MS&E448: Statistical Arbitrage

—

Group 5: Carolyn Soo, Zhengyi Lian, Jiayu Lou, Hang Yang

Statistical Arbitrage (Recap)

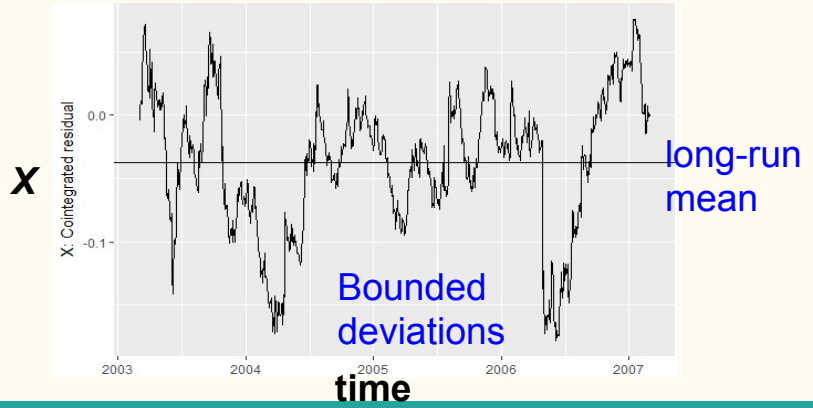
How to systematically identify stationary spreads X(t) and profit from their mean reversion?

- Stock returns - systematic and idiosyncratic components

$$\frac{dP_t}{P_t} = \alpha dt + \sum_{j=1}^{n} \beta_j F_t^{(j)} + dX_t$$

- Form beta-neutral long-short positions using "group ($n \geq 1$)" trades:
  - long: for every \$1 long in asset P, short \$$\boldsymbol{\beta}_j$ of ETF$^{(j)}$ for j = 1 to n
  - short: for every \$1 short in asset P, long \$$\boldsymbol{\beta}_j$ of ETF$^{(j)}$ for j = 1 to n

- Profit from mean-reversion of spread X
  - test for stationarity
  - model as OU process (mean-reverting)
  - check for fast mean reversion ($\boldsymbol{\tau} < 30$ days)
  - compute trading signals
  - execute



Stationarity Test & Fitting an OU-Model

- **Augmented Dickney Fuller (ADF) Test for Stationarity** of X

$$\Delta X_t = \alpha + \gamma X_{t-1} + \sum_{j=1}^{p-1} \delta_j \Delta X_{t-j} + \epsilon_t$$

H_0: $\gamma = 0$ (X_t is non-stationary)
test-statistic: $\hat{\gamma}/S.E.(\hat{\gamma})$
reject if p-value < 0.05

- **X(t)** for each stock modelled as an **Ornstein-Uhlembeck (OU) process**

$$dX(t) = \kappa\Big(m - X(t)\Big)dt + \sigma dW(t)$$

W_t : Wiener process

- Assume κ, m, σ for each stock stays constant over 60-day trailing window,

$$X(t+\Delta t) = a + bX(t) + \zeta$$

$$\begin{cases} \kappa = -ln(b)/\Delta t \\ m = \dfrac{a}{1-b} \\ \sigma = \sqrt{\dfrac{\text{Variance}(\zeta)\cdot 2\kappa}{1-b^2}} \end{cases}$$

accept only if mean reversion time-constant

$\tau = 1/\kappa < 30 days$

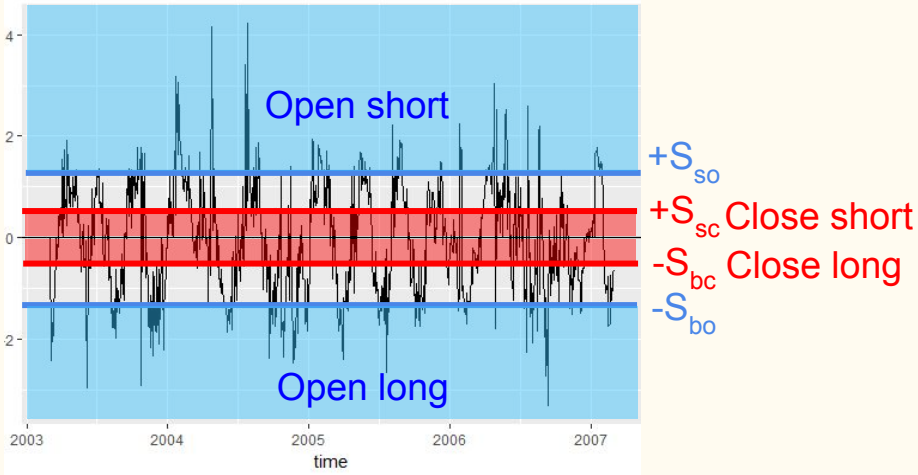
Compute Trading Signals

- *In equilibrium* (i.e. as $\Delta t \to \infty$),

$$\mathbb{E}(X) = m$$
$$\mathbb{V}(X) = \sigma_{eq} = \frac{\sigma^2}{2\kappa}$$

- Signal:

$$S = \frac{X - \mathbb{E}(X)}{\mathbb{V}(X)} = \frac{X - m}{\sigma_{eq}} = \frac{X - m}{\sigma/\sqrt{2\kappa}}$$



Model Adjustments

60 day trailing window to estimate *X* and parameters (*κ, m, σ*)

- reject if mean reversion time constant = 1/*κ* > 30 days
- adjust for bias in *m*

$$bias(\hat{m}) \approx \frac{1}{N_s}\sum_{s=1}^{N_s}\hat{m}_s$$

$$\hat{m}_{adj} = \hat{m} - bias(\hat{m})$$

- volume adjusted returns to reduce signal on high trading volumes.

Reference: Avellaneda & Lee, 2008

Midterm Recap (2003-2012)

$100k, static choice of 2003's top 100 largest-cap

Net return: 115.3%, Sharpe 0.99

We remained beta neutral as desired over the entire period, however our returns showed stagnation after 2008...

Post-2008: Why Stagnate?

- Avellaneda & Lee's paper published → strategy exploited, no more arbitrage! Greater competition for the same opportunities
- Avellaneda & Lee didn't even cover the impact or aftermath of crisis... perhaps the strategy is too simple to succeed in complex environments
- 2008 financial crisis → Changed market environment
- Small capital base, small universe → Not enough opportunities
- Struggling to take the other side of the spread (illiquid stocks)

Post-2008: Delving Down For Improvements

- Period: 2008-2012 focused training ||| 2013-2016 testing
- Capital base: $1 Million
- Costless transactions
  - Costs were obfuscating the predictive power of our signals
  - In practice, funds with large capital pools can trade via intermediaries to avoid driving up the prices for themselves

Post-Midterm Investigation

We have:

- Tested spreads for stationarity before applying our model (prev. discussed)
- Expanded the trading universe
- Addressed incompletely-filled orders (3 ways)
- Optimized signal cutoffs
- Optimized daily trading weights
- *Experimented with* mapping stocks to ETFs (pre-selecting regressors for stocks)

Expanded Universe (Training 2008-2012)

- Previously: fixed universe with top 100 market cap stocks in 2003
- Now: every year, update the universe to the top 500 market cap stocks
  - Max limit of 100 long and 100 short stocks at any one time
- Problem with larger universe (1000)
  - Illiquidity of small-cap stocks leads to frequent unfilled orders and bias in pair trading

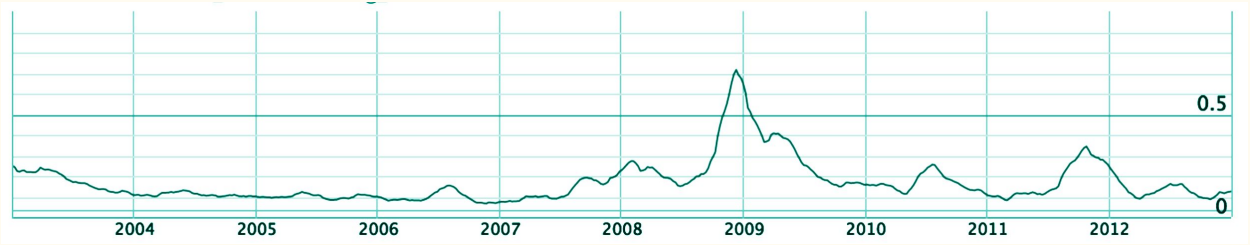
Universe Size	Max Limit	Open Cutoff	Close Cutoff	Sharpe Ratio	Net returns	Annual Return
100	20	1.2	0.6	0.57	21.4%	3.9%
500	100	1.2	0.6	1.23	42.1%	7.1%
1000	200	1.2	0.6	0.58	16.4%	3.0%

Unfilled Orders: Can’t fill our orders on time...

- Persist incompletely-filled stock orders on day 2, 3, etc. until they are filled
  - But violates our beta neutrality in the long term
  - Large position of ETF, small incremental positions in stocks
- Delay the purchase of ETFs to day 2
  - Based on how many shares were successfully filled on the 1st day, we will purchase corresponding ETFs on the 2nd day
  - Overcorrection, results in large-scale 1-day lag
- **Sell overbought or buy back oversold ETFs on day 2**
  - Make sure the consistency is restored on day 2

Cutoff Optimization (Training 2008 - 2012)

- Different combination of static open and close cutoffs
- Dynamic cutoff
  - Incorporates market volatility
  - Higher threshold in volatile market environment



- Conclusion: difficulty to find universally applicable parameters for drastically different pre- and post-crisis market landscapes; all combinations faced issues with lack of return after 2009

Cutoffs: Sensitivity Testing (Training 2008 - 2012)

Open Cutoff	Close Cutoff	Sharpe Ratio	Return	Beta
1.2	0.4	1.20	45.1%	0.00
1.2	0.5	1.18	42.2%	0.00
1.2	0.6	1.23	42.1%	0.00
1.3	0.4	0.86	27.5%	0.01
1.3	0.5	0.83	25.4%	0.02
1.3	0.6	1.01	30.1%	0.01
$1.2 + \frac{1}{5} * \text{sqrt}(\sigma)$	$0.55 - \frac{1}{5} * \text{sqrt}(\sigma)$	1.16	40.5%	0.0

Weights Optimization (Training 2008 - 2012)

- Assign weights proportional to the signal magnitude
  - Pairs with stronger signals will receive higher weights, and vice versa
  - Modest improvement
- Magnify the signal
  - Square the signal, etc.
  - Best Sharpe Ratio: 1.25
- Conclusion: lack of variation between signals above the threshold; magnifying the signal too much means lack of diversification

Variant: Use Sector ETFs as regressors?

- Previously: Forecast expected returns by regressing on market...

- What if we try regressing the stock on a set of logical ETFs?
  - “Regress out” the market from each stock’s returns
  - Regressions and F-testing/ANOVA to select the ETFs which have predictive power for that stock

Variant: Use Sector ETFs as regressors?

- 3 different approaches to map stocks to ETFs:

1. Map each stock to its sector ETF
   - i.e., Microsoft to HHH
2. Map each stock to the most significant sector ETF
   - Regress stock returns against ETF returns and find the ETF with positive beta and smallest p-value
3. Map each stock to all significantly correlated ETFs
   - Identify all ETFs with regression p-values < 0.05

Risk Management

- **Do not "reuse" signals**
  - Change our position only when we see the appropriate signal
- **In a crisis: starved of shorting / profit-taking opportunities**
  - Scale by VWAP
- **Leg risk when trading illiquid tickers**
  - Next-day adjustment to regain beta neutrality

Risk Management 2

- **Capital constraints**
  - Restrict # of longs and shorts that we could hold at any instant
- **Metrics:**
  - Overall beta (systematic risk)
  - Rolling beta (systematic risk)
  - Volatility
  - Max drawdown (worst-case downside risk)
  - VAR (tail risk)
  - Sharpe ratio (risk-to-reward ratio)
  - Sortino ratio (downside risk-to-reward ratio)

Best Models (Training 2008-2012)

2008 - 2012 (5 yr): No Transaction Costs

Universe	Max Stocks	Strategy	Performance Metrics		Risk Metrics		
			Annual Return	Sharpe	Beta	Volatility	Max Drawdown
100	20	Baseline (SPY)	9.5%	0.78	0.01	0.13	-14.4%
		Baseline (SPY) + stationary	4.0%	0.57	-0.01	0.07	-7.7%
		SPY + multiple ETF + stationary	4.1%	0.65	-0.01	0.07	-9.0%
		SPY + best sector ETF + stationary	5.7%	0.86	0	0.07	-10.2%
500	**100**	Baseline (SPY)	6.3%	0.67	0.01	0.1	-11.3%
		Baseline (SPY) + stationary	**7.3%**	**1.23**	**0**	**0.06**	**-8.4%**
		SPY + multiple ETF + stationary	2.6%	0.54	0.01	0.05	-12.2%
		SPY + best sector ETF + stationary	4.2%	0.84	0.01	0.05	-6.6%
1000	200	Baseline (SPY)	8.1%	0.92	-0.01	0.09	-14.6%
		Baseline (SPY) + stationary	3.1%	0.58	-0.01	0.06	-10.4%
		SPY + multiple ETF + stationary	3.6%	0.81	-0.01	0.04	-5.9%
		SPY + best sector ETF + stationary	3.4%	0.77	-0.01	0.04	-5.8%

Training (2008-2012)

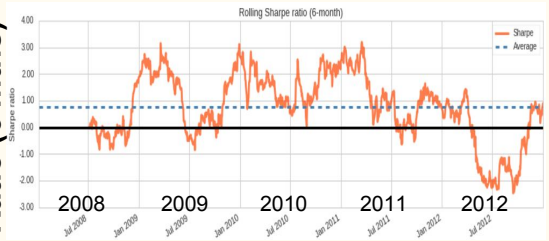
Baseline (SPY)

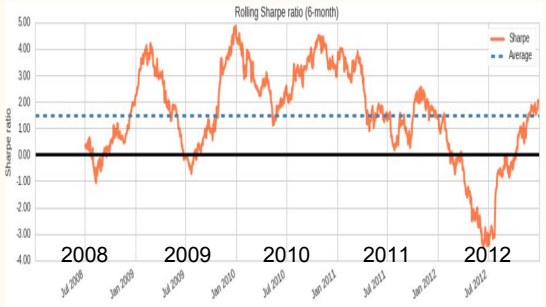
Baseline (SPY) + Stationary

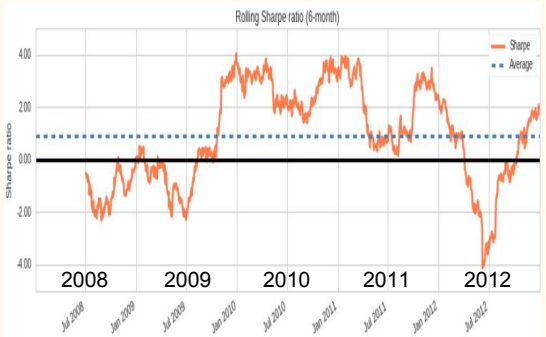
SPY + Multiple Sector ETF(s) + Stationary

Cumulative Return

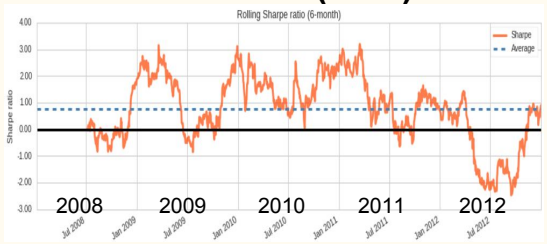
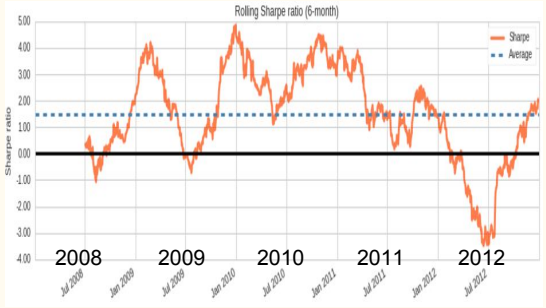
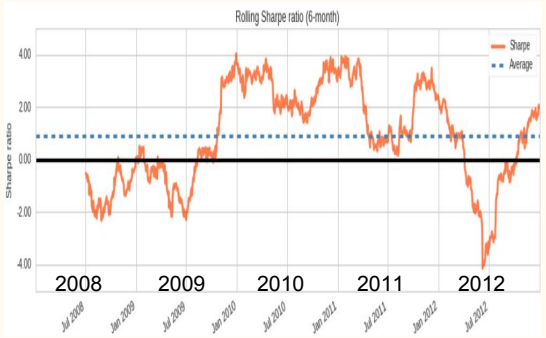
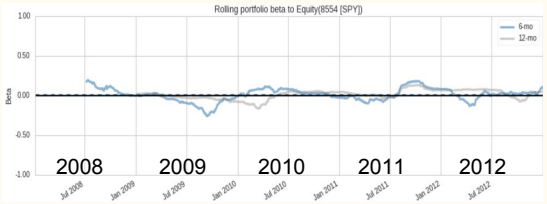
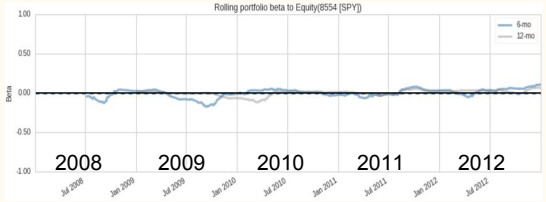
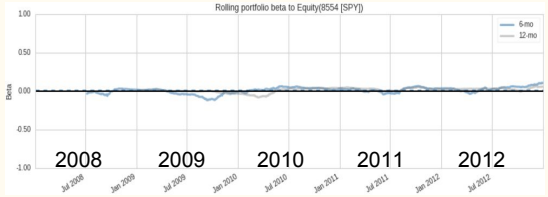
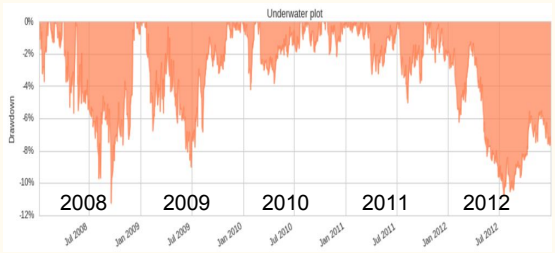
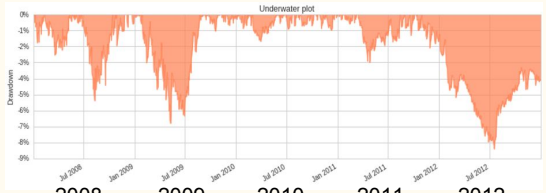
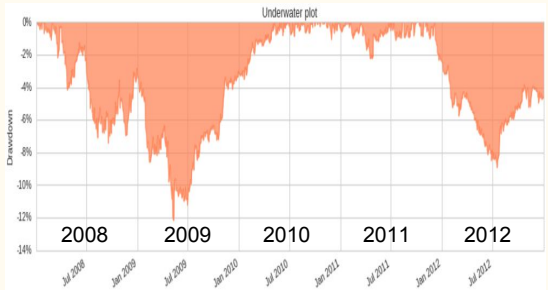
Rolling Sharpe Ratio (6 mths)







Training (2008-2012)

	Baseline (SPY)	Baseline (SPY) + Stationary	SPY + Multiple Sector ETF(s) + Stationary
Sharpe Ratio (6 mths)			
Beta (to SPY)			
Drawdown			

Out-of-Sample Testing (2013-2016)

2013 - 2016 (4 yr): No Transaction Costs

Universe	Max Stocks	Strategy	Performance Metrics			Risk Metrics			
			Annual Return	Sharpe	Sortino	Beta	Volatility	Max Drawdown	VaR*
500	100	Baseline (SPY)	2.2%	0.31	0.44	0.07	0.08	-18.7%	-0.9%
		Baseline (SPY) + stationary	**2.2%**	**0.54**	0.75	0.01	0.04	**-6.6%**	-0.47%
		SPY + Multiple ETF + Stationary	2.4%	0.59	0.84	0.01	0.04	-6.9%	-0.53%
		SPY + Best Sector ETF + stationary	2.5%	0.59	0.84	0.01	0.04	-6.9%	-0.53%

* VaR = bottom 5th percentile of daily returns

Out-of-Sample Testing (2013-2016)

Baseline (SPY)

Baseline (SPY) + Stationary

Cumulative Return

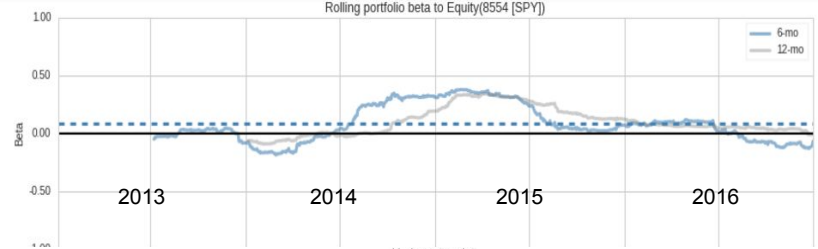
Sharpe Ratio (6 mths)

Out-of-Sample Testing (2013-2016)

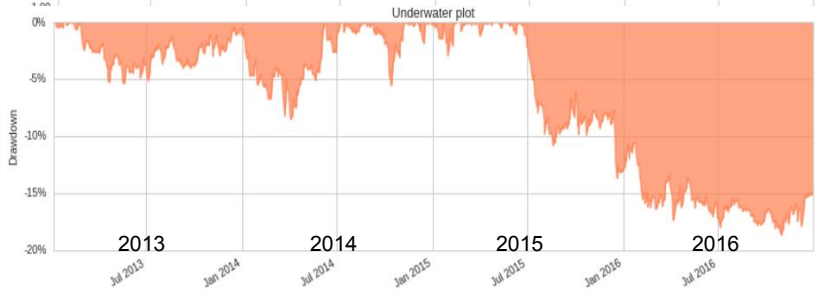
Baseline (SPY)

Sharpe Ratio (6 mths)

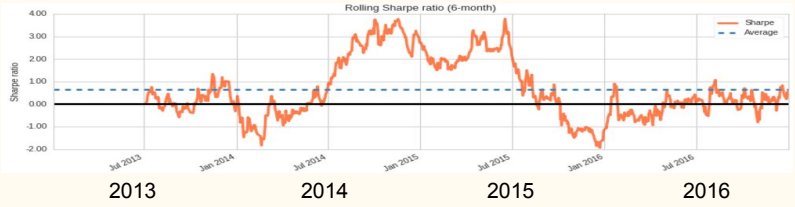
Beta (to SPY)

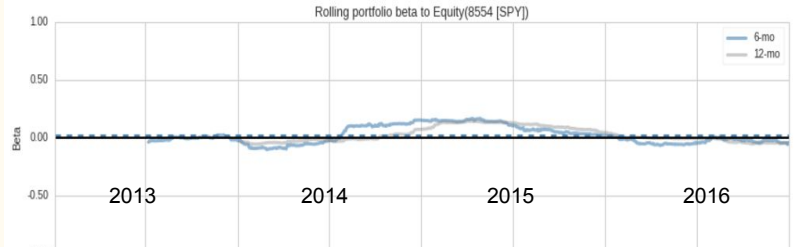


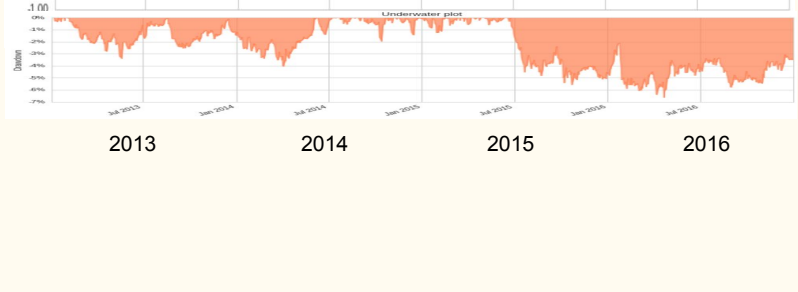
Drawdown



Baseline (SPY) + Stationary







Mediocre Performance --- Seeking Reasons Why

- Avellaneda & Lee ended their study before the crisis hit...
  - Maybe overfitted to pre-crisis period?
  - Or core change in how markets move / how traders think

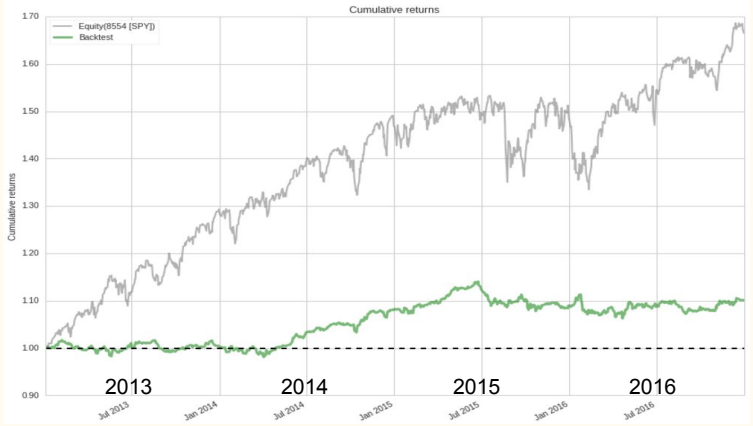
- 2012 Fed quantitative easing →Liquidity pushes up markets → Kill off short positions and opportunities

Sector ETFs as regressors? Slight improvement...

SPY + Sector ETF(s) + Stationary

Baseline (SPY) + Stationary

Cumulative Return



Rolling Sharpe Ratio (6 mths)

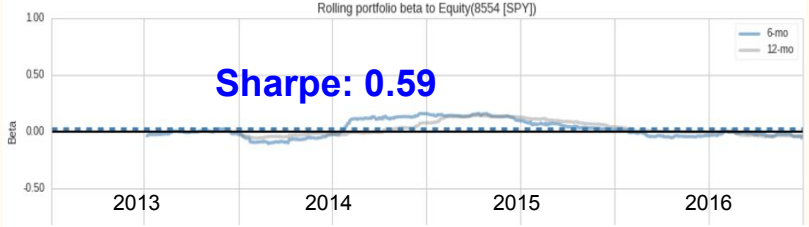
Sector ETFs as regressors? Slight improvement...

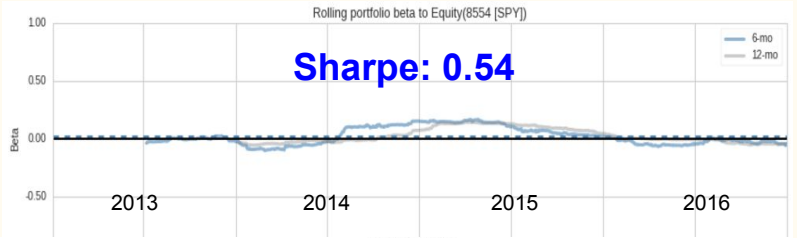
SPY + Sector ETF(s) + Stationary

Baseline (SPY) + Stationary

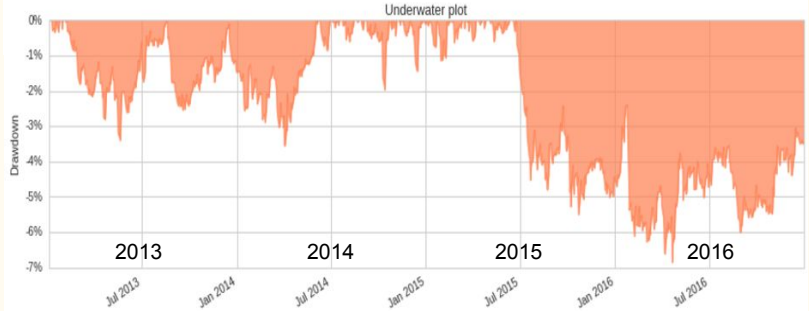
Sharpe Ratio (6 mths)

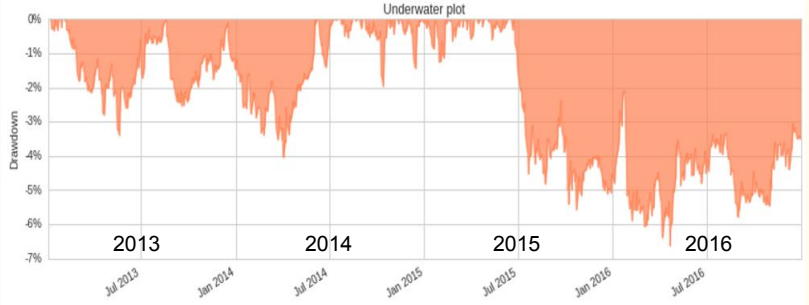
Beta (to SPY)





Drawdown





Sector ETFs as regressors? Slight improvement...

Probably because the predictive power of all sector ETFs with respect to the market-independent component of returns is very poor...

```
2008-01-02 05:45   PRINT [0.17615110675479895, 0.21292001791750881, 0.10559174965145179, 0.116387381742597,
0.31730883303349888, 0.10235640960478565, -0.010488595813865187, nan, 0.036133988537152439, 0.13471998665705565,
0.02932108880812645, 0.0046414121112593465, 0.23498668162787395, 0.26240133271311838, -0.037755924096271487,
0.35509295071839331, 0.00088601098821161628, 0.23389611259810905, 0.12135079373091151, 0.11478372422200467,
0.083619105180570807, 0.09699429389619818, 0.020022104719073286, 0.3263075507366523, 0.08321223116084675,
0.011565577663073756, 0.044936067735830587, -0.033971805441864289, 0.34545597351568824, 0.021359974103224433,
0.038183208506070709, 0.0016046261257525174, 0.038936656510531575, 0.039479713967984553, 0.012564491592312432,
0.12801146657203277, -0.034175387200893192, 0.024992386975435599, 0.053514303546109199, 0.17367130705382194,
0.17472771456431235, 0.1551551370572618, 0.14806486625086657, 0.19652735869744853, 0.051714765642686622,
0.024794706699998614, 0.12196745129070219, 0.094689040890982734, 0.098...
```

Adjusted R^2 values for various tickers regressed on all sector ETFs (2008-2012)

Risk Metrics

Universe	Max Stocks	Strategy	Performance Metrics			Risk Metrics			
			Annual Return	Sharpe	Sortino	Beta	Volatility	Max Drawdown	VaR*
500	100	Baseline (SPY)	2.2%	0.31	0.44	0.07	0.08	-18.7%	-0.9%
		Baseline (SPY) + stationary	**2.2%**	**0.54**	0.75	0.01	0.04	**-6.6%**	-0.47%
		SPY + Multiple ETF + Stationary	2.4%	0.59	0.84	0.01	0.04	-6.9%	-0.53%
		SPY + Best Sector ETF + stationary	2.5%	0.59	0.84	0.01	0.04	-6.9%	-0.53%

Conclusion

- With stationary check, great pre-2008 before crisis hit
- Stationarity check controls max drawdown, preserves profit in spite of unpredictable markets
- Not very robust when economic environment fundamentally changes - wildly unpredictable when we vary starting year
- Too slow, not competitive enough? Rise of HFT
- Need better factors, to know environment

Avenues for Exploration

- Alternative Factor Selection
  - A model which includes sector ETFs performs only slightly better
  - Simple linear regression: Too crude?
- More Robust Parameter Optimization Procedures
- Conditionally-Heteroskedastic / Non-Stationary Models

Q+A