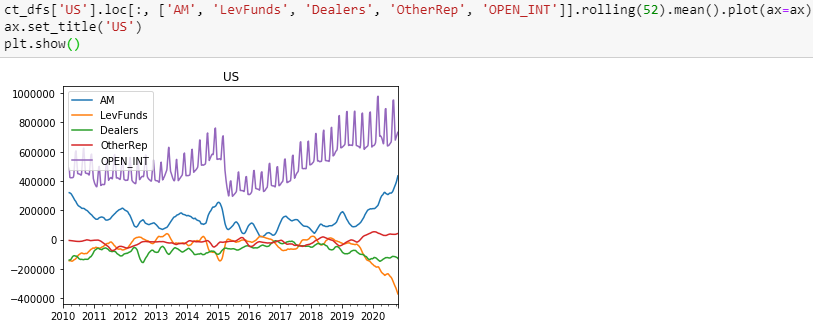
Brendan O’Neil 11/11/2020

## ExodusPoint Case Study – CFTC Bond Futures Data

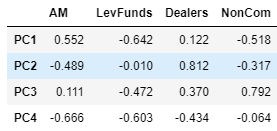
In this document I work through my process for analyzing a dataset for predictive features, constructing signals and backtesting them.

### Reading and Cleaning Data

First step is to read the data out of the excel file[[1]](#footnote-1) and inspect the data for oddities using basic timeseries plots and scatterplots. The only thing that stood out to me was a sharp decline in the open interest for the US contract in early 2015, it sounded vaguely familiar and I found the CME document explaining why[[2]](#footnote-2).



Next I consider the swap data, given what happens to the deliverable basket for US I decide to adjust it’s relevant swap data to be the 15y swap rate 2010 – March 2015, and then linearly interpolate the swap maturity from 20y starting March 16, 2015 down to the 15y rate at current (using weighting determined by that interpolation, and subtract the ‘roll’ from all subsequent datapoints to get a clean series)[[3]](#footnote-3). I use a simple average between the 20y and 30y rate for WNs and accept 2y, 5y and 7y are close enough for TU, FV and TY, respectively.

Next, I look at a PCA and correlation matrices of the weekly changes in positioning to get a sense for the relationships involved.

PC Loadings:

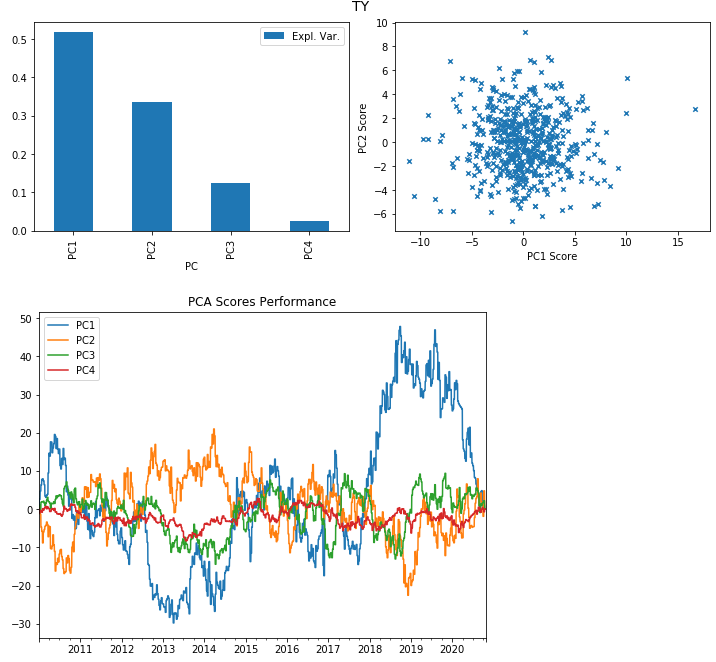
I consider Asset Managers, Leveraged Funds, Dealers and NonCommercial data since Other Reportables are relatively small and Commercial is close to just the other side of NonCommercial positioning. NonCommercial is considered as a proxy for speculators in general with respect to the CoT data.

With the TY contract, for example, PC1 and PC2 explain around 85% of the variance in positioning changes. One might interpret them as:

PC1: Asset Manager changes in positioning are offloaded to speculators

PC2: Buyside manager changes in positioning are taken down by dealers

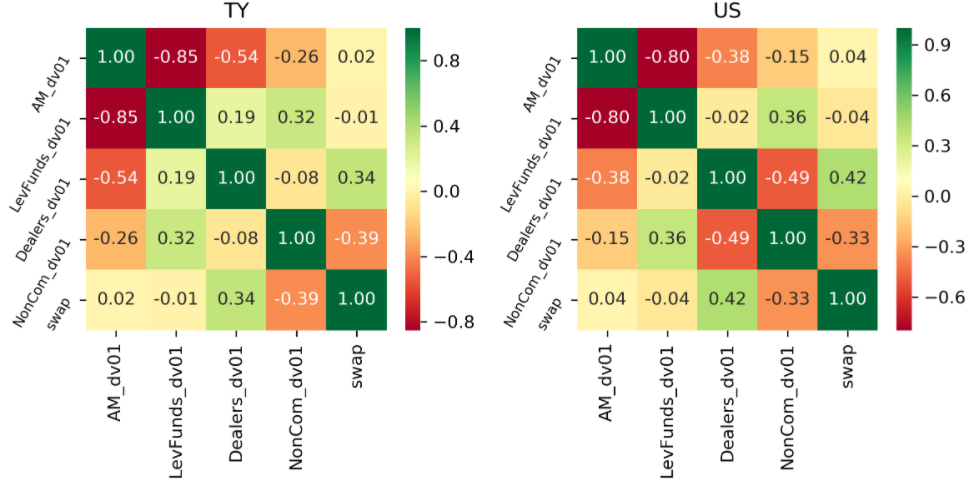
We will find later issues with collinearity in the data, even after removing some of the market participants, and techniques like principal components regression are commonly used to deal with this issue. In the interest of time and maintaining a better intuition in the data I decided to go another route.



### Initial Contemporaneous Analysis

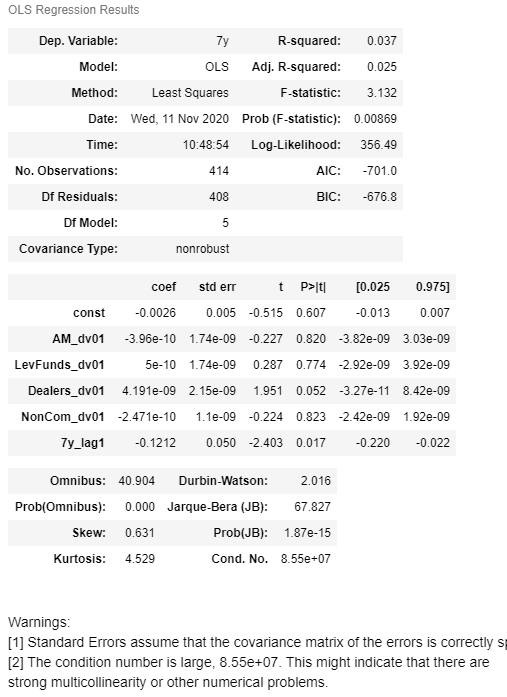
The first question I would like to answer is whether the CoT positioning data has significant explanatory value in analyzing contemporaneous changes in yields (as proxied by swaps). Before I start looking at the data more closely I remove all of the samples from 2018 onward- I will want an untouched out of sample period to test my in-sample findings on unseen data, about 25% of the data is reasonable.[[4]](#footnote-4)

For starters I want to normalize my data to some extent to make it more comparable across contracts. I try both weighting the positions by dv01[[5]](#footnote-5) and as a percentage of the average trailing open interest[[6]](#footnote-6). For the rest of the work I consider the dv01 weighted values[[7]](#footnote-7). Second, CoT data is collected on Tuesdays, but generally reported on Fridays. For this to be of value in predicting future changes I need to align the data with when it is reported, for simplicity I shift it 3 business days forward in time[[8]](#footnote-8). For all contracts looking at weekly changes in positioning, we get the relationships we might expect- changes in AM positioning are taken down by speculators and dealers

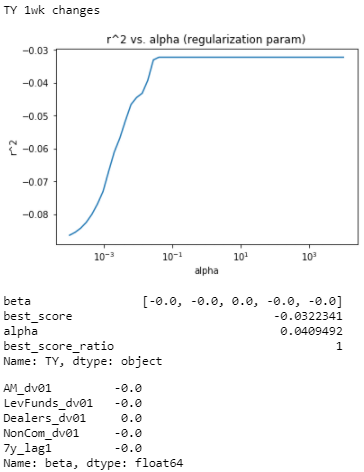


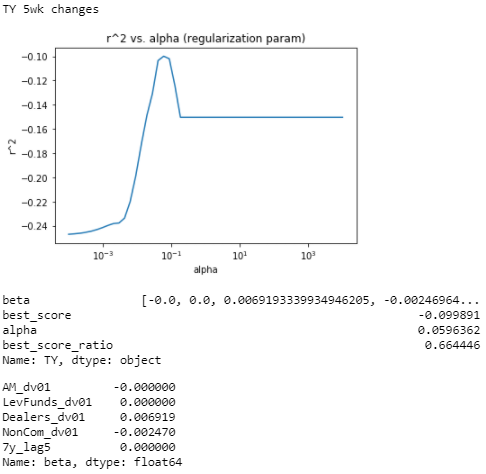
#### Ordinary Least Squares

I regress 1wk Friday changes in the relevant swap series for each contract against its corresponding futures contract positioning changes from AM, LevFunds, Dealers and the swap change lagged 1 week, to account for autocorrelation of changes in yield. As we see below, there might be some relationship with changes in dealer positioning and the lagged change, but there is clearly a lot of collinearity issues, so we can only conclude that the t-statistics are not robust and the betas are very unlikely to be numerically stable. As discussed before, we could try principal components regression to deal with this, but I prefer to use a regularization technique like Ridge or Lasso.



#### Lasso

I like to use Lasso in these cases. It is different from OLS in that it penalizes each beta in the objective function so that the betas are shrunk until their explanatory power outweighs the penalization. This helps significantly in cases such as this, where for example AM and LevFunds positioning are highly negatively correlated. Lasso also behaves like a simple feature selector in that it will shrink betas to 0 for data it does not find useful out of sample. How much penalization to use is evaluating using a time series cross validation process, where the model is fit on a rolling window and evaluated on a subsequent period of (somewhat) out of sample data. The penalization parameter that best fits the out of sample periods on average is selected. The penalization (regularization) parameter is called alpha, in the charts to the right increasing alpha means more penalization of betas. For this to work properly we first need to normalize our features to z-scores so that the penalization affects them equally. [[9]](#footnote-9)

 In the case of 1wk changes in TY positions and swap rates, nothing does better out of sample than just taking the simple average of swap changes over the test period. This is incidated in a negative r^2 and ‘best\_score\_ratio’ of 1 (i.e. the best score is equivalent to the ‘dummy’ that just guesses the fitting period average)

For 5wk changes Dealers and Specs are selected but the results are very poor.

1. Code in eptest\_main.py, primary function read\_ep\_data [↑](#footnote-ref-1)
2. https://www.cmegroup.com/trading/interest-rates/mar-15-jun-15-roll-analysis.html [↑](#footnote-ref-2)
3. Code in eptest\_main.py, function get\_interp\_swaps [↑](#footnote-ref-3)
4. Preferably I would also want samples from different ‘regimes’, e.g. in the beginning and middle of the dataset, but in the interest of time and since the data is weekly and only ~11 years post-GFC, I just reserve the end. I use another technique called cross-validation to ‘stretch’ the in-sample period to consider quasi-out of sample periods. [↑](#footnote-ref-4)
5. Contracts\*Dur\*px/100\*ct\_size/10e4 [↑](#footnote-ref-5)
6. To mollify the periods of open interest on the front contract that are after the roll is mostly complete [↑](#footnote-ref-6)
7. Obviously does not matter within contract at the same time but hopefully introduces some consistency over time. Also, I wanted to explore cross-contract dv01 relationships, but ended up settling on z-score changes anyway [↑](#footnote-ref-7)
8. We should consider the contemporaneous relationship on Tuesday’s close data to be precise [↑](#footnote-ref-8)
9. This code is housed in classification.py, main function run\_cv [↑](#footnote-ref-9)