

Option Data Set:

Option open interest is defining major tops and bottoms in the markets. The option writer is providing temper to the market; shaping momentum as they attempt to write for premium, (time and vol decay) while staying delta neutral. Their response to market movement is asymmetrical to their counter party who is long premium.

Option use has grown, and dormant open interest in the derivative market is very large. Episodic moments in markets can push markets into soaring rallies and tailspin breaks, as option writers are forced to buy at increasing numbers as markets move higher, or sell more as markets move lower to some terminal limit as defined by the option open interest, they are obligated too. These price points and behavior can set risk parameters and predictable behaviors when the trade is on or when the derivative option open interest is beyond definable thresholds.

Most of the open interest in the derivative markets comes into and out of markets within a 4 to 6-month parameter in forward futures, while options written expire weekly. In ag financial markets—markets that trade agricultural commodities bringing together speculators and natural hedgers—or financial markets, the larger expirations are quarterly and seasonally can build and ebb. All have characteristics which are at some level predictable. Open interest builds and must resolve itself while conforming to fundamental arbitrage into delivery (delivery of the underlying contract), but also resolving financial obligations that are specific to the derivative trading.

The basic tenant of our exploration of option open interest is that it is a very dynamic and highly leveraged component of the market. That derivative markets are defined by the open interest that exists already in the market. In other words, the open interest will mostly have to resolve upon itself.

Having delta and Black-scholes calculations can aid in the analysis. Things to consider in EDA:

1. look back at the data and see how many times the markets achieved the avg put price.
 - a. what happened, did the market stop going down once the average put price was achieved.
2. what was the largest delta position achieved in each episode and what was the put position relative to the total open interest.
3. at what price between neutral and the low did the delta change the most.
 1. what was the characteristic of the day's momentum at this price

It represents where the option writer should be most panicked or forced to hedge their delta neutral position

(panic days. sessions where higher opens or lower opens will follow through when you can look for double or triple the avg range)

4. Where does the position of the option writer change the most from neutral to 99% delta as a percent of their total hedge position.
5. In the above questions we are looking for peak momentum in the market and where it abates.
6. Look at the same for the average call price.
7. how often does the market move from the avg put or call back to neutral without going back to the other extreme.
8. how often does the market resolve back to neutral near expiration (within a week or 10 days)
(experience is that the trend moderates and pulls back into expiration, its often post panic and remember the average call or put is often defining where markets have discounted the now known and trade fundamental change)
9. If the market is nearing expiration after having run the avg call or avg put does it retrace into expiration, and if so, how far. For instance, back to our peak Delta. *(experience similar to the comment above.)*

Data explanation and primer:

Here's a bit of a quick primer for this data.

- All the data is option data of underlying futures of the respective commodities, highlighted in the files
- By being Option Futures data, it makes the data very reactive as market participants can really leverage the positions
- the relevant strikes are the 4 front months ... meaning ... if we are in October then the relevant data will be options data for futures expirations in October, November, December, Jan ... or if there's no open interest (meaning there's no demand for that specific option futures month) which ever the next expiration month that has the highest open interest (OI) which in my example will most likely be March.
- each data file denotes the day and the time when the market exchange pulls were made
- The column headers of all files are:
 1. Month/Strike
 2. Open day price

3. High day price
4. Low day price
5. Last day price
6. Settlement day price
7. percentage change
8. Estimated volume
9. Prior day Settlement day price
10. Prior Day Volume
11. Prior day Open Interest

Things to know:

- **OHLL** : Open, High, Low, Last is usually how market data gets quoted. it's the first price exchange registers when open, the highest traded (could be either bid or ask) the lowest traded price of the day, and the last traded price of the day

- **Settlement price** some times due to electronic exchange data and compliance and clearance, the last registered is not the official settlement price of the day, hence the distinction

- **Volume vs Open Interest OI**: It's not the same. *Volume* is how much was traded that day, *OI* is super important, usually included by smart quants in modeling and it denotes the open positions the market is registering up to that day. This open interest which denotes the institutional, the commercial and the speculative positions available for market structure that could potential get closed and throw a wrench in the market, for example given a Trump tweet.

- **any file 10-Year-T-Note_*.txt** has data from 10 year treasury bills. This is interest or yield data. This would be the mid-to-short end of the yield curve, and very sensitive to Trump's tweets.

- **any file *_ags_settlements.txt** has data from soybeans futures, australian wheat futures, black sea wheat futures, different commodity indices, general grain Option Futures data. These files are big. Super sensitive to trade war words Trump Tweets.

- **any file *_eonly_settlements.txt** has data from exchange indices options futures, so that will be S&P500, small caps, midcap, Dow Jones futures, Nikkei, Hang Seng, mini S&P futures which are used a lot for hedging institutional strategies so they will pick up idiosyncratic anomalies per Trump tweets. General market reaction to tweets, these files are big too.

- **any file *_int_settlements.txt** has yield data, different parts of the short end of the yield curve (3 years too included), you also have Eurodollar data (a super mean contract to trade which is mostly traded by sovereign funds which gives a clear perspective of international reaction to tweets), Swap data (mostly internationally used in very large contracts), FED Funds data that

give indication to what the funds reaction is, bond yield data. Anything interest yield related is here, and as the previous files these files are big too.

- **any file *_nymex_settlements.txt** this is the energy data because all the oil/petrol/gasoline(RBOBB)/platts/ethanol, basically all energy is traded in the NYMEX exchange. You have london and Singapore data here, you have crude oil, sweet oil, any energy producing oil data here, natural gas etc. So you have heard in the news “Energy markets are tanking because of this war or that tweet”, here’s where you are picking that behavior up.

- **ALL the other files started with the name of a commodity or future like Corn_*.txt or Kansas City Wheat_*.txt or SnP_*.txt or Soybean Meal_*.txt etc** are future options that have been the most reactive to any trade war /Trump anything tweet

- **Soybeans** mean the Soybeans futures contracts

- **Bean oil**, is the futures contract of the crushing process by-product of processing SOYBEANS, same with meal. Super sensitive to anything Trump/China tweet

- **Any dataset with Wheat** has been super sensitive to Trump Tweets too because of Russian and Australian Wheat production. Hence super elastic to Trump’s policies and tweets.

- **Corn** has been extremely sensitive to trump because the US has a strong production and you see movement with Trump tweets. Also because there’s an acreage shift in the US between Corn and Soybeans so Trade war tweets really affect these markets. Additionally to consider, Corn is related to energy with ethanol production.