Prime Trading

Project Lab - The University of Chicago
April 11th meeting

Futures Basis Model

Presented by

Armand Dang, Haoshu Wang, Kevin Xu,

Mooseok Kang, James Chen

1. **Theoretical value for basis**
    - Cost Carrying Model
    - Market Implied Pricing
    - Machine Learning Models

2. **Trading Strategy proposal**
    - Market basis prediction
    - Strategy Design

3. **Data Retrieval**

1. Theoretical value for basis

Cost Carrying Model: Equations

(1) Definition of Basis

$$\boxed{B_t = F_t^{\text{mkt}} - S_t}$$

(2) Cost-of-Carry Model (Fair-Value Futures Price)

$$\boxed{F_t^{\text{fair}} = S_t\, e^{(r-q)(T-t)}}$$

(3) Definition of Fair-Value Basis

$$\boxed{B_t^{\text{fair}} = F_t^{\text{fair}} - S_t = S_t\left(e^{(r-q)(T-t)} - 1\right)}$$

(4) Net Carry

$$\boxed{\text{Carry} = q - r}$$

Cost Carrying Model: Implementation

1. Data Acquisition and Cleaning:
 - Ingest the necessary daily market data (spot, futures, rates, and dividends)
 - Ensure that all datasets are aligned (e.g., same timestamps)

2. Daily Calculations and Modeling:
 - Potentially model interest rates (r) and dividend yield (q)
 - Calculate the observed basis
 - Compute the theoretical fair-value future price
 - Derive the fair-value basis

3. Validation:
 - Confirm that the inputs (r, q, and $T–t$) correctly reflect current market conditions
 - Compare calculated fair-value basis with observed basis

Market Implied Pricing

- Can apply some Volume-Weighted Average Price(VWAP) method on market data to find market-implied Theoretical Basis

- Apply a VWAP on tick data
    - Calculates the weighted average price of trades over a specified time window
    - Reflects actual demand and supply, based on executed trades
    - May contain noise due to odd-lot trades

- Apply a VWAP on quote data
    - Calculates the average price based on the available liquidity
    - Reflects the current state of the market, including how much liquidity is available at different price levels.
    - However, quotes can be manipulated (e.g., spoofing), and not all orders represent true trading intent
    - Requires more data storage and processing due to frequent updates and multiple levels

Machine Learning Models

- A machine learning model can capture nonlinear interactions between liquidity, volatility, and structural factors.
- Dynamically adapt to shifting market regimes. (e.g. when volatility rise, basis may widen)
- Use a wide array of real-time features to predict short-term deviations and convergence.
- Potential features: ETF/futures prices, interest rates, dividend yields, market microstructure data, volatility, volume, macro/sentiment
- Potential models: linear regression, random forest/XGBoost, Neural networks.
- Can be used in conjunction with the other models.

2. Trading Strategy Proposal

Literature

Literature

- It is a well known fact that trading on index products affects the index itself
- **How Index Futures And ETFs Affect Stock Return Correlations (2022)** shows us that ETF trading affects the correlation between index basket even more than index futures contract despite smaller volume

HOW INDEX FUTURES AND ETFS AFFECT STOCK RETURN CORRELATIONS

MARKUS LEIPPOLD[*] LUJING SU[†] ALEXANDRE ZIEGLER[‡]

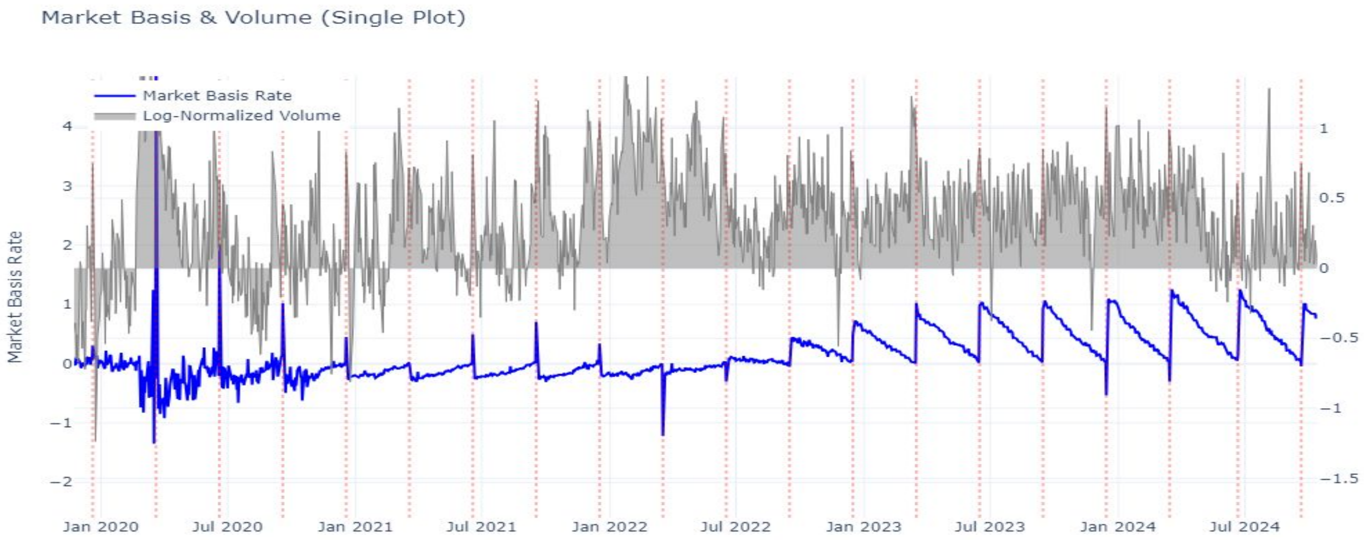
Motivation and Visualization

ES Market Basis

1. ETF Trades in dollar volume hit local maximum at futures expiration dates
2. Market basis rate change in daily level sounds far from lucrative strategy

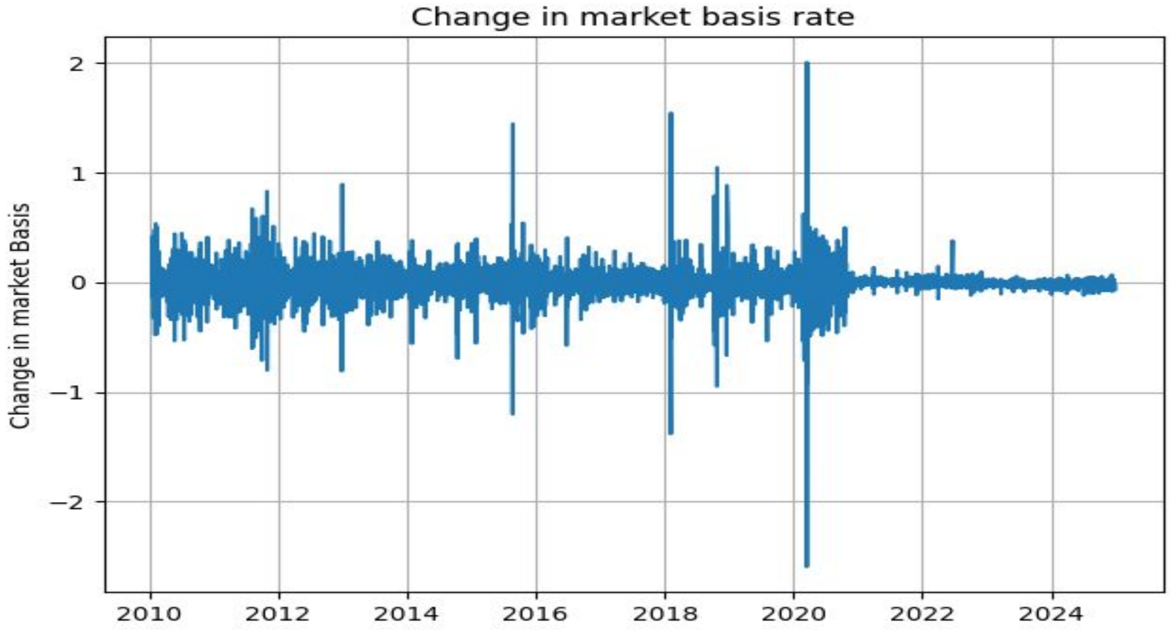
Motivation and Visualization

ETF Trades has some correlation with index futures trading activity



Motivation and Visualization

Daily or lower frequency doesn't seem to be an appropriate strategy for basis trading



	E-Mini SP500
Current Index	5400
Contract Multiplier	50
Minimum tick size	0.25
Tick Value	12.5
Tick Value in %	0.005%

Motivation and Visualization

Motivation

1. Literature and quick visualization tells us that ETF trades should be incorporated when pricing spot / futures
2. Daily level or lower is too long rather we should dig into intraday anomalies

Brief Research Plan

Time Series Modeling

1. **Construct VAR model**
   - Objective is to construct VAR model on market basis rate, ETF net trading DV or (ETF buy DV, ETF sell DV) respectively
2. **Sanity Check & Interpretation**
   - Stationarity
   - Granger Causality / IRF
3. **Prediction**
   - One Step / Multiple Step forward prediction
4. **Backtesting**
   - If the basis change is expected to broaden enter vice versa

Brief Research Plan

1. Construct VAR model

- Both considering the lag of itself and feedback effect

VAR(1) Model:

$$\begin{pmatrix} x_t \\ y_t \\ z_t \end{pmatrix} = \begin{pmatrix} \alpha_x \\ \alpha_y \\ \alpha_z \end{pmatrix} + \begin{pmatrix} \phi_{xx} & \phi_{xy} & \phi_{xz} \\ \phi_{yx} & \phi_{yy} & \phi_{yz} \\ \phi_{zx} & \phi_{zy} & \phi_{zz} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{x,t} \\ \varepsilon_{y,t} \\ \varepsilon_{z,t} \end{pmatrix}$$

- x:ETF Buy Dollar Volume, y: ETF Sell Dollar Volume, z: Basis Rate

Brief Research Plan

2. Sanity Check & Interpretation

- **Stationarity Check**
- **Portmanteau test** : See if the autocorrelation of errors are insignificant, which is assumption of VAR
- **Granger Causality Check** : Take out ETF variables and compare the prediction MSE to confirm that adding ETF variables add info
- **Impulse Response function:** How 1 stdev shock in one variable affects itself and others

Brief Research Plan

3. Prediction

Let $E_t(\mathbf{y}_{t+k})$ denote the k-step ahead forecast of $\mathbf{y}_{t+k}$ i.e. .

The k-step ahead forecast is given by:

$$E_t(\mathbf{y}_{t+k}) = \boldsymbol{\beta}_0 + \sum_{j=1}^{p} \boldsymbol{\beta}_j E_t(\mathbf{y}_{t+k-j})$$

$$E_t(\mathbf{y}_{t+k}) = \boldsymbol{\beta}_0 + \sum_{j=1}^{k-1} \boldsymbol{\beta}_j E_t(\mathbf{y}_{t+k-j}) + \sum_{j=k}^{p} \boldsymbol{\beta}_j \mathbf{y}_{t+k-j}$$

Prediction on Market Basis

Required Data

1. ETF related
   - SPX : SPY , IVV, VOO intraday trade data
   - NDX : QQQ
2. Futures related
   - ES E-mini, E-mini micro intraday trade data
   - NQ E-mini, E-mini micro intraday trade data
3. Indices related
   - SPX, NDX intraday price data
4. Investor Classification - (if available)

Strategy Design

- Compute theoretical basis with our model periodically (depending on the frequency of the strategy) and compare it with market basis.
- Consider factors such as transaction fees, slippage, and expected convergence timing to determine profitability and whether to enter into positions.
- Hold positions until basis converges to theoretical value, or when predefined stop-loss conditions are met.
- Potential stop-loss conditions include further deviation of the basis, updated theoretical value agree with the observed value, convergence hasn't occur after a certain period, etc.

3. Data Retrieval

Data Retrieval

- Cost of Carry Model:
  - Historical Data for model input: historical interest rates(overnight or 1-month rate), storage costs, and dividend yields(no need for total return index)
  - Spot and Futures trade data:
    - SPX or NDX ETFs intraday trade data
    - E-mini, E-mini micro trade data (1-min interval?)
  - Index Price data as reference
  - Window: at least one year, 3- 5 years
- Market-Implied Pricing:
  - Tick data or Quote data
  - Index Price data as reference
  - Window: a few months
  - Need to tackle with microstructure noise

Thank you!