ETF Pair Trading Strategy

Master of Quantitative Economics

University of California, Los Angeles

Xinyue Wu

Faculty Advisor: Samuel Borghese

May 29th, 2022

Abstract

Pair trading strategy is one of the quantitative analysis methods in finance applying in short-term securities. It means long and short two securities with positive correlation at the same time. This article chose ETF to build a pair trading strategy and used two methods to construct the strategy: the spread strategy and the price ratio strategy. After creating an ETF pair trading strategy and backtesting these three ETFs, the result concludes that: Pair trading strategy has a positive return when using the Spread. It performs relatively well compared with the method using price ratio. The annual and cumulative returns are positive, and the alpha of the spread strategy is 0.8, indicating that its excess return on the investment is 0.8 after considering the market's volatility.

Keywords: pair trading, ETF, quantitative analysis

Table of Contents

1 Introduction

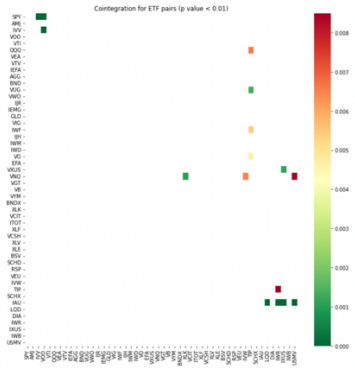
In quantitative economics, the pair trading strategy is one of the quantitative analysis methods in finance. A pair trading strategy means long and short two securities with a positive correlation at the same time. It aims to find the excess profit in two pairs of securities in a short-term by recognizing market mispricing. The pair trading strategy has been applied in many securities, like stocks, exchange rates, and ETFs.

This article chose ETF to build a pair trading strategy. The top 25 ETFs are selected from May 1st, 2021, to May 1st, 2022. All data are downloaded by yahoo finance. The dataset is split into two parts: the training set and the testing set 70 % of the data belongs to the training set, and the rest belongs to the testing set. The training set is used for training the algorithm and trading strategy, and the testing set aims to approach backtesting.

2 Cointegration

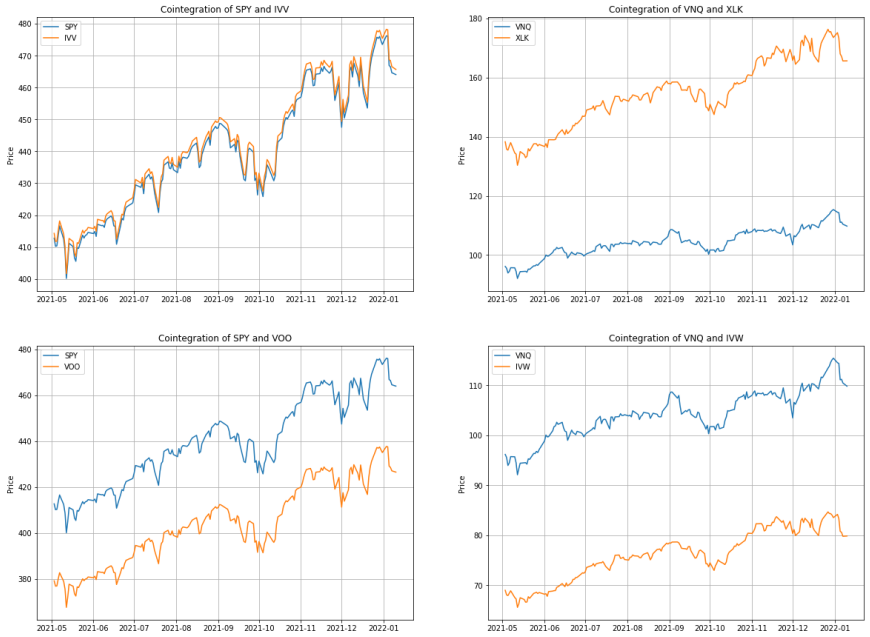
There are two steps in the cointegration test. First, test for a unit root in each component series individually, using the univariate unit root tests, says ADF test. If the unit root cannot be rejected, the next step is to test cointegration among the components, i.e., test whether the dependent variable is I(0).

Sixteen pairs are cointegrated based on the results since the p-value is less than 0.01 in the ADF test. They are SPY/IVV, SPY/ VOO, IVV/VOO, QQQ/TIP, VUG/TIP, IWF/TIP, VO/TIP, VXUS/IXUS, VNQ/XLK, VNQ/IVW, VNQ/USMV, TIP/IWR, IAU/LQD, IAU/IWR, IAU/IXUS, IAU/USMV.



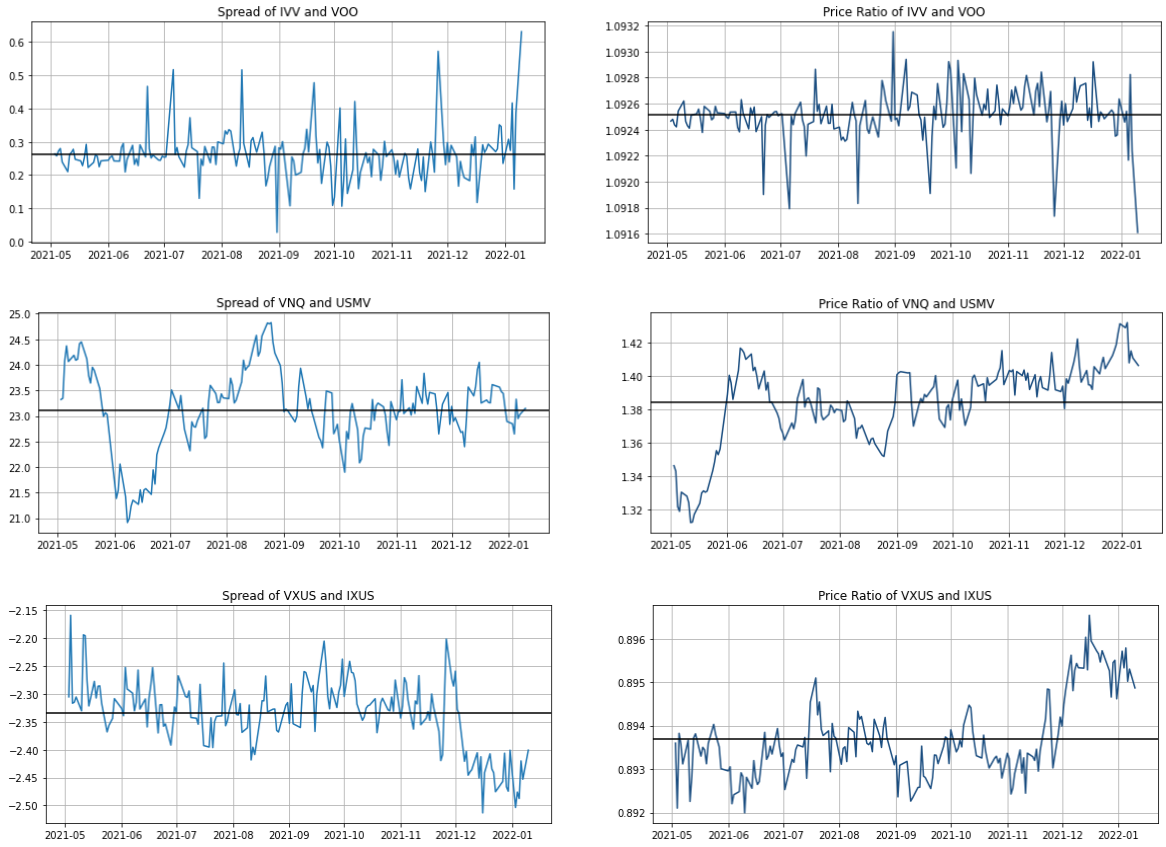
The cointegration plots aim to check the trend of the two ETFs further. Some pairs, like SPY/IVV, their cointegration results are too close since IVV is the exchange-traded fund iShares S&P 500 Index. Therefore, it is hard to find the deviation between these two pairs and gain profit. For some pairs, like IAU/LQD, and IAU/USMW, their cointegration plots include some outliners, affecting future prediction and excess profit. Thus, these pairs aren't chosen as the pairs in the portfolio. Some pairs, like QQQ/TIP, VUG/TIP, and IWF/TIP, are not evident in their positive correlation shown in the plots and thus have been excluded from the portfolio.

Finally, three pairs IVV/VOO, VNQ/USMV, and VXUS/IXUS, are chosen for the trading strategy.



3 Features construction

In this thesis, two methods are chosen to construct the features: Spread and price ratio. First is the Spread of two ETFs. To calculate the Spread, linear regression is used to get the coefficient for the linear combination to construct between two ETFs, and the Spread equals the asking price of one ETF minus the bid price of another ETF. The second feature price ratio equals the price of one ETF divided by the price of another ETF.

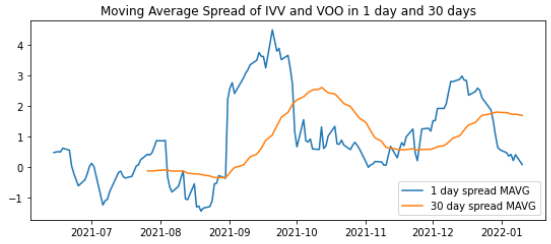


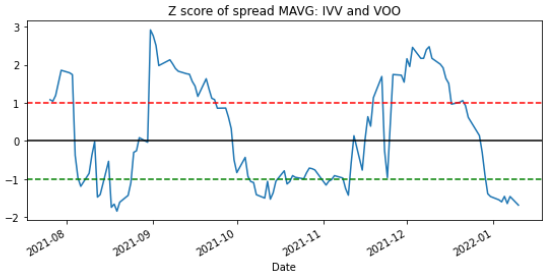
One set of criteria to measure the entire history is not a good strategy. If the market or society changes, the baseline of this Spread composed of X and Y may change. If we calculate a baseline based on the data of the past three years, then it is very likely that it does not perfectly express the moment. Therefore, our optimization only focuses on a continuously rolling window and calculates the baseline (average) within this window, which we call moving averages.

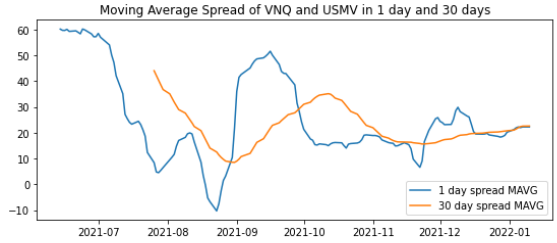
There are two methods used in this thesis: one is moving average Spread for one day and 30

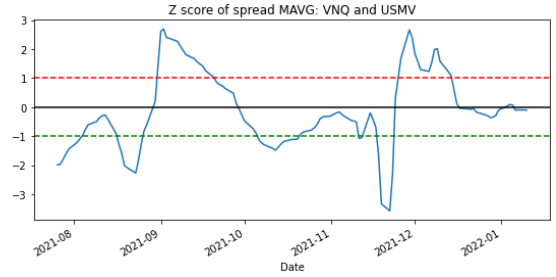
days, and another is moving average price ratio for one day and 30 days. After applying the moving average function to the Spread and price ratio, the z-score can be calculated for each pair. The z-score equals the moving average of 1 day minus the moving average of 30 days and then divided by the standard deviation of the moving average of 30 days.

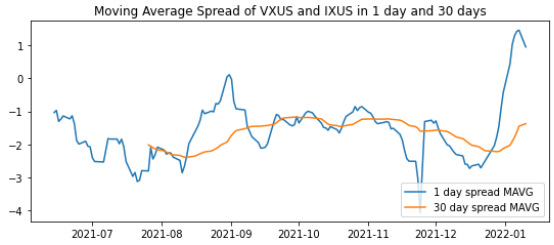
The results below show the moving average spread of chosen pairs and their z-score trend.

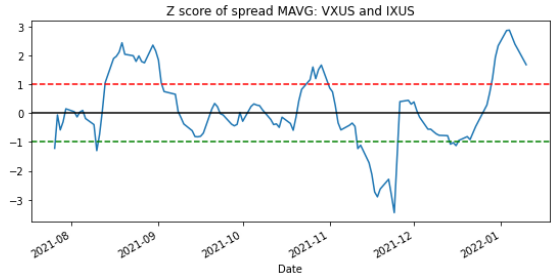




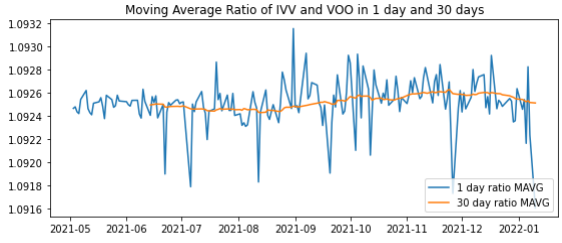


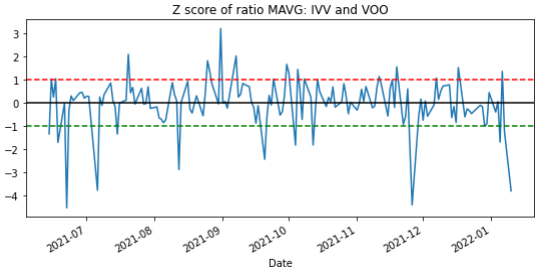


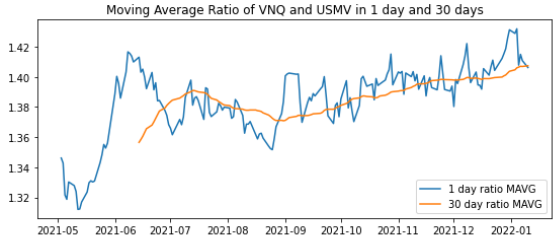


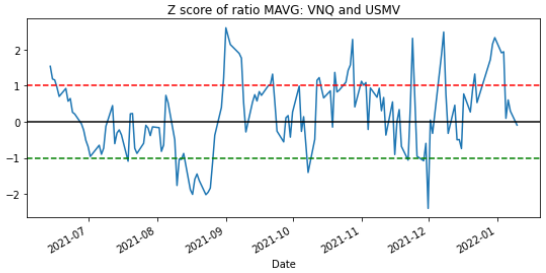


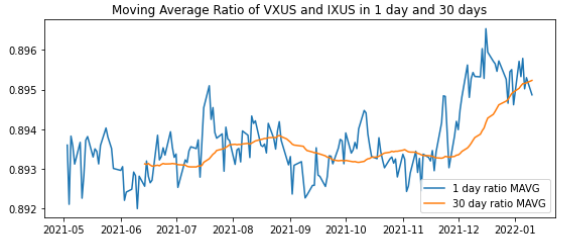
The results below show the moving average price ratio of chosen pairs and their z-score trend.

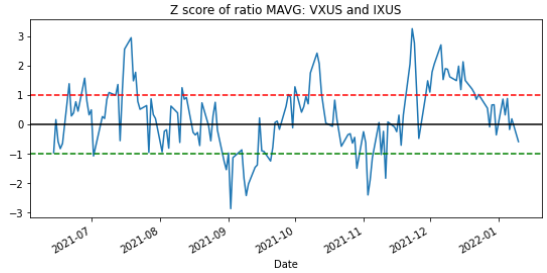










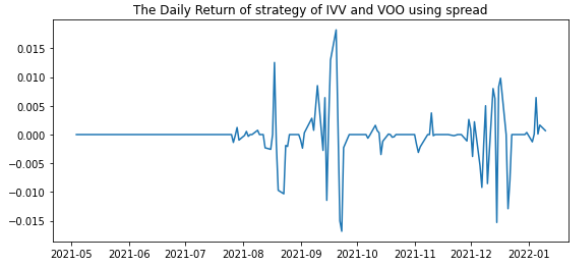


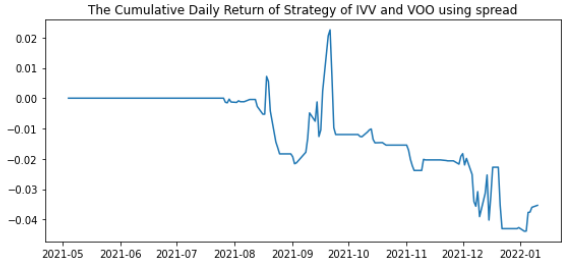
4 Strategy Construction

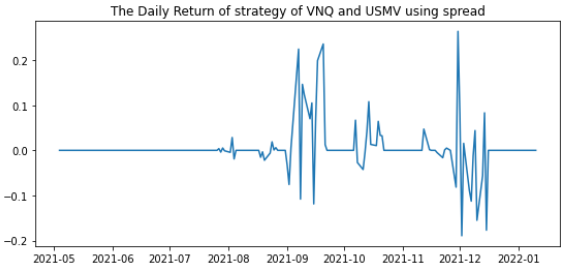
After calculating the moving average of Spread, moving average of price ratio, and z-scores, the trading strategy can be built based on these features. The trading rule is bought 1 when the z-score is below -1, meaning the Spread or ratio is expected to increase; sell -1 whenever the z-score is above 1, meaning we expect the Spread or ratio to decrease.

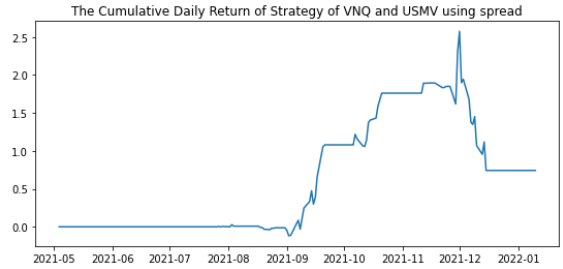
4.1 Construct a Trading strategy using Spread

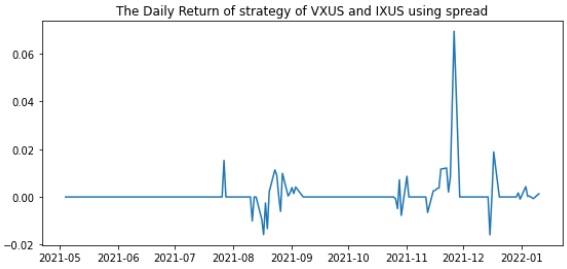
After applying the trading strategy to three pairs of ETFs, the daily return and the cumulative daily return of the spread trading strategy are calculated and plotted below. The daily return of the three pairs is mainly positive, and their cumulative daily returns keep increasing from 2021 to 2022, indicating a good result.

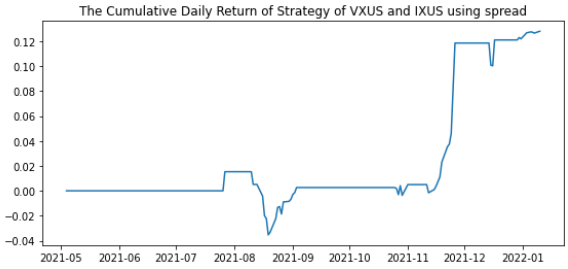






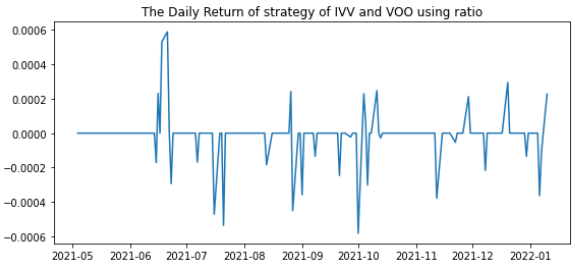


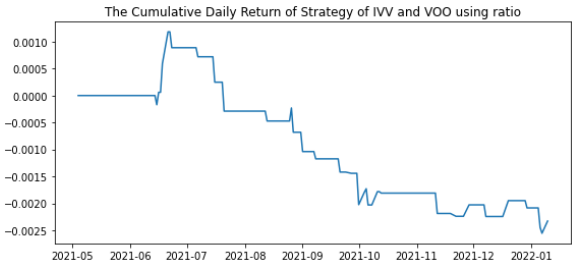


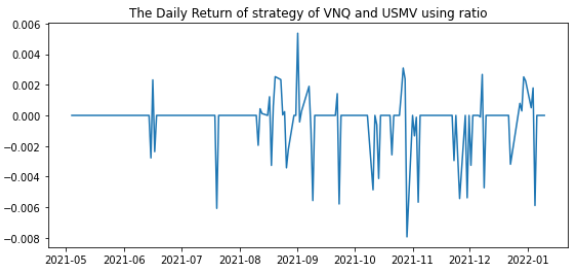


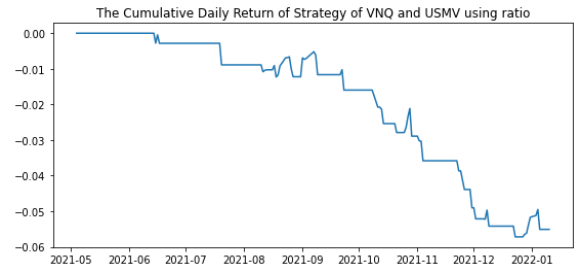
4.2 Construct a Trading strategy using the Price Ratio

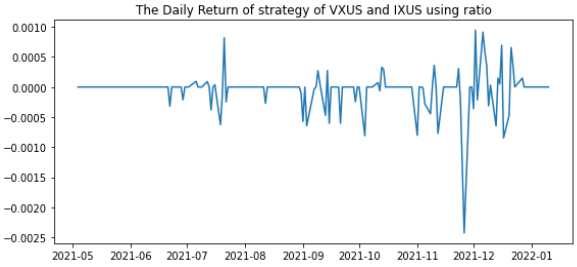
After applying the trading strategy to three pairs of ETFs, the ratio trading strategy's daily and cumulative daily returns are calculated and plotted below. The daily return of three pairs fluctuates around 0, and their cumulative daily returns decrease from 2021 to 2022, meaning that the strategy doesn't perform quite well.

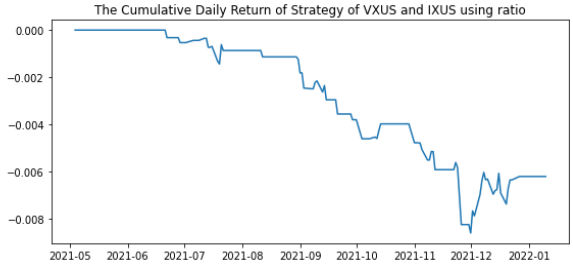












5 Trading Strategy Backtesting

5.1 Trading Strategy using Spread

First, the trading strategy was built based on the Spread. The technical indicators, like annual returns, cumulative returns, annual volatility, Sharpe ratio, skew, kurtosis, daily value at risk, alpha, and beta, are calculated below, indicating the basic information of the strategy.

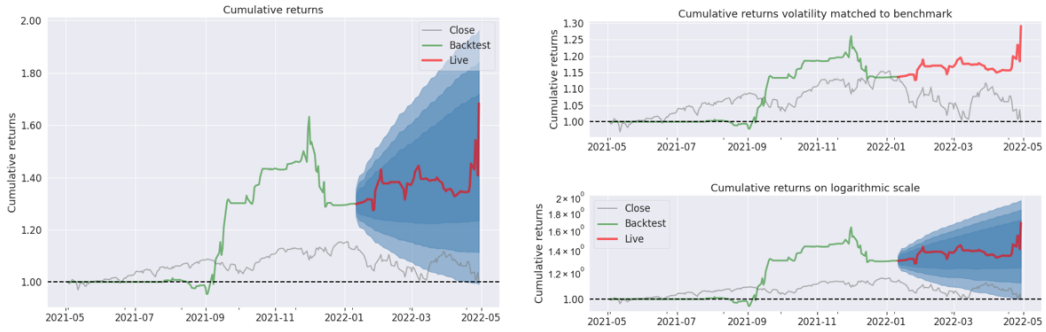
Looking at the technical indicators, we can conclude that the annual and cumulative returns are positive, even though the volatility is relatively high. The skewness and kurtosis indicate that the distribution of this strategy isn't normal. However, the alpha is positive, which means that the method can earn excess strategy profit based on the market profit.

	IN-SAMPLE	OUT-OF-SAMPLE	ALL
ANNUAL RETURN	45.6%	135.6%	68.3%
CUMULATIVE RETURNS	30.0%	29.5%	68.3%
ANNUAL VOLATILITY	28.9%	47.0%	35.3%
SHARPE RATIO	1.44	2.04	1.65
SKEW	1.13	3.65	3.07
KURTOSIS	9.47	22.81	25.63
DAILY VALUE AT RISK	-3.5%	-5.5%	-4.2%
ALPHA	0.78	1.05	0.80
BETA	-0.92	-0.65	-0.76

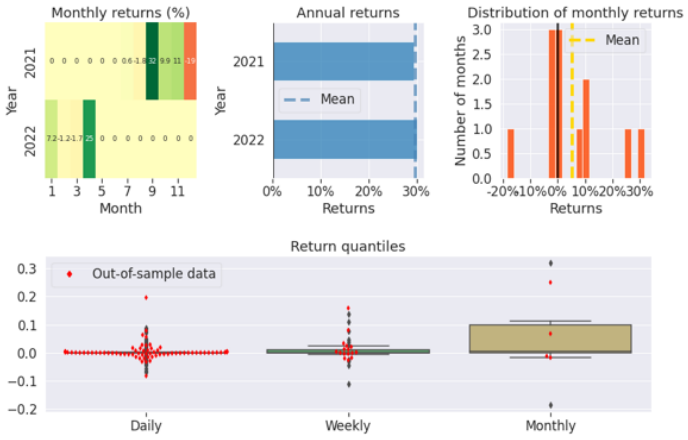
Second, the profit of one strategy is a vital part of strategy analysis. Cumulative returns return

volatility and distribution of returns discover more information about the strategy return. The results show that the return fluctuates around 0, ranging from -0.05 to 0.20. The return volatility has increased recently, especially from January 2022 to May 2022.

From the cumulative return plot, we can see that the cumulative returns keep increasing and reached their peak in May 2022. The return volatility also indicates the high volatility of the strategy compared to the result of the benchmark ETF, the S&P 500.

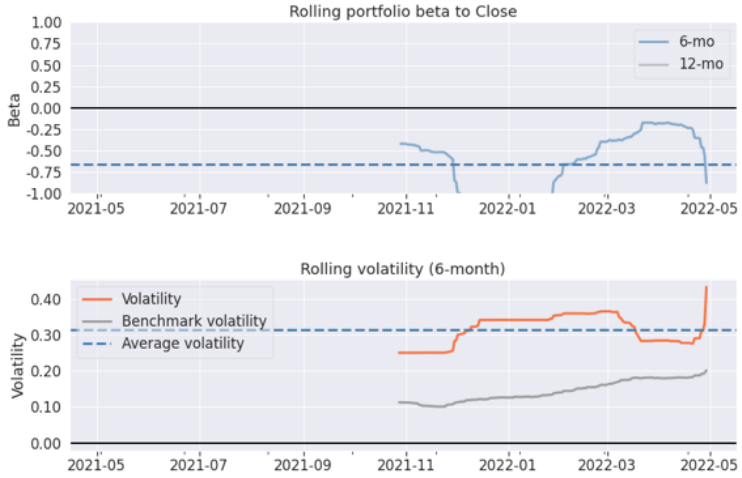


The distribution plots for the strategy present that the overall return is positively skewed, like the plots of the monthly returns, annual returns, and the return quantiles are shown below. Specifically, the recovery in 2021 and 2022 are positive in most months. However, some months have a negative return.



In the end, rolling beta, volatility, and Sharpe ratio help us better understand the strategy's risk. The plot of rolling beta shows that the market return from May 2021 to May 2022 is always negative, whether in the 6-month or 12-month window, indicating a bear market. Therefore, investors expect prices to decline. They prefer selling a security or a commodity and repurchasing it later at a lower price.

The rolling volatility plot in the 6-month window also increases the volatility. The strategy and benchmark volatility are higher than 0.1, and the strategy volatility reached 0.4 in May 2022. Thus, higher volatility is one of the essential conditions to earn excess profit.



The rolling Sharpe ratio plot also reflects the risk of earning excess profit. The Sharpe ratio equals the average return earned more than the risk-free rate per unit of the volatility of actual risk.

From the plot, we can conclude that the Sharpe ratio has decreased recently. Yet, the strategy's return isn't normally distributed, violating the Sharpe ratio assumption. Therefore, we should focus more on the other technical indicators, like skewness and kurtosis.

5.2 Trading Strategy using Price Ratio

First, the technical indicators, like annual returns, cumulative returns, annual volatility, Sharpe ratio, skew, kurtosis, daily value at risk, alpha, and beta, are calculated below, indicating the basic information of the strategy.

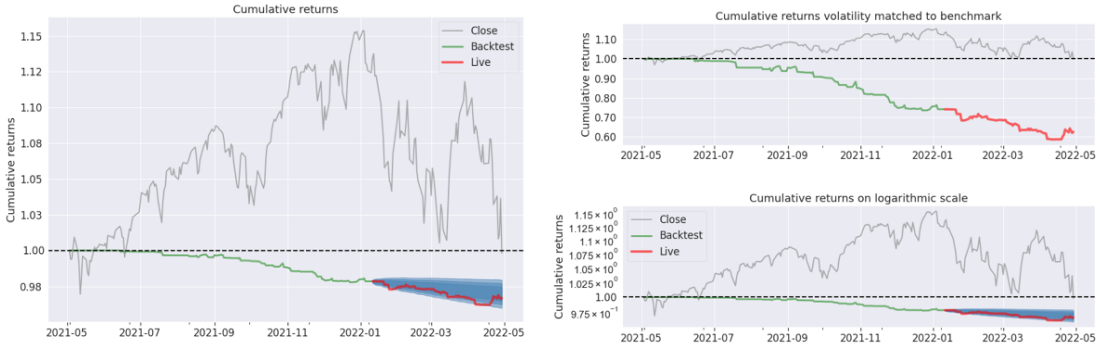
Looking at the technical indicators, we can conclude that the annual and cumulative returns are negative, even though the volatility is relatively low. The skewness and kurtosis indicate that the distribution of this strategy is close to normal distribution. However, the alpha is negative, which means that the method cannot earn excess profit based on the market profit.

	IN-SAMPLE	OUT-OF-SAMPLE	ALL
ANNUAL RETURN	-3.1%	-3.9%	-3.3%
CUMULATIVE RETURNS	-2.1%	-1.2%	-3.3%
ANNUAL VOLATILITY	0.9%	1.7%	1.2%
SHARPE RATIO	-3.36	-2.37	-2.81
SKEW	-1.91	-0.38	-0.96
KURTOSIS	5.65	2.76	5.71
DAILY VALUE AT RISK	-0.1%	-0.2%	-0.2%
ALPHA	-0.03	-0.04	0.03
BETA	0.01	-0.01	-0.00

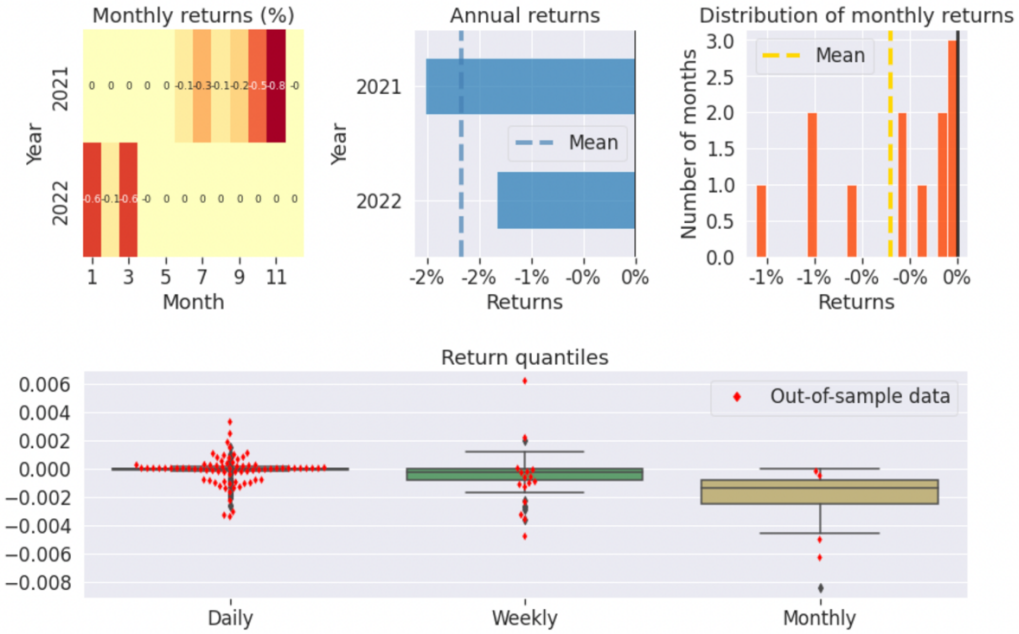
Second, cumulative returns, volatility, and distribution of returns discover more information

about the strategy return. The results show that the return fluctuates around 0, ranging from -0.05 to 0.20. The return volatility has increased recently, especially from January 2022 to May 2022. We can also conclude that the negative return is more than the positive return from the plot.

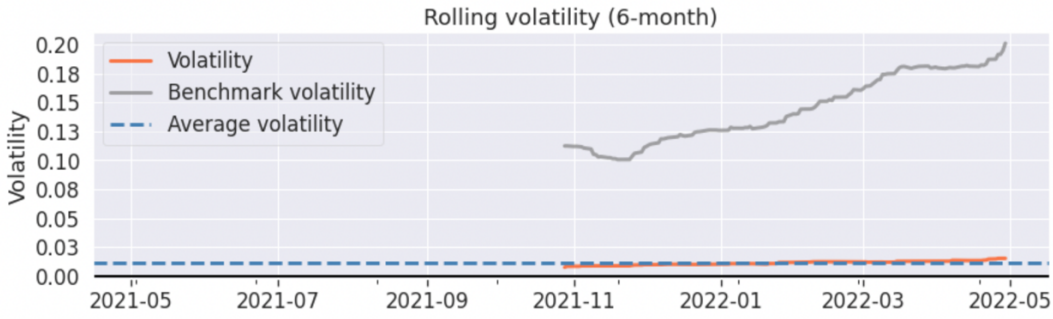
We can see from the cumulative return plot that the cumulative returns keep decreasing and become negative from July 2021 until now. Yet, the return volatility also decreases as the strategy's cumulative returns decrease compared to the benchmark ETF, the S&P 500.



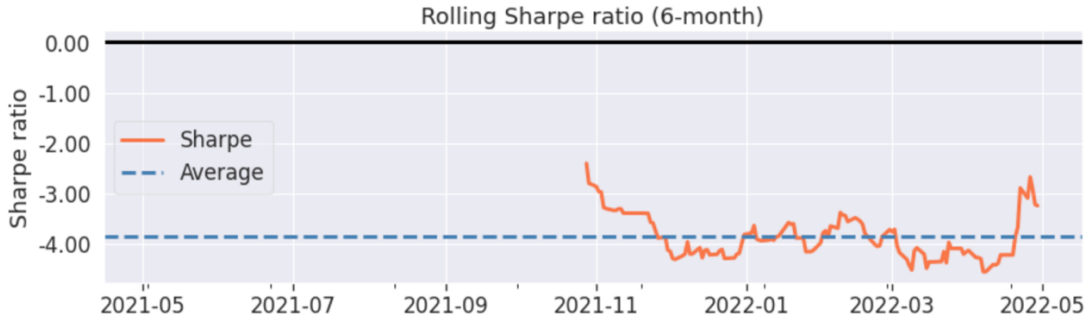
The distribution plots for the strategy present that the overall return is negatively skewed. Specifically, the recovery in 2021 and 2022 are negative in most months. Even though some months have positive returns, their positive returns are never more than 1%.



Finally, it's vital to analyze the risk of one strategy using risk indicators, like rolling volatility and Sharpe ratio. The rolling volatility plot in the 6-month window shows that the strategy's volatility is close to 0.02, indicating a stable approach. In contrast, the benchmark volatility increases from November 2021 to May 2022.



The rolling Sharpe ratio plot also reflects the risk of earning excess profit. From the plot, we can conclude that the Sharpe ratio is negative, indicating that its extra risk is lower than the market risk. However, the return of the strategy performs not very well. Therefore, the overall performance of the price ratio strategy is worse than the spread trading strategy.



6 Conclusion

After building an ETF pair trading strategy and backtesting these three ETFs, the result concludes that: Pair trading strategy has a positive return when using Spread. It performs relatively well compared with the strategy using a price ratio.

The annual return and cumulative return are positive. Specifically, the annual return is 68.3% though the volatility is 35.3%, which is relatively high. Yet, the annual return of strategy two using the price ratio is only -3.3%, showing a negative return. The alpha of the spread strategy is 0.8, indicating that its excess return on the investment is 0.8 after considering the market's volatility. The daily risk of strategy one is -4.2%, which means that it has a 5% probability of losing more risk than -4.2%. Plus alpha of the price ratio strategy is -0.04, which means that this strategy performs worse than the market, even though the value at risk is only -0.03%.

References

a) Chen, C. W., Wang, Z., Sriboonchitta, S., & Lee, S. (2017). Pair trading based on quantile forecasting of smooth transition GARCH models. The North American Journal of Economics and Finance, 39, 38-55.

b) DeClue, T. H. (2003). Pair programming and pair trading: effects on learning and motivation in a CS2 course. Journal of Computing Sciences in colleges, 18(5), 49-56.

c) Huck, N., & Afawubo, K. (2015). Pairs trading and selection methods: is cointegration superior?. Applied Economics, 47(6), 599-613.

d) Krauss, C. (2017). Statistical arbitrage pairs trading strategies: Review and outlook. Journal of Economic Surveys, 31(2), 513-545.

e) Lin, Y. X., McCRAE, M. I. C. H. A. E. L., & Gulati, C. (2006). Loss protection in pairs trading through minimum profit bounds: A cointegration approach. Advances in Decision Sciences, 2006.

f) Miao, G. J. (2014). A two-stage correlation and cointegration approach is highly frequent and dynamic pairs trading based on statistical arbitrage. *International Journal of Economics and Finance*, 6(3), 96-110.

g) Rad, H., Low, R. K. Y., & Faff, R. (2016). The profitability of pairs trading strategies: distance, cointegration, and copula methods. Quantitative Finance, 16(10), 1541-1558.

h) Ramos-Requena, J. P., Trinidad-Segovia, J. E., & Sánchez-Granero, M. A. (2017). Introducing Hurst exponent in pair trading. Physica A: statistical mechanics and its applications, 488, 39-45.

i) Wang, J., Rostoker, C., & Wagner, A. (2009, May). A high performance pair trading application. In 2009 IEEE International Symposium on Parallel & Distributed Processing (pp. 1-8). IEEE.