MSDS 451 Programming Assignment 2
Portfolio Optimization using the Monte Carlo Method

Abhijit Joshi
MSDS-451-DL: Financial Engineering
Northwestern University
Dr. Thomas Miller
October 19, 2025

1. Problem Description

Portfolio optimization is an essential concept in finance that aims to identify the optimal allocation of assets within an investment portfolio to achieve the best possible balance between risk and return. From an investor's point of view, the goal is typically to maximize returns for investments made; however, from a portfolio manager's point of view, it is vital to balance risks and returns, in which the portfolio analysis plays an important role. This project assignment explores a Monte Carlo method to assess risk-return trade-offs across a multitude of different weights for assets within the portfolio through a simulation-based approach. A key aspect of this experiment is also to compare and contrast Long Only and Shorts OK strategies and their impact on the efficiency frontier and overall portfolio returns.

Key Words: *Monte Carlo, portfolio, optimization, efficient frontier, Sharpe ratio, risk, returns, long, short*

2. Data Preparation and Pipeline

2.1. Scenario 1

Recreate the R experiment in Python (Jupyter notebook). The mean, standard deviation vectors, and the covariance matrix for the four assets were the same as those used in the jump-start R program.

2.2 Scenario 2

Personal ETF. Four specific Exchange Traded Funds **(FXAIX, QQQM, SCHD, SFLNX)** were selected for analysis. These were picked from my personal HSA portfolio to see if the portfolio was balanced.

2.2.1. **Data Acquisition:** Historical price data for these ETFs was obtained using the **yfinance** library. The adjusted closing prices were specifically extracted to ensure the returns accurately accounted for corporate actions such as dividends and stock splits.

2.2.2. **Return Calculation:** Daily simple returns were calculated from the adjusted closing prices using the percentage change method, and any resulting initial missing values were removed.

2.2.3. **Annualization:** The annualized mean returns were calculated by multiplying the mean of the daily returns by 252 (the assumed number of trading days in a year). The annualized standard deviations were calculated by multiplying the standard deviation of the daily returns by the square root of the number of trading days.

2.3.4. **Covariance Matrix Construction:** The correlation matrix was computed directly from the daily returns. Finally, the custom function **cov_from_sd_corr** was utilized to construct the necessary **covariance matrix** from the estimated annualized standard deviations and the correlation matrix. This matrix is a critical input for accurately simulating portfolio risk and correlated returns.

3. Research Design
The research design uses a Monte Carlo simulation approach (provided by Dr. Miller) to find the potential portfolio compositions and their corresponding risk-return profiles. This method involves generating a large number of random weights using multivariate normal functions. Two strategies are considered: one allowing only long positions (weights between 0 and 1, summing to 1) and another allowing both long and short positions (weights that can be negative but still sum to 1).

3.1. Portfolio Generation: The make_weights function was used to generate a large number of random portfolio weight vectors. Two primary investment regimes were simulated:
◦ **Long Positions Only**: Asset weights were restricted to be non-negative (between 0 and 1) and sum to 1.
◦ **Shorts OK**: Asset weights were allowed to be negative (representing short positions) but still constrained to sum to

3.2. Return Simulation: For each experiment, the simulate_returns function was used to generate realistic asset returns. This function relies on the multivariate normal distribution (from scipy.stats) and uses the estimated annualized mean vector and the covariance matrix as inputs to generate correlated asset returns over a specified sample size (e.g., 700)

3.3. Portfolio Evaluation: The evaluate_portfolios function calculates key performance metrics for every generated portfolio under simulated returns. These metrics include: mean return, standard deviation (as a measure of risk), Sharpe Ratio, Value at Risk (VaR), and Conditional Value at Risk (CVaR).

3.4. Experiment Orchestration: The run_experiment function integrates these steps, executing the simulation for both the "Long Positions Only" and "Shorts OK" regimes to allow for a direct comparison of the resulting feasible sets and efficient frontiers

4. Programming
This project leverages several key Python libraries for data handling, simulation, and analysis. The libraries are Pandas, numpy, and multivariate_normal from scipy.stats. The following tables list the functions created in the Jupyter Notebook.

Function Name	Description
cov_from_sd_corr	Calculates the covariance matrix from a vector of standard deviations and a correlation matrix

make_weights	Generates random portfolio weight vectors, ensuring weights sum to 1 and incorporating the option for negative weights (shorts_ok=True)
simulate_returns	Generates simulated time series of correlated asset returns using the multivariate normal distribution
evaluate_portfolios	Calculates performance metrics (mean return, standard deviation, Sharpe Ratio, VaR, CVaR) for all portfolios defined by a matrix of weights
run_experiment	Orchestrates the entire simulation process, including data preparation, return simulation, weight generation for both regimes, and portfolio evaluation
plot_results	Visualizes the output as scatter plots showing portfolio risk versus return, helping to identify the efficient frontier and compare the performance of different scenarios and strategies

5. Exposition

The analysis of the Personal ETF scenario (FXAIX, QQQM, SCHD, SFLNX) provided significant insights into understanding portfolio performance, risks, and returns, and their interrelationships.

5.1. Long only vs. Shorts OK

The risk-return scatter plots clearly visualize the efficiency frontier of simulated portfolio returns and weights. A key finding across the simulations was that the **"Shorts OK"** regime generated a significantly larger efficient frontier compared to the **"Long Positions Only"** regime.
The **portfolio with the highest Sharpe Ratio** is a single critical point located on the efficient frontier. This point represents the **tangency portfolio** under the assumption that the risk-free rate is zero (the default assumption for the Sharpe Ratio calculation in this simulation)

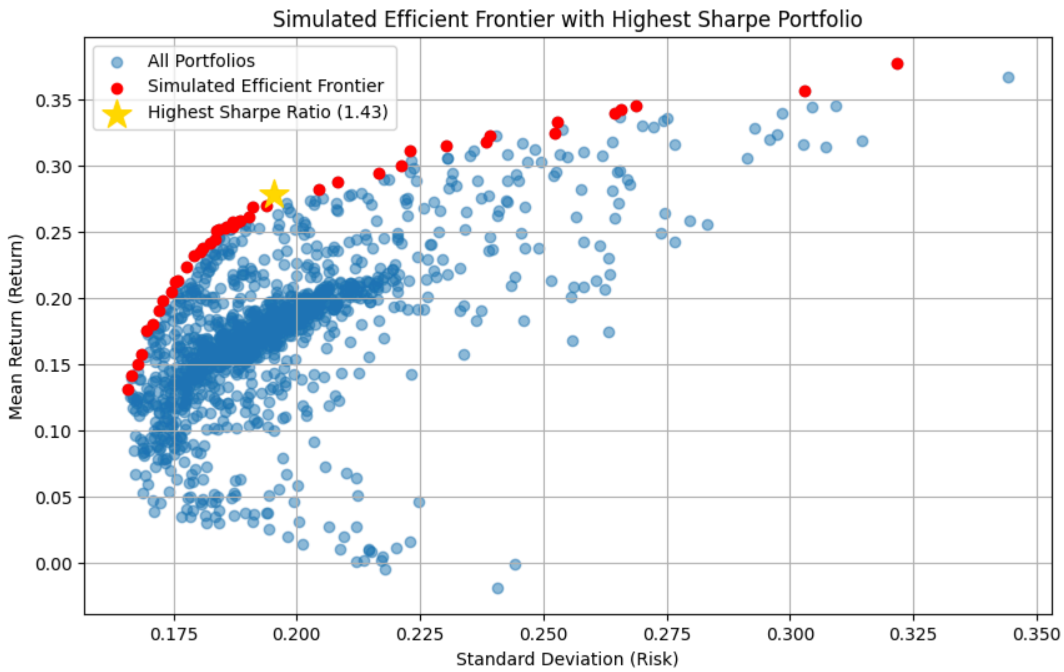

Fig.1 - Efficient frontier highlighted for personal ETF Portfolio. The Tangency Portfolio (1.43) is shown on the efficient frontier.

5.2. Best Sharpe Ratio and Weights

The highest Sharpe Ratio portfolio in the "Shorts OK" strategy exhibited extreme, non-diversified weights: a large **leveraged long position** in SFLNX (2.82) and substantial **short positions** in FXAIX (−0.78), SCHD (−1.00), and QQQM (−0.04)

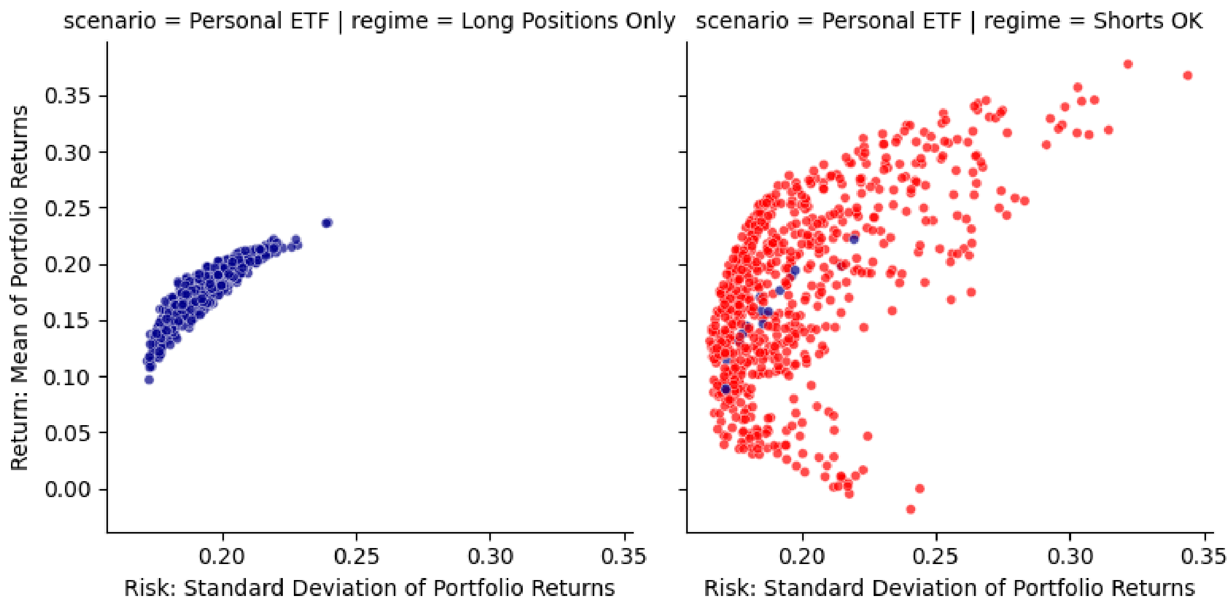

Fig. 2 - Shorts OK strategy generated a broader efficiency frontier, indicating higher returns for the same risks.

The highest Sharpe Ratio achieved in the "Shorts OK" strategy was significantly higher (approximately 1.425) than the maximum achieved in the "Long Positions Only" regime (approximately 1.017). Furthermore, **all of the top 10 portfolios** identified by the Sharpe Ratio belonged to the "Shorts OK" strategy.

5.3. Risk and Return Metric Relationships

For the portfolio simulations returns/risks lying on the efficient frontier, an analysis of the correlation between performance and risk metrics revealed strong relationships:

- There was a **strong positive correlation** between the Sharpe Ratio and the 95% Value at Risk (VaR) (~0.927).
- There was also a **positive correlation** between the Sharpe Ratio and the 95% Conditional Value at Risk (CVaR) (~0.677).

The positive correlations suggest that for the optimal portfolios in the simulation, improvements in overall efficiency (higher Sharpe Ratio) generally coincided with a reduction in estimated downside loss metrics (less negative VaR and CVaR).

In summary, the results demonstrated that while traditional diversification is fundamental, in a portfolio with highly correlated assets, such as my Personal ETF, shorting provides a flexible balancing method. Shorting allows the weights to be adjusted such that positive positions can be extreme in the highest performing assets. In contrast, the positions in the other correlated assets can be shorted as a way to hedge against those extreme long positions to provide the best returns/risks performance. This experiment was indeed insightful in understanding portfolio optimization.