Exploring the Impact of Factor Timing Strategy: Final Project

Group 8 - Jinny Kim, Kondareddy Thanigundala, Raymond David, Hyoju Kang, Yukta Butala

Introduction

In recent years, the investment community has witnessed an intriguing debate surrounding the concept of factor timing. This discussion hinges on whether investors can strategically adjust their exposure to various investment factors—such as value, momentum, or size—based on their current valuations to enhance returns and manage risks. At the heart of this debate are two contrasting perspectives from leading figures in the field of quantitative finance. Rob Arnott ***supports*** factor timing, suggesting that factors, like stocks, fluctuate between overvaluation and undervaluation. By timing these fluctuations—buying low and selling high—investors can potentially enhance returns and reduce risks. On the other hand, Cliff Asness ***argues*** that factor timing is as **challenging** as traditional market timing, highlighting its difficulties and questioning the **reliability** of its supporting evidence. This project will investigate these perspectives, replicating some of the methodological aspects discussed that fuel this debate, aiming to clarify the potential and limitations of factor timing as an investment strategy.

Variable Selection & Exploratory Analysis (EDA)

The Fama-French factors enhance the traditional Capital Asset Pricing Model (CAPM) by incorporating multiple sources of risk that better explain the returns in a diversified portfolio. While CAPM relies solely on the market risk premium (the excess return of the market over the risk-free rate), the Fama-French model includes additional factors: size (SMB, Small Minus Big), value (HML, High Minus Low), profitability (RMW, Robust Minus Weak), and investment (CMA, Conservative Minus Aggressive). These factors address anomalies in market data that CAPM overlooks, such as the tendency for smaller-cap stocks to outperform larger caps. For stocks with high book-to-market ratios to outperform those with low ratios and for firms with conservative investment strategies to outperform those that invest aggressively. Analyzing these factors for overvaluation or undervaluation involves assessing their price relative to historical norms, akin to evaluating individual stocks. Investors might consider these valuations to strategically adjust their exposures to these factors, aiming to buy when undervalued (expecting higher returns) and sell when overvalued (to mitigate lower returns). This valuation-based approach suggests that the systematic risks associated with these factors also experience cycles of higher and lower expected returns, influenced by broader economic and market conditions. Here we are using the Fama/French 5 Factors dataset found in this link. Just like traditional market timing analysis, it is crucial to select significant predictors and develop refined models to ensure that we can make accurate predictions out of sample and develop strong signals to "time" the market, or factors.

The '***S&Pdata.xlsx***' provides us with data on the S&P 500 index, as well as dividend and risk-free rate information. This allows us to compute the price-to-dividend ratio or the market excess return as predictors. The '***PredictorData2019.csv***' includes several predictors found on Welch and Goyal's website.

In addition to these predictors, we have independently incorporated several macroeconomic variables, specifically, the '*10-Year Real Interest Rate*', '*10-Year Breakeven Inflation Rate*', '*Unemployment Rate*', '*Consumer Confidence Index*', and the '*CBOE Volatility Index (VIX)*'. Incorporating additional variables such as the VIX, Inflation, Consumer Confidence Index (CCI), and Unemployment Rate into factor timing analysis enhances the model's robustness by providing broader economic context. The VIX, often referred to as the “fear gauge”, reflects market risk and investors' sentiments, which can influence factor performance. Inflation rates impact the real returns of investments, making them a critical variable for predicting future market conditions. Similarly, the CCI serves as a barometer of consumer optimism towards economic conditions, which can affect consumer spending and, consequently, market performance. Lastly, the Unemployment Rate is a fundamental indicator of economic health, influencing consumer confidence and spending. By integrating these macroeconomic variables, the model not only captures specific market behaviors but also accounts for wider economic trends, potentially improving the accuracy and timing of factor-based investment strategies.

We initiate the analysis with individual simple linear regressions, assigning each factor in turn as the dependent variable. The primary aim is to swiftly identify the significant predictors for each factor. This process will provide preliminary insights into the historical correlations between factors and predictors and the proportion of variance that is elucidated (*Appendix 2)*. This initial phase revealed a particularly strong link between the market factor (Mkt-RF) and several macroeconomic variables, suggesting potential predictors for timing decisions. However, in-sample R-squared[1] for other factors are very minimal. In reality, common practice is to utilize as many significant predictors and to explain adjusted R-squared as well as possible. This involves carefully selecting predictors and determining what stays in our model. Due to the potentially time-consuming nature of the process, here, we are replicating the overall process, using our final 11 predictive variables, and conducting a **forward stepwise regression** graded using the Akaike Information Criterion (AIC) (*Appendix 1*). In the context of stepwise regression, employing AIC assists in iteratively adding or removing predictors based on their contribution to the model's overall efficiency and effectiveness. This method enhances the model's predictive accuracy while avoiding redundancy in variables, ultimately leading to a more streamlined and effective model. Our Stepwise Regression has provided additional insights into historical relationships by retaining only those predictors that add value to the model's predictive performance, while penalizing for excessive complexity (*Appendix 3)*. This stepwise approach serves to hone in on essential variables, ensuring our models are not overcomplicated and are likely to hold up when applied to fresh data.

[1] R-squared measures the proportion of variability in a dataset that is accounted for by a statistical model, indicating how well the model predicts the outcome.

However, *Appendix 4* confronted us with the models' limitations in predicting future performance. The negative out-of-sample R-squared values signal the complex nature of market dynamics that our models must navigate. For instance, while our initial regression model for Mkt-RF accounted for the most variance, the AIC-optimized stepwise model yielded a negative out-of-sample R-squared. Nevertheless, this does not conclusively affirm Asness's perspective. This serves as a reality check to steer us towards the application of more advanced machine learning algorithms, which promise to enhance predictive accuracy by discerning subtle patterns within the data. Here, the regression work lays the foundation, providing a benchmark and a selection of variables that are economically meaningful, which we aim to build upon with more sophisticated machine learning techniques. These next steps involve rigorous testing against the benchmark, refining models to not just fit historical data but to also capture the essence of market movements, improving our factor timing strategy.

Machine Learning Implementation

To advance beyond our initial linear regression models, we are set to evaluate three distinct machine learning algorithms. Our first method employs a **decision tree**, applying it to each factor separately with the same predictors identified in *Appendix 1*. For model validation, we employ an 80:20 data split, leveraging data from the outset of 2003 and projecting returns for the concluding three years (2016-2019). This starting point is chosen because it aligns with the inception of our predictors' dataset (since we merged multiplied predictor datasets). Furthermore, we operate under the premise that models trained on data stretching back over two decades may not effectively capture recent market dynamics. In *Appendix 5*, we compared the performance of decision trees with the results from stepwise regression outlined in *Appendix 4*. The decision tree model demonstrated an improvement for the market factor (Mkt-RF), with an out-of-sample R-squared of 0.1982, a stark contrast to the negative value of -0.7585 previously observed. This suggests that decision trees may better capture the predictive signals for market trends.A higher out-of-sample R-squared suggests the model reliably predicts future factor behaviors, essential for effective factor timing strategies.

However, the model's effectiveness varied across other factors, with the decision tree yielding lower R-squared values for SMB, HML, RMW, and CMA compared to their stepwise regression counterparts. Such outcomes indicate the model's limitations in explaining these factors and underscore the necessity to examine more sophisticated machine learning approaches like random forests and neural networks. In the context of factor timing, decision trees may have improved predictions for the Mkt-RF factor due to their ability to model complex, non-linear relationships that are typical in market dynamics. However, their tendency to overfit is evident from the negative out-of-sample R-squared values for factors like RMW and CMA, suggesting that while the model fits the training data well, it does not generalize effectively to new data. Additionally, the simplistic decision boundaries and high variance of decision trees can be limiting when dealing with the intricate patterns of financial factors, potentially leading to suboptimal performance where more nuanced modeling is required. To address this issue, our immediate strategy is to shift towards using **random forests** as our second model.

In our exploration of machine learning models for factor timing, the random forest approach demonstrated improved performance relative to both the baseline stepwise regression models and decision trees. According to *Appendix 4*, stepwise regression yielded predominantly negative or marginal out-of-sample R-squared values across factors, with Mkt-RF at -0.7585 and RMW at -0.4001, highlighting significant challenges in predictive accuracy. In contrast, our random forest model showed enhancements in these areas; notably, Mkt-RF improved to an out-of-sample R-squared of 0.2322, and even though RMW remained negative at -0.3784, this was an improvement over the decision tree's -1.6233. *(Appendix 5)*. The random forest model also displayed a more consistent performance across different factors compared to decision trees, which exhibited a wide variance in results—from a positive R-squared of 0.1982 for Mkt-RF to extremely negative values for other factors. In summary, the random forest provided a more stable and generally improved predictive performance, suggesting its suitability for handling the complexities of financial data where simpler models like decision trees may falter.

However, we further decided to test out our third model, **Neural networks**. We believe that it is worth exploring beyond random forests due to their ability to model even more complex non-linear relationships and interactions between variables, which can be crucial in capturing the underlying patterns in financial data. Additionally, neural networks' flexibility in architecture design allows for fine-tuning and optimization to specific characteristics of financial time series, potentially yielding superior predictive performance. While neural networks offered some improvements over stepwise regression, such as a notably higher R-squared for Mkt-RF at 0.0831 compared to -0.7585, they generally underperformed relative to random forests. For instance, neural networks resulted in lower R-squared values for factors like RMW and CMA compared to random forests, which demonstrated more stable and less negative outcomes across these factors. *(Appendix 5)*.

Despite the advanced capabilities of neural networks in capturing complex patterns, **random forest remains a preferable choice** in this context. The stability, robustness, and consistently less negative or higher R-squared values offered by random forests suggest that they are better suited to manage the diversity and complexity of financial data. Moreover, these findings suggest that machine learning models, particularly random forests, can significantly enhance predictive accuracy beyond traditional statistical models.

Trading Signal & Strategy

Upon identifying our optimal model, we advance to implementing our factor timing strategy by utilizing **trading signals**. Our factor timing strategy employs a machine learning-based approach using Random Forest Regressor to predict the behavior of various financial factors individually as well as in combination. For each factor, such as 'Mkt-RF', 'SMB', 'HML', 'RMW', and 'CMA', and their combinations, the predictive model is trained using a set of key economic and market predictors. These predictors include inflation, interest rate, unemployment rate, consumer confidence index, market volatility (VIX), and historical dividend and earnings data. The process begins with a train-test split, allocating 80% of the data for training and 20% for testing, carefully avoiding data shuffling to maintain the chronological integrity of the time-series data.

Once the model is trained, it generates predictions for the test dataset, which are then used to create trading signals. A positive prediction, suggesting an increase in the factor value, triggers a 'buy' signal (1), while a negative prediction, indicating a potential decline, triggers a 'sell' signal (-1). To align these signals with realistic trading scenarios, they are shifted forward by one period. This shift reflects the practical lag between signal generation and execution, ensuring that each signal is applied to the subsequent period's trading session. This methodological rigor in generating and applying trading signals is crucial for evaluating the strategy's real-world applicability and effectiveness. Through this approach, we aim to harness the predictive power of machine learning to optimize factor timing strategies, enhancing the potential for informed, data-driven investment decisions.

Our trading signal generation method, using Random Forest, fundamentally differs from the approach described in 'PythonPractice8.ipynb', which primarily relies on technical indicators such as moving averages and Relative Strength Index (RSI) within a linear regression framework. While the *PythonPractice8* notebook uses these technical indicators to predict future returns and generate straightforward long and short trading signals based on the direction of these predictions, our method employs a more sophisticated Random Forest model that considers a broader set of economic and financial predictors beyond just price-derived indicators. This approach not only allows for capturing complex non-linear interactions among variables but also involves shifting the generated signals forward by one period to account for real-time trading execution delays. This adjustment enhances the realism and applicability of our trading strategy, ensuring that signals are actionable and aligned with practical market dynamics, unlike the immediate application of signals in the 'PythonPractice8.ipynb' approach, which may not fully account for execution lag and might rely too heavily on historical price data without considering broader market influences. Moreover, we use a signal shifting method in contrast to the lookahead return calculations in 'PythonPractice8.ipynb'.

Results

The detailed outcomes for each factor are presented in *Appendix 7*, where it's noted that the 'Mkt-RF' factor achieved the highest Sharpe ratio[2] at 0.733. This ratio, while the highest, is still relatively low and is associated with significant drawdowns. Traditionally, a Sharpe Ratio above 1 is considered acceptable. A Sharpe Ratio below 1 suggests that the investment returns are not adequately compensating for the risk taken. In an effort to refine our approach, we extended our analysis to include various combined factor models, including *Equal, Performance, Confidence, and Volatility Weighted* combinations of the five Fama French factors, with an aim to enhance our strategy's efficacy.

In the context of factor timing, transitioning from individual factor models to combination factor models is often justified by the benefits of diversification and risk management. Combining factors such as momentum, value, and size into a unified model capitalizes on the strengths of each while mitigating their individual weaknesses. This strategy can enhance performance stability across various market conditions, as opposed to relying solely on a single factor which may be more susceptible to specific economic cycles or market trends. Ultimately, a combo factor model can offer a more consistent performance with reduced portfolio volatility and drawdowns, thereby potentially improving the risk-return profile of an investment strategy.

Equal Combo Factor Strategy combines predictions from multiple factors using equal-weighted signals. Performance Weighted Strategy applies weights based on historical Sharpe ratios of factors to generate signals. Confidence Weighted Strategy utilizes confidence scores derived from predictions to create weighted signals. Volatility Weighted Strategy calculates weights based on historical volatilities of factors and generates signals.

The findings, detailed in *Appendix 9*, reveal that all but the Volatility Weighted combination fell short of the best individual factor strategy regarding expected returns and Sharpe ratio. The Volatility Weighted factor alone, however, yielded a Sharpe ratio exceeding 1 but also exhibited a considerable maximum drawdown. To mitigate this, we integrated a stop-loss mechanism into the Volatility Weighted factor, which not only lowered the maximum drawdown but also improved our strategy's Sharpe Ratio to 1.808. The mechanism implements a stop-loss mechanism to limit losses beyond a predetermined threshold, enhancing performance by reducing potential downside risk.

Furthermore, when comparing our modified random forest Volatility Weighted combined factor to the approach in ‘PythonPractice8.ipynb’, which uses lookahead returns, our method outperformed with better results and a higher Sharpe ratio (1.808 vs 1.308), as documented in *Appendix 11*. Additionally, market timing strategies applying optimal weights to each combination factor, as explored in ‘Exercise 4’,

---

[2] The Sharpe Ratio measures the risk-adjusted return of an investment, indicating how much excess return you receive for the extra volatility you endure holding a riskier asset.

did not outperform our individual factor strategy in terms of overall performance, as detailed in *Appendix 12*.

Transaction Costs

Transactional cost testing is significant for evaluating the actual profitability and risk of a trading strategy. Particularly for a factor timing strategy, which adjusts investment weights based on various market factors such as market risk, size, and value, the transaction costs incurred can impact the overall returns. The overall goal for testing transaction costs is to quantitatively measure how much these costs reduce investment returns and degrade the efficiency of the strategy. As we utilized three different transaction costs ranging from 10 bps to 50 bps (*Appendix 8),* the following changes are observed:

1. **Annualized Mean Returns**: There is a decrease from 0.63% to 0.61%. This indicates that higher transaction costs result in lower net profits.
2. **Sharpe Ratio**: This decreases from 0.154 to 0.147. The Sharpe Ratio, an indicator of return per unit of risk, shows that the investment efficiency decreases as transaction costs increase.

Thus, it is evident that higher transaction costs significantly impair the profitability and efficiency of the strategy, **especially in the long term**, as the effect may be compounded. This demonstrates that high transaction costs can substantially diminish the performance of the strategy, highlighting the importance of selecting as low a transaction cost as possible. However, *Appendix 8* also suggests that the trading strategy has a certain degree of sensitivity to transaction costs, though not a disproportionate one. The changes in annualized mean returns and the Sharpe Ratio in response to increased transaction costs from 10 bps to 50 bps are relatively minor. This indicates that while the strategy does experience a decline in performance with higher transaction costs, the impact is not overly dramatic. The consistency in the maximum drawdown implies that the risk profile of the strategy remains relatively stable despite the varying costs.

Conclusion

Our exploration into factor timing with various machine learning models underscores the intricate nature of predicting market behaviors. While decision trees showed limited efficacy, our investigations with random forests offered marginal improvements, supporting the skepticism about factor timing's feasibility echoed by Cliff Asness. The enhancements, albeit modest, did not dramatically elevate performance metrics such as the Sharpe ratio, which remained below 1 for most individual factors. This suggests that even sophisticated models struggle with the unpredictable dynamics of financial markets.

The most promising results emerged from a volatility-weighted combination of factors, which achieved a Sharpe ratio above 1, indicating that a more integrated approach to factor combination and risk adjustment might hold potential. Yet, to mitigate substantial drawdowns, this method must be integrated with additional strategies, like employing stop-loss measures. Furthermore, the introduction of transaction costs highlighted the practical challenges, as even modest costs significantly eroded performance benefits. This practical insight aligns with Asness's views on the complexities and potential pitfalls of factor timing.

In summary, our findings straddle the debate between Rob Arnott's support for factor timing and Asness's caution, reflecting the nuanced reality that while certain strategies may show theoretical promise, their practical application remains fraught with challenges. The results affirm the need for cautious optimism and further research in the field of factor timing, emphasizing a balanced approach to harnessing the benefits while acknowledging the inherent risks.

Appendix

Appendix 1: Final Set of Predictors

Column	Description	Reasoning
Index_x	The index level of the S&P 500.	Provides a baseline for market performance; changes can indicate broader market movements.
Dividend	The dividend yield of the S&P 500.	Higher yields may signal undervaluation of the market and a cue for timing decisions.
Rfree	The risk-free rate, often based on government securities.	Influences the discount rate and valuation of assets; a key component in pricing models.
Inflation	The rate at which the general level of prices for goods is rising.	Affects the real returns of stocks and can influence investor sentiment and market dynamics.
InterestRate	The yield on the 10-year Treasury note.	Indicates the long-term risk-free rate influencing cost of capital and equity valuation.
UNRATE	The unemployment rate.	An economic indicator that can predict market cycles and factor performance.
CCI	Consumer Confidence Index, a measure of consumer sentiment.	Reflects consumer confidence which can drive economic activity and impact markets.

VIX	CBOE Volatility Index, a measure of market risk and investor sentiment.	A high VIX suggests increased risk aversion which could signal market turning points.
D12	The past 12 months of dividends for the S&P 500.	Used to estimate dividend growth and gauge market expectations.
E12	The past 12 months of earnings for the S&P 500.	Helps in evaluating company performance and overall market health.
svar	Stock variance, a measure of stock price volatility.	Indicates market uncertainty which can be a factor in timing market entries and exits.

Appendix 2: Simple Regression on Each Factor based on Predictors

Factor	R-squared	Predictor < p-value (0.05)
Mkt-RF	0.575	UNRATE, CCI, VIX, E12
HML	0.149	N/A
SMB	0.198	Inflation
RMW	0.181	Index_x, Inflation,
CMA	0.088	N/A

Appendix 3: Results from Forward Stepwise Regression (Lowest AIC & the most significant predictor)

Factor	AIC	Selected Predictors
Mkt-RF	616.8	'svar', 'Inflation', 'D12', 'VIX', 'UNRATE', 'E12', 'CCI', 'InterestRate'
SMB	594.7	'E12', 'svar', 'Inflation', 'VIX'
HML	568.4	'E12', 'VIX', 'InterestRate'
RMW	498.9	'svar', 'Inflation', 'D12', 'VIX'
CMA	457.8	‘UNRATE'

Appendix 4: Out of Sample R^2 for Each Factors

- Out-of-Sample R^2 for Mkt-RF: -0.7585
- Out-of-Sample R^2 for SMB: 0.0128
- Out-of-Sample R^2 for HML: 0.0967
- Out-of-Sample R^2 for RMW: -0.4001
- Out-of-Sample R^2 for CMA: 0.0170
- Combined Out-of-Sample R^2 across all factors: -0.2064

Appendix 5: Comparison between Decision Trees, Random Forest, and Neural Networks

- **Decision Trees**
  (Mkt-RF) Out-of-Sample R^2 : 0.19823402447058613
  (SMB) Out-of-Sample R^2 : -0.7284583901839252
  (HML) Out-of-Sample R^2 : -0.22380369707040382
  (RMW) Out-of-Sample R^2 : -1.62335469967288
  (CMA) Out-of-Sample R^2 : -0.9526342417138858

- **Random Forest**
  (Mkt-RF) Out-of-Sample R^2 : 0.2322423143589778
  (SMB) Out-of-Sample R^2 : 0.06154033932941361
  (HML) Out-of-Sample R^2 : -0.08378033680544839
  (RMW) Out-of-Sample R^2 : -0.3783800230030685
  (CMA) Out-of-Sample R^2 : 0.028756811018629813

- **Neural Networks**
  (Mkt-RF) Out-of-Sample R^2 : 0.08308530418337057
  (SMB) Out-of-Sample R^2 : -0.15708472476963542
  (HML) Out-of-Sample R^2: 0.03321518101715937
  (RMW) Out-of-Sample R^2: -0.36705774987305273
  (CMA) Out-of-Sample R^2: -0.10870345692053607

Appendix 6: Calculation of Metrics for Each Factor (without implementing a trading strategy)

	Annualized Returns	Annualized StDev	Max Drawdown	Sharpe Ratio
Mkt-RF	17.465492	3.244433	-0.279100	5.382706
SMB	-2.514323	1.372590	-0.999990	-1.833023
HML	2.517185	2.529168	-0.996554	0.994603
RMW	-0.771646	2.005076	-0.999697	-0.385678
CMA	-1.195200	0.976175	-0.993360	-1.226077

Appendix 7: Random Forest (the best model) on Each Factor Separately

Factor	Max Drawdown	Annualized Mean Returns	Annualized Standard Deviation	Sharpe Ratio

HML	-12.34%	4.17%	10.53%	0.386
Mkt-RF	-213.94%	8.15%	10.96%	0.733
SMB	-761.53%	4.65%	9.43%	0.482
CMA	-4.61%	0.65%	6.01%	0.091
RMW	-0.10%	0.64%	4.11%	0.155

Appendix 8: Transaction Costs on Each Basis Points (bps)

Transaction Cost (bps)	Max Drawdown (%)	Annualized Mean Returns (%)	Annualized Standard Deviation (%)	Sharpe Ratio
10 bps	-0.10%	0.63%	4.11%	0.154
20 bps	-0.10%	0.63%	4.11%	0.152
50 bps	-0.10%	0.61%	4.11%	0.147

Appendix 9: Experiment Combo Factor Strategies using Different Weights

Weighting Strategy	Max Drawdown (%)	Annualized Mean Returns (%)	Annualized Standard Deviation (%)	Sharpe Ratio
Equal Weights	-3.25%	1.43%	4.77%	0.278
Performance Weighted	-18.70%	2.56%	6.98%	0.351
Confidence Weighted	-2.93%	10.18%	16.78%	0.600
Volatility Weighted **without** Loss Function	-21731.87%	20.26%	12.26%	1.64
Volatility Weighted **with** Loss Function	-0.19%	21.72%	11.95%	1.808

Appendix 10: Quadratic Costs Test

Initial Capital	Alpha	Final Cumulative Return (%)
$100,000	1e-06	6.92%
$100,000	5e-06	6.91%
$100,000	1e-05	6.89%
$500,000	1e-06	6.91%
$500,000	5e-06	6.85%
$500,000	1e-05	6.77%
$1,000,000	1e-06	6.89%
$1,000,000	5e-06	6.77%
$1,000,000	1e-05	6.61%
$5,000,000	1e-06	6.77%
$5,000,000	5e-06	6.13%
$5,000,000	1e-05	5.35%

Appendix 11: Reviewing the PythonPractice8 (Predictions using lookahead method)

Strategy	Max Drawdown (%)	Annualized Mean Returns (%)	Annualized Standard Deviation (%)	Sharpe Ratio
Mkt-RF	0.0%	6.83%	9.63%	0.698
Volatility-based weighting	-27.97%	12.86%	9.75%	1.308
Performance weighting	-7.96%	0.98%	15.33%	0.057

Appendix 12: Reviewing the Exercise 4 (Timing with Optimal Weights)

Factor	Max Drawdown (%)	Annualized Return (%)	Sharpe Ratio
Mkt-RF	0.0%	4.28%	0.634
HML	0.0%	-0.04%	-1.823
SMB	0.0%	1.84%	0.149
RMW	0.0%	0.34%	0.635
CMA	0.0%	-5.33%	-1.453