Proceedings of the 4th International Conference on Business and Policy Studies
DOI: 10.54254/2754-1169/2025.28381

How Markov Chain and Fama-French Be Used in Practical Financial Alpha Strategy

Ruijie Mao[1*†], Xinxing Li[2†], Jiaxuan Li[3†], Yang Yang[4†], Siyuan Fan[5†]

[1]*Graduate School of Art and Science, New York University, New York, USA*
[2]*School of mathematical physics, Xi'an Jiaotong-Liverpool University, Suzhou, China*
[3]*Carroll School of Management, Boston College, Boston, USA*
[4]*School of Art and Science, New York University, New York, USA*
[5]*School of Art and Science, Queen's University, Kingston, UK*
**Corresponding Author. Email: Rm5372@nyu.edu*
[†]*These authors contributed equally to this work and should be considered as co-first authors.*

Abstract. According to Tulchinsky, Igor, et al's Finding Alphas: A Quantitative Approach to Building Trading Strategies, the model that predicts the price of a financial instrument is called alpha. This paper presents two alpha strategies based on the Markov chain model. which predicts future stock price movements by analyzing the transition probabilities between different market states and the Fama-French model, in which market performance. Note that the general possibility of the Market trend, which is the alpha of the Markov Chain is by using variables: price changes, sentiment, and trading volume. This is a commonly used market prediction method; however, innovation can be made in the interpretation of this possibility. The strategy uses historical data from 2015 to 2023 to create market-neutral short- and long-term portfolios. and we make sure the monetary neutrality to give a general and consistent performance of our alpha strategy.

Keywords: Fama-French, Markov Chain, Alpha, Return

1. Introduction

1.1. Markov chain alpha

1.1.1. Explanation

A Markov chain is a mathematical model used to describe a stochastic process, using the current state to predict its future state. And the forecast should be as accurate by using sufficient historical process data. Markov chain-based alpha strategies are designed to capture market dynamics by modeling the relationships between different market states. Every market situation is a combination of price, sentiment, and trading volume. This strategy uses these states to calculate the transition probability (the possibility that the stock price will turn from increasing to decreasing or from decreasing to increasing), making it possible to predict future returns. This method is a probabilistic approach to market forecasting and portfolio construction.

DOI: 10.54254/2754-1169/2025.28381

Market State S

Each market state S is defined by the following variables:

- Price:
  - Decrease: $(-\infty, -0.01]$
  - Neutral: $(-0.01, 0.01]$
  - Increase: $(0.01, \infty)$
- Sentiment:
  - Negative: $(-\infty, -0.1]$
  - Neutral: $(-0.1, 0.1]$
  - Positive: $(0.1, \infty)$
- Trading Volume:
  - Low: [0, 33%quantile]
  - Medium: [33%quantile, 67%quantile]
  - High: [67%quantile, ∞]

1.1.2. Economic intuition

Alpha represents a stock's abnormal return relative to market factors. The economic intuition behind this

Strategy is that if the possibility that the price of one stock is going to increase is high, we can make money by holding long-position on this stock; while if the possibility that the price of one stock is going to decrease is high, we can make money by holding the short-position on this stock. These possibilities are captured by the turning of the sentiment and trending of the market, and we capture the abnormal market state to make money.

1.2. Fama french alpha

1.2.1. Explanation

The second alpha introduces an alpha-based investment strategy using the Fama-French factors and regression analysis. The strategy aims to exploit market trends by ranking stocks based on alpha estimates and constructing a portfolio of long and short positions. The analysis spans both in-sample (2015-2020) and out-of-sample (2021-2024) periods, with performance metrics evaluated in terms of cumulative returns, volatility, and information ratio.

To simplify the research, we use the Fama-French 3 factors model instead of the Fama-French 5 facots model proposed by Eugene F. Fama's "A five-factor asset pricing model." [1-2] And additional reason we use the three factor model is that we wish the model failing to capture the market instantly change.

1.2.2. Economic intuition

Financial markets are inherently probabilistic, where transitions between different states are driven by various factors such as price fluctuations, investor sentiment, and market volume. It is the traditional probabilistic model, which is the foundation of rational expectation model. However, this model failed to capture the instant market change and fundamental change in the market state like the reform of the company and so on. We can make money by capturing the likelihood the firms will reacts opposite from the market.

DOI: 10.54254/2754-1169/2025.28381

1.2.3. Raw alpha

Yearly Rolling Alpha:

Raw alpha is initially estimated by applying the Fama-French regression model to historical data from the first year. Then calculate the correlation between predicted return and actual return in the second year. The correlation is the alpha.

1.2.4. Refined alpha

Monthly-Rolling Alpha and Short Strategy:

Refined alpha builds upon raw alpha by recalculating it on a monthly rolling basis. After estimating raw alpha yearly, the monthly-rolling alpha is used to update the portfolio composition more frequently, allowing the strategy to better capture short-term market movements.

2. Literature review

According to Gupta Aditya's "Stock market prediction using Hidden Markov models," Markov Chains have proven to be an effective tool for portfolio optimization and stock market prediction in a variety of financial models, including Hidden Markov Models (HMM) (Gupta, 1) [3]. Markov Chain models have been widely used in financial research to predict market behaviors, but their application to alpha generation remains novel. Prior studies have focused on modeling economic cycles, market regimes, and state transitions. This research extends the existing literature by applying Markov transition matrices to portfolio management and alpha generation.

According to Davou Nyap Choji's "Markov Chain Model Application on Share Price Movement in Stock Market," I. U Amadi's "Stochastic analysis of stock price changes as markov chain in finite states," and Jui-Chieh Huang's "Applying a Markov chain for the stock pricing of a novel forecasting model," they specifically emphasize the way people can use the Markov Chain Model to predict the stock price, which showing that people can use matrix to capture the markets performance, attitudes, and price range [4-6].And this matrix can be used to calculate the possibility of the increasing or decreasing od the stock price. This is the foundation of the markov chain alpha in this article.

The Fama-French three-factor model expands on the CAPM by including size and value factors, in addition to market risk, to better explain stock returns. The model has been extensively studied in finance literature and is widely accepted for its ability to account for variations in stock returns across different market environments. This research leverages the Fama-French factors to estimate stock alphas and construct a long-short portfolio.

According to Ralitsa Petkova's "Do the Fama-French Factors Proxy for Innovations in Predictive Variables?" "when loadings on the innovations in the predictive variables are present in the model, loadings on HML and SMB lose their explanatory power for the cross section of returns" (Ralitsa, 1) [7]. It shows that while facing some unexpected shock or some progress which cannot be predicted from period before, the effectiveness of Fama-French model will be reduced. Therefore, the stocks quite fit with the prediction got from Fama-French might not be the stocks which follows the market trend because of the failing of fitting the unexpected change due to market of Fama-French prediction. According to Eugene F. Fama's "Common Risk Factors in the Returns on Stocks and Bonds," he also shows that the Fama French method is built on the huge assumption that there is no intermediate change during the model, and it generally has little predictive force [8].Now it is a good way to think of the opposite, if the stocks follow the opposite way Fama-French did, will there be some differences? This means that if we invest on the stocks which has lower level of alpha, we should have some interesting finding on returns. And according to Robert B. Durand's "Fear and the Fama-French Factors," it also talks about how consumer's fear to the unexpected change might influence the effective ness of fama-french model [9].

According to Philip Gharghori's "Are the Fama-French Factors Proxying Default Risk?" they also talk about similar problem [10]. Because the Fama-French model is a model generated from strong assumption of national expectation, which assumed all the risk of the market change does not exist at beginning. This shows how the alpha can work because trading the stock operating opposite the Fama-French model predicts may help capture the market sudden change.

3. Specification

3.1. Data

3.1.1. Universe

The universe of this study consists of 300 companies with the top 300 average trading volume. The dataset includes daily price data, trading volume, and sentiment from 2013 to 2024.

3.1.2. Data sources

• WRDS (Wharton Research Data Services) Daily Stock price, Volume, Adjusted Return, and Sentiment Data.
• Kenneth R. French - Data Library Daily Fama-French Data

3.1.3. Dataset

The data is obtained from reliable financial sources, including:
• Price: Provided by major financial platforms.
• Sentiment: Obtained from social media and news analysis platforms.
• Volume: Sources from stock exchanges and financial information providers.
• Fama-French Factors: Daily Market Excess Return (Mkt RF), Size Factor (SMB), Value Factor (HML) and Risk Free Rate (RF).
• Stock Returns: Daily stock returns within a specified universe.

3.1.4. Date range

• In-Sample Data: January 2, 2013, to December 31, 2020.
• Out-of-Sample Data: January 1, 2021, to December 31, 2024.

3.2. Strategy detail: markov chains Alpha

3.2.1. Alpha calculation: markov chains

The predicted return for future state S_{+1} is calculated using the transition probabilities Pij between states and the historical return data $f(S_{+1})$. The formula for expected return is as follows:

$$\hat{r}t+1=\sum_{j=0}^{N} P_{ij}\cdot f(S_{t+1}^{j}) \tag{1}$$

DOI: 10.54254/2754-1169/2025.28381

Where

• Pij represents the probability of transitioning from state St to state S+1 .

• f(S+1) is the historical average return for state S+1 .

Fama French Alpha

3.2.2. Raw data

Alpha is calculated using the Fama-French three-factor regression model:

$$RETi(t) = \alpha i + \beta i, MktRF \times MktRF(t) + \beta i, SMB \times SMB(t) + \beta i, HML \times HML(t) + \epsilon i(t) \tag{2}$$

Where:

• RETi (t) is the return of stock i at time t,

• αi is the stock's alpha,

• βi,Mkt RF ,βi,SMB ,βi,HML are the factor loadings for the market, size, and value factors,

• ϵi (t) is the error term.

The formula is:

$$Raw\ Alphai \ = Correlation(RETi\ (y),\ Predicted\ RETi\ (y)) \tag{3}$$

Y represents year

3.2.3. Refined data

The formula is:

$$Raw\ Alphai \ = Correlation(RETi\ (m),\ Predicted\ RETi\ (m)) \tag{4}$$

3.2.4. Trade execution costs

Transaction costs, including bid-ask spreads and slippage, are taken into account to estimate the net profitability of the strategy. These execution costs are crucial for evaluating the practical viability of the strategy.

4. Implementation

4.1. Portfolio construction: markov chains

The portfolio is constructed annually by selecting the top 10 stocks with the highest predicted returns for long positions, and the bottom 10 stocks for short positions. An equal-weighted approach is applied, and the portfolio is rebalanced each year. You divide your total fund of $1 million equally into $500,000 for long positions and $500,000 for short positions, and each position is allocated by equal weight.

4.1.1. Strategy detail

The calculate long short portfolio function constructs the portfolio:

DOI: 10.54254/2754-1169/2025.28381

• Long positions: The top N assets with the highest predicted returns are selected for long positions, assuming they will outperform in the future.

• Short positions: The bottom N assets with the lowest predicted returns are selected for short positions, assuming they will underperform in the future.

• Equal weighting: Both long and short positions are equally weighted in the portfolio. The total return for the portfolio is the weighted sum of long and short returns:

$$R_{porfolio} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{R}_{\mathrm{long,,i}} - \frac{1}{N}\sum_{i=1}^{N} \mathrm{R}_{\mathrm{short,i}} \quad (5)$$

where Rlong,i is the return of the i-th long asset, and Rshort,i is the return of the i-th short asset.

4.1.2. Performance metrics

Key performance metrics are as follows:

Table 1. In-sample vs. out-of-sample performance metrics

Metric	In-Sample	Out-of-Sample
Annualized Return	6.56%	5.20%
Volatility	6.82%	7.17%
Maximum Drawdown	-36.88%	-40.12%
Sharpe Ratio	0.96	0.72

4.1.3. In-sample results

The Markov alpha strategy performed well during the in-sample period, delivering an annualized return of 6.56% with a Sharpe ratio of 0.96. The maximum drawdown was controlled at -36.88%.

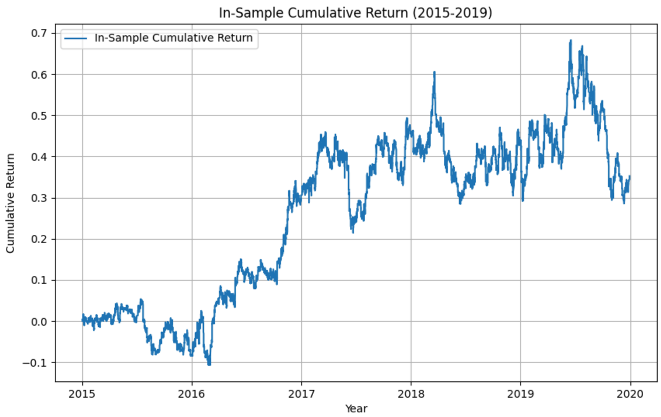

Figure 1. In-sample cumulative PnL results

4.2. Portfolio construction: Fama-French

The portfolio is constructed based on the ranked alpha estimates. Every month, the top 30 stocks are selected for long positions, and the bottom 30 stocks for short positions. Both the long and short portfolios are limited to 1, 200, 000andareequallyweighted.

DOI: 10.54254/2754-1169/2025.28381

4.2.1. Portfolio results

Below are the cumulative return graphs for in-sample periods.

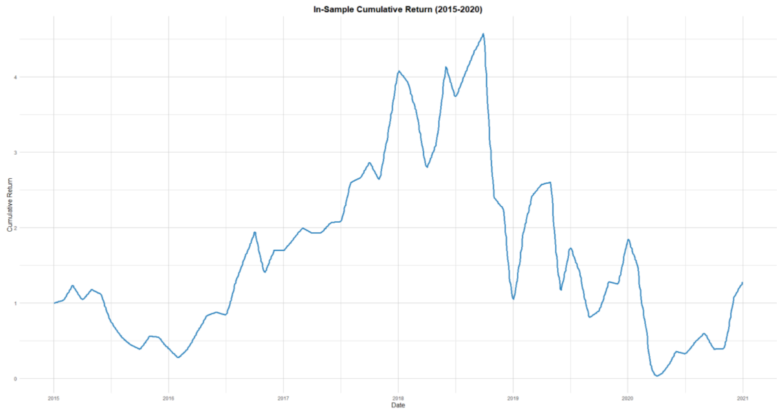

Figure 2. In-sample cumulative return (2015-2020)

5. Refinement

5.1. Markov chains based Alpha refinement

5.1.1. Dynamic discretization

In the discretize data function, the way price changes are discretized was modified. Instead of using fixed bins, the 33rd and 67th percentiles are now dynamically calculated for the price change to create the price category. This makes the categorization adaptive to the distribution of the data in each time period.

5.1.2. Smoothing in transition matrix calculation

In the calculate transition matrix function, a smoothing factor (Laplace smoothing) was added to ac- count for cases where transitions between certain states might not have been observed in the training data. This prevents zero probabilities in the transition matrix and makes it more robust.

5.1.3. Position sizing based on risk control

A new function, adjust position size, was introduced to dynamically adjust the position size based on volatility, maximum drawdown, and a stop-loss mechanism. This allows for better risk control when predict- ing returns. The function considers high volatility and large drawdowns as signals to reduce position sizes, with the possibility of fully exiting the position if the drawdown reaches a stop-loss threshold.

5.1.4. One-year window for training data

For both Markov Chain Alpha and Fama-French Alpha, the training time frame will change from two years to one year. This focuses training on the latest information. This can help predict market behavior based on short-term trends, and this is a better interpretation of the latest information. This will increase the accuracy of the model.

DOI: 10.54254/2754-1169/2025.28381

5.2. Alpha combo

5.2.1. Combining two Alpha strategies

We use both FFAlpha (Fama-French Alpha) and MCalpha (Markov Chain Alpha). These two alphas is what we use in previous sections with both of them standardized by subtracting the mean and dividing by the standard deviation:

$$Std\alpha = \frac{\alpha-\mu}{\sigma} \tag{6}$$

where μ is the mean of the α and σ is the standard deviation of α.

The two alphas are combined using average to create a ranking factor for portfolio construction:

$$CombinedAlpha = \frac{FFAlpha+MCAlpha}{3} \tag{7}$$

This combined alpha is used to rank stocks daily. The top 30 ranked stocks were allocated to the long portfolio, and the bottom 30 ranked stocks were allocated to the short portfolio.

5.2.2. Portfolio construction

For both in-sample and out-sample periods, the portfolios are constructed like this:

- Long Portfolio: Top 30 stocks ranked by the combined alpha.
- Short Portfolio: Bottom 30 stocks ranked by the combined alpha.

$$LongReturn = \frac{PriceonDay3-PriceonDay1}{PriceonDay1} \tag{8}$$

$$ShortReturn = \frac{PriceonDay1-PriceonDay3}{PriceonDay3}$$

The overall portfolio return was calculated as:

$$\text{Portfolio Return} = \frac{\sum \text{Long Returns}}{30} - \frac{\sum \text{Short Returns}}{30} \tag{9}$$

5.2.3. In-sample (2015-2020) cumulative return

The cumulative returns for the sample period are shown in Figure 3, which reflects the performance of the combined alpha strategy from 2015 to 2020.

DOI: 10.54254/2754-1169/2025.28381

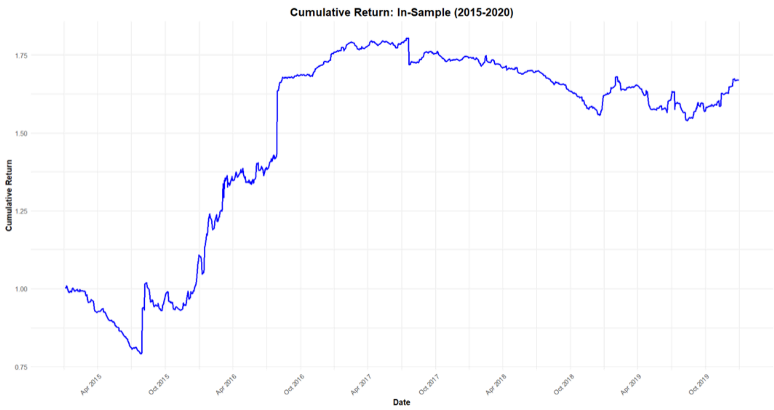

Figure 3. Cumulative return: in-sample (2015-2020)

5.2.4. Out-sample (2021-2024) cumulative return

The cumulative return for the out-sample period is shown in Figure 4. This reflects the performance of the combined alpha strategy from 2021 to 2024.

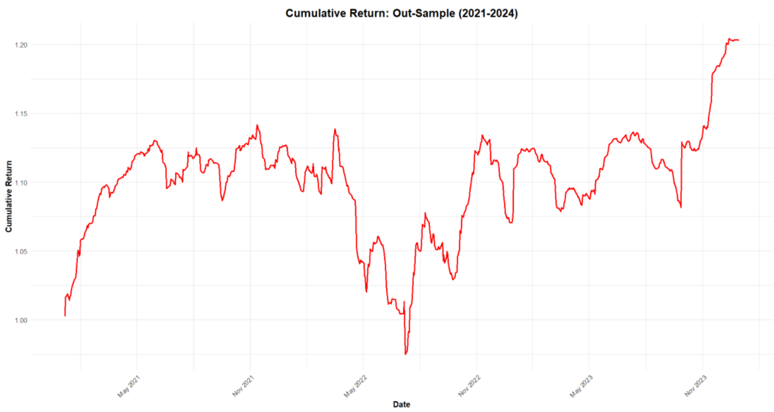

Figure 4. Cumulative return: out-sample (2021-2024)

5.2.5. Statistical summary

The following table provides a summary of the Combined Alpha Strategy's performance based on key financial indicators.

Table 2. Annual returns, volatility, and information ratios in the sample and out-of-sample periods

Period	Annualized Return	Volatility	Information Ratio
In-Sample (2015-2020)	0.2179694	0.161793	1.347206
Out-Sample (2021-2024)	0.3030476	0.074441	4.070938

The table shows that although the information ratio and annual returns are higher in the out-of-sample period, But the volatility has decreased significantly. This indicates a more stable performance after 2021.

6. Conclusion

Markov chain-based alpha strategies are a promising tool for predicting market returns and building profitable long and short portfolios. Both in-sample and out-of-sample testing showed consistent performance. with manageable risks and acceptable pass-through Future research could focus on optimizing state definitions to improve transition probability estimates using machine learning techniques.

Alpha strategy using Fama-French captures market trends effectively. Demonstrates strong performance in favorable markets But it is more sensitive to declines. The strategy's strength lies in its ability to generate returns in a bullish environment. Although there will be more losses during the downturn. Estimates may significantly outperform the market when conditions are favorable but underperform in negative markets.

6.1. Out-of-sample results

Out-of-sample performance shows consistency. It has an annualized return of 5.20% and a Sharpe ratio of 0.72, demonstrating the model's robustness to unseen data.

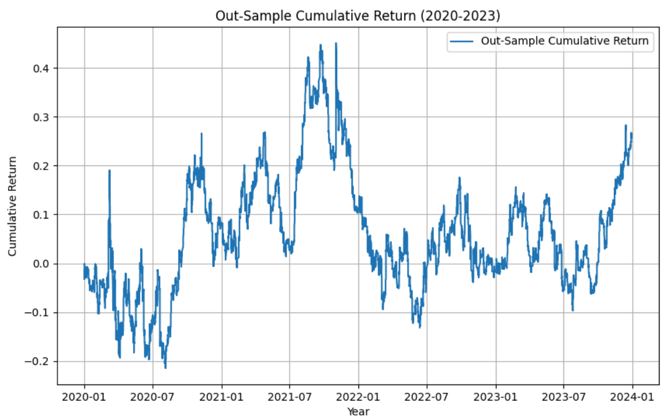

Figure 5. Out-of-sample cumulative PnL results

Table 3. Annualized performance metrics

Period	Annualized Return	Volatility	Information Ratio
In-Sample (2015-2020)	0.0417	0.3666	0.0189
Out-Sample (2021-2024)	1.9663	0.1883	0.3703

Below are the cumulative return graphs for out-sample periods.

DOI: 10.54254/2754-1169/2025.28381

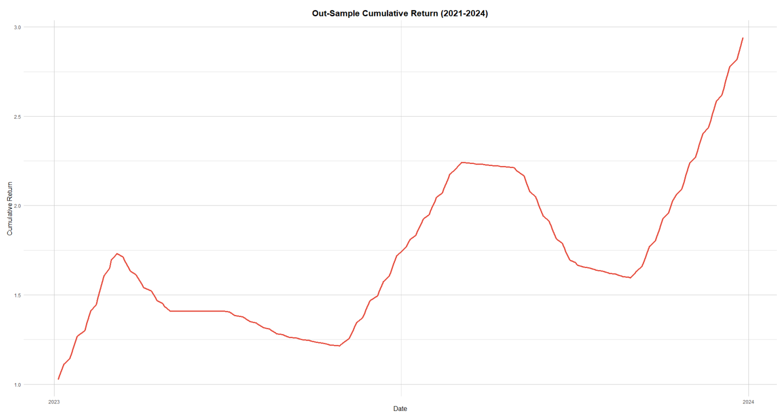

Figure 6. Out-sample cumulative return (2021-2024)

We can see that the out-of-sample data also show a consistent return for this combined alpha portfolio, showing a strong profitability of this model.

The innovation we make from traditional Markov Chain prediction is that we use the yearly rolling alpha to do the transformation matrix to predict the stock which has better chances to have turning point of price change. We did not simply use this possibility to build the portfolio, instead, we make a rank and invest both good possibility and bad possibility by dividing them into long and short group.

And the innovation for Fama-French Alpha is based on the predictive force of the Fama-French model. We capture the turning point and failing point of the national rationality theorem and finding the failure of the performance of the Fama-French. In this way we can see how people's rationality will impact the market change and how people follow the market trend will help make money and also capture the irrationality of how people do not follow as the efficient market do and do the opposite direction as Fama-French predicts. If people's action to one stock is opposite to national rationality, then it should be a rational behavior to react opposite to Fama-French predict either. It is also following this strong assumption.

6.2. Trading recommendations

6.2.1. Fama-French Alpha strategy

The Fama-French Alpha strategy is based on dynamic stock selection using market, size (SMB) and price (HML) risk factors, which are recalculated over a monthly time frame. The optimized alpha generated allows for accurate ranking of stocks. Business advice includes:

• Long Portfolio: Select the top 30 stocks with the highest positive complexity Fama French alpha. These stocks continue to perform outstandingly. It is not explained by traditional market factors. Monthly alpha adjustments ensure timely adjustments to market movements. Prevents alpha degradation.

• Short Portfolio: Select the last 30 stocks with the least negative complex alpha. These stocks are systematically undervalued by the model. This is because the market is inefficient. The strategy extracts returns through mean regression dynamics by targeting these underperformers.

Monthly portfolio rebalancing is essential to accommodate changing factors and improve risk-adjusted returns over time. Fama-French Alpha helps capture unique opportunities driven by structural asymmetries and behavioral bias.

6.2.2. Markov chain Alpha strategy

Markov chain alpha strategies capture market changes using stochastic modeling. The unique value of this strategy comes from its ability to predict state changes based on the likely outcome in a given market state. Business advice includes:

• Long Portfolio: Based on the highest transition probability, select the top 10 stocks that are expected to transition to favorable status. These stores are identified with a higher probability of transitioning from an underperforming state to a superior performing state. which reflects the potential for upward movement.

• Short Portfolio: Select the last 10 stocks that are expected to shift to an unfavorable state. These are the stocks that are most likely to experience changes in sentiment, prices, and volume declines. This makes it an ideal stock that is expected to suffer future losses.

Annual rebalancing allows the strategy to benefit from long-term market trends. At the same time, it reduces short-term volatility. Transaction costs and liquidity are considered factors in portfolio construction to maintain stability and net profitability.

6.2.3. Combined Alpha strategy

The combined strategy combines Fama-French and Markov Chain Alpha in a robust multifactor model. This approach leverages the strengths of two models: Fama-French's ability to systematically detect risk premiums; and Markov chains' predictions about changing market states. The recommendations of the joint strategy are:

• Long Portfolio: Select the top 30 stocks ranked by a combined alpha, which averages both refined Fama-French alpha and the Markov Chain’s transition-based predictions. This hybrid approach cap- tures both systematic factors and probabilistic state transitions, providing a well-rounded selection of stocks poised to outperform across both dimensions.

• Short Portfolio: Select the bottom 30 stocks ranked by the combined alpha. These stocks are expected to underperform from both systematic and state-transition perspectives, making them ideal candidates for short positions.

The hybrid alpha strategy provides a more complex risk adjustment approach by combining the market position dynamics of the Markov chain model with the systematic performance deviations captured by the Fama-French factors. Monthly rebalancing ensures that it adapts to both short-term and long-term market changes. By increasing the efficiency of the exchange between Returns and Volatility

Acknowledgement

Ruijie Mao, Xinxing Li, Jiaxuan Li, Yang Yang, and Siyuan Fan contributed equally to this work and should be considered co-first authors.

References

[1] Tulchinsky, Igor, et al. Finding Alphas: A Quantitative Approach to Building Trading Strategies. Wiley, 2015.
[2] Fama, Eugene F., and Kenneth R. French. A Five-Factor Asset Pricing Model. Journal of Financial Economics, vol. 116, no. 1, 2015, pp. 1-22. https: //doi.org/10.1016/j.jfineco.2014.10.010.
[3] Gupta, Aditya, and Bhuwan Dhingra. Stock market prediction using Hidden Markov models . 2012 Stu- dents Conference on Engineering and Systems, Mar. 2012, https: //doi.org/10.1109/sces.2012.6199099.
[4] Choji, Davou Nyap. Markov Chain Model Application on Share Price Movement in Stock Market.
[5] Amadi, I. U. Stochastic Analysis of Stock Price Changes as Markov Chain in Finite States.

DOI: 10.54254/2754-1169/2025.28381

[6] Huang, Jui-Chieh. Applying a Markov Chain for the Stock Pricing of a Novel Forecasting Model.
[7] Fama, Eugene F., and Kenneth R. French. Common Risk Factors in the Returns on Stocks and Bonds. Journal of Financial Economics, vol. 33, no. 1, 1993, pp. 3-56.
[8] Petkova, Ralitsa. Do the Fama-French Factors Proxy for Innovations in Predictive Variables? The Journal of Finance, vol. 61, no. 2, 2006, pp. 581–612. https: //doi.org/10.1111/j.1540-6261.2006.00846.x.
[9] Durand, Robert B., et al. Fear and the Fama-French Factors. Financial Review, vol. 45, no. 4, 2010, pp. 883–905.
[10] Gharghori, Philip, et al. Are the Fama-French Factors Proxying Default Risk? Australian Journal of Management, vol. 34, no. 1, 2009, pp. 45–71. https: //doi.org/10.1177/031289620903400104.