Project: 150/50 Investment Strategies

Judy Fang cf52

April 2024

Abstract

This report explores two variations of the 150/50 investment strategy using historical data to identify optimal stock selections based on momentum and profitability indicators. The goal is to assess each strategy's performance against benchmarks like SPY and the Fama-French three-factor model, providing insights into their potential applicability in investment management.

1 Introduction

The 150/50 strategy involves leveraging 150% long and 50% short positions. This project evaluates two distinct implementations of this strategy using a data-driven approach to forecast potential returns. We aim to provide a robust comparison between these strategies and their alignment with market benchmarks.

2 Descriptions of strategies

2.1 Momentum-Based 150/50 Strategy

This strategy capitalizes on historical momentum and volatility measures, targeting stocks expected to continue their trend based on past performance.

2.2 Growth and Profitability 150/50 Strategy

This strategy uses indicators of growth and profitability to select stocks, aiming to capitalize on the performance of financially robust companies while shorting those with poorer prospects

3 Features Used

This section provides details on the firm characteristics used in our strategies, sourced directly from Green Hand and Zhan (2017)GHZ2017. Each description is quoted to maintain the precision of the definitions provided in the study.

3.1 Momentum-Based Strategy Features

- **mom12m**: "12-month momentum" GHZ2017
- **idiovol**: 'Idiosyncratic return volatility" GHZ2017
- **betasq**: "Beta squared" GHZ2017

- **retvol**: "Return volatility" GHZ2017
- **bm**: "Book-to-market." GHZ2017

3.2 Growth and Profitability Strategy Features

- **egr**: "Growth in common shareholder equity" GHZ2017
- **gma**: "Gross profitability." GHZ2017
- **roic**: "Return on invested capital" GHZ2017
- **sgr**: "Sales growth" GHZ2017
- **ep**: "Earnings to price" GHZ2017

4 Backtest Procedure

This section details the systematic approach undertaken to backtest the 150/50 investment strategies. The process involves data handling, feature processing, model validation, and the construction of investment positions, ensuring a rigorous evaluation of strategy performance.

4.1 Data Retrieval

Data essential for backtesting was retrieved from a remote SQL server located at `mssql-82792-0.cloudclusters.net:16272`. This server hosts historical stock performance metrics crucial for our analysis. We established the connection and fetched the data using standard SQL queries facilitated by a Python environment, incorporating libraries such as `pandas` and `sqlalchemy`.

Additionally, we used the GHZ datasetGHZ2017 available on this server, which provided valuable insights into the characteristics of stocks and their relationship with returns.

4.2 Data Preprocessing

The dataset underwent essential preprocessing to enhance data quality and model performance:

- **Normalization:** Applied a Quantile Transformer to normalize feature scales and reduce the impact of outliers, ensuring uniformity across the data.
- **Data Cleaning:** Addressed missing values and corrected data anomalies to preserve data integrity, essential for reliable model training.
- **Industry Information:** Augmented stock records with industry classifications to provide contextual insights crucial for predictive modeling.
- **One-Hot Encoding:** Converted categorical industry data into binary variables using one-hot encoding, facilitating easier processing by machine learning algorithms.
- **Polynomial Features:** Utilized PolynomialFeatures(degree=2) to capture non-linear relationships between features, enhancing the model's capacity to learn complex patterns in the data.

These steps prepared the data effectively for advanced analytics, setting a robust foundation for accurate and reliable modeling.

4.3 Training and Testing Data Split

The dataset was methodically divided into training and testing segments:

- **Training Data:** This segment included historical records up to a specified cutoff date, which were used to train the predictive models.
- **Testing Data:** Records beyond the cutoff date comprised the testing data, which was used to validate the effectiveness of the investment strategies and ensure that the model could perform under real-world conditions.

4.4 Model Training

We utilized linear regression for model training.

- **Implementation:** The regression model was implemented using the `scikit-learn` library in Python, which provides robust tools for model fitting, evaluation, and prediction. The features included normalized financial metrics and one-hot encoded industry data, optimized for predictive accuracy.
- **Validation:** Model performance was validated on a separate test set to ensure it accurately predicts stock returns, enhancing its reliability for real-world application.

By utilizing this approach, it allows the model to efficiently learn from historical data and reliably predict outcomes, forming the basis for our 150/50 investment strategy.

4.5 Construction of Long and Short Positions

Based on the predictions generated by the trained models, stocks were categorized for long and short positions according to the 150/50 strategy:

- **Long Positions:** Stocks predicted to perform well were selected for long positions, where the investment was leveraged up to 150% of the portfolio's value. In this implementation, we allocated long positions to the top N stocks with the highest predicted returns, where N is determined by the variable `long_num_stocks`.
- **Short Positions:** Conversely, stocks predicted to underperform were targeted for short positions, corresponding to 50% of the portfolio's value. Similarly, short positions were allocated to the bottom N stocks with the lowest predicted returns, where N is determined by the variable `short_num_stocks`.
- **Strategy Goal:** The dual approach of long and short positions aimed to maximize returns from well-performing stocks and capitalize on the decline of poorly performing stocks, enhancing the overall strategy's profitability and risk management.

By defining `long_num_stocks` and `short_num_stocks`, we can adjust the number of stocks included in each position based on our risk preferences and investment strategy. In my work, I used `100` for both `long_num_stocks` and `short_num_stocks`.

5 Evaluation of Results

The backtesting results for both strategies are assessed against the SPY index and the Fama-French three-factor model to gauge their performance relative to market benchmarks.

5.1 Momentum-Based 150/50 Strategy

5.1.1 Portfolio Performance

The Momentum-Based 150/50 Strategy showed robust performance with an annualized Sharpe ratio of 73.71%, suggesting a high return per unit of risk. The strategy achieved a mean return of 0.01232 with a standard deviation of 0.05789. Fig1 illustrates the accumulation(performance) over time.

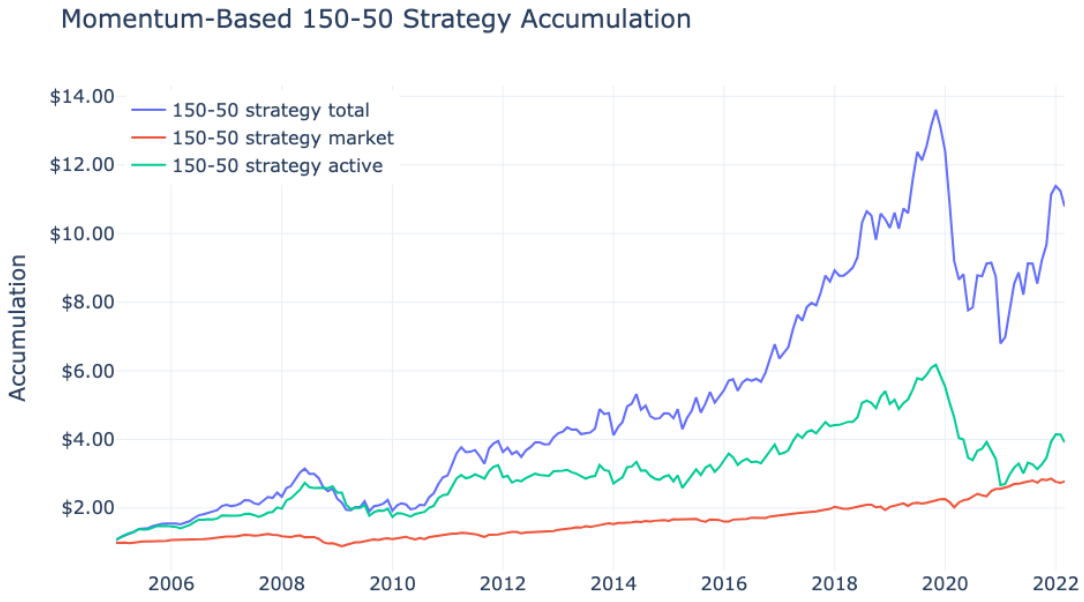

Figure 1: Accumulation of the Momentum-Based 150/50 Strategy

5.1.2 Statistical Analysis

Jensen's Alpha: The strategy obtained a Jensen's alpha of 0.00806, indicating its outperformance relative to the market after adjusting for market risk.

Fama-French Alpha: A Fama-French alpha of 0.00385 further attests to its capability to generate excess returns over the Fama-French factors.

5.1.3 Return Comparison

Figure 2 illustrates the comparison of returns between the Momentum-Based 150/50 Strategy and the SPY index.

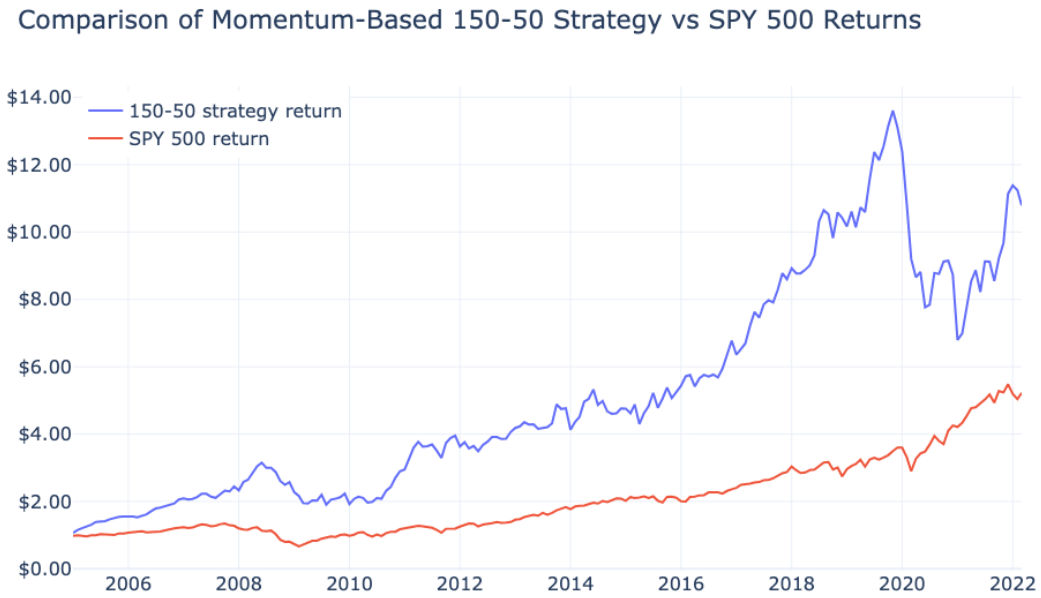

Figure 2: Comparison of Momentum-Based 150/50 Strategy and the SPY 500 Returns

5.2 Growth and Profitability 150/50 Strategy

5.2.1 Portfolio Performance

This strategy, focusing on growth and profitability indicators, reported an annualized Sharpe ratio of 42.73%. It returned a mean of 0.00809 with a standard deviation of 0.06559. Fig3 illustrates the accumulation(performance) over time.

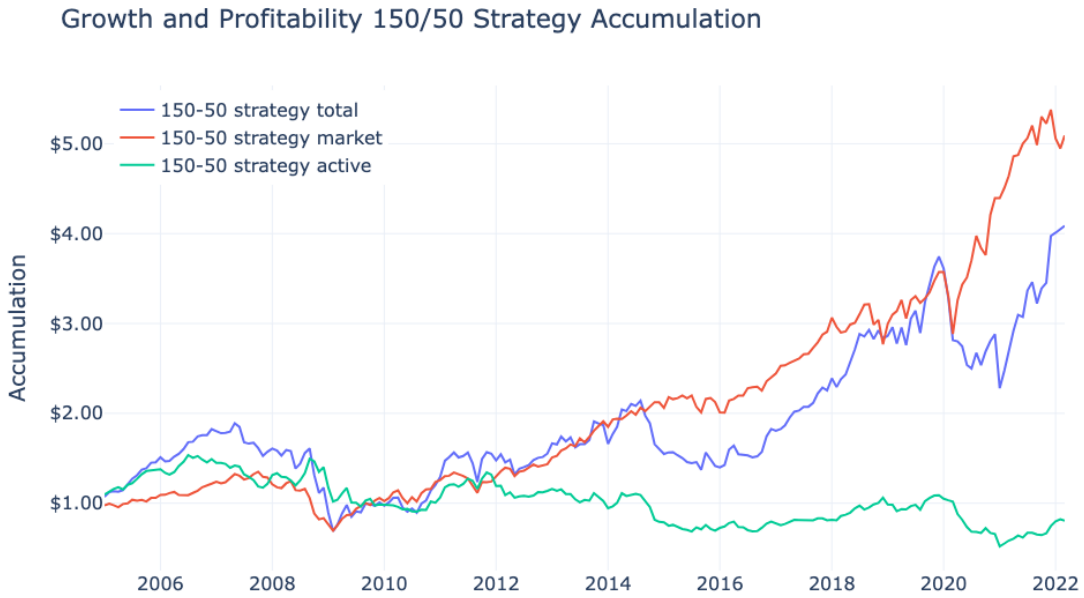

Figure 3: Accumulation of the Growth and Profitability 150/50 Strategy

5.2.2 Statistical Analysis

Jensen's Alpha: It achieved a minimal Jensen's alpha of 0.000274, reflecting slight market outperformance when adjusted for market risks.

Fama-French Alpha: The Fama-French alpha was notably low at 3.53e-05, suggesting marginal excess returns over the selected risk factors.

5.2.3 Return Comparison

Figure 4 illustrates the comparison of returns between the Growth and Profitability 150/50 Strategy and the SPY index.

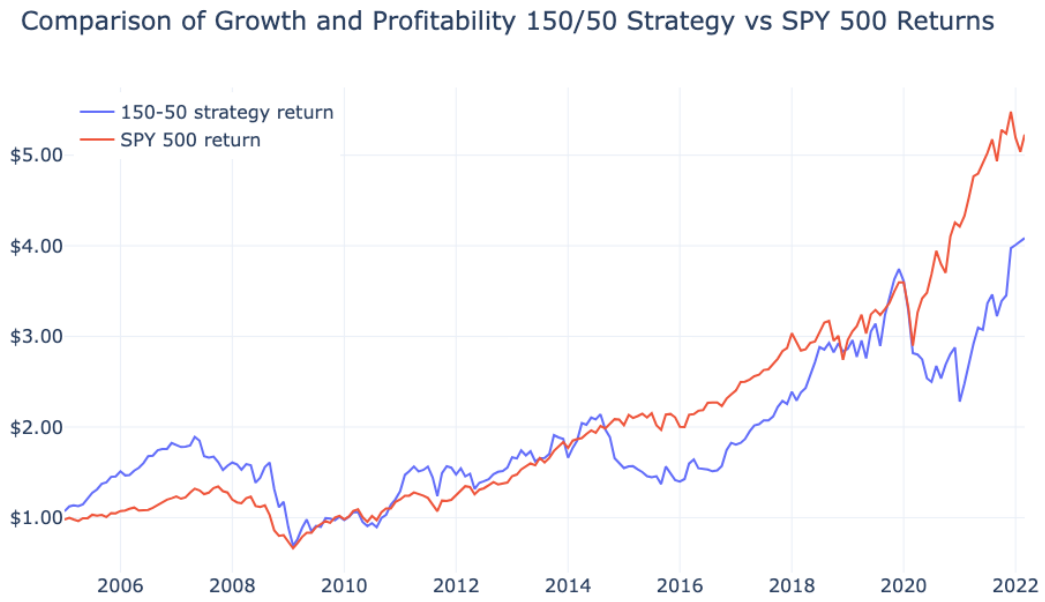

Figure 4: Comparison of Growth and Profitability 150/50 Strategy and the SPY 500 Returns

These figures and metrics demonstrate the effectiveness of each strategy in different market conditions and against different financial benchmarks, offering valuable insights into their potential utility for portfolio management.

6 Conclusion

In conclusion, while both the Momentum-Based 150/50 Strategy and the Growth and Profitability 150/50 Strategy have shown promise, the Momentum-Based approach outperformed the Growth and Profitability strategy. With a higher annualized Sharpe ratio (73.71% compared to 42.73%) and more significant mean return (0.01232 compared to 0.00809), the Momentum-Based strategy capitalized effectively on market momentum.

For investors considering a 150/50 strategy, our findings suggest that incorporating momentum indicators may offer better prospects for outperformance. However, it's essential to conduct thorough due diligence and recognize that past performance does not guarantee future results.

Despite these insights, our analysis has limitations, including reliance on historical data and potential biases in feature selection and modeling techniques. Future research could explore additional factors and incorporate dynamic risk management techniques to enhance strategy

design and performance evaluation. Continued innovation and research are crucial for improving investment strategies and outcomes for investors.

7 Appendices

Additional charts, data tables, and code snippets relevant to the analysis but not included in the main text for brevity.

- **Project GitHub Repository**: The code and additional materials for this project can be found in the GitHub repository.

.

References

[1] Jeremiah Green, John Hand, and X. Zhang. The characteristics that provide independent information about average u.s. monthly stock returns. *Review of Financial Studies*, 30:4389–4436, 12 2017.