Men's Arbitrage

Statistical Arbitrage in the US market

Executive Summary

- Introduction
  - *Motivation*
  - *Backtesting Methodology*
  - *Performance Summary*
- Methodology of Men's Arbitrage
  - *PCA as Risk Factors*
  - *Residual Process*
  - *Trading Signal (Portfolio Selection)*
  - *Mean-reversion Requirement (Portfolio Selection)*
  - *OU process for parameter estimations*
  - *Model Summary and Trading Rules*
- Extensions
  - *Maximum Number of Daily Positions*
  - *Sensitivity Analysis on Long Short Threshold*
  - *Dynamic Number of Principal Components*

Executive Summary

- Introduction
  - *Motivation*
  - *Backtesting Methodology*
  - *Performance Summary*
- Methodology of Men's Arbitrage
  - *PCA as Risk Factors*
  - *Residual Process*
  - *Trading Signal*
  - *Mean-reversion Requirement*
  - *OU process for parameter estimations*
  - *Model Summary and Trading Rules*
- Extensions
  - *Maximum Number of Daily Positions*
  - *Sensitivity Analysis on Long Short Threshold*
  - *Dynamic Number of Principal Components*

Motivation

- Build factor models with PCs factors
- Model idiosyncratic residual process as mean-reverting OU-process
- Derive standardized S-score as trading signals
- Hold ~200 fully-hedged portfolios daily

Goal :

1) Market Neutral

2) Low-volatility, low-risk arbitrage returns

Backtesting Methodology

Platform:
BackTrader
Trading Universe:
US Equity Market(SP500 + Nasdaq100)
Cash:
10 mil USD
Transaction Cost & Slippage:
10bps
Rebalancing Period:
Daily
Back Test Period:
01/01/2013 (Start of data feeding)
01/01/2014 (Start of trading)
28/05/2021 (End of trading)

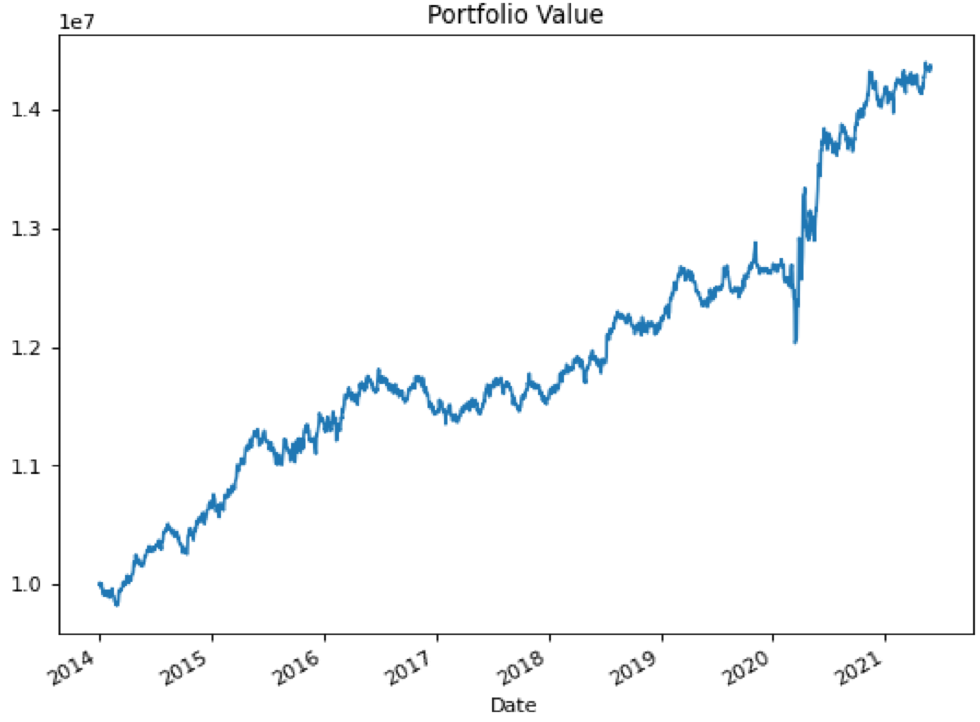


Figure 1. Portfolio Cumulative Return of Men's Arbitrage Strategy

Performance Summary of Men's Arbitrage

Higher Sharpe Ratio than SP500 (2013-2021)

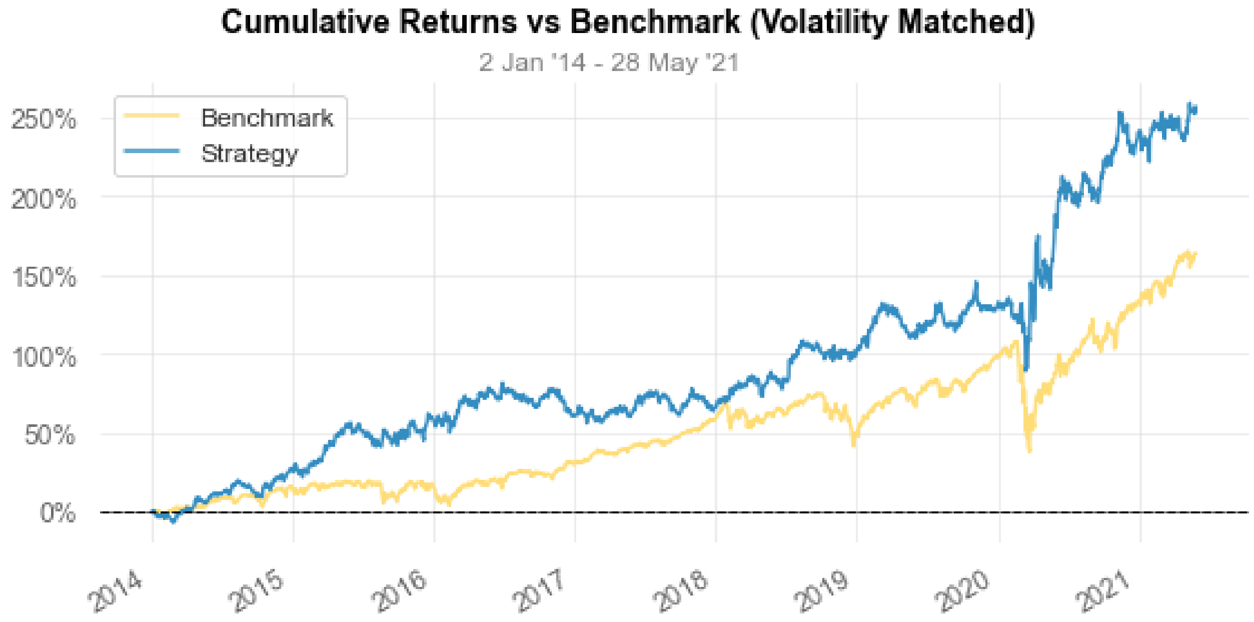


Figure 2. Volatility matched return of Men's Arbitrage Strategy vs SP 500

Market Neutral

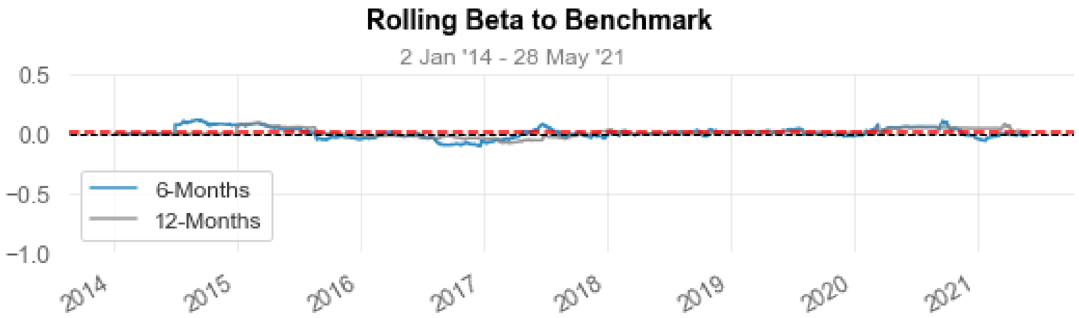


Figure 3. Rolling Betas during the testing periods

Key Performance Metrics

Period: 2013-2021	Men's Arbitrage	SP500
Sharpe Ratio	1.07	0.84
Calmar Ratio	0.76	0.42
Sortino Ratio	1.69	1.17
Maximum Drawdown	-6.59%	-33.72%

Table 1. Comparison of Men's Arbitrage vs other benchmarks

Executive Summary

- Introduction
  - *Motivation*
  - *Backtesting Methodology*
  - *Performance Summary*
- Methodology of Men’s Arbitrage
  - *PCA as Risk Factors*
  - *Residual Process*
  - *Trading Signal (Portfolio Selection)*
  - *Mean-reversion Requirement (Portfolio Selection)*
  - *OU process for parameter estimations*
  - *Model Summary and Trading Rules*
- Extensions
  - *Maximum Number of Daily Positions*
  - *Sensitivity Analysis on Long Short Threshold*
  - *Dynamic Number of Principal Components*

Principal Components as Risk Factors

Risk Factor Decomposition (OLS)**:** systematic & idiosyncratic components

Idiosyncratic: dift, residuals

$$\frac{dS_t}{S_t} = \alpha\, dt + \underbrace{\sum_{i=1}^{n} \beta_i\, F_i}_{\text{Systematic}} + dX_t \xrightarrow{\text{15 PC factors}} \frac{dS_t}{S_t} = \alpha\, dt + \sum_{i=1}^{15} \beta_i\, r_{pc_i} + dX_t$$

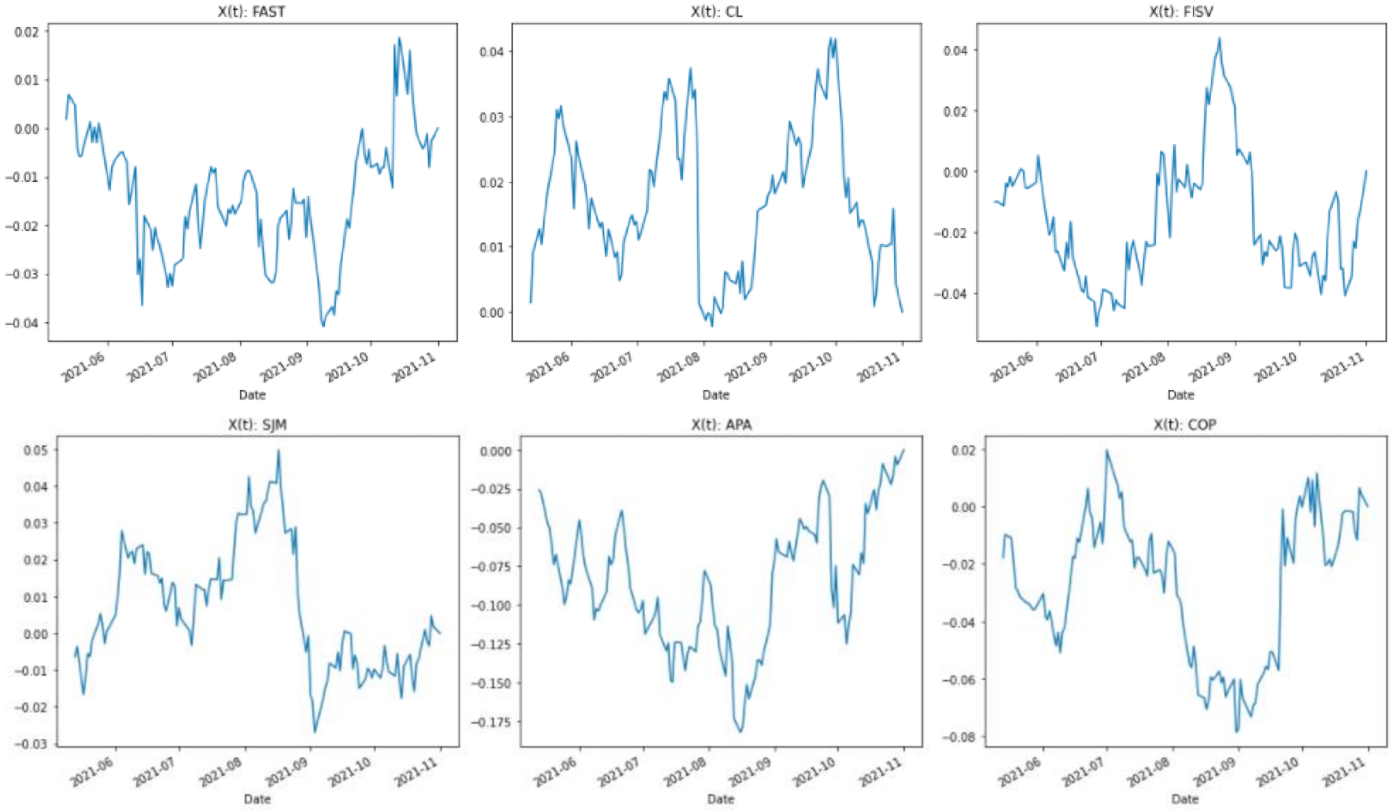
Question: What is X(t) in idiosyncratic part?

-Neither drift nor systematic

-Hypothesis: **Mispricing** due to **market over-reaction** $\longrightarrow$ **Arbitrage Opportunity**

Mean-reverting Residual Process X(t)

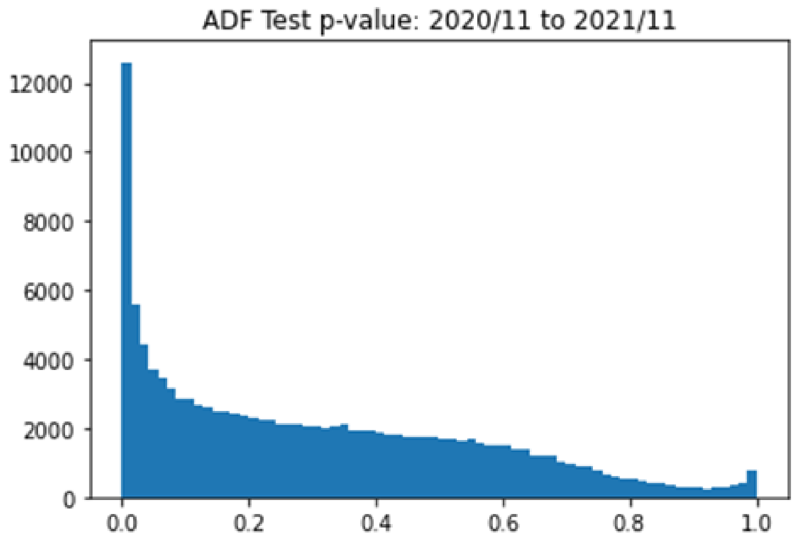
Hypothesis: If X(t) is caused by market over-reaction, logical to assume it **mean-reverts** around an equilibrium mean --> market is rational in the longer run as sentiments fade away

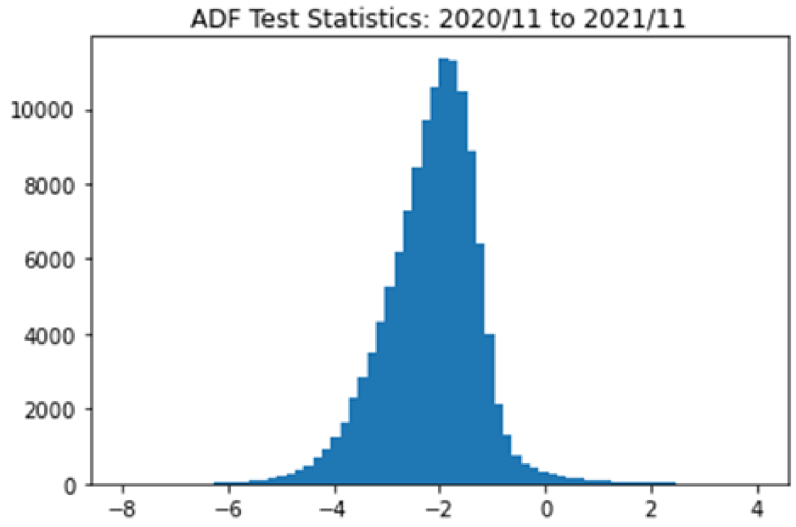


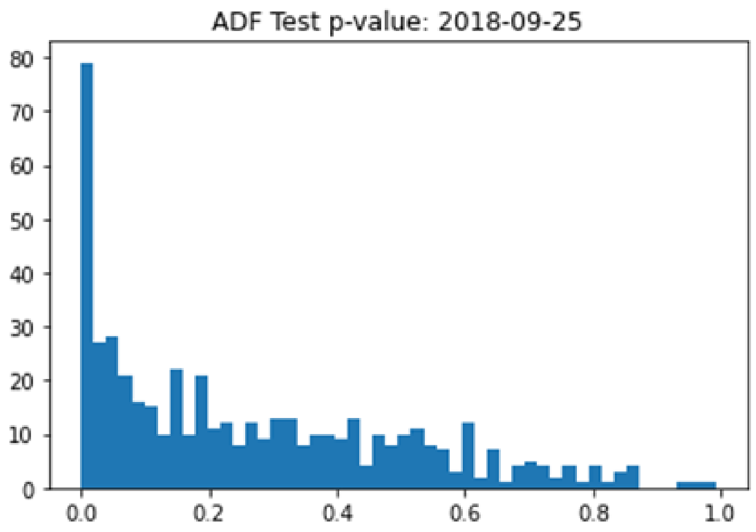
Notebook Demo

Residual process X(t) ADF Test

ADF Test on residual process X(t), SP500:







Result: most of them are mean-reverting, some are not
In practice, we trades every stocks, reasons:
1) equilibrium strategy --> market efficiency effect
2) marginal cost of hedging is low for additional pair (explained later)

X(t) Model

Ornstein-Uhlenbeck process: $dX_t = k(m - X_t)\,dt + \sigma dW_t$

Expected return at time s < t: $E[dX|X(s)] = k(m - X(s))dt$

Hedged Portfolio:

-For every $1 long in S, short $B1 of PC1, $B2 of PC2, …, etc

$$\frac{dS_t}{S_t} = \cancel{\alpha\,dt} + \cancel{\sum_{i=1}^{15} \beta_i\, r_{pc_i}} + dX_t \longrightarrow dP_t = dX_t \longrightarrow$$

Hedged Portfolio value ~~ X(t):

1) **+ve expected return**: X(t) < m

2) **-ve expected return**: X(t) > m

OU Process: Parameter Estimation for generating Trading Signals

OU process:

$$dX_t = km\,dt - kX_t\,dt + \sigma dW_t$$
$$dX_t + kX_t dt = km\,dt + \sigma dW_t$$

Multiply e^{kt} on both sides:

$$e^{kt}(dX_t + kX_t dt) = e^{kt}km\,dt + e^{kt}\sigma dW_t$$

Integrate from 0 to T on both sides:

$$\int_0^T \frac{d(e^{ks}X_s)}{ds} = e^{kT}X_T - X_0 = m\,(e^{kT}-1) + \sigma \int_0^T e^{ks}dW_s$$
$$X_T = X_0 e^{-kT} + m\,(1-e^{-kT}) + \sigma \int_0^T e^{-k(T-s)}dW_s$$

Expectation and Variance:

$$E[X_T] = X_0 e^{-kT} + m\,(1-e^{-kT})$$
$$Var(X_T) = \sigma^2 \int_0^T e^{-2k(T-s)}ds = \frac{\sigma^2}{2}\,(1-e^{-2kT})$$

In terms of each step from $t = t$ to $t = t + \Delta t$:

$$E[X_{t+\Delta t}] = X_0 e^{-k\Delta t} + m\,(1-e^{-k\Delta t})$$
$$Var(X_{t+\Delta t}) = \frac{\sigma^2}{2k}\,(1-e^{-2k\Delta t})$$

Therefore:

$$X_{t+\Delta t} = X_t e^{-k\Delta t} + m\,(1-e^{-k\Delta t}) + \sigma\sqrt{\frac{1-e^{-2k\Delta t}}{2k}} * N(0,1)$$

This is equivalent to AR(1) model:

$$y_{i+1} = b\,y_i + a + \epsilon_{i+1}$$

Thus, parameters can be estimated through linear regression of X(t) with lag 1 X(t-1):

$$b = e^{-k\Delta t},\;\; a = m\,(1-e^{-k\Delta t}),SE = \sigma\sqrt{\frac{1-e^{-2k\Delta t}}{2k}}$$

$$k = -\log(b) * 252$$

$$m = \frac{a}{1-b}$$

$$\sigma = SE * \sqrt{\frac{2k}{1-e^{-2k\Delta t}}}$$

$$\sigma_{eq} = \frac{\sigma}{\sqrt{2k}}$$

Estimation is Simple!
Lag-1 regression on X(t)

*estimation windows is 60 trading days

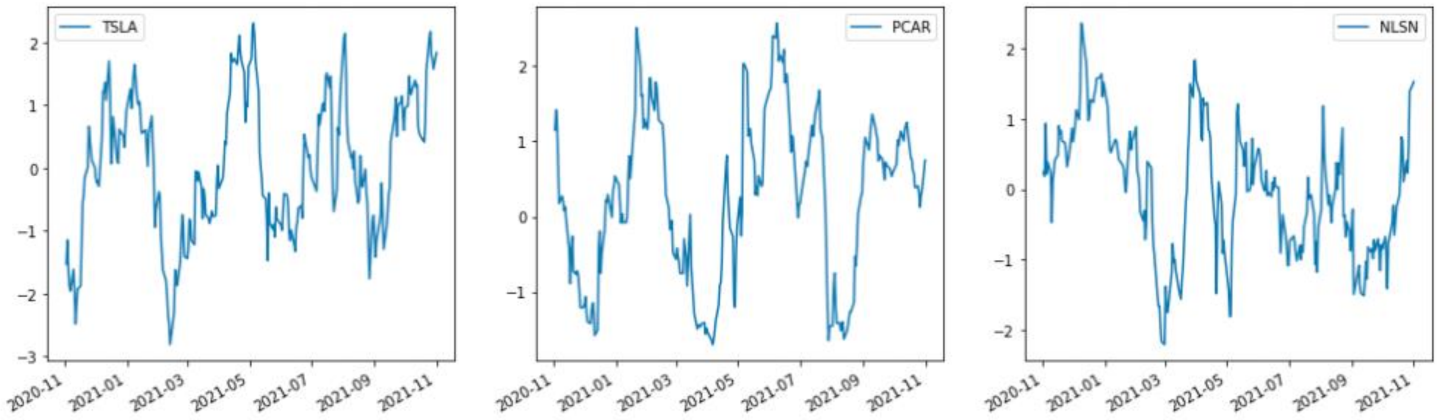
Trading signals: S score (Portfolio Selection)

Standardized measure: extent of deviations from mean

$$s_i = \frac{(X_i(t) - m_i)}{\sigma_{i,eq}} \xrightarrow{\text{Regression: X(T) = 0}} s_i = -\frac{m_i}{\sigma_{i,eq}} \xrightarrow{\text{Centered mean}} s_i = -\frac{\overline{m_i}}{\sigma_{i,eq}}$$

$\overline{m_i} = m_i - \langle m \rangle$ Where <m> : average of m across stocks

Figure 6. S score for TSLA, PCAR and NLSN



Long Short Thresholds on S score (Preset Value):

S_bo: threshold buy open (1.00)

S_cl: threshold close long (0.5)

S_so: threshold sell open (1.25)

S_cs: threshold close short (0.5)

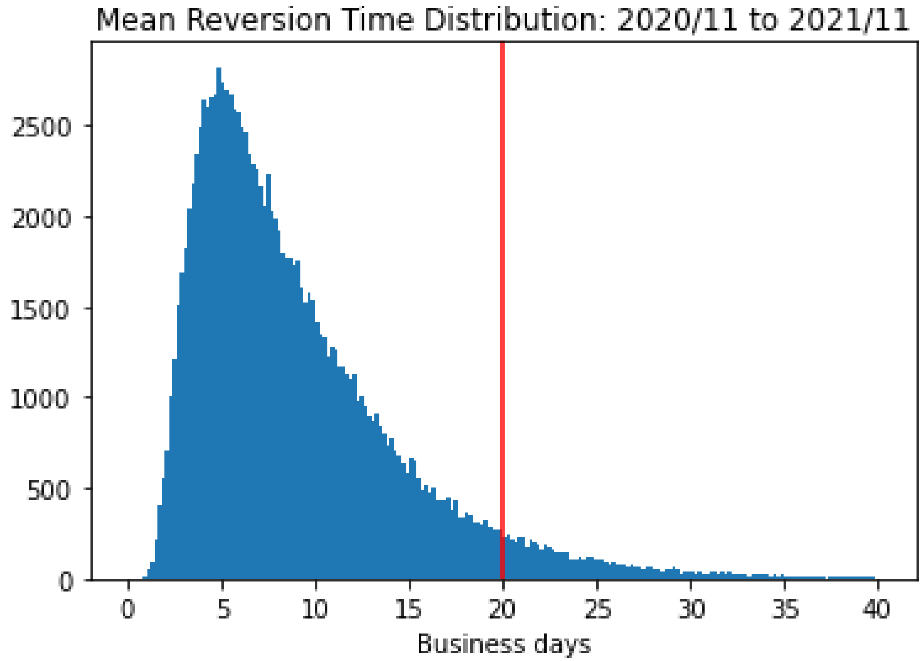
Intuitively, Residual that deviates by -1.00 standard deviations from equilibrium

Portfolio Selection: Mean-reversion requirement: k

Mean reversion speed: k
Mean reversion time in business days: 1/k * 252

Goal: Trade stocks (corresponding hedged portfolios) that have fast mean-reversion:
--> drift effect minimal
--> More trades, more stable returns

Our threshold:
K > 252/20
Mean-reversion time < 20 days



Model Summary and Trading Rules

Model Summary

Regression Estimates of residuals

$$\frac{dS_t}{S_t} = \cancel{\alpha\, dt} + \cancel{\sum_{i=1}^{12} \beta_i\, r_{pc_i}} + dX_t$$

Auxiliary Process to get {X.t}

$$dP_t = dX_t$$

AR(1) process

$$y_{i+1} = b\, y_i + a + \epsilon_{i+1}$$

OU Parameter Estimation

$$b = e^{-k\Delta t},\ \ a = m\left(1 - e^{-k\Delta t}\right), SE = \sigma\sqrt{\frac{1 - e^{-2k\Delta t}}{2k}}$$

$$k = -\log(b) * 252$$

$$m = \frac{a}{1 - b}$$

$$\sigma = SE * \sqrt{\frac{2k}{1 - e^{-2k\Delta t}}}$$

$$\sigma_{eq} = \frac{\sigma}{\sqrt{2k}}$$

S-Score Generation (Trade Signals)

$$s_i = -\frac{\overline{m_i}}{\sigma_{i,eq}}$$

Trading Rules: Pseudo Codes

```
Init vector: trade_size
  For each stock k > 252/20:
     If s > 1.25 (s_so):
          trade_size += short stock, hedge with 15 PCs

     If s < 0.75 (s_cs):
          trade_size += close short.

     If s < -1 (s_bo):
          trade_size += long stock, hedge with 15 PCs

     If s > -0.5 (s_cl):
          trade_size += close long.
Execute(trade_size) next open

Max pos: 200
Invest val per pos: 1/200 PV
```

Theory vs Practice: Txn Cost

In theory:

High Transaction Cost

Each position requires hedging with PCs, which are each portfolio of every single stock. Txn cost is high.

In Practice:

Mild Transaction Cost

Hedging parts of long & short positions cancel out each other. Required hedging for whole portfolio is minimal.

Ticker	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
A	0.002409	0.009975	-0.007671	0.129792	-0.033384	-0.017223	0.202095	-0.008605	-0.266438	0.077373	-0.185659	0.075023	0.105015	-0.059876
AAL	0.001226	-0.002681	-0.004560	0.065476	0.015362	-0.020875	-0.069183	0.004314	-0.027674	0.025384	-0.076443	0.020395	-0.037223	0.021915
AAP	0.002404	0.000644	0.005121	-0.299161	-0.042200	0.032816	-0.149359	-0.014419	0.154812	0.081737	-0.348698	0.030116	0.020116	0.296472
AAPL	0.001164	0.008338	-0.005093	0.123791	0.040500	-0.030271	-0.024034	-0.024252	0.083872	0.005239	-0.047636	-0.043518	0.066935	0.077219
ABBV	0.001775	0.003008	0.001643	0.175814	0.000969	0.182547	-0.010181	0.062513	-0.161859	0.064938	-0.031489	0.015019	-0.034928	0.116536
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
FOX	0.001349	-0.001645	0.003545	-0.126213	0.032921	-0.055927	-0.003403	0.039839	0.086452	0.106330	0.485908	0.013620	0.113391	-0.194918
DOW	0.002841	-0.002698	-0.002543	-0.212104	0.002487	0.014949	0.038337	0.003503	0.032235	-0.042974	0.147986	0.031286	-0.095256	-0.008959
CTVA	0.002784	-0.000081	-0.002804	-0.204992	-0.012152	0.014035	0.025377	0.051163	-0.014989	-0.041642	0.171005	0.012964	0.036041	-0.048443
CARR	0.002453	0.002652	-0.001136	-0.259744	-0.030062	-0.029271	0.036601	-0.000387	-0.188196	-0.006369	0.160044	0.030238	-0.124353	-0.127913
OTIS	0.002497	0.002949	0.004111	-0.231085	-0.020986	-0.016601	0.017728	-0.026479	-0.136130	-0.087704	0.415644	0.011958	0.034688	-0.209929

Figure 1. Eigen portfolios

Example:
1. Long $1 of A share, hedge with short $0.5 PC1
2. Short $1 of B share, hedge with long $0.4 PC1
1+2: Hedge whole portfolio by short $0.1 PC1
--> ~200 positions , overall hedging is less.

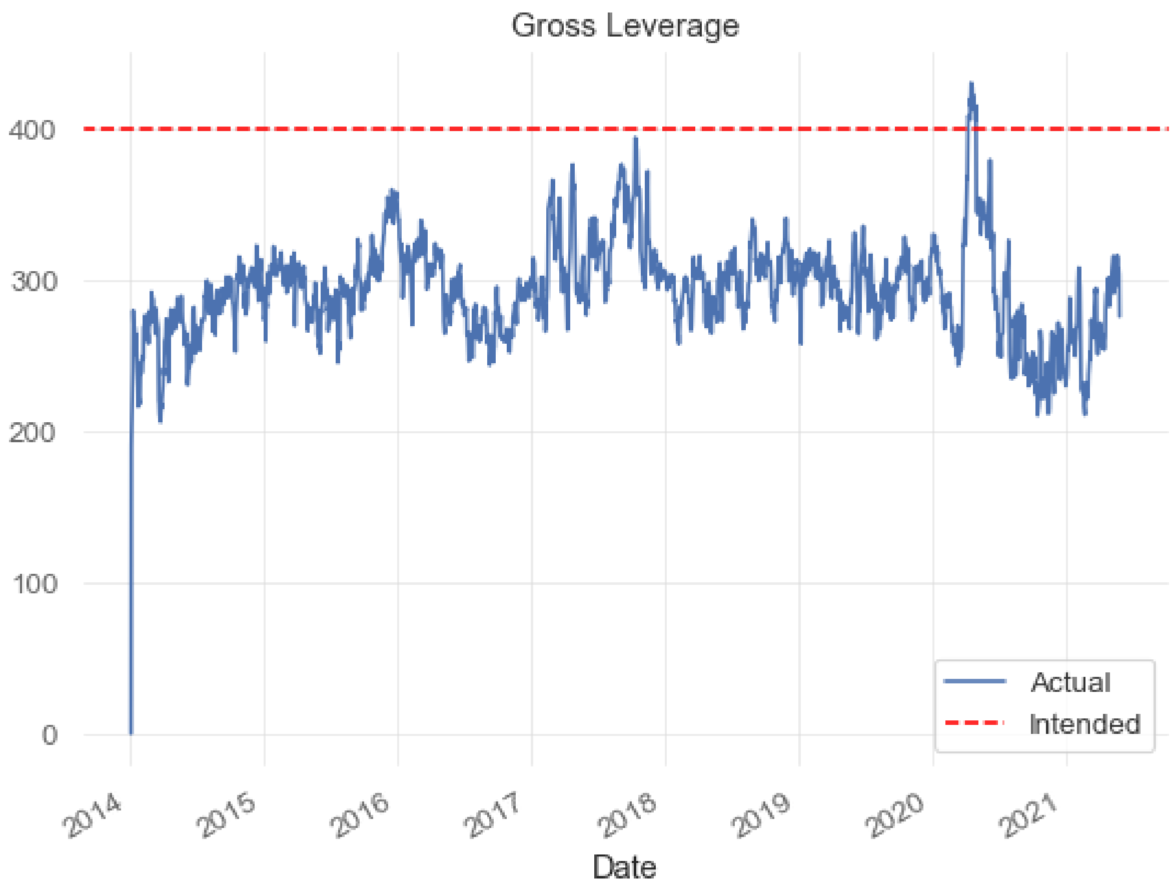
Ticker	
A	0.004119
AAL	-0.003111
AAP	0.067665
AAPL	-0.001638
ABBV	0.004583
...	...
FOX	-0.065664
DOW	-0.004277
CTVA	-0.01578
CARR	-0.007438
OTIS	-0.023671

Figure 2. Resulting weight vector for two pairs

Gross Leverage

Intended:
2x gross both leg, 4x gross leverage

Actual:
~3x gross leverage (hedging parts cancels off)



Executive Summary

- Introduction
  - *Motivation*
  - *Backtesting Methodology*
  - *Performance Summary*
- Methodology of Men’s Arbitrage
  - *PCA as Risk Factors*
  - *Residual Process*
  - *Trading Signal*
  - *Mean-reversion Requirement*
  - *OU process for parameter estimations*
  - *Model Summary and Trading Rules*
- Extensions
  - *Maximum Number of Daily Positions*
  - *Sensitivity Analysis on Long Short Threshold*
  - *Dynamic Number of Principal Components*

Sensitivity Analysis on Long Short Threshold

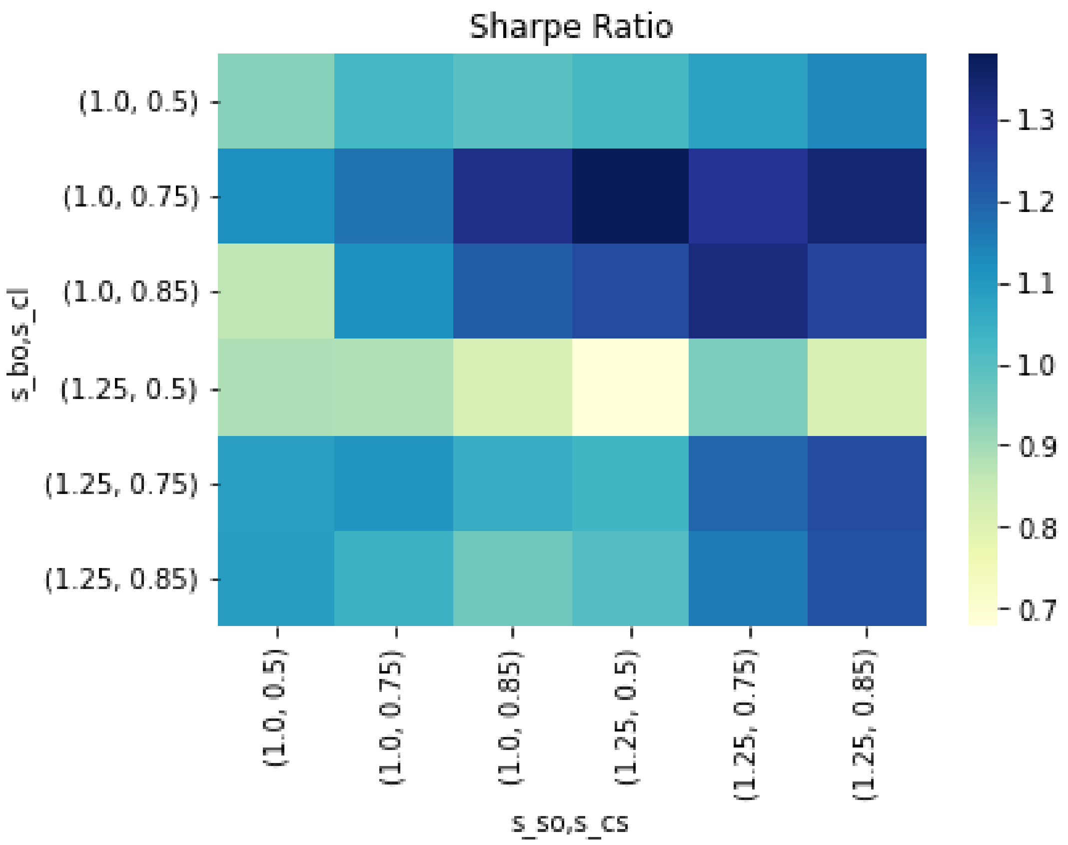
Period
2017/01/01 - 2021/06/01

Tested Parameter
Long Threshold: (s_bo, s_cl)
Short Threshold: (s_so, s_cs)

Model Parameter : (-1 ,-0.5) , (0.75 ,1.25)

Result
Sharpe Ratio: ranged from 0.7 to 1.4
SP500 Sharpe = 0.84

Most of the pairs outperform the market



Maximum Number of Daily Positions

Observation

Limit the maximum number of daily position to 200 generate a better back testing result than 150

Suggestion

Experimenting with the higher number of max daily positions

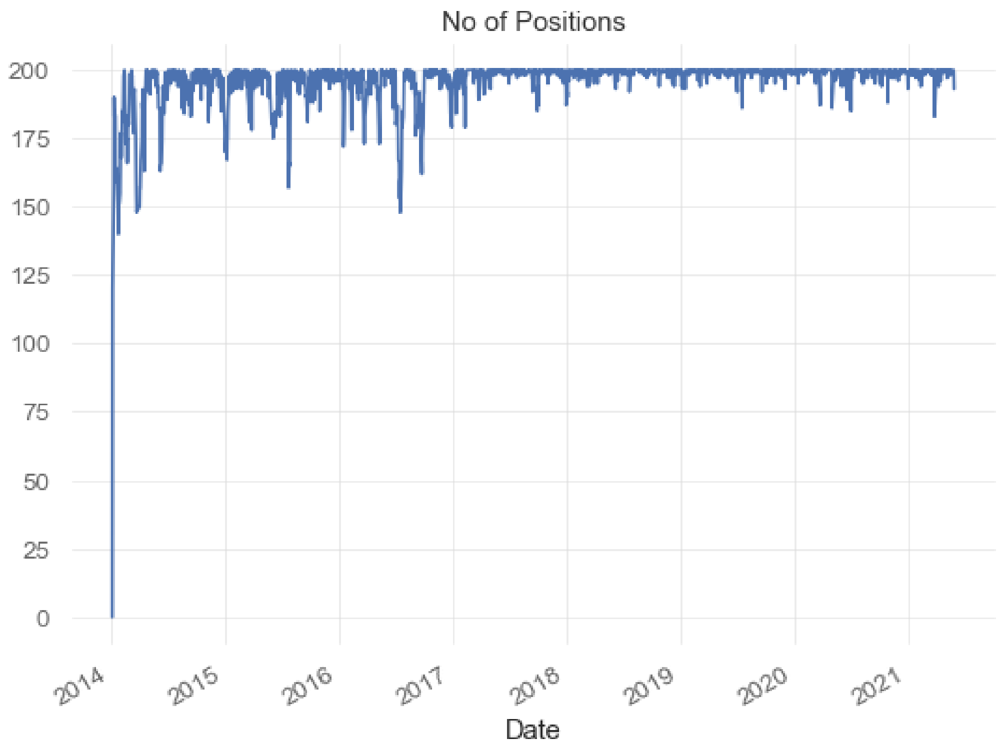


Fig 1. Graph of number of position over backtesting periods

Dynamic Number of Principal Components

Observation

Portfolio cumulative return increases rapidly with high explained variance %

From 2016 to 2018 , cumulative return stop increasing when the explained variance % is low

Suggestion

Potential Performance improvement by using dynamic principal components to maintain a high explained variance

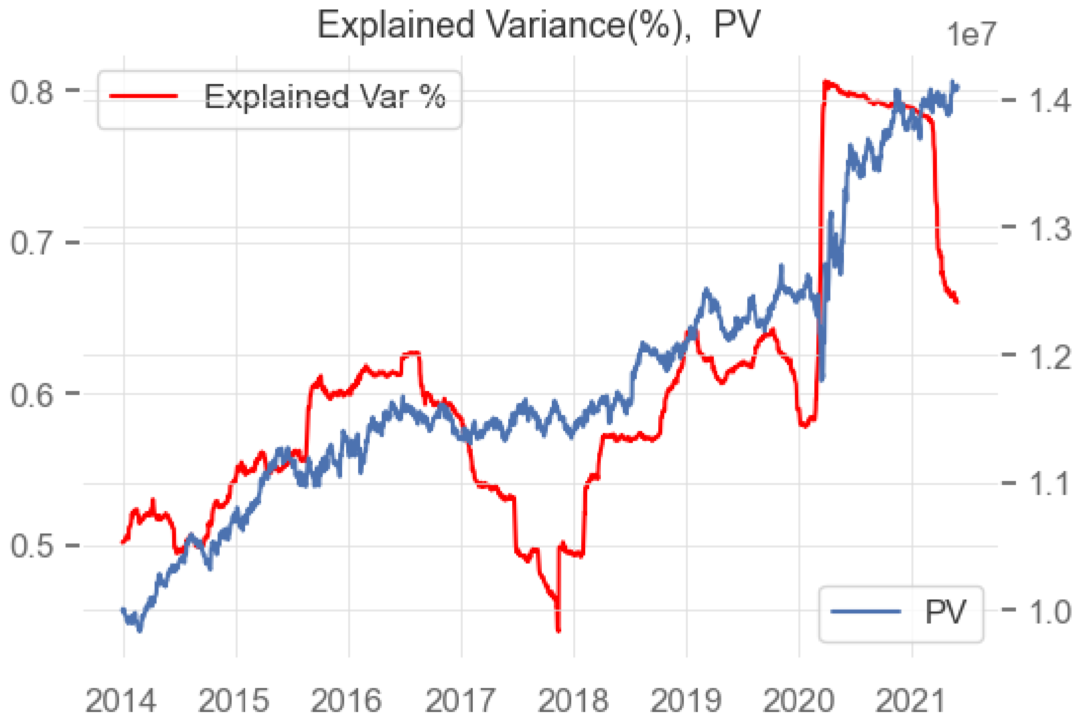


Figure 1. Comparison of explained variance and portfolio return over backtesting periods

Performance Summary of Men's Arbitrage

Higher Sharpe Ratio than SP500 (2013-2021)

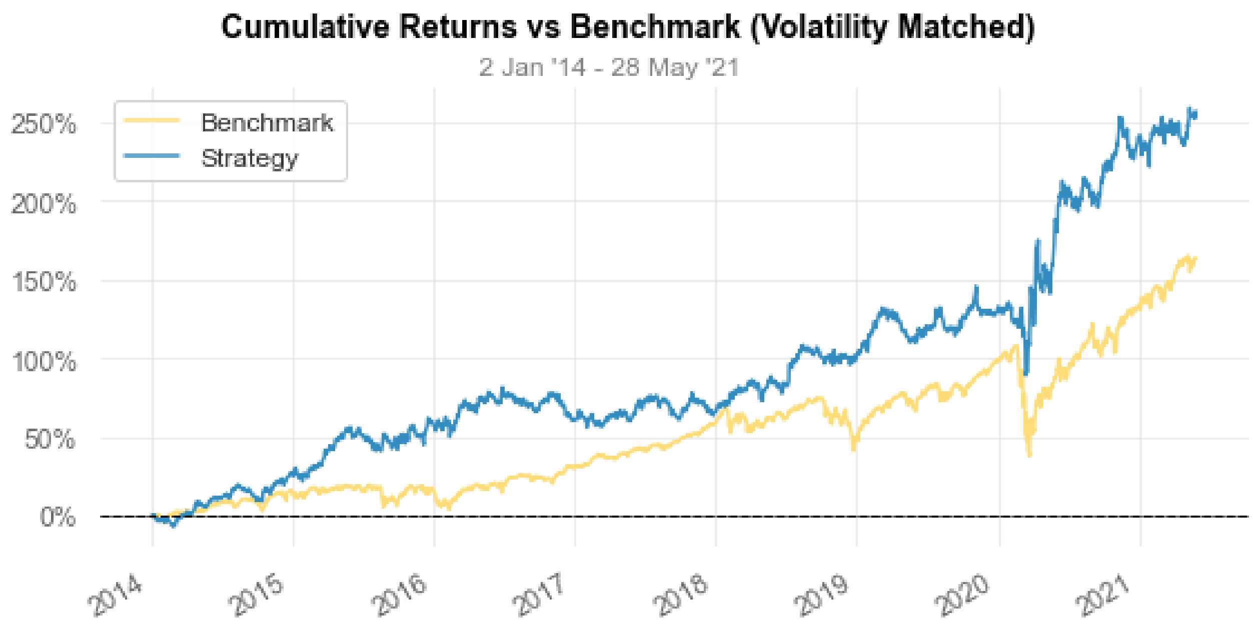


Figure 2. Volatility matched return of Men's Arbitrage Strategy vs SP 500

Market Neutral

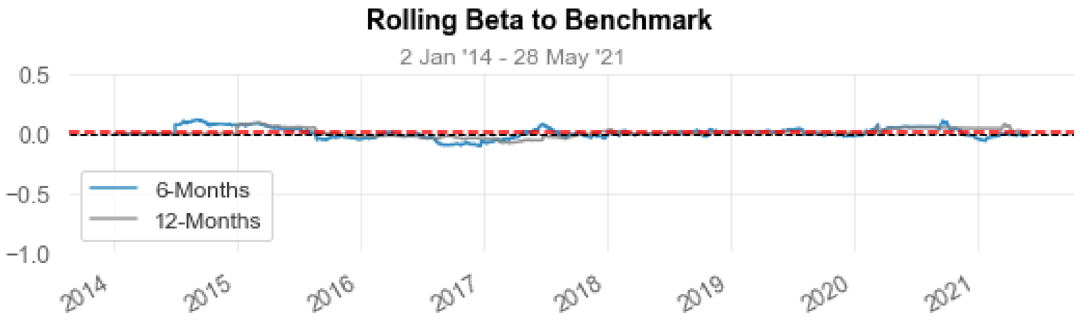


Figure 3. Rolling Betas during the testing periods

Key Performance Metrics

Period: 2013-2021	Men's Arbitrage	SP500
Sharpe Ratio	1.07	0.84
Calmar Ratio	0.76	0.42
Sortino Ratio	1.69	1.17
Maximum Drawdown	-6.59%	-33.72%

Table 1. Comparison of Men's Arbitrage vs other benchmarks