BEM 114 Final Report
Lunar Systematic - Sujit Iyer, Sayuj Choudhari, Patrick Donohoe

Strategy Intuition and Development

This investment strategy takes positions in at-the-money call options based on a classification problem: Will the underlying asset increase by a certain threshold from the strike price over the next two weeks? If the model predicts "yes", then the investor purchases the 10 DTE call option. 10 DTE call options were the target for this strategy because they are relatively cheaper and push the investor to satisfy liquidity constraints. This choice was validated by comparison with other DTE timeframes and it was determined that 10 DTE call options were optimal, producing consistently positive returns with a high volume of trading.

ML models have been used for investment strategies before to predict returns of an asset or specific parameters at the next timestep. This learning problem contains an infinite output space, which poses significant challenges due to the inherent complexity of the task. Coupled with the volatility of financial markets, these regression strategies have steep tail risk. The purpose of this investment strategy is to generate stable returns through ML by simplifying the task to a binary classification problem.

This strategy invests in call options which further reduces tail risk through its stop-loss capabilities: Investors can limit their potential losses to the option premium. Thus, by fixing a success rate on investments and a return threshold for the positions to meet, the investment strategy has a clear path to profit.

$$\mathbb{E}(Profit) = p(R_{asset} - OP_{asset}) + (1-p)(-OP_{asset})$$

$$p(R_{asset} - OP_{asset}) + (1-p)(-OP_{asset}) > 0$$

$$R_{asset} = 2(OP_{asset}) \rightarrow p(2(OP_{asset} - OP_{asset}) + (1-p)(-OP_{asset}) > 0$$

$$2p - 1 > 0 \rightarrow \boxed{p > .5}$$

We can calculate the expected value of the profit of the strategy by considering the two scenarios: The investment succeeds with probability *p*, and yields a profit calculated by the difference between the asset return and the price of the representative option. The investment fails with probability 1 - *p*, and yields a loss determined by the price of the option (since we don't need to exercise it). For the purposes of this strategy, we can set the asset's return to be twice the price of the option. Simplifying the problem, we determine that for the strategy to be profitable, the investment success rate *p*, must be greater than 0.5.

The training set was created by iterating through each daily timestep for each stock in the S&P 500, and generating a binary signal indicating whether or not the underlying asset gained 5% from the strike price over the next two weeks. Historical data analysis of the options with underlying assets that gained 5% indicated that the average price of 10 DTE at-the-money call options was approximately 2.3% of the underlying asset's price. We use pricing data for underlying assets that gained 5% only in order to most accurately estimate the price of options the strategy invests in. For the purposes of backtesting, we conservatively used a value of 2.5% to account for fees, slippage, and other frictions. This measure can only limit returns, which makes it reasonable to implement.

Factor Development

To develop an accurate trading classification, meaningful factors need to be input to the model that can identify profitable trades. For profitability in options, this is mostly dictated by the momentum of the underlying asset as well as its volatility. We evaluate the following 4 factors as inputs to the model:

1) Standard deviation of returns: A rolling standard deviation of returns (30-day window) is tested as a factor for indicating volatility of the underlying equity
2) Rate of change in moving average: The percentage rate of change over the previous 10 day-period in a long and short term (20 and 80 day) moving average is calculated as a factor for indicating momentum in the underlying equities movement. We use the rate of change in a moving average instead of direct price as the moving averages are a smoother curve and thus their rate of change is a more stable statistic.
3) Moving average percentage difference (MAPD): The percentage difference between a short and long term moving average (12 and 26 day periods) is used as a momentum indicator in the underlying equity. A positive difference can indicate an existing trend of consistent higher than average recent returns, while a highly negative difference could indicate underperforming conditions and the possibility of an undervaluing/ trend reversal creating gains in the equity. The percentage difference is used instead of raw price difference in order to normalize the factor values across assets.
4) Average True Range (ATR): The ATR is a volatility indicator calculated by a rolling average (30-day window) of the True Range calculated at each day by the following formula: TR = Max(|High - Low|, |High - Previous Closing|, |Low - Previous Closing|) which is essentially the range size of the equity price over that given day. For normalization purposes, this range is divided by the closing price each day for this factor development so that the factor is normalized across assets and a predictive model can thus train on and apply to all asset data aggregated. In general, this indicator aims to capture the average percentage range of the equity price as a measure of its volatility.

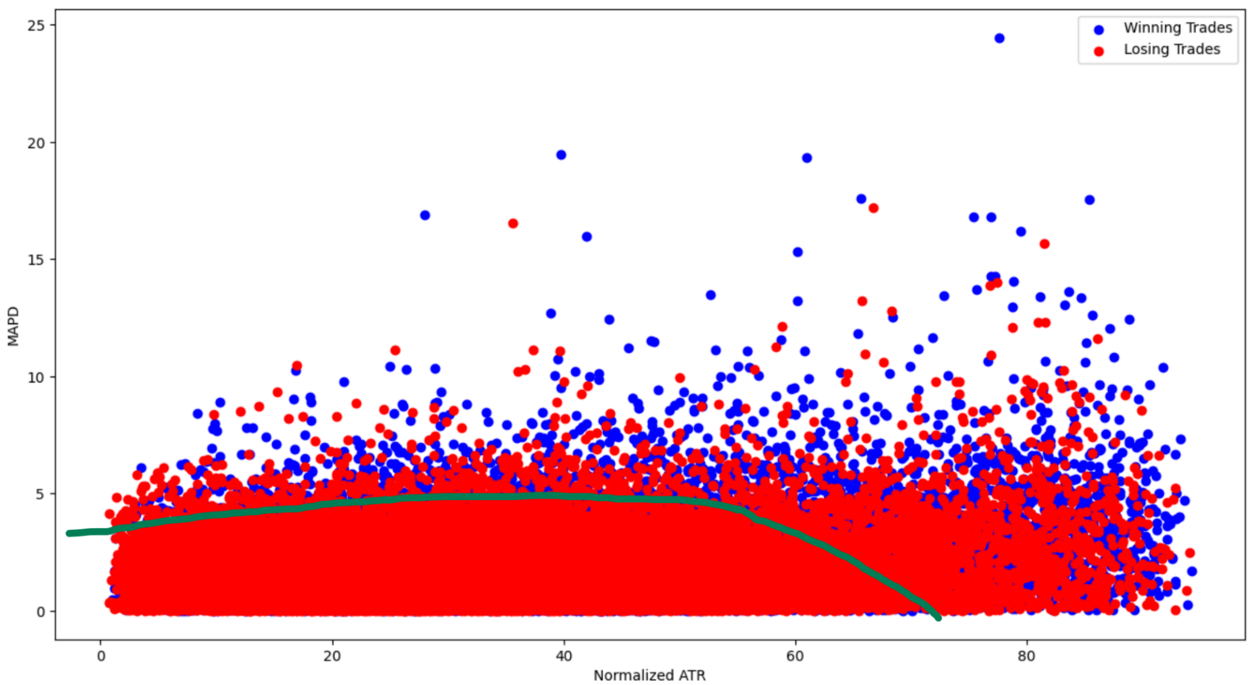
Given our ML problem from the previous section is designed to predict a return threshold of 5% from some set of input factors, statistical tests can be used to evaluate the usage of various factors. For simplicity of the model and reduced overfitting, we choose to only predict on the top 2 factors we identify in order to achieve better generalization. Firstly, biserial correlations can be used to identify trends in factor values against an indicator variable on whether the 5% threshold was reached. Biserial correlations are the ideal method for testing correlations between a nonbinary input and a binary output. One augmentation was made to the factors calculated, in taking the absolute value of the MAPD, as it resulted in a statistically significant biserial correlation compared to the standard MAPD. Running the biserial correlations lead to the absolute MAPD and ATR having the largest correlations to the 5% return threshold:

```
statistic=0.18679454821007954, pvalue=0.0)
statistic=0.11994904918600682, pvalue=4.850204781816039e-146
```

While these correlations seem insignificant in context of the entire dataset, it does signify that on the right edge of the distribution of these factors, concentrations of profitable trades likely exist and these concentrations can be learned by an ML model. Another significant result from these two factors was that they also exhibit a low correlation of ~.2 with each other, meaning that both of the factors provide different information on each trade and thus the combination of them will provide a diverse space of information for the ML model to learn on, hopefully improving accuracy. Taking the absolute value of the MAPD may have seemed like making the indicator a simple volatility measure of how far recent returns are from a historical average, but the lack of correlation between the MAPD and ATR, indicates that the MAPD does not track volatility but likely is just capturing recent excess returns and trend reversals/ undervalued conditions. The MAPD is also advantageous in its use of both short and long term moving averages, meaning it likely incorporates information from both factors on rates of change in the moving average, improving the models robustness by maximizing the information space it can learn on. Statistically, the goal of our factor identification process was to find 2 factors that are significantly correlated with returns in the underlying equity where concentrations of profitable trades exist, while also having both factors be relatively uncorrelated with each other to ensure each factor contributes independent information to the model, expanding the information space and ability of the model to learn. Through this process, we identify the absolute MAPD and ATR to be these 2 factors, and will use them as the input space for model predictions.

Signal Development and Classification Model Design

The biserial correlations for MAPD and ATR indicated that on the right edge of the distribution of the factors, concentrations of profitable trades could be learned by an ML model. To confirm this finding, a plot of ATR and MAPD was generated.



From this plot, we can see that there are indeed areas of profitable trades that an ML model would be able to learn. For visual purposes, we added the green line separating the plot into two distinct regions: points below the green line which are overwhelmingly a majority losing trades and points above the green line which represent a healthy medium between winning and losing trades. The points below the green line are noisy data; there is no need to waste training weights and decision boundaries to classify this region because any position taken in this area will likely lead to a loss. It is more reasonable to train on the points above the green line so that the model can accurately determine the regions of winning trades. Therefore, both the training and test sets were filtered to include only those instances where the MACD is greater than 3.75.

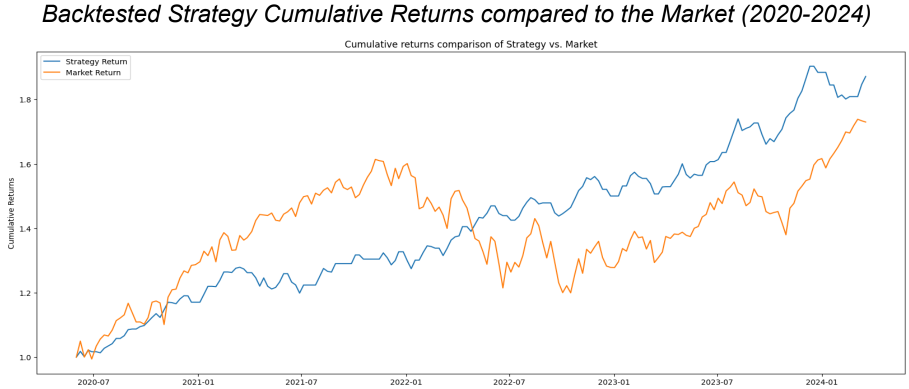
The model architecture that was landed on for training was a bagged network of decision trees. Decision Tree Classifiers can accurately capture non-linearities, which is essential for our dataset as shown in the above plot. The bagged network structure reduces overfitting which is a major concern for volatile financial data. Lastly, the model architecture allows for confidence measurements. The confidence measurement in this case represents the probability of a profitable investment, or p, as written in the earlier mathematical equations. Therefore, using this confidence measurement, we decided to only take on positions which predicted a winning trade with greater than 60% confidence. This 60% parameter was comfortably larger than the minimum profitable p = 50% parameter calculated earlier. These features allowed us to be reasonably confident that the investment strategy would generate consistent positive returns.

Before training, the dataset is collapsed to a weekly timeframe. Classifications run each Monday using the data generated from previous trading weeks (MACD and ATR signals are shifted down to avoid lookahead bias). It is determined which options are likely profitable investments and an equal weighted position is taken. An important portfolio mechanism is that options are exercised immediately if the underlying asset price reaches a 5% gain from the strike price. This mechanism assuages liquidity constraints, ensuring that roughly 65% of capital is available at each week for position-taking. Finally, although “winning trades” are desired, where the asset reached a 5% gain, it is still possible to exercise options on their last day till expiry if the loss can be mitigated or a lesser return can be achieved. Therefore, options are also exercised if the closing asset price on the final day of the 10 DTE is higher than the strike price.

Results

To showcase the validity of our strategy, we have backtested our strategy on a 4-year period from 2020 to the present. The results from our backtest indicate that our strategy returns are largely uncorrelated with the market and we are generating significant alpha. The results from comparing our strategy to the market performance with CAPM and FF3 are largely the same, so we are going to focus on the CAPM comparison. The results are displayed on the following page:

Backtested Strategy Cumulative Returns compared to the Market (2020-2024)



Backtested Strategy performance compared with the Market-RF via CAPM

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.2876	0.069	4.139	0.000	0.151	0.425
Mkt-RF	0.1029	0.028	3.669	0.000	0.048	0.158

Omnibus:	4.011	Durbin-Watson:	1.681
Prob(Omnibus):	0.135	Jarque-Bera (JB):	3.060
Skew:	-0.173	Prob(JB):	0.217
Kurtosis:	2.501	Cond. No.	2.50

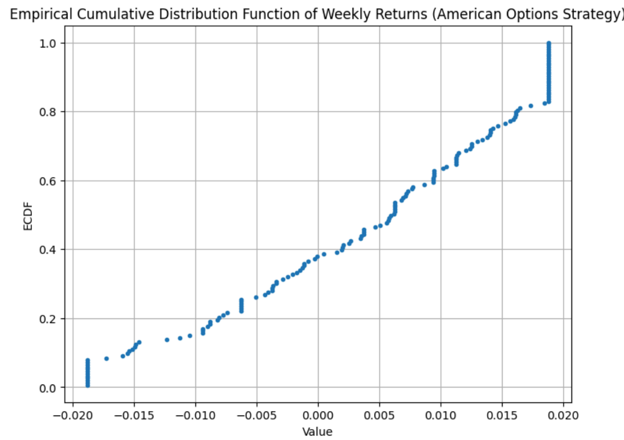
Backtested Strategy performance compared with the Market-RF via FF3

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.2855	0.069	4.118	0.000	0.149	0.422
Mkt-RF	0.0959	0.030	3.235	0.001	0.037	0.154
SMB	0.0782	0.045	1.728	0.085	-0.011	0.168
HML	0.0406	0.028	1.476	0.141	-0.014	0.095

Omnibus:	2.477	Durbin-Watson:	1.724
Prob(Omnibus):	0.290	Jarque-Bera (JB):	2.249
Skew:	-0.171	Prob(JB):	0.325
Kurtosis:	2.607	Cond. No.	2.83

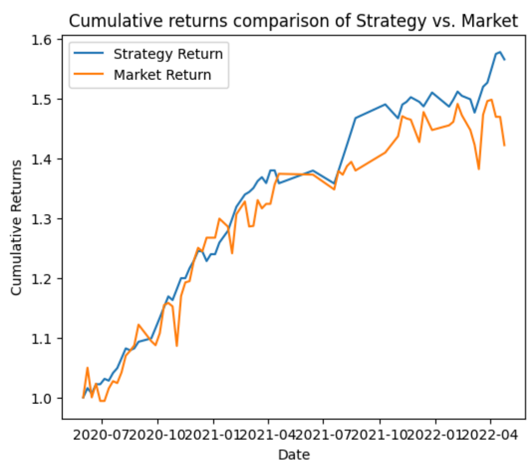
In the CAPM framework, our strategy has a market beta of 0.10, with the 95th percentile beta ranging from 0.05-0.16. This low beta value means we have successfully developed a strategy that provides a differentiated risk to the market. This is valuable for investors who are looking to diversify their exposures away from systematic risk of the market. Furthermore, our strategy produces a significant alpha of 0.288% per week, with a 95th percentile range from 0.151%-0.425% per week. This weekly alpha translates to 15.9% annual alpha. The returns for our strategy backtest produce an annualized Sharpe ratio of 1.993. This provides evidence that we have developed a strategy that provides significantly more reward per unit of risk to investors than investing in the market.

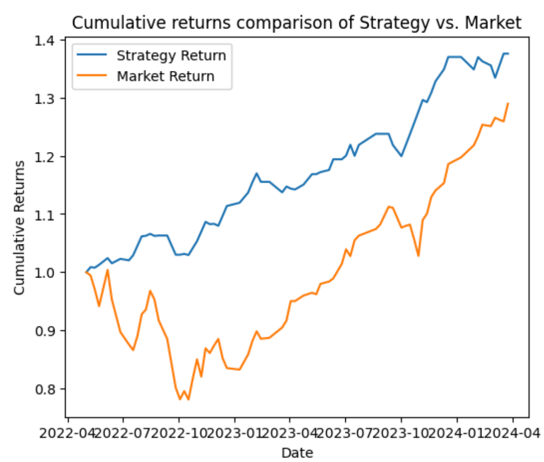
At the heart of our strategy is our models ability to find profitable trades based on the analysis of the factors that were previously discussed. In order for the strategy to be successful for American-style call options, we need predict a 5% gain in the underlying equities with a greater than 50% accuracy (utilizing our assumption that the ATM-call option price is 2.5% of the underlying). Our model is able to predict such gains with a 60.8% accuracy. In the chart below, we can see that over the backtesting period, our model generates positive weekly returns.



Risk Analysis and Extensions of Work

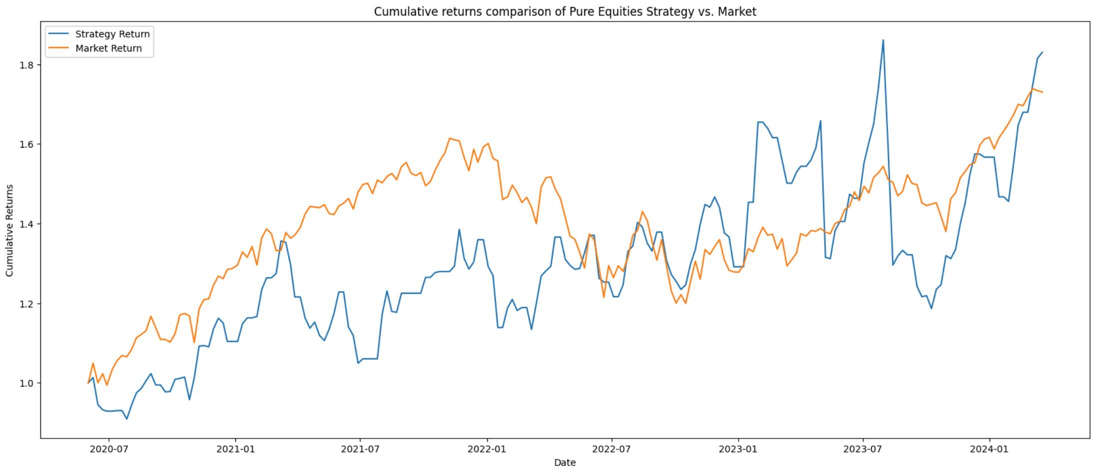
It is clear from the derivation of the strategy's expected profit per trade that classification accuracy of *p* > .5 is a crucial threshold to meet for a positive expectancy on trades. Thus, if the model we developed does not consistently provide such accuracy, our strategy is statistically expected to lose money. To ensure this risk is mitigated, we analyzed the model's training by evaluating returns on two different time periods of training. If the model is effective, profits should be realized across both testing periods and support the conclusion that the model is robust and can effectively train and predict on datasets independent of timeline. For this evaluation, we use a dataset that trains from 2018-2020 and predicts from 2020-2022, as well as a training period from 2020-2022 that predicts from 2022-2024:



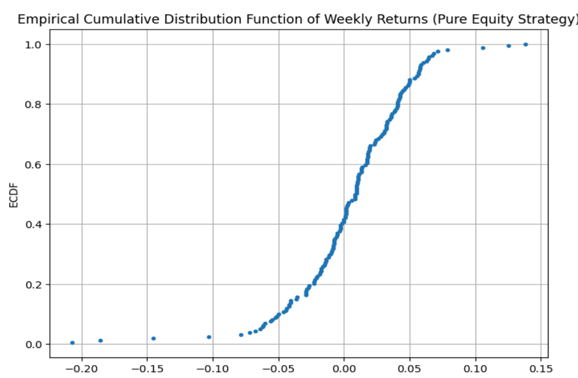


It is clear from the results above that the model's predictions produce consistent profits under both evaluation periods. Given the training period of both evaluations are different, it provides reasonable evidence that the model is robust and effectively generalizes over different time periods.

While the model predicted returns are strong, the use of an option to limit downside risk also negates returns to some degree, raising a question on whether it may be more effective to directly long the underlying asset. We test this strategy of just holding the predicted assets over the 10-day period and produce the following returns:

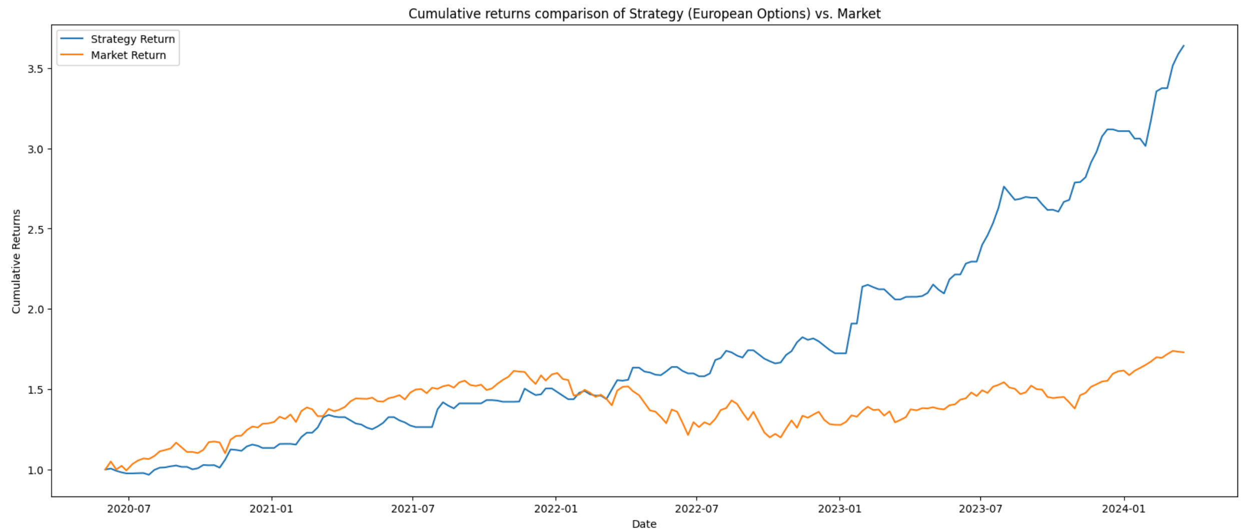


We observe that the returns of this holding strategy are roughly the same as the options strategy, but portfolio movements are significantly more volatile. Specifically, the holding strategy yields an annualized Sharpe ratio of ~.58 and market beta of ~.6, compared to the options strategy Sharpe ratio of ~1.993 and market beta of ~.1, statistics that significantly support the usage of the options strategy for trade execution with the model predictions rather than pure long trades. Use of options also allows for increased leverage, which under the case of the options strategy with low volatility could potentially yield higher returns with reasonably stable portfolio movements. Furthermore, the empirical cumulative distribution of weekly returns for the pure equity strategy (pure long trades) is given below:

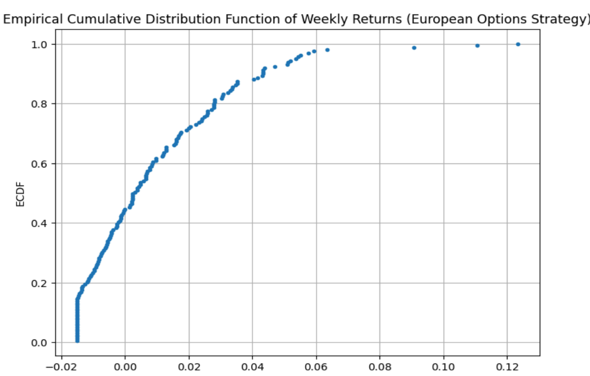


This distribution closely replicates a normal distribution with a slightly positive mean shift. However, we do observe some significantly realized tail risks including 4 weeks where the portfolio loses over 10% of its value. While the right tail also provides instances of large gains, the returns even out over the training period providing similar cumulative returns as the options strategy but with much less stability. Additionally, the options strategy has a profitability rate >60% for each week whereas the pure equity strategy falls below 60% accuracy (~57%), meaning the options strategy also more consistently provides positive returns on a week-by-week basis. It is evident from the statistically significant improvements of the options strategy versus the pure equity strategy, that the limited downside risk of the options strategy outweighs the limiting upside and that the options strategy provides a better portfolio in terms of common portfolio evaluation metrics such as market beta and Sharpe ratio.

Although we have shown the pure equity strategy in which we capture better upside to be unfavorable to the upside-limiting options strategy, we would like to extend this work to investigate possible mechanisms to add to the trading process so that the options strategy could capture additional upside. One possibility is the use of European options versus American options when trading, only executing contracts on the day of expiry. Given that equities have positive drift, we would expect that if the equity reaches the 5% return threshold that it has some significant probability of going beyond that until expiry. The returns of the European option-style trading process is given in the following plot:



The returns of the strategy using European call options yields nearly 4x magnitude greater returns and capture some largely profitable trades. The strategy also yields consistent returns and has no significant or extended down periods. This result in consistency of returns can also be observed in the empirical distribution of weekly returns below:



The profitability rate of the strategy remains nearly the same as the pure equity strategy with around 55% trades yielding positive returns, but the downside risk is limited and the upside trades are still captured. Mainly, comparing against the empirical distribution of the American options trading process with a maximum return of ~2%, nearly 30% of trades on the European options strategy are above that maximum return and over 10% are twice (>4% return), which is a statistically significant improvement in capturing the right tail of the distribution. As seen in the FF3 model estimates below with a low market beta of ~.2 and the Sharpe ratio of ~2.3, the use of European options for trading (including their relatively lower price to American options) is a modification worth continuing to research.

```
FF3 ESTIMATES
                             OLS Regression Results
==============================================================================
Dep. Variable:     Strat Model Returns   R-squared:                       0.071
Model:                             OLS   Adj. R-squared:                  0.057
Method:                  Least Squares   F-statistic:                     4.998
Date:                 Fri, 31 May 2024   Prob (F-statistic):            0.00232
Time:                         23:28:20   Log-Likelihood:                -418.99
No. Observations:                  199   AIC:                             846.0
Df Residuals:                      195   BIC:                             859.1
Df Model:                            3
Covariance Type:             nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.6494      0.144      4.513      0.000       0.366       0.933
Mkt-RF         0.1970      0.062      3.201      0.002       0.076       0.318
SMB            0.1017      0.094      1.082      0.280      -0.084       0.287
HML            0.0103      0.057      0.181      0.857      -0.102       0.123
==============================================================================
Omnibus:                       91.521   Durbin-Watson:                   1.797
Prob(Omnibus):                  0.000   Jarque-Bera (JB):              326.520
Skew:                           1.899   Prob(JB):                     1.25e-71
Kurtosis:                       7.995   Cond. No.                         2.83
==============================================================================
```

Further improvements include backtesting with exact options data versus average estimates, as well as incorporating an analysis of the options chain on each predicted asset to take a position through an option of optimized expectancy. For example, a slightly more in-the-money or out of the money option may have a slightly higher expected value given its price compared to the at-the-money option and may be a more ideal position. We hope to continue testing and implementing these changes to highly optimize the portfolio.