Trading Dislocations Between Fixed Income ETF Fundamentals and Fund Flow

Working Paper
Diego Alvarez
diego[dot]alvarez[at]colorado[dot]edu

February 2025

1 Introduction

This model seeks to find trading opportunities within fixed income ETFs by finding changes between flow and their underlying fundamentals. The idea is that dislocations between ETF flow and fundamentals exist due to a lead-lag (or lag-lead) relationship. This dislocation, while economically explainable, is robust and can be exploited for trading. While this model will sheds light on the lead-lag (or lag-lead) relationship, it does not assume a specific direction. Instead, it uses linear regression to *infer* trading signals from the dislocation. The model use 12 ETFs and generates substantial returns; $\sim$ 1.3 sharpe in-sample and $\sim$1.0 sharpe out-of-sample. Ideally, more fixed income ETFs would be included, but due to data constraints prevent further testing. All of the code for this model can be found on GitHub.[1]

2 Data

The data is collected from a group of data sources. The flow data is collected from a GitHub repo[2], which scrapes flow data from etf.com. The price data is collected from Yahoo Finance and uses the adjusted historical prices. The fixed income fundamentals were collected from Bloomberg Terminal. Below are the ETFs that are used.

Ticker	ETF
AGG	iShares Core US Aggregate Bond ETF
ANGL	VanEck Fallen Angel High Yield Bond ETF
VCSH	Vanguard Short-Term Corporate Bond Idx Fd ETF
IGSB	iShares 1-5 Year Investment Grade Corporate Bd ETF
LQD	iShares iBoxx $ Inv Grade Corporate Bond ETF
VCLT	Vanguard Long-Term Corporate Bond Idx Fund ETF
USHY	iShares Broad USD High Yield Corporate Bond ETF
EMB	iShares JPMorgan USD Emerging Markets Bond ETF
SJNK	SPDR Bloomberg Short Term High Yield Bond ETF
VCIT	Vanguard Intermediate-Term Corp Bond Idx Fund ETF
FALN	iShares Fallen Angels USD Bond ETF
HYG	iShares iBoxx $ High Yield Corporate Bond ETF
JNK	SPDR Bloomberg High Yield Bond ETF

[1] `https://github.com/diegodalvarez/FundamentalFlow`
[2] `https://github.com/yieldcurvemonkey/ETF_Fund_Flows`

3 Wisdom of Crowds vs. Economic Reflexivity

In this case, the primary driver of the model's returns is the dislocation between fundamental factors and market flow. There are two perspectives on this relationship.

1. Wisdom of crowds $\rightarrow$ The market accurately predicts fundamentals before they materialize and allocates capital accordingly

2. Reflexivity $\rightarrow$ Fundamentals materialize first, and the market allocates capital afterward

This model does not determine which perspective is more likely; instead, it focuses on identifying and trading the relationship. It is not particularly important whether the relationship follows reflexivity or the wisdom of crowds. Regardless of the underlying mechanism, each scenario creates a tradable dislocation.

4 Signal Generation

The model is constructed by analyzing the trend of the ETF flow in comparison with the trend derived from the *fundamental* pricing. The trend windows are calculated as the exponential moving average (EMA). In this case the *fundamentally-priced* trend is done using the bond ETF's attributes such as OAS, I-spread, or yield-adjusted which will be referred as *yield*.
The idea behind the model works under this framework,

1. If wisdom of crowds $\rightarrow$ the EMA of ETF flow leads ahead of the of EMA of yield (*fundamentally-priced*).

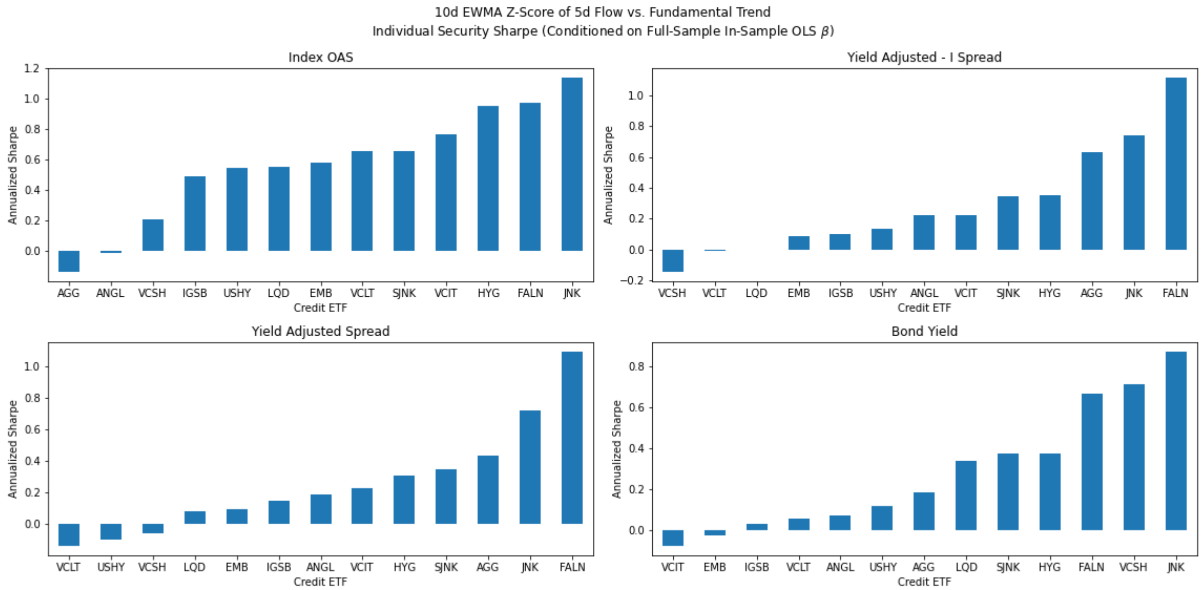
2. If Reflexivity $\rightarrow$ the EMA of yield (*fundamentally-priced*) leads ahead of EMA of ETF Flow

For simplicity, the look back window of the flow is set to 5 days, approximately 1 week, which makes it a medium-frequency trend. In this case, since two trend signals are being compared an additional step is needed to normalize the trend signals. A 10 day z-score is used to normalize both the flow and the yield trend. Next, the difference between the two trends is calculated. In this case, the yield trend is subtracted from flow trend. The final step is to run a regression of the signal against the returns. In this case, the choice of which trend gets subtracted is arbitrary since the signal is conditioned on the *sign* β of the OLS parameter. The β of the OLS parameter also tells which framework the model is trading under: wisdom of crowds or price reflexivity.

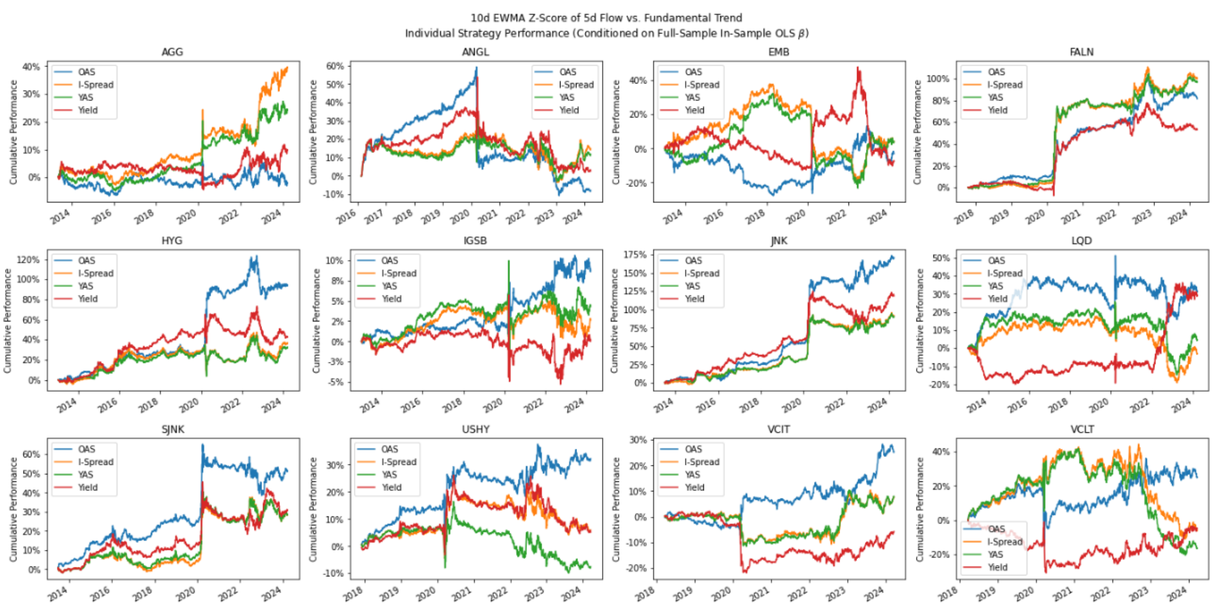
5 Signal Performance

5.1 Full-Sample In-Sample Performance

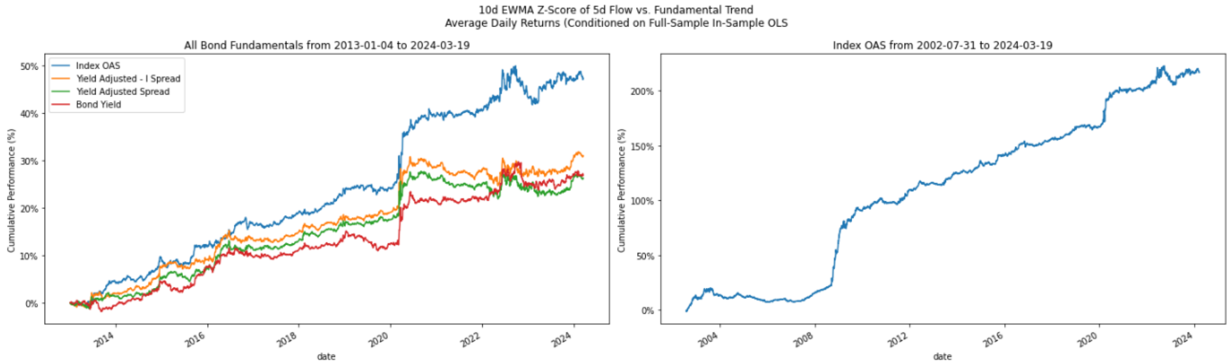
Start by calculating the sharpe for each individual ETF using the signal defined above. In this case, condition the signal on the sign of full-sample β. The results are reasonable.



Below is a plot showing the performance of each ETF based on each different fixed income attributes.

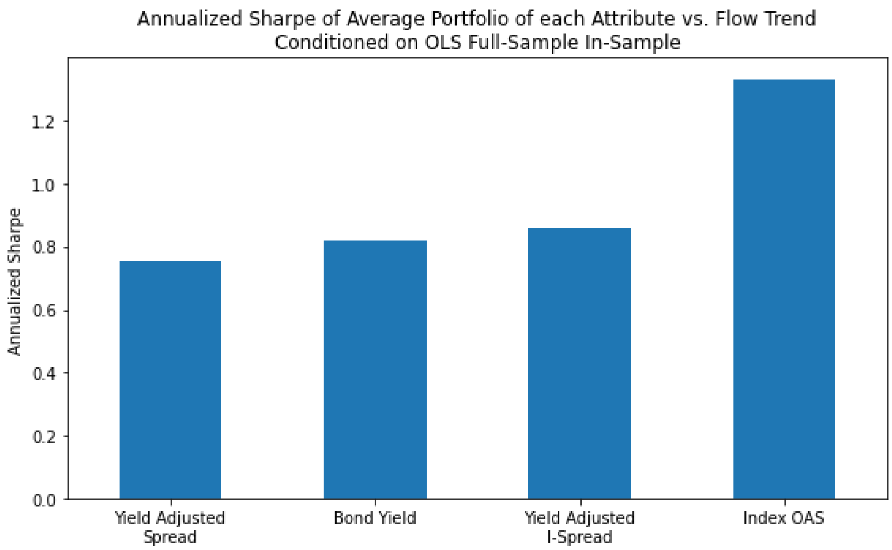


Although some securities have clear underperformance, they will still be included in the portfolio. In this case, equal weighted returns will be used instead of other portfolio optimization. When reporting portfolio performance, two methods will be presented: all fixed income attributes and the underlying index OAS solely. ETF yield data within Bloomberg Terminal goes back to 2014, while Index OAS goes back 2002, making it worth examining on its own.

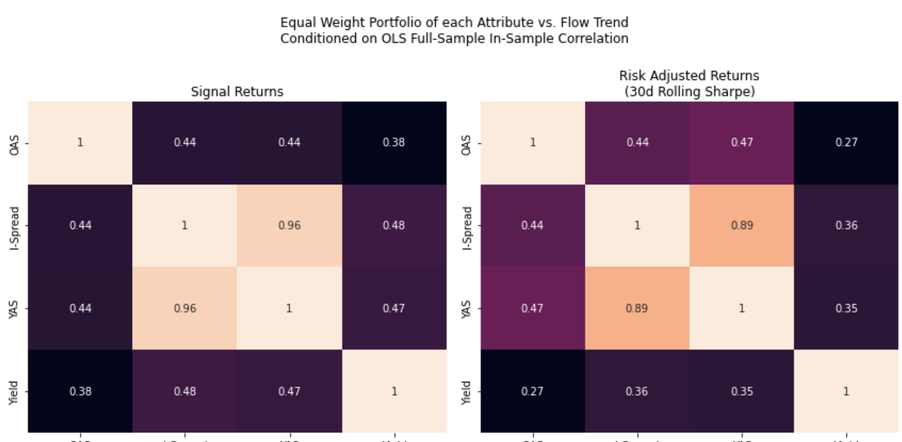


The returns are relatively stable and consistent across all yield measurement. The performance of the OAS reveals some interesting aspects, the performance has clear crisis-alpha attributes, as it generates returns during times of market stress such as in 2008 and 2020.

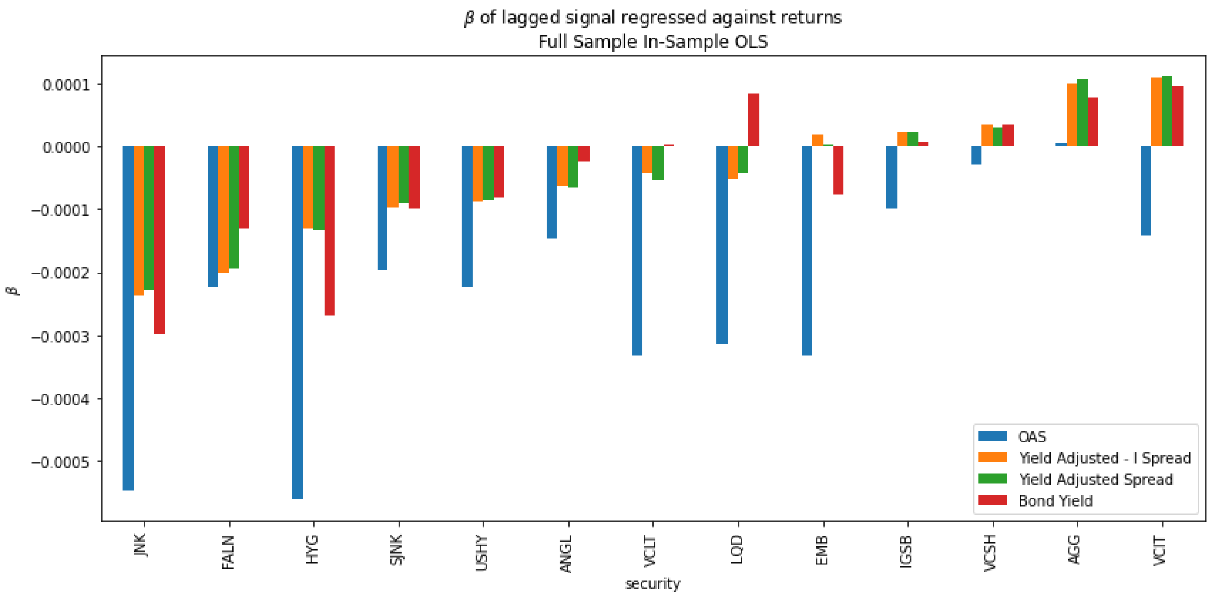
The sharpes of the portfolio are consistently positive.



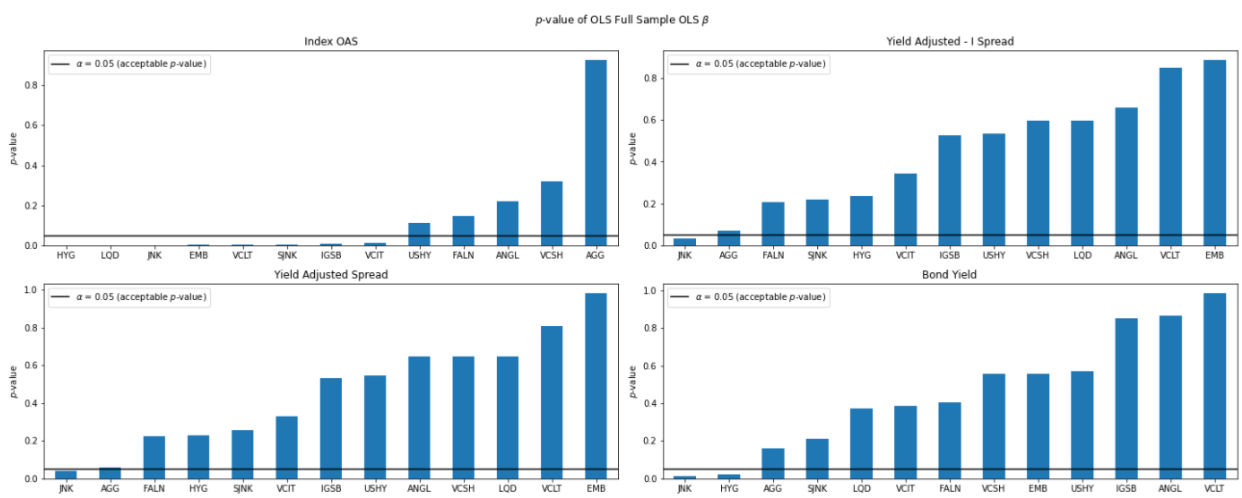
The correlation between each equal weight portfolio is reasonable low, which implies that they can all be blended into one large portfolio since they all exhibit positive sharpes.



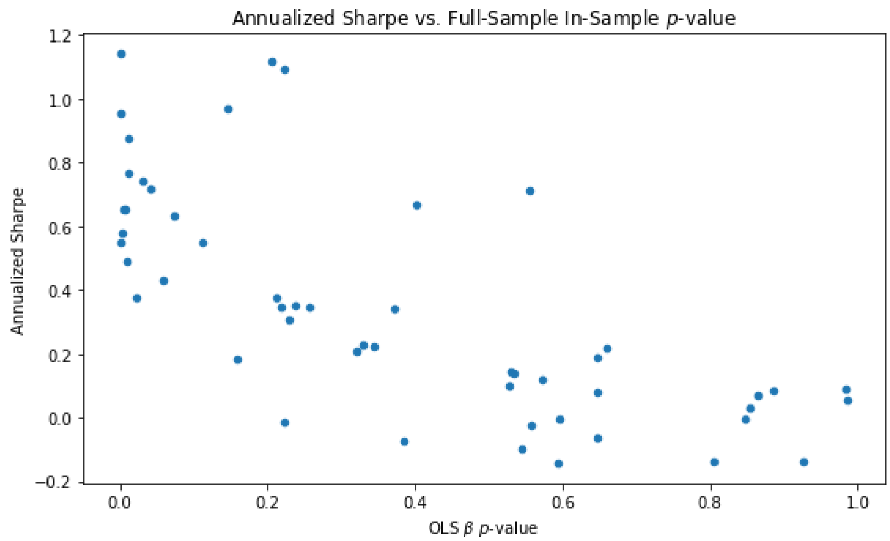
Below is a plot of the βs of the regression. In this case, they are mostly negative, which follows the reflexivity argument. This implies that the flow trend lags the fundamentals trend. What is notable is that OAS tends to be the best signal, which is a common theme in-sample and out-of-sample.



Below is a plot of p-values for the β coefficients, many of which fail to meet the threshold, although some pass, specifically OAS.



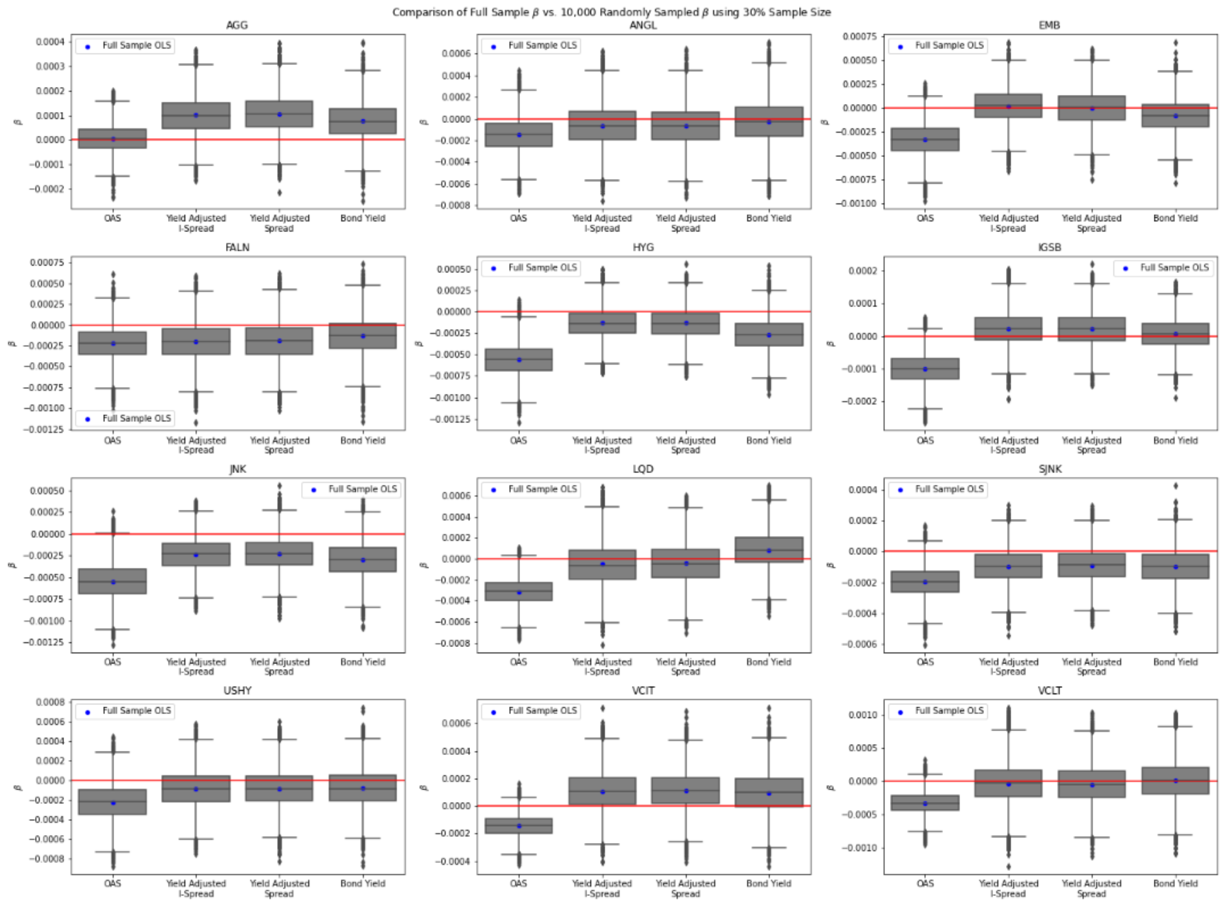
To ensure the sharpe is actually being generate by the regression below is a plot comparing the regression β p-value to the sharpe of that ETF. In this case, it is clear that that lower p-values which imply better regression lead to higher sharpes.



5.2 Bootstrapped Results

Expanding on the prior results is to use a sampled OLS rather than a full-sample OLS. In this case, using 50% sample size and 10,000 samples, calculate the OLS. The idea is that the 10,000 samples generate a distribution of βs which can be scrutinized. Although the distribution still incorporates the data from the full sample, it provides insight into the variation that can result from sampling. Below is a plot of the distribution of βs. The blue dots are the full-sampled βs. The goal is that the β coefficients should

1. Lie on one side of the spectrum either majority positive or majority negative
2. Lie on the same side as the full-sample β

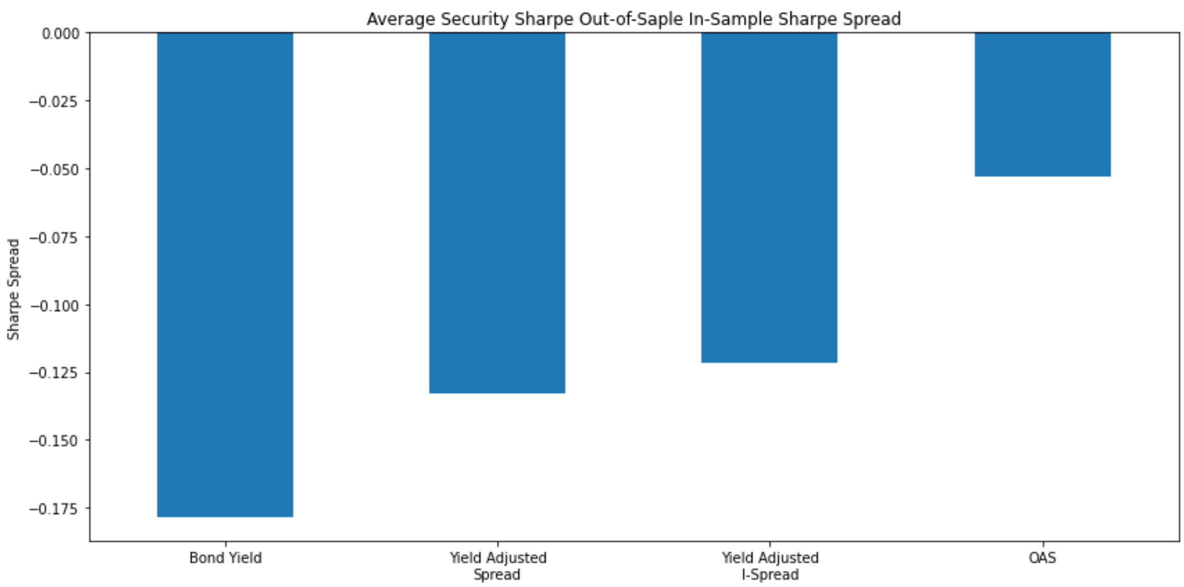


The results are not strong across the board, as many of the β distributions lie in between positive and negative values. From a categorical standpoint, OAS tends to be the best variable as its 25th (or 75th) percentile is above (or below) 0, respectively. This is a common theme, as OAS tends to be the best inputted statistic and covers the longest history.

6 Expanding out-of-sample

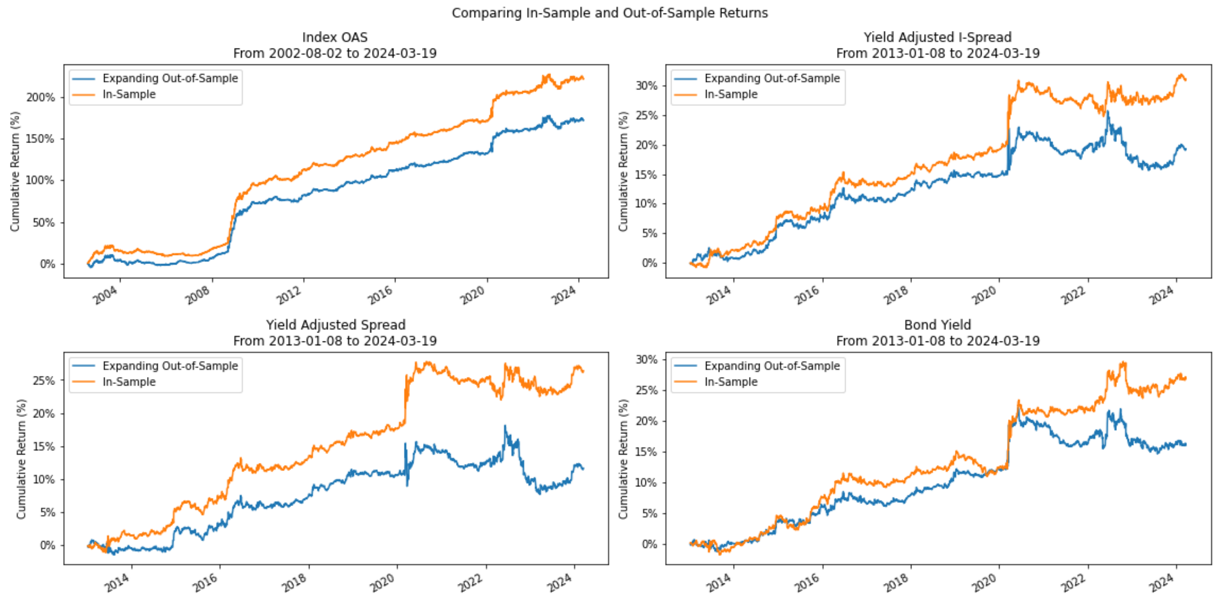
Now, apply the same approach to a rolling expanding OLS. The bootstrapped method provides some insights into how the out-of-sample results will perform. OAS is likely to be the best indicator and should experience the least sharpe reduction. Some ETFs, such as JNK which had βs distributed to one side should perform well across the board.

The out-of-sample results are in line with the bootstrapped OLS. OAS is the best performer across the board. ETFs that had consistent sharpes across variables, such as JNK perform well. Now that in-sample and out-of-sample sharpes have been calculated, calculate the sharpe lost going from in-sample to out-of-sample.



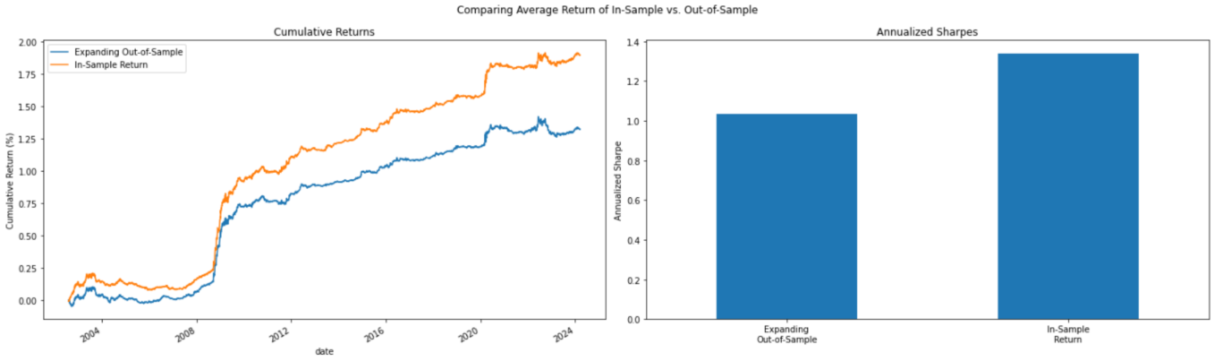
Surprisingly, the sharpe lost to going out-of-sample is quite low, with yield at approximately ∼0.175 and OAS being around ∼0.025 which is very low. Taking into account OAS has a longer history than the other statistics continues to make it the best performing variable.

Now, plot the returns of the in-sample versus the out-of-sample data. In this case, left for another task, using an equal weight portfolio across each ETF.



Below is the sharpe of each equal weight portfolio for each portfolio. Again, OAS tends to be the best performer with the least amount of sharpe reduction.

Blending these strategies using equal weights generates the portfolio return with a sharpe of around ∼1.4 in-sample and ∼1.1.



7 Possible Link to Equities

An attempt was made to expand this model to equities using a fundamental approach, such as Price-to-book trend or price trend vs flow trend, but it ultimately failed. The leading hypothesis is that within equities, the *fundamentals* can deviate far away from price which throws off the relationship.