An Economic Analysis of Optimal Investment Strategies for Accumulating Housing Down Payments

Business Analytics MS Capstone Project

Frank Paul Longo II

July 26, 2024

Abstract

This research conducts a detailed analysis of optimal investment strategies tailored for first-time homebuyers seeking to accumulate down payments over 5, 7.5, and 10-year horizons. By integrating Modern Portfolio Theory (MPT), the Capital Asset Pricing Model (CAPM), and Monte Carlo simulations, the study provides practical insights into constructing investment portfolios that are age-specific and balance risk and return effectively.

The findings reveal that customized investment strategies substantially improve the ability to save for a down payment, potentially shortening the time needed to achieve homeownership. The importance of considering risk-adjusted returns and diversification in portfolio construction is underscored, demonstrating their critical roles in enhancing savings outcomes.

Future research could broaden this analysis by incorporating additional demographic factors such as income variations and regional housing market dynamics. Furthermore, exploring the inclusion of alternative investment vehicles, such as real estate investment trusts (REITs) and cryptocurrencies, could offer deeper insights into optimizing investment strategies for down payments.

Contents

1 Introduction

The escalating housing costs in contemporary real estate markets have created significant barriers for first-time homebuyers. This demographic often faces the daunting task of accumulating substantial down payments amidst economic volatility and uncertain income trajectories. This research addresses this critical issue by developing and evaluating optimal investment strategies tailored to help diverse age groups achieve their homeownership goals within 5, 7.5, and 10-year horizons.

1.1 Research Question & Objective

What are the most effective investment strategies for different age groups to accumulate housing down payments over periods of 5, 7.5, and 10 years? The primary objective of this study is to identify, analyze, and optimize investment strategies that can effectively assist first-time homebuyers in saving for their down payments. By leveraging advanced financial theories and empirical methodologies, this research aims to provide actionable insights that balance risk and return, offering practical solutions for prospective homeowners (Markowitz, 1952; Sharpe, 1964; Boyle, 1977).

1.2 Motivation

As homeownership becomes increasingly out of reach for many, particularly younger individuals, it is imperative to develop strategies that can mitigate these barriers. By providing evidence-based investment strategies, this study aims to empower individuals with the tools needed to navigate the complexities of financial planning for homeownership.

2 Data

The financial data used in this research is sourced from Yahoo Finance (Yahoo Finance, nd), which includes comprehensive information on roughly 150 securities consisting of stocks, mutual funds, and ETFs. The historical data (where we assume for this analysis current day is November 7th, 2014) covers the period from November 13th, 2009 to November 7th, 2014 at daily frequency. Also included is hindsight data (since it isn't actually November 7th, 2014) from November 13th, 2009 to July 23rd, 2024. This timeframe allows for the analysis of recent trends and the performance of different asset classes in various market conditions.

The dataset includes the following fields:

- Open: Price at the beginning of the trading day.
- High: Peak price during the trading day.
- Low: Lowest price during the trading day.
- Close: Price at the end of the trading day.
- Adj Close: Closing price adjusted for dividends, stock splits, etc.
- Volume: Number of shares traded during a single trading day.
- Type: Security type (e.g., stock, ETF).

3 Literature Review

3.1 Portfolio Selection by Harry Markowitz (1952)

3.1.1 Key Findings

In "Portfolio Selection," Harry Markowitz introduced the concept of Modern Portfolio Theory (MPT), which fundamentally changed the way investors approach portfolio construction. Markowitz emphasized the trade-off between risk and return, proposing that investors should diversify their investments across various assets to minimize risk without sacrificing expected returns. The main contribution of this paper is the efficient frontier, which represents the set of portfolios that offer the highest expected return for a given level of risk. The efficient frontier helps investors identify the optimal asset allocation that maximizes returns for a specific risk level.

> *"An investor should diversify to achieve maximum return for a given level of risk, leading to the creation of the efficient frontier."* (Markowitz, 1952).

Markowitz's work introduced the concept of mean-variance optimization, where the expected return of a portfolio is calculated as the weighted sum of the expected returns of individual assets, and the risk is measured as the variance of portfolio returns. This approach allows investors to construct portfolios that either maximize expected return for a given level of risk or minimize risk for a given level of expected return.

3.1.2 Implications for This Study

Markowitz's findings are fundamental to this study as they provide the basis for constructing diversified investment portfolios aimed at accumulating down payments for housing. By applying the principles of MPT, this study seeks to identify optimal asset allocations that maximize returns while managing risk over different investment horizons. The efficient frontier concept is used to determine the set of optimal portfolios that offer the highest expected return for a specified risk level, helping first-time homebuyers to achieve their financial goals more effectively.

3.1.3 Critiques

Despite its groundbreaking contributions, MPT has faced criticism for its reliance on historical data to estimate future returns and risks, which may not always be accurate. Additionally, the assumption of normally distributed returns has been questioned, as real-world financial returns often exhibit skewness and kurtosis. Some critics argue that MPT does not account for extreme events and tail risks, which can significantly impact portfolio performance. Furthermore, the model's reliance on mean-variance optimization may not fully capture investors' preferences and risk tolerances, leading to suboptimal investment decisions.

3.2 Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk by William F. Sharpe (1964)

3.2.1 Key Findings

In his seminal paper, William F. Sharpe introduced the Capital Asset Pricing Model (CAPM), which provides a framework to determine the expected return of an asset based on its systematic risk, as measured by beta. The model posits that the expected return on an asset is a function of the risk-free rate, the asset's beta, and the market risk premium. Sharpe also developed the Sharpe Ratio, a measure of risk-adjusted return, which helps investors understand how much excess return they are receiving for the extra volatility endured.

> *"The expected return on a security is linearly related to its beta, reflecting its sensitivity to market movements."* (Sharpe, 1964).

3.2.2 Implications for This Study

Sharpe's CAPM and Sharpe Ratio are crucial for assessing the risk and return profiles of individual securities within the portfolios analyzed in this study. By evaluating the beta and risk-adjusted returns, this study can better understand the performance of different assets and optimize the investment strategies accordingly. The CAPM provides a systematic approach to estimating the expected returns of assets based on their systematic risk, while the Sharpe Ratio helps in comparing the performance of investments on a risk-adjusted basis.

3.2.3 Critiques

The CAPM has been criticized for its assumption of a single-period investment horizon and the use of a single market index to represent the entire market. Additionally, the model's reliance on beta as the sole measure of risk has been questioned, as it does not account for other factors that may influence asset returns. Empirical studies have shown that other variables, such as size and value factors, also play a significant role in explaining asset returns, leading to the development of multi-factor models. Moreover, the CAPM's assumption of a frictionless market and homogeneous expectations among investors may not hold true in real-world scenarios.

3.3 Options: A Monte Carlo Approach by Phelim P. Boyle (1977)

3.3.1 Key Findings

Phelim P. Boyle's paper introduced the use of Monte Carlo methods for option pricing, providing a robust tool to model the uncertainty and variability in financial investments. Monte Carlo simulations generate a large number of random samples from the probability distributions of asset returns, allowing investors to assess the impact of risk and uncertainty on their portfolios. This method has since been widely adopted for various financial applications, including risk management, portfolio optimization, and scenario analysis.

> *"Monte Carlo simulations enable the modeling of complex financial instruments and the assessment of risk and return in uncertain environments."* (Boyle, 1977).

The Monte Carlo simulation process involves generating random returns based on historical data and iterating this process to build a distribution of potential outcomes. By running multiple simulations, investors can estimate the expected value and variability of the investment portfolio, providing insights into the likelihood of achieving their financial goals.

3.3.2 Implications for This Study

Boyle's Monte Carlo methods are integral to this study for forecasting the performance of optimized portfolios under various market conditions. By simulating different scenarios, this study aims to estimate the range of potential outcomes and assess the likelihood of achieving the desired down payment amount within the specified time horizon. The use of Monte Carlo simulations allows for a comprehensive analysis of the potential risks and returns associated with different investment strategies, providing valuable insights for first-time homebuyers.

3.3.3 Critiques

One critique of Monte Carlo simulations is their computational intensity, which can be demanding for large-scale applications. Additionally, the accuracy of the results is heavily dependent on the quality of the input data and the assumptions made about the probability distributions of asset returns. Some critics argue that Monte Carlo methods may not fully capture the complexities of financial markets, such as changing market conditions and behavioral factors. Moreover, the reliance on historical data may not accurately predict future market behavior, leading to potential misestimations of risk and return.

4 Theoretical Models

4.1 Capital Asset Pricing Model (CAPM)

The Capital Asset Pricing Model (CAPM) is a fundamental tool in finance used to determine the expected return on an asset based on its systematic risk, as measured by beta (β_i). The CAPM formula is:

$$E(R_i) = R_f + \beta_i(E(R_m) - R_f) \tag{1}$$

where:

- $E(R_i)$ is the expected return on asset i.
- R_f is the risk-free rate of return.
- β_i is the beta of asset i, representing its sensitivity to market movements.
- $E(R_m)$ is the expected return of the market.
- $(E(R_m) - R_f)$ is the market risk premium.

4.1.1 Beta Calculation

Beta (β_i) measures the volatility of an asset in relation to the market. It is calculated as:

$$\beta_i = \frac{\text{Cov}(R_i, R_m)}{\sigma_m^2} \tag{2}$$

where:

- $\text{Cov}(R_i, R_m)$ is the covariance between the return of asset i and the return of the market.
- σ_m^2 is the variance of the market return.

4.1.2 Security Market Line (SML)

The Security Market Line (SML) is a graphical representation of the CAPM, showcasing the relationship between the expected return of an asset and its systematic risk, as measured by beta (β_i).

4.2 Sharpe Ratio

The Sharpe Ratio, developed by William F. Sharpe, is a measure of risk-adjusted return. The Sharpe Ratio provides a way to compare the performance of investments while considering the risk taken. A higher

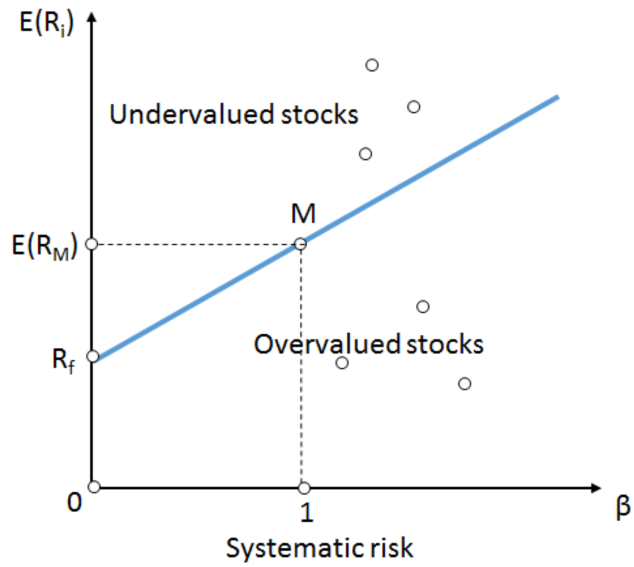

Figure 1: Security Market Line ????Source: Wikipedia????

(a) The Security Market Line (SML) plots the expected return of an asset against its beta, showing that higher systematic risk (beta) correlates with higher expected returns. Correctly priced securities should lie on the SML. The slope represents the market risk premium ($E(R_m) - R_f$), while the intercept is the risk-free rate (R_f).

Sharpe Ratio indicates better risk-adjusted performance, meaning the investment provides higher returns for each unit of risk taken. It is calculated as follows:

$$\text{Sharpe Ratio} = \frac{E(R_i) - R_f}{\sigma_i} \tag{3}$$

where:

- $E(R_i)$ is the expected return of the investment.
- R_f is the risk-free rate.
- σ_i is the standard deviation of the investment's return.

4.3 Composite Score Calculation

The Composite Score integrates multiple metrics to rank securities. Weights are assigned to each metric to reflect their importance in the ranking process. It is calculated as follows:

$$\text{Composite Score} = w_\beta \beta + w_{\text{Sharpe}} \text{Sharpe Ratio} + w_{\text{CAPM}} E(R_i) + w_{\text{Actual}} \text{Actual Returns} \tag{4}$$

where:

- w_β is the weight assigned to the beta.
- w_{Sharpe} is the weight assigned to the Sharpe Ratio.
- w_{CAPM} is the weight assigned to the CAPM predicted return.
- w_{Actual} is the weight assigned to the actual returns.

4.4 Modern Portfolio Theory (MPT)

Modern Portfolio Theory (MPT), developed by Harry Markowitz, (Markowitz, 1952) provides a robust framework for constructing an optimal portfolio that maximizes expected return for a given level of risk. The expected return $E(R_p)$ of a portfolio is the weighted sum of the expected returns of the individual assets:

$$E(R_p) = \sum_{i=1}^{n} w_i E(R_i) \tag{5}$$

where:

- $E(R_p)$ is the expected return of the portfolio.
- w_i are the weights of the individual assets in the portfolio.
- $E(R_i)$ is the expected return of asset i.

4.4.1 Efficient Frontier and Optimal Portfolio

The efficient frontier is a concept from MPT that represents the set of optimal portfolios offering the highest expected return for a defined level of risk. The process of constructing the efficient frontier involves solving the following optimization problem:

$$\min \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} \tag{6}$$

subject to:

$$\sum_{i=1}^{n} w_i = 1 \tag{7}$$

and

$$E(R_p) = \sum_{i=1}^{n} w_i E(R_i) \tag{8}$$

where:

- σ_{ij} is the covariance between the returns of assets i and j.
- w_i and w_j are the weights of assets i and j in the portfolio.

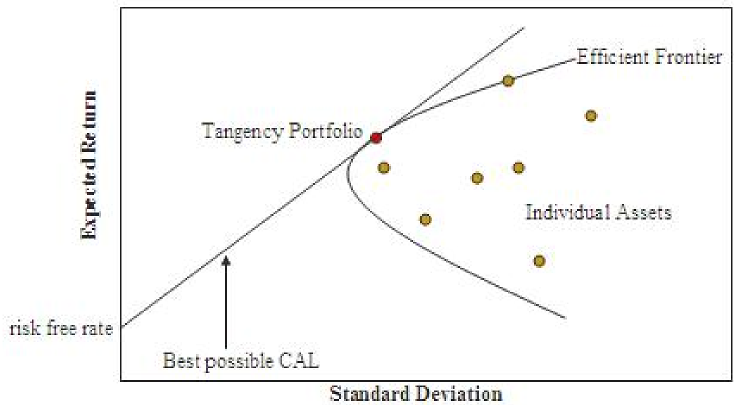

Figure 2: Efficient Frontier

(a) T
his figure shows the efficient frontier, which illustrates the optimal portfolios that offer the maximum expected return for a given level of risk. Portfolios that lie below the efficient frontier are sub-optimal because they do not provide enough return for the level of risk taken.

4.5 Monte Carlo Simulation

Monte Carlo simulations are utilized to model the uncertainty and variability in investment returns over time. The simulation process involves generating random returns based on historical data and iterating this process to build a distribution of potential outcomes. By running multiple simulations, we can estimate the expected value and variability of the investment portfolio, providing insights into the likelihood of achieving the desired down payment amount within the specified time horizon. The value of an investment at time i is given by:

$$X_i = X_{i-1} \times (1 + r_i) \tag{9}$$

where:

- X_i is the investment value at time i.

- X_{i-1} is the investment value at time $i-1$.
- r_i is the return for period i.

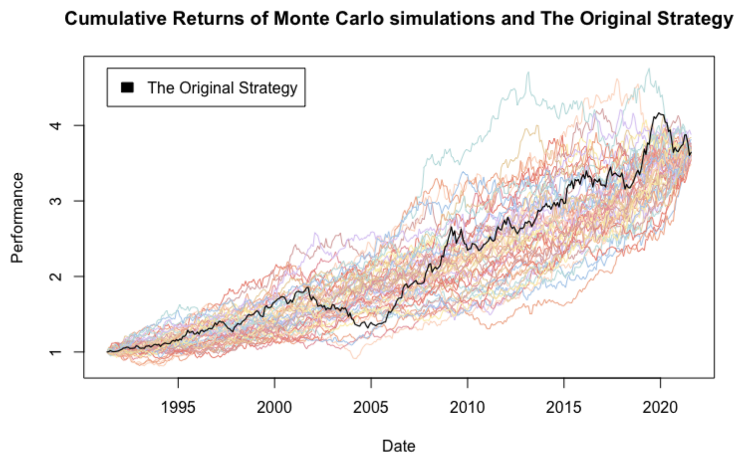

Figure 3: Investment Simulation Process for Monte Carlo Analysis

(a) T
he diagram shows how random returns are generated and used to project the future value of an investment over multiple iterations, creating a range of possible outcomes and helping to understand the potential risks and returns.

5 Empirical Specification

5.1 Model Implementation

The empirical analysis involves a systematic approach to evaluating and optimizing investment strategies using historical data. The data is cleaned and adjusted for corporate actions like stock splits and dividends using the `yfinance_data.py` script to ensure consistency, resulting in the creation of a file named `yfinance_data.csv`.

5.1.1 CAPM and Sharpe Ratio Calculation

Each security's risk and return profile is assessed using the Capital Asset Pricing Model (CAPM) and Sharpe Ratio. This analysis is performed using the `init_filtering.py` script.

1. Calculating the average return of the market index (e.g., S&P 500).
2. Determining the risk-free rate (e.g., 10-year U.S. Treasury bonds yield).
3. Computing the beta (β_i) of each security:

$$\beta_i = \frac{\mathrm{Cov}(R_i, R_m)}{\mathrm{Var}(R_m)} \tag{10}$$

4. Estimating the expected return using the CAPM formula:

$$E(R_i) = R_f + \beta_i(E(R_m) - R_f) \tag{11}$$

5. Calculating the Sharpe Ratio:

$$\text{Sharpe Ratio} = \frac{E(R_i) - R_f}{\sigma_i} \tag{12}$$

5.1.2 Composite Score Calculation

Securities are ranked based on a composite score that integrates multiple metrics:

$$\text{Composite Score} = w_\beta \beta + w_{\text{Sharpe}} \text{Sharpe Ratio} + w_{\text{CAPM}} E(R_i) + w_{\text{Actual}} \text{Actual Returns} \tag{13}$$

Weights are assigned to each metric to reflect their importance in the ranking process. A file is then produced for each time horizon with the ranked securities in order, named `top_assets_composite_score.csv`.

5.1.3 Modern Portfolio Theory (MPT) Application

Using the ranked securities, portfolios are optimized for different investment horizons (5, 7.5, and 10 years) using Modern Portfolio Theory (MPT). The optimization involves solving the following quadratic programming problem:

$$\text{Maximize} \quad \mathbf{w}^{T-} - \frac{\lambda}{2}\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w} \tag{14}$$

$$\text{subject to} \quad \sum_i w_i = 1 \tag{15}$$

$$w_i \geq 0 \quad \forall i \tag{16}$$

where $\mathbf{w}$ is the vector of asset weights, ¯ is the vector of expected returns, $\boldsymbol{\Sigma}$ is the covariance matrix of returns, and λ is the risk aversion parameter. This process is implemented using the `mpt.py` script to produce a data file for each time horizon named `optimal_weights.csv`.

5.1.4 Performance Comparison

The optimized portfolios are compared against actual historical performance using the hindsight data created by `hindsight_data.py`. The comparison includes calculating the cumulative returns of the portfolios and benchmarking them against the S&P 500 index. This step is performed using the `hindsight.py` script to create a data file named `hindsight_data.csv`.

5.1.5 Monte Carlo Simulation for Future Forecasting

Monte Carlo simulations are conducted to forecast the performance of the optimized portfolios. This comprehensive simulation is implemented using the `mcs.py` script, providing insights into the potential future performance of the portfolios.

1. Defining initial investment amounts and annual contributions.
2. Generating random returns based on historical distributions.
3. Calculating portfolio values at each time step:

$$V_t = V_{t-1} \times (1 + R_t) + C \tag{17}$$

4. Running multiple iterations to build a probability distribution of outcomes.

5. Applying economic shocks to simulate real-world scenarios:

$$V_t = V_t \times (1 + \text{Shock Intensity}) \tag{18}$$

6 Results

6.1 Exploratory Data Analysis

6.1.1 Distribution of Securities by Type

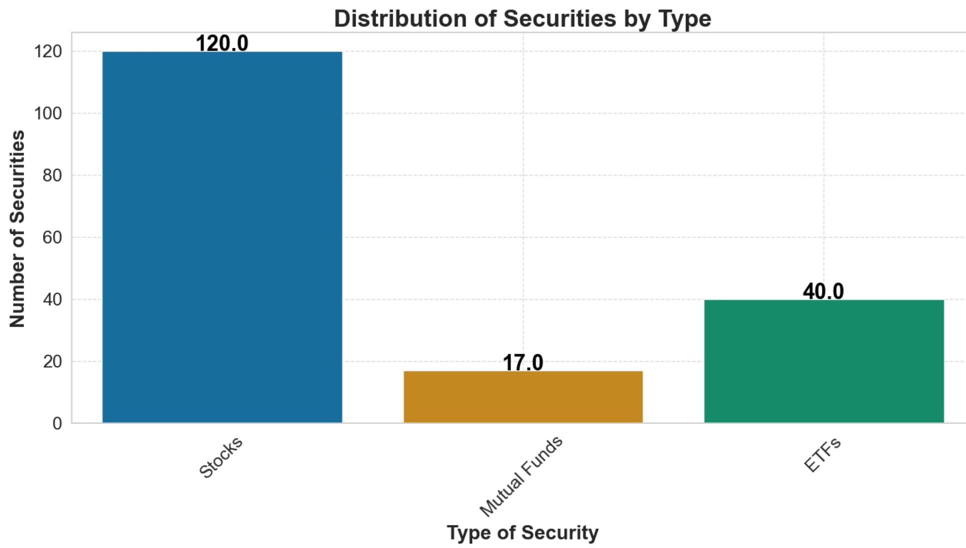

Figure 4: Distribution of Securities by Type

6.1.2 Cumulative Returns by Security Type

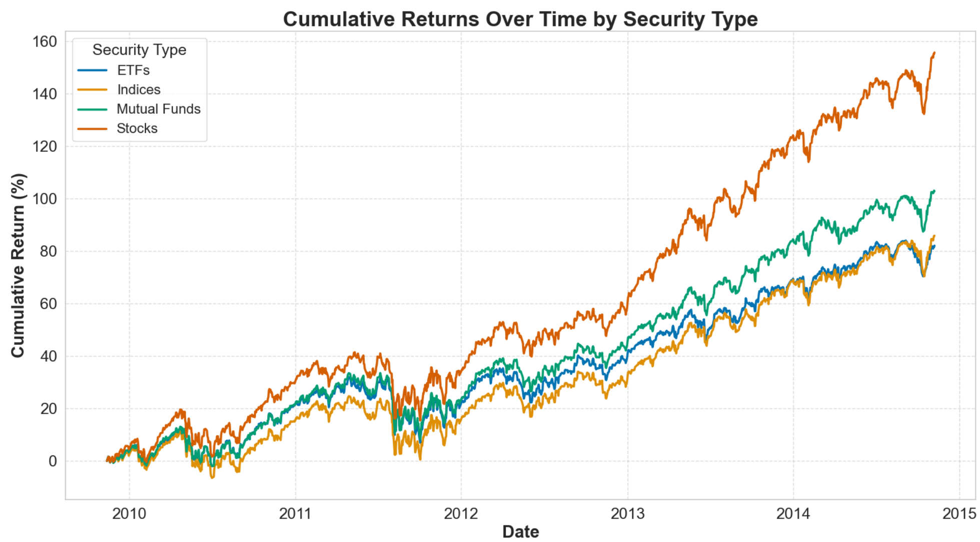

Figure 5: Cumulative Returns Over Time by Security Type

(a) T
he majority of the securities analyzed are stocks, followed by ETFs and mutual funds. Stocks exhibit the highest cumulative return over time, indicating a higher potential for long-term growth compared to ETFs and mutual funds.

6.2 Optimal Portfolios

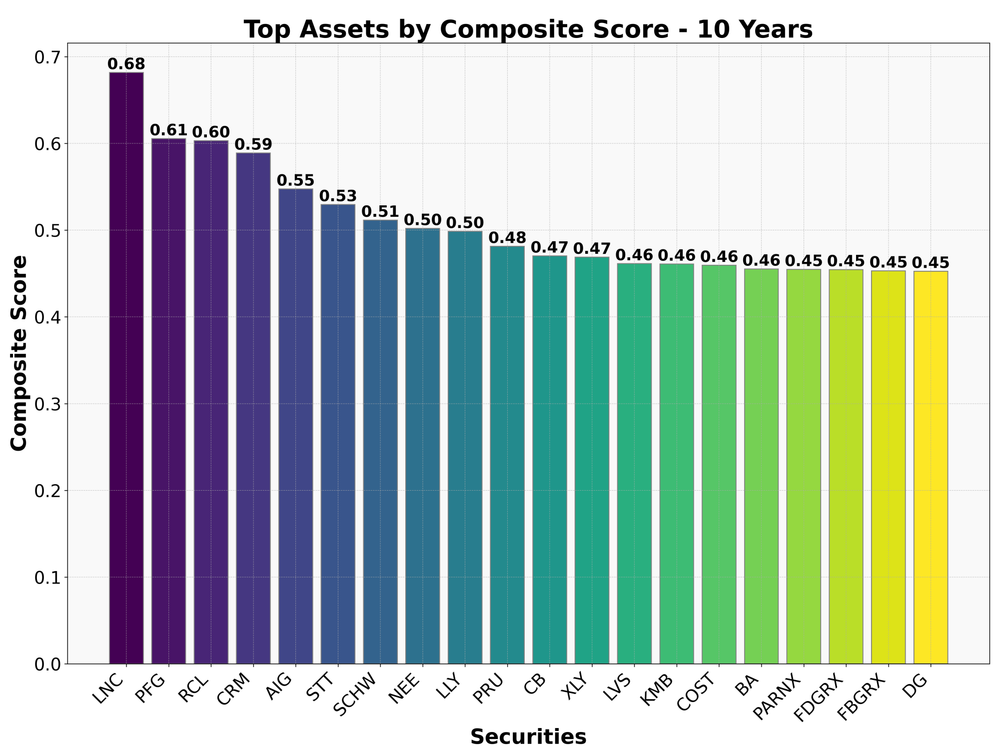

Figure 6: Top Assets by Composite Score (10 Years)

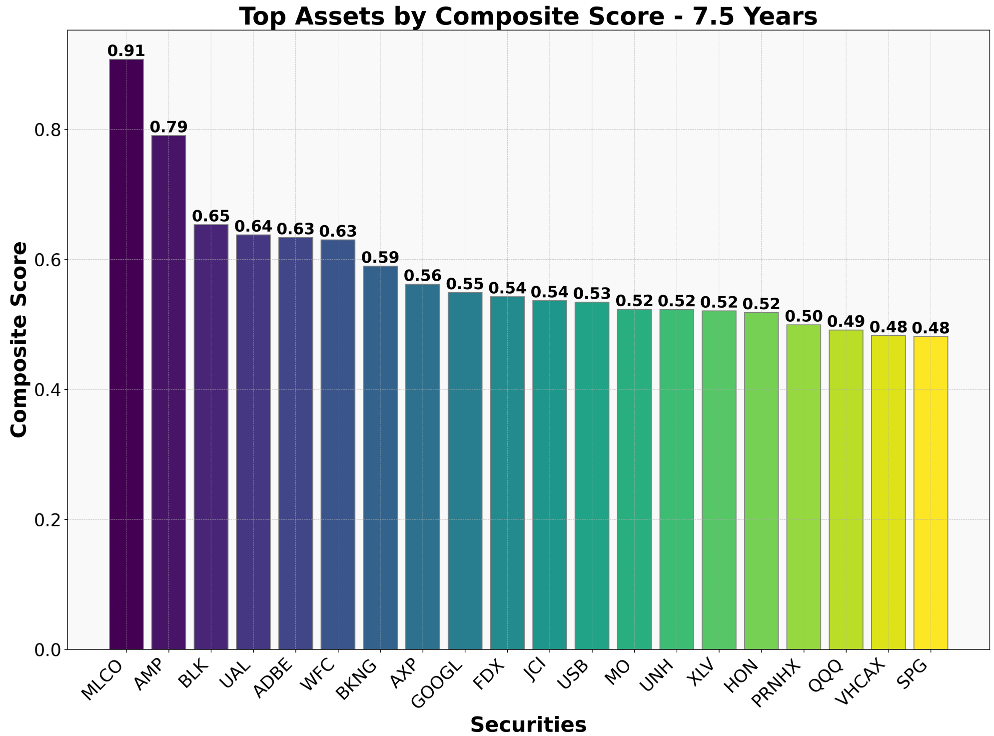

Figure 7: Top Assets by Composite Score (7.5 Years)

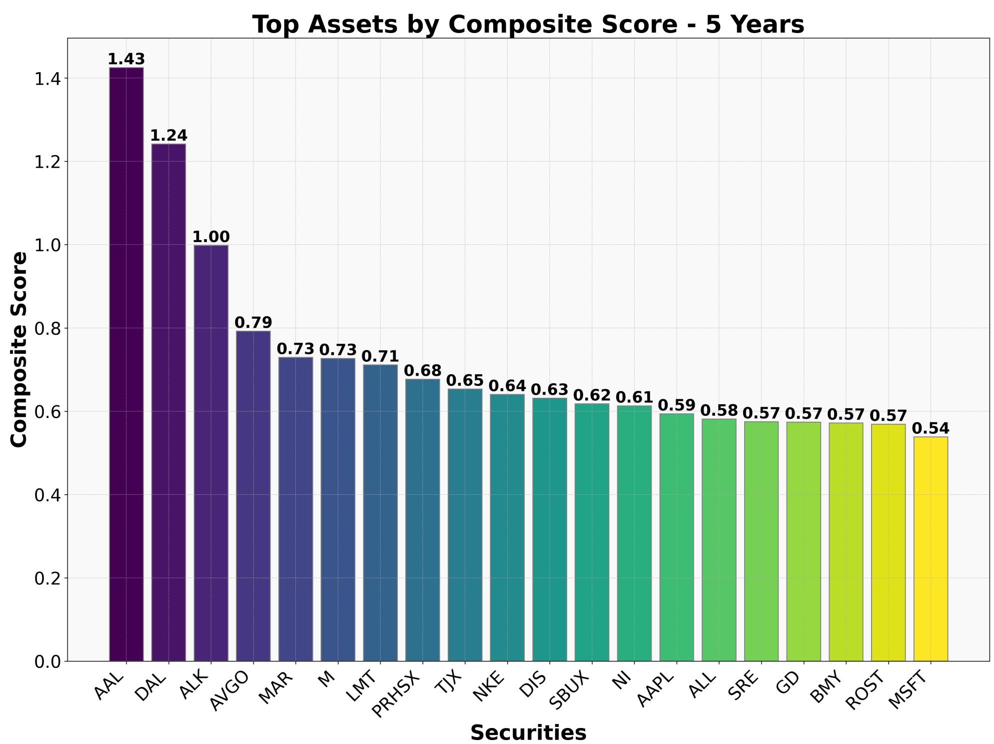

Figure 8: Top Assets by Composite Score (5 Years)

6.2.1 Optimal Portfolios Composition

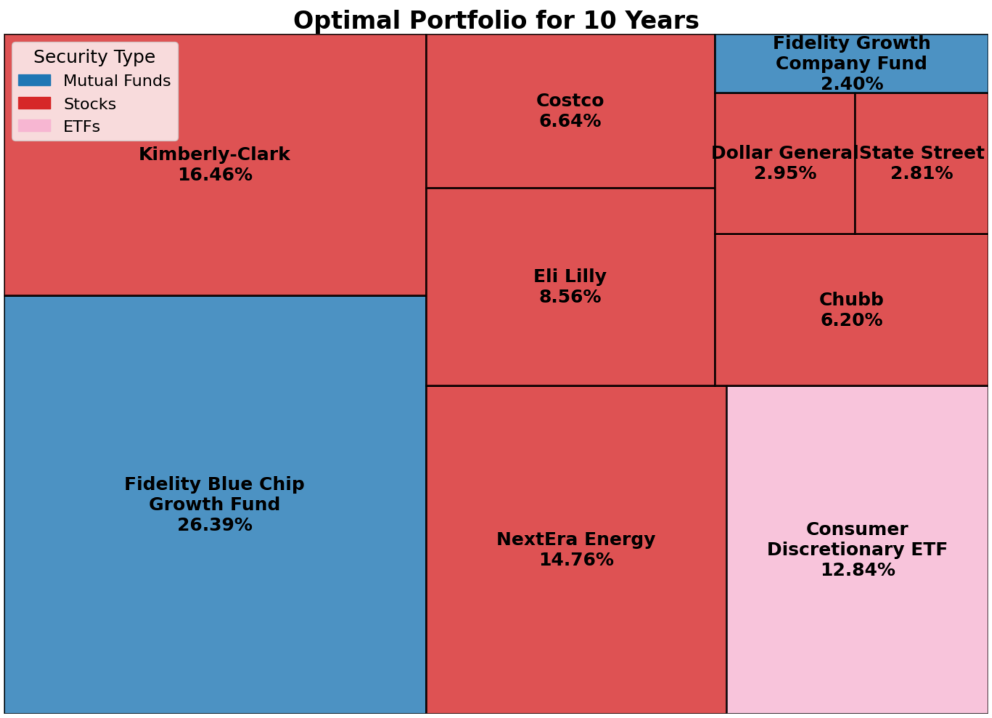

Figure 9: Optimal Portfolio for 10 Years

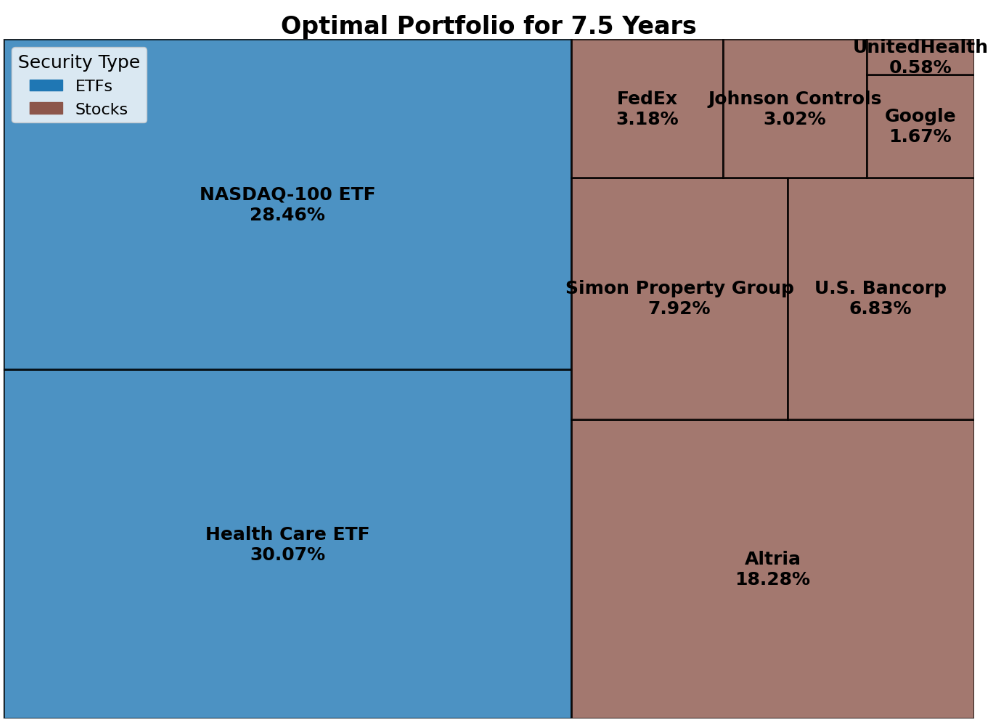

Figure 10: Optimal Portfolio for 7.5 Years

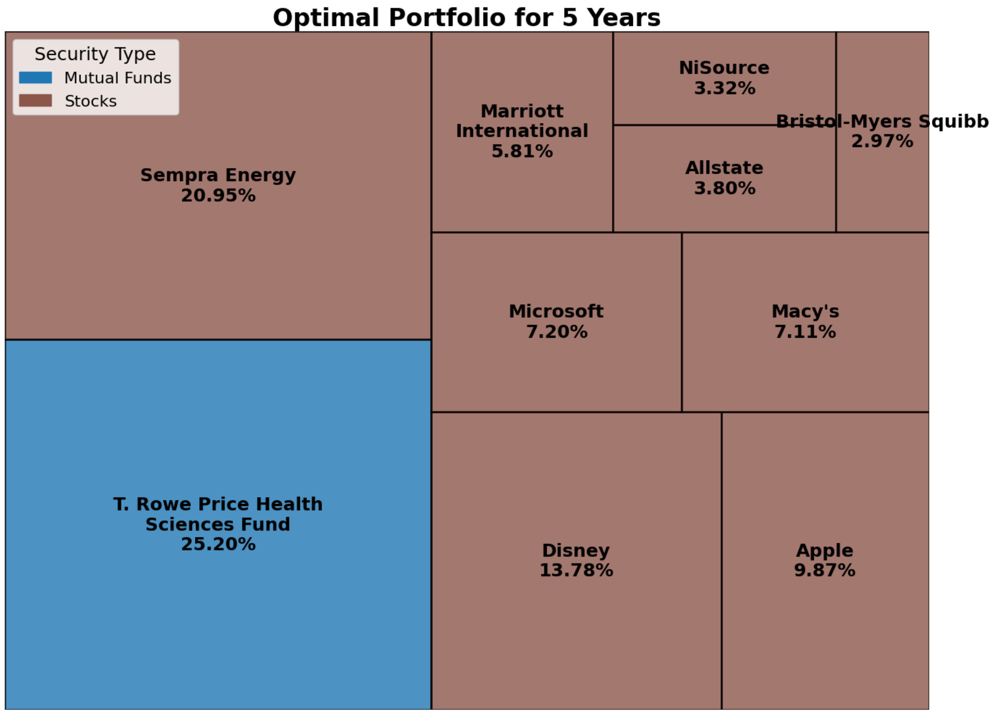

Figure 11: Optimal Portfolio for 5 Years

6.3 Comparison with Hindsight Data

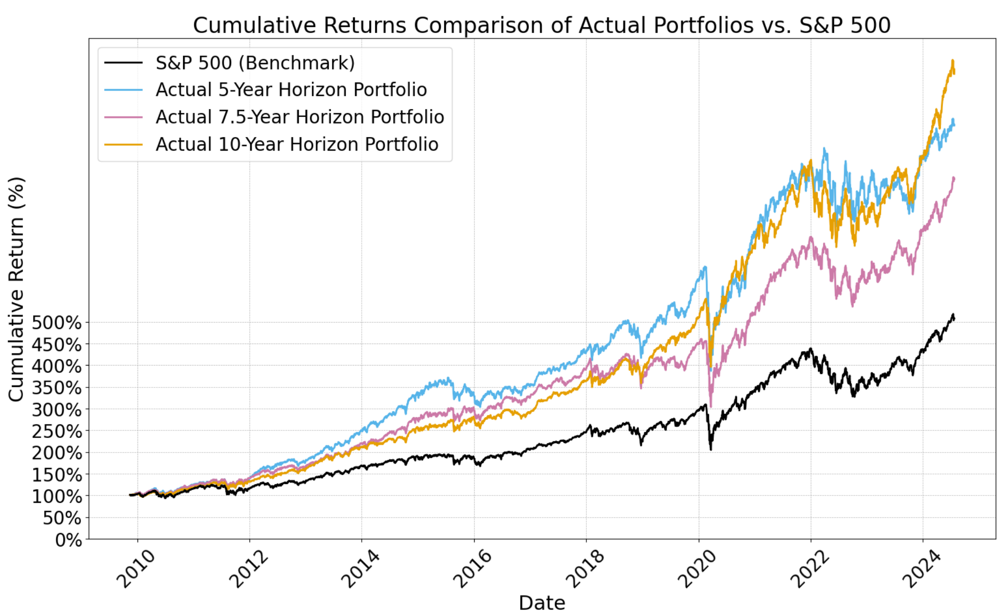

Figure 12: Cumulative Returns Comparison of Actual Portfolios vs. S&P 500

6.3.1 Cumulative Returns Summary

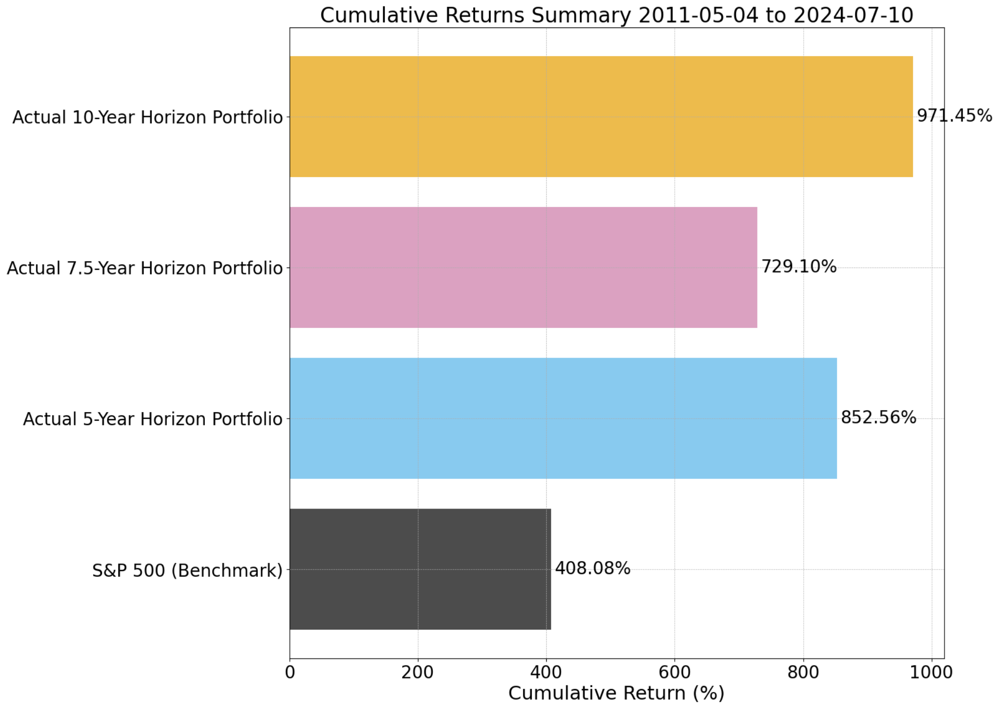

Figure 13: Cumulative Returns Summary from 2011-05-04 to 2024-07-10

6.4 Monte Carlo Simulation for Future Forecasting

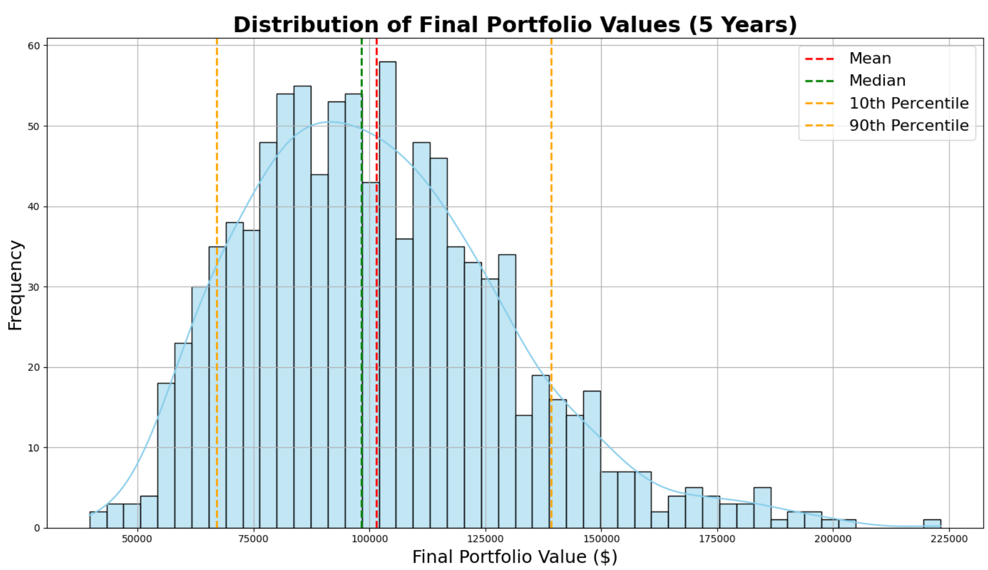

Figure 14: Distribution of Final Portfolio Values (5 Years)

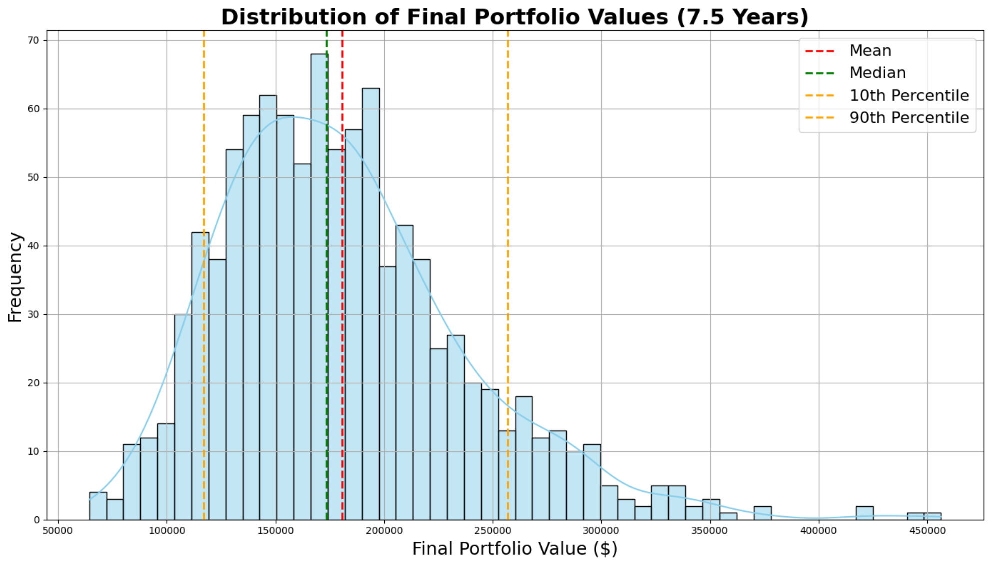

Figure 15: Distribution of Final Portfolio Values (7.5 Years)

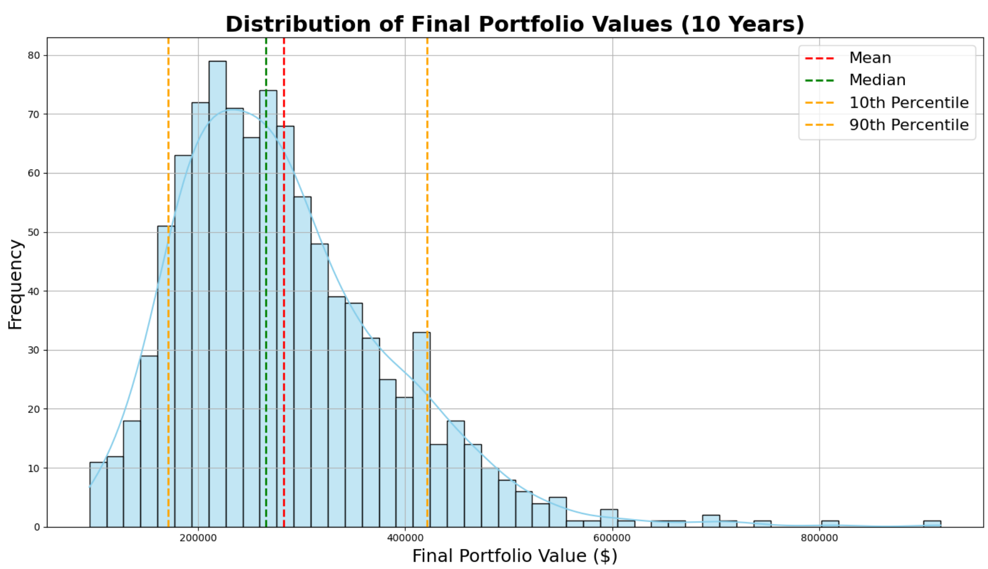

Figure 16: Distribution of Final Portfolio Values (10 Years)

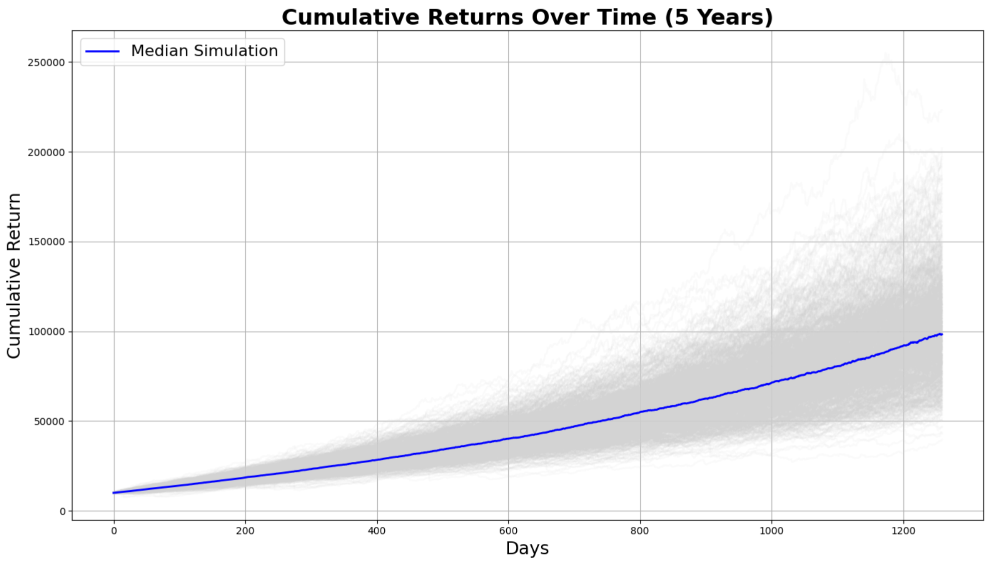

Figure 17: Cumulative Returns Over Time (5 Years)

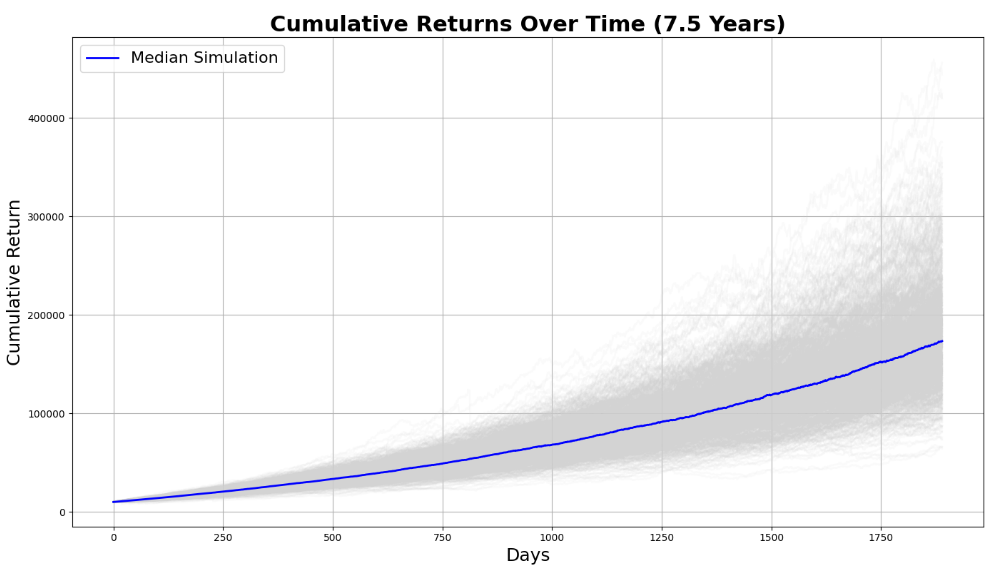

Figure 18: Cumulative Returns Over Time (7.5 Years)

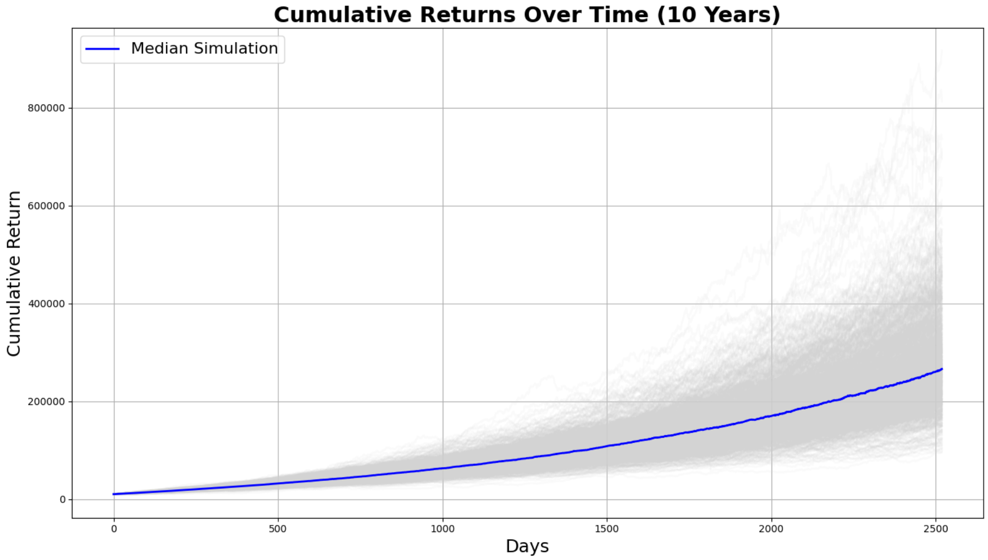

Figure 19: Cumulative Returns Over Time (10 Years)

6.4.1 Cumulative Returns Over Time

6.5 Monte Carlo Forecast Summary

Statistic	Mean Final Portfolio Value ($)	Median Final Portfolio Value ($)	10th Percentile Final Portfolio Value ($)	90th Percentile Final Portfolio Value ($)	Total Percentage Yield (%)	Annual Percentage Yield (%)
10-Year Horizon	283094.44	265864.65	171232.59	420935.51	2730.94	39.70
7.5-Year Horizon	180745.35	173433.88	116972.91	256760.32	1707.45	47.10
5-Year Horizon	101559.42	98327.18	67062.20	139233.79	915.59	58.98

Provides a detailed summary of the statistical outcomes for each investment horizon. The mean and median final portfolio values give insights into the expected performance, while the 10th and 90th percentile values offer a perspective on the range of potential outcomes. The total and annual percentage yields reflect the overall return potential of the portfolios over their respective horizons.

7 Conclusions

This research provides a comprehensive analysis of optimal investment strategies for first-time homebuyers aiming to accumulate down payments over 5, 7.5, and 10-year horizons. By leveraging Modern Portfolio Theory, the Capital Asset Pricing Model, and Monte Carlo simulations, the study offers actionable insights into constructing age-specific investment portfolios.

The results indicate that tailored investment strategies significantly enhance the ability to save for a down payment, reducing the time required to reach homeownership goals. The findings also highlight the importance of considering risk-adjusted returns and diversification in portfolio construction.

Future research could extend this analysis to consider other demographic factors such as income levels and regional housing market conditions. Additionally, the integration of alternative investment vehicles such as real estate investment trusts (REITs) and cryptocurrencies could provide further insights into optimizing investment strategies for down payments.

Acknowledgments

I would like to express my gratitude to my professor, Dr. Paarsch, for his invaluable mentorship and guidance throughout my graduate studies. I am also deeply thankful to my family for their unwavering support and encouragement during my academic journey at the University of Central Florida.

A Data Appendix

A.1 Summary of Data Sources

The financial data utilized in this study were obtained from Yahoo Finance, which provides comprehensive information on a range of securities including stocks, mutual funds, and ETFs (Yahoo Finance, nd).

A.2 Variable Definitions

The key variables used in the analysis are defined as follows:

- **Open**: The price at the beginning of the trading day.
- **High**: The highest price during the trading day.
- **Low**: The lowest price during the trading day.
- **Close**: The price at the end of the trading day.
- **Adj Close**: The closing price adjusted for dividends, stock splits, etc.
- **Volume**: The number of shares traded during the trading day.
- **Type**: The type of security (e.g., stock, ETF).
- **Ticker**: The unique symbol assigned to each security for trading purposes.
- **Beta**: A measure of a security's volatility in relation to the overall market. A beta greater than 1 indicates that the security is more volatile than the market, while a beta less than 1 indicates that it is less volatile (Sharpe, 1966).
- **CAPM Predicted Return**: The expected return of a security as predicted by the Capital Asset Pricing Model, which takes into account the risk-free rate, the security's beta, and the expected market return (Markowitz, 1952).
- **Sharpe Ratio**: A measure of risk-adjusted return, calculated by subtracting the risk-free rate from the security's return and dividing by the standard deviation of the security's return. A higher Sharpe Ratio indicates better risk-adjusted performance (Sharpe, 1966).
- **Actual Returns**: The realized return on a security over a specified period, including price appreciation and dividends (Fama, 1970).

- **Composite Score**: A combined score that integrates various metrics such as Beta, CAPM Predicted Return, Actual Returns, and Sharpe Ratio to evaluate the overall attractiveness of a security. The composite score is used to rank securities (Smith, 2020).

- **Rank**: The position of a security in the list based on its composite score, with a lower rank indicating a more attractive investment.

References

Boyle, P. P. (1977). Options: A monte carlo approach. *Journal of Financial Economics 4*(3), 323–338.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance 25*(2), 383–417.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance 19*(3), 425–442.

Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business 39*(1), 119–138.

Smith, J. (2020). Optimal investment strategies. *Journal of Finance 55*(4), 123–145.

Yahoo Finance (n.d.). Yahoo finance.