

Risk Control of Mean-Reversion Time in Statistical Arbitrage

George Papanicolaou
Stanford University

CDAR Seminar, UC Berkeley

April 26, 2018

with Joongyeub Yeo

"Risk Control of Mean-Reversion Time in Statistical Arbitrage",
J. Yeo and G. Papanicolaou, Risk and Decision Analysis, vol 6, 2017,
p. 263-290.

Outline:

Statistical Arbitrage: It is a widely practiced general form of pairs trading

1. The Merton problem and pairs trading
2. Factor model for US equity returns and its implementation
3. Controlling the mean reversion of residuals
4. Trading: optimal allocation and PnL
5. Results with daily S&P500 data (2004-2014) exactly as they would be obtained in real time
6. Academic studies of statistical arbitrage: Andrew Lo¹ (MIT), and Marco Avellaneda² (NYU-Courant)

¹Hedge Funds: An Analytic Perspective, Princeton University Press, 2010

²Statistical arbitrage in the US equities market, Quantitative Finance, 2010

Mean-variance optimal allocation portfolio in continuous time

This is the now classic portfolio allocation problem formulated by Markowitz (1952), and by Merton (1972) in continuous time.

S_t is the price of a risky asset, $\frac{dS_t}{S_t}$ is the (infinitesimal) return assumed to be Brownian motion with drift

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t, \quad \mu = \text{mean return}, \quad \sigma = \text{volatility}$$

Portfolio value is W_t and a fraction u_t goes into the risky asset while $1 - u_t$ into a bank account with interest rate r . This leads to the portfolio returns

$$\frac{dW_t}{W_t} = u_t \frac{dS_t}{S_t} + (1 - u_t)r dt = ((1 - u_t)r + \mu u_t) dt + u_t \sigma dB_t$$

The problem is to find the optimal allocation ratio u_t^* that maximizes the expected return $E\{U(W_T)\}$, where U is a given utility function

Merton portfolio, log utility, zero mean return

Optimal allocation ratio u^* is a constant (fundamental portfolio theorem) and wealth log growth rate are

$$u^* = \frac{\mu - r}{\sigma^2}, \quad r + \frac{(\mu - r)^2}{2\sigma^2} \quad (= r + \text{Sharpe ratio squared}/2)$$

Since $d \log S_t = (\mu - \sigma^2/2)dt + \sigma dB_t$, when $\mu = \frac{\sigma^2}{2}$ mean log return of stock is zero. Then

$$u^* = \frac{1}{2} \left(1 - \frac{2r}{\sigma^2}\right)$$

and portfolio log growth rate is

$$r + \frac{\sigma^2}{8} \left(1 - \frac{2r}{\sigma^2}\right)^2$$

If $\frac{2r}{\sigma^2}$ is small and $\sigma > .5$, say, then the constant portfolio allocation strategy will realize considerable growth above the basic interest rate even when the stock or index is not growing at all. The portfolio allocation ratio in stock will never be more than 50%. The best environment for this kind of portfolio strategy is: low real interest rates and highly fluctuating stocks that have no (log) growth.

Pairs trading: Realizing pure fluctuations

Suppose P_t and Q_t are the prices of two stocks (or a stock and an index or ETF) that are strongly correlated, two energy stocks, two bank stocks, a technology stock and a technology sector ETF, ...

Regress their returns over some time window (sixty days, for example) so that

$$\frac{dP_t}{P_t} = \beta \frac{dQ_t}{Q_t} + X_t$$

(Rough) Trading signal: The residual X_t . If it is positive then Q is momentarily undervalued relative to P ($\beta > 0$). Action: Long (buy) Q and short (sell) P . When later X changes sign, switch position ...

(Invented in the early nineties at Morgan-Stanley (M. Avellaneda)).

Compare with Merton portfolio when there is no underlying growth. Pairs trading **creates** such a context by focusing on the residual.

Role of mean reversion of the residual, an issue not visibly present in the Merton problem. **Andrew Lo: Betting on mean reversion.**

Generalized "pairs" trading: Factor decomposition

We start with the $N \times T$ matrix of stock returns, R and decompose it into factors:

$$R = LF + U \quad (1)$$

where F is a $p \times T$ matrix of factors L is an $N \times p$ matrix of factor loadings, U is a $N \times T$ matrix of residuals.

Only R is observable, so L , F , and U must be estimated. We obtain factors using principal components analysis (PCA) so only their number p is to be determined. Next, the factor loadings (L) are estimated by multivariate regressions of R on the subspace spanned by F . Finally, the residual (U) is expressed as the remaining parts, after subtracting the estimated common factors from the original returns:

$$\hat{U} = R - \hat{L}\hat{F}. \quad (2)$$

where the hat ($\hat{\cdot}$) notation indicates estimators, since only R is observable.

All this in a time window that is rolling forward in time.

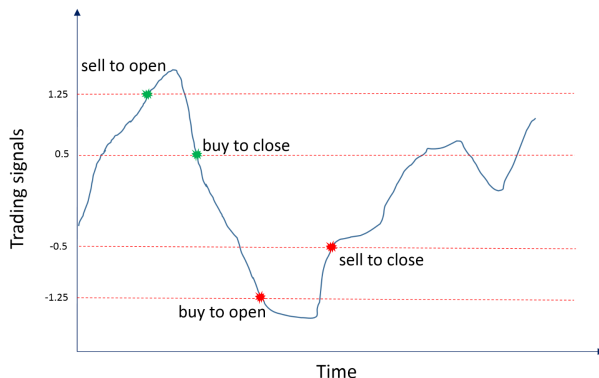
What do we know about the factors?

- In Economics: The three-factor models for stock returns of French and Fama (1992).
- In Econophysics: How to extract factors using PCA so that the residuals have a correlation matrix that has Marchenko-Pastur (purely random) statistics. Many papers in late nineties. Survey of Bouchaud and Potters (2009) presents the ideas and methods in perspective.
- Some startling empirical facts: The principal (left) singular vector of the return matrix R when ordered by size lines up with the market capitalization of the stocks.
- The correlations of returns "uncover" market capitalization and market sectors without knowing anything other than the (time series of the) returns themselves.

The trading strategy and mean reversion of residuals

- Since the "Market" has been taken out by going to residuals, it is essential that these residuals have statistical regularity.
- Market risk has been transformed into risk of "residual irregularity over time", that is, risk from anomalies in the assumed mean reverting behavior of the residuals.
- Investment portfolios are not bets on returns normalized by volatility (Sharpe ratios). They are now bets on mean reversion of residuals (A. Lo).
- Integrated (summed) residuals are modeled by OU processes (discrete time): $dX = -\kappa X dt + \sigma dW$.
- Estimate the rate of mean reversion κ over a past (rolling) window using MLE or least squares for each residual.
- Use residuals for trading only when the estimated mean reversion time is relatively short and the error in the estimation is not too large. This reduces risk substantially!

The long/short trading strategy with normalized residuals as signals: Buy low - sell high



When the trading signal hits 1.25, it is presumed to be over-valued, so we sell the stock to open the short position (green dots), expecting it to go down soon. If it hits 0.5, we buy the stock to close the short position, which leads us to obtain profits. Similar strategy can be applied to the open/close of long position (red dots).

The trading algorithm: Yes, buy low and sell high. But with what allocation?

The dollar amount q_i , $i = 1, 2, \dots, N$ (signed for long/short) to invest is determined at each time t by the following:

$$\text{Market-neutrality : } \sum_i L_{ik} q_i = 0, \text{ for each } k=1, \dots, p \quad (3)$$

$$\text{Dollar-neutrality : } \sum_i q_i = 0 \quad (4)$$

$$\text{Total Leverage : } \sum_i \|q_i\| = I \quad (5)$$

$$\text{Long/Short : } \text{sgn}(q_i) \text{ is given by trading signals.} \quad (6)$$

Here L is the factor loading matrix, p is the number of factors, and I is the total leverage level.

Profit and Loss

The performance from trading can be measured by the cumulative profit-and-loss (PnL):

$$E_{t+\Delta t} = E_t + E_t r \Delta t + \sum_{i=1}^N q_{it} R_{it} \Delta t - \sum_{i=1}^N q_{it} r \Delta t - \sum_{i=1}^N |q_{i(t+\Delta t)} - q_{it}| \epsilon \quad (7)$$

E_t is the value of the portfolio at time t ,

r is the reference interest rate,

ϵ is for frictional cost (measured in basis points).

Note that the fourth term on the right hand side must be zero if the dollar neutral condition is completely satisfied. However, there is numerical tolerance imposed for the feasibility, so we keep that term here.

The US equities data used in backtesting

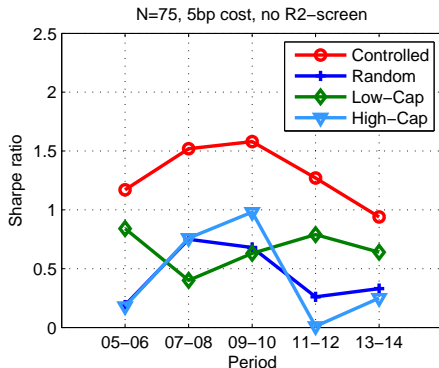
- SP500 daily data from 2005 to 2014, split into 5 time periods: 2005-2007 (before crisis), 2007-2008 (in crisis), 2009-2010 (crisis recovery), 2011-2012 (post crisis I), 2013-2014 (post crisis II).
- Use only 378 stocks from the SP500 to have uninterrupted records and avoid necessary preprocessing of data.
- Use 30-120 day rolling windows in the analysis
- Use 1-20 PCA factors in generating residulas

Avellaneda (2010 and later) uses a much larger set of equities, beyond the SP500 (the over one billion capitalization group). He also uses ETFs for factors, both real and synthetic ones.

Risk control with estimated mean reversion time of residuals

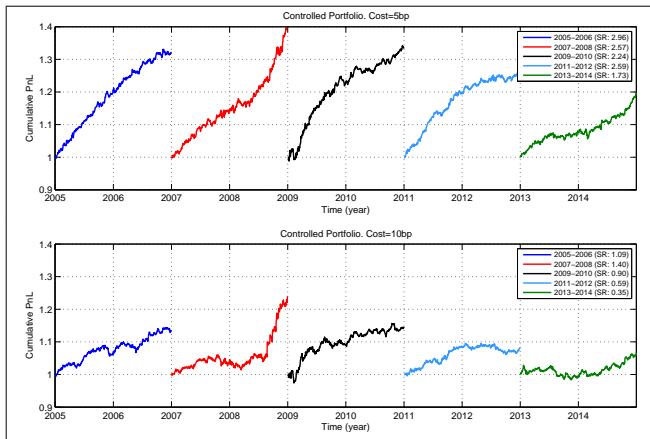
- Since risk has migrated into the mean reverting behavior of the residuals, we need to control it through the estimates of mean reversion times.
- The error (in MLE) is proportional to the reciprocal square root of the time window. And we cannot use large time windows because information in the residuals becomes "stale" (non-stationary).
- Typical tradeoff in applied science (channel estimation, geosciences, biology, genomics, etc): We want to use a large enough data window to have statistical stability but also one that is small enough to have approximately stationary statistical properties.
- Bottom line: Use residuals that have (a) relatively short mean reversion times and (b) relatively high statistical confidence score: Use diversity by proper selection in a high-dimensional setting.

The effect of mean-reversion control on Sharpe ratios



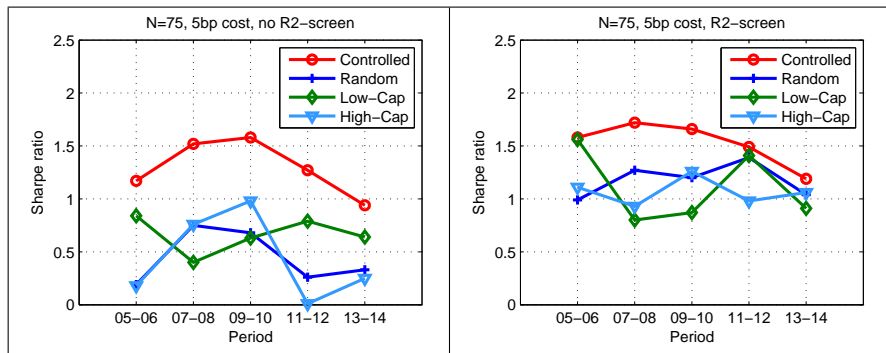
The size of portfolio is 75, the transaction cost is 5bp per volume, and R^2 -screening is not applied here. It is clear that the portfolio selection using the estimated mean-reversion speed (red line) performs better than the other portfolios.

Sample results of cumulative PnL for each period; Controlled



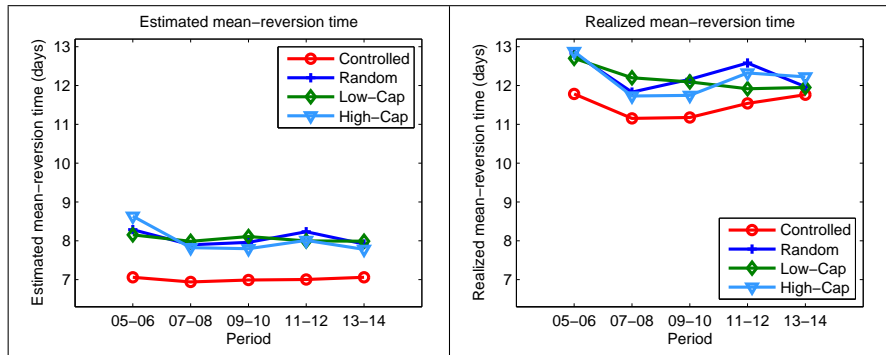
Two transaction costs (of 5bp and 10bp). Portfolio size is $N = 75$; Number of factors is 5; Window length is 30 days. R^2 -screening is applied.

The out-of-sample performance from four different portfolios



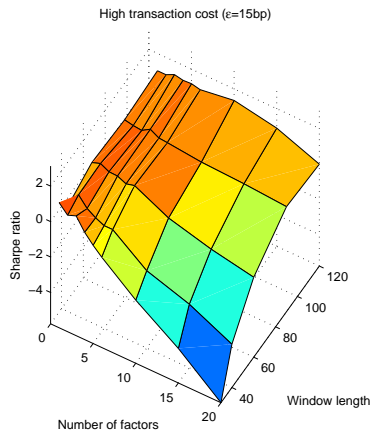
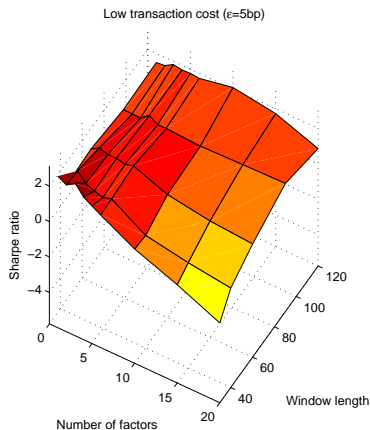
Number of selected stocks is 75; Transaction costs are 5bp; Left: Without R^2 -screening. Right: With R^2 -screening. The overall level of performance is enhanced by selective trading based on R^2 -screening.

Estimated and realized mean-reversion times (in days)



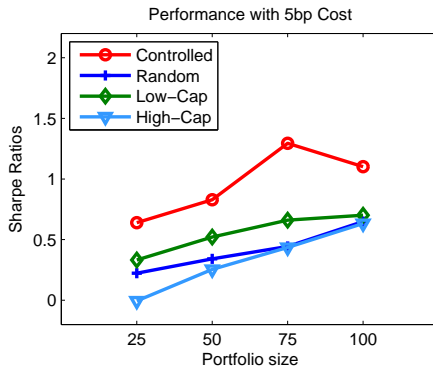
Estimated (Left) and realized (Right). The mean-reversion controlled portfolio has significantly shorter trading intervals compared to other three portfolios. (The values are averaged).

The performance surface as a function of number of factors (p) and window length (T)



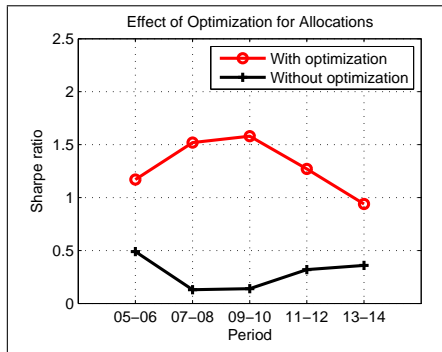
The performance surface as a function of number of factors (p) and window length (T), for two different transaction costs.

Sharpe ratios with respect to portfolio sizes



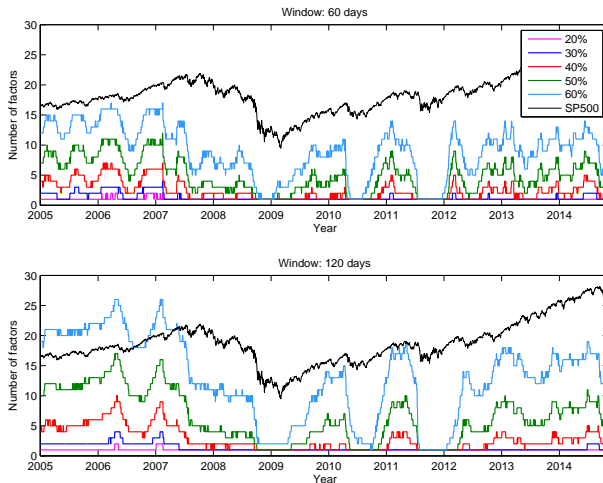
The horizontal axis is portfolio sizes (25, 50, 75, and 100 stocks). The vertical axis is Sharpe ratios. The results are presented with transaction costs of 5bp. Averaged results over all number of factors, estimation windows, and periods. In general, the performance is better with larger portfolio sizes. However, for controlled portfolio, using $N = 75$ is better than using $N = 100$.

Performance with and without optimization for asset allocation



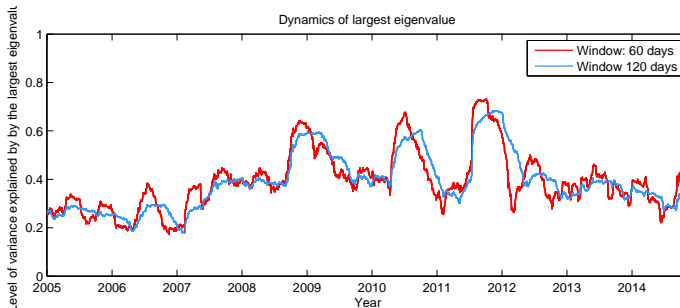
Without optimized asset allocations, the trading strategy is unstable and is not market-neutral.

Dynamics of number of factors that explain different levels of variance



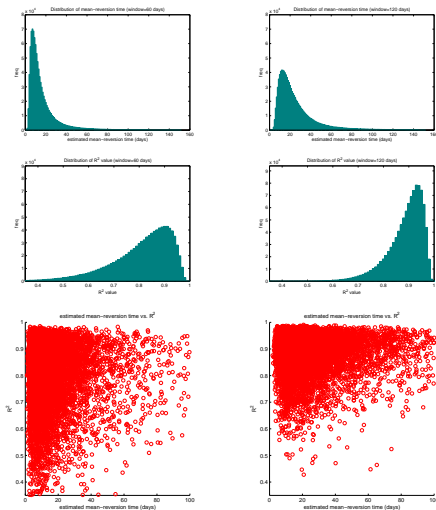
Window: 60 days (top) and 120 days (bottom). $N = 378$ stocks.

The largest eigenvalue of correlation matrix of normalized original returns with moving window



The shorter windows response is more sensitively to the market.

Statistics of mean reversion time and R^2 values (20 factors)



Top: Mean-reversion times. Middle: R^2 values. Bottom: Scatter plots for the two. Window is 60 (Left) and 120 days (Right).

Concluding remarks

- A carefully implemented statistical arbitrage portfolio has remarkably stable returns even in very hostile financial environments.
- This comes out using real data exactly as they would be used in practice (out of sample).
- You do not need to know anything about markets or finance. It is a purely quantitative approach. Just "signal processing".
- An innovation in this study: Mean reversion based control and optimal allocation of trading amounts. Otherwise we follow the Avellaneda-Lee 2010 analysis.
- A very large number of deep and largely untouched mathematical problems come up, in the factor analysis, in the mean reversion analysis, and in the optimal allocation algorithm...
- The basic idea of the trading strategy is simple and intuitive but the details in the implementation are all important and make a huge difference in performance.
- Statarb is not meant to be used exclusively. It should be used together with traditional (long only) value portfolios and other diversified instruments in a carefully integrated strategy.