Equity Statistical Arbitrage

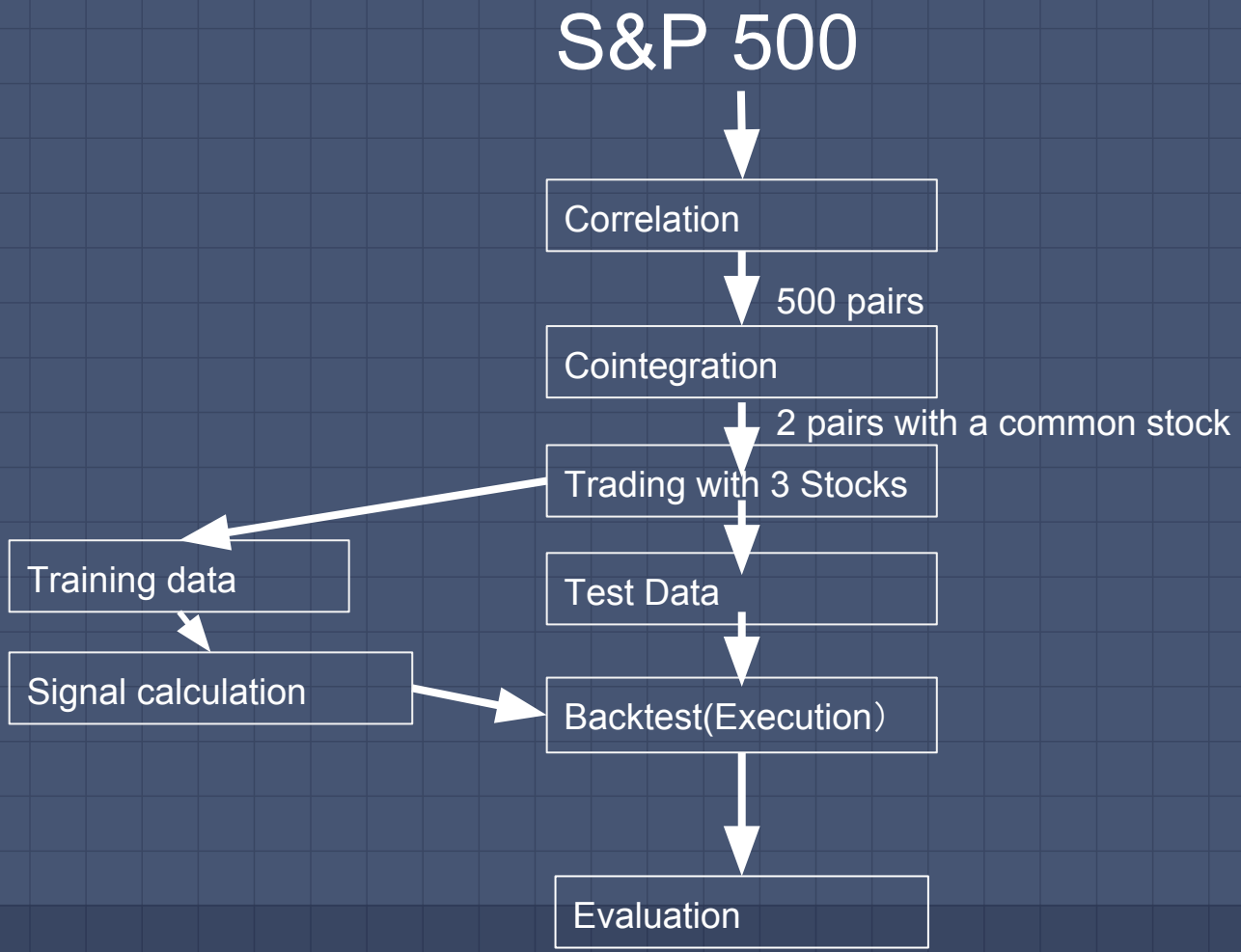
Presented by Yating Liu, Weiqi Tong, Yu Sheng, Yichen Song

Introduction to Statistical Arbitrage

- Statistical Arbitrage: Exploit relative price movements by analyzing price patterns and differences between instruments to generate alpha
- Economic Intuition: pricing inefficiencies between securities
- Not risk free（expected gain is greater than the risk）

Equity Statistical Arbitrage

- Medium or High -frequency strategy
- Trading signals are systematic, rules-based
- Trading book is market-neutral(zero beta)
- The mechanism for generating excess returns is statistical.

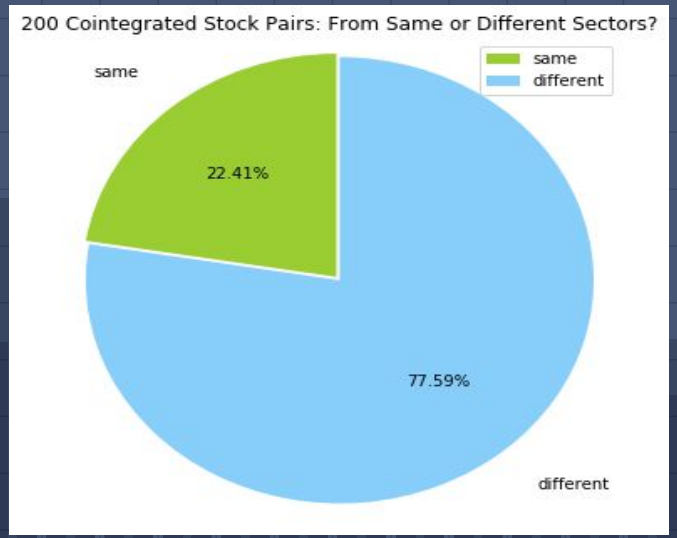

S&P 500
Correlation
500 pairs
Cointegration
2 pairs with a common stock
Trading with 3 Stocks
Training data
Signal calculation
Test Data
Backtest(Execution）
Evaluation

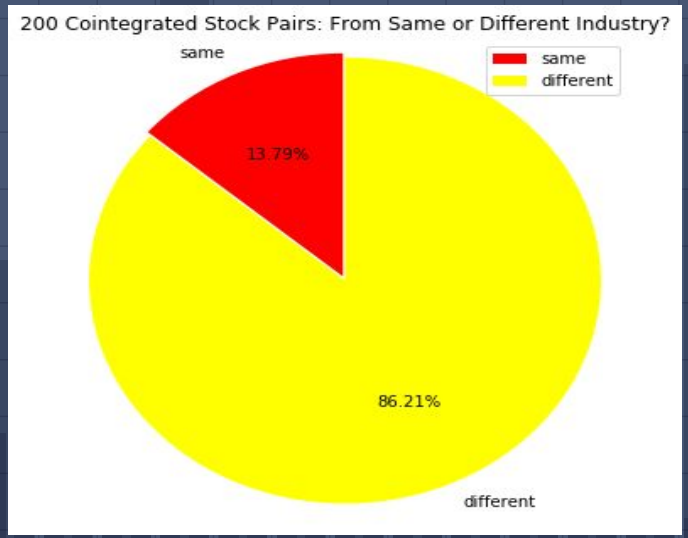
Cointegration Test

- Use method: Statsmodels.tsa.stattools.coint
- Test for no-cointegration of a univariate equation
- This uses the augmented Engle-Granger two-step cointegration test.
- H0: No cointegration exists v.s Ha: Cointegration exists
- Alternative: Johansen test

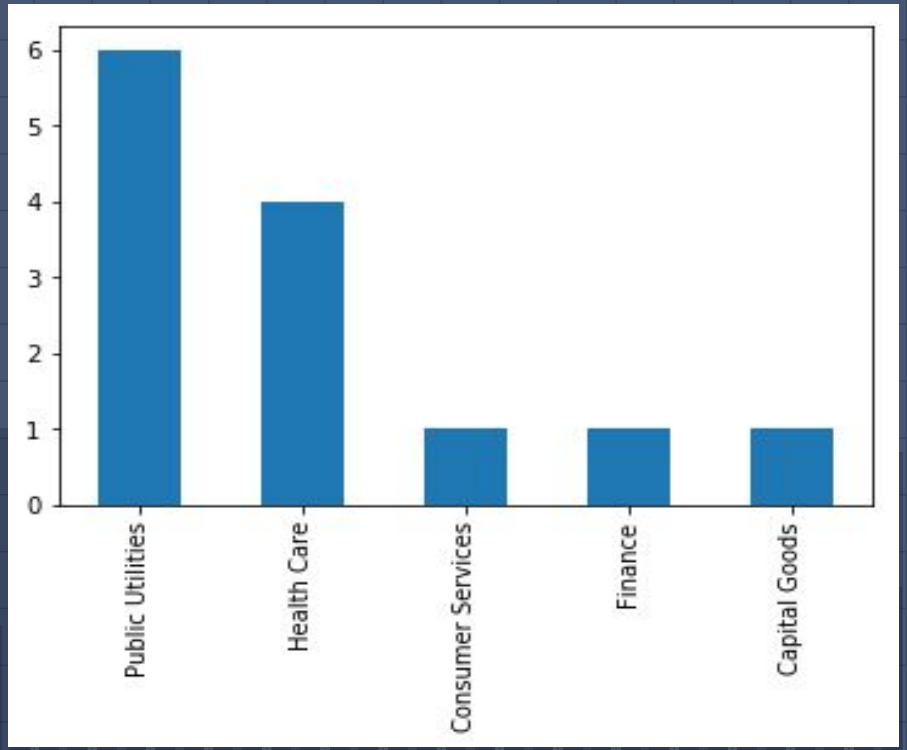
Insights from Cointegration

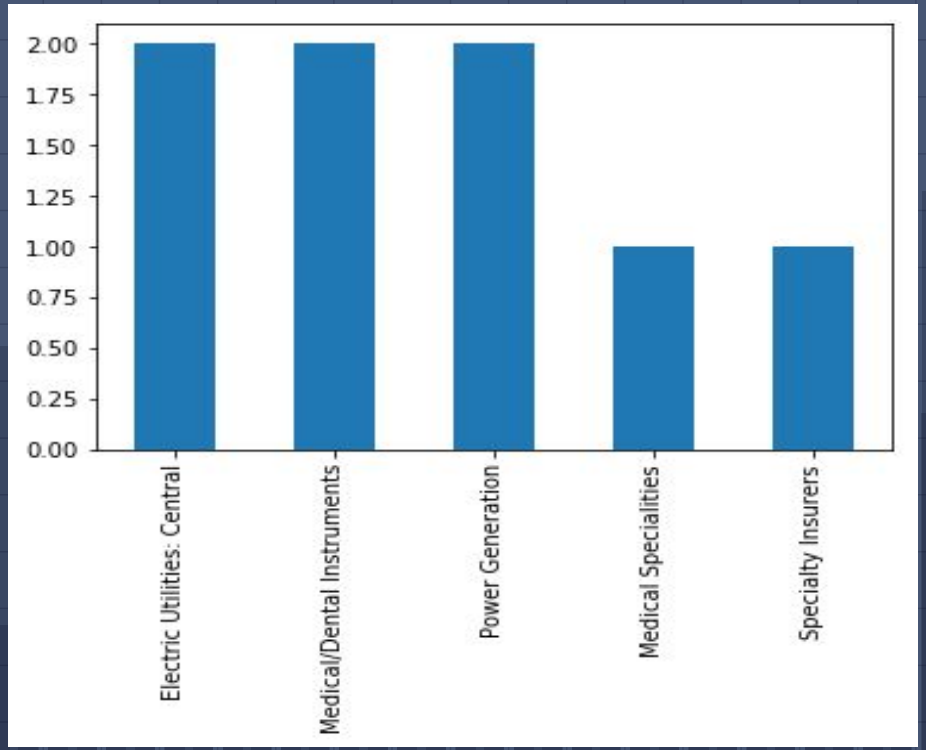
- Stocks that have related / similar business may tend to be cointegrated
- Sector (Healthcare, Energy, Finance…)
- Industry - more specialized (Major Banks, Property-Casualty Insurers…)





Close Look At Stock Pairs from Same Sector/Industry

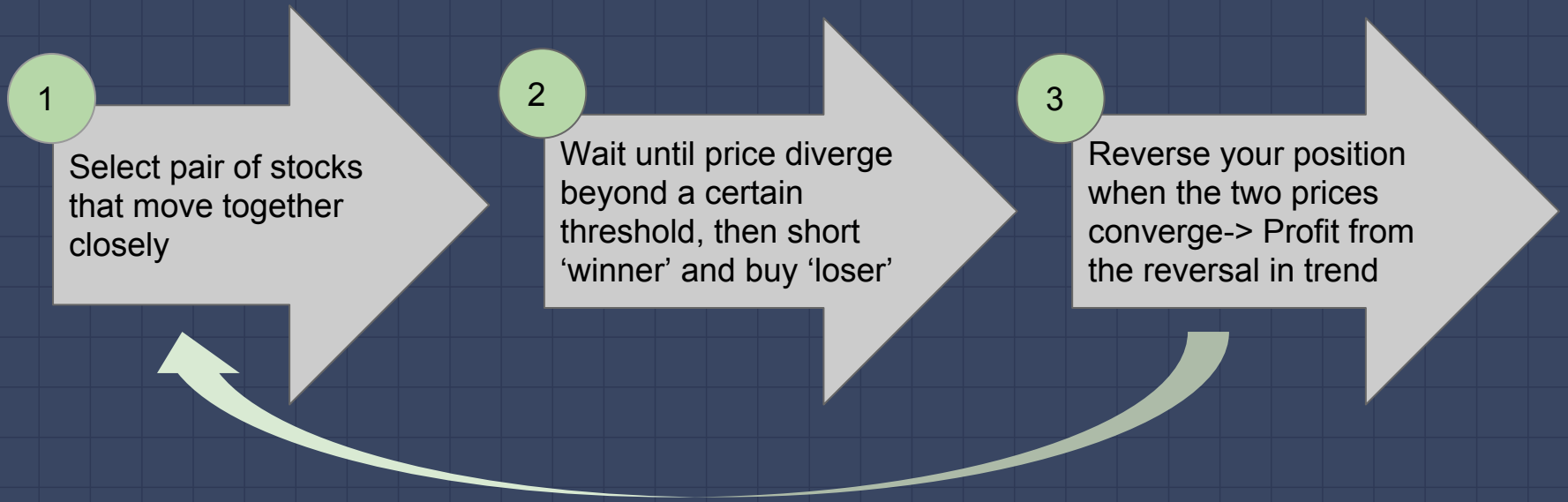




Example for Cointegrated Stocks

	stock1	stock2	name1	name2
0	COST	HD	Costco Wholesale Corporation	Home Depot, Inc. (The)
1	FIS	MMC	Fidelity National Information Services, Inc.	Marsh & McLennan Companies, Inc.
2	APH	FIS	Amphenol Corporation	Fidelity National Information Services, Inc.
3	AWK	WEC	American Water Works	WEC Energy Group, Inc.
4	AEP	WEC	American Electric Power Company, Inc.	WEC Energy Group, Inc.
5	AGN	CAH	Allergan plc.	Cardinal Health, Inc.

Pair Trading



Example of pair trading

Signal

- Residual from cointegration regression
- run cointegration regression on training dataset

$$Y_t = b_0 + b_1 \times X_{1t} + b_2 \times X_{2t}$$

- calculate signal for backtesting data: the residual of fitting the stock price with regression

$$e_t = Y_t - \hat{Y}_t = e_t - (b_0 + b_1 \times X_{1t} + b_2 \times X_{2t})$$

Multi-names Portfolio

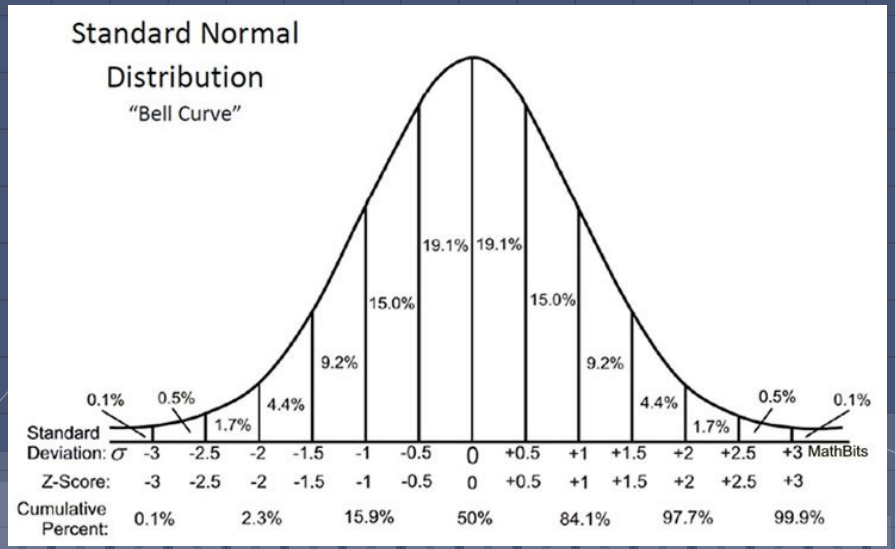
- Large diversified trading universe of US equities
- Select AGN(Allergen plc), CAH(Cardinal Health Inc), CVS(CVS Health Corp)
- Calculate signals every day, and open trade when signal exceeds a threshold(based on z score)
- Monitor for closing trade
- Keep the trade market neutral

Implementation

- Data from yahoo finance use pandas_datareader
- Training data: 2010-01-01 to 2017-01-01
- Test data: 2017-01-01 to 2018-01-01
- Calculate correlation, perform cointegration test and choose 2 pairs with a common stock, which constitute our portfolio
- Calculate signal
- Backtest
  - entry point:
    - $signal_y$>mean+z*std && $signal_t$>mean+z*std
    - OR $signal_y$<mean-z*std && $signal_t$<mean-z*std
  - exit point: signal crosses over mean
  - z score for each pair: go through a list of z scores to check which one gives the best result

Evaluation

- Transaction cost: 15 bps
- Trading with AGN(Allergen plc), CAH(Cardinal Health Inc), CVS(CVS Health Corp) and z-score=0.9 gives the highest return



Return

- Sharpe ratio: 1.05
- Annualized return: 19.42%

Risk - Our Implementation

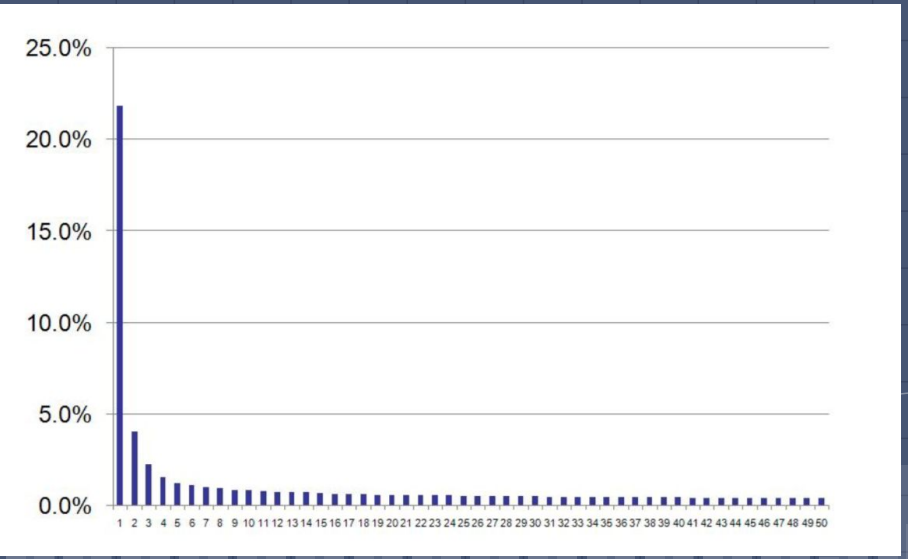
- Volatility: 4.65%
- Maximum drawdown: 6.82%
- Model risk: parameter z-score

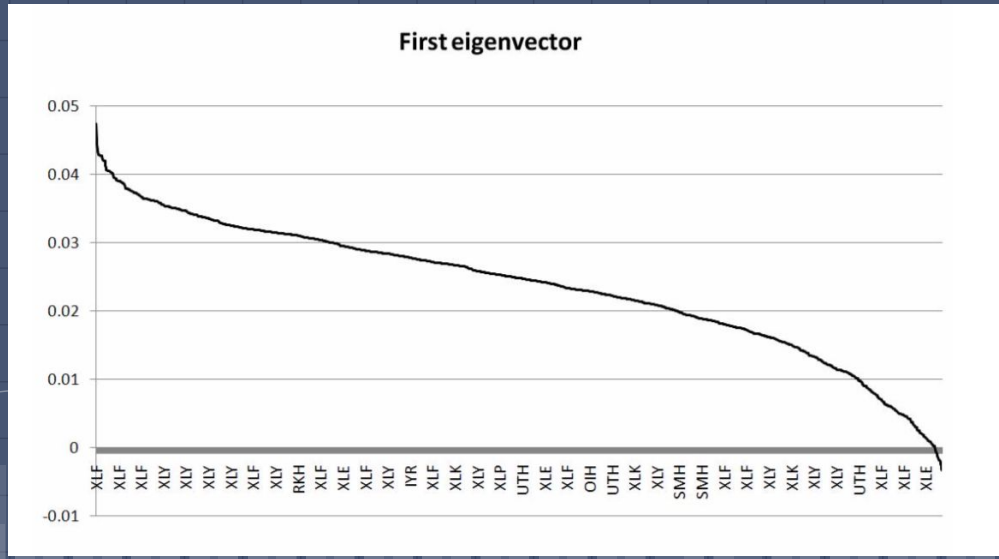
Risk - Pair trading in general

- Heavily depends on mean reversion of price movements, which may not happen in some cases (LTCM)
- Large positions are needed to generate profits from small price movements
- Leg risk: liquidity problem in execution
  - a well-designed pair may instead become a risky, one-sided bet caused by its un-hedged positions
  - can conduct passive trading for liquid stocks, aggressive orders for illiquid ones

Extension - Eigen portfolio

- Construct eigen portfolio to stocks within same sector, pair trade eigen portfolio with a stock in the sector.





Evaluation-Operation

- Data: yahoo finance
- Accessibility: can easily access equity market
- Complexity: Not complex, but require some time to select best portfolio from large universe
- Operational risk: difficult to trade several stocks at the same time

Evaluation-Correlation

- Macro factor
  - Uncorrelated to market trend

- Other strategies
  - Low correlation to other strategies

Evaluation-Competition, Evolution, Marketability

- Competition:
  - Crowded
- Evolution:
  - Potential for complex algorithms: machine learning
- Marketability:
  - Easy to explain

Thank You

Q&A