

Intertemporal Optimal Portfolio Allocation under Regime Switching Using Artificial Neural Networks

Uri Carl¹, Yaacov Kopeliovich², Michael Pokojovy^{*3}, and Kevin Shea⁴

¹Blue Frontiers Partners, LLC

²University of Connecticut, Finance Department

³Old Dominion University, Department of Mathematics and Statistics

⁴Disciplined Alpha, LLC

Monday 24th February, 2025

Abstract

We consider the Merton’s optimal portfolio problem for a financial market responsive to macroeconomic regime changes, as characterized by the VIX, and lay out an allocation strategy that maximizes the expected isoelastic utility of the terminal wealth. Contrary to the traditional paradigm, we neither derive nor solve the Hamilton-Jacobi-Bellman partial differential equation typically used to compute the optimal asset weights. Instead, we employ an artificial neural network (ANN) to represent the latter as a feedback function and directly train the ANN model to solve the resulting stochastic maximization problem. Benchmarking our strategy against the usual regime-agnostic Merton’s portfolio, we show that a relatively simple ANN model adopted with one hidden layer is capable of producing superior expected utility results based on synthetic market returns data. Moreover, selecting a portfolio from twelve common assets plus cash subject to realistic diversification constraints while excluding borrowing/short selling, our back-testing study over the 18-year period spanning 2007-2024 provides evidence that incorporating regimes in portfolio allocation improves both the average wealth utility and average rate of return.

Keywords: Merton’s problem; portfolio optimization; market regimes; deep learning; artificial neural network; multi-layer perceptron.

1 Introduction

Markowitz (1952) revolutionized the financial industry by publishing his celebrated “*Allocation under Uncertainty*” article, in which he proposed a framework for optimal investing adjusted for risk. This paper caused a revolution in the financial industry by providing a framework for pricing models like CAPM. However, Markowitz’ framework was static. It was Merton (1969) who restated and solved this problem in a dynamic setting. He used stochastic control methods to produce a trading strategy that maximizes the expectation of certain utility functions of terminal wealth (known as bequest) and running consumption. The resulting partial differential equation (PDE), known as Hamilton-Jacobi-Bellman (HJB) PDE, was also solved by Markowitz (1952). This gave rise to his famous Merton’s ratio – a simple rule of investing that provides guidance about what percentage of one’s wealth should be invested in each respective asset. The importance of this rule was recently highlighted in The Economist. Buttonwood (Online Column) (2023).

Countless other papers were written since Merton’s seminal paper extending his methods to different stock dynamics and other asset classes (e.g., fixed income products). Merton also proposed a dynamic portfolio theory, the intertemporal CAPM. However, while Markowitz (1952) framework is widely adopted and used in the financial industry to design portfolio weights allocation (for example, the Black-Letterman method),

*Corresponding author. Email: mpokojovy@odu.edu

Merton’s framework has not achieved a comparable level of notoriety on the practitioners side. One reason is the presence of mathematical challenges once one implements Merton’s framework in practice. In many cases, one is forced to first use advanced stochastic calculus to derive the complicated HJB equation and then solve this fully nonlinear PDE with non-trivial boundary conditions. In addition to being computationally challenging, the complexity of this approach is prohibitively high making it inaccessible to most financial professionals.

The goal in this work is to argue that one can still make Merton’s framework feasible and accessible to practitioners by using recent advances in machine learning. Aside from extensive utilization of deep learning to model and forecast future market returns, broad swaths of financial machine learning literature on the Merton’s problem (Bachouch et al., 2022; Davey and Zheng, 2022; Gim and Park, 2021; Grohs et al., 2023) are predominantly concerned with how deep learning can facilitate efficient solution of the HJB equation as a means to derive the optimal portfolio. It is only very recently that some researchers (Kopeliovich and Pokojovy, 2024; Reppen and Soner, 2023; van Staden et al., 2023) opted to completely forego the complicated HJB PDE. Instead, they utilized artificial neural network (ANN) models to represent the feedback allocation strategy. In our case, treating the ANN as a black box, the latter is trained on vast amount of synthetic data resembling naturally occurring market returns to produce the optimal weights. Utilization of synthetic data has been a long-standing successful practice in quantitative finance followed by both academic researchers and investors (Zuckerman, 2019, p. 84). The approach adopted in our paper neither involves derivation nor solution of HJB PDEs or other abstract mathematical optimality conditions. In addition to being essentially market-agnostic, it is extremely user-friendly as no knowledge of stochastic analysis, calculus of variations or numerical PDEs is required.

To exemplify the power of our approach, we consider the Merton’s problem of wealth allocation under regime changes. The topic of regime changes has been extensively studied in econometric literature (Hamilton (1989) followed by over three decades of subsequent research) and successfully utilized in the financial industry. Yet, despite recognized importance for portfolio allocation (Ang and Timmermann, 2012), no concrete suggestions have been as to how to allocate weights under regimes changes. It also seems challenging to produce analytical solutions for such problems in practically relevant scenarios. In fact, the initial motivation for the present work was the question whether our ANN framework, initially developed in a stochastic volatility context (Kopeliovich and Pokojovy, 2024), is robust enough to handle the arguably more complex and less conventional problem of regime changes within the Merton’s framework. Having successfully applied our framework in the latter scenario, the main practical finding of our paper can be summarized in two words: “*Regimes count.*” Specifically, we show that the regime-specific ANN-based optimal portfolio we developed under the Merton’s framework clearly outperforms the usual regime-agnostic allocation, at least for the class of isoelastic utilities typically studied in the literature.

The structure of the paper as follows: In Section 2, we present a formal setup for the classical Merton’s problem and how semi-Markovian regime changes can be incorporated into this framework. In Section 3, we describe the dataset used in our study and protocol the calibration procedure. In Section 4, introduce a class of feedback controls implemented as artificial neural networks (ANN) and reformulate the dynamic asset allocation problem under regime switching subject to a set of portfolio constraints to the usual ANN training with stochastic gradient descent (SGD). All empirical results are presented in Section 5, where our regime-specific ANN-based optimal portfolio is compared to the classical Merton’s weights based on synthetic data and backtested on historical market returns. A summary of our findings and suggestions for future research are presented in the last Section 6.

2 Intertemporal Optimal Portfolio Allocation

The goal of this section is to explain the mathematical background of the optimal allocation problem we consider. The classical portfolio theory pioneered by Markowitz (1952) was a static problem. That is, no trading was allowed in his set-up. The crucial insight of Merton (1969) was to consider dynamical allocation strategy to maximize a given utility function specified by the user. This was a generalization of Markowitz (1952) that was essentially reduced to using quadratic utilities in the dynamic setup.

We formalize the setup of the classical Merton’s optimal portfolio problem is as follows: Given p risky assets that follow a multivariate geometric Brownian motion (GBM) process and risk-less asset (typically, a

short-term U.S. Treasury bond viewed as cash) with a constant return, the question is how one can develop a strategy of how much of one's current wealth be allocated to each of the $p + 1$ assets as to maximize a certain expected utility expression. The portfolio is subject to re-balancing in real time with the aim of optimizing the latter utility function, which may or may not include ongoing consumption and/or bequest at the end of a given time horizon, while accounting for the stochastic nature of future asset trajectories. In our study, we aim to maximize the expectation of the expected isoelastic utility of the terminal wealth at the end of the investment period T .

2.1 Classical Merton Problem

To formalize the allocation problem, let the vector $\mathbf{S}_t = (S_t^1, S_t^2, \dots, S_t^p)'$ denote the prices of the p stocks at time t , while R_t stand for the value of the cash account. The dynamics is then expressed as

$$d\mathbf{S}_t = \text{diag}(\mathbf{S}) (\boldsymbol{\mu} dt + \boldsymbol{\Sigma}^{1/2} d\mathbf{B}_t), \quad (1)$$

$$dR_t = rR_t dt \quad (2)$$

with a p -variate standard Brownian motion $(\mathbf{B}_t)_{t \geq 0}$, where $\boldsymbol{\mu} \in \mathbb{R}^p$ is the vector of expected returns and $\boldsymbol{\Sigma} \in \mathbb{R}^{p \times p}$ is a symmetric, positive definite covariance matrix between the p -variate innovations.

With π_t^k denoting the proportion of wealth invested in the k -th risky asset, $\pi_t^0 = 1 - \sum_{k=1}^p \pi_t^k$ is the the proportion of wealth invested in the riskless asset, which furnishes the usual budget constraint

$$\pi_t^0 + \mathbf{1}'_p \boldsymbol{\pi}_t = 1 \text{ for } t \geq 0 \quad (3)$$

with $\mathbf{1}_p = (1, 1, \dots, 1)' \in \mathbb{R}^p$ and $\boldsymbol{\pi}_t = (\pi_t^1, \dots, \pi_t^p)'$. This leads to the usual wealth dynamics

$$dW_t = W_t \left(\boldsymbol{\pi}'_t (\boldsymbol{\mu} - r\mathbf{1}_p) dt + \boldsymbol{\pi}'_t \boldsymbol{\Sigma}^{1/2} d\mathbf{B}_t \right). \quad (4)$$

The stochastic differential equation (SDE) solely depends on asset returns so that the price dynamics in Equations (1)–(2) becomes redundant.

Let $\mathcal{K} \subset \mathbb{R}^p$ be a non-empty closed set describing constraints imposed on stock weights at any time t . Allowing for unlimited short selling and/or cash borrowing, $\mathcal{K} = \mathbb{R}^p$ (Merton, 1969). If both short selling and borrowing are to be excluded, we get

$$\mathcal{K} = \{ \boldsymbol{\pi} \in [0, 1]^p \mid \mathbf{1}'_p \boldsymbol{\pi} \leq 1 \}.$$

Later in this paper we will additionally adopt a more sophisticated set of assumptions referred to as diversification constraints.

Consider the iso-elastic utility function

$$U(w) = \begin{cases} \frac{w^{1-\eta}-1}{1-\eta}, & \eta \geq 0, \eta \neq 1 \\ \log(w), & \eta = 1 \end{cases} \quad \text{for } w > 0. \quad (5)$$

For an admissible allocation $\boldsymbol{\pi} \in \mathcal{A}_{0,w}$ (cf. Equation (42) in Appendix B), let $W_s^{\boldsymbol{\pi};0,w}$ denote the associated strong solution to Equation (4). Consider the terminal expected utility functional

$$J(U(W_T^{\boldsymbol{\pi};0,w})) = \mathbb{E}_{\mathcal{F}_{0,T}} [U(W_T^{\boldsymbol{\pi};0,w})]. \quad (6)$$

With this notation, the optimal portfolio problem reads as

$$J(U(W_T^{\boldsymbol{\pi};0,w})) \rightarrow \max_{\boldsymbol{\pi} \in \mathcal{A}_{0,w}}, \quad \boldsymbol{\pi}^* = \arg \max_{\boldsymbol{\pi} \in \mathcal{A}_{0,w}} J(W_T^{\boldsymbol{\pi};0,w}), \quad W^* = (W_s^{\boldsymbol{\pi}^*;0,w})_{0 \leq s \leq T}. \quad (7)$$

The stochastic control problem in Equation (7) is typically solved by employing the Bellman's dynamic programming principle. To this end, one can define the value function

$$V(t, w) := \sup_{\boldsymbol{\pi} \in \mathcal{A}_{t,w}} \mathbb{E}_{\mathcal{F}_{t,T}} [U(W_T^{\boldsymbol{\pi};t,w})]. \quad (8)$$

Under suitable regularity assumptions, the well-known Hamilton-Jacobi-Bellman (HJB) can be derived

$$\partial_t V(t, w) + \mathcal{H}(t, w, \nabla V, \nabla^2 V) = 0 \quad \text{for } (t, w) \in (0, T) \times (0, \infty), \quad (9)$$

$$V(T, w) = U(w) \quad \text{for } (t, w) \in \partial_P((0, T) \times (0, \infty)) \quad (10)$$

with the usual elliptic HJB operator

$$\mathcal{H}(t, w, \nabla V, \nabla^2 V) = \sup_{\pi \in \mathcal{K}} H(\pi, t, w, \nabla V, \nabla^2 V) \quad (11)$$

(for a suitably defined function $H(\cdot)$) and the parabolic boundary operator ∂_P . See Fleming and Soner (2006). Using the usual ansatz for the value function

$$V(t, w) = C(t)(w^{1-\eta} - 1) \text{ for some unknown } C = C(t),$$

the optimal weight π^* can be shown to be constant and solve the following quadratic programming problem

$$-\frac{\eta}{2} \pi' \Sigma \pi + \pi' (\mu - r \mathbf{1}_p) \rightarrow \max_{\pi \in \mathcal{K}}. \quad (12)$$

In the unconstrained case, i.e., $\mathcal{K} = \mathbb{R}^p$, the Fermat's optimality conditions associated with solving the algebraic optimization problem in Equation (11) furnish the usual time- and wealth-independent Merton portfolio weights

$$\pi^* = \frac{1}{\eta} \Sigma^{-1} (\mu - r \mathbf{1}_p). \quad (13)$$

Plugging the vector of optimal weights back into HJB equations (9)–(10), the latter PDE can be explicitly solved. Further, using the usual Verification Theorem, the unique solution can be shown to be the value function associated with the optimal control problem in Equation (7).

While the Bellman's dynamic programming principle can theoretically be applied to any Markovian market and wealth dynamics (and, in turn, any dynamics can be rendered Markovian on an extended phase space), theoretical derivation of the HJB are not only tedious and laborious, but may lead to untractable fully nonlinear PDEs or other exotic operator equations, where the regularity issues can further be exacerbated by non-smoothness of the HJB operator. Should one attempt to solve the HJB numerically to circumvent these difficulties, computational issues known as “curse of dimensionality” typically arise.

2.2 Merton Problem with Regime Switching

In this work, we consider a different dynamics for the asset prices S_t and the cash account R_t . Both processes are still assumed pathwise continuous, the drift and volatility in Equation (1) as the short rate in Equation (2) are now subject to a switching dynamics regulated by some independent semi-Markov process $(\alpha_t)_{t \geq 0}$.

Later in Section 4, we will use an ANN model to explore whether asset allocation can be improved if we assume two natural regimes associated with the process $(\alpha_t)_{t \geq 0}$ – referred to as risk-on and risk-off in this paper – in which the economy is doing well and there is an optimism in the market or, on the contrary, the economy is experiencing a downturn and, as such, the market feels pessimistic. Thus, we assume the market dynamics is a “mixture” of two processes with the following properties, for which the notion of portfolio optimality is formalized below:

1. The market is always in one of the two regimes.
2. Even though we do not know when the switching from one to another regime (and vice versa) may occur, we know the infinitesimal probabilities of this switch.
3. We know exactly or can statistically infer which regime we are currently in.

For $\mathcal{S} := \{1, 2, \dots, m\}$ equipped with the discrete metric, let α_t be an \mathcal{S} -valued stochastic process representing one of the m -regimes the market can be in at time t . In this paper, α_t is modeled as a continuous-time Markov chain with a given rate matrix $\mathbf{Q} = (q_{ab})_{a,b=1}^m \in \mathbb{R}^{m \times m}$ satisfying the usual assumption

$$q_{ab} \geq 0 \text{ for } a \neq b \quad \text{and} \quad q_{aa} = - \sum_{b \neq a} q_{ab}. \quad (14)$$

Infinitesimally,

$$\mathbb{P}(\alpha_{t+h} = b \mid \alpha_t = a) = \mathbf{1}_{\{a=b\}} + q_{ab}h + o(h) \text{ as } h \rightarrow 0 \text{ uniformly in } t \geq 0. \quad (15)$$

In accordance to Kolmogorov forward equations, the matrix exponential $\exp(h\mathbf{Q})$ specifies the (approximate) transition probability matrix similar to that associated with a usual discrete-time Markov chain with the time step h .

The process $(\alpha_t)_{t \geq 0}$ possesses a càdlàg realization with a pathwise locally finite number of jumps. Let $(\tau_n)_{n \geq 0}$ with $0 < \tau_1 < \dots < \tau_n < \dots$ denote the sequence of (random) jump times so that $(H_n)_{n \geq 1}$ with

$$H_1 := \tau_1, \quad H_n := \tau_n - \tau_{n-1} \quad \text{for } n \geq 2 \quad (16)$$

corresponds to the usual holding time sequence.

Let $\boldsymbol{\mu}(a) \in \mathbb{R}^{p \times p}$, $\boldsymbol{\Sigma}(a) \in \mathbb{R}^{p \times p}$ and $r(a) > 0$ be the market drift vector, positive definite squared volatility matrix and the riskless rate in the a -th regime ($a \in \mathcal{S}$). Assuming \mathbf{S}_t is a pathwise continuous process that follows an geometric Brownian motion on each of the intervals (τ_{n-1}, τ_n) , the overall market dynamics follows the SDE

$$d\mathbf{S}_t = \text{diag}(\mathbf{S}_t) \left(\boldsymbol{\mu}(\alpha_{t-}) dt + \boldsymbol{\Sigma}^{1/2}(\alpha_{t-}) d\mathbf{B}_t \right), \quad (17)$$

$$dR_t = r(\alpha_{t-}) R_t dt. \quad (18)$$

Similar to Equation (4), this furnishes the wealth dynamics

$$dW_t = W_t \left(\boldsymbol{\pi}'_{t-} (\boldsymbol{\mu}(\alpha_{t-}) - r(\alpha_{t-}) \mathbf{1}_p) dt + \boldsymbol{\pi}'_{t-} \boldsymbol{\Sigma}^{1/2}(\alpha_{t-}) d\mathbf{B}_t \right). \quad (19)$$

For any regular control process $(\boldsymbol{\pi}_t)_{t \geq 0}$, Equation (19) admits a suitably defined strong solution (cf. Appendix B). Let $(W_s^{\boldsymbol{\pi}; t, w, a})_{t \leq s \leq T}$ denote the unique solution to Equation (19) subject to the initial condition

$$W_t = w. \quad (20)$$

As before, consider the terminal expected utility functional

$$J(U(W_T^{\boldsymbol{\pi}; 0, w, a})) = \mathbb{E}_{\mathcal{F}_0, T, a} [U(W_T^{\boldsymbol{\pi}; 0, w, a})]. \quad (21)$$

The Merton's optimal portfolio problem reads then as

$$J(U(W_T^{\boldsymbol{\pi}; 0, w, a})) \rightarrow \max_{\boldsymbol{\pi} \in \mathcal{A}_{0; w, a}} , \quad \boldsymbol{\pi}^* = \arg \max_{\boldsymbol{\pi} \in \mathcal{A}_{0; w, a}} J(W^{\boldsymbol{\pi}; 0, w, a}), \quad W^* = (W_s^{\boldsymbol{\pi}^*; 0, w, a})_{0 \leq s \leq T}. \quad (22)$$

with the set of admissible controls $\mathcal{A}_{t, w, a}$ defined in Equation 45 in Appendix B.

Introduce the value function

$$V(t, w, a) := \sup_{\boldsymbol{\pi} \in \mathcal{A}_{t, w, a}} \mathbb{E}_{\mathcal{F}_t, T, a} [U(W_T^{\boldsymbol{\pi}; t, w, a})], \quad (23)$$

where one can argue similarly to Zariphopoulou (2001) to conclude that the former function does not depend on initial stock prices. The value function is expected to satisfy an HJB equation

$$\partial_t V(t, w, a) + \mathcal{H}(t, w, a, V, \nabla V, \nabla^2 V) = 0 \quad \text{for } (t, w, a) \in (0, T) \times (0, \infty) \times \mathcal{S}, \quad (24)$$

$$V(T, w, a) = U(w) \quad \text{for } (t, w, a) \in \partial_P((0, T) \times (0, \infty)) \times \mathcal{S} \quad (25)$$

with a properly defined HJB operator

$$\mathcal{H}(t, w, a, V, \nabla V, \nabla^2 V) = \sup_{\boldsymbol{\pi} \in \mathcal{K}} H(\boldsymbol{\pi}, t, w, a, V, \nabla V, \nabla^2 V) \quad (26)$$

for a suitably defined algebraic function $H(\cdot)$. (Note that, unlike Equation (11), the HJB operator not only depends on the regime a , but the value function V itself.)

Asset	Description
Cash	Barclays US Treasury (1-3 months) Index
US Government bonds	Barclays US Treasury Bellwethers (10Y) Index
US Fixed Income	Fixed Income Barclays Aggregate
Municipal Bonds	Barclays Municipal Bond Blend (1-17 years) Index
High Yield	Barclays High Yield Index
HFRI Market Defensive	HFRI Fund of Funds Market Defensive Index
HFRI Diversified	HFRI Fund of Funds Market Diversified Index
HFRI Strategic	HFRI Fund of Funds Market Strategic Index
US Core	Russell 1000 Core
US Growth	Russell 1000 Growth
US Value	Russell 1000 Value
Commodities	Dow Jones UBS Commodity Index
Gold	Gold

Table 1: Cash plus twelve risky assets grouped by classes.

3 Dataset Description and Model Calibration

The dataset analyzed in this paper contains monthly returns (on the last trading day of each month) of a cash-like asset (Barclays US Treasury (1-3 months) Index) and twelve risky assets listed in Table 1 for the 35-year period from 1/1/1990 to 12/31/2024. Additionally, the monthly close value of the S&P 500 volatility index, VIX, was used as a risk-regime barometer. All data were sourced from FactSet.

3.1 Regime Segmentation

We chose to use VIX, the CBOE Volatility Index measuring the stock market’s expectation of volatility reflected by S&P 500 index options, as an indicator of the prevailing market regime. It is well known among market participants that when the VIX is alleviated, the greed is prevalent and the market is likely to be in the risk-off regime, while when VIX is low, the fear is dominant and the risk-on regime is typically assumed. Below, we use monthly close VIX values to perform regime segmentation.

To detect the regimes, the 12-month VIX geometric averages are computed according to the equation

$$MA_{t_k} = \left(\prod_{i=1}^{12} VIX_{t_k-i} \right)^{1/12} = \exp \left(\frac{1}{12} \sum_{i=1}^{12} \log(VIX_{t_k-i}) \right), \quad (27)$$

where the first twelve MA values obtained via the usual “padding” at the left of the time interval (cf. Figure 1). Applying the logarithmic transform to the geometric moving averages in line with curved Riemannian nature of volatilities (Pourahmadi, 2013, Chapter 3), the historical $\log(MA_{t_k})$ ’s values from Equation (27) are clustered with the well-known K -means algorithm (MacKay, 2003, Chapter 20) for $K = 2$ and the usual maximum-margin separation threshold $\log(\vartheta_{VIX})$ (Vapnik, 1997) is computed (on the log-scale). For example, using the entire 35 years results a cutoff value of $\vartheta_{VIX} = 18.71$ for MA_{t_k} (on original the pre-log-scale). This corresponds to about 52% and 48% of time spent in the risk-on and the risk-off regime, respectively. However, note that we adopted a rolling 17-year time window in the backtesting study reported in Section 5 below. This means that both the cutoff values and average holding times may and do vary throughout the backtesting process.

The time grid is then subdivided into risk-on and risk-off segments consisting of contiguous “intervals” comprised of time periods t_k with $MA_{t_k} < \vartheta_{VIX}$ and $MA_{t_k} > \vartheta_{VIX}$, respectively. Next, the average lengths of risk-on and risk-off time segments are computed producing estimates of average holding times in respective regimes. Over the 35-year time horizon, this translates to 4.56 and 4.19 years on average in risk-on and risk-off regimes, respectively. Therefore, with $q_{11} = -4.56$ and $q_{22} = -4.19$, we arrive at an estimate

$$\hat{Q} = \begin{pmatrix} -4.56 & 4.56 \\ 4.19 & -4.19 \end{pmatrix} \quad (28)$$

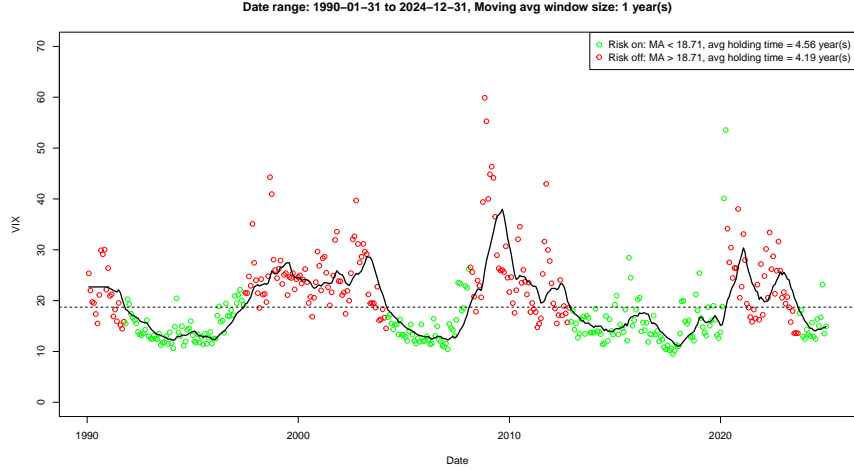


Figure 1: VIX, VIX 12-month geometric moving averages and maximum-margin separation threshold in risk-on (green) and risk-off (red) regimes over the period from 1/1/1990 to 12/31/2024 along with the average holding times in each regime (legend).

of the transition rate matrix \mathbf{Q} in Equations (14)–(15). Taking the matrix exponential $\exp((\Delta t)\mathbf{Q})$ gives the (approximate) transition probability matrix for switching between the risk-on and risk-off regime over an infinitesimally short time period Δt . For example, letting $\Delta t = \frac{1}{12}$, the approximate transition probability matrix reads as

$$\exp\left(\frac{1}{12}\mathbf{Q}\right) = \begin{pmatrix} 0.7302 & 0.2698 \\ 0.2479 & 0.7521 \end{pmatrix}.$$

3.2 Market Calibration

The developments in Section 3.1 enable us to segment the calibration time horizon into risk-on and risk-off months

$$\mathcal{T}(1) = \{t_k \mid \text{MA}_{t_k} < \vartheta_{\text{VIX}}\} \text{ and } \mathcal{T}(2) = \{t_k \mid \text{MA}_{t_k} > \vartheta_{\text{VIX}}\}, \quad (29)$$

respectively. See the time series plots in Figure 2 for the returns from cash (first panel) as well risky assets (subsequent twelve panels).

Next, we compute the usual unbiased estimates of the risk-free rate, drift vector and squared volatility matrix

$$\hat{r}(a) = \frac{1}{|\mathcal{T}(a)|} \sum_{t_k \in \mathcal{T}(a)} \text{Riskless Return}_{t_k}, \quad (30)$$

$$\hat{\boldsymbol{\mu}}(a) = \frac{1}{|\mathcal{T}(a)|} \sum_{t_k \in \mathcal{T}(a)} \text{Risky Return Vector}_{t_k}, \quad (31)$$

$$\hat{\boldsymbol{\Sigma}}(a) = \frac{1}{|\mathcal{T}(a)| - 1} \sum_{t_k \in \mathcal{T}(a)} (\text{Risky Return Vector}_{t_k} - \hat{\boldsymbol{\mu}}(a)) (\text{Risky Return Vector}_{t_k} - \hat{\boldsymbol{\mu}}(a))' \quad (32)$$

with $a = 1$ and $a = 2$ corresponding to risk-on and risk-off regimes, respectively, where

$$\begin{aligned} \text{Riskless Return}_{t_k} &\approx \frac{1}{\Delta t} (\log(R_{t_k}) - \log(R_{t_{k-1}})), \\ \text{Risky Return Vector}_{t_k} &\approx \frac{1}{\Delta t} (\log(\mathbf{S}_{t_k}) - \log(\mathbf{S}_{t_{k-1}})) \end{aligned}$$

denote the annualized historical monthly returns and $\Delta = \frac{1}{12}$ years.

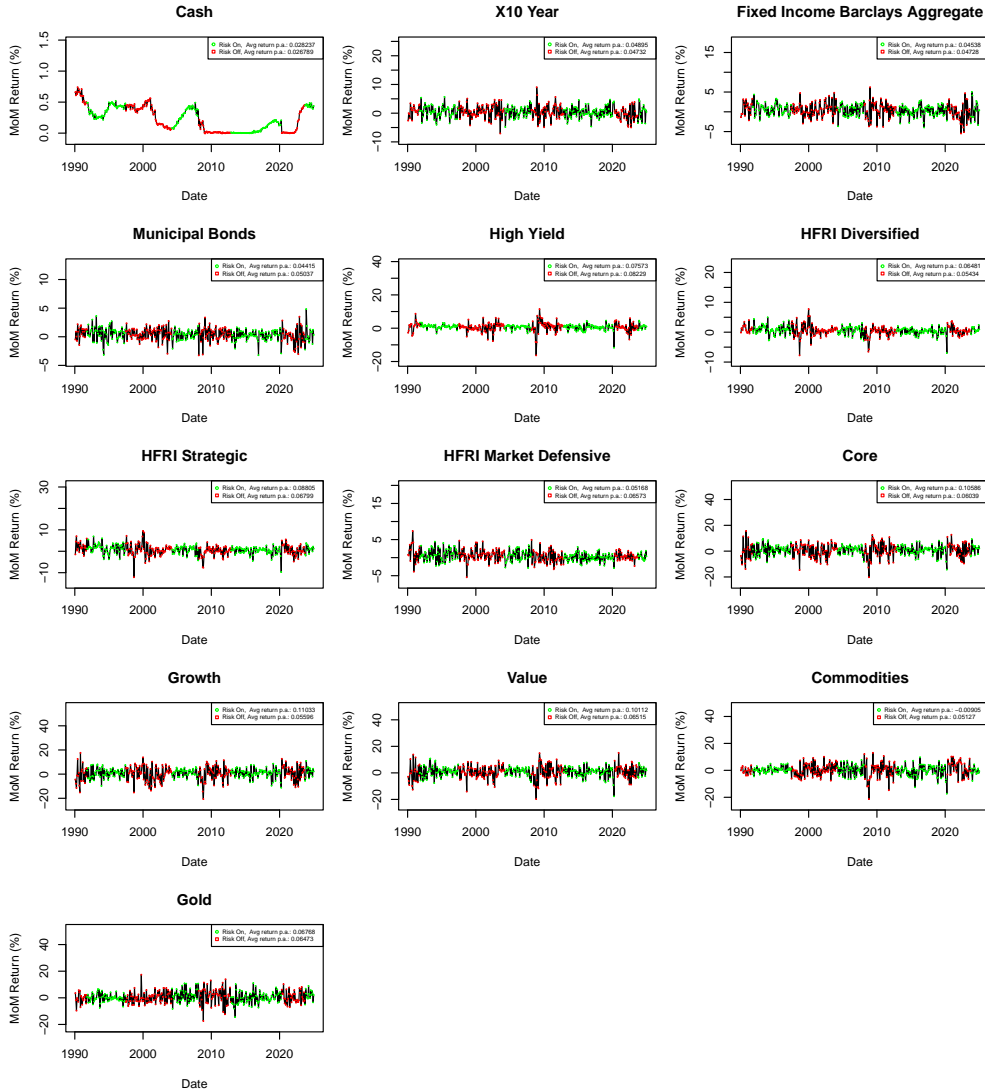


Figure 2: Historical market returns between 1/1/1990 and 12/31/2024 in risk-on (green) and risk-off (red) regimes.

4 Optimal Portfolio Allocation via Deep Learning

Instead of solving the HJB equation to compute the optimal weights for the portfolio allocation, we propose a framework where the optimal portfolio value and weights are regular functions, which can be modeled as an ANN. That is, to solve the weights optimization problem, we represent it as a neural network with activation functions. This ANN is trained over mini-batches of data, and a standard data split of train, test, and validation is used to protect against over-fitting. It is important to note that the value of the portfolio does not to be computed in this model but rather the weights through the optimization routine of the ANN. Instead of solving the HJB equation to compute the optimal weight, a simple yet rigorous alternative is to impose appropriate regularity assumptions. These assumptions ensure that V and, therefore, π^* are sufficiently regular functions of their arguments. In such a case, the usual ANN approximation results hold (Hornik, 1991), implying that the unknown feedback control π^* can be learned with DL over any compact set in (t, y) without having to compute the value function $V(\cdot)$.

4.1 Portfolio Constraints

In our study, both short selling and borrowing are excluded. Additionally, realistic diversification constraints are imposed as detailed in Table 2. This produces the constraint set \mathcal{K} and the associated orthogonal projector $\Pi_{\mathcal{K}}: \mathbb{R}^p \rightarrow \mathcal{K}$ defined in Appendix C and used in the ANN construction in Section 4.2 below.

Class	Asset	Constraints
\mathcal{I}_0	Cash	$l_0 = 0.00 \leq \sum_{i \in \mathcal{I}_0} \pi^i \leq 0.20 = u_0$
\mathcal{I}_1	US Government bonds US Fixed Income Municipal Bonds	$l_1 = 0.20 \leq \sum_{i \in \mathcal{I}_1} \pi^i \leq 0.70 = u_1$
\mathcal{I}_2	High Yield	$l_2 = 0.00 \leq \sum_{i \in \mathcal{I}_2} \pi^i \leq 0.30 = u_2$
\mathcal{I}_3	HFRI Market Defensive HFRI Diversified HFRI Strategic	$l_3 = 0.04 \leq \sum_{i \in \mathcal{I}_3} \pi^i \leq 0.30 = u_3$
\mathcal{I}_4	US Core US Growth US Value	$l_4 = 0.20 \leq \sum_{i \in \mathcal{I}_4} \pi^i \leq 0.70 = u_4$
\mathcal{I}_5	Commodities Gold	$l_5 = 0.04 \leq \sum_{i \in \mathcal{I}_5} \pi^i \leq 0.30 = u_5$

Table 2: Portfolio constraints involving cash plus twelve risky assets grouped by classes. Note that the cash weight $\pi^0 = 1 - \mathbf{1}'_p \mathbf{w}$ is uniquely determined through the remaining twelve weights.

4.2 Feedback Control via Artificial Neural Network

To develop a feedback control, we observe that the relevant variables associated with the dynamics in Equation (19) are the time $t \in [0, T]$, the state $a \in \mathcal{S}$ and the wealth $w > 0$. Since the latter is irrelevant under the isoelastic utility (Zariphopoulou, 2001), only the former two remain. Since $a \in \mathcal{S}$ is a categorical variable with m levels, we can use the usual encoding

$$\mathbf{D}(a) \equiv (D_1(a), D_2(a), \dots, D_{m-1}(a))$$

with $m - 1$ dummy variables

$$D_1(a) = \mathbb{1}_{\{a_1\}}(a), \quad D_2(a) = \mathbb{1}_{\{a_2\}}(a), \quad \dots, \quad D_{m-1}(a) = \mathbb{1}_{\{a_{m-1}\}}(a),$$

where the m -th state a_m selected as the baseline (with all dummy variables being 0). In the two-regime (risk-on and risk-off) scenario of Section 3, a single dummy variable

$$\mathcal{D}(a) \equiv \mathbb{1}_{\{1\}}(a)$$

is sufficient, where $a = 1$ and $a = 2$ correspond to the risk-on and risk-off regimes, respectively.

Choosing a feedback function

$$\boldsymbol{\pi}: \mathbb{R}^m \rightarrow \mathcal{K} \subset \mathbb{R}^p, \quad (t, \mathbf{D}) \mapsto \boldsymbol{\pi}(t, \mathbf{D}) \quad (33)$$

with m inputs (1 for time t plus $(m - 1)$ for regime dummy variables) and p outputs (weights for each risky asset), the feedback control is given via

$$\boldsymbol{\pi} \equiv \boldsymbol{\pi}(t, \mathbf{D}(\alpha_t)). \quad (34)$$

Let $(W_t^{\boldsymbol{\pi}; 0, w^0, a^0})_{t \in [0, T]}$ denote the wealth process given as the unique solution to Equation (19) driven by the feedback control in Equation (34) and subject to the initial condition

$$W_0 = w^0, \quad \alpha_0 = a^0$$

for given initial wealth $w^0 > 0$ and initial regime $a^0 \in \mathcal{S}$. The control problem in Equation (22) reduces then to maximizing the functional

$$J(\boldsymbol{\pi}) = \mathbb{E}[U(W_T^{\boldsymbol{\pi}=\boldsymbol{\pi}(t, \mathbf{D}(\alpha_t)); 0, w^0, a^0})] \quad (35)$$

over all sufficiently regular feedback controls $\boldsymbol{\pi}: \mathbb{R}^m \rightarrow \mathcal{K}$.

Following Kopeliovich and Pokojovy (2024), instead of invoking the dynamic programming principle to compute $\boldsymbol{\pi}$ via solving the HJB for the value function, we represent the unknown feedback function as a feedforward artificial neural network (ANN), always known as a multi-layer perceptron (Goodfellow et al., 2016, Chapter 6). This type of deep learning models are expressed as a function

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\theta}) = \left(\bigcirc_{i=1}^d a_i(\mathbf{W}_i \cdot + \mathbf{b}_i) \right)(\mathbf{x}) \quad (36)$$

consisting of $d \geq 2$ layers, namely, one input, one output and $(d - 2)$ hidden layers, and associated bias vectors $\mathbf{b}_i \in \mathbb{R}^{n_i}$ and weight matrices $\mathbf{W}_i \in \mathbb{R}^{n_i \times n_{i-1}}$ that can be flattened to a parameter vector

$$\boldsymbol{\theta} = ((\mathbf{b}_i, \mathbf{W}_i))_{i=1, \dots, d}$$

as well as activation functions $a_i: \mathbb{R} \rightarrow \mathbb{R}$, $i = 1, \dots, d$, where “ \bigcirc ” is the usual composition operator. Specifically, we choose $n_0 = m$ (m inputs) and $n_d = p$ (p output). As for activation functions, to facilitate universal approximation and mitigate the vanishing gradients problem, we choose linear sigmoid units (SiLU) $a_i(x) = x/(1+e^{-x})$ ($i = 1, \dots, d-1$) for all layers but the output layer, where the sigmoid activation function $a_d(x) = 1/(1+e^{-x})$ to force the weights into $(0, 1)$ interval, thus, excluding borrowing or short-selling. Figure 3 illustrates the shallow ANN architecture with only $d = 3$ layers (an input layer with $m = 2$ neurons, one hidden layer with 3 neurons and an output layer with $p = 12$ neurons) used in the present paper. The weights displayed in the figure are obtained from a training procedure described in subsequent sections.

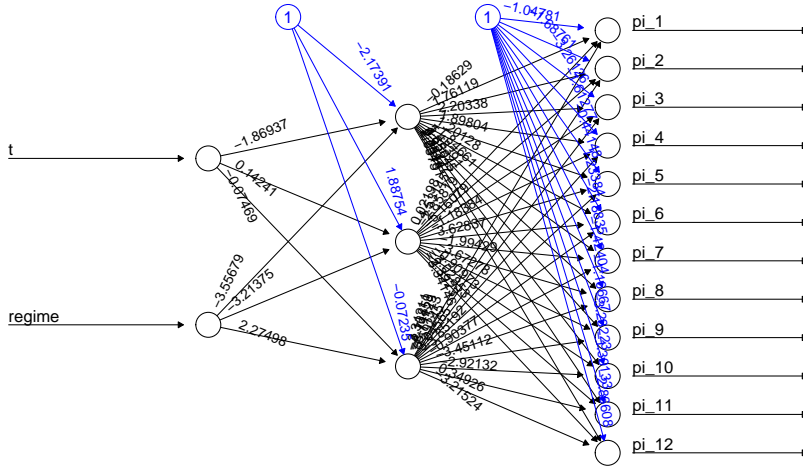


Figure 3: An illustration of a feedforward ANN.

With this notation, the feedback control in Equation (34) can be expressed as

$$\boldsymbol{\pi}_t(\mathbf{D}(a)) := \mathcal{F}(\mathbf{x}|\boldsymbol{\theta}) \text{ with } \mathcal{F}(\mathbf{x}|\boldsymbol{\theta}) = \Pi_{\mathcal{K}}(\mathbf{S}(\mathcal{N}(\mathbf{x}|\boldsymbol{\theta}))) \text{ and } \mathbf{x} = (t, \mathbf{D}(a)) \quad (37)$$

with m inputs (1 for time t plus $(m - 1)$ for regime dummy variables) and p outputs (weights for each risky asset). Here, \mathbf{S} is the normalization mapping

$$\mathbf{S}: ([0, 1]^p \setminus \{\mathbf{0}\}) \rightarrow \mathbb{R}^p, \quad \boldsymbol{\pi} \mapsto \boldsymbol{\pi} / \min\{\mathbf{1}'_p \boldsymbol{\pi}, 1\}$$

and $\Pi_{\mathcal{K}}(\cdot)$ is the projection mapping defined in Appendix C. Note that no division by zero can occur since $\mathcal{N}(\cdot|\boldsymbol{\theta})$ only outputs positive vectors.

Using this ansatz, the expected utility reduces to a function

$$J(\boldsymbol{\theta}) = \mathbb{E}[U(W_T^{\boldsymbol{\pi}=\mathcal{F}(t, \mathbf{D}(\alpha_t)|\boldsymbol{\theta});0,w^0,a^0})] \quad (38)$$

that solely depends on the ANN weights $\boldsymbol{\theta} \in \Theta$ in the “flattened” parameter space

$$\Theta := \prod_{i=1}^d \mathbb{R}^{n_i \times n_{i-1}} \times \mathbb{R}^{n_i}.$$

For the computational treatment of the problem, we adopt the Euler-Maruyama discretization (see Appendix A) to replace the continuous wealth process W_t with its discrete counterpart $W_t^{\Delta t}$. Replacing the expectation operator in Equation (38) with the sample average, we obtain the empirical form of the objective functional

$$\hat{J}(\boldsymbol{\theta}) = \frac{1}{B} \sum_{b=1}^B U(W_T^{\Delta t, b, \boldsymbol{\pi}=\mathcal{F}(\cdot|\boldsymbol{\theta});0,w^0,a^0}), \quad (39)$$

where the average is taken over a “minibatch” of (independent) B solutions $(W_{t_k}^{\Delta t, b, \boldsymbol{\pi}=\mathcal{F}(\cdot|\boldsymbol{\theta});0,w^0,a^0})_{t_k}$ to Equations (40)–(41) driven by B independent Wiener increments $(\Delta \mathbf{B}_{t_k}^b)_{t_k}$ for $b = 1, \dots, B$. As proposed by Kopeliovich and Pokojovy (2024), partial derivatives

$$\partial_{w_i} \hat{J}(\boldsymbol{\theta}) \approx (1/\epsilon)(\hat{J}(\boldsymbol{\theta} + \epsilon \mathbf{e}_i) - \hat{J}(\boldsymbol{\theta}))$$

are approximated using numerical differentiation, where \mathbf{e}_i is the i -th $\dim(\Theta)$ -dimensional unit basis vector and $\epsilon = 10^{-6}$, furnishing the approximate gradient

$$\nabla \hat{J}(\boldsymbol{\theta}) \approx \left((1/\epsilon)(\hat{J}(\boldsymbol{\theta} + \epsilon \mathbf{e}_i) - \hat{J}(\boldsymbol{\theta})) \right)_{i=1, \dots, \dim(\Theta)}.$$

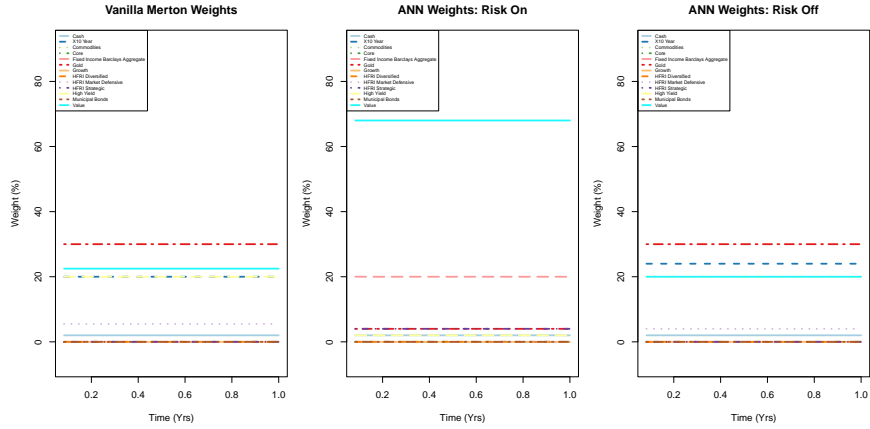
The maximization problem for $\hat{J}(\boldsymbol{\theta})$ is then reduced to minimizing the “loss function” $-\hat{J}(\boldsymbol{\theta})$ using one of the conventional (batch) stochastic gradient descent techniques (SGD). In this paper, we chose the popular Adam (Adaptive Moment Estimation) algorithm (Kingma and Ba, 2014) to train the neural network. Adam SGD is one of the prominent optimization techniques preferred by ML researchers and practitioners for its ability to adopt adaptive learning rates in individual weights in the ANN based on exponentially weighted moving averages of previous gradients and their squares. Since unlimited supply of synthetic data is available through the calibrated discrete dynamics described in Appendix A, no overtraining can occur so that the usual training, validation and test split is not required unless the user opts for using pre-computed market dynamics (viz. Equations (17)–(18)) for compute time reasons.

5 Results

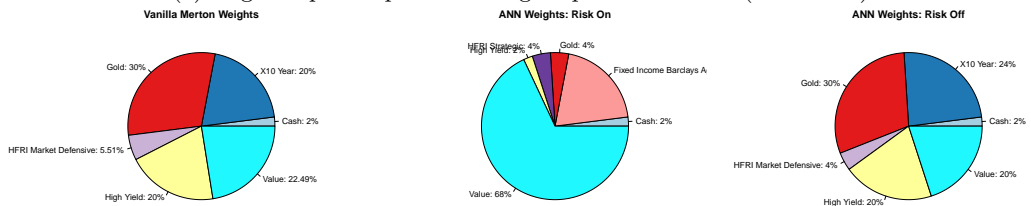
This section details the results of a backtesting study that documents the performance of our methodology and benchmarks it against the “vanilla” single-regime geometric Brownian motion used in the conventional Merton (1969) approach. We specifically focus on the isoelastic utility (viz. Equation (5) with power $\eta = 2$, which is within the typical range adopted by most investors. The remaining choices ($\eta = 1, 3, 4$) are given in the Supplement.

With 35 years’ worth of data from 1/1/1990 to 12/31/2024 described in Section 3, the backtesting was performed on the latter 18 years between 1/1/2007 and 12/31/2024. Starting in 2007 as the first backtesting year, the 17 years preceding 1/1/2007 (i.e., 1/1/1990 through 12/31/2006) were used to detect the regimes, estimate the semi-Markov switching process (viz. Section 3.1) and estimate the geometric Brownian motion model in the risk-on and risk-off regimes (viz. Section 3.2).

The calibrated discrete dynamics in Equations (40)–(40) in Appendix A with the ANN feedback control from Equation (37) was used to simulate the monthly wealth dynamics over a one-year investment horizon ($T = 1$ year with $\Delta t = \frac{1}{12}$ years) subject to portfolio rebalancing at the beginning of each month. Choosing the ANN architecture displayed in Figure 3 (with 47 trainable parameters) and adopting the portfolio

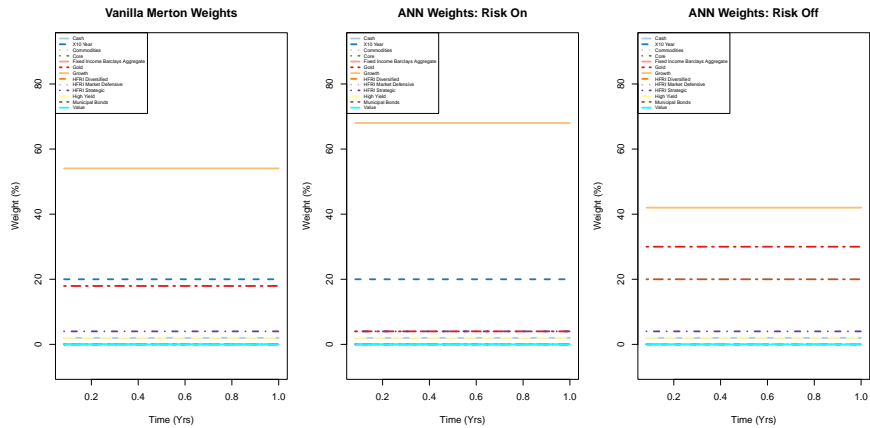


(a) Regime-specific portfolio weights plotted vs time (in months).

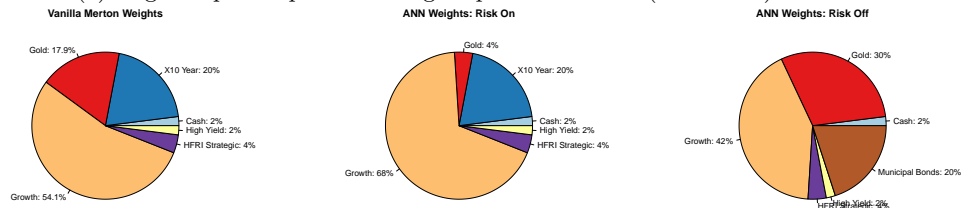


(b) Vanilla vs ANN-based regime-specific portfolios displayed as pie-charts.

Figure 4: Vanilla vs ANN-based regime-specific portfolios based on 17 years preceding 2013.



(a) Regime-specific portfolio weights plotted vs time (in months).



(b) Vanilla vs ANN-based regime-specific portfolios displayed as pie-charts.

Figure 5: Vanilla vs ANN-based regime-specific portfolios based on 17 years preceding 2022.

constraints described in Section 4.1, the ANN weights were trained to minimize the loss function $-\hat{J}(\theta)$ (viz. Equation (38)) using the Adam batch SGD method (Kingma and Ba, 2014) with default tuning parameter settings ($\rho_1 = 0.9$, $\rho_2 = 0.999$ and $\delta = 10^{-8}$). The weights were initialized as i.i.d. $\mathcal{N}(0, 0.1^2)$ random variate (meaning that nearly all wealth is allocated into cash). The following adaptive training schedule was adopted to speed-up the convergence: 2000 steps with minibatch size $B = 10$ and $\varepsilon = 0.1$, 500 steps with $B = 20$, $\varepsilon = 0.01$, 500 steps with $B = 20$, $\varepsilon = 0.005$, and 100 steps with $B = 50$, $\varepsilon = 0.005$. Once the ANN has been trained, the resulting optimal portfolio was evaluated both on synthetic data (see Section 5.1) and real historical data (see Section 5.2) for the respective year. The procedure was repeated for each of the 18 backtesting years (2008 through 2024), while recalibrating the regime and market models and retraining the ANN portfolio.

All computations were implemented in plain R code and run in RStudio under 64-bit Windows 11 Pro OS on a Dell Precision 3581 laptop with 13th Gen Intel(R) Core(TM) i7-13800H 2.50 GHz CPU and 32 GB RAM. No third-party DL libraries or GPU acceleration were employed. The run time (including training and performance evaluation) for each of the 18 years over the backtesting period 2007-2024 was about 5 hrs, i.e., less than 20 mins per each year of backtesting.

Figure 4 and 5 display Merton’s vanilla vs our ANN-based regime-specific portfolios for the years 2013 and 2022, each calibrated on 17 years preceding respective backtesting year. The vanilla portfolio weights were computed by solving the quadratic programming problem in Equation (12). The top plots (viz. Figures 4a and 5a) suggest the optimal weights are time-independent, but regime-specific. Here, instead of assuming it *a priori*, we followed the more prudent approach of letting the ANN discover this property. The bottom plots (viz. Figures 4b and 5b) compare both allocation strategies displaying the weights as pie-charts. The full set of graphs (for all backtesting years 2007-2024 and all $\eta = 1, 2, 3, 4$ values) are available in the Supplement.

5.1 Benchmarking Results on Synthetic Data

First, we present a set of benchmarking comparisons based on synthetic data. As previously described, for each of the 18 years 2007-2018, the regime-specific model was trained on synthetic data simulated using the market dynamics on the 17 regime and returns data preceding each respective year. See Figure 6.

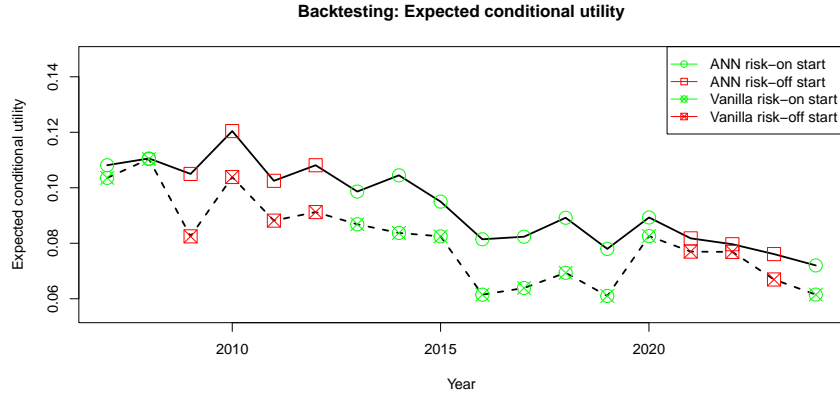
Using a Monte Carlo simulation to obtain $N = 100,000$ synthetic market paths, the wealth dynamics over 12 months and the annual rate of return as the end of the 12-month investment periods were computed assuming a monthly rebalancing based on the portfolios obtained with our regime-specific ANN method and the vanilla GBM approach. Figure 6a displays the estimated expected utility from the terminal wealth resulting from \$1 invested at the beginning of each respective year using both portfolio allocation approaches based on synthetic data. Comparing the two curves, one can see that regime-specific expected utility is never below its vanilla counterpart providing some basic evidence that our ANN-based optimization approach is a *bona fide* maximizer of the expected terminal wealth. Similarly, Figure 6b compares the expected rate of return (in %) under both allocation strategies. The pattern is generally similar to the previous plot. Finally, Figure 6 plots the *ex ante* empirical Sharpe ratio for each of the two methods computed from synthetic rates of return at the end of each 12-month period

$$\hat{S} = \frac{\hat{\mathbb{E}}[\text{RoR} - r]}{\widehat{\text{Var}}^{1/2}[\text{RoR} - r]} = \frac{\frac{1}{N} \sum_{b=1}^N (\text{RoR}^b - r)}{\sqrt{\frac{1}{N-1} \sum_{b=1}^N (\text{RoR}^b)^2 - \frac{N}{N-1} \left(\frac{1}{N} \sum_{b=1}^N \text{RoR}^b \right)^2}}$$

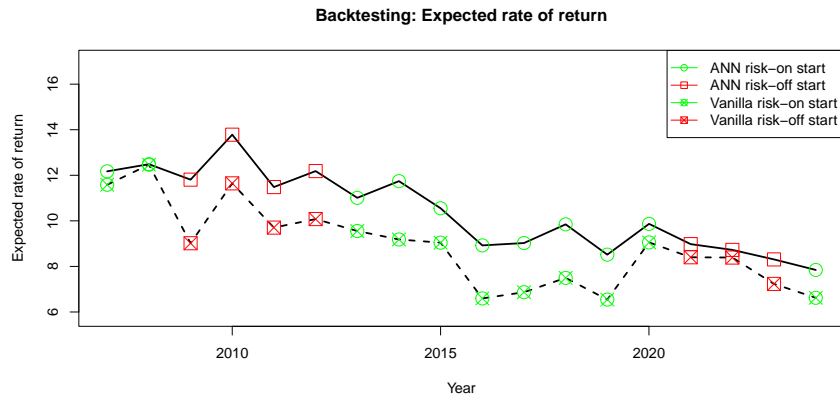
based on the N independent Monte Carlo replicates RoR^b ($b = 1, \dots, N$) of the 12-month rate of return, where r is the historical average annualized rate of return on cash over the previous 17 years. Most of the time, the regime-specific approach is on par or outperforms its vanilla competitor. Note that the ANN-based portfolio was trained to optimize the expected utility, not the Sharpe rate. Thus, occasional outperformance of the vanilla approach does not constitute any type of contradiction.

5.2 Backtesting Results on Historical Data

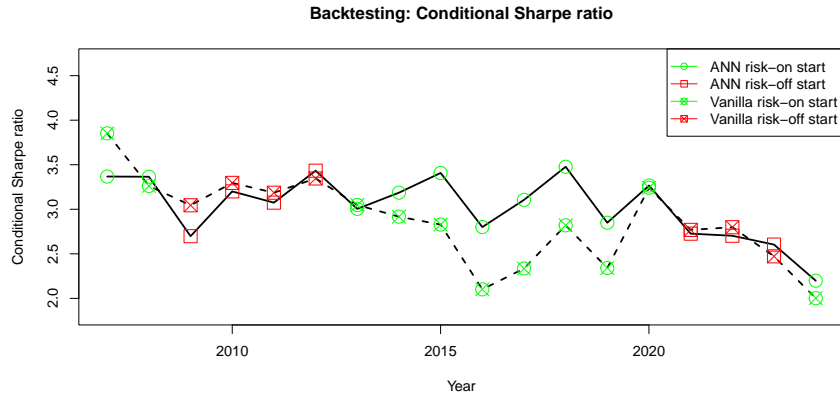
In a similar fashion, both methods were also backtested on historical data. For each of the 18-years of the backtesting time horizon 2007-2018, both our regime-specific ANN-based and the vanilla GBM-based portfolio allocation were computed using the 17 years preceding each given backtesting year. Then, recalibrating



(a) Average utility YoY for 2007-2024.



(b) Average rate of return (in %) YoY for 2007-2024.



(c) Sharpe ratio YoY for 2007-2024.

Figure 6: Average annual *ex ante* utility, annual rate return and Sharpe ratio YoY for 2007-2024.

the portfolio monthly based on these optimal weights, both monthly and annual returns on a 1\$ initial wealth for each of the two methods were computed for each year. For our regime-specific ANN approach, the regime expected to prevail was estimated and used for portfolio rebalancing at the beginning of each month based on VIX readings recorded at the end of each of the twelve months immediately preceding the current month as described in Section 3.1. Thus, any “prescience” about the future market regime is eliminated.

η	Average utility		Average rate of return	
	Regime-specific	Vanilla	Regime-specific	Vanilla
1	0.0682	0.0440	8.26%	5.42%
2	0.0540	0.0315	8.18%	5.00%
3	0.0406	0.0237	8.13%	4.98%
4	0.0250	0.0171	7.35%	4.86%

Table 3: Comparison of YoY performance metrics for our regime-specific ANN-based method and the vanilla approach over the 18-year backtesting period 2007-2024.

Table 3 summarizes the average terminal utility as well as annual return over the 18 backtesting years for the two methods across the four $\eta = 1, 2, 3, 4$ values investigated in this study. As the empirical evidence suggests, the regime-specific portfolio weights always outperform the vanilla ones in terms of both performance metrics across all η 's considered.

In addition to pooled performance metrics based on annual rates of return, we also computed *ex post* YoY performance metrics for each of the two methods using monthly portfolio returns associated with historical market data. The procedure is essentially identical to that outlined in Section 5.1 except that twelve historical portfolio returns in lieu of 100,000 simulated ones were considered. Additionally, the prevailing market regime was estimated empirically using the latest twelve monthly close VIX readings. Compared to the vanilla GBM approach, Figure 7 suggest that the regime-based portfolio allocation not only outperforms on average but also in large chunks of the historical period discussed. Over this period, the regime-based model seems to be either on par or consistently outperform (except with respect to the Sharpe ratio in the year 2024) the regime-agnostic vanilla competitor. For the post-pandemic period, it is possible that the simplistic two-regime model based on VIX lacks the ability to adequately capture the market regimes in this period. Indeed, inspecting the 2024 returns of the growth indices listed in Table 1, one can readily conclude that their performances was essentially independent of the VIX dynamics, leading to a behavior different from that in the past years. More sophisticated regime detection mechanisms have the potential to provide a remedy going forward.

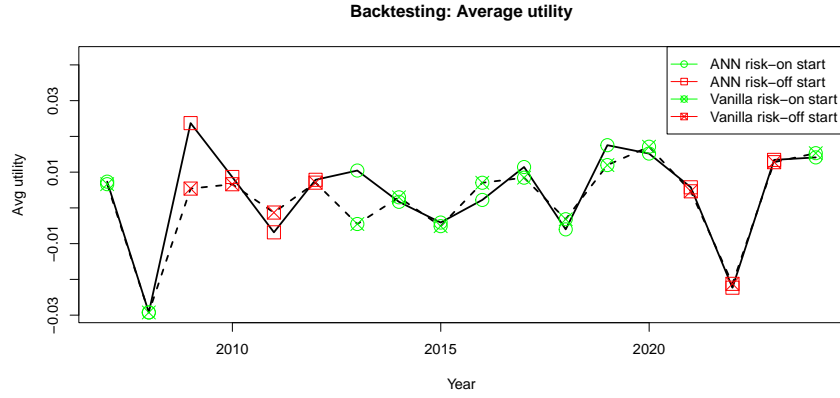
6 Summary and Conclusions

In this paper, we proposed an ANN-based optimal allocation strategy under market regime changes. Our method consists of training an ANN that regulates asset allocation based on the utility theory proposed by Merton (1969). To showcase the power of our method, we applied it to intertemporal wealth allocation across multiple assets subject to suitable portfolio constraints – a problem that has not been previously addressed in the literature. As a byproduct, we conclude that regimes should be taken into account by investors when making portfolio allocation decisions. Indeed, over the 18-year historic period 2007-2024 considered, the ANN-based regime-specific portfolio on average outperforms the vanilla Merton's allocation in terms of annual rate of return by approximately 2.49 to 3.18 percentage points (51% to 64%) across all power utilities studied. This outperformance is not limited to the period of the crisis but also persists during the 2010-2020 period. During 2020-2024, both approaches perform on par in terms of the average utility and average rate of return. However, the vanilla model never outperforms the regime-specific ANN model with respect to the expected utility. Therefore, adoption of regime-based models can reduce the overall risk.

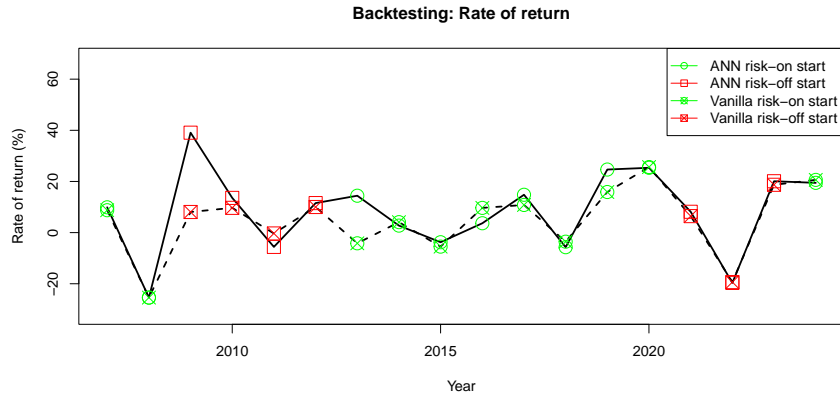
Supplementary Materials

The full set of figures produced as part of this study are available under the temporary link:

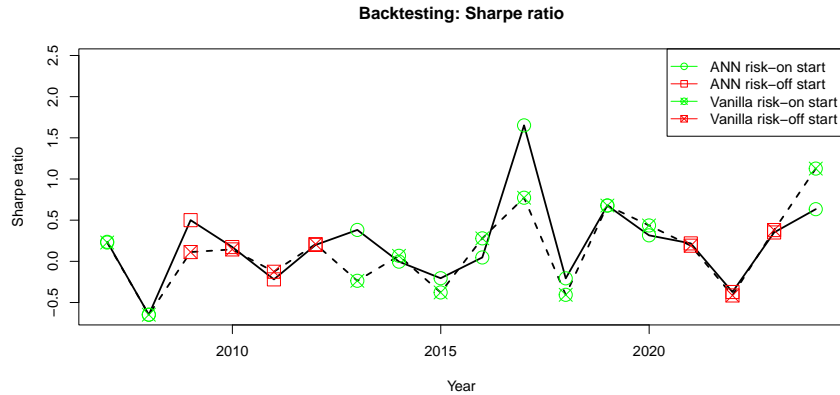
https://olddominion-my.sharepoint.com/:b:/g/personal/mpokojovy_odu_edu/EUcMkvtWV1dMvTagtef0jSvBiQQUsDcX.UJ1EWL-2MBiXw?e=7gm10h



(a) Average monthly utility YoY 2007-2024.



(b) Average monthly rate of return YoY 2007-2024.



(c) Sharpe ratio YoY 2007-2024.

Figure 7: Average monthly *ex post* utility, annual rate return and Sharpe ratio YoY 2007-2024.

A Euler-Maruyama Discretization

Consider an equispaced grid $I^{\Delta t} = \{t_k \mid t_k = (\Delta t)k, k = 0, \dots, n\}$ with a time step $\Delta t = T/n > 0$ for some $n \in \mathbb{N}$. The jump times H_k of the semi-Markov process $(\alpha_t)_t$ are generated using the principle of competing exponentials. Starting in a state $a \in \mathcal{S}$, the process remains in that state for $E_a \sim \text{Exp}(-q_{aa})$ years, after

which it jumps into a state b with the smallest value of $E_b \sim \text{Exp}(q_{ab})$, where all random variables are stochastically independent. Once generated, the jump time is truncated to the smallest grid point on $I^{\Delta t}$, and the procedure is repeated, thus giving rise to a discrete process $(\alpha_t^{\Delta t})_t$.

Further, the Euler-Maruyama approximation $W_{t_k}^{\Delta t}$ for the wealth process in Equation (19) reads as

$$W_{t_{k+1}}^{\Delta t} = W_{t_k}^{\Delta t} + (\Delta t)(\boldsymbol{\pi}_{t_k}^{\Delta t})'(\boldsymbol{\mu}(\alpha_{t_k}^{\Delta t}) - r(\alpha_{t_k}^{\Delta t})\mathbf{1}_p) + (\boldsymbol{\pi}_{t_k}^{\Delta t})'\boldsymbol{\Sigma}^{1/2}(\alpha_{t_k}^{\Delta t})\Delta\mathbf{B}_{t_k} \quad \text{for } k \geq 0 \quad (40)$$

$$W_{t_0}^{\Delta t} = w^0, \quad (41)$$

where $\Delta\mathbf{B}_{t_k} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mathbf{0}, (\Delta t)\boldsymbol{\Sigma})$ with the discretized stock weight process $\boldsymbol{\pi}_{t_k}^{\Delta t}$.

B Some Formal Definitions

Similar to Albosaily and Pergamenchtchikov (2021), the set of admissible controls for the classical Merton problem can be defined as

$$\begin{aligned} \mathcal{A}_{t,w} &:= \{ \boldsymbol{\pi} \mid (\boldsymbol{\pi}_s)_{s \in [t,T]} \text{ is } (\mathcal{F}_{t,s})_s\text{-adapted, progressively measurable with c\^adl\^ag paths,} \\ &\boldsymbol{\pi}_s \in \mathcal{K} \text{ for a.e. } s \in [t, T] \text{ and } \int_t^T |\boldsymbol{\pi}_s|^2 ds < \infty \text{ } \mathbb{P}\text{-a.s.,} \\ &(W_s^{\boldsymbol{\pi};t,w})_{s \in [t,T]} \text{ is a unique strong solution such that } W_t^{\boldsymbol{\pi};t,w} = w, \\ &W_s^{\boldsymbol{\pi};t,w}(\omega) > 0 \text{ holds } \mathbb{P} \otimes \nu \text{ a.e. in } [t, T] \times \Omega, \quad \mathbb{E} \left[\text{ess sup}_{s \in [t,T]} U(W_s^{\boldsymbol{\pi};t,w})_- \right] < \infty \}, \end{aligned} \quad (42)$$

where $x_- = \max\{0, -x\}$ denotes the negative part of x , ν is the Borel measure on $[0, \infty)$ and the filtration

$$\mathcal{F}_{t,s} = \sigma(\{\mathbb{E}[\mathbf{B}_\zeta \mid \mathbf{B}_t] \mid t \leq \zeta \leq s\}) \quad (43)$$

on $(\Omega, \mathcal{F}, \mathbb{P})$ is completed by all null sets.

Under regime changes, consider instead the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{t \geq 0}, \mathbb{P})$ with the filtration

$$\mathcal{F}_{t,s;a} = \sigma(\{\mathbb{E}[\mathbf{B}_\zeta \mid \mathbf{B}_t] \mid t \leq \zeta \leq s\}) \otimes \sigma(\{\mathbb{E}[\alpha_\zeta \mid \alpha_t = a] \mid t \leq \zeta \leq s\}) \quad (44)$$

completed by all null sets and introduce the set of admissible controls $\mathcal{A}_{t,w,a}$ defined as

$$\begin{aligned} \mathcal{A}_{t,w,a} &:= \{ \boldsymbol{\pi} \mid (\boldsymbol{\pi}_s)_{s \in [t,T]} \text{ is } (\mathcal{F}_{t,s;a})_s\text{-adapted, progressively measurable with a.s. continuous paths,} \\ &\boldsymbol{\pi}_s \in \mathcal{K} \text{ for a.e. } s \in [t, T] \text{ and } \int_t^T |\boldsymbol{\pi}_s|^2 ds < \infty \text{ } \mathbb{P}\text{-a.s.,} \\ &(W_s^{\boldsymbol{\pi};t,w,a})_{s \in [t,T]} \text{ is a unique strong solution such that } W_t^{\boldsymbol{\pi};t,w,a} = w, \\ &W_s^{\boldsymbol{\pi};t,w,a}(\omega) > 0 \text{ holds } \mathbb{P} \otimes \nu \text{ a.e. in } [t, T] \times \Omega, \quad \mathbb{E} \left[\text{ess sup}_{s \in [t,T]} U(W_s^{\boldsymbol{\pi};t,w,a})_- \right] < \infty \}. \end{aligned} \quad (45)$$

C Portfolio Constraints and Orthogonal Projector

With the diversification constraints in Table 2, the definition of the constraint set reads as

$$\mathcal{K} = \left\{ \boldsymbol{\pi} \in [0, 1]^p \mid \sum_{i=1}^p \pi^i \leq 1, l_k \leq \sum_{i \in \mathcal{I}_k} \pi^i \leq u_k \text{ for } k = 0, 1, \dots, 5 \right\}.$$

Further, the orthogonal projector operator is defined as

$$\Pi_{\mathcal{K}}(\boldsymbol{\pi}) = \arg \min_{\tilde{\boldsymbol{\pi}} \in \mathcal{K}} \|\tilde{\boldsymbol{\pi}} - \boldsymbol{\pi}\|^2 \quad (46)$$

and can be efficiently computed by solving the underlying quadratic programming problem (Goldfarb and Idnani, 1983). The projection mapping $\Pi_{\mathcal{K}}(\cdot)$ is known to be Lipschitz with a constant 1 and differentiable in an open dense subset of \mathbb{R}^p (Wolfe, 1976).

References

- Albosaily, S. and Pergamenchtchikov, S. (2021). Optimal investment and consumption for multidimensional spread financial markets with logarithmic utility. *Stats*, 4:1012–1026.
- Ang, A. and Timmermann, A. (2012). Regime changes and financial markets. *Annual Review of Financial Economics*, 4:313–337.
- Bachouch, A., Huré, C., Langrené, N., and Pham, H. (2022). Deep neural networks algorithms for stochastic control problems on finite horizon: Numerical applications. *Methodology and Computing in Applied Probability*, 24:143–178.
- Davey, A. and Zheng, H. (2022). Deep learning for constrained utility maximisation. *Methodology and Computing in Applied Probability*, 24:661–692.
- Fleming, W. H. and Soner, H. M. (2006). *Controlled Markov Processes and Viscosity Solutions*, volume 25. Springer Science & Business Media, New York, NY.
- Gim, D. and Park, H. (2021). A deep learning algorithm for optimal investment strategies.
- Goldfarb, D. and Idnani, A. (1983). A numerically stable dual method for solving strictly convex quadratic programs. *Mathematical Programming*, 27(1):1–33.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press, Cambridge, MA.
- Grohs, P., Hornung, F., Jentzen, A., and Von Wurstemberger, P. (2023). *A proof that artificial neural networks overcome the curse of dimensionality in the numerical approximation of Black–Scholes partial differential equations*, volume 284 of *Memoirs of the American Mathematical Society*. AMS Press, Providence, RA.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57:357–384.
- Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4:251–257.
- Kingma, D. and Ba, J. (2014). Adam: A method for stochastic optimization. <https://arxiv.org/abs/1412.6980>.
- Kopeliovich, Y. and Pokojovy, M. (2024). Portfolio optimization with feedback strategies based on artificial neural networks. *Finance Research Letters*, 69:106185.
- MacKay, D. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, Cambridge, UK.
- Markowitz, H. (1952). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Journal of Finance*, 7(1):77–91.
- Merton, R. (1969). Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51:247–257.
- Pourahmadi, M. (2013). *High-Dimensional Covariance Estimation: With High-Dimensional Data*, volume 882 of *Wiley Series in Probability and Statistics*. John Wiley & Sons, Hoboken, NJ.
- Reppen, A. M. and Soner, H. M. (2023). Deep empirical risk minimization in finance: Looking into the future. *Mathematical Finance*, 33:116–145.
- The Economist. Buttonwood (Online Column) (2023). How to avoid a common investment mistake. <https://www.economist.com/finance-and-economics/2023/09/21/how-to-avoid-a-common-investment-mistake>.

- van Staden, P., Forsyth, P., and Li, Y. (2023). A parsimonious neural network approach to solve portfolio optimization problems without using dynamic programming.
- Vapnik, V. N. (1997). The support vector method. In W. Gerstner et al., editor, *Artificial Neural Networks – ICANN '97*, pages 261–271, Berlin, Heidelberg. Springer Verlag.
- Wolfe, J. M. (1976). Differentiability of nonlinear best approximation operators in a real inner product space. *Journal of Approximation Theory*, 16(4):341–346.
- Zariphopoulou, T. (2001). A solution approach to valuation with unhedgeable risks. *Finance and Stochastics*, 5(1):61–82.
- Zuckerman, G. (2019). *The Man Who Solved the Market: How Jim Simons Launched the Quant Revolution*. Portfolio / Penguin, New York, NY.