DAFI REPORT

Introduction

The Fama-French five-factor model has been a cornerstone in asset pricing and risk management, providing a robust framework for understanding cross-sectional stock returns. However, recent advancements in financial research have introduced new factors, such as industry-specific factors, momentum (MOM), short-term and long-term reversal factors, which aim to capture additional sources of risk and return. This report evaluates whether these new factors outperform the traditional Fama-French factors by analyzing their risk metrics both in-sample and out-of-sample. The analysis is based on the variance-covariance matrices, ex-ante and ex-post betas, and key risk metrics such as volatility, Value-at-Risk (VaR), and Expected Shortfall (ES).
We then create six different portfolio strategies and evaluate the performance with proper metrics. We also conduct a cross-portfolio analysis and evaluate the risk exposures of portfolios.

Methodology

The analysis is conducted using two datasets that have been called Dataset1 and Dataset3. These two datasets have in common the financial return of the SP500. In order to create the datasets, we have developed on python different functions that were able to load the data, read its contents and perform some operations such as create a date column or remove the missing values.

1. Dataset1: Contains the traditional Fama-French factors (Mkt-RF, SMB, HML, RMW, CMA):
    - Market (MKT-RF): Market return minus the risk-free rate
    - Size (SMB - Small Minus Big): Difference in returns between small-cap and large-cap companies
    - Value (HML - High Minus Low): Difference in returns between value and growth stocks
    - Profitability (RMW - Robust Minus Weak): Difference in returns between firms with high and low profitability
    - Investment (CMA - Conservative Minus Aggressive): Difference in returns between firms with conservative and aggressive investment strategies

2. Dataset3: Includes the some of the Fama-French factors (Mkt-RF, RMW) augmented with new factors (Industry_Factor_1, Industry_Factor_2, MOM):
    - Momentum (MOM): Stocks with positive trends tend to outperform
    - Industry-Market: Sector return relative to the overall market (on the website, there are not the factors, but only the portfolios. For this reason, we did the pca of the 17 industries, and we kept the 5 most important)

The following steps were performed:

1. Linear Regression: Computed three different times. The first time on the dataset with all the variables, the second time on the dataset with only the Fama French factors, and finally, on the dataset3.

2. Lasso Regression: The lasso regression has been useful to identify the most relevant factors to use in the further analysis.
3. Variance-Covariance Matrices: Computed for both datasets to understand the relationships and risk structures of the factors.
4. Ex-Ante and Ex-Post Betas: Estimated to evaluate the sensitivity of the factors to market movements.
5. Risk Metrics: Calculated in-sample and out-of-sample to assess the performance and stability of the factors.

Lasso Regression

Lasso regression (Least Absolute Shrinkage and Selection Operator) is a type of linear regression that includes L1 regularization, which helps in feature selection by shrinking some coefficients to zero. The lasso regression has been useful in selecting the most relevant factors among the full set of factors. It showed that value, size and investment were not relevant factors among the Fama French. The relevant industry factors were only the first and second component, and the short and long term factors were not significant. The relevant factors were used to create the Dataset3.

Linear Regression

We performed linear regression on the two different datasets (Dataset1 and Dataset3) and analyzed the mean squared error (MSE) to evaluate how well the factors estimated the daily return of the S&P 500. The results indicated strong predictive performance, as the MSE for Dataset3 was significantly lower than that of Dataset1 (0.0000018385 vs. 0.0003476591). In addition, the linear regression of Dataset3 was able to explain better the training data.

Analysis of Variance-Covariance Matrices

High variance values indicate that a factor is highly volatile and contributes significantly to portfolio risk.

Covariance values indicate the degree to which factors move together. Positive covariance suggests that the factors tend to move in the same direction, while negative covariance suggests they move in opposite directions.

In the context of the report, the variance-covariance matrices for Dataset1 and Dataset3 help identify the risk structure and relationships between the factors.

<u>Dataset1 (Fama-French Factors)</u>

The variance-covariance matrix for Dataset1 reveals the following:

- Mkt-RF exhibits the highest variance (0.000129), indicating its dominant role in explaining market risk.
- SMB and HML show moderate variances, suggesting they capture size and value premiums effectively.
- RMW and CMA have lower variances, indicating weaker explanatory power for profitability and investment factors.
- Covariances between factors are generally low, implying minimal multicollinearity.

Dataset3 (New Factors)

The variance-covariance matrix for Dataset3 highlights:

- Mkt-RF remains the dominant factor with a variance of 0.000131.
- MOM has a variance of 0.000082, indicating its potential to explain momentum-driven returns.
- Covariances between new factors and traditional factors are relatively low, except for Industry_Factor_1 and Mkt-RF (0.000582), which may indicate overlapping risk exposures.

Beta Analysis

A beta greater than 1 indicates that the factor is more volatile than the market, while a beta less than 1 indicates lower volatility.

A positive beta means the factor tends to move in the same direction as the market, while a negative beta means it moves in the opposite direction.

In the report, ex-ante and ex-post betas are calculated to evaluate the expected and realized sensitivity of the factors to market movements.

Ex-Ante vs. Ex-Post Betas

- Dataset1: The ex-ante betas for Mkt-RF, SMB, and HML are positive, indicating their expected sensitivity to market movements. The ex-post betas are consistent but slightly lower, suggesting a marginal overestimation of risk in the ex-ante analysis.
- Dataset3: The ex-ante beta for Mkt-RF (0.1786) is lower than in Dataset1, reflecting the dilution of market risk by the new factors. The ex-post beta for Industry_Factor_1 is positive, confirming its relevance in explaining returns.

Risk Metrics

1. Volatility measures the standard deviation of returns, representing the risk or uncertainty of an investment. Higher annualized volatility indicates greater risk, as the returns are more dispersed around the mean.Annualized volatility scales the daily volatility to a yearly basis by multiplying by the square root of the number of trading days in a year (typically 252).
2. VaR is a risk metric that estimates the maximum potential loss of a portfolio over a specified time horizon at a given confidence level
3. ES provides a more comprehensive measure of tail risk than VaR, as it considers the severity of losses beyond the VaR threshold.

In-Sample Risk Metrics

- Volatility: 0.3933
- VaR 95%: 0.2539
- Expected Shortfall (ES): 0.4162

Out-of-Sample Risk Metrics

- Volatility: 0.13930. This indicates that the new factors maintain stability in unseen data.
- VaR 95%: 0.1785, suggesting similar downside risk in out-of-sample testing.
- Expected Shortfall (ES): 0.2176, confirming the robustness of the new factors.

Comparative Analysis

In-Sample vs. Out-of-Sample

- Volatility: Reduction in volatility from in-sample to out-of-sample, indicating that the models generalize well to unseen data.
- VaR and ES: The out-of-sample VaR and ES are lower than their in-sample counterparts, suggesting that the model is not overfitting and maintains predictive power.

Fama-French vs. New Factors

- The new factors in Dataset3 do not significantly alter the risk profile compared to Dataset1. However, the inclusion of Industry_Factor_1 and MOM provides additional explanatory power, as evidenced by their positive betas and moderate variances.
- The new factors exhibit lower covariances with traditional factors, reducing multicollinearity and enhancing model interpretability.

Conclusion

The new factors introduced in Dataset3 demonstrate comparable risk metrics to the traditional Fama-French factors, both in-sample and out-of-sample. While they do not significantly reduce portfolio risk, they provide additional explanatory power and reduce multicollinearity. This suggests that the new factors are a valuable extension to the Fama-French framework, particularly for capturing industry-specific and momentum-driven risks. Future research could explore the performance of these factors in different market conditions and asset classes to further validate their robustness.

Portfolio Construction

We construct and evaluate six different portfolio strategies using different approaches.

1. <u>Equal-Weight Portfolio</u>: It assigns an identical weight to each asset within the portfolio, ensuring that no single stock dominates returns. This strategy is simple yet effective, leveraging diversification benefits while avoiding concentration risk.
2. <u>Equal-Weight Dynamic Portfolio:</u> It starts with an equal allocation of weights and then dynamically adjusts them based on realized returns, ensuring the portfolio remains balanced relative to the returns of the individual stocks.
3. <u>Minimum Variance Portfolio</u>: It minimizes the overall portfolio volatility by optimizing asset weights through covariance matrix estimation. This approach employs a quadratic programming approach constrained by non-negative weights and full investment (weights summing to one).
4. <u>Mean-Variance Portfolio:</u> It maximizes expected returns for a given level of risk by optimizing asset weights. The optimization problem incorporates an objective function that maximizes the trade-off between expected return and risk, using a risk-aversion coefficient to adjust the trade-off.
5. <u>Market-Capitalization Weighted Portfolio:</u> It assigns weights to assets based on their relative market capitalization. Here larger companies hold greater influence on portfolio returns.
6. <u>Risk Parity Portfolio</u>: It equalizes risk contributions across assets, ensuring that each asset contributes proportionally to total portfolio risk rather than capital allocation. This

methodology assigns lower weights to highly volatile assets and higher weights to less volatile ones, promoting stability in diverse market conditions.

Portfolio Evaluation

We evaluate the performance of each portfolio using these metrics:

1. Annualized Total Return: it measures the compounded rate of return over a given time horizon. It is calculated by taking the cumulative product of daily returns and computing the growth rate over the total period. The formula adjusts returns based on the number of trading days in a year (261)

$$ATR = \left(\frac{cumulative\ return_{end}}{cumulative\ return_{start}}\right)^{ann\ const/nr\ days} - 1$$

2. Annualized Volatility: it measures the risk of a portfolio's returns. It is computed as the standard deviation of daily returns, scaled by the square root of the annualization constant.

$$AV = \sigma_{returns} \times \sqrt{ann\ const}$$

3. Efficiency Ratio: it measures the risk-adjusted performance of a portfolio by comparing its mean daily return, annualized by the number of trading days, to its annualized volatility. This metric indicates how effectively a portfolio generates returns relative to the amount of risk it undertakes.

$$Efficiency = \frac{\mu_{returns} \times ann\ const}{AV}$$

4. Drawdown: It measures the decline from a portfolio's peak value to its lowest point before a new high is achieved.

$$Drawdown = \frac{cumulative\ return}{cumulative\ return_{max}} - 1$$

5. Maximum Drawdown: It measures the most severe drawdown over the entire period

$$Maximum\ drawdown = min\ (drawdown)$$

6. Value at Risk: it measures the potential loss in portfolio value at 95% confidence level with the historical simulation method assuming normal market conditions. The calculation involves centering and sorting returns to find the 5%-percentile return and then take the negative value scaled by the square root of the annualization constant.

$$VAR\ (95\%) = -\ Percentile(returns,\ 5\%) \times \sqrt{ann\ const}$$

Considerations

The Min Variance Portfolio has the lowest returns but also the lowest volatility, making it a conservative choice for risk-averse investors. The Equal-Weight and Risk Parity portfolios have similar behavior, with low returns and moderate volatility. The Market-Cap Weighted Portfolio shows slightly higher volatility due to exposure to large companies. The Dynamic Equal-Weight Portfolio offers higher returns but at the cost of much higher volatility, due to frequent rebalancing. The Mean-Variance Portfolio provides the highest returns but also high volatility.

All portfolios were impacted by the 2008 crisis, with the Mean-Variance showing the deepest drawdown. The Min Variance portfolio, focused on risk reduction, showed the smallest decline. The Mean-Variance Portfolio has the highest Efficiency Ratio, reflecting its superior return relative to risk, while the Dynamic Equal-Weight also ranks high in efficiency. The VaR analysis confirms that portfolios with higher risk, like Mean-Variance and Dynamic Equal-Weight, face larger potential losses.

Cross-Portfolio Analysis

We perform a cross-portfolio analysis to have valuable insights into the risk exposures of the portfolios. In the first phase of the analysis, we perform factor regressions for each portfolio using our dataset of factors. These factors capture the systematic risk exposure of each portfolio. The results from the regression show the betas, which represent the portfolio's sensitivity to the individual factors, indicating how much the portfolio's returns move in response to changes in each factor.

Then we calculate the residuals for each portfolio, which are the differences between the observed portfolio returns and the predicted returns based on the factor model. These residuals represent the idiosyncratic risk of each portfolio—risk that cannot be explained by the factor model. Then the covariance matrix of the residuals is calculated to evaluate the correlations between the unexplained risks of the portfolios. From the covariance matrix we can see that most residuals are uncorrelated. This indicates that the factors are explaining the returns perfectly or we are just overfitting.

We also calculate the correlation between portfolios' factor exposures to see the relationships between portfolios based on their factor sensitivities. This is important because portfolios with high correlation in their factor exposures may behave similarly in response to changes in the factors, leading to a potential concentration of risk. Equal-Weight and Risk Parity portfolios have very high correlations with others, indicating they have similar factor exposures, which leads to less diversification when combined with other portfolios. Dynamic Portfolio shows low correlations with most portfolios, suggesting it behaves differently from others, likely due to its adaptive nature.

The **risk evaluation** involves calculating three types of risk metrics:

1. **Historical Risk**: This is measured as the standard deviation of the portfolio returns, which indicates the total volatility of the portfolio. It provides a general understanding of how much a portfolio's returns deviate from the mean over time.
2. **Factorial Risk**: This is calculated as the standard deviation of the residuals (unexplained risk), representing the portion of the risk that is not explained by the factors. A higher factorial risk suggests that the portfolio's return is more driven by factors not captured by the model.
3. **Robust Risk**: This is calculated as the Median Absolute Deviation (MAD), to account for extreme values and outliers in the data. This is an alternative to the standard deviation, which can be overly sensitive to outliers.

Dynamic and Mean-Variance have the highest factorial risks, indicating that these portfolios' returns are not entirely explained by the chosen factors. In contrast, the other portfolios have much lower factorial risks, indicating that their returns are better explained by our factors. These portfolios also have higher robust risks, indicating that they experience more outliers or large swings that aren't captured by the model. The other portfolios show much lower robust risks, implying greater stability.