498Captial – Project Goals

**Implementation Initiative**: Gasoline and Heating Oil Price Future Predictions

**Project Sponsor:** Donald Wedding, CEO

**Project Team: Tate Bolick, Andrius Markvaldas, Brandon Moretz, Joshua Wood**

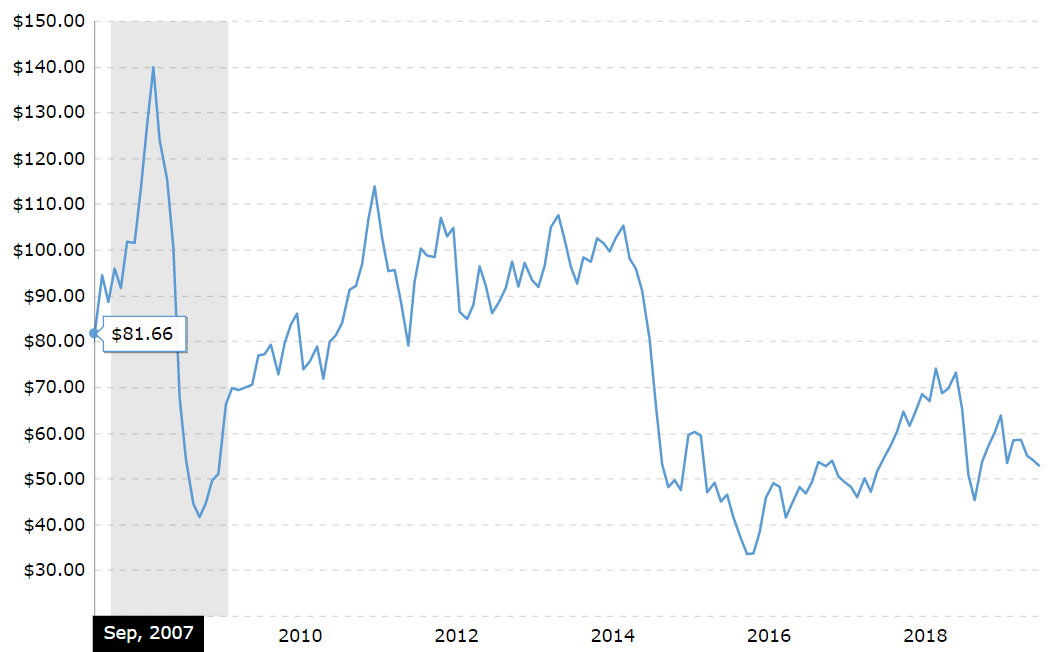


Project Background

Commodities are highly speculative, and offer massive potential upside along with the prospect of staggering losses. For example, Tampa-based Optionseller.com recently cited such high volatility in natural gas and crude oil trading that over $150,000,000 USD was wiped out in one day’s trading, causing all accounts to be liquidated (1). In many cases, commodities pricing is heavily linked to uncertain government policies, inaccurate data and producer reporting, poor research, and structurally flawed investment products (described below) of the environment. Combined, these challenges have led to a number of people moving assets into managed accounts, or passive investment vehicles (there has been an explosion in these types of structures on the equity side). These managed accounts are similar to investment vehicles, such as exchange traded funds (ETF’s) and exchange traded notes (ETN’s), and are, for derivative markets, structurally flawed. This is very different from an experience in the equity space. In the next section, we will offer a solution because this means that these current vehicles massively underperform the underlying commodities over time.

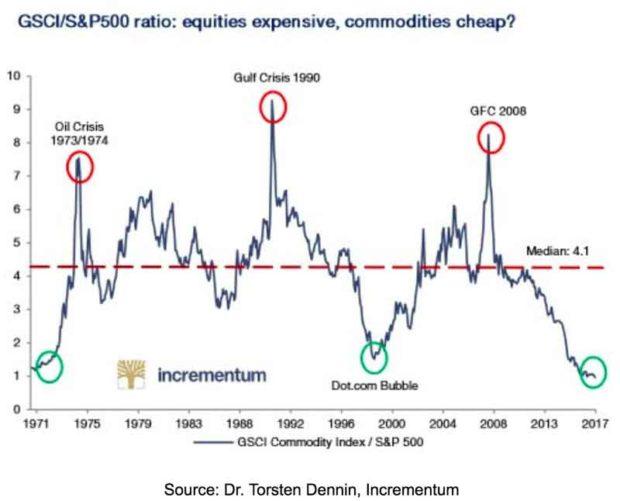
Despite all the aforementioned challenges, there seems to be renewed interest in commodity investing. Since peaking at $140 per barrel of West Texas Intermediate (WTI) in June of 2008, oil has seen a great deal of volatility, with general trend in the negative price direction (3). See *Figure 1* below for the price per barrel of WTI since Setpember 2007:

*Figure 1*



A recent example of such volatility occurred last month (September 2019) when there were attacks on Saudi oil fields. *However*, a lot of research likes to exhibit charts like the one in *Figure 2*, below. It shows the relative relationship between the levels of the S&P GSCI commodity index and the S&P 500 index of US stocks.

*Figure 2*



The idea is that when the line is ultra-low, such as now, it's a signal that commodities are cheap relative to stocks. The problem is that this is a ratio. It could just be signaling that stocks are expensive overall. (There are plenty of other flaws in it, but that's a whole project in itself.)

A popular way to invest in commodities is via some kind of index fund, such as an ETF or ETN. The commodity indices are all very different. Between different types of commodities, the supply and demand dynamics have nothing in common. However, speculative money flows can push all prices up or down, as investors pile into index funds, or dash for the exits. In the same day, one person makes a great case for why oil is heading higher while another uses solid reasoning to explain why it's going lower. Too often, commodities investing just seems like guesswork.

It's beyond the scope of this project to explain the full, technical details. However, the problem is that index-tracking ETFs and ETNs do it by way of the futures market, a type of financial derivative. A future is a contract to buy something at a later date at an agreed price.

Investment funds don't want to take delivery of physical commodities, so they roll over the futures contracts before they expire. This means they sell the contracts that are close to expiry and buy longer-dated ones. Futures prices bake in something called "basis" on top of the current "spot" price of the underlying asset. Basis is the difference between the cost of owning an underlying asset (e.g. lost interest on foregone cash deposits invested and any storage costs) less the benefits of owning something directly (e.g. dividends on stocks and coupons on bonds).

In the case of commodity futures, there's no foregone income yield from owning the future instead of the physical. So, basis mainly consists of the cost side, or "cost of carry." This is because investors in commodities have a choice. They can either own the physical commodity, paying out cash and taking on storage expenses. Or they can enter into a futures contract, keeping the cash and any interest on it, and avoiding the storage costs. The basis element of the future's price keeps the economic pros and cons of both situations in balance.

Basis shrinks over time, as a future approaches expiry. As a result, the price of the future gets closer to the spot price, being the price to buy the underlying asset straight away.

When the commodity funds roll over their futures positions, this means they sell something with little basis in the price and buy something with more basis in the price. This means the funds are constantly locking in small losses at each rollover.

It's actually more complex than that, involving things called "contango" and "backwardation." Therefore, funds that rely on futures contracts have a flawed structure that accumulates losses over time.

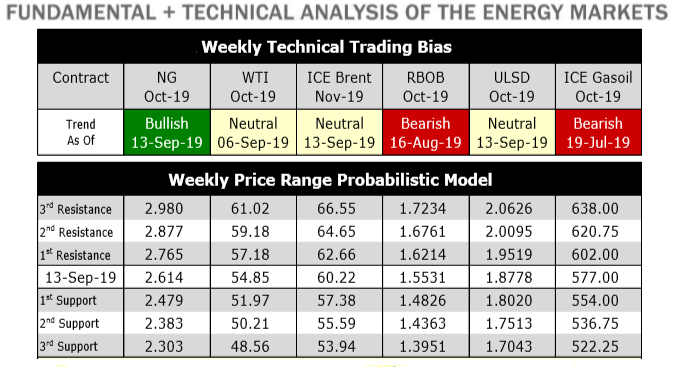
Project Objective

The main objective of the project is to model the price movements and forecasts of gasoline, and heating oil. We want to be able to answer the following questions:

1. Based on the data available, can we develop predictive models to predict future gasoline and heating oil prices?
2. What type of trading strategy do the models inform?
3. How well does the trading strategy perform versus relevant benchmarks/indices?

The project team will use a variety of cutting edge and sophisticated machine learning algorithms along with comprehensive technical and fundamental review of the energy markets. The first goal is to produce a simple technical chart forecasting the trading ranges of energy futures products. Given the time constraints, we will focus only on gasoline and heating oil futures. A dashboard will be created to not only give resistance guidelines on price, but trend indicators such as bullish or bearish. Below, in *Figure 3*, is a possible example of a future dashboard:

*Figure 3*



The next step will be to create a manageable number of different trading strategies with the inputs being the probabilistic models developed, and creating a back-testing harness to evaluate. The main goal is to deliver a strategy that is both quantifiable in terms of profitability, but also within risk guidelines.

Methodology

The initial workflow is outlined as follows, but is subject to change based on complexity and our initial findings.

1. Acquire the data
2. Preprocess the data
3. Process the data
4. Build the models – Linear regression and neural network (LSTM) seems to be popular on this type of data set.  Random forest to potentially classify the direction of the move in prices (bearish/bullish).
5. Test the model – train/test data sets
6. Build some type of back-testing system to test out a simple trading scenario.
7. Dashboard build
8. Build application

Scope

Functions in Scope

1. Exploratory data analysis
2. Model development in R, Python, and/or Excel
3. Model testing and validation
4. Trading strategy recommendation(s)
5. Trading strategy benchmarking versus relevant indices/benchmarks
6. Dashboard development using R
7. Application development using R

Functions Out of Scope

1. A plan for updating the base data over time (no API’s or data automation efforts)
2. Predicting the impact of potential future macroeconomic factors (such as changes in elected officials, potential military conflicts, etcetera)
3. Text analytics, such as scraping electronic news articles and/or websites to formulate a target price

Assumptions

1. It is possible to predict the future price of gasoline and heating oil with the data given
2. The project team has the required technology and available information to build a mobile application

Constraints:

1. Time: the CEO has given only 10 weeks for the project to finish
2. Money: we will not purchase any additional data from Bloomberg, Morningstar, MSCI, etcetera to supplement our current data
3. Human Resources: the project team is limited to four members, though collaboration with the CEO and cross-functional teams is possible, and expected

Dependencies:

1. Availability of resources across the necessary core team will impact the timeliness and quality of the final deliverables
2. The modeling is dependent on finding statistically significant relationships in the base data
3. The recommended trading strategies are dependent on finding meaningful outcomes from the predictive models

References

1. <https://www.bloomberg.com/news/articles/2018-11-19/hedge-fund-s-accounts-liquidated-amid-energy-market-volatility>
2. <https://seekingalpha.com/article/4174551-problems-commodity-funds>