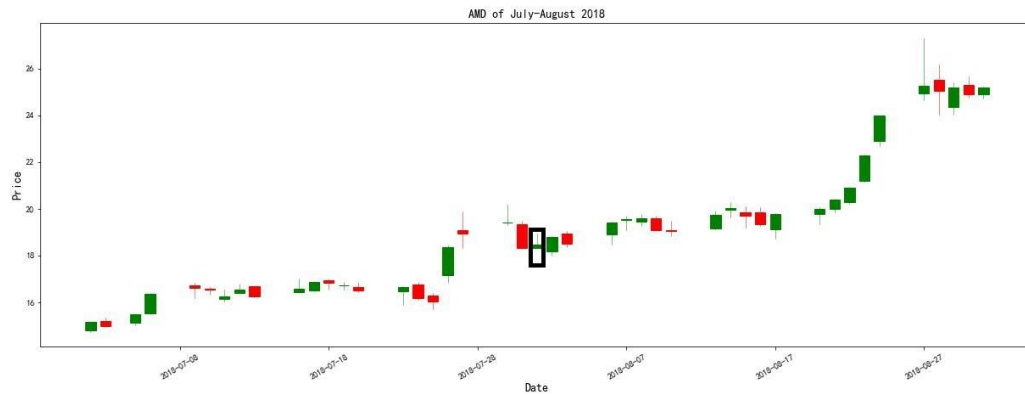


Use the K chart to build a stock selection model



Chuyi Wang Chuyiwww@bu.edu

Yuxuan Zhou zyxuan@bu.edu

Linglan Xu xll@bu.edu

Yiming Ding yimingd@bu.edu

Yitao Huang yhuang99@bu.edu

12.2022

MF 703: Programming for Math Finance

BACKGROUND

The K chart of the stock price is also known as the candle chart. After years of development, it has become a very mature technical analysis method of stock price. Its function is to reflect the mentality of market participants and the changing process of participants' mentality with the continuous development of the market through the shape of a K chart.

As an important part of the candle chart, many studies have shown that the upper and lower shadow lines play a greater role in market timing. For example, if there is a long upper shadow line in the rising price of a stock, it is believed that the stock is under great selling pressure and the shorts are dominant.

Our team thought that since this factor can be used for market timing, can it also be used for stock selection strategy? Therefore, our team wants to form a stock selection factor through the upper and lower shadow lines of stocks and conduct a backtest on it. Then compare its excess return rate with that of SPY500 and other indicators such as the value of IC and RankIC to measure the performance of the factor.

INTRODUCTION

We will implement our project with python.

Our group will calculate the upper and lower shadow lines for every trading day and standardize the data by dividing the 5-day rolling mean of upper and lower shadow lines:

$$\text{upper shadow lines} = \text{High} - \text{Max}(\text{Open}, \text{Close})$$

$$\text{lower shadow lines} = \text{Min}(\text{Open}, \text{Close}) - \text{Low}$$

$$\text{standardized upper shadow lines} = \frac{\text{upper shadow lines}}{\text{5 days rolling mean}}$$

$$\text{standardized lower shadow lines} = \frac{\text{lower shadow lines}}{\text{5 days rolling mean}}$$

Then we will take the mean and the std of standardized shadow lines over the month as our factors and divide our data into 5 groups with an ascending order of factors. Then buy the stocks with equal-weighted and record returns of 5 portfolios by the end of the month. We will also similarly construct factors by W%R.

William's upper shadow lines = High - Close

William's lower shadow lines = Close - Low

standardized William upper shadow lines = $\frac{\text{William upper shadow lines}}{5 \text{ days rolling mean}}$

standardized William lower shadow lines = $\frac{\text{William lower shadow lines}}{5 \text{ days rolling mean}}$

Then repeat the process done before that takes the mean and std of monthly factors data and separates factors into 5 groups with an ascending order. Then record the P&L with 5 portfolios.

PROCEDURE

1. Archiving all S&P 500 stocks data and S&P 500 index in the required period.
2. Deleting the nan data and suspended stocks.
3. Using the formula introduced above, construct the two factors and use DataFrame to store it.
4. Take the mean and the std of standardized shadow lines over the month as our factors and divide our data into 5 groups based on monthly ranking. Then buy the stocks with equal-weighted and record returns of 5 portfolios by the end of the month.
5. Repeat the process that was done before that takes the mean and std of monthly factors data and separates factors into 5 groups with an ascending order. Then record the P&L with 5 portfolios. Also, use this to filter out the stocks with the best performance in a certain portfolio in a single month.
6. Calculate the annual return, annual volatility, Sharpe ratio, max drawdown, IR and alpha of the portfolios, benchmark, and excess return to analyze the performance of the factors in building portfolios.

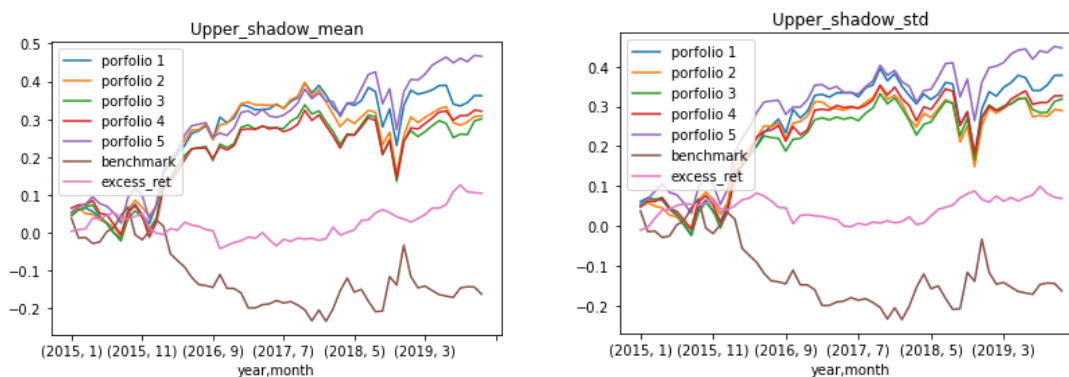
7. Use the monthly return as the benchmark, and calculate the excess return through the subtract between portfolio 5 and portfolio 1, which means long and short hedging.
8. Plot out the performance of the portfolios together with the benchmark and excess return based on the eight factors.
9. Plot out the candlestick of the stocks with the best performance in a single portfolio of each factor.

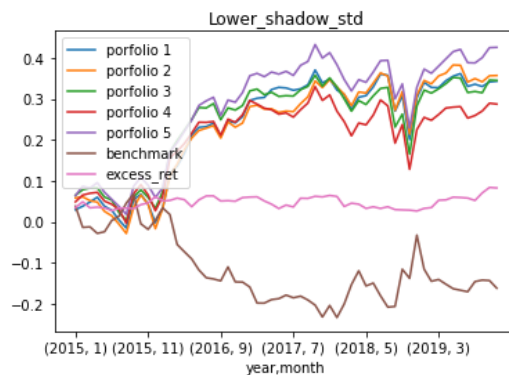
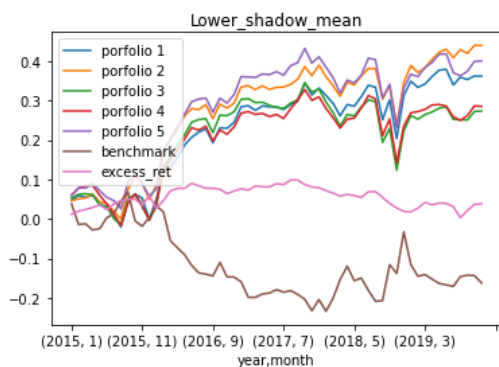
DATA

Since S&P 500 is a stock market index tracking the stock performance of 500 large companies listed on stock exchanges in the United States, We will use all the stocks in S&P 500 as our research sample. The backtesting time period is set from January 2015 to December 2019. The reason chose this time period is to try to avoid the uncertainty caused by Covid-19 in the Stock market starting in 2020. Using the yfinance API in Python, we archive all S&P 500 stocks data from Yahoo Finance in our required period. S&P 500 index as the benchmark. After deleting the nan data and suspended stocks, we start to calculate our factors.

RESULTS

Graph for Shadow Line Factors





Upper_shadow_mean

	portfolio 1	portfolio 2	portfolio 3	portfolio 4	portfolio 5	benchmark	excess_ret
Annual_Return	6.83%	5.69%	5.52%	5.99%	9.12%	-3.88%	2.01%
Annual_Volatility	12.25%	12.12%	12.20%	11.90%	12.08%	11.54%	5.09%
sharpe_ratio	60.10%	51.79%	50.17%	54.91%	78.52%	-28.61%	41.61%
Max_Drawdown	-16.40%	-22.72%	-19.56%	-17.57%	-15.08%	-26.87%	-9.33%
IR	20.77%	17.86%	17.85%	18.71%	23.56%		11.38%
Alpha	8.46%	6.99%	6.95%	7.32%	10.36%		1.76%

Upper_shadow_std

	portfolio 1	portfolio 2	portfolio 3	portfolio 4	portfolio 5	benchmark	excess_ret
Annual_Return	7.10%	5.28%	5.95%	6.13%	8.72%	-3.88%	1.32%
Annual_Volatility	12.77%	12.10%	11.78%	11.76%	11.91%	11.54%	4.37%
sharpe_ratio	60.16%	48.60%	54.97%	56.53%	76.28%	-28.61%	32.06%
Max_Drawdown	-20.93%	-19.52%	-16.38%	-16.94%	-14.40%	-26.87%	-8.24%
IR	20.74%	17.57%	18.61%	19.00%	22.94%		10.04%
Alpha	8.80%	6.72%	7.22%	7.42%	9.91%		1.03%

Lower_shadow_mean							
	portfolio 1	portfolio 2	portfolio 3	portfolio 4	portfolio 5	benchmark	excess_ret
Annual_Return	6.87%	8.58%	4.92%	5.22%	7.61%	-3.88%	0.71%
Annual_Volatility	11.90%	11.75%	12.27%	11.91%	12.54%	11.54%	3.91%
sharpe_ratio	61.80%	76.01%	45.31%	48.70%	64.86%	-28.61%	19.94%
Max Drawdown	-14.29%	-15.15%	-20.98%	-18.19%	-19.50%	-26.87%	-9.20%
IR	21.02%	23.66%	16.56%	17.08%	20.82%		9.03%
Alpha	8.41%	9.96%	6.29%	6.46%	8.97%		0.51%

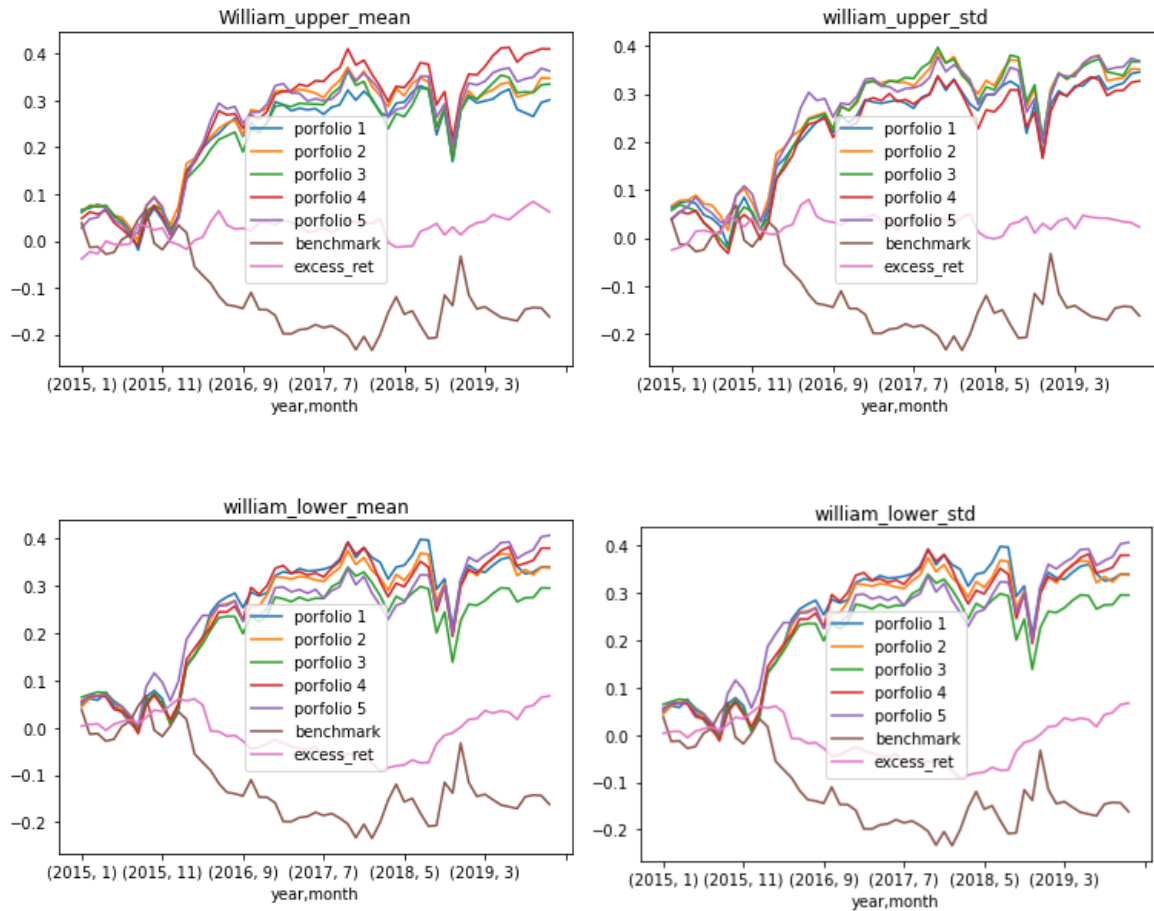
Lower_shadow_std							
	portfolio 1	portfolio 2	portfolio 3	portfolio 4	portfolio 5	benchmark	excess_ret
Annual_Return	6.49%	6.76%	6.48%	5.23%	8.21%	-3.88%	1.63%
Annual_Volatility	11.71%	11.91%	12.13%	12.24%	12.28%	11.54%	3.44%
sharpe_ratio	59.60%	61.00%	57.87%	47.74%	70.52%	-28.61%	48.81%
Max Drawdown	-16.78%	-14.80%	-18.57%	-19.58%	-19.65%	-26.87%	-3.78%
IR	19.72%	20.19%	19.62%	17.44%	21.99%		11.69%
Alpha	7.79%	8.13%	7.91%	6.71%	9.52%		1.61%

Upper-shadow mean portfolio is the best performing mode among all upper-shadow factors. The excess return of the hedging portfolio is calculated by P5-P1.

The Sharpe ratio compares the return of an investment with its risk. It's a mathematical expression of the insight that excess returns over a period of time may signify more volatility and risk, rather than investing skill. The information ratio is similar to sharpe ratio but using the excess return to the benchmark rather than a risk-free rate. A maximum drawdown (MDD) is the maximum observed loss from a peak to a trough of a portfolio before a new peak is attained. Show investors the max loss the portfolio may have.

Even though our result is not performing well, which excess_ret has a 2.01% annual return, 41% with Sharpe ratio, -9% maximum drawdown, and 11.38% on Information Ratio, the hedging portfolio still beats the market. It was caused by two huge systematic disasters in the US market. Once it happened by the end of 2015, and the other one happened in 2018.

Graph for William Shadow Line Factors



william_upper_mean							
	porfolio 1	porfolio 2	porfolio 3	porfolio 4	porfolio 5	benchmark	excess_ret
Annual_Return	5.49%	6.50%	6.27%	7.94%	6.90%	-3.88%	1.10%
Annual_Volatility	12.55%	12.45%	12.19%	11.80%	11.88%	11.54%	5.78%
sharpe_ratio	48.84%	56.84%	56.00%	70.77%	62.19%	-28.61%	21.81%
Max Drawdown	-16.08%	-18.78%	-18.77%	-18.73%	-16.37%	-26.87%	-7.75%
IR	18.16%	20.43%	19.42%	21.64%	18.86%		8.76%
Alpha	7.17%	8.29%	7.77%	9.09%	7.73%		0.53%

william_upper_std

	porfolio 1	porfolio 2	porfolio 3	porfolio 4	porfolio 5	benchmark	excess_ret
Annual_Return	6.43%	6.53%	6.95%	6.17%	7.02%	-3.88%	0.31%
Annual_Volatility	12.59%	12.72%	12.11%	11.54%	11.75%	11.54%	5.56%
sharpe_ratio	55.82%	56.16%	61.68%	57.64%	63.74%	-28.61%	8.27%
Max Drawdown	-16.86%	-18.26%	-18.86%	-16.71%	-17.41%	-26.87%	-8.06%
IR	20.27%	19.99%	19.98%	18.79%	19.61%		7.31%
Alpha	8.26%	8.28%	8.23%	7.29%	8.00%		-0.24%

william_lower_mean

	porfolio 1	porfolio 2	porfolio 3	porfolio 4	porfolio 5	benchmark	excess_ret
Annual_Return	7.51%	7.11%	7.52%	5.15%	5.86%	-3.88%	-1.83%
Annual_Volatility	12.63%	12.04%	12.02%	12.51%	11.29%	11.54%	4.83%
sharpe_ratio	63.77%	63.14%	66.43%	46.44%	56.06%	-28.61%	-35.79%
Max Drawdown	-19.53%	-17.42%	-14.89%	-20.41%	-16.46%	-26.87%	-20.54%
IR	20.75%	20.57%	21.73%	16.88%	18.92%		3.65%
Alpha	8.94%	8.46%	8.99%	6.58%	7.11%		-1.69%

william_lower_std

	porfolio 1	porfolio 2	porfolio 3	porfolio 4	porfolio 5	benchmark	excess_ret
Annual_Return	6.35%	6.31%	5.46%	7.17%	7.84%	-3.88%	1.27%
Annual_Volatility	12.13%	12.18%	11.67%	12.61%	12.01%	11.54%	4.65%
sharpe_ratio	56.89%	56.40%	51.44%	61.25%	68.89%	-28.61%	29.46%
Max Drawdown	-18.13%	-17.62%	-19.34%	-19.31%	-13.88%	-26.87%	-14.34%
IR	18.67%	19.10%	18.06%	20.86%	22.02%		11.44%
Alpha	7.54%	7.69%	6.81%	8.79%	9.23%		1.58%

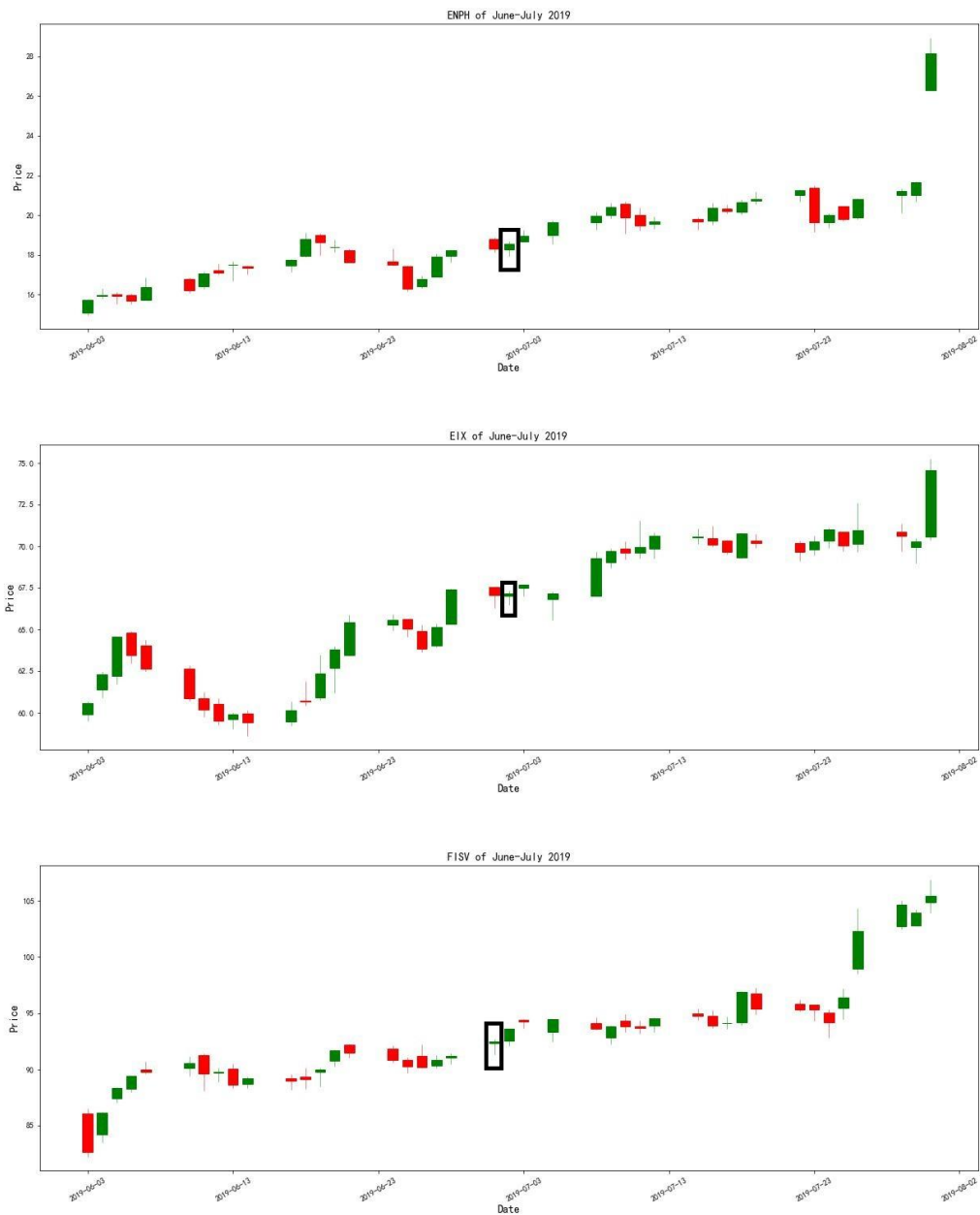
William-lower-stander portfolio is the best performing portfolio with the William factor. Similar to the factor before. It is not doing that well. The portfolio has a 1.27% annual return, 29.46% with Sharpe ratio, -14% maximum drawdown, and 11.44% on Information Ratio, and it is also better than our benchmark.

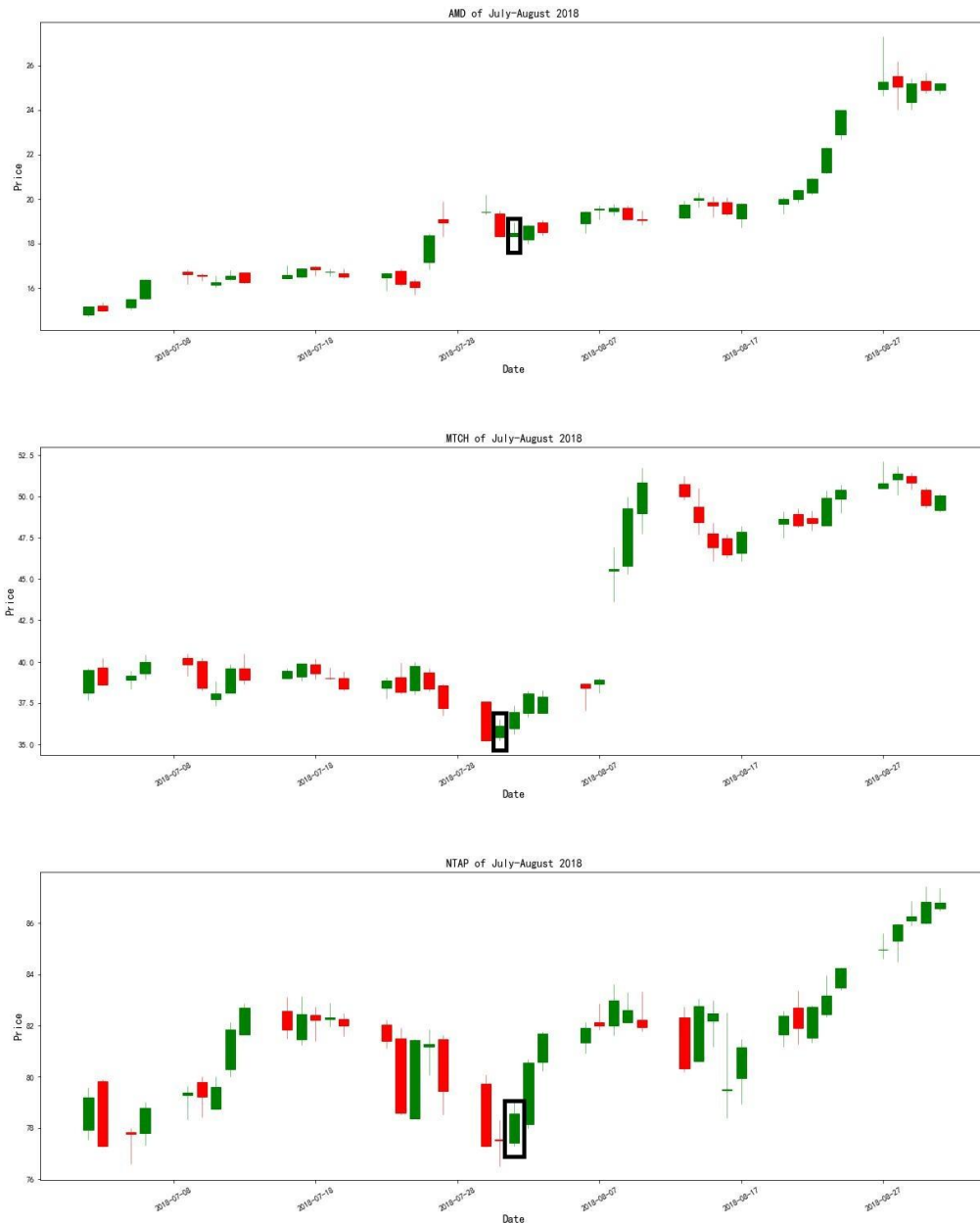
Next part, we plot out the candlestick of the stocks with the best performance.

First, we choose portfolio's that are based on the factors with the best performance, which is upper shadow lines mean and standardized William lower shadow lines. Then we found out the month with the highest excess return, which is July, 2019 and August, 2018. After that, we choose the top 3 stocks in return in these two months and plot out their candlestick graphs.

It should be noted that the period we choose when plotting is the month with the best performance and the month before this month. In this way, we can see the candlestick of the date at the start of the best month and the following trend. In the graphs, we point out this candlestick.

We use `candlestick_ohlc` from the `mpl_finance` package to show the results.





We can see that the candlestick of the certain date of the stocks all bodes well for an upward tendency in the following month. They all have a relatively large value of the factor, which means high pressure on buying and selling. That's how our strategy works.

CONCLUSION

We construct stock selection factors according to the definition of the upper and lower shadow lines of the candle chart. The backtest results show that the factor based on the upper shadow line of the candle chart has good stock selection ability, while the lower shadow line factor of the candle chart has poor stock selection ability.

Among them, the average factor of "upper_shadow mean" of the candle chart has the best effect, with an excess return of 2.10%, Sharpe ratio of 41.61%, the maximum drawdown of -9.33%, and IR of 11.38%.

In addition to the upper and lower shadow lines of the candle chart, the William Index is also often used to measure the overbought and oversold phenomenon of the market, but its understanding of the market buying and selling forces is different from that of the candle chart. Research has found that the William index can sometimes correct the incorrect judgments of candle charts.

According to the understanding of the buying and selling force of the William index, we define the William upper and lower shadow line and construct the stock selection factor. The backtest results show that the stock selection effect of William_lower shadow line factor is better, which just makes up for the deficiency of the normal lower shadow line factor of the candle chart.

Among them, the standard deviation factor "William_lower_std" has the best effect, with an excess return of 1.27%, sharpe rate of 29.46%, a maximum drawdown of -14.34%, and IR of 11.44%.

Although the two factors of the stock selection effect are not outstanding, compared to the current year index, we still think that the upper and lower lines of the candle chart can help us when picking stocks.

IMPROVEMENTS

- 1.Frequency: We only test a small scope of data to test the daily exchange strategy, it does not mean that at every period of time the daily exchange strategy is not better than monthly exchange strategy.
- 2.Windows: When we standardize the upper shadow and lower shadow, we use the value of the past 5 days to calculate the mean as the denominator. For further study, we will try different windows.
- 3.Weight: Now we use equal weight to buy each stock and close the position at the end of each month, next time we might use volume to create a different weight of each asset.
- 4.New factor: The best factor is William lower std and shadow line upper mean, those two factors are complementary and can be used to create a new factor , which might also be effective and efficient.

REFERENCES

1. Williams, L. (n.d.). Original Williams %R. Williams Percent R Indicator (%R). Retrieved October 13, 2022, from <https://williamspercentr.com/the-original-percent-r>
2. Nison, S. (2021, February 12). Learn to use Candlestick charts from the master. Retrieved October 13, 2022, from <https://candlecharts.com/candlestick-training/>

DIVISION

Yuxuan Zhou:responsible for constructing factors, conducting long/short hedging tests, backtesting some factors and completing the report;

Linglan Xu: responsible for constructing some factors and index, collecting coding results and

writing slides for presentation, and completing the report;

Chuyi Wang: responsible for calculating IC, checking and helping all the procedures, aggregating and modifying the final code, and completing the report;

Yiming Ding: responsible for archiving all required data , reformatting and cleaning data, plotting out the candlestick graph and completing the report;

Yitao Huang: responsible for filtering data of factors by months and rank. Then divide stocks to 5 portfolios every month. Also constructed the result table and graph the portfolios and max drawdown.