TWO-PHASE PREDICTION OF AGRICULTURAL PRODUCTION AND PRICES

FINM 33150

Eric Chu (12334173), Jin Li (12283551), Luca Michelacci (12337696), Tony Zu (12334622)

Strategy Description

The Two-Phase Prediction of Agricultural Production and Prices strategy (the "Strategy") uses a two-step process to predict prices of agricultural commodity futures. The Strategy achieves this by leveraging a combination of quantitative methods along with the supply and demand relationships for each commodity

PHASE 1

- The first phase of the Strategy interpolates gaps in the United States Department of Agriculture ("USDA") production and consumption reporting data and also predicts supply and demand using various quantitative methods
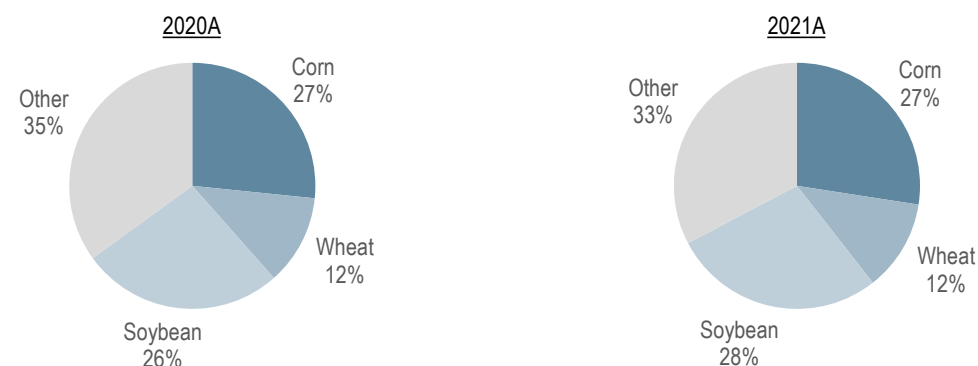
PHASE 2

- The second phase uses the predicted supply and demand curves to predict the prices of the respective commodity futures. Using the predicted prices, the Strategy enters into outright positions in front and back month contracts

- Based on the predictions above, the Strategy rebalances on a monthly basis
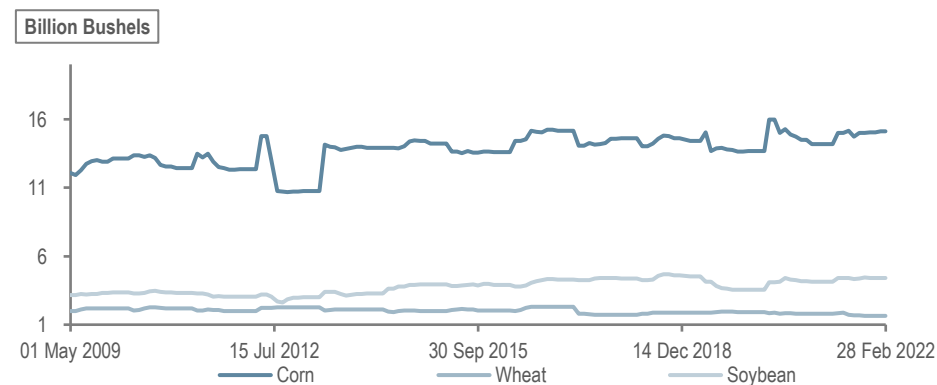
Strategy Motivation

MOTIVATION

- The total area planted in the United States is finite, therefore the amount of each crop planted is a function of the other crops

- Soybean, corn, and wheat have consistently been the most popular crops in the United States, typically accounting for 2/3 of the total crops planted. The production of these crops are related by the total area planted. This relationship can be leveraged to forecast the production and prices for the next growing cycle

- The supply of each crop is closely tied to the total area planted. The demand for each crop can be predicted based on past demand patterns. Given the supply and demand predictions, the Strategy forecasts future prices for the commodity

US Planted Acres Breakdown[1]

2020A

Corn 27%
Other 35%
Wheat 12%
Soybean 26%

2021A

Corn 27%
Other 33%
Wheat 12%
Soybean 28%

US Selected Crop Production[1]

Billion Bushels

| | 01 May 2009 | 15 Jul 2012 | 30 Sep 2015 | 14 Dec 2018 | 28 Feb 2022 |

Corn — Wheat — Soybean

Investment Universe and Securities

TRADED PRODUCTS

- The Strategy trades in a core basket of CME agricultural commodity futures, which account for ~67% of the area planted for crops in the United States and are the most liquid contracts traded on the exchange:

- Based on the predicted prices, the Strategy enters positions in both the front and back month futures contracts based on assumed mispricings

Futures Contracts

Product	Symbol	Contract Specs	Tick Size	Termination of Trading	Vol (2/25/2022)
Corn	ZC	5000 bushels	¼ cent per bushel ($12.5)	15th day of the contract month	346K / 160K
Soybean	ZS	5000 bushels	¼ cent per bushel ($12.5)	15th day of the contract month	234K / 87K
Wheat	ZW	5000 bushels	¼ cent per bushel ($12.5)	15th day of the contract month	112K / 67K

- The investment universe can be expanded to other commodity futures as shares of acreage grow in importance and/or liquidity becomes available. Other commodities to consider include:

  - Live Cattle (LE)

  - Canola (RS)

  - Cotton (CT)

Competitive Edge and Risks

COMPETITIVE EDGE

- The **Quantitative Methods** employed by the Strategy are the main competitive edge. The Strategy uses regression techniques and Bayesian methods to forecast crop production and prices

- Additionally, the commodities market prices can be modelled with **Fundamental Supply And Demand Dynamics**

- The Strategy executes at multiple delivery dates. This allows for scaling of the strategy

RISKS

The Strategy is exposed to the following risks:

1. **Weather Risk** – weather conditions outside of historical ranges could lead to supply and demand dynamics outside of the range of predictions

2. **Global Factors** – impediments to global trade (i.e., trade deals, trade wars, etc.) are not modeled in the Strategy. These events can lead to actual supply and demand dynamics outside of the range of predictions

3. **Storage Costs** – if the cost of storage changes drastically, within the contract month, the "carry" cost of the contract can lead actual prices to fall outside of the range of predictions

4. **Interest Rate Risk** – the Strategy assumes month-over-month changes in interest rate are small enough not to factor into price predictions. Extreme changes in month-over-month interest rates can invalidate these predictions.

5. **Other Environmental Factors** – the Strategy does not attempt to model unforeseeable circumstances (*force majeure* or acts of god). Circumstances such as global pandemics, biological hazards, etc., could lead to supply and demand dynamics outside of our range of predictions

Model Construction

PHASE 1 – PREDICTING SUPPLY AND DEMAND

The Strategy constructs supply and demand for each commodity using data points from USDA reports

- **Supply**

  1. Supply is constructed from ending stocks, total planted area, and production (yields and acres planted) from USDA reporting

  2. Monthly production yields are predicted from historical yields

  3. In most of the United States, the growing season is not year-round. Production in non-growing months is assumed to be 0.

  4. Supply models are adjusted for seasonality

- **Demand:**

  1. Demand is constructed using exponential weighting averages

  2. The Strategy applies the moving averages on quantity of crop consumed to create a forecast

PHASE 2 – PREDICTING PRICES

- The Strategy utilizes the constructed supply and demand from phase 2 to predict commodities' prices through fundamental economic relationships.

- The supply and demand curves are be related to price through regression and Bayesian methods

- The Strategy updates the prices based on the most recent data from USDA reports. The differences between the predicted and actual prices are traded on by executing positions in the front and back month contracts.

Model Construction – Methods and Constraints

METHODS

- **Preprocessing of data** – metrics are standardized prior to analysis

- **Data Interpolation** – data is forward-filled to obtain a point in time estimate of missing data points

- **Bayesian Methods –** exponential moving average was used to forecast supply and demand

- **Linear Regression** – ordinary least squares is used to relate supply and demand to price

CONSTRAINTS

- The supply schedule of a crop in a growing season is constrained by the total planted area. The total planted area is projected based on historical crop planted area, yields, historical precipitation, and temperature data.

- As a result of growing times, actual crop production supply schedule is rigid during the growing season..

- The model incorporates movements in expected crop supply schedule[1] only prior to a new growing season

(1) Refers to production

Trading Strategy

EXECUTION

- The Strategy executes on monthly intervals. Positions are opened after USDA report releases and closed prior to new ones using daily open and close prices

- The front and back month contracts are defined as the contracts with the most open interest (front-month) and the following month contract (back-month)

- Position sizing is proportional to the difference between the predicted and actual price as measured by standard deviation

RISK PARAMETERS / TRADING COSTS

- Margin maintenance requirements for the outright contracts are as follows [1]:

  - **ZC** - $1750

  - **ZS** - $3300

  - **ZW** - $2700

- CME Agricultural Product Fees [2] - $0.51 per contract per side

- Futures margin does not carry interest. However, any leverage on the portfolio is assumed to accrue interest at the retail rate [3]

- The Strategy trades at most 1% of daily volume for each contract

- Stop losses are considered at 50% of margin maintenance requirements for outright contracts

- The Strategy crosses the spread to take positions (assume 3 tick additional cost on open/close prices)

Back Test / Walk Forward Optimization

IN-SAMPLE

- A minimum period of 1 year trains the Strategy to optimize model parameters and increase prediction accuracy

- In-sample back test assesses the Strategy profit and loss as well as risk characteristics during the training period

OUT-OF-SAMPLE

- Out-of-sample back test assesses the performance (PnL & risk characteristics) of the Strategy in the year immediately following the training period

- Model parameters are not updated during out-of-sample back test

ADJUSTMENT

- When an in-sample and out-of-sample back test is completed, the back test periods is expanded by 1 year to train and predicts on the following testing period

Benchmarks

The Strategy uses both general benchmarks and sector-specific benchmarks. For general benchmarks, we chose to compare our Strategy to two of the most popular US equity indexes, the S&P 500 and the Dow Jones Industrial Average. For sector-specific benchmarks, we focused on funds that trade heavily in the three crops that our Strategy focuses on, using two popular US agriculture funds (DBA and RJA) and a constructed fund.

General Benchmarks	
SPDR S&P 500 ETF Trust (SPY)	Broad market equities
Dow Jones Industrial Average (DJI)	Dow Jones price-weighted measurement of 30 prominent companies
Sector Specific Benchmarks	
Invesco DB Agriculture Fund (DBA)	Futures based broad agriculture fund with weightings: Corn - 12.5% Soybean - 12.5% Wheat - 6.25%
Elements Rogers International Commodity Index – Agriculture Total Returns ETN (RJA)	Futures based broad agriculture fund with weightings: Corn - 13.61% Soybean - 8.6% Wheat - 20.06%
Constructed Even Weighted Portfolio	Teucrium Corn Fund (CORN) – futures based corn fund (33.3%), Teucrium Wheat Fund (WEAT) – futures based wheat fund (33.3%), Teucrium Soybean Fund (SOYB) – futures based soybean fund (33.3%)

Dataset Construction

BLOOMBERG

- Bloomberg provides historical times-series on crop production and planted acreage data

- The strategy relies on monthly historical data between Jan-2009 and Feb-2022 for each crop in our universe

- The key metrics imported were production, yield, ending stock and area harvested

QUANDL

- The Stevens Analytics Continuous Futures library provides a collection of long-term continuous price history for U.S. and international futures contracts traded on the CME

- For all commodities in the universe, historical time-series data on front and back month contracts are considered

USDA REPORTS

- The US Department of Agriculture publishes monthly reports[1] on crop production volumes and annual reports on total harvested land

- USDA time series data has been cross-matched with Bloomberg data for validation purposes

OTHER SOURCES

- The National Weather Service collects monthly temperature and precipitation averages for each US state for the period Jan-2009 to Jan-2022

- The strategy considers temperature and precipitation data for the major crop producing states – i.e., Iowa, North Dakota, Kansas, Illinois, Minnesota, Nebraska, South Dakota, Missouri, and Indiana[2]