

**Optional Project (6 points)**

**You may choose to undertake this project in lieu of, or in addition to, Assignment 1.**

**Due Saturday, April 19<sup>th</sup>, 11:59pm**

This exercise investigates whether the long and short positions of various types of futures market participants can predict futures returns. Denote the price of a futures contract in period  $t$  by  $F_t$ . Then the “unlevered” return from period  $t$  to  $t + 1$  is  $\frac{F_{t+1}}{F_t} - 1$ ; you can also use natural log return:  $\ln\left(\frac{F_{t+1}}{F_t}\right)$ . A period can be a minute, a day, a week, etc.

The CFTC publishes weekly “Commitment of Traders” report on its website:

<https://www.cftc.gov/MarketReports/CommitmentsofTraders/index.htm>.

The historical data are available here:

<https://www.cftc.gov/MarketReports/CommitmentsofTraders/HistoricalViewable/index.htm>.

There are six categories of commodities: Agriculture; Petroleum and Products; Natural Gas and Products; Electricity; Metals and Other; and Financial. For this exercise, you may focus on “Futures Only” data.

There are five categories of traders: (1) Producer, Merchant, Processor, or User; (2) Swap Dealer; (3) Managed Money; (4) Other Reportables; and (5) Non-Reportables. You should explore these pages to get more background information about these categories.

While this exercise is open-ended, here are some suggestions to get you started:

1. Pick a few futures contracts of interest to you. Find historical prices of these futures contracts. You can focus on the “front month” contract, that is, the contract that has the shortest expiration date. Note that once the front-month contract expires, the next month contract becomes the front month. For this reason, around the rollover date (when the front month contract expires), you can use the return of the second-month contract; alternatively, you can focus on futures returns when you know the front-month contract is not rolling over.
2. Line up the futures prices data with the positions data of various categories of traders. You can extract whichever position signal you want. Note that the raw positions tend to be numerically large. It might make more sense to convert raw positions into percentage changes in positions or “market share” of positions or something else that’s more well behaved. Think about how you use long positions and short positions.
3. Once the data are prepared, you can explore whether positions information as of week  $t$  can predict futures returns realized in week  $t + 1$ ,  $t + 2$ , etc.

4. You can explore if return predictability (or lack thereof) depends on how storable the underlying asset is.
5. You can adjust your test to the actual date of the release of the COT report.
6. You can also explore if the futures returns are predicted by the information contained in the term structure of the futures curve (rates + storage costs - convenience yield). How does that predictability (if any) vary across different product types? How does that predictability (if any) interact with the information contained in the positions?

The writeup should summarize your data source, empirical methodology, main findings, and any literature or references (including AI). Please attach the source code. The writeup will be graded on these dimensions, even if your result is that there is little return predictability.