Time Series Modelling Effects on Pairs Trading Profitability

Among US Sector SPDR ETFs

FE 542: Group 5

Yuanyuan Liu

Wyatt Marciniak

Jimit Sanghvi

Muye Zhang

Objective

Our project analyses time series modeling effects on pairs trading strategies between highly correlated US sector Exchange Traded Funds (ETF). Our hypothesis is that modeling a time series' trends and behaviors away will offer a less biased correlation effect between ETFs and allow us to create pairs trading strategies that will yield a greater return. This benefit will be observed in the model residuals, which when modeled correctly should show no correlation or discernible trend.

This implies that the residuals are just the excess "noise" of the underlying ETF and not the result of a deeper market factor that could have been driving their behaviors in the first place. Once that factor has been modelled out, correlations of the ETFs will be based more purely on internal factors and those pairs are the truer "best pair" to trade on.

Data Source and Description

We have collected 10 years of daily data from the SPDR ETFs for the 10 sectors of the US economy, which are Biotech (s.bio), consumer Discretionary (s.cd), Consumer Staples (s.cs), Energy (s.energy), Financial (s.fin) Healthcare (s.hc), Industrials (s.ind), Materials (s.mat), Technology (s.tech), and Utilities (s.util), and using the 'quantmod' package in R. For the daily price, we use open price, which is more appropriate for trading strategy compared with close price. Then we calculated continuously compounded daily return for each sector ETF as,

$$r_t = \ln \frac{P_t}{P_{t-1}}$$

	s.bio	s.cd	s.cs	s.energy	s.fin	s.hc	s.ind	s.mat	s.tech	s.util
2008-05-14	19.43667	33.14	28.14	85.30	21.275387	31.48	39.10	44.63	25.02	40.66
2008-05-15	19.39333	33.17	28.20	85.40	21.299757	31.55	39.15	45.04	24.97	40.52
2008-05-16	19.57000	33.76	28.50	87.18	21.592201	31.61	39.46	45.66	25.50	40.53
2008-05-19	19.41000	33.35	28.51	88.97	21.316004	31.58	39.35	46.07	25.48	40.74
2008-05-20	19.41667	33.13	28.31	89.62	21.015434	31.67	39.36	45.47	25.25	41.20
2008-05-21	19.25667	32.91	28.29	90.02	20.682371	31.91	39.04	46.05	24.95	41.35

	s.bio	s.cd	s.cs	s.energy	s.fin	s.hc	s.ind	s.mat	s.tech	s.util
2008-05-15	-0.0022319351	0.0009048108	0.0021299973	0.0011716345	0.0011447995	0.0022211337	0.0012780577	0.0091447050	-0.0020004408	-0.0034491288
2008-05-16	0.0090684342	0.0176308178	0.0105820739	0.0206288226	0.0136365205	0.0019000006	0.0078870017	0.0136716351	0.0210033879	0.0002467366
2008-05-19	-0.0082093842	-0.0122188981	0.0003508156	0.0203242987	-0.0128740326	-0.0009495490	-0.0027915512	0.0089393378	-0.0007846215	0.0051680439
2008-05-20	0.0003433723	-0.0066184659	-0.0070398391	0.0072792973	-0.0142010319	0.0028458517	0.0002541736	-0.0131091891	-0.0090676750	0.0112278207
2008-05-21	-0.0082744829	-0.0066626835	-0.0007066431	0.0044532917	-0.0159754253	0.0075495797	-0.0081633104	0.0126749512	-0.0119522934	0.0036340925
2008-05-22	-0.0172854594	-0.0289786312	-0.0074508415	-0.0126321955	-0.0299005366	-0.0129318945	-0.0157483826	-0.0337892146	-0.0173847512	0.0048251241
2008-05-23	0.0085928106	0.0049922725	0.0042644342	-0.0046230673	0.0080612829	0.0028530690	-0.0057411693	-0.0101592476	-0.0008159935	-0.0009631834
2008-05-27	-0.0012229581	-0.0169495894	-0.0060466126	-0.0294762885	-0.0153724154	-0.0082645098	-0.0113173997	-0.0011352028	-0.0024519423	-0.0258676457
2008-05-28	0.0145784252	0.0157038403	0.0106459130	-0.0078294994	0.0181785743	0.0057288508	0.0162776761	-0.0146455476	0.0234543298	0.0210379574

Figure1: 10-year ETF Sector Price and Return Data

Methodology

The project consists of multiple modeling and testing stages.

First, we found the best pairs of ETFs for trading based on market open price using correlation and simple linear regression analysis. For linear regression analysis, we used adjusted R-squared value for comparison.

Second, we found the best pairs based on continuously compounded daily return using the same methodology as we did for the price.

Then, the best pairs based on ARMA model residuals were found. In this step, we constructed AR(p), MA(q) and ARMA(p,q) models by performing the following steps:

- Calculate the significance level for lag spikes.
- Extract all lags that touch or cross this threshold.
- Validate their significance with Ljung-Box dependency tests.
- Create final order lists based on lags with dependency.
- Construct all AR(p), MA(q) and ARMA(p,q) models with significant lag orders.
- Select the best model by Akaike Information Criterion (AIC).

After Building ARMA model, we obtained residuals from all 10 model. Then we found the best pairs based on residuals using the same methodology as we did for the price. Then we plotted ACF of squared residuals and performed Ljung–Box test to test if there is any autoregressive conditional heteroscedastic (ARCH) effect.

In order to find the best pairs of ETFs based on ARMA+GARCH modeled residuals, we built 8 GARCH models with orders (1,1), (1,2), (2,1), and (2,2), and then choose the best ARMA+GARCH model based on AIC.

After we have all the highly correlated pairs of sectors, we created a pairs trading strategy to backtest all the pairs then optimized the parameters of the strategy. To constructed a pairs trading strategy, we calculated a ratio of the stock price, return or model residual as the distance or relationship series and then calculate the mean of the series. We then set the signal lines to be 'k' standard deviations away from that mean line. These act as signals to open and close positions by denoting when the distance (spread) of the pairs prices/returns/residuals is abnormally large. This is possibly caused by picking 2 assets that are highly correlated, we know that their price and return movements in the markets will be closely related. This means up and down movements, as well as trends, are very similar and when the assets' relationship diverges from its normal behavior (the spread increases) we can capitalize on the hypothesis that this relationship will return to normal.

In the strategy, we long/short the winner/loser depending on which asset moves too high and which moves too low. When their spread normalizes (come within the signal lines), we reverse our trade and take our profit. We have also included a stop loss parameter for a more realistic test when considering an investors risk tolerance. What's more, since short sale can only be traded in marginal account, we add marginal function to our pair trading which is: set a margin account with initial capital, then investor can borrow money which has a maximum value from broker-dealer to short ETF with a particular interest rate and investor should pay coupon per month. Next, the margin account must be maintained at a special level which requires the investor to charge account if the current level is below that. To make our trading process more closely to the real world, we also add transaction cost which is 5 dollar per trade.

To optimize the strategy, we created a function that optimizes the backtest strategy 'k' and stop loss percentage to give the best parameters for maximized return.

For each step of the analysis, we closely monitor how the correlations between the ETFs change and how the strongest correlated pairs are selected. We also analyze how changing pair relationships based on the modeling techniques do or do not improve the profitability of the strategies.

Results

1. Pairs based on Market Price Data and Daily Return Data

```
Best Pair: s.bio + s.hc based on corr = 0.9795 and r^2 = 0.9595
Best Pair: s.cd + s.cs based on corr = 0.9867 and r^2 = 0.9736
Best Pair: s.cs + s.cd based on corr = 0.9867 and r^2 = 0.9736
Best Pair: s.energy + s.mat based on corr = 0.6056 and r^2 = 0.3668
Best Pair: s.fin + s.ind based on corr = 0.9604 and r^2 = 0.9224
Best Pair: s.hc + s.cd based on corr = 0.9855 and r^2 = 0.9711
Best Pair: s.ind + s.cd based on corr = 0.9831 and r^2 = 0.9665
Best Pair: s.mat + s.ind based on corr = 0.9713 and r^2 = 0.9435
Best Pair: s.tech + s.ind based on corr = 0.9812 and r^2 = 0.9627
Best Pair: s.util + s.cs based on corr = 0.9581 and r^2 = 0.918
```

Figure 2: Best Pairs based on Price

We have identified the best pairs to be tested in a pairs trading strategy based on price (Figure 2) using correlation and adjusted R-squared value. Biotechnology ETF pairs with health care, which seems obvious. Consumer discretionary and consumer staples form pair also seem obvious. Also, we can note that Energy has the least correlation value pairing with the material sector. Also, we can see most of the sectors form the pair with the industrial sector, this may be because most of the sectors are dependent on goods and products manufactured by industrial sector.

```
Best Pair: s.bio + s.hc based on corr = 0.7091 and r^2 = 0.5028
Best Pair: s.cd + s.ind based on corr = 0.8608 and r^2 = 0.7409
Best Pair: s.cs + s.ind based on corr = 0.7468 and r^2 = 0.5577
Best Pair: s.energy + s.mat based on corr = 0.827 and r^2 = 0.684
Best Pair: s.fin + s.ind based on corr = 0.8022 and r^2 = 0.6434
Best Pair: s.hc + s.ind based on corr = 0.7565 and r^2 = 0.5724
Best Pair: s.ind + s.tech based on corr = 0.8628 and r^2 = 0.7444
Best Pair: s.mat + s.ind based on corr = 0.8597 and r^2 = 0.739
Best Pair: s.tech + s.ind based on corr = 0.8628 and r^2 = 0.7444
Best Pair: s.util + s.cs based on corr = 0.642 and r^2 = 0.4121
```

Figure 3: Best Pairs based on Return

Similarly, we identified best pairs based on daily returns data. Here, as well most of the sectors (6 out of 10) are pairing with industrial sector may be because of same reason mentioned above. We can also notice that the correlation value and adjusted R-squared value decrease from price to return. This makes sense as price movements swing much more greatly then actual prices and those movements appear to reveal a stronger correlation indication. But in case of energy pairing with the material sector, the correlation increases. We will need to do some further analysis as to why this behavior is occurring, however in terms of our project objective at this time this observation will not affect our final results.

2. Timeseries ARMA Model

We have fit ARMA(p,q) models to the 10 sectors ETFs. The goal of doing this step separately from the full modeling step (with volatility) is to increase the granularity of our analysis with respect to modeling effects. We have constructed these models and tested them for pairs creations as well as testing their residuals to identify models that will need volatility modeling as well.

```
-------------------------------------------------------------------------------------------
Building Models for Time Series: s.bio
-------------------------------------------------------------------------------------------
(1) Checking for Unit Roots (ADF Test)...
Unit root removed with 0 differences
(2) Selecting + Validating AR/MA orders up to 20 lags to build models...
* User set 'significance booster' (Additional Crossover Length needed to be truly significant) at 20 %
AR order = ( 1,4 ) with max = 4
MA order = ( 1,2,5 ) with max = 5
-------------------------------------------------------------------------------------------
(3) Creating ARMA(0,0,0) model...
ARMA model (0,0,0) has AIC = -12341.7702 and variance = 0.0004081
(4) Creating AR models...
ARMA model (1,0,0) has AIC = -12349.6727 and variance = 0.00040648
ARMA model (4,0,0) has AIC = -12351.9188 and variance = 0.00040513
(5) Creating MA models...
ARMA model (0,0,1) has AIC = -12349.4636 and variance = 0.00040651
ARMA model (0,0,2) has AIC = -12347.904 and variance = 0.00040644
ARMA model (0,0,5) has AIC = -12350.0935 and variance = 0.0004051
(6) Creating ARMA models...
ARMA model (1,0,1) has AIC = -12347.815 and variance = 0.00040645
ARMA model (1,0,2) has AIC = -12345.9085 and variance = 0.00040644
ARMA model (1,0,5) has AIC = -12348.0747 and variance = 0.0004051
ARMA model (4,0,1) has AIC = -12349.9519 and variance = 0.00040512
ARMA model (4,0,2) has AIC = -12349.491 and variance = 0.00040487
ARMA model (4,0,5) has AIC = -12347.0677 and variance = 0.00040429
(7) Adding an Auto Calibrarted Model Selection NOT SELECTED by user...
-------------------------------------------------------------------------------------------
Model Forcing was not allowed by user...
-------------------------------------------------------------------------------------------
Total Models Constructed = 12 which was 92.31 % of total model consideration
Model Choosen = ARMA(4,0,0) with AIC = -12351.9188 and variance = 0.00040513
```

Figure 4: Constructing AR, MA, and ARMA Models based on AIC

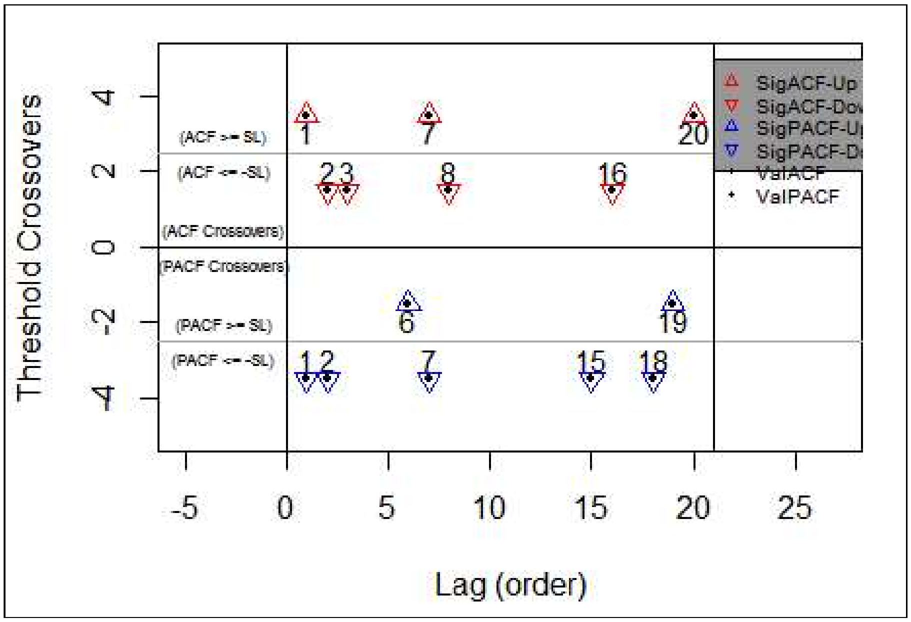

Figure 5: Utilities Sector ETF Return Data ACF/PACF plot with Dependency Verification

Observation	Model
Biotech	ARMA(4,0,0)
Consumer Discretionary	ARMA(15,0,14)
Consumer Staples	ARMA(14,0,14)
Energy	ARMA(13,0,14)
Financials	ARMA(17,0,14)
Healthcare	ARMA(15,0,3)
Industrials	ARMA(15,0,15)
Materials	ARMA(14,0,15)
Technology	ARMA(14,0,14)
Utilities	ARMA(15,0,16)

Table 1: Best ARMA(p,q) Models Selected

The table 2 shows the best ARMA(p,q) models selected for further analysis. After extracting the residuals from these models we find the best pairs based on residuals.

```
Best Pair: s.bio + s.hc based on corr = 0.7082 and r^2 = 0.5015
Best Pair: s.cd + s.ind based on corr = 0.8416 and r^2 = 0.7083
Best Pair: s.cs + s.ind based on corr = 0.7326 and r^2 = 0.5366
Best Pair: s.energy + s.mat based on corr = 0.7934 and r^2 = 0.6296
Best Pair: s.fin + s.ind based on corr = 0.7774 and r^2 = 0.6044
Best Pair: s.hc + s.ind based on corr = 0.7419 and r^2 = 0.5504
Best Pair: s.ind + s.tech based on corr = 0.8427 and r^2 = 0.7102
Best Pair: s.mat + s.ind based on corr = 0.834 and r^2 = 0.6956
Best Pair: s.tech + s.ind based on corr = 0.8427 and r^2 = 0.7102
Best Pair: s.util + s.cs based on corr = 0.6373 and r^2 = 0.4061
```

Figure 6: Best Pairs based on ARMA Residuals

We identified best pairs based on residuals like we did for price and return data (Figure 6). They form the same pair as pairs obtained using return data. The correlation and r-squared value are almost same as well. This may be because residuals are directly calculated using daily return value.

3. Timeseries ARMA+GARCH

(a) ARMA Residuals Analysis

After Building ARMA model, we obtained residuals from all 10 modelled residuals. Then we write function to plot ACF of residuals squared and do Ljung–Box test to test if there is any arch effect.

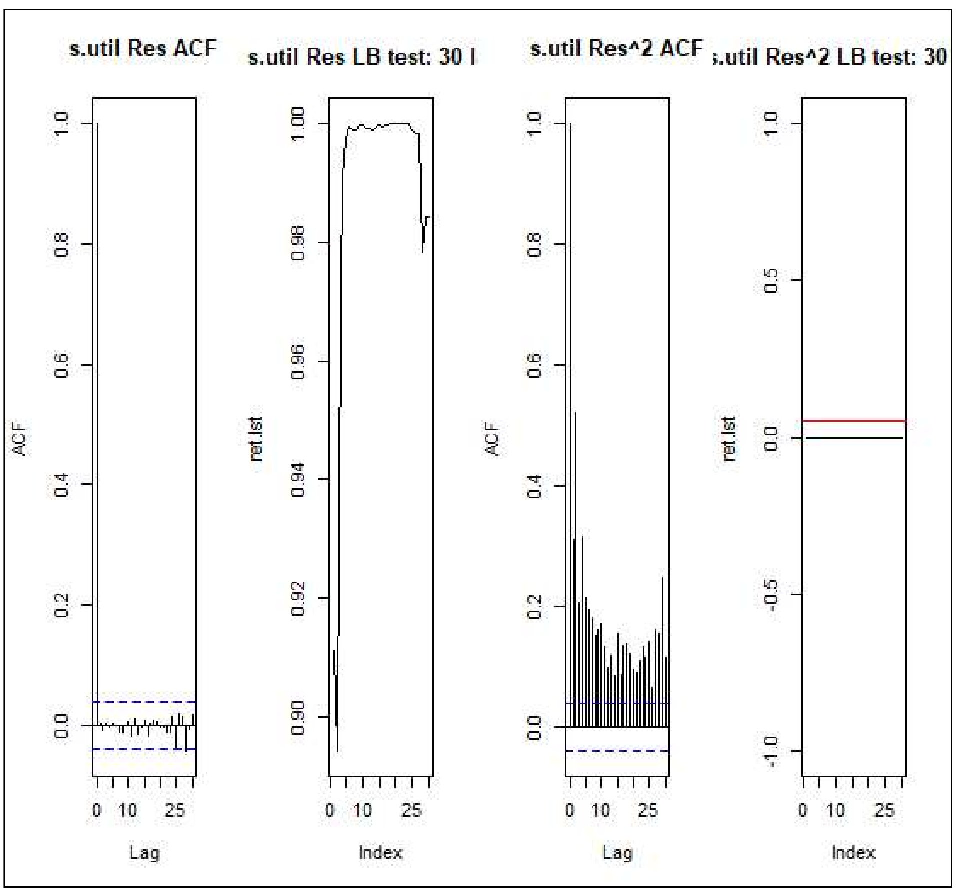

Figure 7: Utilities Sector ETF ARMA Model Residuals

Take Utilities ARMA(15,16) model for example. The first plot shows residuals of our model are behaving like white noise which is good and second plot is the p-value of Ljung-Box test in 30 lags which proved the residuals have no dependence among them. For ARCH effect ACF plot and Ljung-Box test of residuals squared prove that there exists arch effect so GARCH model is required to be considered.

(b) Fit ARMA+GARCH model

In our project, we built 8 GARCH models with orders (1,1), (1,2), (2,1), (2,2) because these orders are universal in major researches and the conditional distributions are standard normal distribution and student-t distribution. After we obtained the 8 models, we choose the best one based on AIC. The function works like following:

```
Building arma(p,q)+garch(m,s) models for: s.bio
(1) Distribution = norm
arma(4,0)+garch(1,1) has AIC/LLH = -5.1761 / -6441.9046
arma(4,0)+garch(1,2) has AIC/LLH = -5.1775 / -6444.6391
arma(4,0)+garch(2,1) has AIC/LLH = -5.1751 / -6441.6513
arma(4,0)+garch(2,2) rejected based on NaNs in error testing phase
(2) Distribution = std
arma(4,0)+garch(1,1) has AIC/LLH = -5.2002 / -6472.8987
arma(4,0)+garch(1,2) has AIC/LLH = -5.201 / -6474.8742
arma(4,0)+garch(2,1) has AIC/LLH = -5.1993 / -6472.681
arma(4,0)+garch(2,2) rejected based on NaNs in error testing phase
--------------------------------------------------------------------
Model Selection Criteria = aic
Models compared = 6 which is 75 % of possible 8 models
Best model: arma(4,0)+garch(1,2) with dist = std where aic = -5.201
```

Figure 8: Constructing ARMA+GARCH Model based on AIC

```
Best Pair: s.bio + s.hc based on corr = 0.6882 and r^2 = 0.4736
Best Pair: s.cd + s.ind based on corr = 0.7833 and r^2 = 0.6135
Best Pair: s.cs + s.cd based on corr = 0.6521 and r^2 = 0.4252
Best Pair: s.energy + s.mat based on corr = 0.7049 and r^2 = 0.4968
Best Pair: s.fin + s.ind based on corr = 0.7637 and r^2 = 0.5832
Best Pair: s.hc + s.bio based on corr = 0.6882 and r^2 = 0.4736
Best Pair: s.ind + s.cd based on corr = 0.7833 and r^2 = 0.6135
Best Pair: s.mat + s.ind based on corr = 0.7819 and r^2 = 0.6113
Best Pair: s.tech + s.cd based on corr = 0.7664 and r^2 = 0.5873
Best Pair: s.util + s.cs based on corr = 0.5556 and r^2 = 0.3086
```

Figure 9: Correlation in ARMA+GARCH Residuals

Like we did for the price, return and ARMA residuals data we identified best pairs based on GARCH+ARCH (Figure 9). We can see the pairing actually change. The healthcare ETF price was highly correlated with consumer discretionary price, whereas the residuals of GARCH model are highly correlated with biotechnology. Healthcare pairing with biotechnology makes sense. But at the same time technology's GARCH Residuals changed pairing from industrial to Consumer discretionary doesn’t make sense, which can be verified with this pair making so less profit. But, for most of the times, it gives sensible pairs when compared to price and return. The pair trading backtesting will further analyze the countability of these pairs.

4. Correlation

Pair trading is based on two highly correlated equities, so the most important part of pair trading is to choose best pairs. We now have price, return, residuals of ARMA, and residuals of ARMA+GARCH dataset then we can analyze the correlation between them and find best pair based on different datasets. Correlation calculation is based on:

$$\text{Correlation} = \rho = \frac{Cov(x, y)}{\sqrt{VAR(x)VAR(y)}}$$

After we analyzed all correlations, we can choose top one to make our trading pair as it's in table 2. Generally, the correlation tend to be higher in price and decrease in return, ARMA residuals and ARMA+GARCH residuals

Observation	Price	Return	ARMA	ARMA + GARCH
Biotech	Healthcare (0.98)	Healthcare (0.71)	Healthcare (0.71)	Healthcare(0.69)
Consumer Discretionary	Consumer Staples (0.99)	Industrials (0.86)	Industrials (0.86)	Industarials(0.78)
Consumer Staples	Consumer Discretionary (0.99)	Industrials (0.75)	Industrials (0.74)	Consumer Dis(0.65)
Energy	Materials (0.60)	Materials (0.83)	Materials (0.82)	Materials(0.70)
Financials	Industrials (0.95)	Industrials (0.80)	Industrials (0.80)	Industrials(0.76)
Healthcare	Consumer Discretionary (0.99)	Industrials (0.76)	Industrials (0.75)	Biotech(0.69)
Industrials	Consumer Discretionary (0.98)	Technology (0.86)	Technology (0.86)	Consumer Discretionary (0.78)
Materials	Industrials (0.97)	Industrials (0.86)	Industrials (0.85)	Industrials (0.78)
Technology	Industrials (0.98)	Industrials (0.86)	Industrials (0.86)	Consumer Staples (0.77)
Utilities	Consumer Staples (0.95)	Consumer Staples (0.64)	Consumer Staples (0.64)	Consumer Staples (0.56)

Table 2: Best Pairs

According to table 2, correlations for the price, returns, ARMA Residuals, and ARMA+GARCH residuals decrease significantly. Return and ARMA Residual Correlations show little difference, and ARMA+GARCH residuals correlations are lower than all other categories.

5. Pair Trading & Backtest

We chose 40 pairs in total to do our backtest(10-year), and the results are shown in table 3: For Biotech ETF, the best-paired ETF is Healthcare. When we used ARMA residuals as for trading, we made 800% return in 10 years, which is much larger than others. When we chose Consumer Discretionary and Industrial as a pair, the return is 246% based on ARMA residuals. In fact, as shown in tables 3, pair trading with ARMA residuals always get the highest return out of all the categories, and all the annual return could reach to at least 9.3% level which is very outstanding. The only loss we had is pair trading healthcare and industrial ETFs.

ETF	Pair Trading with Price	Pair Trading with Return	Pair Trading with ARMA Residuals	Pair Trading with GARCH Residuals
Biotech	Health Care (61.36%)	Health Care (142.74%)	Health Care (800.93%)	Health Care (16.14%)
Consumer Discretionary	Health Care (17.14%)	Industrial (125.37%)	Industrial (245.84%)	Industrial (28.95%)
Consumer Staples	Consumer Discretionary (39.27%)	Industrial (12.85%)	Industrial (93.12%)	Consumer Discretionary (21.74%)
Energy	Material (21.74%)	Material (119.33%)	Material (521.14%)	Material (61.32%)
Financial	Industrial (120.64%)	Industrial (283.6%)	Industrial (534.74%)	Industrial (73.25%)
Health Care	Consumer Discretionary (66.23%)	Industrial (-7.43%)	Industrial (156.29%)	Biotechnology (58.26)
Industrial	Consumer Discretionary (55.73%)	Technology (32.86%)	Technology (120.13%)	Material (68.53%)
Material	Industrial (50.94%)	Industrial (132.69%)	Industrial (391.69%)	Industrial (32.77%)
Technology	Industrial (73.76%)	Industrial (32.86%)	Industrial (120.13%)	Consumer Discretionary (19.75%)
Utility	Consumer Staples (65.1%)	Consumer Staples (102.46%)	Consumer Staples (291.69%)	Consumer Staples (17.44%)

Table 3 : Return on Pair Tradings using Back-testing

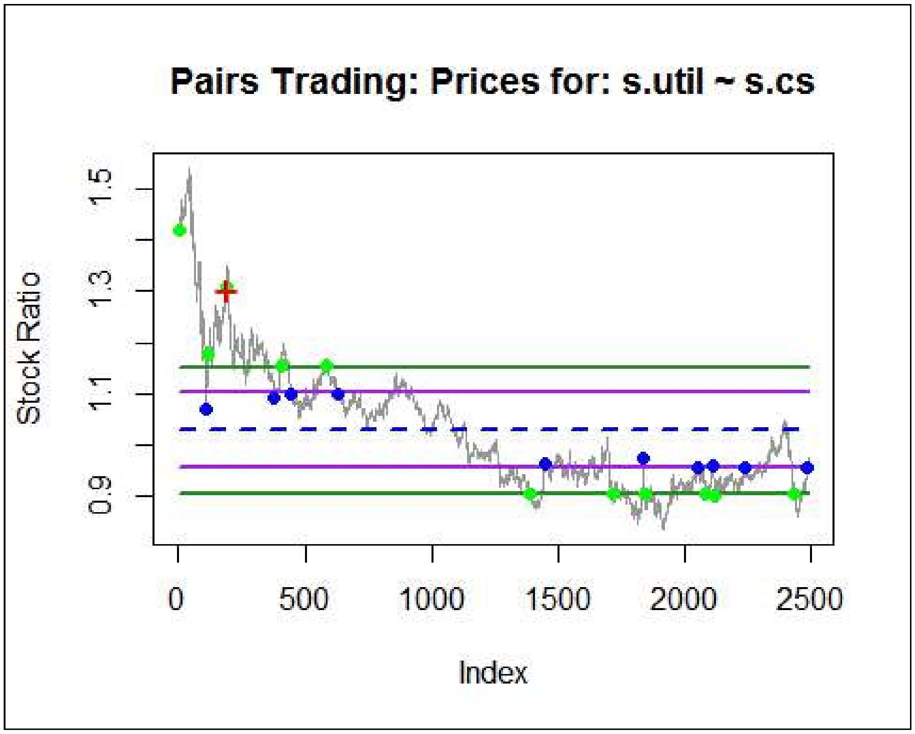

Figure 10: Back-test Visualization using Daily Price for Utilities Sector ETF

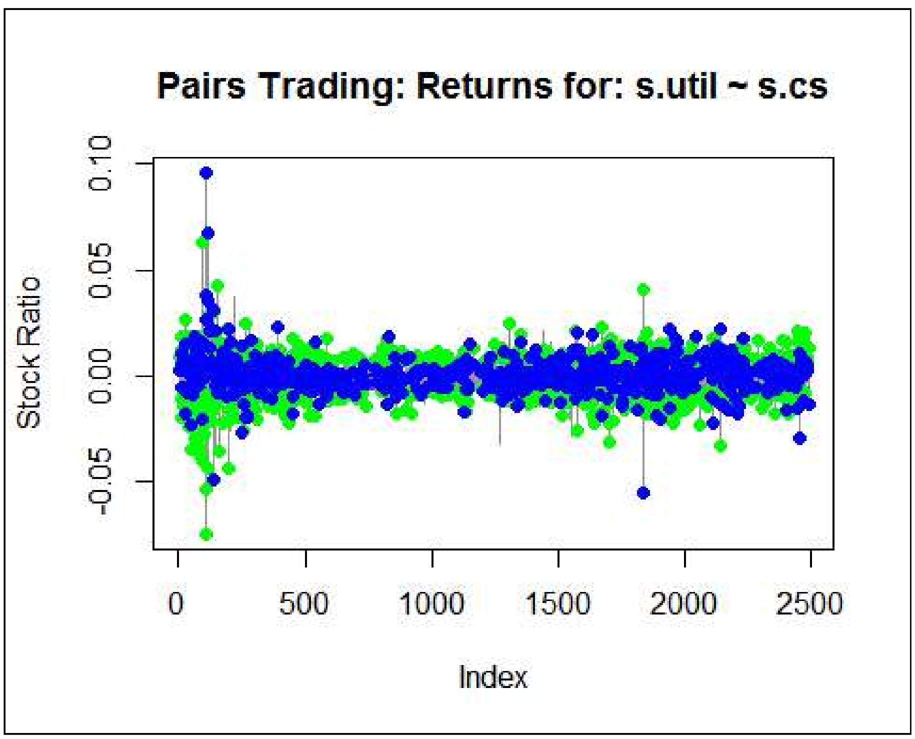

Figure 11: Back-test Visualization using Daily Return for Utilities Sector ETF

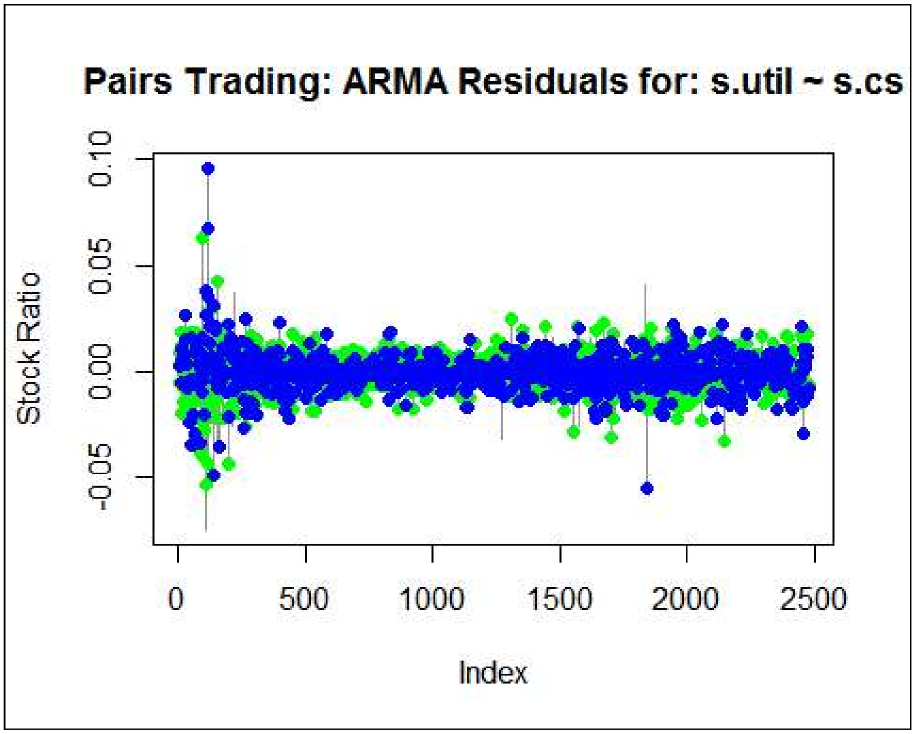

Figure 12: Back-test Visualization using ARMA Residuals for Utilities Sector ETF

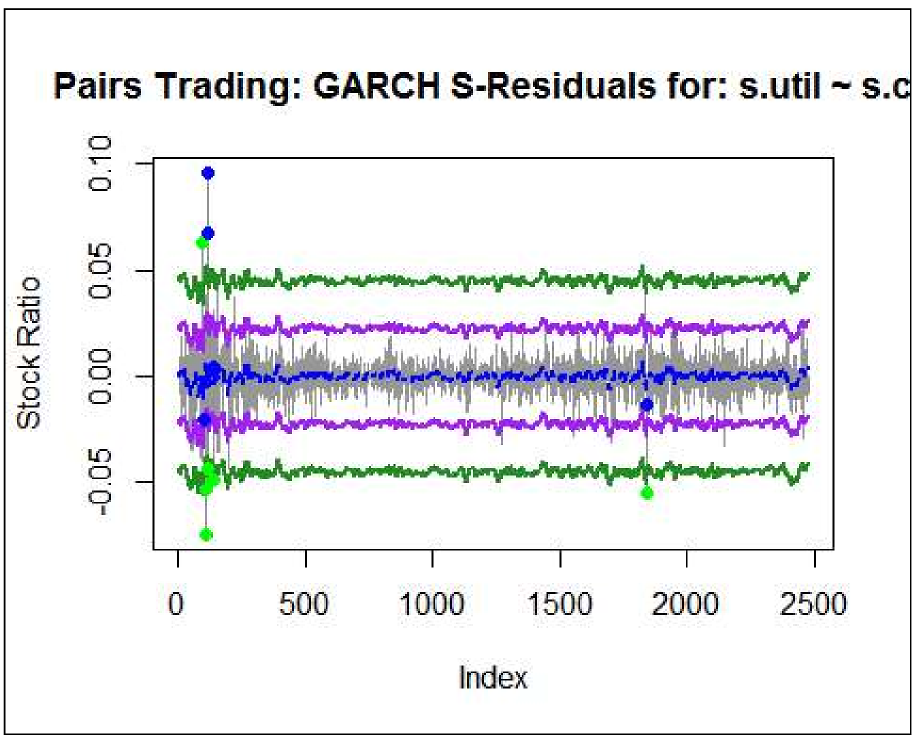

Figure 13: Back-test Visualization using ARMA+GARCH Residuals for Utilities Sector ETF

Next we look at the numbers of trades in the backtest as shown in tables 4:
Though the “Return pair tradings” do not behave as well as ARMA ones, but the numbers of trades are really close. Interestingly, the numbers of trade are similar but the return is significantly different, which can prove that when we trade with ARMA residuals data the signals of opening and existing are more accurate to make profit.

We can also see that the numbers of trade in price and GARCH term are lower which explain the why returns are lower.

ETF	Numbers of Trades for Price	Numbers of Trades for Return	Numbers of Trades for ARMA Residuals	Numbers of Trades for ARMA+GARCH Residuals
Biotech	Health Care (19)	Health Care (760)	Health Care (853)	Health Care (45)
Consumer Discretionary	Health Care (9)	Industrial (873)	Industrial (867)	Industrial (14)
Consumer Staples	Consumer Discretionary (13)	Industrial (398)	Industrial (691)	Consumer Discretionary (9)
Energy	Material (7)	Material (859)	Material (881)	Material (134)
Financial	Industrial (23)	Industrial (843)	Industrial (873)	Industrial (27)
Health Care	Consumer Discretionary (12)	Industrial (835)	Industrial (871)	Biotechnology (124)
Industrial	Consumer Discretionary (6)	Technology (717)	Technology (809)	Material (141)
Material	Industrial (8)	Industrial (848)	Industrial (875)	Industrial (49)
Technology	Industrial (25)	Industrial (717)	Industrial (809)	Consumer Discretionary (34)
Utility	Consumer Staples (11)	Consumer Staples (591)	Consumer Staples (774)	Consumer Staples (6)

Table 4 : Number of Trades

We compared our return to 3 important economical indexes in the United States and we could see even if the whole market is good, we can still make more profit than the market when we choose ARMA residuals to pair trading.

Other	S&P 500	DJIA	Long ETFs
10 - Year Return	98.24%	84.72%	81.56

Table 5 : Market Return

Conclusion

- Time-series behavior significantly affects the pairing between two ETFs. The pairs obtained after applying GARCH+ARCH time-series model makes more sense.
- For pair-trading using price financial sector and industrial sector form the best pair, as it gives maximum return i.e. 120.64%. So, a trader can think on trading with this pair if he/she want to trade using price.
- For pair-trading using continuously compounded daily return again financial sector and industrial sector form the best pair when we compare the return on investment i.e. 283.6%. So, a trader can think on trading with this pair if he/she want to trade using daily return.
- For pair-trading using ARMA residuals healthcare sector and biotechnology sector form the best pair, as it gives maximum return i.e. 800.93%, which is indeed the highest return compared to any other pair or method used. So, a trader can think on trading with this pair if he/she want to trade using ARMA residuals.
- For, pair-trading using GARCH+ARCH residuals again financial sector and industrial sector form the best pair when we compare the return on investment i.e. 73.25%. So, a trader can think on trading with this pair if he/she want to trade using the GARCH+ARCH residuals.
- Pair-trading using time-series modeled ARMA residuals can provide us better returns when compared to returns, price, and GARCH+ARCH residuals. Most of the time it has given 6 to 10 times more returns as compared to the overall market return.
- If a trader has to trade with respect to daily return and ARMA modeled residuals, he/she may have to execute more trades.
- Pair-trading using ARCH+GARCH modeled residuals shows worst performance over the years. So, we can conclude that one should not think of implementing trades using pair-trading strategy using ARCH+GARCH modeled residuals.
- Overall, we can conclude that our objective of using time-series model to optimize pair-trading seems satisfied, as back-testing results show that time-series ARMA model gives maximum return on investment.

Appendix

FE542_Testing_rmd.pdf
FE542_Function.r
FE542_Testing.Rmd