Statistical Arbitrage with Sector ETFs and Index Arbitrage

Ventseslava Encheva, Jen Tsung Hu, Nina Kuklisova,
Swarna Ramineni, Henry Shin *

ABSTRACT

In this project, we implement two trading strategies. One is a statistical arbitrage strategy with sector ETFs. The other one involves S&P 500 index additions and deletions. Our statistical arbitrage strategy with sector ETFs utilizes the multiple linear regression model and outperforms the S&P 500 index.

In the S&P 500 index additions and deletions strategy, we devise a set of trading strategies which seek profitability based on event-driven arbitrage. The strategies exploit price movements from stock additions to the S&P 500 Index.

*The Final Project for the Mathematics of Finance Course was implemented under the guidance of Professor Mikhail Smirnov, Mathematics Department, Columbia University ...

I. Statistical Arbitrage with Sector ETFs

A. Introduction

Our statistical arbitrage strategy is based on the fact that a sector ETF cointegrates with its constituent stocks. In this strategy, we construct a multiple linear regression model using a subset of the constituents and the ETF. The constituents will be used as predictors of the ETF.

We pick the top 5 holdings of the ETF as our basket. The higher the number of the constituents, the better the basket cointegrates with the ETF, but the smaller the profits will be. Having a smaller set of basket also reduces the capital requirement. With the multiple linear regression model, we long the ETF and short the basket when the ETF price is smaller than the predicted ETF price by more than 2 standard deviations. We short the ETF and long the basket when the ETF price is larger than the predicted ETF price by more than 2 standard deviations.

Our stop loss strategy is that we close our position if the ETF/basket divergence does not mean revert after 10 half-lives time. The half life is calculated based on the assumptions that the spread follows the Orenstein-Uhlenbeck model, which means the spread mean reverts, and the spread decays exponentially.

B. Strategy Models:

We assume the ETF and the constituents follow a multiple linear regression model:

$$y = b_0 + b_1 \times x_1 + b_2 \times x_2 + b_3 \times x_3 + b_4 \times x_4 + b_5 \times x_5 + \epsilon;$$

where y is the ETF price, x_2, x_3, x_4, x_5 are the top 5 constituents. Using the least squares method we construct our predictor:

$$\hat{y} = \hat{b_0} + \hat{b_1} \times x_1 + \hat{b_2} \times x_2 + \hat{b_3} \times x_3 + \hat{b_4} \times x_4 + \hat{b_5} \times x_5; \tag{1}$$

We also assume that the spread $y - \hat{y}$ follows the Orenstein-Uhlenbeck model:

$$dz(t) = \theta \times (z(t) - u) + dW \tag{2}$$

where $z(t) = y(t) - \hat{y}(t); u$ is the mean of $z(t)$; dW is random noise; θ can be found with a simple linear regression between $z(t)$ and $z(t) - u$. The half life can then be calculated as:

$$T_{1/2} = -\frac{\ln(2)}{\theta}.$$

The first assumption is intuitive because an ETF price is a linear combination of its constituent prices. The second assumption is true if the spread mean reverts. The half life calculation is based on the assumption that the spread decays exponentially which serves only as a stop loss control.

C. *Strategy Implementation*

Our data source is Yahoo Finance. We download the daily adjusted prices via an R program:

```
xle = read.csv("http://ichart.finance.yahoo.com/table.csv?
s=xle&ignore=.csv",stringsAsFactors=F)
```

We use an in-sample data set of 252 days of close prices of an ETF and its top 5 constituents to fit our model, equation (1). We apply the F-test to ensure the overall model is at least 95% statistically significant. We also apply the T-test to make sure each constituent predictor is at least 95% statistically significant.

Due to the fact that we only use 5 of the constituent set which often consists of several dozen stocks, the residuals are not normally distributed. However, our strategy is not based on any normal distribution. Furthermore, the model can be improved by including more constituents. However, this will make the ETF and the basket more tightly correlated and decrease profits. In essence, by incorporating fewer predictors we are taking on more risk in exchange for higher returns.

Finally, we use Augmented Dickey-Fuller test on the residuals to check for cointegration with at least 95% confidence. If the residuals fail the test, we will refrain from trading. We refit the model daily until cointegration resumes.

Given a satisfactory fit, we start out-of-sample trading. We use the fit to predict the ETF price and calculate the zscore, $(y(t) - \hat{y}(t))/\hat{\sigma}$, where $\hat{\sigma}$ is the estimated standard error of the model. We long the ETF and short the basket when the ETF price is smaller than the predicted ETF price by more than 2 standard deviations. We short the ETF and long the basket when the ETF price is larger than the predicted ETF price by more than 2 standard deviations. We refit the model every 6 months and repeat the same process.

After we initiate a position, we wait for the spread to mean revert to zscore equal to 0 to close the position. If the spread does not mean revert after 10 half lives time (2 - 3 months), we close the position.

One major advantage of our strategy is that it is parameterless in the sense that we do not optimize any parameter value to maximize performance. This way we completely avoid the possibility of model overfitting and data snooping.

D. *Trading Results*

We apply our strategy to 9 SPDR sector ETFs. Our initial investment is 1,000 \$. We show how it changes from the inception of the ETFs to date and show some performance metrics.

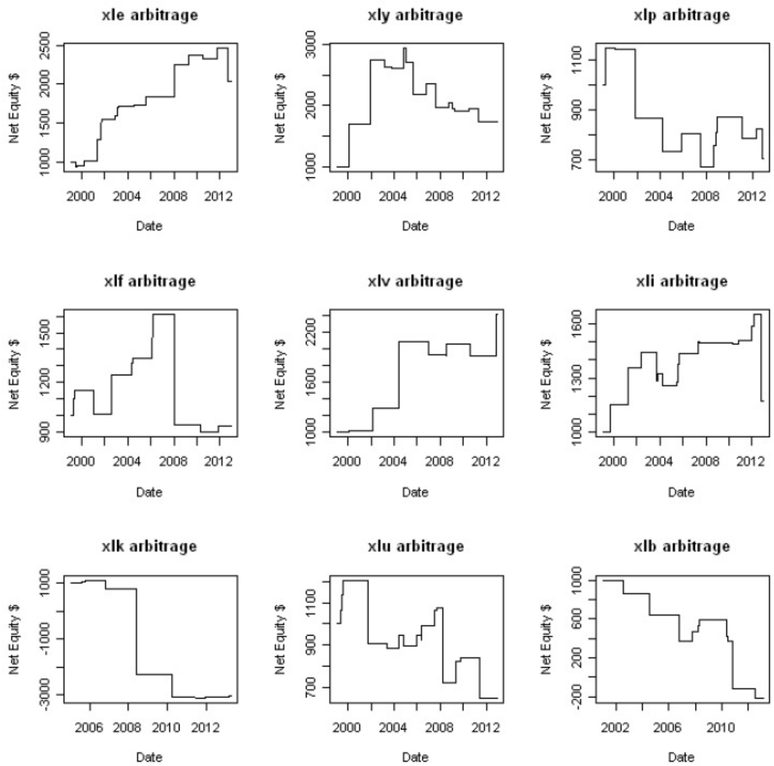

Figure 1. Results of trading by statistical arbitrage

	index	startDate	annReturn	spyReturn	annStdev	sharpe	spySharpe	drawdown	spyDrawdown	trades	holdDays	perctWinner
1	xle	19991222	0.051312784	0.03436523	0.1119831	0.45821916	0.1633712	-0.16964773	-0.5519296	17	21.0	0.8125000
5	xlv	19991223	0.063554654	0.03328902	0.1597581	0.39781792	0.1582361	-0.08190848	-0.5519296	11	81.0	0.6000000
2	xly	19991223	0.039834962	0.03329855	0.2146965	0.18554080	0.1583293	-0.40868355	-0.5519296	15	49.0	0.4285714
6	xli	19991222	0.011530535	0.03435541	0.1194260	0.09654966	0.1633792	-0.28956798	-0.5519296	18	25.5	0.7058824
7	xlk	20050818	0.007176437	0.06797818	0.1601949	0.04479815	0.3115451	-3.81322676	-0.5519296	9	28.0	0.5000000
4	xlf	19991222	-0.004938698	0.03435541	0.1681409	-0.02937237	0.1633792	-0.44170541	-0.5519296	12	49.5	0.7272727
9	xlb	20011030	-0.036674911	0.06341504	0.2815139	-0.13027746	0.3069577	-1.21517000	-0.5519296	11	81.0	0.3000000
8	xlu	19991222	-0.031059621	0.03441438	0.1596657	-0.19452902	0.1635763	-0.46092532	-0.5519296	17	19.0	0.6250000
3	xlp	19991222	-0.024954003	0.03441438	0.1275297	-0.19567207	0.1635763	-0.41244119	-0.5519296	13	81.0	0.5000000

Figure 2. Results of trading by statistical arbitrage

From the above trading results, we can see that XLE and XLV are the top 2 performers.

We can further improve the performance by constructing a portfolio with XLE and XLV which will have the highest sharpe ratios in such a way that the combined portfolio has minimum volatility, i.e. we minimize

$$\sigma_p^2 = w_{XLE}^2 \times \sigma_{XLE}^2 + w_{XLV}^2 \times \sigma_{XLV}^2 + 2\rho w_{XLE} \times \sigma_{XLE} \times w_{XLV} \times \sigma_{XLV}$$

With $\rho = -0.0008$, we get $w_{XLE} = 0.67$ and $w_{XLV} = 0.33$. The low correlation of XLE and XLV arbitrage returns results in enhanced performance:

	XLE + XLV	SPY
return	5.53%	3.43%
stdev	9.16%	11.20 %
sharpe	0.60	0.16
max drawdown	-11.37%	-55.19%

E. Conclusion

Our statistical arbitrage strategy outperforms our benchmark SPY both in terms of annualized return, sharpe ratio, and maximum drawdown. One desirable characteristic of our strategy is that the correlation between the returns of different ETF arbitrage is low. This makes it possible to construct a portfolio of multiple arbitrage implementations with different ETFs and enhance performance.

One caveat with the strategy is that trading opportunity is infrequent. However, the average holding period is around 3 months. It is possible to combine with other strategies that have shorter holding periods to more efficiently utilize capital. Another way to increase trading opportunity is to apply the same methodology to intraday prices.

II. Index Arbitrage based on Additions and Deletions in S&P Indices

A. Introduction

S&P Game represents arbitrage based on temporary price imperfections and corrections which occur as a result of announced additions or deletions of stocks to the S&P500 index. This kind of arbitrage was explored by Honghui Chen for S&P 500 (1), Rajesh Chakrabarti for MSCI (2) and Shinhua Liu for Nikkei 225 (3). We base our Strategy on the published article Intraday and Night Index Arbitrage (4) which gives empirical evidence of abnormal returns on additions to S&P 500 Index between 1999 and 2002 . Based on the conclusions of this article, we created multiple trading strategies for the suggested price adjustment periods around announcement (AD) and effective (ED) dates and tested our hypothesis on a stock additions sample covering the period January 2003 and October 2013. We used the additions between January 2003 and October 2013 to test our strategies because of the the information act which made the information on index additions available to all retail investors post 2002. The temporary over-pricing of additions happens when the index arbitrageurs game the ETF managers. They are required to buy added stocks only during closing hours of the day prior to the effective date to avoid tracking error even though the announcement of the addition is made public 3 - 5 days before the effective date.

Index funds have an objective to minimize tracking error and therefore, they must buy (sell) the stock only at the time of its addition to (deletion from) the index.

However, with the growth in indexing and ETF launches, orders from index funds at the opening bell have created order imbalances and volatility. As a result from October 1, 1989 onward S&P initiated preannouncement process of all stock additions to alleviate the traction and increased volatility around the effective date of each added stock. The gap between preannouncement and effective inclusion in the index leads to temporary over-pricing of additions a phenomenon usually exploited by arbitrageurs who attempt to game the ETF managers. An excess return upon an index change, in the absence of alternative explanations is justified through the imperfect substitutability or a downward sloping demand curve. Other suggested hypotheses for the price increase and respectively abnormal returns include increased investor awareness hypothesis, also called the shadow cost hypothesis, and the liquidity effect hypothesis.

B. Data

- Index Selection for Case Study

  We selected S&P500 for our study for the following reasons

  1. There are 3 ETFs benchmarked to it. They have over $ 200bn of AUM.

• Symbol	Name	Assets *	Average Vol	YTD
SPY	SPDR S&P 500 ETF	$163,348,928	112,722,555	+27.10%
IVV	iShares Core S&P 500 ETF	$51,597,427	4,270,880	+27.12%
VOO	Vanguard S&P 500 ETF	$13,675,253	1,326,135	+27.09%

Source: ETF Database (http://etfdb.com/) Note: Assets as of Dec 2, 2013

Figure 3. ETFs used in this trading strategy

  2. There is a lag between the announcement date and effective date.
  3. The added stocks are usually highly liquid, tradable, keeping the transaction costs low.

- Data Selection

  We used the additions from January 2003 to October 2013. Initial sample set of 232 additions was obtained and further reduced based on the number of trading days between the announcement and the effective dates according to the requirements of each tested Strategy.
- Datasources

  All price data and effective dates of additions was obtained from Bloomberg API. The announcement dates were programmatically derived from Market Watch news wires and via Launchpad news functions in Bloomberg.
- Data Processing

  All data processing was done using excel. You will find the calculation sheets attached with the project material.

C. Methodology

Following the paper we test the following strategies Strategy 1: Enter long position at the opening price of immediate trading day after the announcement day which we denote as the AD+1 Open. All announcements are made after the trading hours on AD making AD+1 Open the earliest possible entry option. We exit on the closing price of the day which is 2 days prior the effective date which we shall denote as ED-2 Close. Strategy 2: Enter long position on AD+1 Open and exit on ED-1 Open. (The returns were better on ED-1 Open as compared to those on ED -1 close since the fund managers take their necessary positions just before ED-1 Close) Strategy 3: Enter short position on ED Open and exit at ED + 2 Close. This was an attempt to capture the price inversion effect. Strategy 4: Combine Strategy 2 and 3 to accentuate the returns and diversify risk. Strategy 5: Some stocks showed major information leakage. We define information leakage based on the return between AD-1 Close and AD 0 Close. If this return is beyond 2.5% we follow only the shorting Strategy 3, else we follow Strategy 4. The threshold was selected based on average daily price moves of major stocks like Apple, IBM etc.

D. Results

Based on negative annual returns, Strategy 1 was excluded from further analysis: (All the returns are log returns.)

Year	Returns on Strategy 1	Returns on Strategy 2	Returns on Strategy 3	Returns on Strategy 4	Returns on Strategy 5
2003	-0.68%	-0.37%	-6.61%	-6.99%	-6.99%
2004	1.38%	3.26%	2.44%	5.69%	5.69%
2005	-6.05%	-4.99%	-1.11%	-6.10%	-6.10%
2006	24.67%	21.07%	6.62%	27.69%	22.45%
2007	5.54%	9.81%	-1.70%	8.10%	8.10%
2008	4.74%	12.62%	43.23%	55.85%	54.84%
2009	-2.97%	2.18%	6.39%	8.57%	16.37%
2010	-9.20%	-6.34%	9.23%	2.89%	3.85%
2011	-2.26%	-2.59%	5.16%	2.57%	1.95%
2012	-9.79%	-6.97%	-8.06%	-15.03%	-15.03%
2013	-6.34%	-5.16%	-5.35%	-10.51%	-10.51%
Total	-0.96%	22.52%	50.23%	72.75%	74.63%

Figure 4. Annual Returns

2003 - 2013	Strategy 1	Strategy 2	Strategy3	Strategy 4	Strategy 5
Average Returns	(0.087%)	2.048%	4.566%	6.614%	6.785%
St deviation	0.0967	0.0903	0.1408	0.2004	0.1953
Sharpe Ratio	(0.0090)	0.2267	0.3243	0.3301	0.3475

Figure 5. Annual returns analysis between 2003 and 2013

The highest number of additions in 2008 was coupled with highest returns across all implemented strategies. The high turnover of stocks was explained though the high level of turnover in the corporate structure of firms during the crisis period including hostile takeover, bankruptcies, etc. For instance, the Seattle-based Washington Mutual Inc. (NYSE: WM) was removed from the S&P 500 index in September 2008 after the bank was closed by the Office of Thrift Supervision and taken over by the Federal Deposit Insurance Corp. Standard & Poor's also removed Freddie Mac from its S&P 500 indexin September 2008 and replaced it with Salesforce a company largely

2005 - 2009	Strategy 1	Strategy 2	Strategy 3	Strategy 4	Strategy 5
Average Returns	0.2275%	0.326%	0.427%	0.753%	0.765%
St deviation	0.0216	0.0211	0.0252	0.0342	0.0340
Max DD	(17.82%)	(17.41%)	(12.84%)	(15.15%)	(15.15%)
Sharpe Ratio	0.1052	0.1544	0.1694	0.2205	0.2254

Figure 6. Annual returns analysis between 2005 and 2009

Year	Strategy 1	Strategy 2 to 5
2003	4	4
2004	15	15
2005	11	12
2006	23	27
2007	27	30
2008	29	32
2009	23	23
2010	15	16
2011	14	17
2012	17	17
2013	12	12
Total	190	205

Figure 7. Number of eligible additions for each strategy based on the lag between AD and ED

responsible for the growing software-as-a-service movement, and in important player in the broader, emerging market of cloud computing.

2003 - 2013	Strategy 1	Strategy 2	Strategy 3	Strategy 4	Strategy 5
Average Returns	(0.005%)	0.109%	0.244%	0.353%	0.362%
St deviation	0.0197	0.0195	0.0235	0.0322	0.0322
Max DD	(42.06%)	(31.63%)	(23.43%)	(39.34%)	(38.66%)
Sharpe Ratio	(0.0025)	0.0561	0.1038	0.1095	0.1127

Figure 8. Per addition returns analysis from 2003 to 2013

E. Conclusions

- Our best strategy was Strategy 5 which implied that diversification and conditioning can result in better Sharpe ratios. Smoothing of drawdowns is also clearly visible
- Both our long only and short only strategies eroded post 2009 although the shorting strategy persisted till a little later. This was a result of use of sophisticated lead indicators like momentum and index methodology factors which predicted the additions prior to the announcement. S&Ps index methodology is made publicly available and researchers can forecast additions much earlier than announcement date. When evaluating companies for inclusion on the S&P 500, Standard & Poor's considers such things as corporate structure, sector representation, accounting standards and exchange listings, financial viability, and whether the company has

2005 - 2009	Strategy 1	Strategy 2	Strategy 3	Strategy 4	Strategy 5
Average Returns	0.2275%	0.326%	0.427%	0.753%	0.765%
St deviation	0.0216	0.0211	0.0252	0.0342	0.0340
Max DD	(17.82%)	(17.41%)	(12.84%)	(15.15%)	(15.15%)
Sharpe Ratio	0.1052	0.1544	0.1694	0.2205	0.2254

Figure 9. Per addition returns analysis from 2005 to 2009

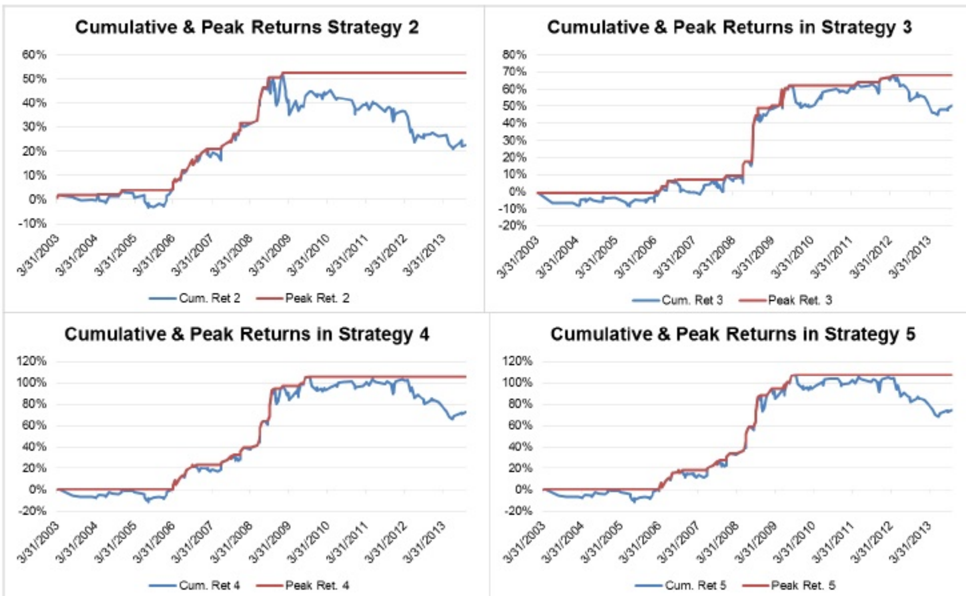

Figure 10. Cumulative returns and peak returns

adequate liquidity and a reasonable stock price, according to the S&P Web site. At least half of an S&P 500 company's outstanding shares must be publicly owned.

- The index arbitrage strategies based on stock additions worked really well during the 2008 crisis due to higher turnover in the changes made to the S&P 500 Index. This was largely due to an increase in the number of M&As and bankruptcies etc., when failing firms were removed and high-growth companies were included
- For entities with existing news based research systems and index design expertise, this strategy can be easily implemented at a low initial cost in the form of a cash neutral tactical signal.
- Overall, the proliferation of information around additions and deletions and the increased high frequency trading activity have made it difficult to pin point the exact entry and exit point to achieve profitability.

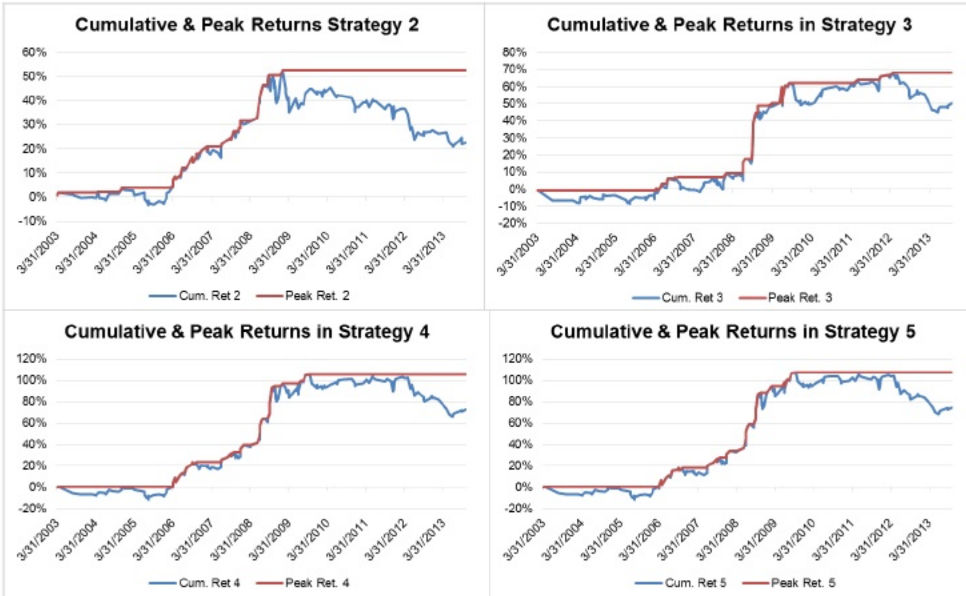

Figure 11. Drowdown

REFERENCES

[1] The Price Response to S&P 500 Index Additions and Deletions: Evidence of Assymetry and a New Explanation; Honghui Chen, Gregory Noronha, Vijay Singal; The Journal of Finance Vol. LIX, no.4 ; August 2004.

[2] Price and Volume Effects of Chnages in MSCI Indices - Nature and Causes; Wei Huang, Rajesh Chakrabart, Narayan Jayaraman, Jinsoo Lee; Journal of banking & Finance; Vol.29.2005, 5, p.1237-1264.

[3] Index membership and predictability of stock returns: The case of the Nikkei 225; Shinhua Liu; Pacific-Basin Finance Journal; Volume 17, Issue 3, June 2009, p. 338-251.

[4] Lee, Chun I; Gleason, Kimberly C; Madura, Jeff; Intraday and Night Index Arbitrage; Quarterly Journal of Finance and Accounting; Spring 2008; 47, 2; ProQuest pg. 3

[5] Lynch, Anthony, Mendenhall, Richard; New Evidence on Stock Price Effects Associated with Changes in the S&P 500; Journal of Business 1997