Statistical Arbitrage

Neelesh Upadhye

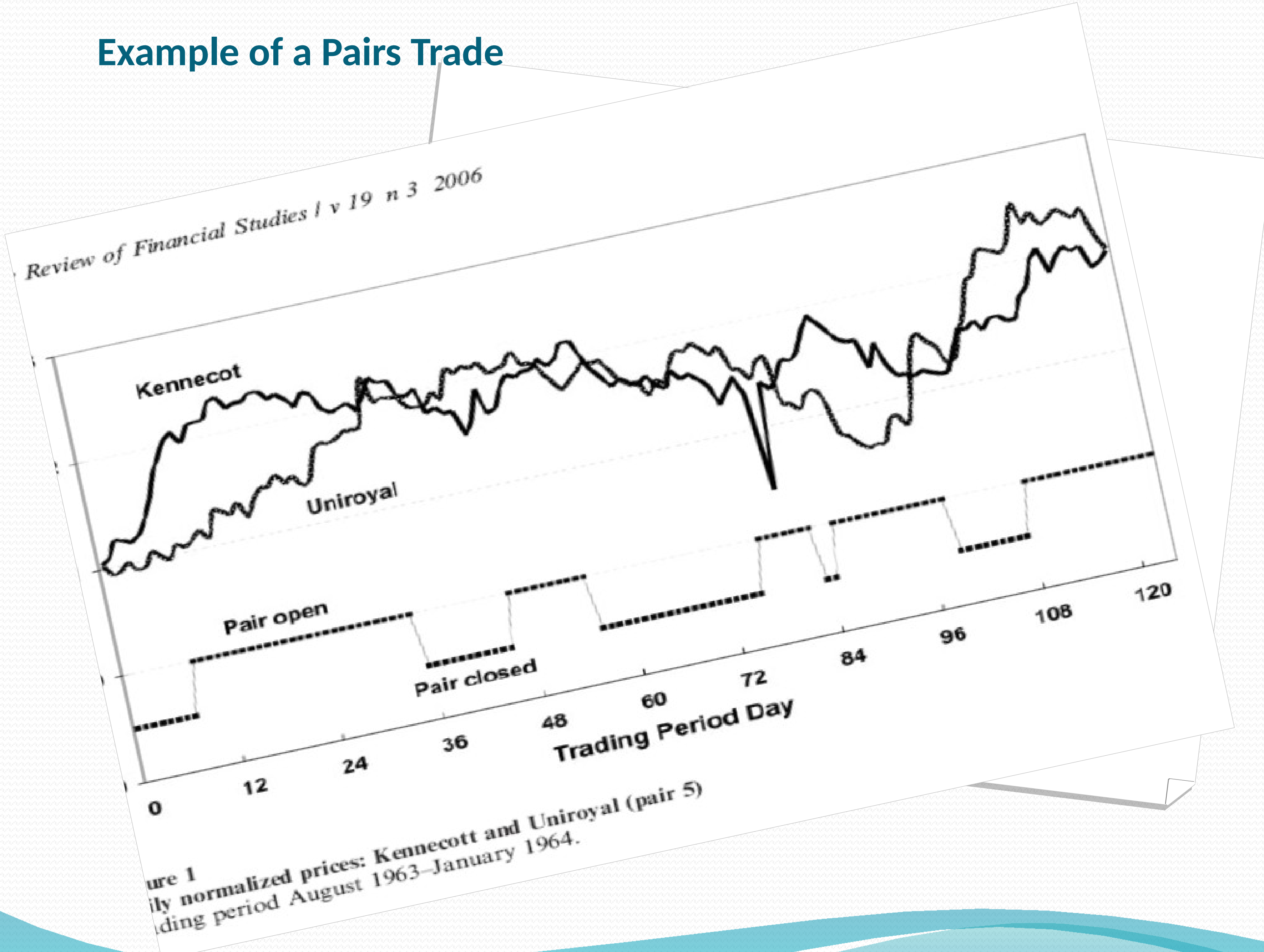
Background

- Developed in the 1980’s by a group of Quants at Morgan Stanley, who reportedly made over $50 million profit for the firm in 1987
- A contrarian strategy that tries to profit from the principles of mean-reversion processes
- In theory, one could expand the strategy to include a basket of more than a pair of related stocks

Main Idea

- Choose a pair of stocks that move together very closely, based on a certain criteria (i.e. Coke & Pepsi)
- Wait until the prices diverge beyond a certain threshold, then short the “winner” and buy the “loser”
- Reverse your positions when the two prices converge --> Profit from the reversal in trend

Example of a Pairs Trade



Investor Decisions

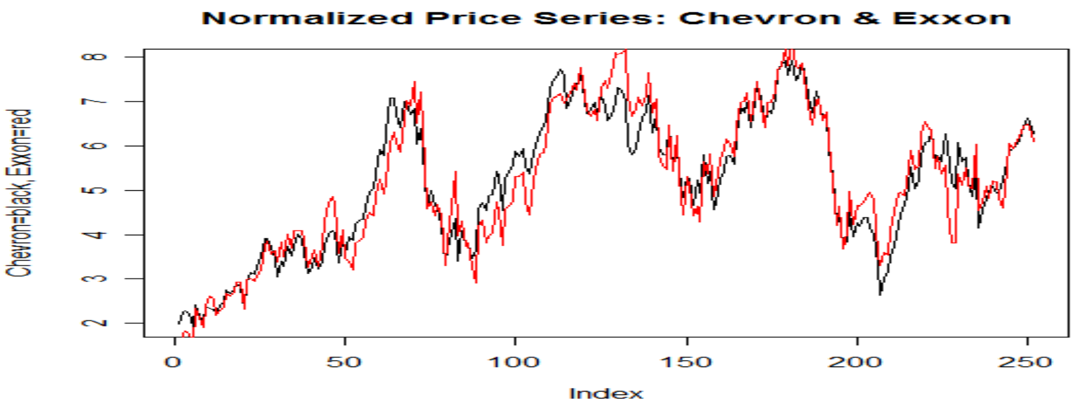
- Pair Selection Criteria
  - Correlation (Parametric & Non-Parametric Spearman’s Rho)
  - Dickey-Fuller Test Statistic (Cointegration)

- Trading Threshold (areas of consideration)
  - Volatility of the Market
  - Historical returns
  - Cost of each transaction

Normalization of Stock Data

- METHOD:
  - Find pair that has maximal correlation
  - Normalize price series, plot spread over 1 year “formation period”
  - Generate optimal threshold non-parametrically: choose a threshold **T_i=c*sd(spread),** calculate profit for each T_i, choose T_i generating max profit
    - Calculate profit by going $1 short on winner, $1 long on loser; close position when prices converge, i.e. spread=0
  - Normalize price series in 6 month “trading period” using mean and sd from formation period
  - Plot spread using optimal threshold found from formation period, calculate profit
  - Lower thresholds✉ More transactions✉ Higher transaction costs ✉ Lower Returns
  - Higher transaction costs✉ Smaller Returns

Chevron & Exxon



Formation Period Corr=0.93

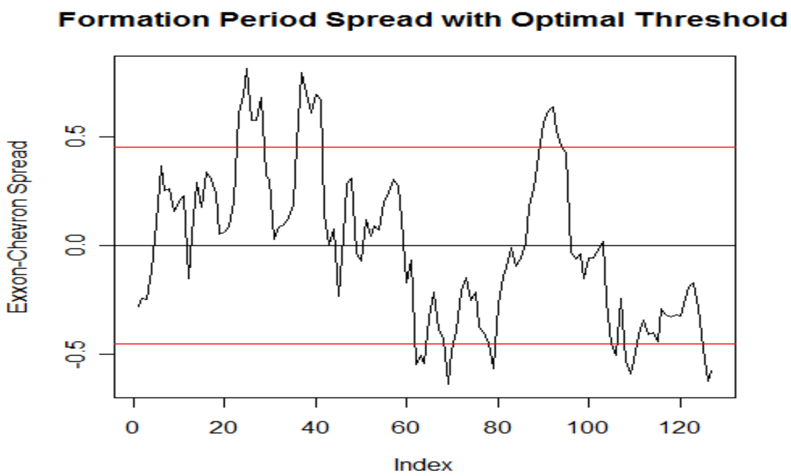
Trading Period Corr=0.96

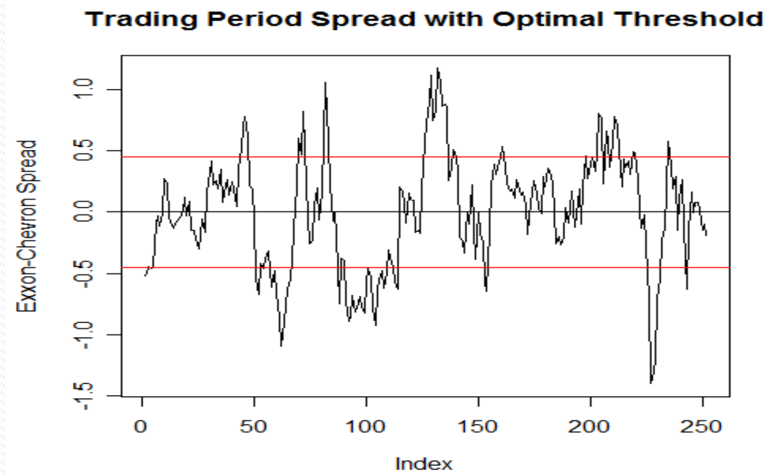
Optimal Threshold=1.25*sd's

Transactions=10

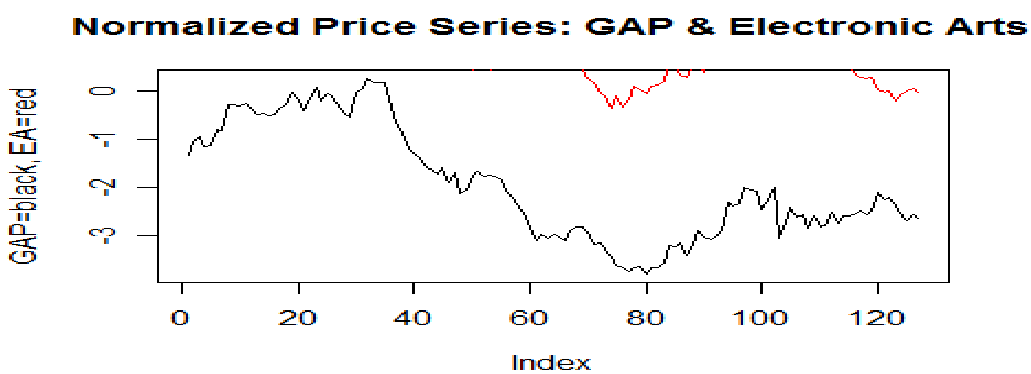
Returns=15%

Win.

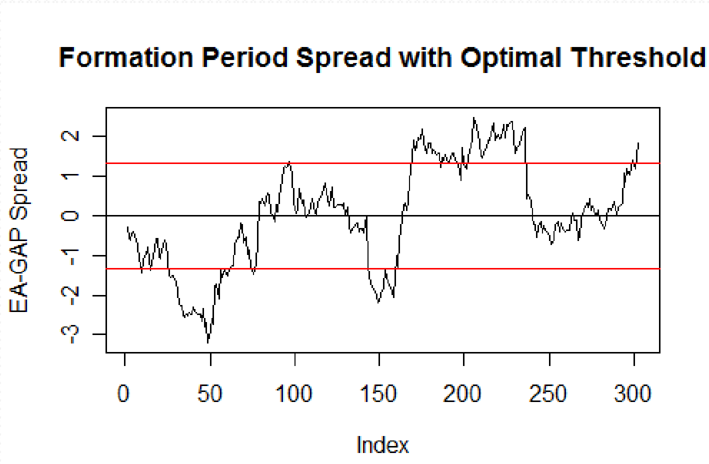


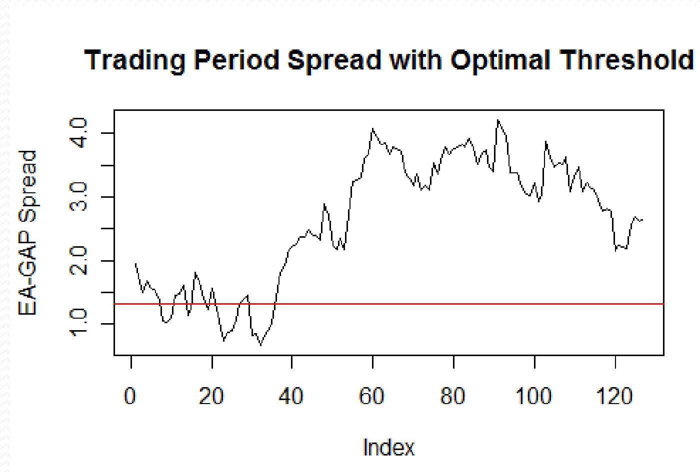


Electronic Arts & GAP

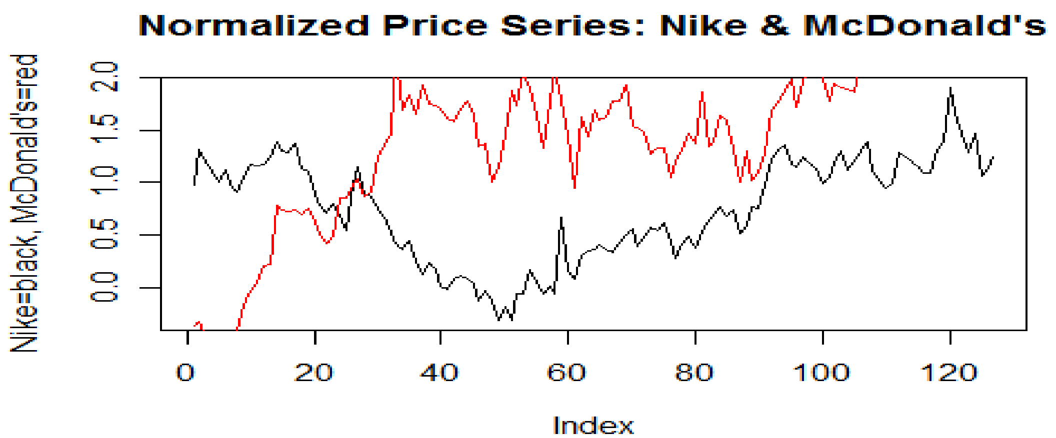


- Formation Corr=0.12
- Trading Corr=0.56
- Optimal Threshold=1 sd
- # Transactions=0 (Open a position, but spread never returns to 0)
- Return= -0.04
- **Lose.**

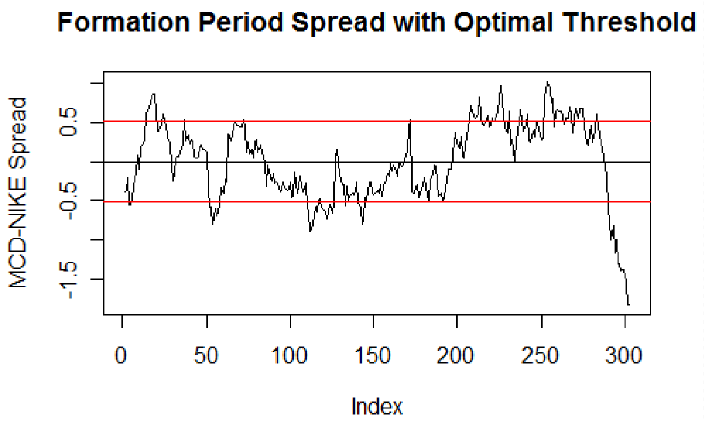


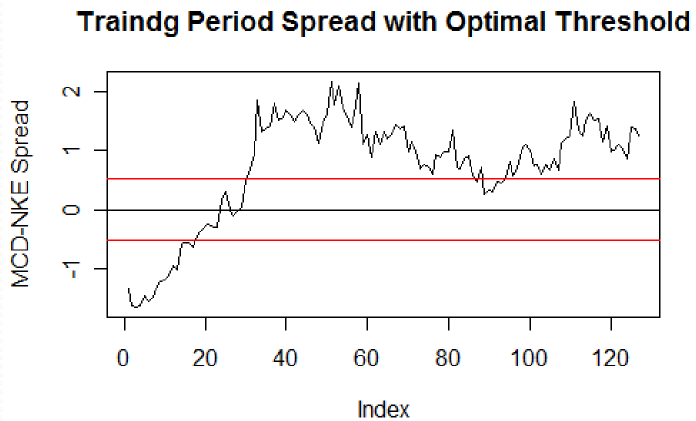


Nike & McDonald's



- Formation Corr=0.87
- Trading Corr=0.02
- #Transactions=1
- Return= -0.05
- Lose.
- Correlation is imperfect criteria for selecting pairs.





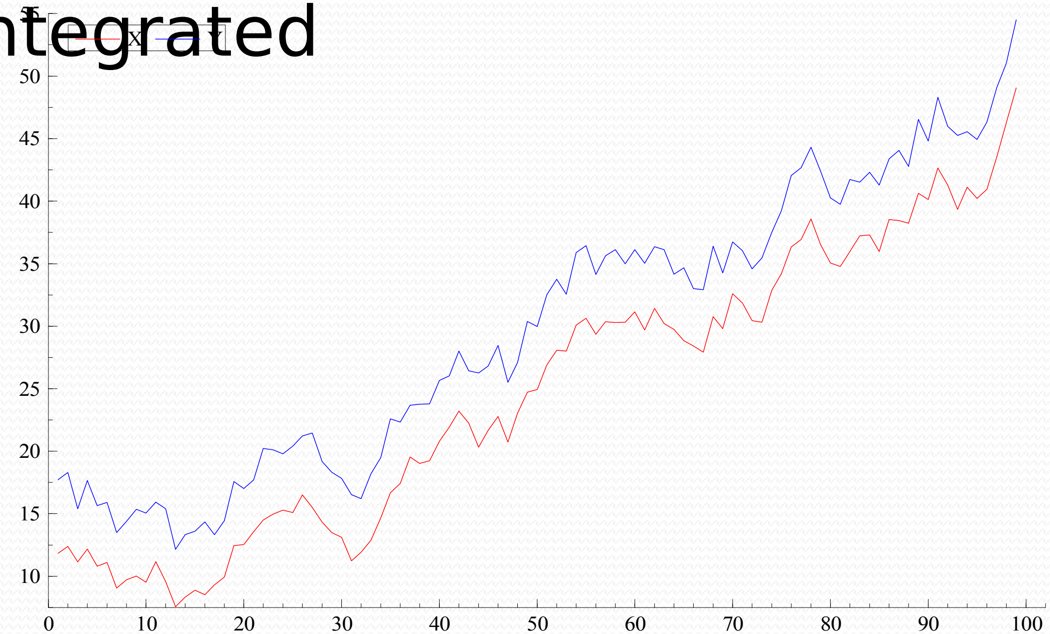
Cointegration

- If there exists a relationship between two non-stationary I(1) series, Y and X, such that the residuals of the regression

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

are stationary, then the variables in question are said to be cointegrated

Note: X and Y here are clearly not stationary, but they seem to move together. In fact, they are cointegrated --> ($Y - \beta_1 X - \beta_0$)should be stationary



Application to Pairs Trading

- If we have two stocks, X & Y, that are cointegrated in their price movements, then any divergence in the spread from zero should be temporary and mean-reverting.

- The important issues here are: 1) how to test for cointegration between prices and 2) estimating the constant

Testing For Cointegration

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

- Many Methods – most of them focus on testing whether the residuals of are stationary processes
- We use the Cointegrating Regression Dickey-Fuller Test, which essentially operates the following regression:

  $\boldsymbol{\Delta u_t = \varphi\, u_{t-1} + e_t}$

- H_0: $\boldsymbol{\varphi} = 0$ => no cointegration*
- H_a: $\boldsymbol{\varphi} < 0$ => cointegration*
- To obtain the cointegration factor estimates, we must regress the de-trended Y_t on the de-trended X_t

* We must use critical values different from Gaussian ones due to non-symmetric properties of the Dickey-Fuller distribution

Results of Test

- NO PAIR OF PRICES ARE COINTEGRATED!
- No surprise there
- Alternative: take the “most cointegrated” pair & optimize thresholds as we did with normalized data
- Compare the results against normalized thresholds in the same time period

Auto-Regressive Time Series

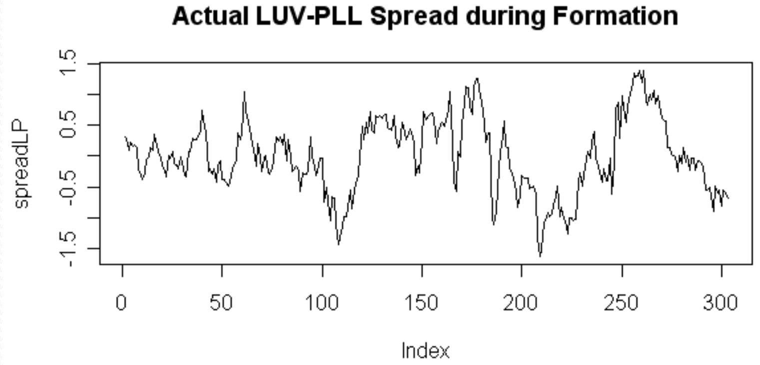
- Cointegration is an ideal construct for pairs trading
- But Dickey-Fuller Hypothesis Test is inconclusive
- Instead we can fit a time series to the spread data
  - AR(1): $Y_t = \beta Y_{t-1} + \varepsilon_t$
- Looking for a spread that produces an AR(1) with $|\beta| < 1$, so that will be stationary.

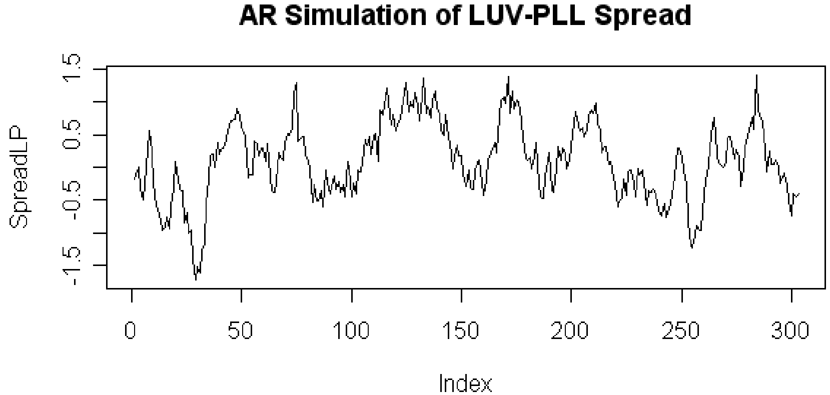
Choosing thresholds with AR(1)

- For the interest of time, we are only going to focus our most cointegrated pair: LUV and PLL.
- We will fit an AR(1) to the data by estimating β and the standard deviation of each iid white noise $\varepsilon_{t.}$
- Then we will run one thousand simulations of this AR(1) model and estimate each of their optimal benchmarks
- The average of the optimal benchmarks from each simulation will serve as our estimate for the optimal benchmark in the formation period.

Results of AR(1) Thresholds

AR(1) Coefficient estimate (µ-hat)	0.8605
Optimal Threshold estimate	1.046
SD of Optimal Threshold	0.2597
Number of Transactions	12
Returns over Trading	17.7%





Alternative Strategies

- Conditional correlation or some other measure of “relatedness”, such as Copulas
- Modeling the spread as GARCH processes
- Optimize profits w.r.t. certain global indicators (i.e. market volatility, industry growth, etc.)
- Factor Analysis on the spread

Bibliography

- Gatev, Evan, William N. Goetzmann, and K. Geert Rouwenhorst, "Pairs Trading: Performance of a Relative-Value Arbitrage Rule," *Review of Financial Studies* (2006): 797-827.

- Vidyamurthy, Ganapathy, *Pairs Trading: Quantitative Methods and Analysis* (New Jersey: John Wiley & Sons, Inc., 2004).

- Wooldridge, Jefferey M., *Introductory Econometrics, A Modern Approach, Third Edition* (Ohio: Thomson South-Western, 2006).

Thank You!

Questions?