can be interpreted as being associated with an implicit target of \( \left[ \gamma e^{-r(T-t)} + \hat{W}(t) \right] \) for \( W_{ir}^*(t) \) (see 3.22), (3.26) and (3.29). Similarly, ignoring contributions for the moment, the QD-optimal strategy can be interpreted as having an implicit target of \( e^{rT} \hat{W}(t) \) for \( W_{qd}^*(t) \) (see 3.39), (3.43) and (3.47). 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 \( t \).

This observation has significant implications for the allocation profile of the resulting investment strategies. In particular, 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. This is clearly demonstrated in the numerical results (Section 5) see for example Figures 5.2 and 5.3, regardless of whether realistic investment constraints are applicable or not.

(iii) Investment risk profile over time: The preceding observation also implies that 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.5), but the numerical results in Section 5 show that this observation remains true in even more general cases.

4 Numerical solutions

In this section, we consider the IR and QD problems under more realistic assumptions, and as a result we discuss the numerical solutions of the problems in more detail. We emphasize that the assumptions of this
section deviate substantially from the assumptions of Section 3 and include the following: (i) Not requiring any information regarding the underlying benchmark investment strategy (i.e. the proportion of the benchmark wealth invested in each underlying asset) when rebalancing the portfolio, but instead requiring only the returns of the benchmark strategy, which are usually publicly available even if the underlying strategy is not (see Remark 2.1). (ii) Rebalancing the portfolio only at discrete time intervals. (iii) Making no assumptions regarding the use of short-selling and leverage in formulating the investor’s investment strategy, and thus also requiring the parametric process dynamics (such as (3.5) and (3.10)) for the underlying assets. (iv) Simultaneously restricting the use of short-selling and leverage in formulating the investor’s investment strategy, and thus also requiring that the investor’s wealth remains non-negative at all times, $W(t) \geq 0$, $t \in [t_0 = 0, T]$. Note that assumptions regarding the constraints applied by the benchmark strategy are not required.

We therefore do not require Assumption 3.1 or Assumption 3.2 as in Section 3 so that the subsequent solution techniques also do not require the main elements of the standard assumptions in the benchmark outperformance literature concerning analytical solutions (see for example Bajeux-Besnainou et al. (2013); Basak et al. (2006); Bo et al. (2021); Brown (1999); 2000; Davis and Lee (2008); Lim and Wong (2010); Nicolosi et al. (2018); Oderda (2015); Tepla (2001); Yao et al. (2006); Zhang and Gao (2017); Zhao (2007)).

4.1 Realistic assumptions for numerical solutions

The first realistic assumption we introduce in this section is 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 (see Section 2) is now given by

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

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 rebalancing time $t_n \in \mathcal{T}$.

Note that the basic aspects of the formulation remains as in Section 2 including the use of $N_a$ assets. In particular, since $\mathcal{T}$ is given by (4.1), the investor strategy (2.1) is therefore now given by

$$\mathcal{P} = \{ p(t_n, X(t_n)) = (p_i(t_n, X(t_n)) : i = 1, ..., N_a) : t_n \in \mathcal{T} \}. \quad (4.2)$$

The second realistic assumption of this section is that no parametric model specification is required for the dynamics of the underlying assets, since we will discuss a numerical solution methodology below that is entirely data-driven. While further details are provided below (as well as in Appendix C), at this point it is sufficient to note that $R_i(t_n)$ denotes the return on asset $i \in \{ 1, ..., N_a \}$ over the time interval $[t_n, t_{n+1}]$, then the investor’s wealth dynamics in this discrete setting using a control of the form (4.2) is given by

$$W(t_{n+1}) = \left[ W(t_n) + q(t_n) \right] \cdot \sum_{i=1}^{N_a} p_i(t_n, X(t_n)) \cdot [1 + R_i(t_n)], \quad n = 0, ..., N_{rb} - 1, \quad (4.3)$$

with $W(t_0) = W(t_0) := w_0 > 0$.

The third and arguably most important realistic assumption of this section concerns the information known by the investor regarding the benchmark strategy, formulated as Assumption 4.1 below. Specifically, in contrast to Section 3 (and in particular Assumption 3.2), we now assume that the investor can observe the benchmark returns, so we do not assume any knowledge regarding the underlying benchmark asset holdings.

**Assumption 4.1.** *(Numerical solutions: Information known about the benchmark strategy)* For the numerical solution approach of this section, we assume that the investor can observe, at time $t_{n+1}$, the value of $\tilde{R}_W(t_n)$, which denotes the return to the benchmark portfolio over the time interval $[t_n, t_{n+1}]$. As a result, the benchmark wealth dynamics are given by

$$\tilde{W}(t_{n+1}) = \left[ W(t_n) + q(t_n) \right] \cdot [1 + \tilde{R}_W(t_n)], \quad n = 0, ..., N_{rb} - 1, \quad (4.4)$$

where $\tilde{W}(t_0) := w_0$. At each $t_n$, the investor therefore wishes to derive $p(t_n, X(t_n))$ as per (4.2), where the minimal form of the information $X(t_n)$ taken into account by the investor strategy is given by $X(t_n) = \left( W(t_n), W(t_{n-1}) \right)$. For deriving the investor strategy, while we do not assume parametric models for $\tilde{R}_W$ or for...
the underlying $N_a$ assets, we do assume that joint returns data (for example, joint historical return observations of $R_{it}$, $i \in \{1, ..., N_a\}$) is available to the investor.

With regards to Assumption 4.1 we highlight the following:

(i) There is now no need to observe and include the benchmark strategy $\hat{P}$ (i.e. the benchmark’s underlying asset holdings as per (2.2)) in the specification of $X$, as was required in Section 3. The reason for this is that information of the benchmark returns, together with the corresponding returns data for the set of $N_a$ underlying assets, constitute sufficient information for the neural network approach (presented below) to learn the appropriate outperformance strategy. While we use historical data in our examples in Section 5 “forward-looking” joint returns data obtained by using for example an economic scenario generator could also be used.

(ii) The availability of returns data for the benchmark is reasonable and practical, since the investor would presumably select a benchmark precisely because for example it has performed reasonably well historically, or is widely-used as a benchmark and therefore has ample associated current and historical data (the S&P 500 being a typical example).

(iii) The investor is no longer limited (as in Assumption 3.2) to investing in the same set of underlying assets.

(iv) The minimal form of $X$ specified in Assumption 4.1 is based on the characteristics of the analytical solutions of Section 3 as well as the further technical results presented in Subsection 4.2 below (see Proposition 4.2). Note that $X(t_n)$ can be augmented to contain additional relevant information such as trading and economic indicators.

If the underlying asset holdings of the benchmark strategy $\hat{P}$ are in fact known, which we emphasize is the case for many popular benchmarks used in practice (see Remark 2.1), then given

$$\hat{P} = \left\{ \hat{p} \left( t_n, \hat{X} (t_n) \right) = \left( \hat{p}_i \left( t_n, \hat{X} (t_n) \right) : i = 1, ..., N_a \right) : t_n \in \mathcal{T} \right\},$$

(4.5)

we can of course simply calculate $\hat{R}_{it} (t_n)$ using (4.5) as

$$\hat{R}_{it} (t_n) = \sum_{i=1}^{N_a} \hat{p}_i \left( t_n, \hat{X} (t_n) \right) \cdot [1 + R_t (t_n)], \quad n = 0, ..., N_{rb} - 1.$$  

(4.6)

For illustrative purposes, in the examples presented in Section 5 below, we derive investor strategies to outperform popular constant proportion benchmark strategies (in which case $\hat{P}$ is obviously known), so that (4.6) is used to calculate $\hat{R}_{it}$ in Section 5.

The fourth and final realistic assumption we introduce is 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

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

(4.7)

where

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

(4.8)

which also ensures that the investor’s wealth (with dynamics (4.3)) remains non-negative.

In summary, in this section we are therefore concerned with the numerical solution of the investment problems

$$\inf_{\mathcal{P} \in \mathcal{A}} \mathcal{L}^{\Delta t, w_0}_{\mathcal{P}} \left[ \left(W (T) - \left[ \hat{W} (T) + \gamma \right] \right)^2 \right], \quad \gamma > 0.$$  

(4.9)

$$\inf_{\mathcal{P} \in \mathcal{A}} \mathcal{L}^{\Delta t, w_0}_{\mathcal{P}} \left[ \left(W (T) - e^{\beta T \hat{W} (T)} \right)^2 \right], \quad \beta > 0.$$  

(4.10)

where $T = N_{rb} \cdot \Delta t$, the set of rebalancing times $\mathcal{T}$ is given by (4.1), the wealth dynamics are given by (4.3) and (4.4). Assumption 4.1 is applicable, and where $\mathcal{P}$ and $\mathcal{A}$ are given by (4.2) and (4.7), respectively.
4.2 Selected challenges with a dynamic programming-based solution approach

In this subsection, we provide some technical motivations for our preferred neural network-based solution methodology discussed below in Subsection 4.3.

We begin by noting that the standard approach to solve problems of the form (4.9)-(4.10) is to use dynamic programming (DP). For example, we could use the Q-learning algorithm, which is arguably the most popular data-driven Reinforcement Learning algorithm (see for example Dixon et al. (2020); Gao et al. (2020); Lucarelli and Borrotti (2020); Park et al. (2020)) that fundamentally relies on the DP principle to solve (4.9)-(4.10). However, as we discuss in this subsection, there are many practical challenges associated with DP-based solution techniques. While some of the challenges, such as those related to error amplification, do enjoy some recognition in the literature (as discussed below), we now present an additional motivation for avoiding the use of the DP principle to solve problems specifically of the form (4.9)-(4.10).

In the following proposition (Proposition 4.2), we demonstrate that the DP approach is, in a sense, unnecessarily high-dimensional in our setting. Therefore, significant computational advantages may follow if the investor is able to solve for the optimal control (i.e. the optimal investor strategy $P^*$ according to the chosen objective) without the need to solve for the corresponding performance criterion which is typically required (either implicitly or explicitly) by DP-based solution methods (see for example Oksendal and Sulem (2019)).

For concreteness and illustrative purposes, note that Proposition 4.2 incorporates some assumptions which we emphasize are not required in the rest of this section. These assumptions are simply made for purposes of concreteness, 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 4.2. (Discrete rebalancing: Dimensions of the dynamic programming solutions to the IR and QD problems) Suppose the IR($\gamma$) and QD($\beta$) problems in (4.9)-(4.10) are solved using dynamic programming in the case where the portfolio is only rebalanced at the set of discrete rebalancing times $T$ 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 (4.10). (ii) The benchmark’s asset allocation strategy is an adapted feedback control of the form $\hat{p}(t_n, X(t_n)) = \hat{p}(t_n, W(t_n))$, $t_n \in T$. (iii) At each rebalancing event, the investor can observe the benchmark asset allocation vector $\hat{p}(t_n, W(t_n))$.

Then at each fixed rebalancing time $t_n \in T$, regardless of the number of underlying assets $N_a$, the optimal controls of problems IR($\gamma$) and QD($\beta$) in (4.9)-(4.10) 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 of the form (4.2) for each problem consists of the vectors $\hat{p}^*_{ir}(t_n, X^*_{ir}(t_n))$ and $\hat{p}^*_{qd}(t_n, X^*_{qd}(t_n))$, respectively, where $X^*_{ir}(t_n) = (W^*_ir(t_n), W(t_n))$ and $X^*_{qd}(t_n) = (W^*_qd(t_n), W(t_n))$.

However, in using dynamic programming to obtain the optimal controls $\hat{p}^* : \mathbb{R}^{(2N_a + 1)} \rightarrow \mathbb{R}^{N_a}, k \in \{ir, qd\}$, which are only two-dimensional controls at each fixed rebalancing time $t_n \in T$, 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 $t_n, t_{n+1} \in T$.

Proof. Due to the importance of this proposition in motivating the neural network-based approach presented below, the proof is presented in detail in Appendix A.2.

Therefore, given the stated assumptions, Proposition 4.2 shows that the case of discrete rebalancing, the investor has 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 $t_n \in T$. In contrast, in the case of continuous rebalancing, the results of Section 3.3 show that that the investor only requires the solution of a 2-dimensional value function at every given time $t \in [t_0, T]$.

In the more realistic case of discrete rebalancing, we can therefore interpret Proposition 4.2 as highlighting the observation that the DP solution approach is potentially unnecessarily high-dimensional in the case of benchmark outperformance problems of the form (4.9)-(4.10). This follows since the investor is expected to be more concerned with the (comparatively low-dimensional) optimal control instead of the high-dimensional performance criterion.

Many of the published concerns with using DP techniques, including in multi-asset portfolio optimization settings (see Li et al. (2020); Tsang and Wong (2020)), follow from the fact that once the above-mentioned
high-dimensional performance criterion (i.e. an approximation to the conditional expectation) is required, then value iteration (as in for example the Q-learning algorithm, or in Appendix A.2 where it takes the form of solving a high-dimensional PIDE ) implies that an optimization problem has to be solved to determine the value function using the performance criterion at each rebalancing time \( t_n \in \mathcal{T} \) recursively backwards. As a result, value iteration can cause significant problems with regards to the stability and convergence associated with the estimated value function and estimated optimal control, since there occurs a potentially significant amplification of the estimation errors over each iteration (see for example Li et al. (2020); Tsang and Wong 2020; Wang and Foster (2020)).

4.3 Overview: neural network (NN) approach

We now discuss the numerical solution of problems (4.9)-(4.10) 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 the dimensionality and error amplification issues outlined in Subsection 4.2 above.

The numerical solution methodology for (4.9)-(4.10) presented here is inspired by an approach that has been successfully applied to other portfolio optimization problems in some of our previous work (see Li and Forsyth (2019); Ni et al. (2022); Van Staden et al. (2021b)). This section aims to give the reader an intuitive overview of this approach as specifically applied to (4.9)-(4.10), leaving the technical details and rigorous definitions for Appendix C.

Our basic task in solving problems (4.9)-(4.10) is to determine the control \( \mathcal{P} \) (see (4.2)) in feedback form \( p(t, X(t)) \), where we assume that \( p(t, X) \in \mathcal{Z} \) is a continuous function of \((t, X)\), an assumption that will be explained in more detail below. Appealing to the Universal Approximation Theorem (see e.g. Hornik et al. (1989)), we approximate this control function by a neural network \( F(t, X(t); \theta) \equiv F(\cdot, \theta) \), where \( \theta \in \mathbb{R}^{n_\theta} \) is the set of NN parameters (i.e. the NN weights and biases), so that

\[
p(t, X(t)) \approx F(t, X(t); \theta) \equiv F(\cdot, \theta) \tag{4.11}
\]

We therefore approximate problems (4.9)-(4.10) by

\[
\inf_{\theta \in \mathbb{R}^{n_\theta}} E_{\mathcal{F}(\cdot, \theta)}^{\varepsilon, n_0} \left( (W(T; \theta) - \hat{W}(T) + \gamma)^2 \right), \quad \text{and} \quad \inf_{\theta \in \mathbb{R}^{n_\theta}} E_{\mathcal{F}(\cdot, \theta)}^{\varepsilon, n_0} \left( (W(T; \theta) - e^{\beta T} \hat{W}(T))^2 \right), \tag{4.12}
\]

where the optimization problems (4.12) are unconstrained if the NN \( F(\cdot, \theta) \) satisfies certain structural assumptions (see Assumptions C.1 in Appendix C.1).

From a computational point of view, the expectations \( E_{\mathcal{F}(\cdot, \theta)}^{\varepsilon, n_0}(\cdot) \) in (4.12) are approximated using a finite set of samples \( Y \), which in the usual terminology (see e.g. Goodfellow et al. 2016) serves as the training data set of the NN. Specifically, in the case of problems (4.12), if \( X \) in (4.11) is of the minimal form \( X(t_n) = (W(t_n), \hat{W}(t_n)) \) as per Assumption 4.1, then \( Y \) is assumed to be of the form \( Y = \{ Y^{(j)} : j = 1, \ldots, N_d \} \), where each \( Y^{(j)} \) represents a path of joint return observations \( R_{W,i} \) and \( R_{i,j} \in \{ 1, \ldots, N_a \} \) observed at each \( t_n \in \mathcal{T} \). As discussed in Appendix C, we implement a stationary block bootstrap resampling approach to construct \( Y \) using historical data. Note that if \( X \) is not of the minimal form, the relevant additional information will be included in the training dataset \( Y \).

Two special cases for constructing the training data set \( Y \) are worth highlighting: (i) If the underlying benchmark strategy (4.5) is in fact known, only the underlying asset returns \( R_{i,j} \in \{ 1, \ldots, N_a \} \) are needed in \( Y \), since \( R_{W,i} \) can be calculated using (4.6). (ii) In the special case where the underlying assets follow specified dynamics (as is the case in the ground truth results presented in Appendix C.3 where we show that the NN approach recovers the analytical results of Section 3), the data set \( Y \) can be obtained via Monte Carlo simulation using the given dynamics.

For a given \( \theta \in \mathbb{R}^{n_\theta} \) in (4.12) and a given training dataset \( Y \), each sample \( Y^{(j)} \in Y \) can therefore be associated with corresponding wealth outcomes \( W^{(j)}(T) \) and \( \hat{W}^{(j)}(T) \) calculated using (4.3), (4.4) and (4.11) - please refer to Appendix C.2 for more technical details. Our final computational problems for (4.9)-(4.10) can
therefore be expressed as

\[
IR(\gamma) : \min_{\theta \in \mathbb{R}^{N_\theta}} \left\{ \frac{1}{N_d} \sum_{j=1}^{N_d} \left( W^{(j)}(T; \theta) - \left[ \hat{W}^{(j)}(T) + \gamma \right] \right)^2 \right\}, \tag{4.13}
\]

\[
QD(\beta) : \min_{\theta \in \mathbb{R}^{N_\theta}} \left\{ \frac{1}{N_d} \sum_{j=1}^{N_d} \left( W^{(j)}(T; \theta) - e^{\beta T} \hat{W}^{(j)}(T) \right)^2 \right\}. \tag{4.14}
\]

The optimal NN parameter vectors for problems (4.13)-(4.14), denoted by \( \theta^*_k, k \in \{\text{ir}, \text{qd}\} \) respectively, can then be obtained using standard (unconstrained) optimization methods. For more details, see Appendix C regarding both the chosen optimization algorithm details, as well as the relevant gradient computations.

The resulting optimal investment strategies \( p^*_k(\cdot, X(\cdot)) \approx F(\cdot, \theta^*_k), k \in \{\text{ir}, \text{qd}\} \) can then be implemented on a testing data set \( Y^{test} \) to assess the out-of-sample performance of the resulting strategies. The contents of \( Y^{test} \) is expected to differ from that of the training dataset \( Y \), for example it might be based on a different historical time period or different data generation assumptions, but it is assumed to have a similar structure to the training dataset.

Emphasizing that the underlying technical details to solve problems (4.13)-(4.14) are provided Appendix C, we highlight two important properties of our approach:

(i) Using \( t \) simply as a parameter (i.e. feature of the NN) in the approximation to the control \( p(t, X(t)) \approx F(t, X(t); \theta) \) enforces the condition that, in the limit as \( \Delta t \rightarrow 0 \), the approximate control is a continuous function of time. We believe that this is a necessary practical constraint to any investment policy. 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 \( X(t) \) is a smooth function of time. This aspect distinguishes our approach from that of Tsang and Wong (2020), where the control to a multi-asset portfolio optimization problem is also represented by a NN, but where the smooth behavior of the control as \( \Delta t \rightarrow 0 \) is not automatically guaranteed.

(ii) Observing that alternatives to DP-based numerical solution techniques enjoy significant recent interest in the context of portfolio optimization literature (see for example Duarte et al. (2021), Chen and Langrene (2020) and Tsang and Wong (2020)), we highlight that our approach also avoids using the DP principle to solve problems (4.9)-(4.10). In the case of our approach, we emphasize that we solve only a single optimization problem for each objective ((4.13)-(4.14)), and since the time \( t \) is used as a feature (input) into the NN, while the optimal parameters \( \theta^*_k, k \in \{\text{ir}, \text{qd}\} \) are independent of time, the optimal investment strategy \( p^*_k(\cdot, X(\cdot)) \approx F(\cdot, \theta^*_k), k \in \{\text{ir}, \text{qd}\} \) gives the required optimal asset allocation at any rebalancing time. We therefore also avoid the problems associated with value iteration over \( t_n \in T \) outlined in Subsection 4.2.

For further details, the reader is referred to Appendix C. Finally, we note that under some conditions (for example when investments constraints are not binding), solving (4.13), (4.14) results in numerical solutions that indeed converge to the analytical solutions of problems (2.5) and (2.8) given in Section 3. More information on these ground truth comparison results can be found in Appendix C.4.

5 Illustration of investment results

In this section, we illustrate the results from investing according to the optimal strategies associated with problems (2.5) and (2.8), using both analytical solutions (Section 3), as well as numerical solutions (Section 4).

We formulate a realistic 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 Ang (2014)).

5.1 Investment scenario

In this subsection, we briefly outline the investment scenario details on which all the subsequent results are based. Table 5.1 summarizes the general assumptions. Note that we choose an investment time horizon of 5This follows since a different set of NN parameters is applicable to each rebalancing time \( t_n \) in Tsang and Wong (2020).
In the subsequent analyses, we will compare the results from investing according to the IR- and QD-optimal investment strategies. Since there are many possibilities for the basis of comparison, we make the practical assumption that the investor wishes to achieve an expected terminal wealth of $\hat{E}$ regardless of whether the IR or QD strategy is followed. In more detail, if the benchmark investment strategy $\hat{P}$ results in an expected value of terminal wealth $E^{\hat{P}, \text{no} w_0}_{\text{ir}} \left[ \hat{W} (T) \right] = K$, we assume the investor chooses some value of $\hat{\beta} > 0$ in (5.1) below to achieve an expected terminal wealth of $\hat{E}$:

$$E_{\hat{P}^{\text{ir}} w_0} \left[ W_{\text{ir}}^E (T) \right] \equiv E^{\hat{P}^{\text{ir}} w_0}_{\text{ir}} \left[ W_{\text{ir}}^E (T) \right] := \hat{E} := e^{\hat{\beta} T} \cdot K = e^{\beta T} \cdot E^{\hat{P}, \text{no} w_0}_{\text{ir}} \left[ \hat{W} (T) \right].$$

The desired target expectation (5.1) can be achieved by solving numerically (or in some cases, analytically - see Appendix A) for values of $\gamma = \hat{\gamma}$ in the IR ($\gamma$) problem and $\beta = \hat{\beta}$ in the QD ($\beta$) problem such that the associated IR- and QD-optimal strategies $\hat{P}_{\text{ir}}^E$ and $\hat{P}_{\text{qd}}^E$, respectively, each result in the desired expected value of terminal wealth $\hat{E}$. Note that (5.1) implies that we always have the strict inequality $\hat{E} > K$ (since $\hat{\beta} > 0$), which is required since if $\hat{E} = K$, then the IR- and QD-optimal strategies will be identical to the benchmark strategy.\[6\]

### 5.2 Underlying assets and source data

Table 5.2 summarizes the combinations of candidate assets considered by the investor for investment (combinations are identified by the label “P$x$”, $x \in \{0, 1\}$), as well as the benchmarks under consideration (benchmarks are identified by the label “BM$x$”, $x \in \{0, 1\}$). Both benchmarks portfolios are equally-weighted between stocks and bonds.

As noted in Subsection 4.1, the investor is not necessarily limited to investing in the same assets that are used by the benchmark. While more details are provided in subsequent sections, note that we will 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).

Table 5.2 also indicates the data sources used to obtain the underlying data. In summary, data for the basic assets such as the T-bills/bonds and the broad market index were obtained from the CRSP, whereas factor data for Size and Value (see Fama and French (2015, 1992)) were obtained from Kenneth French’s data library. All data was obtained for the period from 1963:07 to 2020:12, which includes the period of significant market volatility experienced during 2020.

As a result of the reasonably long time horizon (Table 5.1), we will assume as in for example Forsyth and Vetzal (2019) and Forsyth et al. (2019) that the investor is primarily interested in the real (or inflation-adjusted)\[6\]

\[6\] While intuitive, the fact that $\hat{E} = K$ implies $P_{\text{ir}}^E = P_{\text{qd}}^E = \hat{P}$ can also be shown analytically by setting $\hat{E} = K$ in the expressions for $\hat{\gamma}$ and $\hat{\beta}$ in Lemma A.3 in Appendix A and then substituting the resulting values into the optimal controls (3.22) and (3.10).

\[7\] 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.

\[8\] See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
Table 5.2: Underlying assets and data sources for the investor portfolios and the benchmarks. Portfolios of candidate assets considered by the investor are abbreviated by “P<sub>x</sub>”, x ∈ {0,1}. The tick mark “✓” indicates the inclusion of the asset in the portfolio optimization problem, with N<sub>a</sub> denoting the total number of candidate assets. Two constant proportion benchmarks are considered, abbreviated by “BM<sub>x</sub>”, x ∈ {0,1}, with asset holdings as a percentage of wealth \( \hat{p}_i \) as indicated. CRSP refers to the Center for Research in Security Prices, and KFDL refers to Kenneth R. French’s Data Library.

<table>
<thead>
<tr>
<th>Label</th>
<th>Asset description</th>
<th>Data source and definition</th>
<th>Investor portfolios</th>
<th>Benchmarks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>P0</td>
<td>P1</td>
</tr>
<tr>
<td>T30</td>
<td>30-day Treasury bill</td>
<td>CRSP: Monthly returns for 30-day Treasury bill.</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>B10</td>
<td>10-year Treasury bond</td>
<td>CRSP: Monthly returns for 10-year Treasury bond.</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Market</td>
<td>Market portfolio (broad equity market index)</td>
<td>CRSP: Monthly returns, including dividends and distributions, for a capitalization-weighted index (the VWD index) consisting of all domestic stocks trading on major US exchanges.</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Size</td>
<td>Portfolio of small stocks</td>
<td>KFDL, “Portfolios Formed on Size”: 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.</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Value</td>
<td>Portfolio of value stocks</td>
<td>KFDL, “Portfolios Formed on Book-to-Market”: 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.</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

Number of candidate assets (N<sub>a</sub>): 2 5 2 3

5.3 Illustration of analytical solutions

For the illustration of the analytical results of Section 3, we assume that the investor portfolio P0 is constructed to outperform benchmark BM0 as per Table 5.2. As discussed in Section 3, we limit our attention to N<sub>a</sub> = 2 assets for illustrative purposes in this subsection, since considering only a diversified market index representing the risky asset basket (the “risky asset”) in conjunction with a risk-free asset are sufficient to illustrate the key behavioral aspects of the closed-form optimal controls. In the terminology of Section 3, T10 and Market (Table 5.2) are therefore associated with the risk-free and risky assets, respectively.

In order to parameterize (3.5) and (3.8), we use the same calibration methodology as outlined in Dang and Forsyth (2016); Forsyth and Vetzal (2017). For illustrative purposes, we assume the risky asset evolves according to the dynamics of the Kou (2002) model, with log \( \xi \) having an asymmetric double-exponential distribution,

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

where \( \nu \) denotes the probability of an upward jump given that a jump occurs. Table 5.3 summarizes the resulting parameters obtained using the filtering technique for the calibration of jump diffusion processes - see Dang and Forsyth (2016); Forsyth and Vetzal (2017) for the relevant methodological details.

We now compare analytical investment results on the basis of (5.1), using 10<sup>6</sup> Monte Carlo simulations of asset dynamics (3.5), (3.8), and (5.2) with parameters as in Table 5.3.

\[\text{http://www.bls.gov.cpi}\]
Table 5.3: Analytical solutions: Calibrated, inflation-adjusted parameters for asset dynamics (3.5) and (3.8), with \( f_t(\xi) \) given by (5.2). For calibration purposes, a jump threshold equal to 3 has been used in the methodology of Dang and Forsyth (2010).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( r )</th>
<th>( \mu )</th>
<th>( \sigma )</th>
<th>( \lambda )</th>
<th>( \nu )</th>
<th>( \zeta_1 )</th>
<th>( \zeta_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.0074</td>
<td>0.0749</td>
<td>0.1392</td>
<td>0.2090</td>
<td>0.2500</td>
<td>7.7830</td>
<td>6.1074</td>
</tr>
</tbody>
</table>

Table 5.4 presents the key results using two expected value targets as in (5.1), namely \( E = 370 \) (\( \hat{\beta} \approx 1\% \) in (5.1)) and \( E = 400 \) (\( \beta \approx 2\% \)). Results are shown for both the terminal wealth (absolute performance) and the wealth ratio (relative performance).

Since the results in Table 5.4 include both jumps (see Table 5.3) and nonzero contributions (see Table 5.1), two key assumptions underlying Lemma 3.15 are violated, but it is interesting that the probability of underperformance (i) is larger for the QD strategy than for the IR strategy, and (ii) remains insensitive to the aggressiveness of the outperformance target. Note that these results are obtained via an implementation of optimal investment strategies (3.22) and (3.39) in a Monte Carlo simulation, and thus the probabilities of underperformance are expected to exhibit some variability.

Table 5.4 shows that as the outperformance target becomes more aggressive, differences between the outcomes associated with the IR and QD strategies are magnified, with the IR strategy resulting in improved downside wealth outcomes, as well as improved target outperformance (as illustrated by the wealth ratio results).

Table 5.4: Analytical solutions, no constraints, investor portfolio \( P_0 \), benchmark \( BM_0 \): Selected quantities associated with the distributions of the investor’s target terminal wealth \( W_j^{E_j}(T) \) and ratio \( W_j^{E_j}(T)/W(T) \), for \( j \in \{ir, qd\} \). 10\(^6\) Monte Carlo simulations, with \( \bar{E} = 370 \) and \( \bar{E} = 400 \) in (5.1), “CExp 5\%” refers to the average of the lowest 5\% of outcomes, and “Prob. underp.” is the probability of underperformance, \( \mathbb{P}(W_j^{E_j}(T)/W(T) \leq 1) \), for \( j \in \{ir, qd\} \).

<table>
<thead>
<tr>
<th>Quantity</th>
<th>BM0</th>
<th>( E_j = 370 )</th>
<th>( E_j = 400 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( W_j^{E_j}(T) ) mean</td>
<td>330</td>
<td>370</td>
<td>370</td>
</tr>
<tr>
<td>( W_j^{E_j}(T)/W(T) ) mean</td>
<td>400</td>
<td>400</td>
<td>1.21</td>
</tr>
<tr>
<td>5th pctile</td>
<td>208</td>
<td>244</td>
<td>236</td>
</tr>
<tr>
<td>5th pctile</td>
<td>228</td>
<td>244</td>
<td>236</td>
</tr>
<tr>
<td>95th pctile</td>
<td>454</td>
<td>504</td>
<td>518</td>
</tr>
<tr>
<td>Prob. underp.</td>
<td>2.62%</td>
<td>3.35%</td>
<td>2.61%</td>
</tr>
</tbody>
</table>

Figure 5.1 illustrates the simulated probability density functions (PDFs) associated with the benchmark and investor wealth outcomes using the analytical solutions. The corresponding CDFs are shown in Appendix B. Figure 5.1(b) shows that the IR- and QD-optimal strategies share many (qualitative) similarities in terms of the wealth PDFs when contrasted with the benchmark, while Figure 5.2(a) illustrates the fat left tails of the ratio PDF for both strategies. This has the result that although the 5th percentile of the ratio distribution is better/higher for the IR strategy than for the QD strategy as per Table 5.4, the average of the lowest 5\% outcomes of the ratio results (see “CExp 5\%” in Table 5.4) are identical for both strategies.

Considering the investment strategy in more detail, Figure 5.2(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 5.2(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 5.2(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 3.6 and Lemma 3.12.

Note that the qualitative aspects of the relative behavior of the optimal investment strategies observed in Figure 5.2(a) is in fact to be expected, as we show rigorously in Appendix A (see Theorem A.5).
Figure 5.1: Analytical solutions, no constraints, investor portfolio P0, benchmark BM0: Simulated PDFs of benchmark and investor’s target terminal wealth \( \hat{W}(T) \) and \( W_{ir}^F(T) \), respectively, as well as the ratio \( W_{ir}^F(T)/\hat{W}(T) \), for \( j \in \{ir, qd\} \). 10⁶ Monte Carlo simulations, \( \xi = 400 \) in (5.1). The corresponding CDFs are shown in Figure B.1 in Appendix B.

Figure 5.2: Analytical solutions, no constraints, investor portfolio P0, benchmark BM0: 80th percentiles of the investment in the single risky asset \( p_j^F(t) \) and probability of insolvency as a function of time \( t \rightarrow P_{j,w}^{\beta,T}[W_j^F(t) \leq 0] \), for \( j \in \{ir, qd\} \). 10⁶ Monte Carlo simulations, \( \xi = 400 \) in (5.1).

5.4 Illustration of numerical solutions

We now consider the more realistic scenario of multiple investment constraints and discrete rebalancing (see Table 5.1). In contrast to the preceding results, the problems are now solved as outlined in Section 4 but the results are still compared on the basis of (5.1). In particular, note that we now require the targeted expected value \( \mathcal{E} = e^{\beta_T \mathcal{K}} \) (see (5.1)) to be achieved on the neural network’s training data set \( Y \). This means that the values \( \gamma = \gamma_{ir} \) and \( \beta = \beta_{qd} \) achieving (5.1) can no longer be derived analytically as in the preceding section (see Lemma A.3 in Appendix A), but are obtained numerically using iterative solutions of problems (4.13)-(4.14). In terms of the NN structure and other parameters used in the implementation of the numerical solution methodology outlined in Section B, please refer to Appendix C.3.

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 a detailed discussion of the rationale as well as a theoretical analysis, see Ni et al. (2022). Note that all subsequent results were also tested using various different assumptions for expected block sizes, and since qualitatively similar results were obtained (as expected based on the robustness assessments presented in Li and Forsyth (2019); Ni et al. (2022)), only results for the data sets as outlined in Table 5.5 are presented. For data sets DS1 and DS2, the relatively shorter expected block sizes 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. However, we emphasize again that our conclusions remain very robust to expected blocksize assumptions. Note that Table 5.5 also provides the mean benchmark terminal wealth \( \mathcal{K} \) obtained on each training data set. In the case of DS0 and benchmark BM0, the mean \( \mathcal{K} = 334 \) in Table 5.5 is slightly higher than the mean reported in Table 5.4, this difference is entirely due to the effects of discrete rebalancing.

Remark 5.1. (Rationale for training data period selections) With regards to the data set parameters given in Table 5.5, we highlight the following. Data set DS0, obtained using simulation of specified asset dynamics, is included to isolate the impact of discrete rebalancing and investment constraints on the results of Subsection 5.3. DS1 and DS2 incorporates 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.
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 an 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 Forsyth (2018)). As a result, 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. Data sets DS1 and DS2 will therefore be the main focus of our analysis. Regardless of these observations, to assess the robustness of our conclusions, 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 Fama and French (1992, 1993).

Note that Table 5.5 also considers scenarios of annual and quarterly rebalancing, to assess the robustness of our conclusions to different realistic rebalancing frequencies.

Table 5.5: Data set combinations, labelled DS_x, x ∈ {0, 1, 2}, used for training and testing the neural network. “SBBR” refers to stationary block bootstrap resampling, with expected blocksize reported in brackets.

<table>
<thead>
<tr>
<th>Label</th>
<th>Rebal. freq</th>
<th>Source data</th>
<th>Data set generation</th>
<th>Benchmark</th>
<th>Source data</th>
<th>Data set generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS0</td>
<td>Continuous</td>
<td>10⁶ Monte Carlo simulations of asset dynamics (3, 5, 2) with parameters as in Table 5.5</td>
<td>BM0: K = 334</td>
<td>N/a</td>
<td>N/a</td>
<td></td>
</tr>
<tr>
<td>DS1</td>
<td>Annually</td>
<td>Historical data, 1963:07 - 2009:12 (6 months)</td>
<td>SBBR</td>
<td>BM1: K = 338</td>
<td>Historical data, 2010:01 - 2020:12 (3 months)</td>
<td>SBBR</td>
</tr>
<tr>
<td>DS2</td>
<td>Annually</td>
<td>Historical data, 1963:07 - 1999:12 (6 months)</td>
<td>SBBR</td>
<td>BM1: K = 364</td>
<td>Historical data, 2000:01 - 2010:12 (3 months)</td>
<td>SBBR</td>
</tr>
<tr>
<td>DS3</td>
<td>Quarterly</td>
<td>Historical data, 1995:01 - 2009:12 (3 months)</td>
<td>SBBR</td>
<td>BM1: K = 352</td>
<td>Historical data, 2010:01 - 2020:12 (3 months)</td>
<td>SBBR</td>
</tr>
</tbody>
</table>

Table 5.5 illustrates the combinations of investor portfolios and benchmarks used in the subsequent results, as well as the targeted level of outperformance. In the case of portfolio P0, benchmark BM0 and data set DS0, we choose \( \mathcal{E} = 370 \) to ensure alignment with the analytical solutions presented in Table 5.4, but note that this still implies that \( \hat{\beta} \approx 1.0\% \). In the case of using portfolio P1 (5 assets) to outperform BM1 (3 assets), we use a slightly more ambitious value of \( \hat{\beta} \approx 1.7\% \), since the investor has more opportunities for outperformance given that factors are available for investment (see Van Staden et al. (2021b)). Note that the \( \mathcal{E} \) values reported are different due to different values of \( K \) (see Table 5.5).

Table 5.6: Numerical solutions, with constraints: Target expected values for combinations of the investor portfolios, benchmarks and data set combinations. As per Table 5.5, 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.

<table>
<thead>
<tr>
<th>Investor portfolio</th>
<th>To outperform benchmark:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BM0 (2 assets)</td>
</tr>
<tr>
<td>P0 (2 assets)</td>
<td>DS0: ( \mathcal{E} = 370 \ (\beta \approx 1.0%) )</td>
</tr>
<tr>
<td>P1 (5 assets)</td>
<td>N/a</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.7, based on using portfolio P0 to outperform benchmark BM0 on training data set DS0, shows the impact of applying realistic investment constraints and discrete rebalancing to the results of Subsection 5.3. For ease of reference, the results for \( \mathcal{E} = 370 \) from Table 5.4 are repeated in the “No constraints” columns, while
the BM0 performance changes slightly in the “With constraints” case due to discrete rebalancing being used (as noted above).

The key observations regarding Table 5.7 are the following: (i) with constraints, QD-optimal probability of underperformance is now lower than the corresponding IR-optimal value, and thus the results of Lemma 5.15 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.

Table 5.7: 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 5.1. NN trained on data set DS0. Since no out-of-sample testing is conducted for DS0 (see Table 5.3), the “With constraints” results are obtained on the training data set.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>No constraints: P0, $\mathcal{E} = 370$ (Table 5.4)</th>
<th>With constraints: P0, $\mathcal{E} = 370$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BM0</td>
<td>W_{T} (T)</td>
</tr>
<tr>
<td>Mean</td>
<td>330</td>
<td>370</td>
</tr>
<tr>
<td>CExp 5%</td>
<td>208</td>
<td>193</td>
</tr>
<tr>
<td>5th pctl</td>
<td>228</td>
<td>244</td>
</tr>
<tr>
<td>Median</td>
<td>325</td>
<td>368</td>
</tr>
<tr>
<td>95th pctl</td>
<td>454</td>
<td>504</td>
</tr>
<tr>
<td>Prob. underp.</td>
<td>2.62%</td>
<td>3.35%</td>
</tr>
</tbody>
</table>

Table 5.8 focuses only on the numerical results obtained when applying realistic investment constraints, and compares the training and testing (out-of-sample) results for portfolio P1 constructed to outperform benchmark BM1 on data set DS1. While the training data results are (qualitatively) as expected given the results of Table 5.7, it is the out-of-sample results in Table 5.8 that are the most significant. Specifically, we see that the QD strategy outperforms the IR strategy in the out-of-sample results regardless of whether we consider the wealth, the wealth ratio or the probability of underperformance. Qualitatively similar results are also obtained for a different data set (DS2), as shown in Appendix B (see Table B.1), so it should be emphasized that the reported out-of-sample performance is robust to the choice of underlying data sets.

We also note the potential risk of underperforming the benchmark is significantly larger out-of-sample (both in terms of probability and the downside statistics of the wealth ratio) than for the training data. This is to be expected, since with the actual market data used here the true underlying data generating process is not known, and thus the training results are not expected to generalize perfectly.

Table 5.8: Numerical solutions, with constraints, investor portfolio P1, benchmark B1, data set DS1: Training and testing results for mean terminal wealth $\mathcal{E} = 400$ ($\beta \approx 1.7\%$ in Table 5.4) on the training data.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BM1</td>
<td>W_{T} (T)</td>
</tr>
<tr>
<td>Mean</td>
<td>338</td>
<td>400</td>
</tr>
<tr>
<td>CExp 5%</td>
<td>199</td>
<td>213</td>
</tr>
<tr>
<td>5th pctl</td>
<td>219</td>
<td>254</td>
</tr>
<tr>
<td>Median</td>
<td>328</td>
<td>394</td>
</tr>
<tr>
<td>95th pctl</td>
<td>490</td>
<td>563</td>
</tr>
<tr>
<td>Prob. underp.</td>
<td>3.07%</td>
<td>2.77%</td>
</tr>
</tbody>
</table>

In order to assist in the explanation of the results of Table 5.8, Figure 5.3 presents the 80th percentiles of the proportion of wealth invested in each candidate asset in P1 over time according to the IR- and QD-optimal investment strategies, on the training data set of DS1. While a detailed analysis of the risk/return analysis of these underlying assets using historical data is provided in Van Staden et al. (2021b), here we simply note that the zero investment in Size, as well as the large investment in Value, are to be expected given their historical performance.

Note that the results in Table 5.8 show a significantly larger probability of underperformance (out-of-sample) compared to the training set. This can be easily explained: the time frame of the training data set was
a historical period where value investing (i.e. a large investment in Value) outperformed, but the testing time frame was a "lost decade" for value investing (see for example Arnott et al. (2020); Israel et al. (2020)).

With regards to Figure 5.3 we observe that the key qualitative observations regarding the analytical solutions discussed in Subsection 5.3 (see also Theorem A.3 in Appendix A) hold even if investment constraints are applied. Specifically, compared to the QD strategy, Figure 5.3 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.

Figure 5.3: 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.

Figure 5.4 and Figure 5.5 illustrate the consequences of applying the investment strategies illustrated in Figure 5.3 on the out-of-sample or testing data of DS1 (annual rebalancing) and DS3 (quarterly rebalancing). Note that the corresponding CDFs are illustrated in Appendix B Figures B.2 and B.3. While the wealth distribution of the QD strategy is possibly preferable (Figures 5.4(a) and 5.5(a)), the different wealth ratio distributions (Figures 5.4(b) and 5.5(b)) provide a dramatic visual illustration of the underlying results giving rise to the selected statistics in for example Table 5.8 with the QD strategy resulting in a much more desirable outperformance profile than the IR strategy.

Figure 5.4: 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 $W^j_{T} \sim (T)$, respectively, as well as the ratio $W^j_{T} \sim (T)/\hat{W}(T)$, 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.

The preceding results presented for data sets DS1 and DS3 also hold for other data sets, for example data set DS2 (see Appendix B).

Finally, Table 5.9 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 until the maturity $T = 10 + \Delta t$ years is reached. Results are shown for strategies trained on the training data of DS1 and DS2 using annual rebalancing, as well as on the training data of DS3 using quarterly rebalancing. Note that there is significant overlap (5 years) between the underlying data of each pair of adjacent rows. It should be emphasized that Table 5.9 presents out-of-sample results, since the probability that the actual historical path appears in the training data set can be shown to be vanishingly small (Ni et al. 2022).
Table 5.9: Terminal wealth $W_j^T(\tau)$ 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.

<table>
<thead>
<tr>
<th>$t_0$ for $[t_0, T + t_0]$</th>
<th>BM1</th>
<th>Annual rebalancing</th>
<th>Quarterly rebalancing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NN trained on DS1</td>
<td>NN trained on DS2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IR, QD, Best</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BM1</td>
<td>IR, QD, Best</td>
</tr>
<tr>
<td>1980:01</td>
<td>463</td>
<td>537</td>
<td>556</td>
</tr>
<tr>
<td>1985:01</td>
<td>400</td>
<td>467</td>
<td>479</td>
</tr>
<tr>
<td>1990:01</td>
<td>497</td>
<td>568</td>
<td>593</td>
</tr>
<tr>
<td>1995:01</td>
<td>384</td>
<td>460</td>
<td>474</td>
</tr>
<tr>
<td>2000:01</td>
<td>260</td>
<td>315</td>
<td>309</td>
</tr>
<tr>
<td>2005:01</td>
<td>342</td>
<td>400</td>
<td>405</td>
</tr>
<tr>
<td>2010:01</td>
<td>370</td>
<td>432</td>
<td>442</td>
</tr>
</tbody>
</table>

The preceding numerical results also provide a vivid illustration of the potential challenges in broadly assuming (as is often implicitly assumed in the literature) that the qualitative aspects of the analytical solutions will also hold in more realistic settings. Specifically, we highlight the following observations.

- **Reliance on trading in insolvency:** We observed that the inequality $\frac{\text{Ratio}}{\text{outperformance}}$ of Lemma 3.15 reverses in a more realistic setting, with the QD strategy consistently having a smaller probability of underperforming the benchmark (in-sample or out-of-sample) than the IR strategy. Numerical evaluations using (3.31) and (3.49) showed that the IR-optimal strategy relies to a far greater extent than the QD strategy on the trading in insolvency (see Figure 5.2(b)) which is permitted under Assumption 3.1. Once this is no longer allowed, the risk profile of the IR strategy (see Figure 5.2 and Theorem A.5 as well as Figure 5.5) can in fact become a liability, since the investor can no longer rely on trading in insolvency if the larger risky asset exposure of the IR strategy early in the investment time horizon results in poor performance.

- **Out-of-sample benchmark outperformance:** Arguably the most fundamental assumption underlying the analytical results is that the underlying process dynamics (3.13)-(3.14) are fully specified. In reality, lack of full knowledge regarding the underlying data generating process creates challenges for the out-of-sample and historical performance of the IR strategy, as our numerical results showed. The relatively lower reliance on the riskiest asset by the QD strategy early in the investment time horizon (Figures 5.2 and 5.3) improves its out-of-sample performance. In this sense, the IR strategy expresses much stronger convictions.
regarding the expected risk/return performance of the underlying assets than the QD strategy. The IR-strategy therefore retains some resemblance to the results from MV optimization, where “high conviction” strategies typically result in out-of-sample performance challenges (see for example Lehalle and Simon (2021)).

6 Conclusion

In this paper, we derived and compared two dynamics investment strategies for outperforming a benchmark using two widely-used objectives of practical interest to the investor, 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.

Using standard assumptions in the literature, we derived and analyzed the optimal investment strategies allowing for jumps in the risky asset processes, thereby extending known results for IR-optimal investment strategies, whereas all results associated with the QD objective are novel. We also presented closed-form comparison results of the strategies which generalize very effectively (in a qualitative sense) to settings where the assumptions underpinning the analytical results no longer hold. Specifically, by applying leverage and short-selling restrictions as well as discrete rebalancing to portfolios of multiple assets, we solved both problems numerically using a data-driven neural network (NN) approach. This also enabled us to compare the results on a more realistic basis in terms of what performance might be expected by an investor in practice.

A key property of our numerical approach, is that (i) in contrast to most previous work, we do not use the dynamic programming (DP) principle; (ii) we approximate the control directly using a NN, and (iii) we use time explicitly as a parameter in the NN. We have argued that this approach is inherently more efficient than any DP methodology, since a DP technique requires approximating a high dimensional conditional expectation (the performance criterion) in order to compute a low dimensional control, which is the object of interest. Using time directly as a parameter in the NN results in a parsimonious control approximation, and enforces desirable (limiting) continuity properties.

We also demonstrated that compared to the IR-optimal strategies, the resulting QD-optimal strategies are typically associated with less extreme positions in the assets with the most and the least risk, respectively, leading to improved benchmark outperformance in out-of-sample testing. Our results therefore demonstrate the attractiveness directly targeting the tracking difference using the proposed QD objective.

As noted in the Introduction, fairly complex objective functions have been formulated in the literature for benchmark outperformance. As discussed, in this paper we have made the deliberate choice to focus instead on objective functions targeting metrics which are valued by investors in practice (see the Introduction for references). A natural question is how the benchmark outperformance of the IR- and QD-optimal strategies presented here compare out-of-sample to that of the optimal strategies associated with more complex objective functions. We leave this for our future work.

References


Business Insider (2020). Most investment pros can’t beat the stock market, so why do everyday investors think they can win? *Business Insider, E. Rosenberg*.


CNBC (2020a). Some mutual funds are pricier than others. here’s when they may benefit investors. *CNBC.com, G. Iacurci*.


CNBC (2020b). Stock picking has a terrible track record, and it’s getting worse. *CNBC, B. Pisani*.


URL https://doi.org/10.1057/s41260-021-00219-z


Appendix A: Additional analytical results and selected proofs

A.1: Equivalent formulations of the IR maximization problem

Recall from Section 2 that the IR (2.3) is maximized by solving a mean-variance (MV) optimization problem with scalarization parameter $\rho$ (see Bajeux-Besnainou et al. (2013)), which is repeated here for convenience:

$$
\sup_{P \in A} \left\{ E^{P_{ir_0}, w_0}_P \left[ W(T) - \hat{W}(T) \right] - \rho \cdot Var^{P_{ir_0}, w_0}_P \left[ W(T) - \hat{W}(T) \right] \right\}, \quad \rho > 0. \tag{A.1}
$$

In this paper, we use the embedding technique of Li and Ng (2000); Zhou and Li (2000) to maximize the IR (2.3) by solving problem (2.5) instead, which is

$$
\inf_{P \in A} E^{P_{ir_0}, w_0}_P \left[ \left( W(T) - \left[ \hat{W}(T) + \gamma \right] \right)^2 \right], \quad \gamma > 0. \tag{A.2}
$$

We claimed without proof in Section 2 that the problem formulations (A.1) and (A.2) are equivalent, in the sense that for any $\rho > 0$ and associated optimal control $P^*_{ir} \in A$ maximizing (2.4), there exists a value of an embedding parameter $\gamma > 0$ such that the same control $P^*_{ir} \in A$ is also optimal for problem (A.2).

Of course, in the absence of optimization relative to a benchmark (for illustrative purposes, simply set $\hat{W}(T) \equiv 0$ in (A.1) and (A.2)), we recover the case of dynamic MV optimization, where the equivalence of the scalarization and embedding formulations are rigorously treated in Li and Ng (2000); Zhou and Li (2000).

In our setting, where the optimization is performed relative to a benchmark, Proposition A.1 below shows that the equivalence of (A.1) and (A.2) holds trivially given the above-mentioned dynamic MV results.

Proposition A.1. (Equivalent formulations of the IR maximization problem). Fix an admissible control set $A$, a value of $\rho > 0$, and assume that there exists control $P^*_{ir} \in A$ such that

$$
E^{P^*_{ir}, w_0}_P \left[ W^*(T) - \hat{W}(T) \right] - \rho \cdot Var^{P^*_{ir}, w_0}_P \left[ W^*(T) - \hat{W}(T) \right] \tag{A.3}
$$

$$
\sup_{P \in A} \left\{ E^{P_{ir_0}, w_0}_P \left[ W(T) - \hat{W}(T) \right] - \rho \cdot Var^{P_{ir_0}, w_0}_P \left[ W(T) - \hat{W}(T) \right] \right\} \tag{A.4}
$$

Then there exists a value of $\gamma > 0$ such that $P^*_{ir}$ also satisfies

$$
E^{P^*_{ir}, w_0}_P \left[ \left( W^*(T) - \hat{W}(T) - \gamma \right)^2 \right] = \inf_{P \in A} E^{P_{ir_0}, w_0}_P \left[ \left( W(T) - \hat{W}(T) - \gamma \right)^2 \right]. \tag{A.5}
$$

Proof. As in Section 3 we can define the auxiliary process $Y_{ir}(t) := W(t) - \hat{W}(t), \forall t \in (t_0, T]$ with $Y_{ir}(t_0) := y_0 = 0$. The proof of the claim then reduces to proving the equivalence of

$$
\sup_{P \in A} \left\{ E^{P_{ir_0}, y_0}_P \left[ Y_{ir}(T) \right] - \rho \cdot Var^{P_{ir_0}, y_0}_P \left[ Y_{ir}(T) \right] \right\}, \quad \inf_{P \in A} E^{P_{ir_0}, w_0}_P \left[ \left( Y_{ir}(T) - \gamma \right)^2 \right]. \tag{A.6}
$$

which is formally the same problem as considered by dynamic MV optimization. For a proof of (A.6), see Li and Ng (2000); Zhou and Li (2000).
We emphasize the following points regarding Proposition A.1

(i) Proposition A.1 does not place any restrictions on either the admissible control set \( \mathcal{A} \) or on the underlying wealth dynamics, so that we simply define \((A.2)\) as the "IR maximization problem" throughout this paper.

For a simple proof explaining why the equivalence holds in such generality, see Dang and Forsyth (2016).

(ii) We also emphasize that the particular equivalence relationship between \( \gamma \) and \( \rho \) is only of academic interest, and has very limited practical value. To clarify, note that for a fixed value of \( \rho > 0 \), problem \((A.1)\) could be solved (without resorting to dynamic programming - see Section 4.3) to yield \( \mathcal{P}^*_\gamma \in \mathcal{A} \).

The resulting control \( \mathcal{P}^*_\gamma \) could then be used to calculate \( E_{\mathcal{P}^*_\gamma, w_0}^{(0)} \left[ W^* (T) - \hat{W} (T) \right] \) and therefore also

\[
\gamma = \frac{1}{2 \rho} + E_{\mathcal{P}^*_\gamma, w_0}^{(0)} \left[ W^* (T) - \hat{W} (T) \right]. \tag{A.7}
\]

The results of Li and Ng (2000); Zhou and Li (2000) and Dang and Forsyth (2016) show that using \((A.7)\) as the value of \( \gamma \) and then solving problem \((A.2)\), we will obtain the identical optimal control \( \mathcal{P}^*_\gamma \). However, it is clear that this is of very little practical value. Instead, different values of \( \rho > 0 \) and \( \gamma > 0 \) can simply be interpreted to encode different relative "risk-return" trade-offs, with formulation \((A.2)\) providing the additional benefit of allowing a straightforward interpretation of \( \gamma \) as an "outperformance target" (see for example Vigna (2014)).

A.2: Proof of Proposition 4.2

In this proof, we consider only the QD problem \((4.10)\), since the proof for the IR problem \((4.9)\) 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 in the description of the dynamic programming (DP) solution approach below. To this end, let \( U (t) = (U_i (t) : i = 1, ..., N_\mathcal{A})^T \) and \( \hat{U} (t) = (\hat{U}_i (t) : i = 1, ..., N_\mathcal{A})^T \) 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 \( W (t) = \sum_{i=1}^{N_\mathcal{A}} u_i (t) \) and \( \hat{W} (t) = \sum_{i=1}^{N_\mathcal{A}} \hat{U}_i (t) \), respectively.

For an arbitrary admissible investor strategy \( \mathcal{P} \in \mathcal{A} \) with discrete rebalancing (see \((4.2)\)), define \( \mathcal{P}_n = \{ p \mid \mathcal{P} \in \mathcal{P}_n \} \) and \( \mathcal{P}_n = \{ p \mid \mathcal{P} \in \mathcal{P}_n \} \).

To solve the QD problem \((4.10)\) using DP, we define the performance criterion (see Oksendal and Sulem (2019)), which at time \( t \in [t_0, T) \) is given by the conditional expectation

\[
J \left( t, u^-, \hat{u}^-, \mathcal{P}_n \right) = E_{\mathcal{P}_n}^{W^-, \hat{W}^-} \left[ \left( W (T) - e^{\beta T} \hat{W} (T) \right)^2 \right] \left( U^-, \hat{U}^- \right) = \left( u^-, \hat{u}^- \right). \tag{A.8}
\]

where \( u^- = (u^-_1, ..., u^-_{N_\mathcal{A}})^T \) and \( \hat{u}^- = (\hat{u}^-_1, ..., \hat{u}^-_{N_\mathcal{A}})^T \). Note that \((A.8)\) 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 \( p_n := p(t_n, X (t_n)) \) and \( \mathcal{P}_n = \{ p \mid \mathcal{P} \} \) so that \( \mathcal{P}_n = \mathcal{P} \cup \mathcal{P}_n \). We also define \( \mathcal{A}_n = \{ p \mid \mathcal{P} \} \).

The investor and benchmark wealth immediately prior to the cash contribution at \( t_n \) is therefore given by \( W (t_n)^+ := w^- + \xi N (t_n) \) and \( \hat{W} (t_n)^+ := \hat{w}^- + \xi \hat{N} (t_n) \), 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 4.2, the investor can observe the benchmark allocation \( \hat{p}_n := \hat{p}(t_n, \hat{w}^+) \), while we have amount dynamics between rebalancing events, i.e. for \( t \in (t_n, t_{n+1}) \), given by

\[
\frac{dU (t)}{dt} = \left( \mu - \lambda \circ \kappa (t) \right) dt + \sigma \cdot dZ (t) + dN (t), \quad U (t)^+ = u^+ = w^- \cdot p_n, \tag{A.9}
\]

\[
\frac{d\hat{U} (t)}{dt} = \left( \mu - \lambda \circ \kappa (t) \right) dt + \sigma \cdot dZ (t) + d\hat{N} (t), \quad \hat{U} (t)^+ = \hat{u}^+ = \hat{w}^- \cdot \hat{p}_n. \tag{A.10}
\]

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

\[
V \left( t_n^-, w^-, \hat{w}^- \right) = \inf_{p_n \in \mathcal{A}_n} E_{p_n}^{W^-, \hat{W}^-} \left[ \left( W (T) - e^{\beta T} \hat{W} (T) \right)^2 \right] \left( W (t_n^-), \hat{W} (t_n^-) \right) = \left( w^-, \hat{w}^- \right) \tag{A.11}
\]

\[
= J \left( t_n^-, \hat{u}^-, \mathcal{P}_n = p_n^* \cup \mathcal{P}_n^* \right). \tag{A.12}
\]
where \( \mathcal{P}_n^* \in \mathcal{A}_n \) denotes the control realizing the infimum in (A.11), whereas the dependence of (A.11) and (A.12) on \( (w^-, \hat{w}^-) \) and \( (u^-, \hat{u}^-) \), 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

\[
V(T, w, \hat{w}) = V(T^-, w^- = \sum_{i=1}^{N_n} u_i^-, \hat{w}^- = \sum_{i=1}^{N_n} \hat{u}_i^-) = J(T^-, u^-, \hat{u}^-, \mathcal{P}_{n+1}^*) \equiv \emptyset = \left[ \left( \sum_{i=1}^{N_n} u_i^- \right) - e^{\beta T} \left( \sum_{i=1}^{N_n} \hat{u}_i^- \right) \right]^2,
\]

(A.13)

From (A.13), 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 \( t_n \in T \), and assume that the function \( J(t_{n+1}, u^-, \hat{u}^-, \mathcal{P}_{n+1}) \) 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.12), we have \( V(t_{n+1}, w^-, \hat{w}^-) = J(t_{n+1}, u^-, \hat{u}^-, \mathcal{P}_{n+1}^*) \), we do require the performance criterion \( J(t_{n+1}, \cdot, \cdot, \cdot) \) as a function of the amounts \( (u^-, \hat{u}^-) \), since \( J(t_{n+1}, u^-, \hat{u}^-, \mathcal{P}_{n+1}^*) \) will serve as the terminal condition to be satisfied by the (at this point, unknown) performance criterion function \( J(t, u, \hat{u}, \mathcal{P}_t), t \in (t_n, t_{n+1}) \). 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 3. The dynamic programming principle, definition (A.8) and dynamics (A.9)

(A.10) therefore imply that \( J(t, u, \hat{u}, \mathcal{P}_t) \) satisfies the following \((2N_n + 1)\)-dimensional PIDE on \( t \in (t_n, t_{n+1}) \)

\[
0 = J_t + \left[ \frac{\partial J}{\partial u} \right] \cdot \nabla J_u + \left[ \frac{\partial J}{\partial \hat{u}} \right] \cdot \nabla J_{\hat{u}} + \frac{1}{2} \text{tr} \left[ \text{diag} (u) \cdot \nabla^2 J_{uu} \right] + \frac{1}{2} \text{tr} \left[ \text{diag}(\hat{u}) \cdot \nabla^2 J_{\hat{u}\hat{u}} \right] + \text{tr} \left[ \text{diag}(u) \cdot \nabla^2 J_{u\hat{u}} \right] - \sum_{i=1}^{N_n} \lambda_i \cdot J(t, u, \hat{u})
\]

\[
+ \sum_{i=1}^{N_n} \lambda_i \int_0^\infty \left[ J(t, u + i \xi - 1) \cdot e_i \cdot \hat{u} + \hat{u} + \xi - 1 \cdot e_i \right] f_{\xi_i}(\xi_i) d\xi_i.
\]

(A.14)

In (A.14), \( \text{tr} (\cdot) \) denotes the trace of a matrix, \( \text{diag}(v) \) denotes the diagonal matrix with vector \( v \) on the main diagonal, \( e_i \in \mathbb{R}^{N_n} \) is the \( i \)-th standard basis vector in \( \mathbb{R}^{N_n} \), and we have gradients \( \nabla J_u = \left[ \frac{\partial J}{\partial u} : i = 1, \ldots, N_n \right]^T \) and \( \nabla J_{\hat{u}} = \left[ \frac{\partial J}{\partial \hat{u}} : i = 1, \ldots, N_n \right]^T \), as well as matrices of second derivatives \( \nabla^2 J_{uu} = \left( \frac{\partial^2 J}{\partial u_i \partial u_j} \right)_{i,j=1,\ldots,N_n} \)

\( \nabla^2 J_{\hat{u}\hat{u}} = \left( \frac{\partial^2 J}{\partial \hat{u}_i \partial \hat{u}_j} \right)_{i,j=1,\ldots,N_n} \),

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

Let \( J \) denote the lower semi-continuous envelope of the function \( J \) obtained by solving (A.14). 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

\[
p_n^* = p^* \left(t_n, w^+, \hat{w}^+\right) = \arg\min_{p_n \in \mathcal{P}_n} J \left(t_n, u^+ = w^+ \cdot p_n, \hat{u}^+ = \hat{w}^+ \cdot \hat{p}_n, \mathcal{P}_n = p_n \cup \mathcal{P}_{n+1}^* \right),
\]

(A.15)

with \( w^+ = \sum_{i=1}^{N_n} u_i^- + q(t_n) \) and \( \hat{w}^+ = \sum_{i=1}^{N_n} \hat{u}_i^- + q(t_n) \). 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

\[
V \left(t_n, w^-, \hat{w}^-\right) = J(t_n, u^+ = w^+ \cdot p_n^*, \hat{u}^+ = \hat{w}^+ \cdot \hat{p}_n, \mathcal{P}_n^* = p_n^* \cup \mathcal{P}_{n+1}^*)
\]

(A.16)

\[
= J(t_n, u^-, \hat{u}^- \cdot \mathcal{P}_n^*),
\]

(A.17)

where \( p_n^* \) is given by (A.15).

The results (A.13) and (A.17) 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 $(2N_0 + 1)$-dimensional performance criterion $J : \mathbb{R}^{(2N_0 + 1)} \to \mathbb{R}$, obtained in this case by solving the PIDE (A.14).

A.3: Additional analytical comparison results

As a supplement to the investment comparison results in Section 3 we present additional analytical comparison results which assist in rigorously understanding some of the observations regarding the numerical results of Section 3.

As noted in Section 3, there are many possibilities for comparing the IR- and QD-optimal results. Since the strategies are compared in Section 3 on the basis of equal expectation of terminal wealth (see (5.1)), we formally introduce Assumption A.2 outlining the basis of the comparison of the subsequent analytical results. Note that these results are again all derived within the setting of Section 3, and in particular require that Assumptions 3.1, 3.2 and wealth dynamics (3.13)-(3.14) hold.

**Assumption A.2.** (Expected value target for terminal wealth) Assume that Assumption 3.1, Assumption 3.2, and wealth dynamics (3.13)-(3.14) hold. Suppose that the benchmark investment strategy, given by the fractions of wealth in the risky assets $\hat{q}(t, \hat{W}(t))$, results in an expected value of benchmark terminal wealth satisfying

$$E_{\hat{q}(tW_0)}^\gamma [\hat{W}(T)] := \mathcal{K}, \quad \text{where } \mathcal{K} > \omega_0 e^{rT}. \quad (A.18)$$

We assume that, regardless of whether the IR- or QD-optimal strategy is implemented, the investor wishes to achieve a given multiple of the benchmark expected wealth (A.18). In other words, the investor chooses the parameters $\gamma = \gamma_{ir}^\varepsilon$ in the IR($\gamma$) problem and $\beta = \beta_{qd}^\varepsilon$ in the QD($\beta$) problem such that the associated IR- and QD-optimal strategies $\varrho_{ir}^\varepsilon$ and $\varrho_{qd}^\varepsilon$, respectively, each result in the desired expected value of terminal wealth,

$$E_{\varrho_{ir}^\varepsilon}^\gamma [W_{ir}^\varepsilon(T)] = E_{\varrho_{qd}^\varepsilon}^\gamma [W_{qd}^\varepsilon(T)] = \mathcal{E} = e^{\beta T} \mathcal{K}, \quad \text{for some } \beta > 0. \quad (A.19)$$

The value of $\mathcal{E}$ (A.19) will be referred to as the expected value target for terminal wealth.

Note that in this setting, where we do have closed-form solutions, the values of $\gamma = \gamma_{ir}^\varepsilon$ and $\beta = \beta_{qd}^\varepsilon$ achieving (A.19) can be derived analytically, as the following lemma shows.

**Lemma A.3.** (Analytical values $\gamma$ and $\beta$ achieving expected value target) Suppose that Assumption 3.1, Assumption 3.2 and wealth dynamics (3.13)-(3.14) hold. The optimal controls of problems IR($\gamma = \gamma_{ir}^\varepsilon$) and QD($\beta = \beta_{qd}^\varepsilon$) achieve the required expected value target $E_{\varrho_{ir}^\varepsilon}^{\gamma_{ir}} [W_{ir}^\varepsilon(T)] = \mathcal{E}, j \in \{ir, qd\}$, provided $\gamma_{ir}^\varepsilon$ and $\beta_{qd}^\varepsilon$ are given respectively by

$$\gamma_{ir}^\varepsilon = \frac{(\mathcal{E} - \mathcal{K})}{(1 - e^{-rT})}, \quad \text{and} \quad \beta_{qd}^\varepsilon = \frac{1}{T} \log \left[ \frac{(\mathcal{E} - \eta \langle 1 - e^{-rT} \rangle + \omega_0)}{\mathcal{K} - \eta \langle 1 - e^{-rT} \rangle + \omega_0} e^{\langle r - \eta \rangle T} \right], \quad (A.20)$$

where $\eta$ is given by (3.12).

**Proof.** Follows from re-arranging (3.24) and (3.41), and using (A.19).

Note that the results of Lemma A.3 allow for jumps in the risky asset processes.

In Subsection 5.3 we presented the results of implementing the closed-form optimal controls in the particular case where a constant proportion benchmark strategy is considered. For reasons as outlined in Section 3, our results focused on a diversified stock index representing the risky asset basket (the “risky asset”), as well as a risk-free asset. To provide further analytical explanations for the particular results observed in Subsection 5.3 we therefore present the following closed-form results for the specific case of 2 assets (a single risky asset and a risk-free asset) in combination with a constant proportion benchmark strategy.

**Lemma A.4.** (Constant proportion benchmark, $N_0 = 2$) Suppose that Assumption 3.1, Assumption 3.2 and wealth dynamics (3.14) with a single risky asset are applicable, so we have $N^r_0 = 1$ and $\mu = \mu > r$. If the benchmark strategy is a constant proportion strategy with $\hat{q}(t, \hat{W}(t)) \equiv \hat{q} > 0$ for $t \in [t_0 = 0, T]$, then the
benchmark terminal wealth has expectation $E_{p}^{\hat{q},w_{0}}[\hat{W}(T)] \equiv K = K(T)$, where $t \to K(t)$ is given in terms $\hat{\varrho}$ by

$$K(t) := E_{p}^{\hat{q},w_{0}}[\hat{W}(t)] = w_{0}e^{(r-(\mu-r)\hat{\varrho})t} + \frac{q}{r + (\mu - r)\hat{\varrho}} \cdot (e^{(r-(\mu-r)\hat{\varrho})t} - 1), t \in [t_{0} = 0,T]. \quad (A.21)$$

In the particular case where contributions are zero ($q = 0$), a given value of $E_{p}^{\hat{q},w_{0}}[\hat{W}(T)] \equiv K$ can therefore be achieved by choosing the constant $\hat{\varrho}$ according to

$$\hat{\varrho} = \frac{1}{(\mu - r)T} \log \left( \frac{K}{w_{0}e^{rT}} \right), \quad (if \ q = 0). \quad (A.22)$$

Proof. Follows from applying standard analysis techniques to (3.14) after setting $\hat{\varrho}(t,\hat{W}(t)) := \hat{\varrho} > 0$.

The preceding results enable the main comparison result, given in Theorem A.5 below.

**Theorem A.5.** (QD-optimal vs. IR-optimal strategies, $N_{a} = 2$: Risky asset exposure over time) Suppose the following assumptions hold: (i) Assumption 3.2 and wealth dynamics (3.13)-(3.14) with a single risky asset ($N_{a} = 1$); (ii) the investor compares investment strategies on the basis of Assumption A.2; (iii) contributions are zero ($q = 0$); (iv) the benchmark strategy is a constant proportion strategy with $\varrho(t,\hat{W}(t)) \equiv \hat{\varrho} > 0$ for $t \in [t_{0},T]$. Note that $X_{ir}^{\ast}(t) := (W_{ir}^{\ast}(t),\hat{W}(t),\hat{\varrho})$ and $X_{qd}^{\ast}(t) := (W_{qd}^{\ast}(t),\hat{W}(t),\hat{\varrho})$.

Then, at inception $t = t_{0} = 0$, the IR-optimal strategy $X_{ir}^{\ast}(t_{0}) := \varrho_{ir}^{\ast}(t_{0};X_{ir}^{\ast}(t_{0}))$ requires a larger investment in the single risky asset than the QD-optimal strategy $X_{qd}^{\ast}(t_{0}) := \varrho_{qd}^{\ast}(t_{0};X_{qd}^{\ast}(t_{0}))$.

**Proof.** Since we are assuming that $\hat{\varrho} > 0$ (see (3.12)) satisfies $\eta > r$, then the function

$$f(t) := E_{p}^{\hat{q},w_{0}}[\varrho_{ir}^{\ast}(t,X_{ir}^{\ast}(T)) \cdot W_{ir}^{\ast}(T)] - E_{p}^{\hat{q},w_{0}}[\varrho_{qd}^{\ast}(t,X_{qd}^{\ast}(T)) \cdot W_{qd}^{\ast}(T)] \quad (A.25)$$

is monotonically decreasing on $t \in [t_{0},T]$.

Note that the additional requirement $\eta > r$ leading to (A.25) is indeed satisfied in the case of typical process parameters, including by the parameters in Table 5.3.

Theorem A.5 suggests that in order to achieve the same expected value of terminal wealth, the IR-optimal strategy requires comparatively more extreme positions in the risky (riskiest) asset than the QD strategy, and in that qualitative sense the IR strategy is expected to be less diversified than the QD strategy at some points during the investment time horizon. Specifically, compared to the QD strategy, the IR strategy relies on a larger investment in the riskiest asset early in the investment time horizon, and once the desired outperformance becomes increasingly likely, the exposure to the riskiest asset is expected to be reduced to a level below that of the QD strategy.
The results of Theorem A.5 are illustrated in Figure A.1. Specifically, Figure A.1(a) shows the expected (average) amount invested in the risky asset over time according to each strategy, while Figure A.1(b) shows that the difference \( f(t) \) as per (A.25) remains monotonically decreasing in this case where there are nonzero contributions to the portfolio (in contrast to the assumptions of Theorem A.5).

Figure A.1: Analytical solutions, no constraints, investor portfolio P0, benchmark BM0: Expected amount \( t \to E_{\tilde{\mu}}^{x_{0},w_{0}}[\tilde{g}^{x}(t), X_{j}^{x}(t)] \cdot W_{j}^{x}(t), j \in \{ir, qd\} \) in the risky asset, as well as the resulting difference \( f(t) \) as per (A.25). 10\(^6\) Monte Carlo simulations, \( \mathcal{E} = 400 \) in (A.19).

Note that the qualitative implications of Theorem A.5 hold even if the underlying assumptions are relaxed and the problems are considered in a more realistic setting (see Section 5), such as when multiple investment constraints are applied and contributions are nonzero. As a result, Theorem A.5 is valuable both for confirming the numerical results of Section 4.

Appendix B: Additional numerical results

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

B.1: Cumulative distribution functions (CDFs)

As a supplement to the results in Subsection 5.3, Figure B.1 illustrates CDFs corresponding to the PDFs presented in Figure 5.1. 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 Van Staden et al. (2021a) for a definition and discussion).

However, the situation changes when realistic 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 5.4 and 5.5 (Subsection 5.4). 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.

B.2: Portfolio P1, benchmark BM1, data set DS2

As a supplement to the results of Subsection 5.4, Table B.1 presents results for using investor portfolio P1 to outperform benchmark BM1 on data set DS2. Compared to Table 5.8 which is based on DS1, we observe that the qualitative aspects of the comparative performance of the IR and QD-optimal strategies also hold on data set DS2.
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 $W^j(T)$, respectively, as well as the ratio $W^j(T)/\hat{W}(T)$, for $j \in \{\text{ir, qd}\}$. $10^6$ Monte Carlo simulations, $E^* = 400$ in (5.1).

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 $W^j(T)$, respectively, as well as the ratio $W^j(T)/\hat{W}(T)$, for $j \in \{\text{ir, qd}\}$.

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 $W^j(T)$, respectively, as well as the ratio $W^j(T)/\hat{W}(T)$, for $j \in \{\text{ir, qd}\}$.

Table B.1: Numerical solutions, with constraints, investor portfolio P1, benchmark BM1, data set DS2, annual rebalancing: Training and testing results for mean terminal wealth $E = 430 (\hat{\beta} \simeq 1.7\%$ in (5.1)) on the training data.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$W(T)$</td>
<td>$W^j(T)$</td>
</tr>
<tr>
<td></td>
<td>BM1</td>
<td>IR</td>
</tr>
<tr>
<td>Mean</td>
<td>364</td>
<td>430</td>
</tr>
<tr>
<td>CExp 5%</td>
<td>212</td>
<td>249</td>
</tr>
<tr>
<td>5th pctile</td>
<td>235</td>
<td>286</td>
</tr>
<tr>
<td>Median</td>
<td>354</td>
<td>422</td>
</tr>
<tr>
<td>95th pctile</td>
<td>331</td>
<td>601</td>
</tr>
<tr>
<td>Prob. underp.</td>
<td>1.22%</td>
<td>1.05%</td>
</tr>
</tbody>
</table>
Appendix C: Neural network (NN) approach - additional details

In this appendix, we discuss a number of topics related to the neural network (NN) numerical solution methodology discussed in Section 4.3, including technical and implementation details, training and testing, as well as ground truth comparison results.

C.1: Technical details: NN approach

In this subsection, we give an overview of the technical details of the NN approach - see Subsection 4.3 for some intuition regarding the main underlying ideas.

Consider a fully-connected, feed-forward NN with \((L - 1)\) hidden layers (see for example Goodfellow et al. 2016). The NN layers are indexed by \(\ell \in \{0, \ldots, L\}\), where \(\ell = 0\) and \(\ell = L\) denote the input and output layers, respectively. Let \(\eta_\ell \in \mathbb{N}\) denote the number of nodes in layer \(\ell\). With the exception of the input layer, each layer \(\ell \in \{1, \ldots, L\}\) is associated with a weights matrix \(x^{[\ell]} \in \mathbb{R}^{\eta_\ell \times \eta_{\ell - 1}}\) into the layer, an optional bias vector \(b^{[\ell]} \in \mathbb{R}^{\eta_\ell}\), as well as an activation function \(\phi^{[\ell]} : \mathbb{R}^{\eta_{\ell - 1}} \to \mathbb{R}^{\eta_\ell}\) which is applied to the weighted inputs into the layer.

Specifically, for a given node \(i\) in layer \(\ell\), \(x_{i}^{[\ell]} \in x^{[\ell]}\) denotes the weight applied to the output from node \(k\) in layer \((\ell - 1)\) into node \(i\) in layer \(\ell\), and \(b_i^{[\ell]} \in b^{[\ell]}\) denotes the bias applied by node \(i\) in layer \(\ell\). Therefore, after applying the activation function \(\phi^{[\ell]}\), the output of node \(i\) in layer \(\ell\) is given by the component \(a_i^{[\ell]} \in a^{[\ell]}\), where

\[
a_i^{[\ell]} := a_i^{[\ell]} \left( \sum_{k=1}^{\eta_{\ell - 1}} x_{ki}^{[\ell - 1]} a_k^{[\ell - 1]} \right) + b_i^{[\ell]}, \quad i = 1, \ldots, \eta_\ell, \quad \ell \in \{1, \ldots, L\}.
\]

The output of NN layer \(\ell \in \{1, \ldots, L\}\) can therefore be abbreviated as the vector \(a^{[\ell]} \in \mathbb{R}^{\eta_\ell}\) using the following short-hand notation,

\[
a^{[\ell]} := \left( a_i^{[\ell]} : i = 1, \ldots, \eta_\ell \right) = a^{[\ell]} \left( x^{[\ell - 1]} a^{[\ell - 1]} \right) + b^{[\ell]}, \quad \ell \in \{1, \ldots, L\},
\]

where we emphasize that \((C.2)\) is only an abbreviation of \((C.1)\).

The parameter vector of the NN, which consists of all weights and biases, is denoted by \(\theta \in \mathbb{R}^{n_\theta}\), where \(n_\theta \in \mathbb{N}\) denotes the total number of weights and biases. In other words, the weights matrices \(\{x^{[\ell]} : \ell = 1, \ldots, L\}\) and optional bias vectors \(\{b^{[\ell]} : \ell = 1, \ldots, L\}\) are flattened\(^1\) into a single vector \(\theta = (\theta_1, \ldots, \theta_{n_\theta})\), where each \(\theta_\ell \in \theta\) can be uniquely mapped to a single weight or bias in some layer.

Note that no activation function is applied at the input layer \((\ell = 0)\), so that the \(\theta_0\) output values of the input layer corresponds to feature (input) vector of the NN, which will be denoted by \(\phi \in \mathbb{R}^{n_\phi}\). Recalling that \(n_c\) is the number of nodes in the output layer \((\ell = L)\) and setting the bias vectors \(b^{[\ell]} \equiv 0\) for convenience, the NN can therefore be written as a single function \(F(\phi; \theta) : \mathbb{R}^{n_\phi} \to \mathbb{R}^{n_c}\), where

\[
F(\phi; \theta) := (F_1(\phi; \theta), \ldots, F_{n_c}(\phi; \theta))
\]

\[
= \left( a^{[L]}, \ldots, a^{[L]} \right), \quad \left( a^{[L]} \right) \text{ is the activation function of the output layer},
\]

\[
= \left( a^{[L]} \circ \left( x^{[L - 1]} a^{[L - 1]} \right) \circ \cdots \circ \left( x^{[1]} a^{[1]} \right) \circ \phi, \quad \phi \in \mathbb{R}^{n_\phi}, \theta \in \mathbb{R}^{n_\theta},
\]

where \(\circ\) denotes the composition of two functions in this appendix, and not the Hadamard product.

For subsequent reference, we highlight that the output of the \(i\)th node in the output layer (i.e. the \(i\)th output of the NN) is given by \(F_i(\phi; \theta) = a_i^{[L]}\).

Given this standard fully-connected, feedforward NN formulation, we introduce the following assumption.

Assumption C.1. (NN approach) Suppose the information taken into account by the investor strategy \(P\) at time \(t \in [t_0 = 0, T]\) is given by the vector \(X(t) \in \mathbb{R}^{n_X}\) with minimal form \(X(t) = (\hat{W}(t), \hat{W}(t))\), so that \(n_X \geq 2, n_X \in \mathbb{N}\). We assume that the investor is subject to the investment constraints of no short-selling and no leverage, and therefore requires \(P \in A\) where \(A\) is given by \((4.7)\). In addition, we assume the following.

(i) The investor strategy or control \(P\) is a continuous function \(P : [t_0, T] \times \mathbb{R}^{n_X} \to [t_0, T] \times \mathbb{Z}\), where \(\mathbb{Z} \subset \mathbb{R}^{n_a}\) is as per \((4.8)\) with \(n_a\) being the number of assets. If the portfolio is rebalanced only at discrete time

---

\(^1\)For example, in Python this could be done with the “numpy.ndarray.flatten” method.
Intervals, the investment strategy can be found by evaluating this continuous function at discrete time intervals, i.e., \((t_n, X(t_n)) \rightarrow \mathcal{P}(t_n, X(t_n)) = p(t_n, X(t_n)), t_n \in T.\)

(ii) The NN has \(\eta_0 = \eta_X + 1\) input nodes, with feature (input) vectors \(\phi \in \mathbb{R}^{\eta_0}\) of the form
\[
\phi(t) := (t,X(t)) \, .
\] (C.5)

In particular, we emphasize that the time \(t \in [t_0, T]\) is an input into the NN as per \([C.5]\).

(iii) The NN has \(\eta_L = N_a\) output nodes, with node \(i\) being associated with asset \(i \in \{1, \ldots, N_a = \eta_L\}\).

(iv) The NN output layer \((\ell = L)\) uses the softmax activation function, so that \(a^{[L]} = \left( a^{[L]}_i : i = 1, \ldots, \eta_L \right)\) is given by
\[
a^{[L]}_i = \frac{\exp \left\{ z_i^{[L]} \right\}}{\sum_{m=1}^{\eta_L} \exp \left\{ z_m^{[L]} \right\}}, \quad i = 1, \ldots, \eta_L, \text{ with } \eta_L = N_a, \quad (C.6)
\]
where \(z_i^{[L]} = \sum_{k=1}^{\eta_{L-1}} \sum_{p=1}^{N_a} \phi_{ik}^{[L-1]} + b_i^{[L]}\) is the weighted input into output node \(i\). Note that \((C.6)\) implies that for any input vector \(\phi(t) = (t,X(t))\), the NN generates outputs as per \((C.3)\) given by \(F(\phi; \theta) = a^{[L]} \in \mathcal{Z}\), where \(\mathcal{Z}\) is as per \((C.3)\).

(v) The continuous function \(\mathcal{P}\) can be represented by a NN \(F(\phi; \theta)\) satisfying (ii)-(iv) above, provided \(F(\phi; \theta)\) has sufficient width (i.e. sufficiently large \(\eta_\ell, \forall \ell = 1, \ldots, L - 1\)) and sufficient depth (i.e. sufficiently large \(L\)),
\[
\mathcal{P}(t,X(t)) = p(t,X(t)) = F(\phi(t); \theta) , \quad \forall \phi(t) = (t,X(t)) \in [t_0, T] \times \mathbb{R}^{\eta_X} . \quad (C.7)
\]
In particular, considering \((C.3)\) and \((C.4)\) together with (i)-(iv) above, \((C.7)\) means that we interpret the output of the \(i\)th node of the NN as the fraction of wealth invested in the \(i\)th asset,
\[
p_i(t_n, X(t_n)) \equiv F_i(\phi(t_n); \theta) , \quad \forall t_n \in T, i = 1, \ldots, N_a . \quad (C.8)
\]
Figure C.1 illustrates the structure of the NN as described in Assumption C.1.

\[\textbf{Figure C.1: Illustration of the structure of the NN used to model the control (investment strategy) as per Assumption C.1} \]

Note the treatment of the rebalancing time as a feature (input).

It should be emphasized that the representation aspect of Assumption C.1(v) is included largely as a matter of convenience, since this should be expected to follow from Assumption C.1(i)-(iv) given the universal approximation results of Diamond and Fomenko (1993); Funahashi (1989); Hornik (1991); Hornik et al. (1989); Leshno et al. (1993); Sonoda and Murata (2017). However, these universal approximation results typically require mild assumptions on for example the properties of the activation functions C.1 with any of the standard activation functions encountered in the literature (see for example Goodfellow et al. (2016)) being acceptable. Instead of repeating these requirements, we simply include \((C.7)\) as an assumption.
We note that Assumption [C.1] does not specify any details regarding hyperparameters such as the width and depth of the NN to be used, nor any details on the choice of activation functions for the hidden layers. These hyperparameters are to be tailored to portfolio optimization problem at hand, and in the case of the QD and CD problems a discussion can be found in Appendix [C.5] below. Here we simply note that if there are ground truth results available under certain assumptions (in this case, the analytical solutions of Section 3), the NN should be able to produce those same results under similar assumptions to an acceptable level of accuracy. If ground truth solutions are not available, numerical tests typically show that if \( \eta \) and \( \zeta \) is sufficiently large, investment results stabilize in the sense that the generalization error (see Goodfellow et al. (2016)) shows no further improvement from an increase in the width or depth of the NN.

The main consequence of Assumption [C.1] is that it allows to us to solve a portfolio optimization problems (4.9)-(4.10) subject to admissibility set \( A \) given by (4.7) as a single, unconstrained optimization problem, without having to resort to dynamic programming. This follows since there is only a single NN \( F (\phi; \theta) \) with a time-independent parameter vector \( \theta \) to be trained, and by (C.6) and (C.8) the output is automatically in the set \( Z \) as per (4.8), thus no further constraints need to be applied. Specifically, this results in the approximations noted in (4.12) repeated here for convenience,

\[
(\text{IR}(\lambda)) : \quad \inf_{\theta \in \mathbb{R}^{2n}} E_{F(\theta)} \left[ (W (T; \theta) - \hat{W} (T) + \gamma)^2 \right], \quad \gamma > 0, \quad (C.9)
\]

\[
(\text{QD} (\beta)) : \quad \inf_{\theta \in \mathbb{R}^{2n}} E_{F(\theta)} \left[ (W (T; \theta) - e^{\beta T} \hat{W} (T))^2 \right], \quad \beta > 0, \quad (C.10)
\]

where \( F (\cdot; \theta) \) is a NN of the form (C.4) subject to the requirements of Assumption [C.1]. In (C.9), the notation \( W (t_n; \theta) := W (t_n) \) is used to highlight the dependence of the investor's wealth on the investment strategy (NN) parameter vector \( \theta \in \mathbb{R}^{2n} \) through the control \( P = F \) (see (4.3), (4.2) and (C.7)). In particular, given Assumption [C.1] the investor wealth dynamics (4.3) can now be written as

\[
W (t_{n+1}; \theta) = W (t_n; \theta) + q (t_n) \cdot \sum_{i=1}^{N} F_i (\phi (t_n); \theta) \cdot [1 + R_i (t_n)], \quad n = 0, ..., N_{rb} - 1, \quad (C.11)
\]

while the benchmark wealth dynamics (4.4) remain unchanged.

C.2: Training and testing the NN

We now discuss the solution of problems (4.9)-(4.10) in detail, which corresponds to the training of the NN \( F (\cdot; \theta) \) on a training data set to obtain \( \theta^*_w (\lambda) \) and \( \theta^*_p (\beta) \), which respectively achieve the minimum in (C.9) and (C.10), for given values of \( \lambda \) and \( \beta \). The resulting NNs \( F (\cdot; \theta^*_k) \), \( k \in \{ ir, qd \} \), corresponding to the investor’s optimal investment strategies as per (C.7), are then applied to testing data to assess out-of-sample performance.

As a concrete example, the training and testing data of \( F (\cdot; \theta) \) in the case of the minimal form \( X (t) = (W (t), \hat{W} (t)) \) together with a known benchmark strategy (for example in the case of the constant proportion benchmark strategies considered in Section 5) consists of the sets \( Y \in \mathbb{R}^{N_d \times N_{rb} \times N_a} \) and \( Y^{test} \in \mathbb{R}^{N_d^{test} \times N_{rb} \times N_a} \), respectively. These sets each consist of \( N_d \in \mathbb{N} \) and \( N_d^{test} \in \mathbb{N} \) sample paths of (joint) returns, respectively, of the \( N_a \) candidate assets observed over \( N_{rb} \) time intervals of length \( \Delta t = T/N_{rb} \). Since the data structure of \( Y \) and \( Y^{test} \) are similar, we focus only on the training dataset \( Y \). In particular, let

\[
Y^{(j)} (t_n) = 1 + R^{(j)} (t_n), \quad j \in \{ 1, ..., N_d \}, \quad n \in \{ 0, ..., N_{rb} - 1 \}, \quad i \in \{ 1, ..., N_a \}, \quad (C.12)
\]

where \( R^{(j)} (t_n) \) denotes the (possibly inflation-adjusted) return along sample path \( j \in \{ 1, ..., N_d \} \) for asset \( i \in \{ 1, ..., N_a \} \) over the time period \( [t_n, t_{n+1}] \), where \( n \in \{ 0, ..., N_{rb} - 1 \} \). A sample path \( j \in \{ 1, ..., N_d \} \) of joint returns is simply the subset \( Y^{(j)} \subset Y \), where

\[
Y^{(j)} = \left\{ Y^{(j)} (t_n) : n = 0, ..., N_{rb} - 1, \quad i = 1, ..., N_a \right\} \in \mathbb{R}^{N_{rb} \times N_a}, \quad j \in \{ 1, ..., N_d \}. \quad (C.13)
\]

In principle, any market data generator (or economic scenario generator) can be used to generate \( Y \) and

---

11Since Assumption [C.1] implies that the NN only models investment strategies that do not allow short-selling or leverage, while the analytical solutions of Section 3 place no such restrictions on the investment strategies, the ground truth comparison should be constructed carefully. For example, below in Appendix [C.3] we consider an investment scenario where the short-selling and leverage constraints are not binding, and thus the NN satisfying Assumption [C.1] can approximate the analytical solution without difficulty.
where, for a fixed sample path $Y^{(j)}$ as per (C.13), the wealth dynamics (C.11) and (4.4) are respectively given by

\begin{align}
W^{(j)}(t_{n+1}; \theta) &= \left[ W^{(j)}(t_n; \theta) + q(t_n) \right] \cdot \sum_{i=1}^{N_d} F_i \left( \phi^{(j)}(t_n); \theta \right) \cdot Y_i^{(j)}(t_n), \\
\hat{W}^{(j)}(t_{n+1}) &= \left[ \hat{W}^{(j)}(t_n) + q(t_n) \right] \cdot \sum_{i=1}^{N_d} p_i^{(j)} \left( t_n, \hat{X}^{(j)}(t_n) \right) \cdot Y_i^{(j)}(t_n),
\end{align}

for $n = 0, \ldots, N_{rb} - 1$ and $j \in \{1, \ldots, N_d\}$. In (C.16), $\phi^{(j)}(t_n) = (t_n, X^{(j)}(t_n); \theta) \in \mathbb{R}^{n_0}$, and we use the superscript $(j)$ throughout (C.14)-(C.17) to highlight the dependence of quantities on the sample path in the underlying data. We emphasize again that the parameter vector $\theta \in \mathbb{R}^{n_0}$ of the NN does not depend on the rebalancing time $t_n$ or on the sample path $j$.

For optimization purposes, or equivalently training the neural network, we use stochastic gradient descent to solve (C.14)-(C.15). Focusing on (C.14) for concreteness, with $\nabla_\theta$ denoting the gradient with respect to the NN parameter vector $\theta$, we have

\begin{align}
\nabla_\theta G_{ir}(\theta; Y) &= \frac{2}{N_d} \sum_{j=1}^{N_d} \left[ W^{(j)}(T; \theta) - \left[ \hat{W}^{(j)}(T) + \gamma \right] \right] \cdot \nabla_\theta W^{(j)}(T; \theta),
\end{align}

where the gradient $\nabla_\theta W^{(j)}(T; \theta)$ can be obtained using (C.11) and iterative computation (i.e. backpropagation) over $t_n \in T$. Specifically, starting from

\begin{align}
\nabla_\theta W^{(j)}(t_0; \theta) &= [w_0 + q(t_0)] \cdot \sum_{i=1}^{N_d} Y_i^{(j)}(t_0) \cdot \nabla_\theta F_i \left( \phi^{(j)}(t_0); \theta \right),
\end{align}

we have, for $n = 1, \ldots, N_{rb}$, the following recursion,

\begin{align}
\nabla_\theta W^{(j)}(t_{n-1}; \theta) &= \sum_{i=1}^{N_d} F_i \left( \phi^{(j)}(t_{n-1}); \theta \right) \cdot Y_i^{(j)}(t_{n-1}) \cdot \nabla_\theta W^{(j)}(t_{n-1}; \theta) \\
&\quad + \left[ W^{(j)}(t_{n-1}; \theta) + q(t_{n-1}) \right] \cdot \sum_{i=1}^{N_d} Y_i^{(j)}(t_{n-1}) \cdot \nabla_\theta F_i \left( \phi^{(j)}(t_{n-1}); \theta \right).
\end{align}

In (C.19) and (C.20), the gradients $\nabla_\theta F_i$ (i.e. the gradients of the NN outputs with respect to the NN parameter vector $\theta$) can be calculated using backpropagation.

Once the training of the NNs in (C.14)-(C.15) are completed, the resulting IR- and QD-optimal investment strategies $F^k(\cdot; \theta^k)$, $k \in \{ir, qd\}$ can be tested by implementing these strategies on the testing data set $Y^{test}$ with $N_{d}^{test}$ sample paths of returns. Using the QD strategy as an example, the dynamics (C.16)-(C.17) can be used together with the optimal vector $\theta^{*_{qd}}(\beta)$ to obtain quantities such as the QD($\beta$)-optimal terminal wealth, $W^{*_{qd}}(T) := W^{*_{qd}}(t_{N_{rb}}; \theta^{*_{qd}}(\beta))$ along each sample path $j$ in $Y$ or $Y^{test}$. Additionally, ratios such as $W^{*_{qd}}(T)/\hat{W}^{(j)}(T)$ can also be calculated without difficulty for each path. Using $Y^{test}$ as an example, the sets $\{W^{*_{ir}}(T) : j = 1, \ldots, N_{d}^{test}\}$ and $\{W^{*_{qd}}(T) : j = 1, \ldots, N_{d}^{test}\}$ can then be used to compare the terminal wealth distributions of the resulting strategies on the testing data set, while $\{W^{*_{ir}}(T)/\hat{W}^{(j)}(T) : j = 1, \ldots, N_{d}^{test}\}$
and \( \left\{ W^{q(j)}(T) / W^{(j)}(T) : j = 1, ..., N_d^{\text{test}} \right\} \) can be used to obtain benchmark outperformance statistics associated with each strategy on the testing data set. Of course, similar calculations can also be performed on the training data set \( Y \). An illustration of these results in a realistic investment scenario is presented in Section 5.

C.3: Implementation parameters and gradient descent algorithm

For the purposes of the training of the NNs in (C.14)-(C.15), we used the Gadam algorithm of Granziol et al. (2020), which combines the Adam algorithm (Kingma and Ba 2015)) with tail iterate averaging for improved convergence properties and variance reduction (Mucke et al. 2019; Neu and Rosasco 2018; Polyak and Juditsky 1992)). Numerical experiments showed that the default algorithm parameters of Kingma and Ba (2015) 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; essentially identical results are obtained when the NN is trained multiple times on the same underlying data.

In for the structure of the NN, note that the input and output layer are constrained by the structural requirements as outlined in Assumption C.1. For the input layer, our results focused on the minimal form of \( X \) as per Assumption C.1 and thus three features were used (time, investor wealth, benchmark wealth). This ensures that the results are realistic while remaining robust against possible overfitting. As per Assumption C.1, the number of nodes in the output layer corresponds to the number of assets, \( N_a \).

However, note that Assumption C.1 imposes no requirements on the depth or width of the NN, in other words the number of hidden layers and number of nodes in each hidden layer are hyperparameters to be chosen according to the complexity requirements of each optimal control. In our setting, using 2 hidden layers, each with \( N_a + 2 \) nodes, were found during testing to be a NN structure that captured sufficient complexity for both benchmark outperformance problems, while ensuring that reliable and stable results were obtained on the numerical solutions as well as the ground truth solutions (see Appendix C).

C.4: Ground truth results

In order to show that the numerical solutions obtained as described in Section 4 can converge under suitable conditions to the closed-form solutions as described in Section 3, we immediately 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. \( \hat{\beta} \approx 1.0\% \) in (5.1)), 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. This latter requirement is readily achieved by using parametric models with parameters as in Table 5.3 to simulate paths of the underlying assets. Analytical investment strategies are calculated based on these parameters, while the numerical approach uses the sample paths based on these parameters as training data for the neural network.

The results, obtained using \( 10^6 \) Monte Carlo simulations, are compared in Table C.1 for investor portfolio \( P_0 \), benchmark BM0. Note that we assume contributions are zero, \( q = q(t_n) = 0 \) to avoid discrete approximation errors when comparing a continuous contribution rate to discrete contribution amounts made at rebalancing times.

Table C.1 shows that in this scenario, the numerical results using the data-driven NN approach as described in Section 4 indeed recover the analytical results obtained as per Section 3.
Table C.1: Ground truth comparison, investor portfolio $P_0$, benchmark $BM_0$, and data set $DS_0$ used for NN training data: $w_0 = 100$, $q = q(t_n) = 0$, $T = 1$ year. Since $BM_0$ results in an expected terminal wealth $K = 104.20$, a value of $\mathcal{E} = 105.25$ implies $\beta \approx 1.0\%$. 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.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Analytical solutions: $P_0$</th>
<th>Numerical solutions (using NN): $P_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$W^*_e(T)$</td>
<td>$W^*_e(T)/W(T)$</td>
</tr>
<tr>
<td>BM0</td>
<td>$W^*_e(T)$</td>
<td>$W^*_e(T)/W(T)$</td>
</tr>
<tr>
<td>Mean</td>
<td>104.2</td>
<td>105.3</td>
</tr>
<tr>
<td>CExp 5%</td>
<td>85.6</td>
<td>80.1</td>
</tr>
<tr>
<td>5th pctile</td>
<td>90.7</td>
<td>87.4</td>
</tr>
<tr>
<td>Median</td>
<td>104.1</td>
<td>105.6</td>
</tr>
<tr>
<td>95th pctile</td>
<td>117.9</td>
<td>121.9</td>
</tr>
</tbody>
</table>