NTNU
Norwegian University of Science and Technology

Pairs Trading in the Aluminum Market

A Cointegration Approach

Oddvar Hallset Reiakvam
Stian Borgen Thyness

Industrial Economics and Technology Management
Submission date: July 2011
Supervisor: Sjur Westgaard, IØT

Norwegian University of Science and Technology
Department of Industrial Economics and Technology Management

Problem Description

Pairs trading have been a popular way of trading securities since the 1980s. It is usually applied to trade different stocks with similar characteristics such as companies within the same industries. This can for instance be two oil producers. It can also be applied to trade commodity sensitive stocks vs the underlying commodity. An extension of pairs trading is statistical arbitrage where one of the stocks may be substituted with a set of risk factors or a portfolio of securities. In the last decade, there has been a boom in commodity investing and a variety of investment vehicles have been introduced. Among these are ETFs and indices. Speculation in futures contract has also become more common. We want to explore the market for different aluminum related securities in order to find securities with characteristics making them suitable for statistical arbitrage trading. Therefore, the problem we want to solve is:

Is it possible to find aluminum securities which co-moves in such a way that it is possible to create statistical arbitrage trading rules yielding higher return than a passive buy-and-hold strategy in the same securities?

Assignment given: 21. February 2011
Supervisor: Sjur Westgaard

Pairs Trading in the Aluminum Market

A Cointegration Approach

MASTER'S THESIS

18th of July 2011

ODDVAR HALLSET REIAKVAM

&

STIAN BORGEN THYNESS

Department of Industrial Economics and Technology Management

Norwegian University of Science and Technology

Abstract:

This paper applies various ways of constructing statistical arbitrage trading rules for aluminum securities. The paper use daily observations of stocks, futures and two securities supposed to mirror the return of physical aluminum. We employ several sophisticated analysis of the statistical properties of these securities and how they relate to each other. This paper applies Engle-Granger and Johansen tests for cointegration to identify suitable securities for pairs trading. The paper is useful for speculators and hedge fund managers who want to increase their risk adjusted returns, as our analysis shows that trading sector neutral positions instead of holding passive long positions in aluminum securities have significantly higher risk adjusted returns. Our methodology is not unique for aluminum and can be transferred to other areas such as oil or precious metals.

Keywords: Commodities, ETFs, Stocks, Cointegration, Correlation, Engle-Granger, Johansen, Statistical Arbitrage, Principal Component Analysis, Trading Strategies, Moving Average, Pairs Trading, Bollinger Band.

01. Introduction

The individual speculators are nowadays taking various positions in the commodity markets and have higher exposure to commodities than in the past. This is not only because the average investor is more enlightened about the financial markets than earlier, but also because it is easier than before. New investment vehicles and online trading platforms make it simple for individual speculators to execute the transactions from their personal computers. The transaction costs have decreased significantly over the years[1].

The first commodity index was the S&P Goldman Sachs Commodity Index (GSCI) launched in 1999. After 2000, there has been created many indices and exchange-traded funds (ETF); both indices that track a wide specter of securities and those who track one single commodity. The first ETF tracking one single commodity was the StreetTRACKS Gold Shares and was created in 2004. By 2008 its market capitalization exceeded 17 billion U.S. Dollars. ETFs have become an inexpensive alternative for commodity investments and can for example be used to exploit future mispricing. The ETFs made it simpler for smaller speculators to be a part of the arbitrage game or to make more sophisticated portfolios.

ETFs track indices, measuring the performance of different asset classes. However, they do not promise to track every single movement, rather they seek to replicate, to the extent possible. Therefore we have times where these funds fall short, called tracking error. A study by Morgan Stanley found that the average tracking error for U.S. listed ETFs was 0.52 percent in 2008. The average mispricing has a high degree of variability. During the volatile period of 2008 the iShares FTSE NAREIT Mortgage REIT was 11.8 percentage points under, Vanguard Telecommunication Services was 5.7 percentage points under, and iShares MSCI Emerging Markets Income ETF 4.1 percentage points over.

We anticipate that there is no long lasting mispricing and hence arbitrage opportunities in the futures, equities and ETF markets due to low transaction costs and the high. However, we believe that there are mispricing between different types of securities, at least in the short-term. The mispricing can for example exist due to speculation by major actors, or by errors in ETFs. Commodity ETFs usually track the price of a commodity through the futures markets; buying a

[1] The online trading platforms have 0.05% and less in commission fee.

contract close to expiry. When the contract gets close to maturity, the ETF providers sell the contract and buy a new one. This procedure is repeated every month to avoid taking delivery of the underlying commodity. This is called rolling contracts. If the next contract has a higher price it incurs a loss. Even if the price of the commodity stays the same, the ETF could still have a small loss due to transactions cost. On the other hand, if the new contract has a lower price, the ETF will have an upward bias. The divergence can be large; in 2006 the United States Oil Fund Index was 13 percent below the West Texas crude oil, the commodity the fund was supposed to replicate.

We believe that commodity prices can be affected by speculators and hence get driven away from their fundamental value. There are cases where governments accuse major actors in the oil market for manipulating the oil prices. Weather speculators are to blame for the increased volatility in the commodity markets is a subject market anticipants disagree on. In late 2007 and early 2008 the oil prices doubled in less than a year to all time high of $147 a barrel. Kenneth Medlock at Rice University claims that in 2008 the speculators holdings marked for short-term trades was 55 percent of total contracts, and that speculators still held around 50 percent in 2009. Medlock concludes that unchecked short-term price speculation is to blame for excessive price volatility over the past two years and that speculators drive up prices that eventually reach consumers. The EDHEC-Risk Institute (2008) also analyzed what caused the sharp increase in oil prices during past years. They concluded that supply-and-demand imbalances were the major cause over the long run, while futures trading by speculators can have only short-term effects.

In May 2011 the Commodity Futures Trading Commission accused oil speculators for driving the prices of crude oil higher in the early months of 2008. Some speculators bought a lot of oil and created a shortage of oil in Cushing (a major point for oil delivery), and thereby drove the price of future contracts higher. The same speculators later shorted future contracts to other investors, while they sold their holdings in Cushing. The speculators made a huge profit.

There has been much work devoted to the field of relationships between commodities and equities. One can question if the speculators' increased exposure to the commodity market has an impact on the prices on the related financial instruments or not. However, we do know that price changes in commodities can affect nation's economies and impact certain sectors. The most affected sectors seem to be the oil-related industries and those who are highly sensitive to oil prices like transportation and manufacturing industries. Commodity prices will affect the companies' earnings, which in turn will affect the stock price. Speculators are aware of these relationships and are making bets on the direction of prices on both the stock and the underlying commodity.

There exist many contributions on the pricing relationships between commodity prices and stock prices. However, as far as we know, these researches have been devoted to oil. No work has been done on aluminum securities. The aluminum sector is less capitalized than the oil sector, but it is still an important sector worldwide. There exists forward and futures contracts on aluminum, and there are also ETFs on these contracts. There are many companies around the world that produce aluminum and the stock prices of these companies are highly affected by changes in aluminum prices.

This paper focuses on mispricing between aluminum securities. We use a portfolio of stocks containing two of the largest aluminum producers in the word, Alcoa (AA) and Rio Tinto (RIO), together with Kaiser Aluminum (KALU) and Century Aluminum (CENX). We investigate whether there are any long-term relationships or co-movements between this portfolio of stocks and future contracts on aluminum and ETFs. We use these results in a statistical arbitrage approach designed to exploit short-term deviations from a long-term equilibrium prices between two securities.

In this study we select trading pairs based on cointegration between aluminum securities. First we test if there is any stochastic trends in the individual times series[2] . Then a cointegration test is conducted to test if the different time-series have common long-term relationships and whether causality exists. The presence of cointegration enables us to combine pairs of securities in a linear combination so that the pair is a stationary process.

We use various techniques to analyze such divergence in returns and see if there exist relationships we can take advantage of and gain a risk adjusted return than a buy and hold strategy. The portfolio of pairs is formed by buying the relative undervalued and shorting the overvalued. Since the combination of securities share a long-run equilibrium relationship the

[2] A shock that hits a security with a stochastic trend will have permanent effect and hence the series is not mean-reverting.

deviation from this equilibrium is temporary and they are expected to return some time in the future.

The investment strategy we aim to implement is a sector neutral strategy, also known as pairs trading. There are two types of pairs trading: statistical arbitrage convergence/divergence trades, and fundamentally driven valuation trades. The fundamentally driven pair trading strategy is a trading strategy where the speculator is tracking two different securities with approximately the same properties, for example two oil producers. If one of the securities has been outperforming the other a pairs trader will short the outperformer and buy the underperformer, waiting for the prices to converge. The Statistical Arbitrage strategy evolved out of the fundamental strategy and is a more quantitative approach where two securities or portfolio of securities are out of equilibrium and are expected to converge. The statistical arbitrage method is a highly technical short-term mean reversion strategy normally involving a large number of securities.

Pairs trading became popular in the 1980's. Tens of thousands of dollars in computer hardware and software were required to be able to trade pairs, limiting pairs trading only to large investment banks. Morgan Stanley's Black Box was the first systematically statistical arbitrage trading system in the world, launched in 1985. As the cost of computers has been reduced substantially since then, more speculators have been able to employ pairs trading strategies. As more speculators have engaged in pairs trading and statistical arbitrage, profits have diminished (Pole 2007).

There is no exact definition of how a statistical arbitrage model and trading rule is set up. Instead of trading two securities with the same characteristics like in conventional pairs trading, Avellaneda & Lee (2008) substitute one of the securities with a set of relevant risk factors which is affecting the return of the security. We deploy several trading strategies on the data set**.**

We employ a relative value statistical arbitrage model with cointegrated pairs of securities and portfolios. We use a Johansen test and an Engle-Granger test to identify cointegrated pairs and the appropriate weights for each security. We then apply inverted moving average rules and modified Bollinger Band rules to generate buy, sell and short signals.

We find that cointegrated pairs trading yields significant higher risk adjusted return than a buy and hold strategy. Our strategies are independent of market direction, and are therefore perfect substitutes for long only strategies in markets that that does not trend. This is highly relevant in today's financial markets with high degree of uncertainty. We believe that our techniques can easily be transformed to other commodity markets where securities are cointegrated.

Our paper is organized as follows: Section two reviews related literature. Section three introduces the dataset and the descriptive statistics for the securities. Section four investigates the long-run relationships between the securities and test these for cointegration. This is the process for selecting trading pairs. In this section we also make weighted portfolio of stocks using Principal Component Analysis (PCA). In section five we introduce the different methodologies used in the statistical arbitrage approach. Here we present the trading models that we are going to apply to our cointegrated pairs. Section six reveals our major findings, empirical results and trading strategies. In section seven we present our conclusions and suggestions for further research.

02. Literature Review

In this section we present the earlier contributions to the subject of interest. There are many contributions on the field of arbitrage and diversification strategies between equities and commodities, while less investigation regarding the ETFs. This is mainly because the ETFs are relatively new investment vehicles and most of the commodity ETFs was introduced in 2006. The relationship between stocks and futures on underlying commodities has been an interesting subject for decades. Lee et. al (1985) investigates the distributional and causal relations between stocks and the commodity futures market indices; the S&P500 and the Commodity Futures Index of 27 commodities. They find no relationship between the two time series and conclude that it is not likely to find arbitrage opportunities between the two indices. However, they do not deny the existence of arbitrage opportunities between individual stocks and individual commodities.

In the 90s and the following decade the interest of the relationship between commodities and stock increased rapidly and resulted in many empirical contributions. Among them were Huang, Masulis and Stoll (1996) which examined the relationship between oil futures and stock returns using daily returns. They find that oil futures return can affect the individual stock returns, but do not have any impact on the broad-based market indices. A recent study by Büyûksahin, Haigh and Robe (2008) applies dynamic correlation and recursive cointegration techniques to examine whether the increased commodity investment have affected commodity price correlations with traditional securities. They find that the relationship between the returns of the different securities have not changed significantly in the period from 1992 to 2008 and conclude that there is no increase in co-movement between the securities.

Hammoudeh, Dibooglu and Eleisa (2002) investigate the relationship among oil prices and the oil industry equity indices. They use cointegration- and spillover analysis to investigate whether the relationship can offer any diversification opportunities. Their results indicate that oil stocks were not able to explain the movements in the futures prices, but the oil futures prices could explain movements in the stock prices of independent companies engaged in exploration and oil refining.

Haigh, Harris, Overdahl, and Robe (2007) analyze the speculator positions in the New York Mercantile Exchange's WTI Crude Oil Futures. They find increased participation by hedge funds and commodity swap dealers. This participation has increased relationship between futures prices and resulted in greater pricing efficiency, which would decrease the possibility for arbitrage and mispricing among securities. MacKinlay and Ramaswamy (1998) use high frequency data and finds that mispricing between futures contracts and the underlying index are persistent in 15-40 minutes, and it is likely to believe that the intraday mispricing has decreased since then. This is because after the introduction of ETFs, index arbitrage should be less costly and easier since there is only one asset transaction. Fewer deviations from the equilibrium pricing relationship should exist.

If mispricing and arbitrage opportunities occur in the financial markets one can question the efficient market hypothesis. Theory states that mispriced financial assets move fast towards equilibrium due to the actions of economically rational market participants. Abreu and Brunnermeir (2002) argue that to eliminate mispricing it requires coordination of speculators rather than single actors. Shleifner and Vishny (1997) and De Long et. al (1990) provide examples about what can lead to mispricing of securities due to costs, risks and constraints. They argue that arbitrage opportunities exist. Thaler et. al (1991) and De Long et. al (1990) also find evidence that speculators have the ability to drive prices away from equilibrium and exacerbate mispricing.

Avellaneda and Lee (2008) investigate the relationship between ETFs and equities. They create two different categories of statistical arbitrage trading models; Principal Component Analysis (PCA) and sector ETFs. In the ETF models they use ETFs as proxies for industry factors. In the PCA approach they extract eigenportfolios from the eigenvectors of the returns correlation matrix. They present a systematic approach to statistical arbitrage and for construction market neutral portfolio strategies based on mean-reversion. That stocks and financial instruments are mean-reverting has been a investigated for decades and there are several previous studies. See for example Poterba and Summers (1988), Lehmann (1990) or Lo and MacKinley (1990). However, Avellaneda and Lee find that the strategies yielding the best results are based on either 15 ETFs or a 15-PCA strategy. If they increased the number of factors, the corresponding residuals got small variances, and the opportunity to make money vanished. Statistical arbitrage as a method is quite popular among fund managers and many quantitative hedge funds. After the recent financial crisis the strategy gained increased attention, due to the problems of convergence in prices and de-leveraging of portfolios. See for example Barr (2007), Rusli (2007) or Khandani and Lo (2007).

Gatev, Goetzmann and Rouwenhorst (1999) deploy a distance approach where they identify pairs of stocks with similar price history. Trades are opened when the relative price is more than two standard deviations away from equilibrium. They use data from 1962 to 1997 and manage to create excess return of up to 12 percent from a portfolio of their top pairs. Nath (2003) uses high frequency data from the U.S. Government bond market. Each time the spread crosses the 15 percentile, a trade signal is generated. The position is then held for a given period of time or until the spread crosses the 5 percentile signaling stop loss. He concludes that it is possible to create simple pairs trading models to exploit short term mispricing.

Vidyamurthy (2004) creates pairs trading models using the Engle-Granger (Engle & Granger, 1987). If two securities are cointegrated they will have a long-run equilibrium relative price and deviations from this price may be exploited by speculators. Trading signals are generated when the relative price is deviating sufficiently from its long-run equilibrium. Because the securities are cointegrated, the speculators believe the relative price will at some point in the future converge to its long-run equilibrium. Vidyamurthy applies two different methods for analyzing how large the deviation from the long-run equilibrium must be in order to trade. First he chooses the deviation which maximizes profits over the sample period. Next he models the residuals as an ARMA process using Rice's formula[3] to calculate the rate of zero crossings and level crossings in order to maximize expected returns. The deviation generating the highest return is then used to trade with. Elliot et al (2005) uses a Vasicek Process to model the spread. An important limitation to this model is that it requires the securities to have approximately the same return series; hence it is usually deployed to trade companies listed at more than one exchange. Do et al (2006) uses a model where the long-run equilibrium price is found through the arbitrage pricing theory and then modeled as a mean reverting continuous time process.

03. Data

Our analysis is based on seven different securities. Four of them are the aluminum producers Alcoa (AA), Century Aluminum (CENX), Kaiser Aluminum (KALU) and Rio Tinto (RIO). We use these four companies to create an equal weight portfolio in order to reduce unsystematic risk. We construct a portfolio using Principal Component Analysis[4] (PCA). The ETFS Aluminum ETF (ETF), the GSCI Aluminum Index (G17Y) and the 3 month aluminum futures contract (3MFUT) different aluminum based investment vehicles. Our analysis therefore consists of all the tradable classes of securities related to aluminum, except options.

The data are obtained through Reuters' EcoWin. It has some missing observations which are filled with data from Yahoo! Finance. There are 995 daily return observations in all, from the 23th of April 2007 to the 27th of May 2011. Ideally, our analysis would contain more observations, but because the aluminum ETF was introduced in 2007 it is impossible to get more observations since we want to include an ETF. Since we use data from both US and European exchanges we have to reduce the number of observations, because the different countries have different holydays. We use European data because aluminum futures are no longer traded on the

[3] See Rice(1945)

[4] The methodology is presented in section 05.01

CME Group's exchanges[5] and aluminum ETFs have only been traded in the US since 2008. The number of observations we remove are small, so the effect is neglectable.

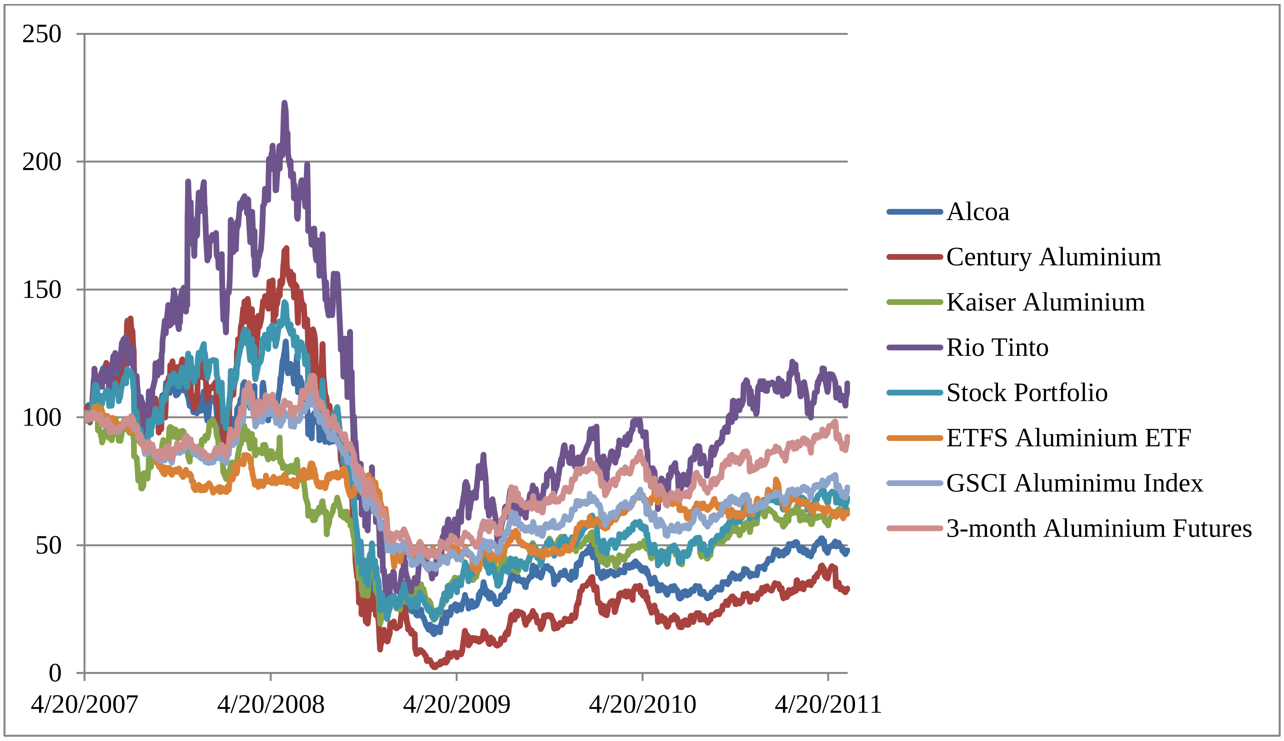


Exhibit 03.01 – Performance of the different securities

RIO and AA are two of the largest aluminum producers traded on the New York Stock Exchange. RIO is the largest aluminum producer in the world, but it is also involved in several other metals and mining businesses. Only ¼ of their revenues in 2010 came from its aluminum operations. Nevertheless, Rio Tinto has approximately the same correlation with 3MFUT as the three other companies[6].

CENX and KALU are very small compared to RIO and AA, but they are nevertheless the 3rd and 4th largest aluminum producers traded at the NYSE with sufficient number of data observations[7]. We choose to include the four stocks in the portfolio to reduce unsystematic risk. The stocks are equally weighted instead of weighted based on market capitalization.

[5] The CME Group consists of the Chicago Mercantile Exchange, The Chicago Board of Trade and the New York Mercantile Exchange (including the COMEX).

[6] See appendix D – Correlation Matrix

[7] Several of the largest aluminum producers in the world are either not traded at the NYSE or have only been traded since 2009/10. We have chosen only NYSE traded stocks to avoid biases caused by different trading hours.

03.01. Descriptive Statistics

The four stocks have a positive mean and median return; however RIO is the only company yielding a positive return over the period (13.2%). The other companies yields a negative return; AA -52.0%, CENX -67.0% and KALU -46.1%. The equal weight portfolio has a return of -31.5% over the period. CENX stands out among these companies as the most volatile stock with a standard deviation (SD) of 6.45. CENX also has the most extreme returns in both tails of the distribution (90.1% and -37.4%). KALU is the least volatile stock with a SD of 3.37. The equal weight portfolio has a SD of 3.65. It is unusual to see that the company with the lowest market capitalization is also the company with the lowest SD. Its SD is even lower than that of the equal weight portfolio (3.75) and the PCA portfolio (3.64).

CENX have the highest kurtosis of the four stocks (36.0), indicating high tail risk, while the others have 8.5 (AA), 7.7 (KALU) and 9.7 (RIO). Both portfolios have kurtosis of 7.5. The distribution of returns for CENX is heavily right skewed (2.03). The other companies' return distributions are mildly skewed; AA has a skewness of .27, KALU -0.42 and RIO 0.21. The equal weight portfolio has a skewness of 0.13 and the PCA generated -0.13. The Jarque-Bera tests rejects normally distributed returns for all four stocks and the portfolio. The distribution of financial returns is rarely distributed normally, but they usually have negative skewness. The stocks' returns therefore deviate some from the general financial returns' distribution functions.

We employ an Augmented Dickey-Fuller (ADF) test to check whether the stock returns unit roots. We also test for autocorrelations by deploying a Ljung-Box test. The Ljung-Box test finds autocorrelation for CENX on a 5% level[8]. The correlation matrix confirms that the stocks are highly correlated. RIO and KALU have the lowest correlation with 0.594 while AA and CENX have the highest correlation (0.735). Because the price of aluminum is so important for the profits of these companies, the high correlation was expected. Financial returns are in general stationary and not autocorrelated. The results from the Ljung-Box tests are therefore in line with the usual properties of financial return series.

The ETF and the G17Y both try to replicate the return of physical aluminum. We therefore expect these two's returns to have approximately identical statistics. However, the analysis

[8] See appendix C – Descriptive Statistics

implies that there are some differences in statistics, which indicates that there might be periods of mispricing. G17Y has a correlation of 0.997 with the 3MFUT. We choose to include both these securities in our analysis because we want to investigate price differences between them.

All the three securities have negative return over the sample period. 3MFUT yields the highest return of -7.7%, while the ETF yields -37.8% and the G17Y -27.4%. All the three securities have a mean return close to zero. 3MFUT have a mean daily return of 0.007%, while G17Y has -0.017% and ETF -0.032%. G17Y and 3MFUT have almost the same maximum and minimum values. 6.11% and -7.22% for the G17Y and 6.91% and -7.18% for 3MFUT. The similar values for the ETF are 9.34% and -6.59%. RIO is the only security yielding a positive return over the sample period. This is probably because the company is well-diversified and involved in several mining businesses. We also notice that stock values declined more than the price of aluminum in our sample period.

The SDs for the three securities are almost identical, 1.75 (G17Y), 1.77 (ETF) and 1.74 (3MFUT), indicating that they have the same level of volatility. However, there are differences in both skewness and kurtosis among the three securities; G17Y and 3MFUT have negative skewness, -.20 and -.18 respectively. The ETF has right skewed distribution with skewness of .34. G17Y and 3MFUT also have about the same kurtosis, 3.93 and 4.00 respectively. The ETF has higher kurtosis (5.20). Investing in physical aluminum appears to have lower tail risk than aluminum stocks. The Jarque-Bera test rejects normally distributed returns for all securities. Likewise, the ADF tests reject unit roots for all securities and indicate stationary returns. None of the securities have autocorrelated returns, according to the Ljung-Box test[9].

[9] See appendix C – Descriptive Statistics

04. Selection of Trading Pairs – A Long-Term Relationship

In this section we present the econometrical methods applied in this paper for investigating the relationships among the aluminum securities. A detailed discussion is conducted about the idea of cointegration and the various tests to identify cointegrated relationships between time series. Before starting the cointegration tests we present a method for making a portfolio using Principal Component Analysis (PCA) that we use throughout this paper. Then we test the order of integration for the time series before applying a simple Engle-Granger test and a more robust Johansen test for each pair of securities. We also apply Granger's Causality tests to decide which security in the pair to be the dependent.

04.01. Principal Component Analysis (PCA)

PCA was invented by Pearson (1901), whilst the best modern reference is Jolliffe (2002). In this paper we utilize PCA to construct a portfolio of stocks. The PCA portfolio is an alternative to the equal weight portfolio in order to reduce the variance and hence reduce risk without reducing expected return. One could have achieved a similar effect using derivatives as a hedging tool, but the PCA approach is chosen because the transaction costs are lower and it does not affect expected return as much as alternative methods.

The objective of the method is to reduce the dimensionality of data whilst preserving as much of the information as possible. The procedure uses an orthogonal transformation to convert a set of data containing correlated variables into a new set of uncorrelated variables called principal components. These components are linear functions of the original data set. The greatest variance by any projection of the data comes to lie on the first coordinate, the second largest variance on the second coordinate and so on. This is achieved by computing a correlation matrix for the data set. We use historical price data on N stocks going back M days in history. The return data is given by the equation

$$R_{i,k} = \frac{S_{i(t_o-(k-1)\Delta \mathrm{t})} - S_{i(t_0-k\Delta\tau)}}{S_{i(t_0-k\Delta\tau)}} \tag{4.1}$$

Where k=1,..,M and i = 1,..,N and S_{it} is the price of stock i at time t.

The elements of the empirical correlation matrix are defined as:

$$\rho_{ij} = \frac{1}{M-1}\sum_{k=1}^{M}\frac{(R_{ik}-\mu_i)(R_{jk}-\mu_j)}{\sigma_i\sigma_j} \tag{4.2}$$

Where μ_i is the mean return of i and σ_i is i's standard deviation. The dimensions of a correlation matrix is usually very high and contains data from years of return history. A problem arising when M is high is that returns occurring years back are considered as important as returns that have occurred over the last weeks. This does not make economic sense and a commonly used solution is to extract the most important data from the data set. Next, the eigenvectors and eigenvalues of the correlation matrix are computed. The eigenvalues are ranked in decreasing order.

$$N \geq \lambda_1 \geq \lambda_2 \geq ......\geq \lambda_N \geq 0 \tag{4.3}$$

The corresponding eigenvectors are denoted by

$$v^{(j)} = (v_1^{(j)}, ..., v_N^{(j)}) \text{, where j = 1,….,N} \tag{4.4}$$

The next step in the procedure is to find out how much to invest in each individual stock; an eigenportfolio must be created.

$$Q_i^{(j)} = \frac{v_i^{(j)}}{\sigma_i} \tag{4.5}$$

The returns from the eigenportfolio is given as

$$F_{jk} = \sum_{i=1}^{N}\frac{v_i^{(j)}}{\sigma_i}R_{ik} \tag{4.6}$$

Laloux, Cizeau, Potters and Bouchaud (2000) points out that the dominant eigenvector is associated with the market portfolio, because all the coefficients $v_i^{(1)}$ (i=1,…,N) are positive. We notice that these weights (eq. 4.5) are inversely proportional to the stock's volatility. This weighting is usually consistent with the capitalization-weighting, since large cap stocks tend to be less volatile than small cap stocks.

The amounts invested in each of the different stocks Q^j are found as shown in equation 4.5. The portfolio is therefore made up of *21.2% Alcoa (short), 32.4% Kaiser Aluminum (long), 57.9% Rio Tinto (long) and a neglectable .1% long-position in Century Aluminum*.

04.02. Long-term Relationship - Cointegration Analysis

Two time series are cointegrated if they share a common stochastic drift and if a linear combination of these variables is stationary. These stationary linear combinations may be interpreted as long-run equilibrium relationships among the securities. The idea is that there is a common force, based on mean-reverting behavior that moves the variables in the same direction over time. Crowder and Wohar (1998) claim that less common trends in a system, the more stable the system is. Cointegration also implies convergence among securities, which might lead to trading opportunities. There is reasonable to believe that aluminum companies' stock prices depend on spot and future prices of aluminum.

The presence of cointegration between securities implies that one of them can be used to forecast the market direction of the other because a valid casual relationship based on the error-correction model exists. Presence of cointegration may therefore limit the benefits from long-run diversification, but accelerate the interest for profitable trading opportunities. There are many different tests for cointegration and most of the work on cointegration relies on Engle and Granger (1987), Johansen (1988), and Johansen and Juselius (1990). In this paper we apply both the Engle-Granger and the Johansen procedure to capture the cointegrated relations.

Before starting the analysis of the long-term co-movements among aluminum securities we need to determine in which levels they are stationary. To investigate stationarity the ADF test is conducted to all price series. According to Banerjee, Hendry, Galbraith and Dolado (1993) among others, the ADF is the most robust test for the presence of autoregressive errors. Appendix H presents the results from the ADF tests for all 7 indices, indicating that all series contain a unit root. However, when tested again in the first differences, all the individual series are stationary. This indicates that the series are integrated of order one, I(1).

04.02.01. The Engle-Granger Cointegration Analysis

The Engle-Granger method tests for cointegration by running an OLS regression and test its residuals for stationarity. Stationary residuals imply that the securities are cointegrated. For the optimal specification one can also include lagged variables; numbers of lags included in the model are selected based on Akaike Information Criteria (AIC). Another important decision is whether to include a trend or not. This paper analyzes both cases, even though it is unlikely that a time trend would be necessary for most financial markets. Equation (4.7) show the regression applied in Engle-Granger. X_i and X_j are the logarithm of the price series for any i and j types of aluminum security.

$$X_{i,t} = \alpha_0 + \beta_1 X_{j,t} + \varepsilon_t. \qquad (4.7)$$

This regression is used supplementary to the Johansen cointegration procedure for comparison. The Engle-Granger approach is easy to apply, but can only estimate up to one cointegration relationship between two securities. Because of this we made various combinations of the aluminum related securities. First we investigate whether there is cointegration relationships among the stocks included in the portfolio. We use AIC to determine the lag length, and test both with and without a linear trend. It might be interesting to take a look at the indices in Exhibit 03.01; the securities are rebased to 100 at the start of the period, simply because this makes them easier to compare. This visualization, which gives an impression that almost all indices are highly cointegrated, can be quite misleading.

Among the four stocks there were slightly changes in the presence of cointegration, due to the inclusion of trend or not. In the model without a linear trend five out of six combinations have cointegration relationship at a significance level between 1 and 10 percent[10]. We assume the model without trend to be more reliable. Since this analysis is built on the concept of autocorrelated residuals it is also interesting look at the residuals graphs. The average value of residuals are zero (always the case for OLS residuals), but it is not easy to spot if the series are slowly mean-reverting or not mean-reverting at all[11].

[10] Cointegration results from Engle-Granger can be found in Appendix I.
[11] Graphs of residuals from Engle-Granger can be found in Appendix J.

The more often the residuals cross the mean, the more likely is it that the series are cointegrated. Exhibit 04.01 illustrates an example of the types of dynamic behavior that we want in our residuals. This is the residuals from RIO and KALU. After 2007 we see no trending data, medium levels of volatility and mean reversion around the equilibrium value of zero.

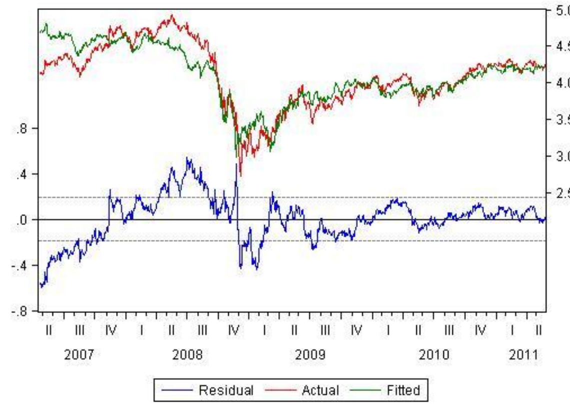


Exhibit 04.01 – Residuals RIO - KALU

These results will not be brought to further investigation, because we do not make trading strategies based on pairs of stocks. There will be too much risk due to company specific event such as bad investments, fines, management trouble etc. This is why we rather use portfolios of stocks in our analysis. In this way some of the unsystematic risk would be diversified, and hence the investment will be safer. The long-run relationships among stocks strengthen our belief that aluminum securities are cointegrated and trading them can be beneficial.

Performing the Engle-Granger test on nine combinations among the portfolio, PCA, ETF, G17Y and 3MFUT gives us four highly significant cointegration relationships; Portfolio is highly cointegrated with G17Y, significant at 1% level. The portfolio is also cointegrated at a significance level of 5% with 3MFUT. The cointegration test is also conducted on the PCA constructed portfolio; the results from these tests show that PCA are strongly cointegrated with both G17Y and 3MFUT.

Exhibit 04.02 illustrates an example of the types of dynamic behavior that we don't want to see. This is the residuals from 3MFUT and G17Y. After mid 2008 we see positive trending data, low levels of volatility and no mean reversion around the equilibrium value. We conclude this section with the four cointegrated pairs below. The Engle-Granger method gives,

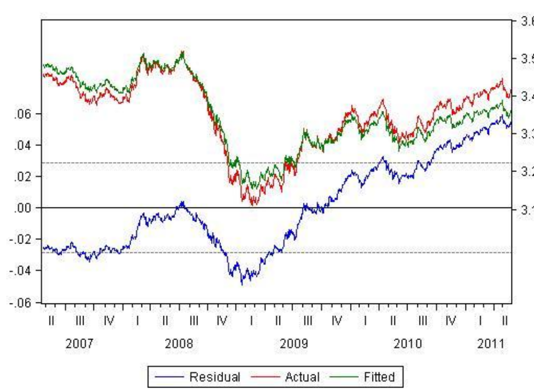


Exhibit 04.02 – Residuals 3MFUT –

$$\text{PORTFOLIO} = 0.01 + 2.00*(\text{3MFUT}) + \varepsilon_t \tag{4.8}$$
$$\text{PORTFOLIO} = 0.09 + 1.88*(\text{G17Y}) + \varepsilon_t \tag{4.9}$$
$$\text{PCA} = 0.12 + 1.58*(\text{3MFUT}) + \varepsilon_t \tag{4.10}$$
$$\text{PCA} = 0.16 + 1.33*(\text{G17Y}) + \varepsilon_t \tag{4.11}$$

Where ε_t is a stationary process. The cointegration coefficient (2.00) in equation 4.8 is the number of units 3MFUT held short, for every unit of the Portfolio held long, so that the pair is mean reverting. The value of the portfolio has an equilibrium value of 0. 01 and fluctuates around this value by forces from ε_t. This approach suffers from several drawbacks, therefore we utilizes the Johansen test which overcome these issues. As mentioned above, these two tests are different; The Johansen test seeks the linear combinations which are the most stationary while Engle-Granger seeks the stationary linear combination that has the minimum variance.

04.02.02. The Johansen Cointegration Analysis

Johansen's procedure uses a unified Vector Autoregressive (VAR) system approach for testing cointegration and can investigate cointegration among several securities. The test relies on the relationship between the rank of a matrix and its eigenvalues. Johansen derived the maximum likelihood estimator of the space of cointegration vectors and the likelihood ratio test of the hypothesis that is has given a number of dimensions.

The first step in Johansen's methodology is to decide the number of lags p in the VAR model. To find the optimal lag length one can use AIC. Then the Johansen procedure estimates the vector error correction model (VECM). This is to determine the number of cointegrating vectors and it is given by

$$y_t = \mu + A_1 y_{t-1} + ... + A_p y_{t-p} + \varepsilon_t \tag{4.12}$$

Where A_t is an nxn matrix of parameters, y_t is the nx1 vector of variables integrated of order one and ε_t is the nx1 vector of Gaussian independently distributed innovations. The VAR equation (4.12) can be reformulated into VECM form, subtracting Y_{t-1} on both sides,

$$\Delta y_t = \mu + \Pi y_{t-1} + \Sigma_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{4.13}$$

Where

$$\Gamma_i = -\Sigma_{j=i+1}^{p} A_j \tag{4.14}$$

$$\Pi = \Sigma_{i=1}^{P} (A_i - I) \tag{4.15}$$

If the coefficient matrix Π has reduced rank (r < n), then there exist nxr matrices α and β each with rank r such that eq. (4.16) and $\beta' y$ is stationary.

$$\Pi = \alpha\beta' \tag{4.16}$$

The elements of α are the vectors of adjustment coefficients, or the adjustment parameters in the VECM, while β is the cointegration vectors. r is the number of cointegration relationships, or linear combinations of y_t. It is important to notice that subtraction of the first differences may not be the most appropriate representation of the data. The numbers of lagged first differences are chosen so that the residuals are not autocorrelated.

The Johansen test is based on the method of maximum likelihood on the equation in (4.13), while the restriction is posed for a given value of r in equation (4.16). The maximum likelihood estimator of β defines the combination of y_{t-1} that yields the r largest canonical correlations of Δy_t with y_{t-1} on the lagged differences.

The next step in the procedure involves testing the hypothesis of the long run relationship. This involves the rank of the long run matrix Π or the number of columns in β, which is equivalent. Johansen proposes two different likelihood ratio tests of the significance. That is the trace test and the maximum eigenvalue test.

Trace test: $$\lambda_{trace}(r) = -T\Sigma_{j=r+1}^{k} \log(1-\hat{\lambda}_j) \tag{4.17}$$

Maximum eigenvalue test: $$\lambda_{\max}(r) = -T \log(1-\hat{\lambda}_{r+1}) \tag{4.18}$$

The trace test tests whether the smallest k-r eigenvalues are significantly different from zero[12]. The null hypothesis of r cointegration vectors against the alternative hypothesis of n cointegration vectors. The maximum eigenvalue test tests the null hypothesis of r cointegration vectors against the alternative hypothesis of r+1 cointegration vectors. T is the sample size and is $\ddot{\lambda}_j$ the i'th largest canonical correlation.

The Johansen test requires more from the analyst, since there are many important decisions regarding selection of the models. To decide the number of lags in the VAR-models we use the Lag Order Selection Criteria based upon statistical tests and criterias such as AIC and Schwartz Information Criteria (SIC). Then we decide if the series have zero means, deterministic trends or stochastic trends. Similarly, the cointegration equation may have intercepts and deterministic trends. In order to carry out the proper cointegration test, we need to decide between five different models of characteristics regarding the VAR equation of securities[13]. The decisions about the deterministic components were based upon the AIC by analyzing the summary of all five models. In this paper we choose between models 2, 3 and 4, since 1 and 5 are rarely used[14]. In some cases we see that small adjustments may have great influence on the conclusions, and therefore we are conservative in our selection process. We only conclude with cointegration where the results are robust and consistent.

The number of cointegrated vectors for the aluminum VARs are as following: It is one vector between the equal weighted portfolio and the 3MFUT and the G17Y. It is also one vector between the PCA generated portfolio and the 3MFUT and the G17Y. These results are the same as for the Engle-Granger test. It is important to notice that one vector is the maximum number of vectors in a model, consisting of two securities, for them to be cointegrated.

An interesting observation is that the 3MFUT and the two replication instruments, G17Y and ETF, are not cointegrated. Even though the G17Y is highly correlated with the 3MFUT, and it seems to be a good tracker of the future price. Since these ETF providers are rolling futures contract to avoid taking delivery they could be considered as an aluminum spot instrument.

[12] The critical values for these tests can be found in Johansen and Juselius (1990).
[13] A summary of the five deterministic trend cases considered by Johansen (1995, p. 80–84) can be found in EViews 7 User Guide II, page 689.
[14] You should use case 1 only if you know that all series have zero mean. Case 5 may provide a good fit in-sample but will produce implausible forecasts out-of-sample. See E-Views guide II for further explanation.

Therefore it could be interesting to analyze the spot price and the future price to see whether the current future price is an efficient predictor of the future spot price, and hence predict the movements of for example G17Y. The future price can be tested against the lagged G17Y and should be cointegrated with a normalized cointegration vector of (1, -1). However, following the AIC for selecting the model, neither the lagged nor the current G17Y are cointegrated with the future price. Given that there were a cointegration relationship the vector would have been something like (1, -0.55), which might indicate that there are some value loss in G17Y and hence the future prices are not an efficient predictor for the expected return from the G17Y. This is consistent with the graph of the trending residuals in Exhibit 04.01 which indicate that G17Y is slowly losing value against 3MFUT. One possible explanation might be the cost of rolling the aluminum contracts to avoid delivery, and the fee to the managers for managing the ETF. However, these pairs will not be analyzed any further as they are disqualified for the pairs trading.

The forces that comove companies in the same sector are underlying fundamental factors. Since the companies in this paper are aluminum producers, it is natural to believe that aluminum prices and inputs for aluminum production link them together. Aluminum production is an energy consuming process, so the aluminum prices are highly exposed to fluctuations in energy prices. Other important factors are the industrial production, bauxite and alumina prices, aluminum storage etc. These factors are perhaps the explanatory variables for both the companies and the aluminum price, and hence the reason for the strong cointegration between the stock portfolios, 3MFUT and G17Y. We expected this result based on the studies of, among others, Lee et al (1985) discover that commodities and stocks pays off in different states, but generates approximately the same return in the long run.

The cointegration vectors can be given an economic interpretation using normalization on the parameters of the cointegrating equations for the different VAR. The normalized vectors in Exhibit 04.03 represent the long run effects imposed by the variables on the selected trading pairs.

VAR	AA	CENX	KALU	RIO	PORTFOLIO	PCA	3MFUT	G17Y	Constant	Trend
PORTFOLIO 3MFUT					1		-1.875*			0.000*
							(15.478)			(-6.195)
PORTFOLIO G17Y					1			-1.878*	1.942*	
								(18.450)	(-9.526)	
PCA 3MFUT						1	-1.702*		-0.135*	
							(11.109)		(6.168)	
PCA G17Y						1		-1.465*	-0.196*	
								(10.118)	(6.985)	

*Rejection of the null hypothesis at 1% significance level

**Rejection of the null hypothesis at 5% significance level

***Rejection of the null hypothesis at 10% significance level

Exhibit 04.03 - Johansen Cointegration Equation with Normalized Parameter Estimation

The parentheses in the Exhibit 04.03 are the t-statistics and all variables are highly significant. Grangers Causality gives an insight into the dynamics of the cointegration relationship for a given pair of securities. It reveals which security is the dependent variable in the cointegration equation. The other security is therefore driving the price changes in the dependent variable. It is important that both securities in the sequences do not Granger Cause each other; we wish to constrain the contemporaneous effects of the series equal to zero so that the equations become identifiable. The results from the Granger Causality are displayed in Exhibit 04.04[15]. Results are robust and significant at a 1 percent level. From the results we see that there is always one security leading the other and we have no occurrences where securities in the pair lead each other.

Trading Pair	Direction of Causality	P-Value	Direction of Causality	P-Value
PORTFOLIO - 3MFUT	L3MFUT does not lead LPORTFOLIO	0.5357	LPORTFOLIO does not lead L3MFUT	0.0049
PORTFOLIO - G17Y	LG17Y does not lead LPORTFOLIO	0.2634	LPORTFOLIO does not lead LG17Y	0.000
PCA - 3MFUT	L3MFUT does not lead LPCA	0.5722	LPCA does not lead L3MFUT	0.000
PCA - G17Y	LG17Y does not lead LPCA	0.5911	LPCA does not lead LG17Y	0.000

Exhibit 04.04 – Results from the Granger Causality test

We have analyzed the long term relationship among aluminum securities and are ready to move on with our analysis. We identified four cointegrated pairs which we use to create trading rules.

[15] The results can be found in Appendix M.

05. Methodology - Statistical Arbitrage

Statistical Arbitrage is a term used to describe a variety of different trading strategies. Common features for these strategies are that they are market or sector neutral, the trading signals are systematic and rules based and designed with econometrical techniques in order to provide signals for execution. Trading strategies are often based on a mean-reversion principle, but can also be designed using factors as momentum, spillovers, lead/lag effect etc. Common for all the trading strategies is that there is a statistical mispricing in one or more securities based on the expected or fundamental value of these securities. However, it is important to notice statistical arbitrage is not true arbitrage because it does not guarantee positive return.

The objective is to create high risk adjusted returns, which is uncorrelated with the stock and commodity markets. Holding periods range from seconds to days, weeks or even longer. Pairs trading is assumed to be the ancestor of statistical arbitrage. If securities P and Q are securities with similar characteristics, like two oil companies, one expects the returns of the two stocks to track each other after controlling for beta. Accordingly, if P_t and Q_t denote the corresponding price time series, then we can model the system as

$$\frac{P_t}{P_0} = \frac{Q_t}{Q_0} + \varepsilon_t \tag{5.1}$$

Where ε_t is a stationary mean-reverting process which will be referred to as the cointegration residual, or residual for short, in the rest of the paper. The model suggests a contrarian investment strategy in which we buy X dollar of P and short X dollars of Q if ε_t is below some predefined value. It also suggest doing the opposite; short P and buy Q if ε_t is above some predefined value. These values are based on statistics and are usually set to be a moving average or η standard deviations from mean. Statistical arbitrage models usually have a rule to determine when to close positions and may also have a stop loss rule.

The trading strategy is expected to produce a positive return as P and Q converge. the mean-reversion paradigm is typically associated with market over-reaction: securities are temporarily mispriced with respect to one or several reference securities[16]. Another possibility is to consider scenarios in which one of the securities is expected to out-perform the other over a significant period of time. In this case the cointegration residual should not be stationary. This paper only focuses on cointegrated time series.

There are several ways of extending a statistical arbitrage pairs trading model from conventional pairs trading. One of them is to replace one of the stocks with a portfolio of stocks, often represented by an ETF, where the stock and the ETF is traded using the same rules as mentioned above. Other extensions may be to trade different integrated sectors, such as oil service and oil producers. One can also trade stocks and commodities, such as aluminum stocks and aluminum futures. The main importance when deciding which securities to trade is to make sure they are cointegrated, or ε_t will not be mean-reverting.

05.01. Statistical Arbitrage for Cointegrated Securities

Throughout this chapter we explain trading strategies that are based on the principles described above. In this section, we design trading strategies to exploit cointegrated securities. From equation 5.1 we know that

$$\frac{P_t}{P_0} = \frac{Q_t}{Q_0} + \varepsilon_t$$

When P_t and Q_t are cointegrated, ε_t is stationary. This implies that

$$\lim_{T\to\infty}\left(\frac{1}{T}\sum_{t=0}^{T}\varepsilon_t\right) = \mu \tag{5.2}$$

and

$$\lim_{T\to\infty}\left(\sqrt{\frac{1}{T-1}\sum_{t=0}^{T}(\varepsilon_t-\mu)^2}\right) = SD \tag{5.3}$$

[16] See Lo and MacKinley (1990)

Because μ is the long-run equilibrium for the relative price of P_t and Q_t we believe that ε_t will at some point in time, regardless of its current value, revert to μ.

Because the data set only consists of about four years of data, μ and SD has to be calculated using the same data set as we use to trade. In order to create unbiased trading strategies we use the following variables instead of μ and SD:

$$\mu_t = \frac{1}{T}\sum_{t=0}^{T}\varepsilon_t \tag{5.4}$$

$$SD_t = \sqrt{\frac{1}{T-1}\sum_{t=0}^{T}(\varepsilon_t - \mu_t)^2} \tag{5.5}$$

Strategy #1 – Modified Bollinger Bands

The first trading strategy is based on the following trading rules:

(I) Buy P – Short Q if $\varepsilon_t < \mu_t - n * SD_t$

(II) Buy Q – Short P if $\varepsilon_t > \mu_t + n * SD_t$

(III) Close position when ε_t crosses μ_t

This is a simple trading rule. It tells the speculator to buy the undervalued security and sell the overvalued security short and then wait until the securities are fairly priced relative to each other before closing positions and wait for a new trading opportunity. Because μ_t is likely to be volatile as T is low, trading usually do not start before T=100.

This is an extension of the trading rule known as Bollinger Bands[17]. The difference between this trading rule and Bollinger Bands is that Bollinger Bands are based on a moving average instead of an average of all the observations. This is because conventional Bollinger Bands are designed to be used on all sorts of securities. The Modified Bollinger Band strategy can only be applied when trading two cointegrated securities.

[17] See Bollinger (2002)

Strategy #2 – Modified Bollinger Bands with Stop Loss

Like Strategy #1 this strategy tries to exploit the properties of cointegrated securities, but in a more cautious way. It is based on the following set of rules:

(I) Buy P – Short Q if $\varepsilon_t < \mu_t - n * SD_t$ and $\varepsilon_{t-1} > \mu_{t-1} - n * SD_{t-1}$

(II) Buy Q – Short P if $\varepsilon_t > \mu_t + n * SD_t$ and $\varepsilon_{t-1} < \mu_{t-1} + n * SD_{t-1}$

(III) Close position when ε_t crosses μ_t or $\varepsilon_t > \mu_t - n * SD_t$ or $\varepsilon_t < \mu_t + n * SD_t$

Like strategy number one, it is a modified version of Bollinger Bands designed for cointegrated securities. The main difference between this strategy and the former is that this includes a stop loss which is activated when ε_t is outside the bands.

Strategy #3 – Moving Averages (MA)

One of the most common ways of trading securities is by using one or two moving averages. According to Taylor (2005), this can be done because security prices tend to move in the same direction for a period of time. A speculator may exploit this by buying a security when the price is higher than the moving average or sell the security short or hold a neutral position when the price is below the moving average.

The relative price of cointegrated securities is expected to behave the opposite way; if the relative price has been rising, it is expected to decline and vice versa. Because the relative price of cointegrated securities has this property, we want to test if an inverted version of a conventional non-exponential MA strategy can be applied. The trading rules are therefore quite simple:

(I) Hold a short position in P and a long position in Q if $\frac{1}{L}\sum_{i=1}^{L}\varepsilon_t < \varepsilon_t$

(II) Hold a long position in Q and a short position in P if $\frac{1}{L}\sum_{i=1}^{L}\varepsilon_t > \varepsilon_t$

A challenge when creating a MA rule is setting the length, L of the moving average. Taylor (2005) uses L=5 while Brock et al (1992) claims that L should be at least 50. Each pair has its own unique properties, hence it is impossible to set a universal value of L applying to all pairs.

There are two reasons why one should be cautious using MA based strategies; (I) The return is very sensitive to changes in L. (II) L is chosen *ex post*.

06. Empirical Results – Statistical Arbitrage

This part of the paper presents the empirical results from our analysis of the different trading strategies, deployed to trade the cointegrated securities. The returns have been calculated for the entire period and for each 100 days sub period.

From Section 4 we know that four pairs are cointegrated and one of the securities leads the other. This relationship was derived from the Granger Causality test. We recall the Engle-Granger equation from Section 4 that for every unit of the dependent security we buy, we need to short β_1 units of the explanatory security. In this way the pair will be mean reverting. Note that we do not use the coefficients (weights) from section 4. We divide the sample set in two parts. In this way we can test our results out of sample and the results will not be biased. The trading strategies are also performed with equal weights.

06.01. Modified Bollinger Bands #1

Because it is difficult to determine how many standard deviations (SDs) that should be used to create trading rules for this strategy we deploy six different numbers of SDs; 0.5, 1, 1.5, 2, 2.5 and 3. This will cause a bias in our analysis because the number of standard deviations has to be chosen *ex ante* when applied to real life trading. Our analysis will therefore only give an indication of how many SDs to use when trading aluminum securities and whether or not this is a desirable strategy to use. The term *strategy* is used to describe the three different approaches (MBB#1, MBB#2, MA), while the term *trading rule* is used to describe the different rules within the strategies. For instance to trade securities when the relative price crosses two SDs. The results presented in this section are from the simulations where we used the weights obtained from the Engle-Granger tests[18]. These weights are obtained from the first 500 observations in the data set. The rest of the data set is used to perform an out-of-sample test of our strategies.

[18] We also performed the same test based on equal weights. These can be found in Appendix N

06.01.01. PCA – 3MFUT

Our analysis shows that the MBB#1 strategy yields very good returns for this pair. We can see from Exhibit 06.01 that the trading rules applying the lowest number of SDs are yielding the highest returns. This is because there is low volatility in this pair and there is no trading above 1.5SD from the mean. The rules perform well since there is only one sub period with negative return. The three trading rules with trading in the period yields on average 17.3%. However, there are only 9 nine trades during the 496 trading days due to the low volatility in the pair[19].

PCA-3MFUT	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
500-600	11.78%	6.08%	12.97%	0.00%	0.00%	0.00%	5.14%
600-700	-0.19%	3.78%	0.00%	0.00%	0.00%	0.00%	0.60%
700-800	3.92%	3.92%	0.00%	0.00%	0.00%	0.00%	1.31%
800-900	7.13%	0.00%	0.00%	0.00%	0.00%	0.00%	1.19%
900-996	2.05%	3.24%	0.00%	0.00%	0.00%	0.00%	0.88%
Tot return	24.37%	14.41%	12.97%	0.00%	0.00%	0.00%	8.63%
SD of returns	4.18%	1.96%	5.19%	0.00%	0.00%	0.00%	1.89%
SharpeRatio	5.834	7.356	2.500	-	-	-	5.230

Exhibit 06.01 – The return for the different trading rules based on MBB#1 utilized on the pair PCA - 3MFUT

06.01.02. PCA –G17Y

We notice that the MBB#1 based rules yields a good return for this pair. The trading rules which generate trading signals are yielding a return of 13.7% on average. Only the trading rules with the three lowest number of SDs generate trade signals and only one of the sub periods yield a negative return for one of the trading rules. This is the same results as the ones obtained for PCA-3MFUT. Because the G17Y is mirroring the 3MFUT we expected to get similar results. The numbers of transactions are also very low for this set; only 8 trades are conducted over the period.

PCA-G17Y	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
500-600	10.96%	5.84%	13.48%	0.00%	0.00%	0.00%	5.05%
600-700	2.98%	0.00%	0.00%	0.00%	0.00%	0.00%	0.50%
700-800	1.13%	0.00%	0.00%	0.00%	0.00%	0.00%	0.19%
800-900	-3.38%	0.88%	5.03%	0.00%	0.00%	0.00%	0.42%
900-997	1.11%	1.11%	1.11%	0.00%	0.00%	0.00%	0.55%
Tot return	12.88%	7.95%	20.51%	0.00%	0.00%	0.00%	6.89%
SD of returns	4.70%	2.18%	5.13%	0.00%	0.00%	0.00%	2.00%
SharpeRatio	2.743	3.639	4.002	-	-	-	3.461

Exhibit 06.02– The return for the different trading rules based on MBB#1 utilized on the pair PCA – G17Y

[19] The number of transactions for each trading strategy and rule can be found in Appendix O.

06.01.03. PORTFOLIO - 3MFUT

When comparing this pair to PCA-3MFUT we notice that this pair is more volatile and has a lower average return than PCA-3MFUT. We also notice that the trading rule using 2SDs is generating trading signals, unlike its PCA adversary. This is because Century Aluminum is assigned a higher weight in the equal weight portfolio than in the PCA portfolio. Century Aluminum is a small aluminum company and its stock is more volatile than the others. Six of the sub periods are yielding a negative return; this pair is less robust than PCA-3MFUT.

PORTFOLIO-3MFUT	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
500-600	17.37%	10.09%	10.09%	10.09%	0.00%	0.00%	7.94%
600-700	-4.71%	-0.04%	0.00%	0.00%	0.00%	0.00%	-0.79%
700-800	4.26%	8.27%	15.65%	0.00%	0.00%	0.00%	4.70%
800-900	-6.05%	-2.14%	1.74%	0.00%	0.00%	0.00%	-1.07%
900-997	-8.56%	-8.56%	-8.56%	4.58%	0.00%	0.00%	-3.52%
Tot return	0.18%	6.62%	18.44%	15.13%	0.00%	0.00%	6.73%
SD of returns	9.50%	6.88%	8.38%	3.99%	0.00%	0.00%	4.79%
SharpeRatio	0.019	0.962	2.199	3.788	-	-	1.742

Exhibit 06.03 - The return for the different trading rules based on MBB#1 utilized on the pair PORTFOLIO – 3MFUT

06.01.04. PORTFOLIO – G17Y

The average return for the trading rules that gives a trade signal is 17.9%. On average all the nine sub periods yields a positive return. We notice that the differences in return deviates substantially the PORTFOLIO-3MFUT. This is because the residuals of this pair are fluctuating around the mean with smaller amplitude than PORTFOLIO-3MFUT. This is because the value of G17Y is declining as a percentage of 3MFUT over time, and hence it will not yield as many trading signals.

PORTFOLIO-G17Y	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
500-600	9.08%	10.76%	10.76%	10.76%	0.00%	0.00%	6.89%
600-700	8.30%	0.00%	0.00%	0.00%	0.00%	0.00%	1.38%
700-800	6.56%	0.00%	0.00%	0.00%	0.00%	0.00%	1.09%
800-900	6.76%	0.00%	0.00%	0.00%	0.00%	0.00%	1.13%
900-997	0.41%	4.09%	0.00%	0.00%	0.00%	0.00%	0.75%
Tot return	34.94%	15.29%	10.76%	10.76%	0.00%	0.00%	11.96%
SD of returns	3.05%	4.20%	4.30%	4.30%	0.00%	0.00%	2.64%
SharpeRatio	11.443	3.637	2.500	2.500	-	-	5.020

Exhibit 06.04 - The return for the different trading rules based on MBB#1 utilized on the pair PORTFOLIO – G17Y

06.02. Modified Bollinger Bands #2

06.02.01. PCA – 3MFUT

Applying MBB#2 to this pair yields a lower return than MBB#1, but the risk adjusted return is almost equal. We notice that three of the trading rules yield exactly the same return for this pair. The is because the trading starts with a buy signal. The "neutral" signal is given when crossing the mean, which is equal for all trading rules. We also notice that one of the trading rules (SD=0.5) yields a negative return because the cointegration residual is crossing the 0.5SD line several times before crossing the mean, triggering stop loss. This is the least volatile pair to trade with a MBB strategy.

PCA-3MFUT	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
500-600	0.00%	0.84%	1.76%	1.76%	1.76%	0.00%	1.02%
600-700	-4.34%	2.99%	2.99%	2.99%	2.99%	0.00%	1.27%
700-800	-2.54%	0.63%	0.63%	0.63%	0.63%	0.00%	-0.01%
800-900	2.54%	2.54%	2.54%	2.54%	2.54%	0.00%	2.12%
900-996	4.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.72%
Tot return	-4.29%	7.17%	8.14%	8.14%	8.14%	0.00%	4.55%
SD of returns	3.18%	1.16%	1.13%	1.13%	1.13%	0.00%	1.29%
SharpeRatio	-1.350	6.190	7.222	7.222	7.222	-	5.301

Exhibit 06.05 - The return for the different trading rules based on MBB#2 utilized on the pair PCA – 3MFUT

06.02.02. PCA – G17Y

This is the trading pair yielding the lowest return. Adjusted for risk it is the second lowest yielding pair. The return is highly dependent on the number of SDs selected. We notice that the rule applying 1.5SDs is yielding the highest return in four of the six sub periods while none of the strategies with a higher SD yields any return at all. The is because when applying 1.5SDs a trading signal is only given twice and is not trigging the stop loss.

PCA-G17Y	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
500-600	7.23%	4.03%	12.57%	0.00%	0.00%	0.00%	3.97%
600-700	1.16%	0.00%	0.00%	0.00%	0.00%	0.00%	0.19%
700-800	-3.08%	0.00%	0.00%	0.00%	0.00%	0.00%	-0.51%
800-900	-7.63%	-0.32%	2.56%	0.00%	0.00%	0.00%	-0.90%
900-996	-4.11%	-0.84%	1.11%	0.00%	0.00%	0.00%	-0.64%
Tot return	-2.56%	3.70%	15.49%	0.00%	0.00%	0.00%	2.77%
SD of returns	5.10%	1.75%	4.76%	0.00%	0.00%	0.00%	1.93%
SharpeRatio	-0.502	2.110	3.256	-	-	-	1.621

Exhibit 06.06 - The return for the different trading rules based on MBB#2 utilized on the pair PCA – G17Y

06.02.03. PORTFOLIO – 3MFUT

This pair is one of the poorest performing. It has the second lowest return and the lowest risk adjusted return. We can see in Exhibit 06.07 that a change of 0.5SDs can change the return by more than 30%. The return is also changing substantially from one time period to the next, making this a high risk strategy which does not seem to pay off accordingly. There are 17 trades for the 1SD rule, indicating high volatility.

PORT-3MFUT	**.5SDs**	**1SDs**	**1.5SDs**	**2SDs**	**2.5SDs**	**3SDs**	**Average**
500-600	6.37%	6.55%	11.73%	13.09%	0.00%	0.00%	6.29%
600-700	-2.53%	-9.88%	0.00%	0.00%	0.00%	0.00%	-2.07%
700-800	1.04%	-0.78%	16.22%	0.00%	0.00%	0.00%	2.75%
800-900	-2.55%	-6.08%	0.06%	0.00%	0.00%	0.00%	-1.43%
900-996	0.00%	-2.11%	-12.20%	0.91%	0.00%	0.00%	-2.23%
Tot return	2.08%	-10.41%	29.92%	13.09%	0.00%	0.00%	5.78%
SD of returns	3.27%	5.52%	10.00%	5.16%	0.00%	0.00%	3.99%
SharpeRatio	0.636	-1.886	2.993	2.538	-	-	1.070

Exhibit 06.07 - The return for the different trading rules based on MBB#2 utilized on the pair PORTFOLIO – 3MFUT

06.02.04. PORTFOLIO – G17Y

In this pair there are 27 trades for the 0.5SD rule. The trading rule applying 0.5SDs is yielding positive returns in all sub periods except the last. In the other rules only a few trades are executed. Because only a few trades are being executed the returns are fairly stable relative to the other pairs for the MBB#2 strategy.

PORT-G17Y	**.5SDs**	**1SDs**	**1.5SDs**	**2SDs**	**2.5SDs**	**3SDs**	**Average**
500-600	1.02%	9.81%	13.10%	13.23%	0.00%	0.00%	6.19%
600-700	5.59%	0.00%	0.00%	0.00%	0.00%	0.00%	0.93%
700-800	1.11%	0.00%	0.00%	0.00%	0.00%	0.00%	0.19%
800-900	5.84%	0.00%	0.00%	0.00%	0.00%	0.00%	0.97%
900-996	-5.45%	0.85%	0.00%	0.00%	0.00%	0.00%	-0.77%
Tot return	13.80%	9.81%	13.10%	13.23%	0.00%	0.00%	8.32%
SD of returns	4.10%	3.85%	5.24%	5.29%	0.00%	0.00%	3.08%
SharpeRatio	3.364	2.546	2.500	2.500	-	-	2.728

Exhibit 06.08 - The return for the different trading rules based on MBB#2 utilized on the pair PORTFOLIO – G17Y

06.03. Moving Average Based Rules

The trading rules based on moving averages yields variable results depending on how many daily observations included in the moving average. For most of the trading rules, there does not appear to be a pattern or a trend regarding which average to choose. Here we present the results from the four cointegrated pairs.

06.03.01. PCA-3MFUT

The average return for these trading rules is 15.0% for the entire period with a standard deviation of 13.0%. Only 8 of the 49 MAs yield a negative return over the period. The return appears to be increasing as the number of observations in the average is increasing. The volatility of the different MAs appears to be constant over time. The average return pr. 100 day period is higher for the equal weight model than the Engle Granger based model.[20]

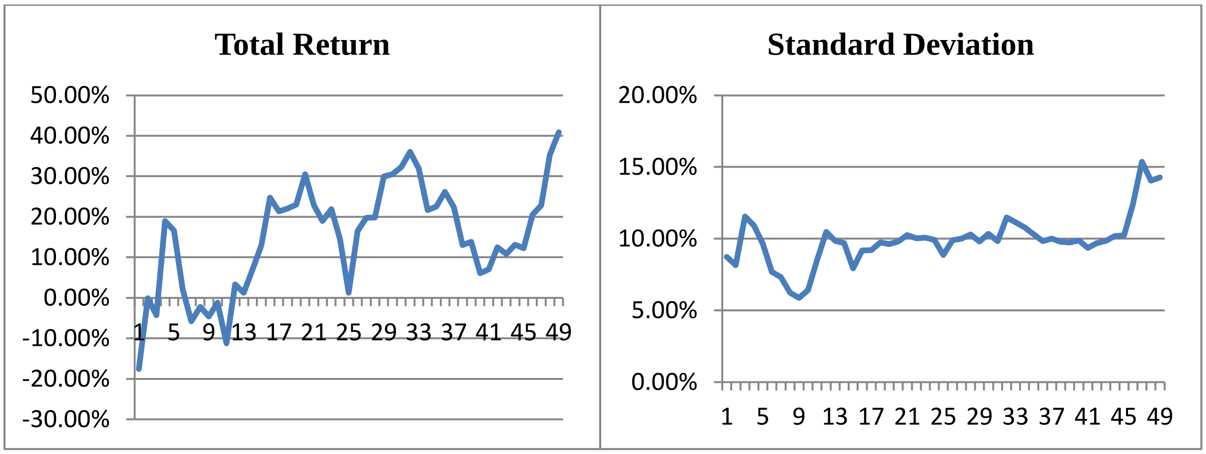


Exhibit 06.09– The return and standard deviations for the different MAs used to trade PCA-3MFUT.

06.03.02. PCA-G17Y

The average return for these trading rules is 14.3% over the entire period with a standard deviation of 13.4%. Seven of the trading rules yield a negative total return, all of them among the shortest averages. The return appears to be increasing as the size of the moving average is increasing, which is also the case for the other rule using the PCA portfolio. The volatility of this pair is varying more across the averages than the other PCA pair. The average return pr. 100 day period is 6.7%, compared to 2.2% for the equal weight strategy.

[20] Average returns for the MA rules can be found in Appendix P (Equal weight) and Q (Engle-Granger weighted). Transaction costs and number of trades can be found in Appendix R (Equal weight) and S (Engle-Granger weighted).

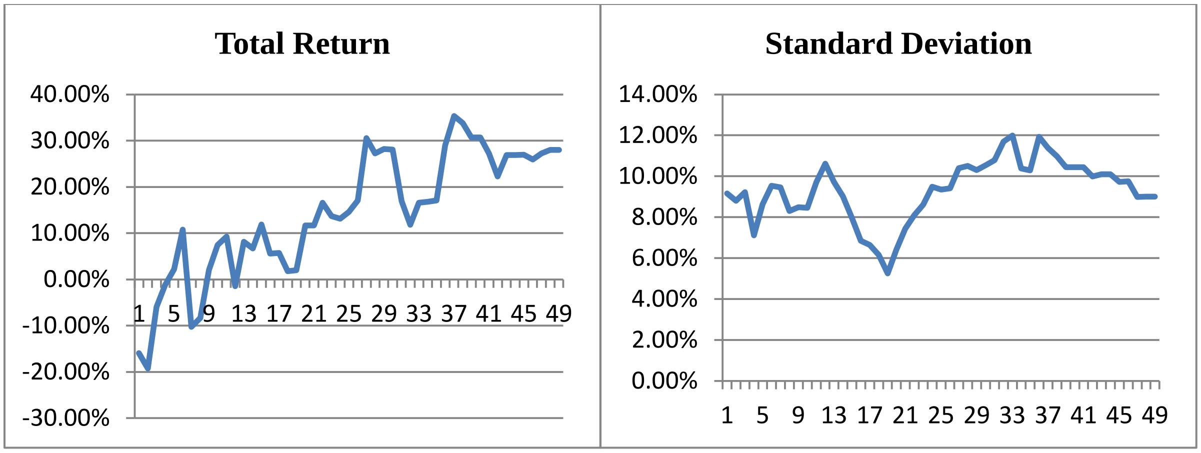


Exhibit 06.10– The return and standard deviations for the different MAs used to trade PCA-G17Y.

06.03.03. PORTFOLIO-3MFUT

The average return from these trading rules is 6.6% over the entire period, with a standard deviation of 6.4%. The average return per 100 day period is 3.2%, compared to 11.8% for the equal weight strategy. Eight of the averages applied yield a negative return. The return is varying substantially from one average to the next, indicating that the strategy is fragile.

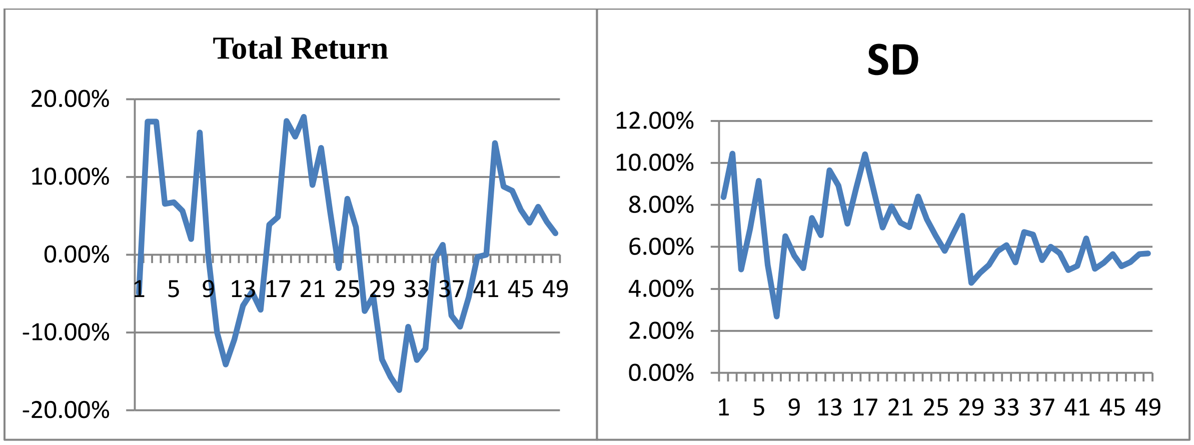


Exhibit 06.11 – The return and standard deviations for the different MAs used to trade PORTFOLIO-3MFUT.

06.03.04. PORTFOLIO – G17Y

This pair yields an average return of 6.6% with a standard deviation of 5.3%. The volatility is low, but it varies substantially depending on the size of the MA. The average return pr. 100 day period is 3.2%, compared to 5.8% for the equal weight rule. The volatility tends decrease as the size of the MA is increasing. The return is varying substantially between rules, indicating a fragile strategy. Eight of the MAs yield a negative return.

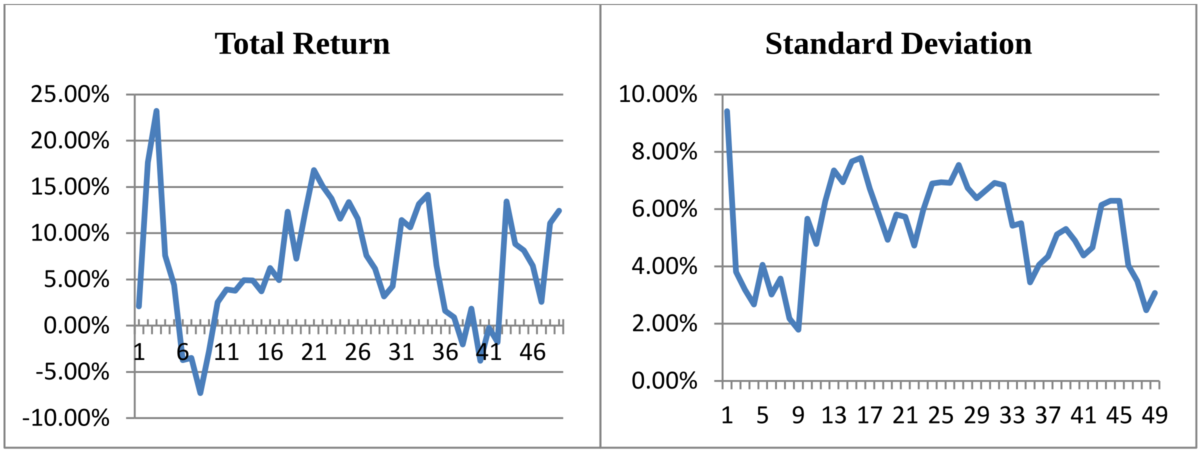


Exhibit 06.12– The return and standard deviations for the different MAs used to trade PORTFOLIO-G17Y.

06.04. The Performance of Aluminum Securities

The peers our trading strategies are compared against are passive unleveraged long positions in the securities which we use to construct trading rules.

SUB PERIOD	PORTFOLIO	ETF	G17Y	3MFUT	Average
0-100	-0.29%	-19.33%	-17.28%	-16.31%	-13.30%
100-200	21.51%	2.29%	15.29%	17.47%	14.14%
200-300	-3.28%	-3.29%	15.65%	17.91%	6.75%
300-400	-75.42%	-33.83%	-55.67%	-53.98%	-54.73%
400-500	20.52%	-12.78%	-3.79%	0.26%	1.05%
500-600	33.25%	2.30%	21.44%	25.07%	20.52%
600-700	11.15%	35.66%	15.36%	18.65%	20.21%
700-800	-4.54%	3.86%	-3.18%	-1.34%	-1.30%
800-900	31.39%	5.08%	9.47%	11.61%	14.38%
900-996	0.05%	-10.78%	3.97%	5.66%	-0.27%
Total Return	-35.48%	-37.77%	-27.37%	-7.67%	-27.07%
SD of returns	29.50%	17.43%	21.68%	22.11%	22.68%
SharpeRatio	-1.2027	-2.1670	-1.2625	-0.3469	-1.1936

Exhibit 06.13 – The return of the different securities used in the trading models

From Exhibit 06.13 we can see that a buy-and-hold investment in any of the securities would have yielded a return lower than all of the trading strategies. The trading strategies in our analysis yield a significantly[21] higher return than a passive investment in the traded securities except for the futures contract. There is some bias in favor of the futures contract; it does not include rolling costs which a speculator would incur because he would not take physical delivery of the aluminum.

[21] At a 99% significance level

07. Discussion

Our motivation to write this paper was to investigate long-term relationships between aluminum securities, and to investigate if we could make profitable trading strategies based on these relationships. In the introduction we talked about how speculation and frictions in the financial markets drive prices on commodity securities away from their fundamental value and that there should be forces that drive these securities back to equilibrium. Our results show that aluminum stocks, GSCI Aluminum Index and the three-month futures contract have a significant long-term relationship. This is in line with previous studies of the relationships between stocks and commodities[22]. Our techniques are not particularly advanced and our procedure is carefully explained throughout the paper. We therefore believe that this research can be useful and applicable for speculators with an intermediate knowledge of statistics that seeks to explore new trading strategies rather than just buy and hold.

After identifying the cointegrated securities we examined whether it was possible to construct mechanical trading rules for the cointegrated pairs that yields significantly higher risk adjusted returns than buy-and-hold investments in the same securities. Buy-and-hold strategies for a broad specter of securities will only generate positive returns in an upward trading market, while we seek to generate profits regardless of the market direction. We used only cointegrated pairs for statistical arbitrage modeling. The models are based on the assumption that the cointegration residual is mean reverting and has long-run equilibrium value. Before applying our trading strategies we made some interesting observations during the empirical analysis of the securities. We would like to discuss these findings before continuing with the conclusion of our trading rules.

07.01. Conclusions

G17Y and 3MFUT are not cointegrated, even though they have a very high correlation (0.997). This pair is therefore not suited for cointegrated pairs trading as their series drift apart in the long run. We also find that the 3MFUT is not an efficient predictor for the lagged G17Y, indicating some value losses in G17Y. The value loss in G17Y relative to 3MFUT is due to transaction

[22] See for example Huang, Masulis and Stoll (1996), Büyûksahin, Haigh and Robe (2008), Hammoudeh, Dibooglu and Eleisa (2002), or Haigh, Harris, Overdahl, and Robe (2007)

costs and management fee of G17Y. However, the G17Y is a good replicator of the physical aluminum return and can be a good substitute in the short run.

The aluminum ETF included in this analysis does not seem to replicate the physical aluminum return. The security is volatile, has lower correlation and has no long run relationship with other aluminum securities. This should be a warning to speculators. The ETF management claims that the ETF replicate the return of physical aluminum, while our analysis indicates that it does not. We therefore advise the speculators to investigate ETFs performance and statistics before investing in such securities.

The two different categories of trading strategies we apply have been chosen based on properties that cointegrated securities have:

(I) The relative price will at some point in time revert to its long-run equilibrium

(II) If one security has been outperforming the other in the past, then the underperformer is likely to outperform the former outperformer in the future.

The modified Bollinger Band (MBB) rules are based on (I), while the moving average rules are based on (II). As far as we know are we the first to perform an empirical research investigating pairs trading opportunities in the aluminum securities market, therefore we do not have any comparable contributions. However, we believe that our models and main findings can be applicable in other commodity markets. Here are our main findings applying our trading models.

The MBB trading strategies with weights obtained from the Engle-Granger test yields higher returns than the equal weight strategies across the board. These strategies have both higher returns and are less volatile than the equal weight pairs. The Engle-Granger weighted strategies are also yielding significantly higher return than a buy-and-hold strategy for three of the four securities/portfolios. For seven of the eight applications of the Engle-Granger weights the first sub period yields the highest return. This implies that speculators employing a trading strategy based on weights from an Engle-Granger test should recalculate the appropriate weights more often than is done in this paper.

The MBB#1 strategies yield both higher returns and risk adjusted returns than the MBB#2 strategies for all four pairs. Both models try to exploit the properties of cointegrated securities, but MBB#2 in a more cautious way. However, our results confirm that the MBB#1 strategy is a better strategy for cointegrated aluminum pairs. We expected lower returns, because the MBB#2 strategies have lower risk due to the stop loss rule. MBB#2 may therefore be more suitable for speculators with higher risk aversion.

The pair including the PCA portfolio has lower average standard deviations for all trading strategies. We can therefore confirm that the Principal Component Analysis procedure is successful in reducing variance and hence reduce risk. However, the pair including the PCA portfolio does not provide a significant higher risk adjusted return. The Sharpe ratio is higher for the pair including the PCA portfolio for about half of the trading rules.

All cointegrated pairs yield a significantly[23] *higher return than a buy-and-hold strategy for the traded securities.* This is the case for all three trading strategies applied in this paper. We therefore believe that cointegrated pairs trading strategies can be applied as a good alternative to a buy-and-hold strategy. This is important in today's markets with lack of trends in either direction. There is much uncertainty in the global financial markets due to highly unstable financial environment and debt situations.

We acknowledge the fact that it is hard to choose the number of standard deviations for the Modified Bollinger Band strategies, but we get strong indications that the number of standard deviations should be 0.5 or 1. Applying a higher number of SDs rarely generates any trading signals. Prior to identifying a long-run equilibrium, our equal weighted MBB strategies usually yield negative returns. We notice that due to the negative average returns, the average loss is decreasing as the number of SDs in the trading rule is increasing. This is because the number of days a speculator is holding securities instead of cash is reduced when the number of SDs is increased.

We also acknowledge the fact that it is hard to choose the correct number of days to include in the moving average. For some of the pairs it appears to be best to use a high number of observations (40-50), but this is difficult to determine in advance. However, the strategy generates highly unstable returns, so without further improvements we suggest the speculator to diversify with applying several moving averages.

[23] At the 99% level.

It is important to find the correct trading strategy for the cointegrated pair. Even though the pairs are cointegrated, the trading strategy will not guarantee positive returns. It is important to investigate the pairs' behavior and test what strategy that will perform good over time. We can notice from our results that the return generated is highly different among the three strategies. Our analysis indicates that MBB#1 yields the highest returns for cointegrated aluminum securities.

Transaction costs do not play a significant role in returns. In this analysis we have included all transaction costs incurring when trading. This makes our results more realistic. The number of trades is low for all MBB strategies and the transaction costs only account for about 1-2 percent. The commission fee for placing trades has almost diminished been substantially reduced after the introduction of online trading platforms one can now execute trades for less than 0.05%.

07.02. Weaknesses and Further Improvements

There are several ways to extend and improve our analysis. There is a bias in the analysis against the equal weight MBB strategies; we do not have enough observations to calculate the long-run equilibrium before we start to trade because the price data for the traded securities do not go further back than 2007. We notice that for two of the traded pairs; PCA-3MFUT and PCA-G17Y, there seems to be an identification of a long-run equilibrium towards the end of the sample period. In the periods after the mean seems to have reached its long-run equilibrium, only seven of the 28 sub periods yield a negative return. We therefore believe that a longer sample period is needed in order to examine whether these strategies will yield good risk adjusted returns.

The ETF included in this paper does not replicate the daily returns of aluminum as well as expected. New aluminum ETFs have been launched after the one by ETF Securities, but the number of observations is too few to include them in an analysis at this point in time.

In order to get a good comparison of our trading strategies with passive "buy-and-hold" investments we should have had a longer sample period. This because stocks and commodities usually yield positive returns in the long run, but in our sample period the return in negative. Unfortunately this was not possible because there does not exist sufficient price history for the ETF and CENX.

The Engle-Granger analysis suffers from several drawbacks. Financial return series violates many of the assumptions for obtaining the best linear unbiased estimator in OLS. However, the coefficients were fairly similar to the ones obtained from the Johansen test and our results were good. We therefore conclude that coefficients are better than the equal weighted. However, we do not deny that there are better methods to obtain these weights.

To improve our analysis we suggest including more securities in the analysis. One should investigate the possibility to make an algorithm that continuously updates the weights of securities. It could also be interesting applying the same models on higher frequency data, since intraday trading will reduce the risk and that there is more likely to be mispricing intraday.

08. References

Abdi, H., and Williams, L.J. (2010):“Principal Component Analysis” Wiley Interdisciplinary Reviews: Computational Statistics, Vol.2, p.433-459

Abreu, D., and Brunnermeier, M. (2002):”Synchronization risk and delayed arbitrage” Journal of Financial Economics, Vol.66, p.457-488

Ackert, L.F., and Tian, Y.S. (2000):”Evidence on the efficiency of index options markets” Federal Reserve Bank of Atlanta Economic Review, Vol.85, p.40-52

Avellaneda, M., and Lee J.H. (2008):”Statistical Arbitrage in the U.S. Equities Market”, Quantitative Finance, Vol.10 (7), p.761-782

Barr, A. (2007):”Quant quake shakes hedge-fund giants; Goldman, Renaissance, AQL see losses, but also sense opportunity”, Marketwatch, August 13

Bhatt, S., and Cakici, N. (1990):”Premiums on stock index futures: Some evidence” Journal of Futures Markets, Vol.10, p.367-375

Bollinger, J. (2002):”Bollinger on Bollinger Bands” McGraw Hill, ISBN:978-0-07-137368-5

Buyuksahin, B., Haigh, M.S., and Robe, M.A. (2008):”Commodities and Equities: A Market of One?” Working Paper, January

Costa, J.M., Cruz L.V., and Leiria P.S. (2004):“Government bond markets: what is the magnitude of volatility-spillover effects in the euro area?” Instituto de Gestão do Crédito Público (IGPC), Working Paper, September

Chan, K., and Chung, P. (1993):”Intraday relationship among index arbitrage, spot and futures price volatility, and spot market volume: A transaction data test” Journal of Banking and Finance, Vol.17 (4), p.663-687

Cheung, Y., and Ng, L. (1990):”The Dynamics of S&P 500 index and the S&P 500 futures intraday price volatilities” Review of Future Markets, Vol.2, p.458-466

Chung, P. (1991):”A transactions data test of stock index futures market efficiency and index arbitrage profitability” Journal of Finance, Vol.46 (5), p.791-809

De Long, B., Shleifer, A., Summers, L., and Waldmann, R. (1990):”Noise trader risk in financial markets”, Journal of Political Economy, Vol.98, p.703-738

EDHEC-Risk Institute (2009):”Commodities; what really drives oil-prices” Editorial

EDHEC Risk and Asset Management Research Center (2008):”The Oil Market: Let the Data Speak For Itself” Editorial

Haigh, M.S., Harris, J.H., Overdahl J.A., and Robe, M.A. (2007):“Market Growth, Trader Participation and Derivative Pricing” Working Paper

Hammoudeh, S., and Aleisa, E. (2002):“Relationship between Spot/Futures Prices of Crude Oil and Equity Indices for the Oil-Producing Economies and US Oil-Related Industry” Arab Economic Journal, Vol.11 (27), p.37-62

Hargreaves, S. (2011):"Oil Speculators charged with price manipulation" CNNmoney.com, May 24.

Hasbrouk, J. (1995):"One security, many markets: Determining the contributions to price discovery" The Journal of Finance, Vol. 50 (4), p.1175-1199

Hasbrouk, J. (2003):"Intraday Price Formation in U.S. Equity Index Markets" Journal of Finance, Vol.58 (6), p.2375-2399

Hotelling, H. (1933):"Analysis of a complex of statistical variables into principal components" Journal of Educational Psychology, Vol.24, p.417-441

Huang, R.D., Masulis, R.W., and Stoll, H.R. (1996):"Energy Shocks and Financial Markets" Journal of Futures Markets, Vol.16 (1), p.1-27

Johansen, S. (1988):"Statistical Analysis of Cointegration Vectors" Journal of Economic Dynamics and Control, Vol.12, p.231–254

Johansen, S. (1991):"Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models" Econometrica, Vol.59 (6), p.1551-80

Jolliffe, I.T. (2002):"Principal Component Analysis" Springer-Verlag New York Inc, 2nd edition, ISBN: 0-387-95442-2

Kawaller, I., Koch, P., and Koch, T. (1988):"The Relationship between the S&P 500 index and S&P 500 index future prices" FRB Atlanta Economics Review, Vol.4, p.2-11

Kim, M., Szakmary, A., and Schwarz, T. (1999):"Trading costs and price discovery across stock index futures and cash markets" The CFA Digest, Vol.19 (4), p.475-498

Kurov, A., and Lasser, D. (2002):"The effect of the introduction of cubes on the NASDAQ-100 Index Spot-futures pricing relationship" Journal of Futures Markets, Vol.22 (3), p.197-218

Lee, Charles M.C., Shleifer, A., Thaler, R.H. (1991):"Investor Sentiment and the Closed-End Fund Puzzle" The Journal of Finance Vol.46 (1), p.75–109

Lee, C.F., Leuthold, R.M, and Cordier, J.E. (1985):"The Stock Market and the Commodity Futures Market: Diversification and Arbitrage Potential" Financial Analyst Journal, Vol.41 (4), p.53-60

Lehmann, B.N. (1990):"Fads, Martingales, and Market Efficiency" NBER Working Paper Series, Vol.2533

Lo, A.W., and MacKinley, A.C. (1990):"When are contrarian profits due to stock market overreaction?" The Review of Financial Studies, Vol.3 (2), p.175-205

Lo, A.W and Khandani, A.E. (2007): "What happened to the quants in August 2007?" Journal of Investment Management, Vol.5 (4), p.29–78

Mackinlay, C., and Ramaswamy, K. (1988):"Index-futures arbitrage and the behavior of stock index futures prices" The Review of Financial Studies, Vol.1 (2), p.137-158

Modest, D., and Sundaresan, K. (1983):"The relationship between spot and futures prices in stock index future markets: some preliminary evidence" Journal of Futures Markets, Vol.3 (1), p.15-41

Morgan Stanley Research North America (2010):"Exchange-traded Funds; Average Tracking Error Rose Meaningfully in 2009", February 12th

Neal, R., (1996):"Direct tests of index arbitrage models" Journal of Financial and Quantitative analysis, Vol.31, p.541-562

Papapetrou, E. (2001):"Oil price shocks, stock market, economic activity and employment in Greece" Energy Economics, Vol.23, p.511-532

Pearson, K. (1901):"On Lines and Planes of Closest Fit to Systems of Points in Space" Philosophical Magazine, Vol.2 (6), p.559-572

Pole A. (2007):"Statistical arbitrage: Algorithmic trading insights and techniques", Wiley Finance, ISBN: 978-0-470-13844-1

Poterba, J.M., and Summers, L.H. (1988):"Mean Reversion In Stock Prices: Evidence and Implications" Journal of Financial Economics, Vol.22, p.27-59

Rusli, E.M. (2007):"Goldman Sachs Alpha to fail?", Forbes.com, August 9th

Sadorsky, P. (1999):"Oil price shocks and stock market activity" Energy Economics, Vol.21 (2), p.449-469

Shleifer, A,. and Vishny, R. (1997):"The limits of Arbitrage" Journal of Finance, Vol.52, p.35-55

Smith, L. (2002):"A Tutorial on Principal Component Analysis"

Solnik, B., Boucrelle, C., and Fur, Y.L. (1996):"International Market Correlation and Volatility" Financial Analyst Journal, Vol.52 (5), p.17-34

Stoll, H., and Whiley, R. (1990):"The dynamics of stock index and index futures returns" Journal of Financial and Quantitative Analysis, Vol.45 (4), p.191-220

Switzer, L., Varson, P., Zghidi, S. (2000):"Standard and Poor's Depository Receipts and the performance of the S&P 500 Index futures Market" Journal of Futures Markets, Vol.20 (8), p.705-716

Yadav, P., and Pope, P. (1994):"Stock index futures mispricing: profit opportunities, or risk premia?" Journal of Banking and Finance, Vol.18 (5), p.921-953

Appendix List

Appendix A – The Different Securities

Appendix B – Return Series Securities

Appendix C – Descriptive Statistics Securities

Appendix D – Correlation Matrix

Appendix E – Dynamic Conditional Correlation (DCC)

Appendix F – Scatter Diagrams Pairs of Securities

Appendix G – Principal Component Analysis

Appendix H – Augmented Dickey Fuller test: Testing Order of Difference

Appendix I – Cointegration Test Results: Engle-Granger

Appendix J – Graph of Residuals from Engle-Granger Test

Appendix K – Cointegration Test: Selecting Model for the Johansen Test

Appendix L – Cointegration Test Results: Johansen

Appendix M – Granger Causality test Results

Appendix N – Return Equal Weight Modified Bollinger Band Strategies

Appendix O – Number of Trades and transaction cost

Appendix P – Moving Average Trading Strategies Equal Weighted

Appendix Q – Moving Average Trading Strategies Engle Granger Weighted

Appendix R – Transaction Cost Moving Averages Equal Weighted

Appendix S – Transaction Cost Moving Averages Engle Granger Weighted

Appendices

Appendix A – The Different Aluminum Securities

Alcoa is the third largest aluminum producer in the world. In 2010 their total revenues amounted to $21.0B. Alcoa was founded in 1888 in Pittsburg, Pennsylvania. Alcoa is traded on the New York Stock Exchange under the ticker AA. It is a component in both the Dow Jones Industrial Average and the S&P 500. They employ 59,000 people worldwide.

Website: www.Alcoa.com

Century Aluminum Company was founded in 1995 in Switzerland as a holding company. Its main business is to invest in aluminum related businesses. The company's total revenue in 2010 was $1,17B. Century Aluminum is traded at the New York Stock Exchange under the ticker code CENX.

Website: www.centuryaluminium.com

Kaiser Aluminum Corporation was founded in 1946 when Henry J. Kaiser bought three aluminum facilities from the United States government. Its revenue in 2010 was $1,08B. The company headquarters are located in Foothill Ranch, California.

Website: www.kaiseraluminium.com

Rio Tinto became the world's largest producer of aluminum when they acquired Alcan in 2007. Rio Tinto generated $60.3B in revenue in 2010. $15.2B of these came from Rio Tinto Alcan, making aluminum Rio Tinto's second largest business after its iron ore operations. The company was founded in 1873 and is traded at the stock exchanges in Sydney and London. It was delisted from the New York Stock Exchange in 2010. Rio Tinto Alcan has its headquarters in Montreal, Canada.

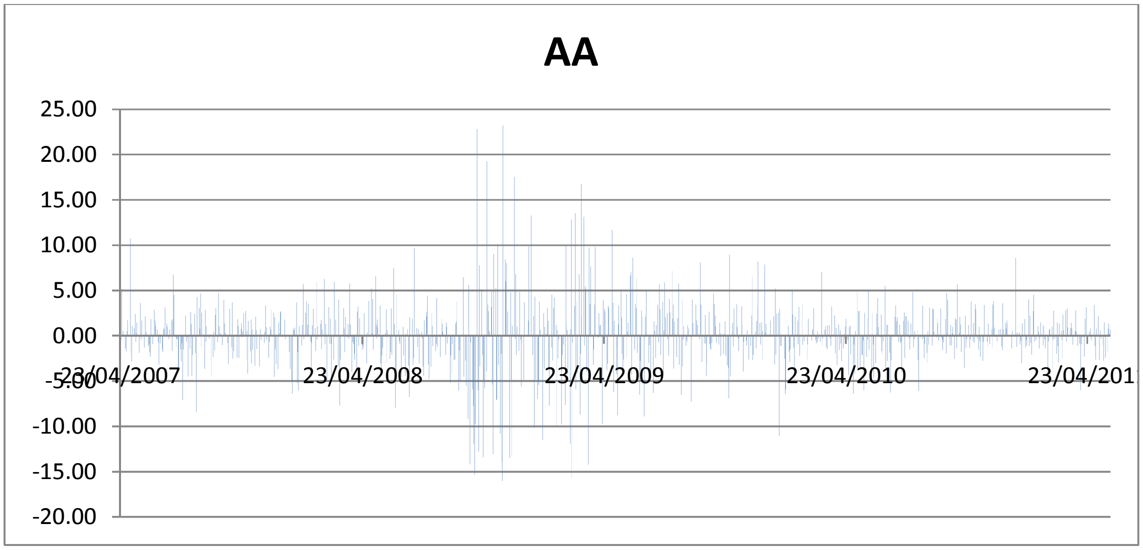
Website: www. Riotinto.com

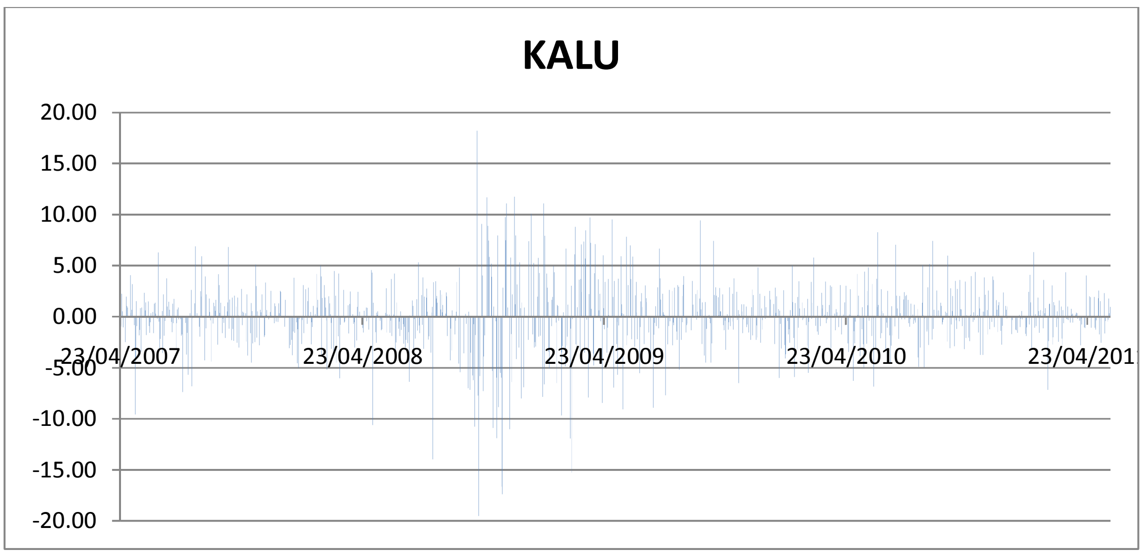
ETF Securities ETF is an ETF managed by ETF Securities. It is designed to replicate the daily price moves of aluminum forward contracts. The ETF is not leveraged.

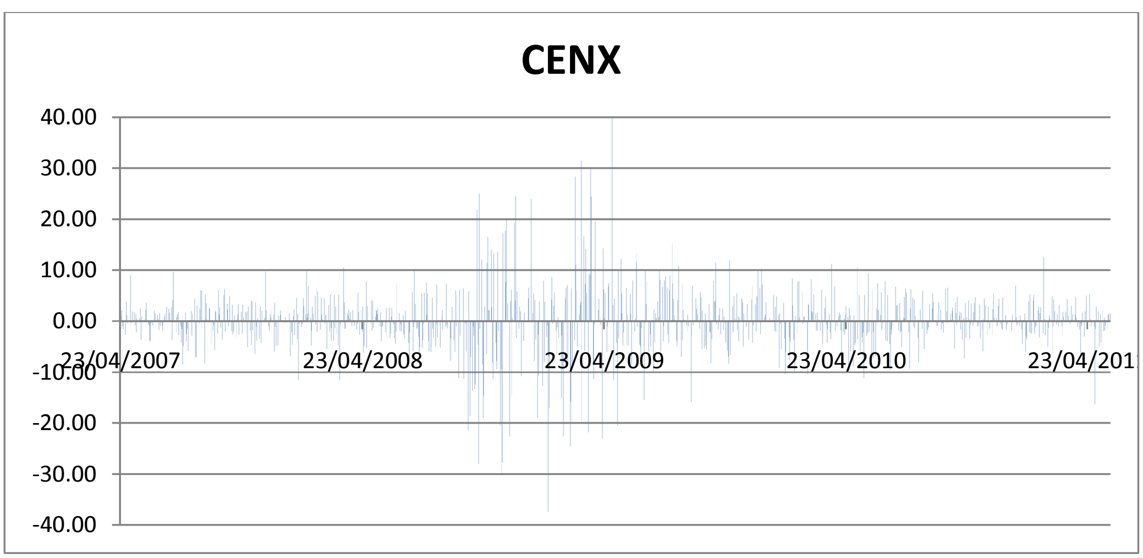
Goldman Sachs Commodity Index Aluminum Index is an index created by Goldman Sachs as a synthetic investment opportunity in aluminum.

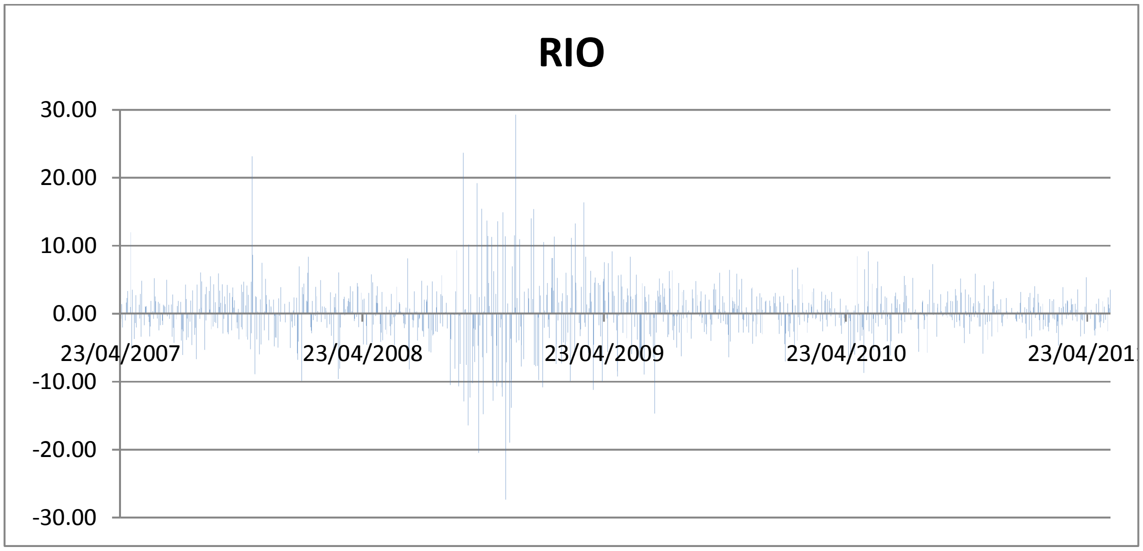
Three Month Aluminum Futures is a futures contract for delivery of physical aluminum in three months.

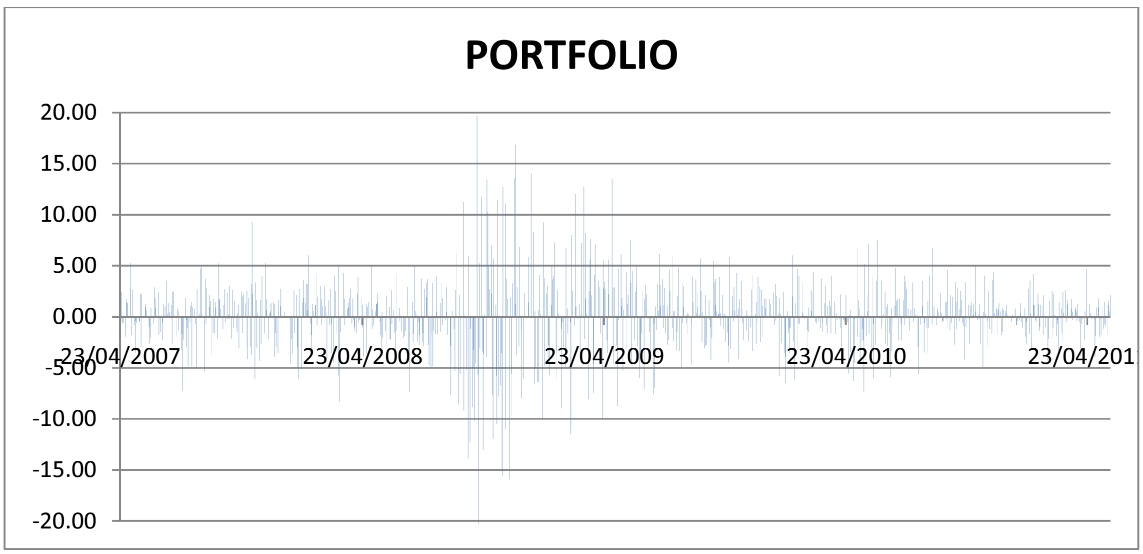
Appendix B - Return Series (20.04.2007 – 27.05.2011)

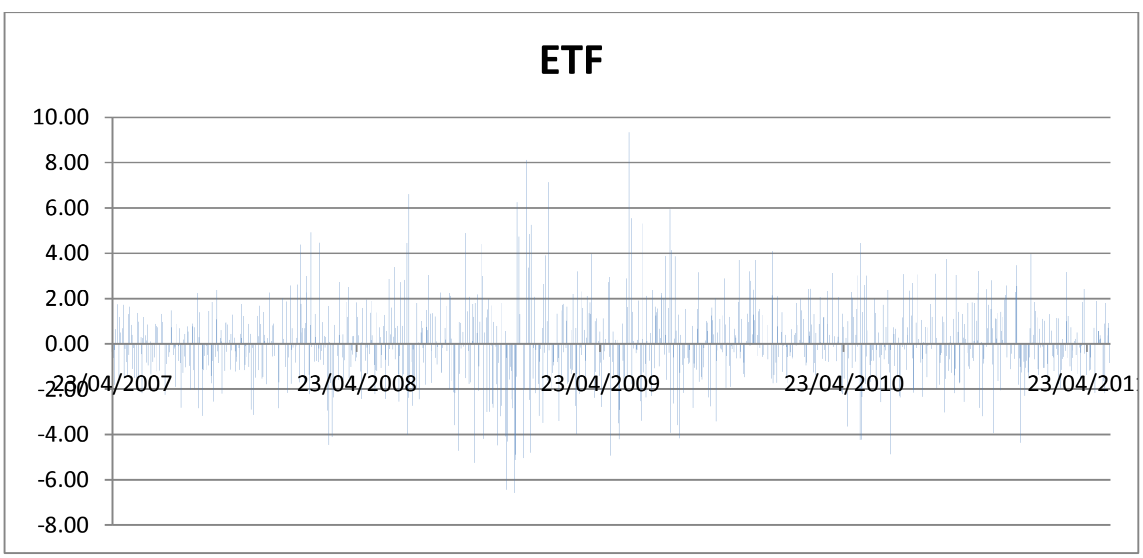


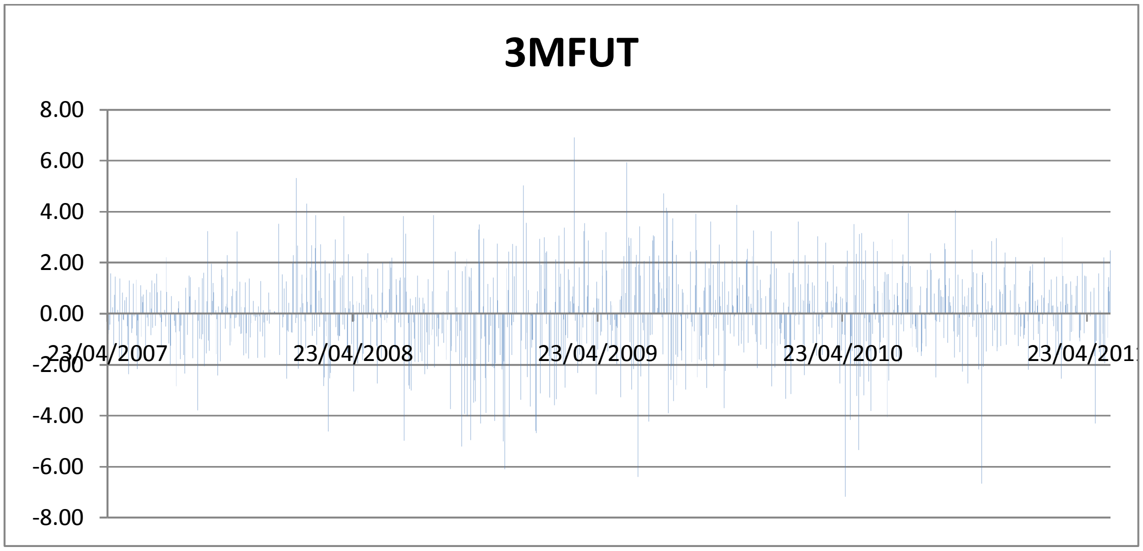


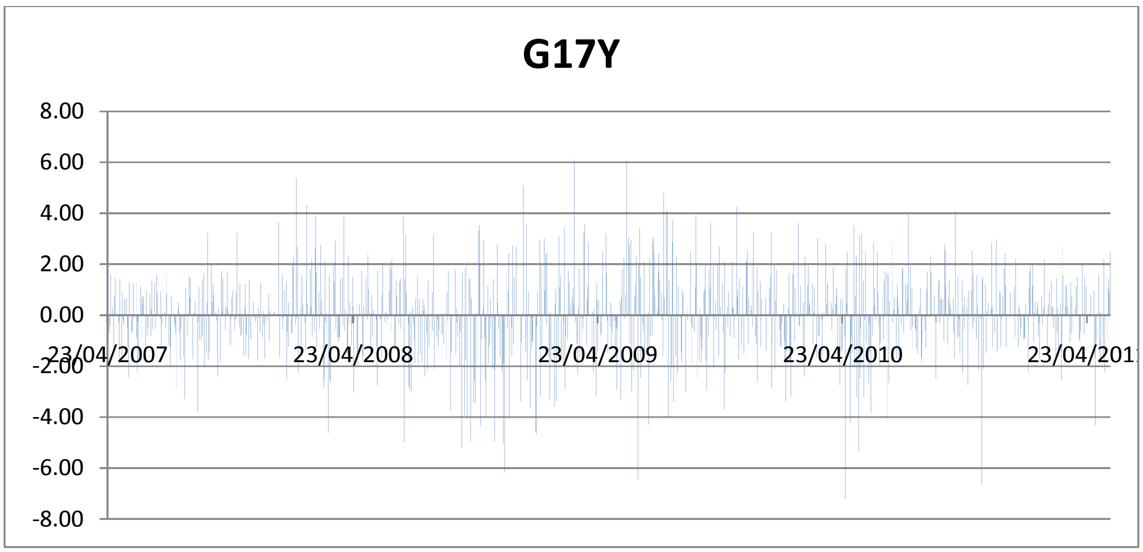


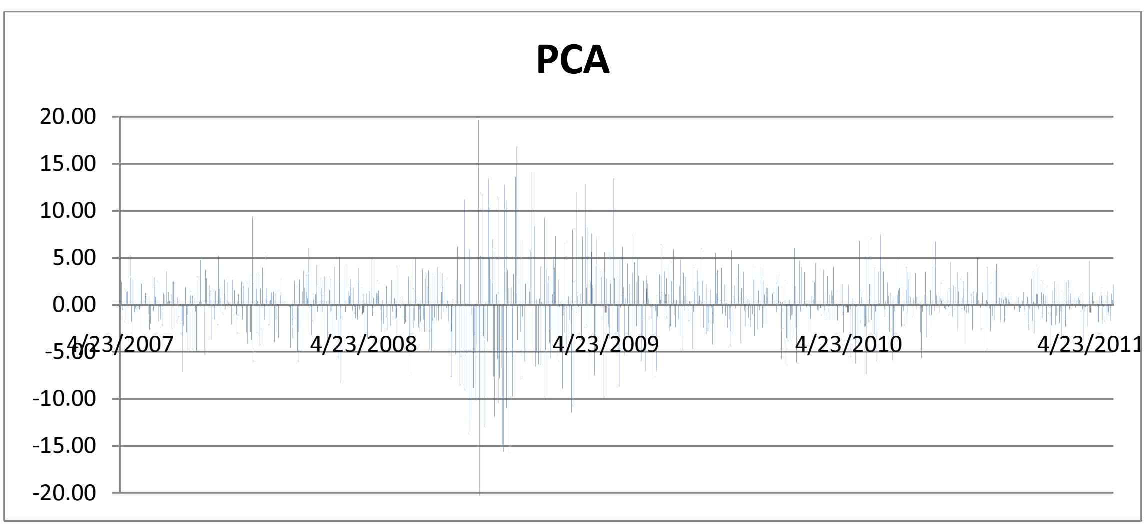

RIO
30.00
20.00
10.00
0.00
-10.00
-20.00
-30.00
23/04/2007
23/04/2008
23/04/2009
23/04/2010


PORTFOLIO
20.00
15.00
10.00
5.00
0.00
-5.00
-10.00
-15.00
-20.00
23/04/2007
23/04/2008
23/04/2009
23/04/2010


ETF
10.00
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
23/04/2007
23/04/2008
23/04/2009
23/04/2010


3MFUT
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
23/04/2007
23/04/2008
23/04/2009
23/04/2010
23/04/2011


G17Y
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
23/04/2007
23/04/2008
23/04/2009
23/04/2010
23/04/2011


PCA
20.00
15.00
10.00
5.00
0.00
-5.00
-10.00
-15.00
-20.00
4/23/2007
4/23/2008
4/23/2009
4/23/2010
4/23/2011

Appendix C – Descriptive statistics for daily returns (in percent) of aluminum securities (20.04.2007-27.05.2011)

	Mean	Standard deviation	Skewness	Kurtosis	Minimum	Maximum	Median	Jarque-Bera test statistics	ADF test statistics	Autocorr. lag 1	Autocorr. lag 2	Ljung Box test statistics	Obser-vations
Alcoa	0.001	3.879	0.274	8.532	-16.05	23.21	0.000	1 280.99**	-31.38**	-0.004	0.066	0.02	995
Century Aluminium	0.114	6.854	2.032	36.044	-37.42	90.10	0.103	45 952.86**	-29.39**	0.069	0.014	4.77*	995
Kaiser Aluminium	0.012	3.367	-0.417	7.705	-19.52	18.22	0.043	946.41**	-30.02**	0.048	-0.031	2.29	995
Rio Tinto	0.110	4.409	0.215	9.662	-27.33	29.29	0.099	1847.47**	-33.45**	-0.060	0.002	3.63	995
EW Portfolio	-0.004	3.746	-0.467	7.718	-22.32	18.34	0.134	959.05**	-32.25**	-0.024	-0.013	0.56	995
PCA Portfolio	0.029	3.648	-0.133	7.509	-20.39	19.67	0.137	845.82**	-32.25**	-0.024	-0.013	0.56	995
GSCI Alu. Index	-0.017	1.752	-0.205	3.927	-7.22	6.11	0.010	42.55**	-32.28**	-0.025	0.013	0.65	995
3 Month Alu. Futures	0.007	1.743	-0.184	3.999	-7.18	6.91	0.027	46.95**	-32.29**	-0.026	0.009	0.66	995
ETFS Alu. ETF	-0.032	1.774	0.335	5.204	-6.59	9.34	-0.800	219.98**	-30.07**	0.047	-0.020	2.20	995

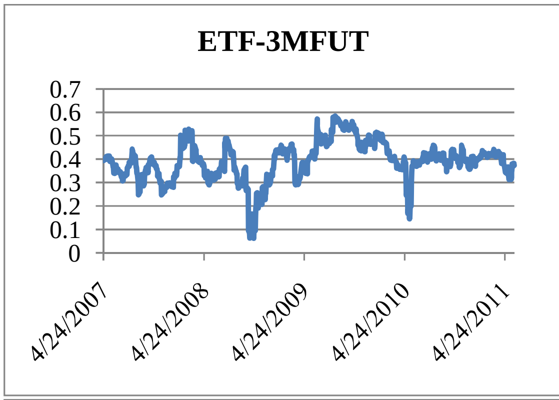
Note: * and ** indicate significance at the 5% and 1% levels, respectively.

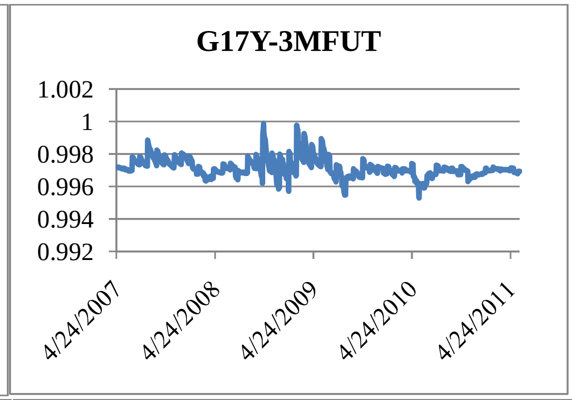
ADF and Ljung-Box test statistics use a one day lag.

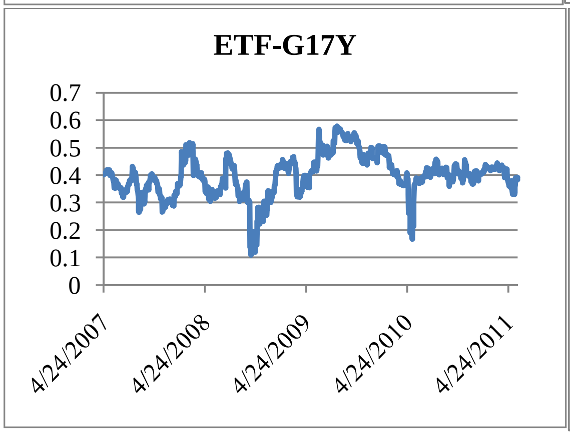
Appendix D – Correlations between aluminum securities (20.04.2007-27.05.2011)

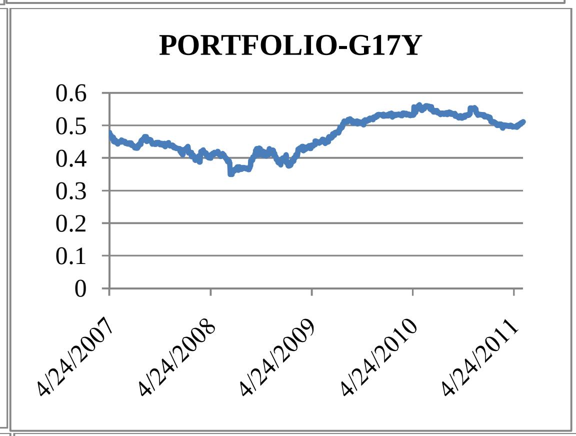
	Alcoa	Century Alu	Kaiser Alu	Rio Tinto	EW Portfolio	PCA Portfolio	ETF	G17Y	3MFUT
Alcoa	1								
Century Alu	0.735	1							
Kaiser Alu	0.722	0.636	1						
Rio Tinto	0.667	0.616	0.594	1					
EW Portfolio	0.873	0.812	0.805	0.900	1				
PCA Portfolio	0.848	0.793	0.824	0.913	0.996	1			
ETF	0.074	0.067	0.036	0.014	0.047	0.041	1		
G17Y	0.424	0.419	0.330	0.411	0.464	0.452	0.423	1	
3MFUT	0.421	0.419	0.324	0.411	0.461	0.450	0.415	0.997	1

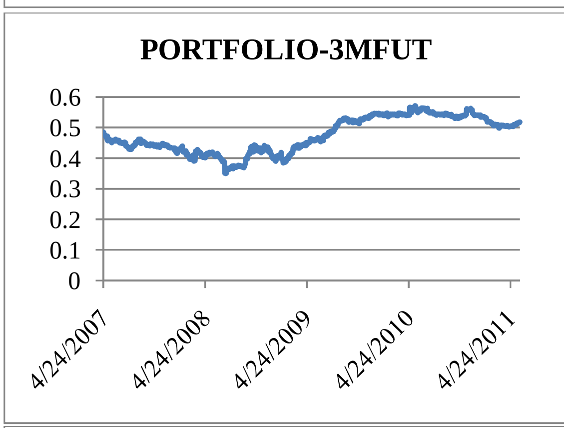
Appendix E – Dynamic Conditional Correlation (DCC)

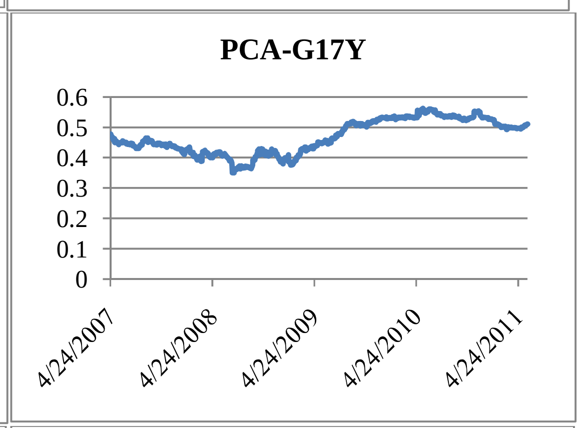


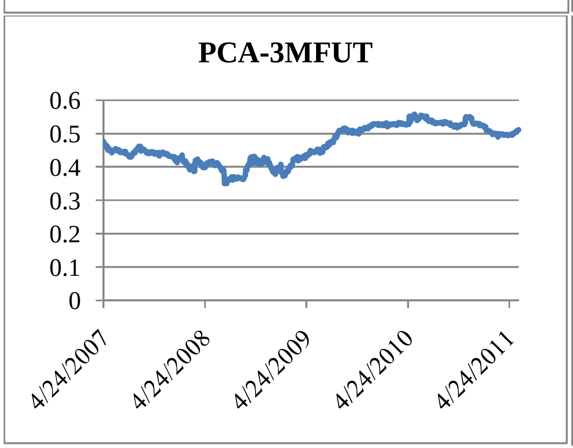


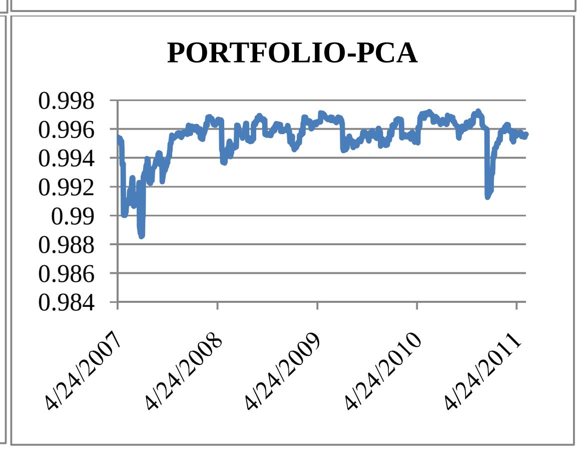




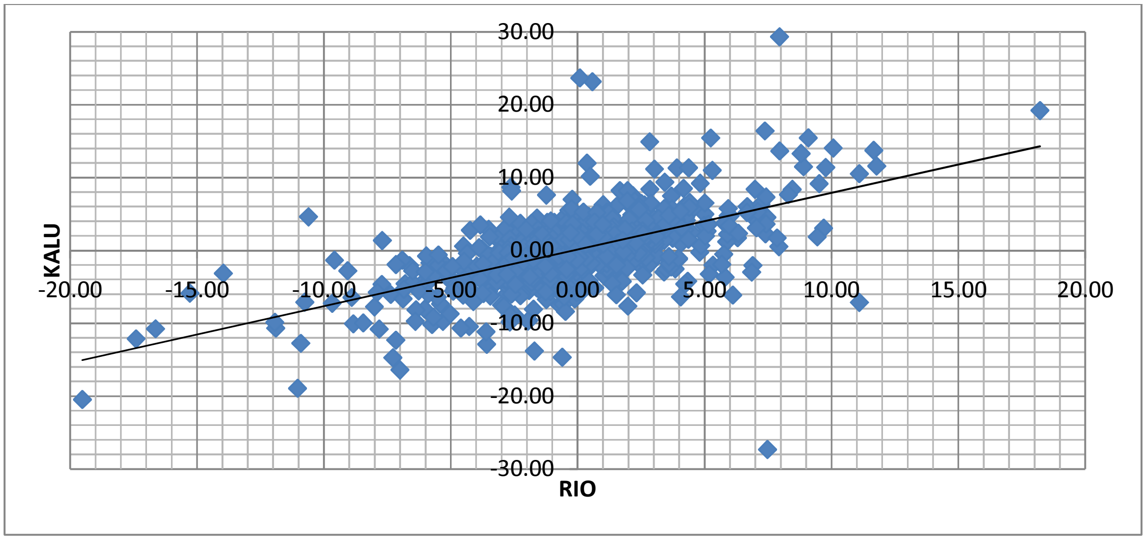


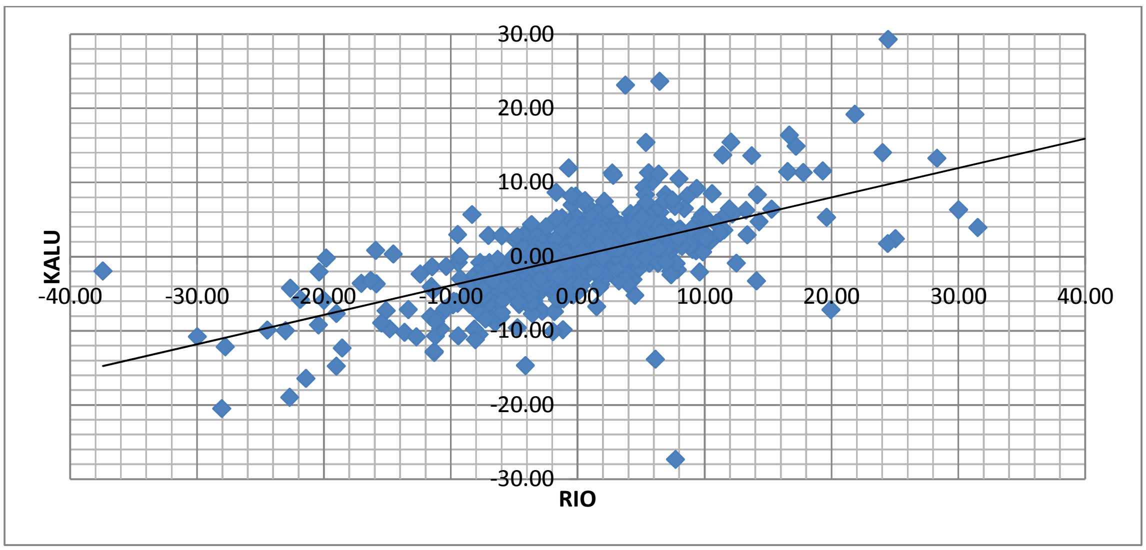


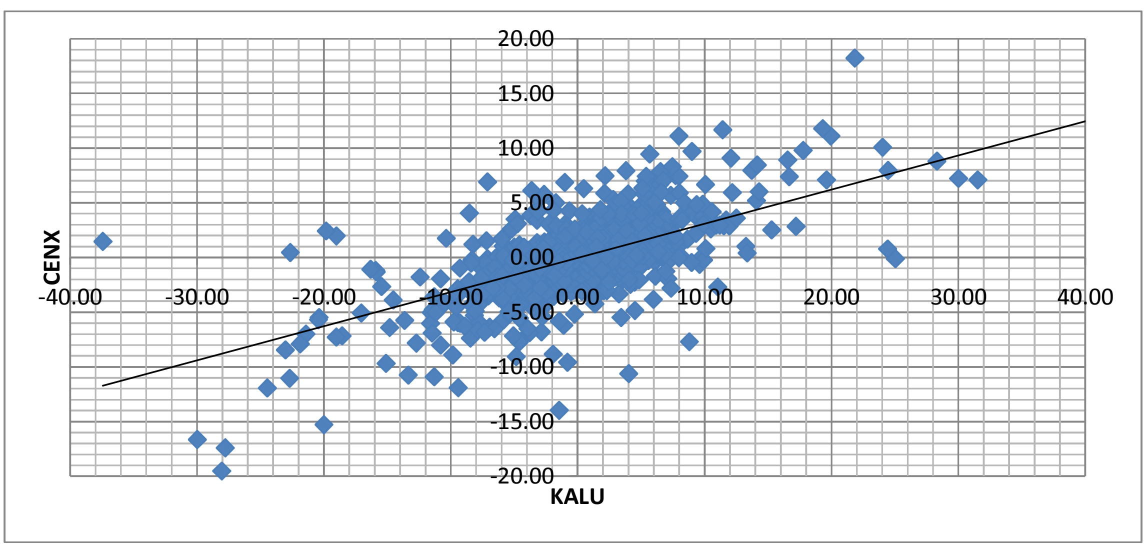


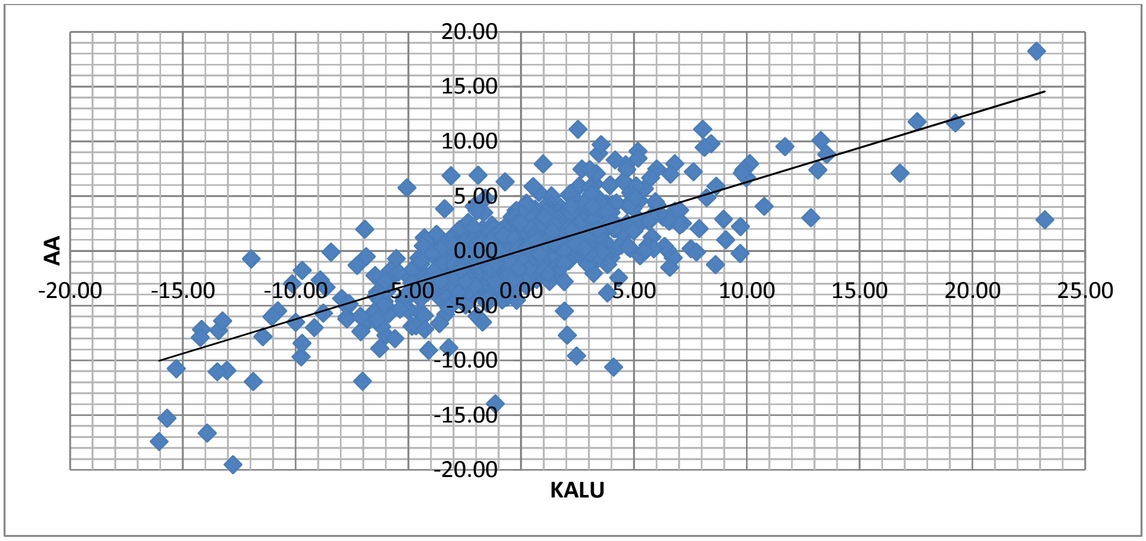


Appendix F - Scatter Diagram Aluminum Securities (20.04.2001-27.05.2011)

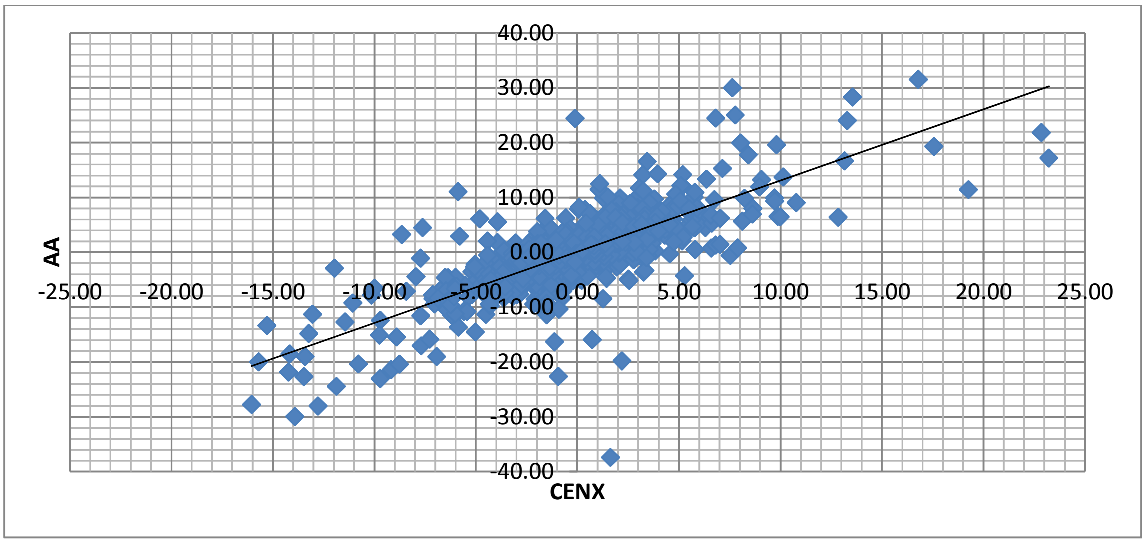




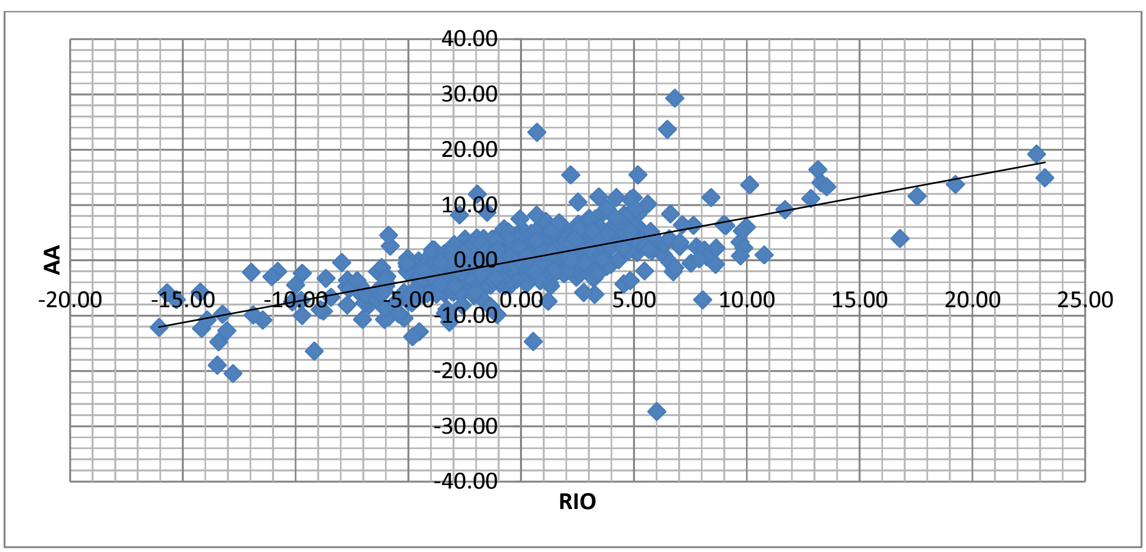




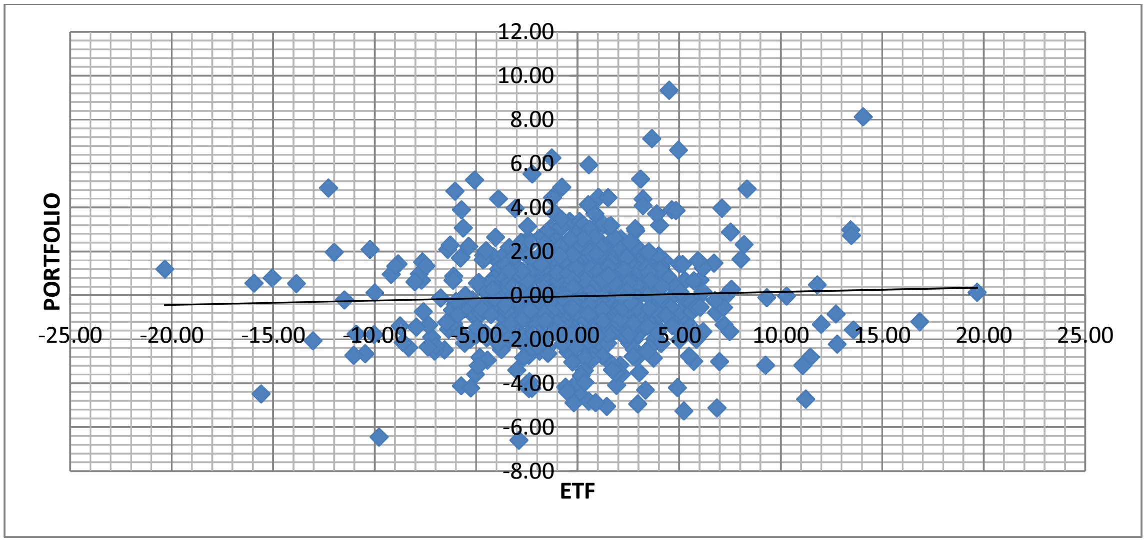
AA
KALU
20.00
15.00
10.00
5.00
0.00
-5.00
-10.00
-15.00
-20.00
-20.00
-15.00
-10.00
-5.00
0.00
5.00
10.00
15.00
20.00
25.00



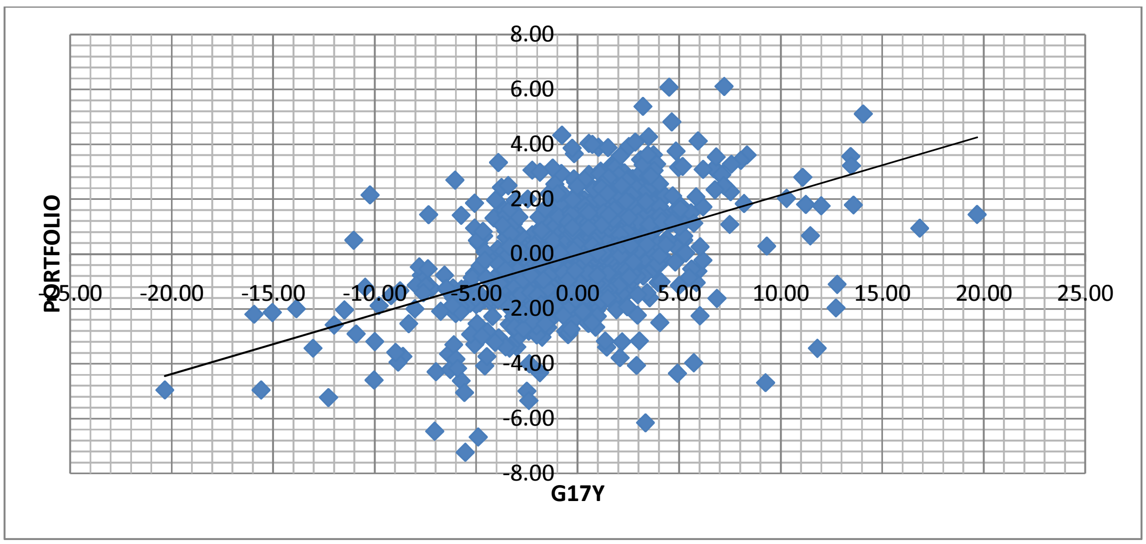
AA
CENX
40.00
30.00
20.00
10.00
0.00
-10.00
-20.00
-30.00
-40.00
-25.00
-20.00
-15.00
-10.00
-5.00
0.00
5.00
10.00
15.00
20.00
25.00



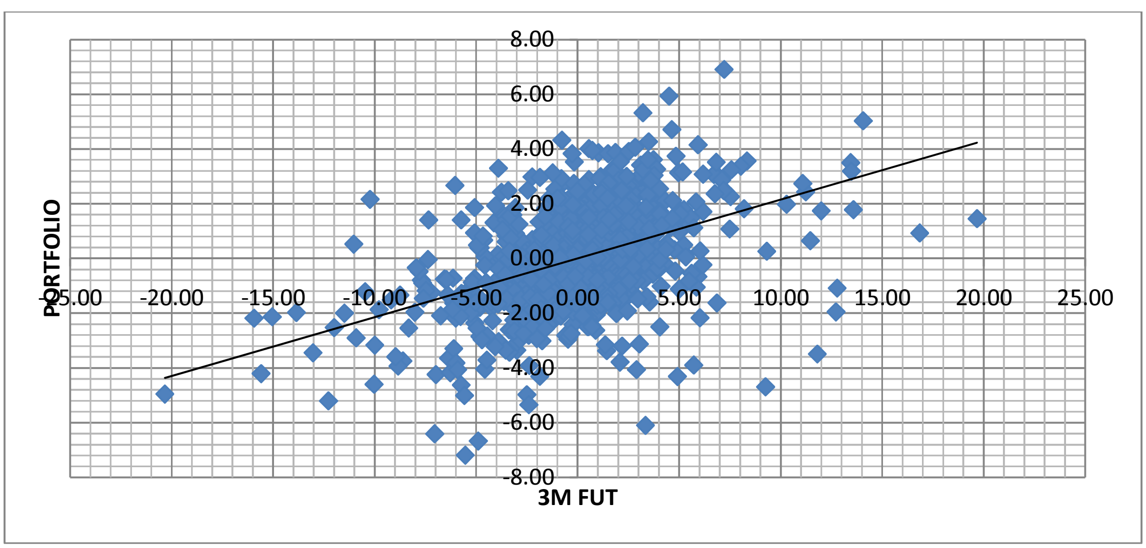
AA
RIO
40.00
30.00
20.00
10.00
0.00
-10.00
-20.00
-30.00
-40.00
-20.00
-15.00
-10.00
-5.00
0.00
5.00
10.00
15.00
20.00
25.00



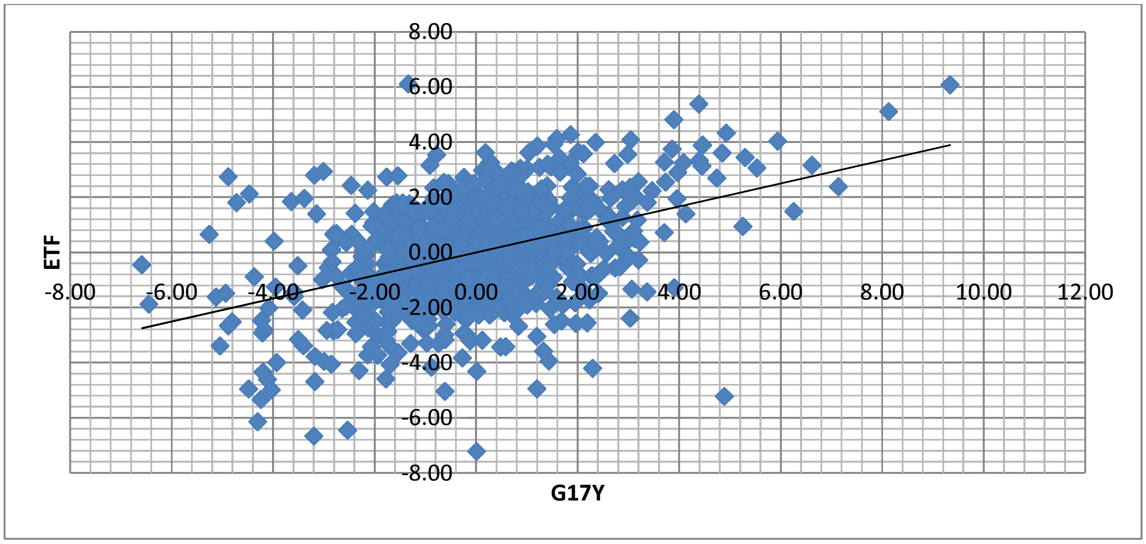
12.00
10.00
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
-25.00
-20.00
-15.00
-10.00
-5.00
0.00
5.00
10.00
15.00
20.00
25.00
PORTFOLIO
ETF

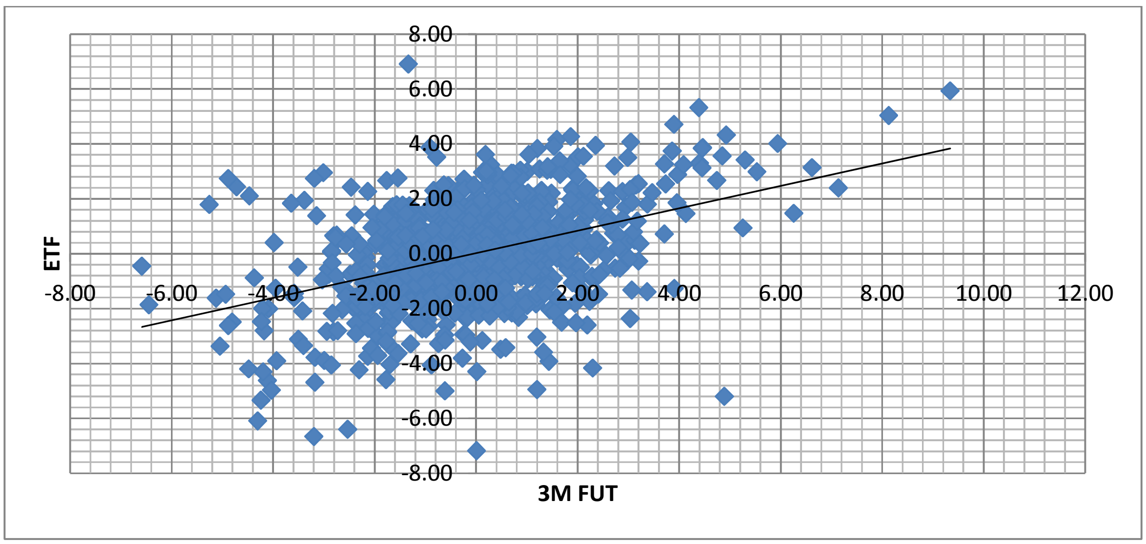


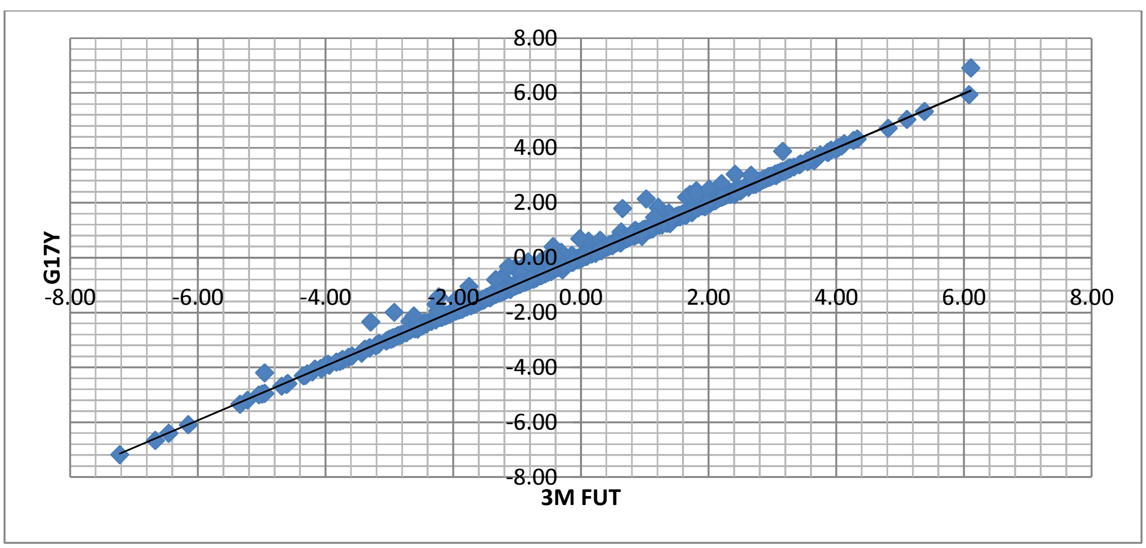
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
-25.00
-20.00
-15.00
-10.00
-5.00
0.00
5.00
10.00
15.00
20.00
25.00
PORTFOLIO
G17Y

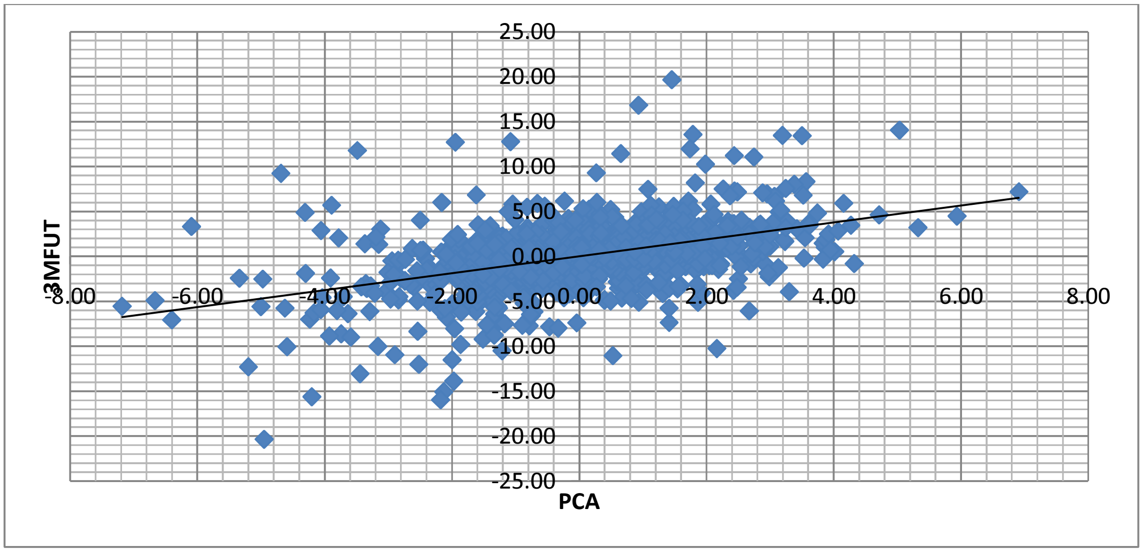


8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
-25.00
-20.00
-15.00
-10.00
-5.00
0.00
5.00
10.00
15.00
20.00
25.00
PORTFOLIO
3M FUT

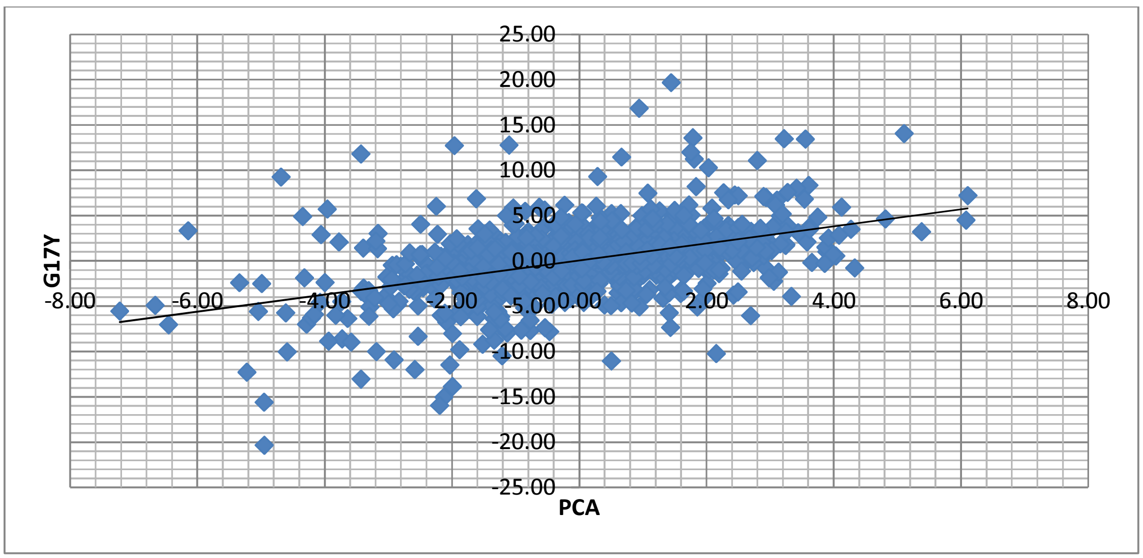

8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
10.00
12.00
ETF
G17Y


8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
10.00
12.00
ETF
3M FUT

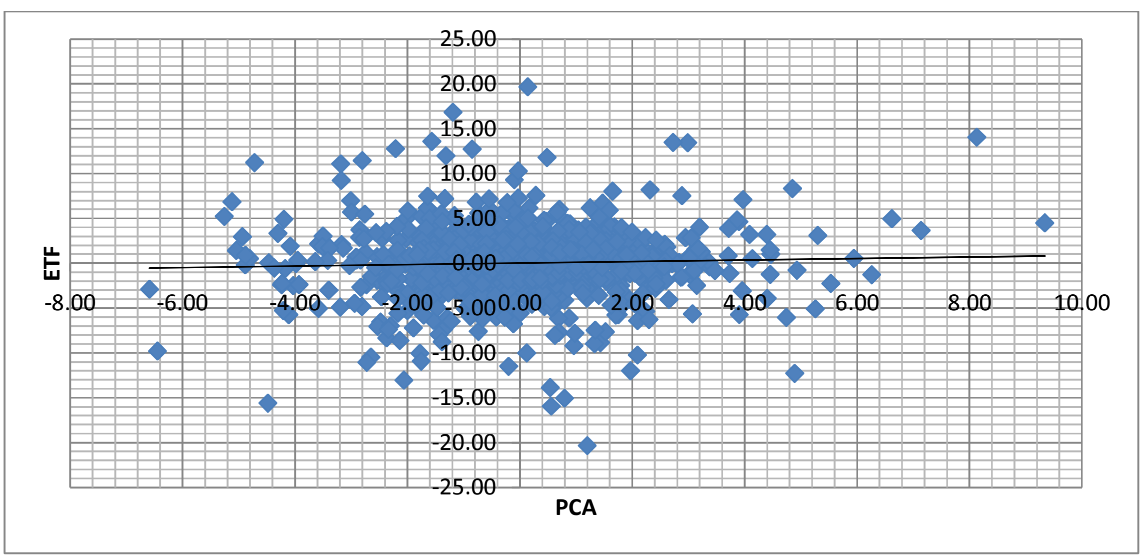

8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-6.00
-8.00
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
G17Y
3M FUT



25.00
20.00
15.00
10.00
5.00
0.00
-5.00
-10.00
-15.00
-20.00
-25.00
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
i3MFUT
PCA



25.00
20.00
15.00
10.00
5.00
0.00
-5.00
-10.00
-15.00
-20.00
-25.00
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
G17Y
PCA



25.00
20.00
15.00
10.00
5.00
0.00
-5.00
-10.00
-15.00
-20.00
-25.00
-8.00
-6.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
10.00
ETF
PCA

Appendix G – Principal Component Analysis (PCA) – Eigenvectors and Eigenvalues

Eigenvalues: (Sum = 4, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	2.9873	2.572173	0.7468	2.987302	0.7468
2	0.41513	0.053367	0.1038	3.402432	0.8506
3	0.36176	0.125958	0.0904	3.764195	0.941
4	0.23581	---	0.059	4	1

Eigenvectors (loadings):

Variable	PC 1	PC 2	PC 3	PC 4
AA	0.525311	-0.153747	-0.073278	-0.833691
CENX	0.500649	-0.161525	-0.744702	0.410704
KALU	0.493923	-0.490006	0.629308	0.346274
RIO	0.478998	0.842713	0.209811	0.127966

PC – Principal Component

Appendix H - Augmented Dickey Fuller test for Unit Root

Panel A: Test on log(index) for security	**ADF t-statistic**	**P-value**
3MFUT	-1.079	0.93
PORTFOLIO	-1.151	0.92
PCA	-1.618	0.47
G17Y	-1.008	0.94
ETF	-1.636	0.78
AA	-1.126	0.92
CENX	-1.180	0.91
KALU	-1.501	0.83
RIO	-1.591	0.80
Panel B: Test on log($index_t$-$index_{t-1}$) for security	**ADF t-statistic**	**P-value**
3MFUT	-32.339	0.00
PORTFOLIO	-32.096	0.00
PCA	-33.368	0.00
G17Y	-32.335	0.00
ETF	-30.067	0.00
AA	-31.167	0.00
CENX	-29.259	0.00
KALU	-29.931	0.00
RIO	-33.239	0.00
Test critical values:	1% level***	-3.967308
	5% level**	-3.414341
	10% level*	-3.129294

Appendix I_1 – Engle Granger Cointegration test (20.04.2007 – 27.05.2011)

Panel A: Test on residuals (exog. Constant)	**Lags**	**ADF t-statistics**	**P-value**
AA - CENX	0	-3.491*	0.008
AA - KALU	0	-2.277	0.180
AA - RIO	0	-2.563***	0.100
RIO - CENX	0	-2.761***	0.064
KALU - CENX	0	-2.643***	0.085
RIO - KALU	0	-3.738*	0.004
Test critical values:		1% level*	-3.436703
-The Lag Length is automatic and based on AIC, maxlag=21		5% level**	-2.864233
		10% level***	-2.568256

Panel B: Test on residuals (exog. Constant, linear trend)		**ADF t-statistic**	**P-value**
AA - CENX	0	-3.787**	0.018
AA - KALU	0	-3.581**	0.032
AA - RIO	0	-3.348***	0.059
RIO - CENX	0	-2.924	0.155
KALU - CENX	0	-2.850	0.180
RIO - KALU	0	-3.645**	0.027
Test critical values:		1% level*	-3.967298
-The Lag Length is automatic and based on AIC, maxlag=21		5% level**	-3.414336
		10% level***	-3.129291

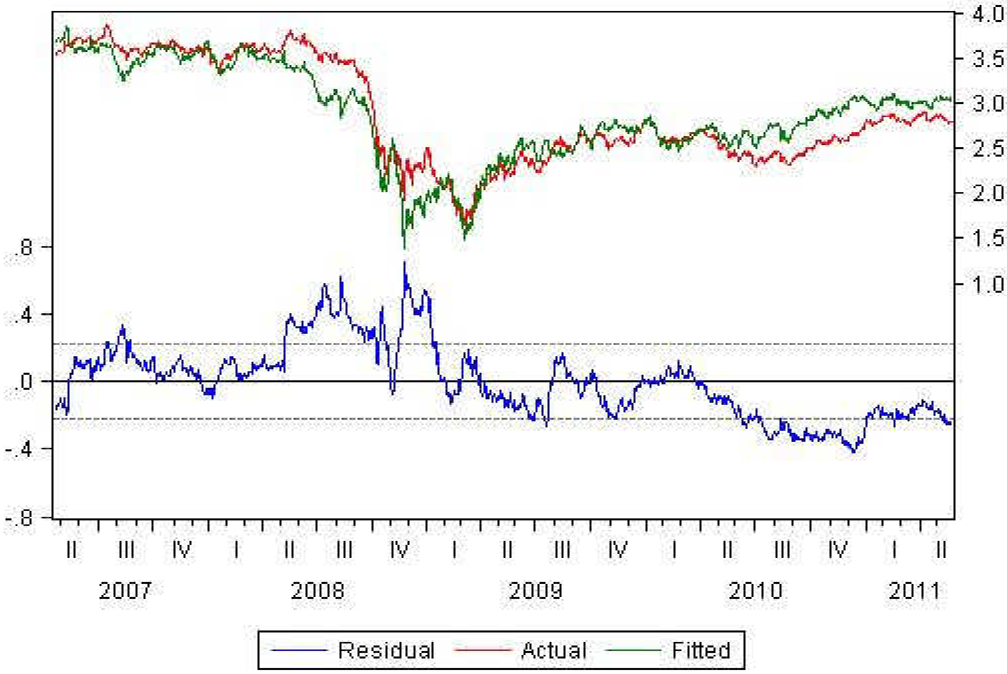
Appendix I_2 – Engle Granger Cointegration test (20.04.2007 – 27.05.2011)

Panel A: Test on residuals (exog. Constant)	**Lags**	**ADF t-statistics**	**P-value**
3MFUT - PORTFOLIO	0	-2.825***	0.055
3MFUT - G17Y	0	0.204	0.973
3MFUT - ETF	1	-1.720	0.421
PORTFOLIO - G17Y	0	-3.969*	0.002
PORTFOLIO - ETF	0	-2.645***	0.083
G17Y - ETF	1	-2.102	0.244
PCA - 3MFUT	0	-3.645*	0.005
PCA - ETF	0	-2.649***	0.083
PCA - G17Y	0	-3.293**	0.016
Test critical values:		1% level*	-3.436703
-The Lag Length is automatic and based on AIC, maxlag=21		5% level**	-2.864233
		10% level***	-2.568256

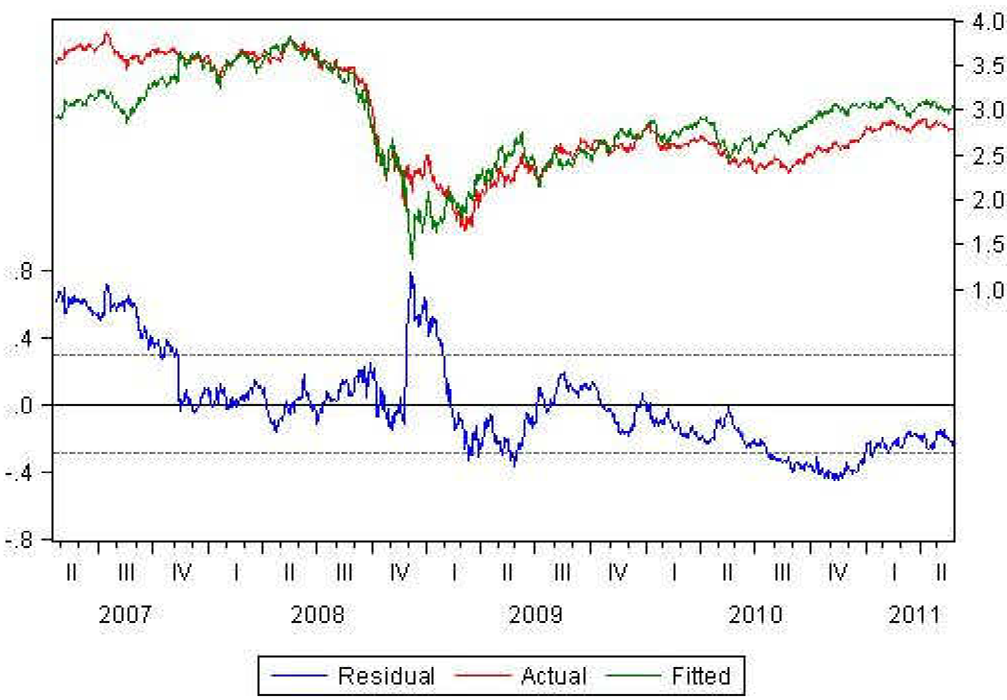
Panel B: Test on residuals (exog. Constant, linear trend)	**Lags**	**ADF t-statistics**	**P-value**
3MFUT - PORTFOLIO	0	-3.614**	0.029
3MFUT - G17Y	0	-1.539	0.816
3MFUT - ETF	1	-1.864	0.673
PORTFOLIO - G17Y	0	-3.972*	0.010
PORTFOLIO - ETF	0	-2.72918	0.225
G17Y - ETF	1	-2.14412	0.520
PCA - 3MFUT	0	-3.696**	0.023
PCA - ETF	0	-2.729	0.225
PCA - G17Y	0	-3.416**	0.050
Test critical values:		1% level*	-3.967298
-The Lag Length is automatic and based on AIC, maxlag=21		5% level**	-3.414336
		10% level***	-3.129291

Appendix J- Residuals from Engle-Granger approach for testing cointegration between securities

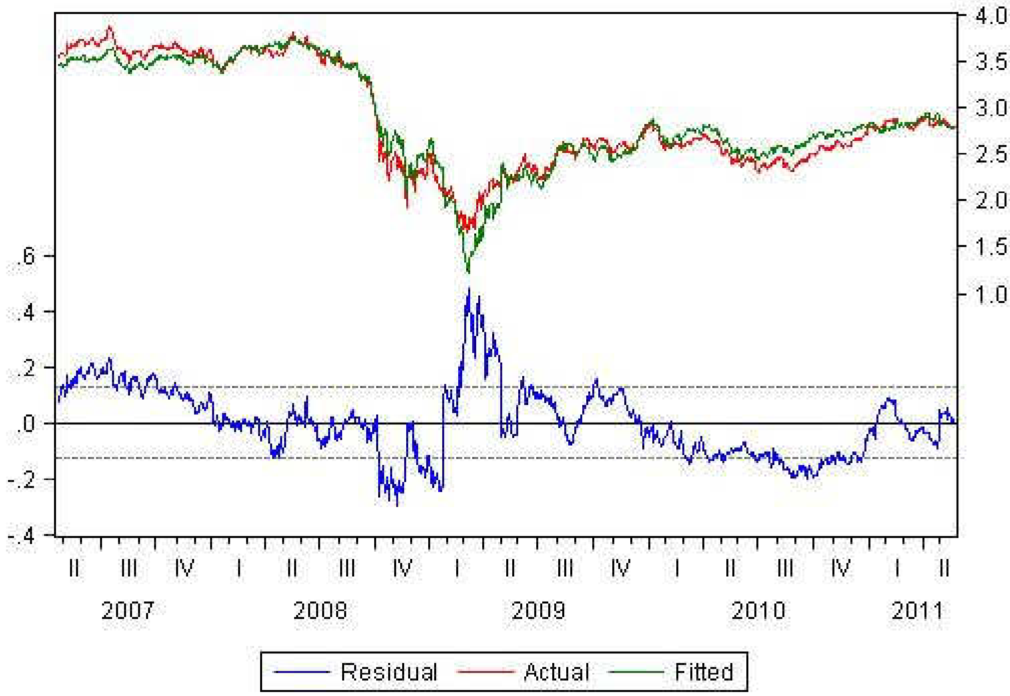
$LOG(AA) = C + LOG(KALU) + u$



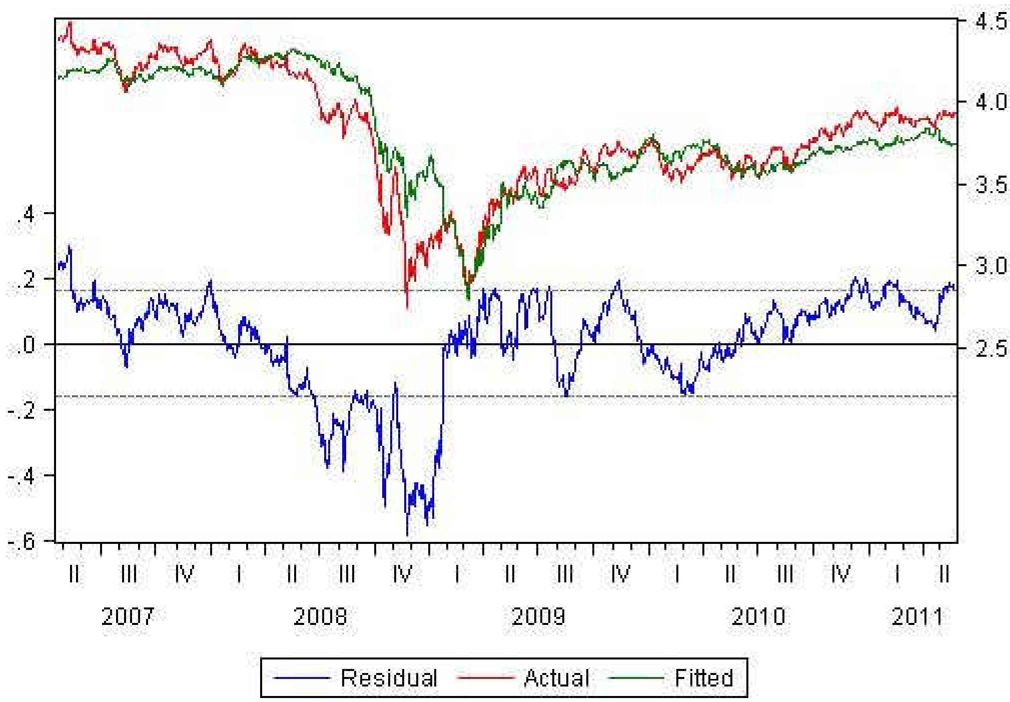
$LOG(AA) = C + LOG(RIO) + u$



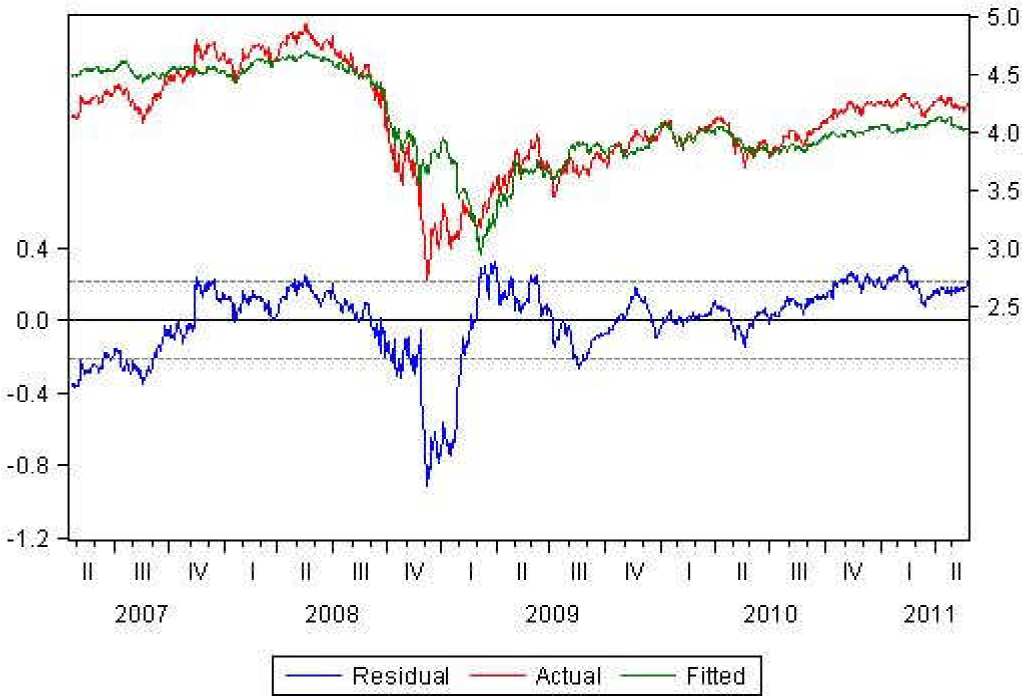
LOG(AA) = C + LOG(CENX) + u



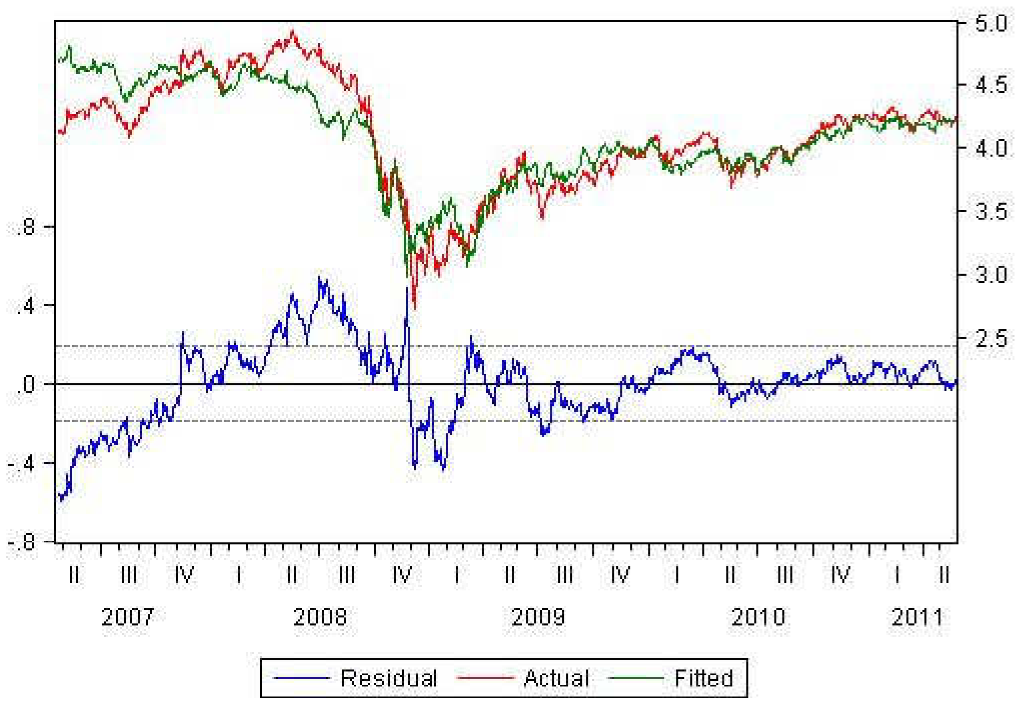
LOG(KALU) = C + LOG(CENX) + u



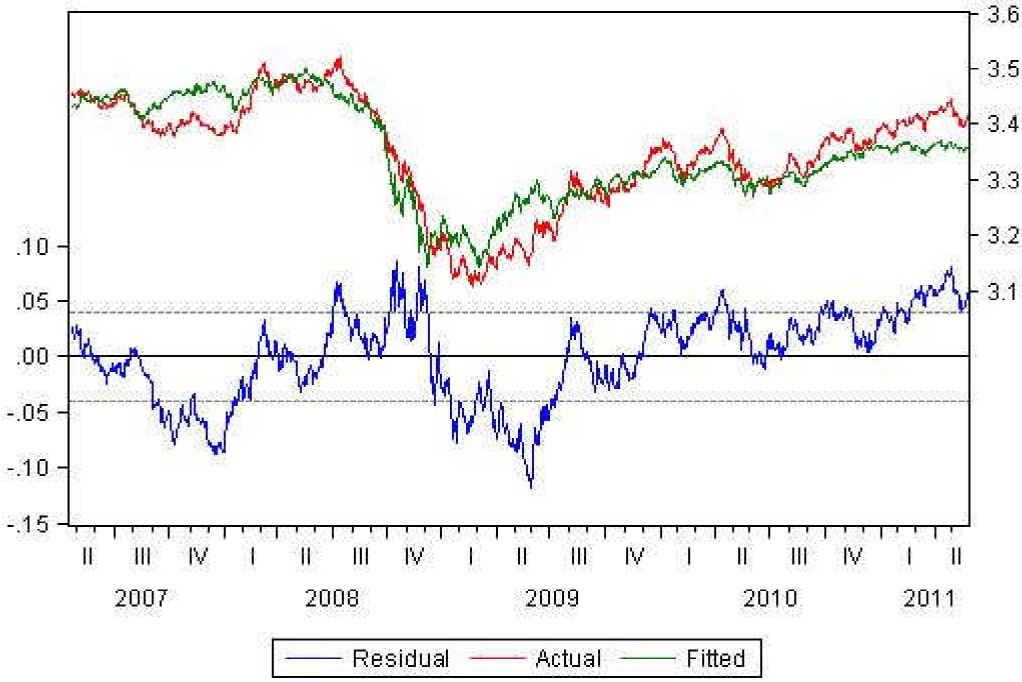
LOG(RIO) = C + LOG(CENX) + u



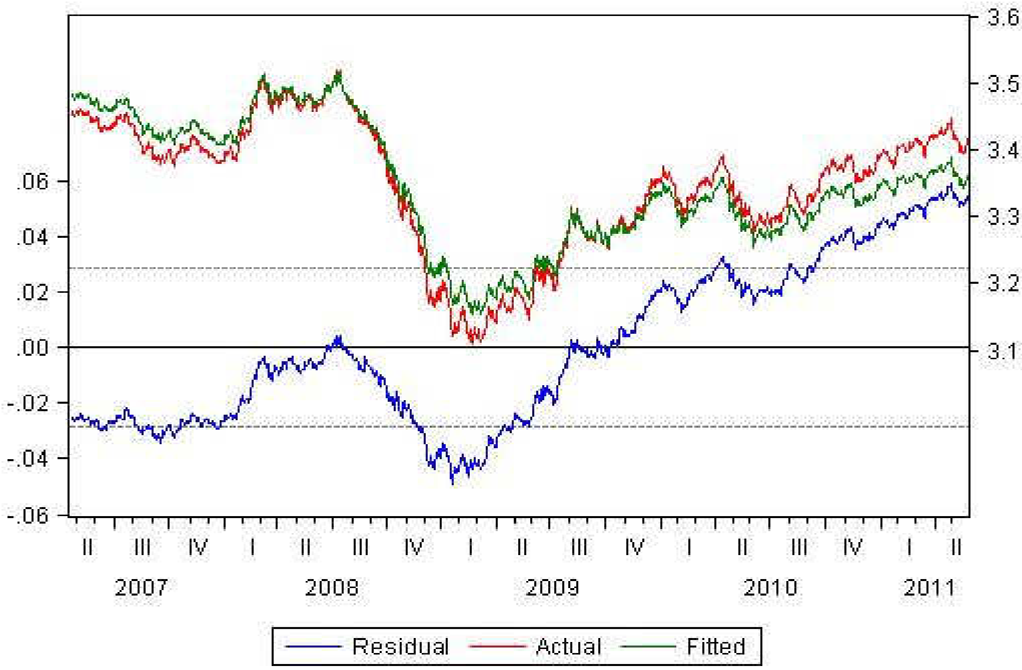
LOG(RIO) = C + LOG(KALU) + u



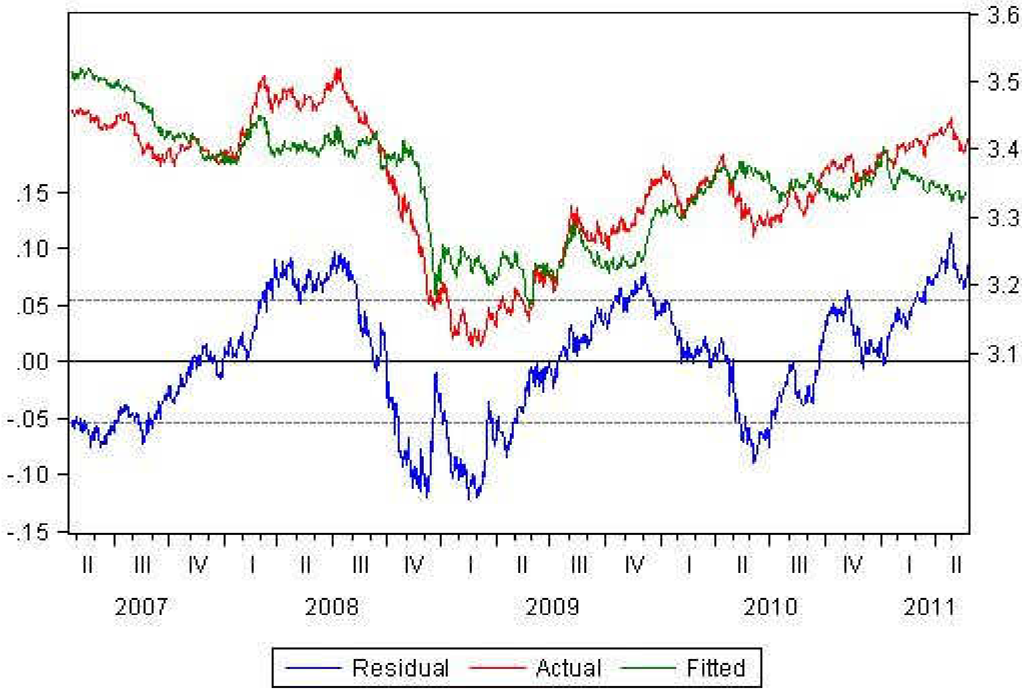
LOG(3MFUT) = C + LOG(PORTFOLIO) + u



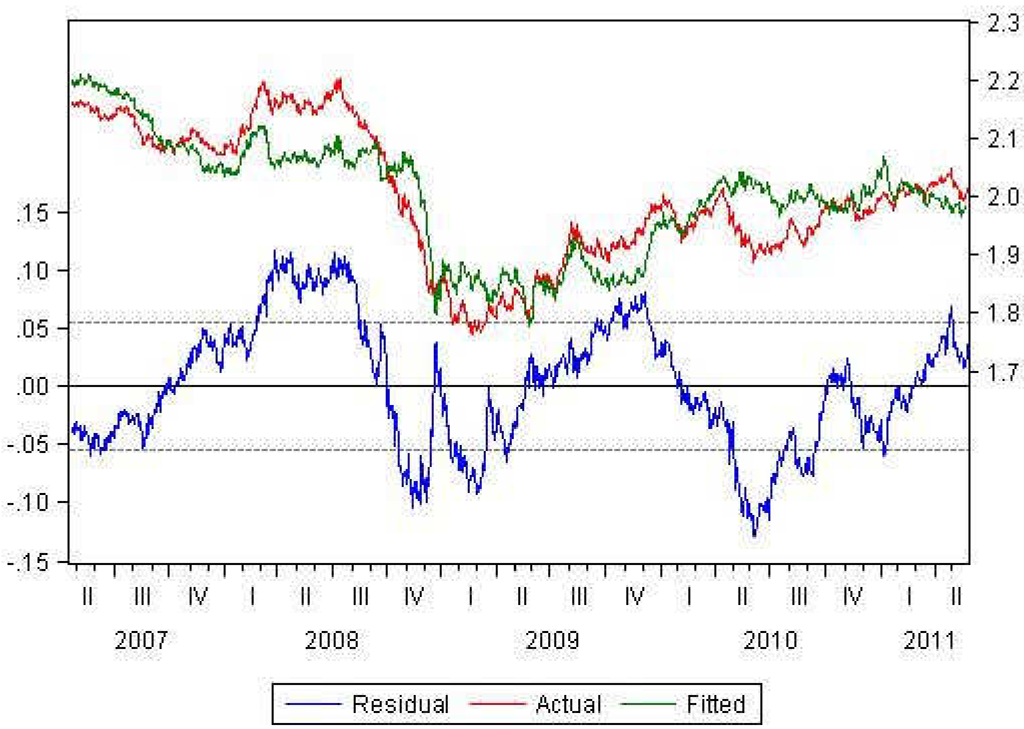
LOG(3MFUT) = C + LOG(G17Y) + u



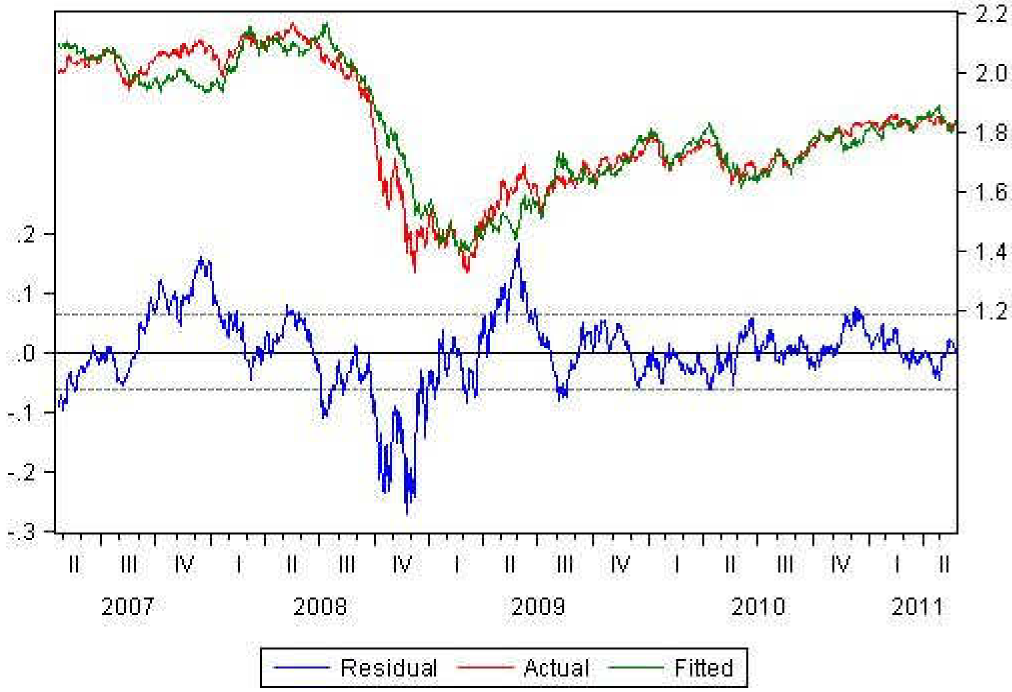
LOG(3MFUT) = C + LOG(ETF) + u



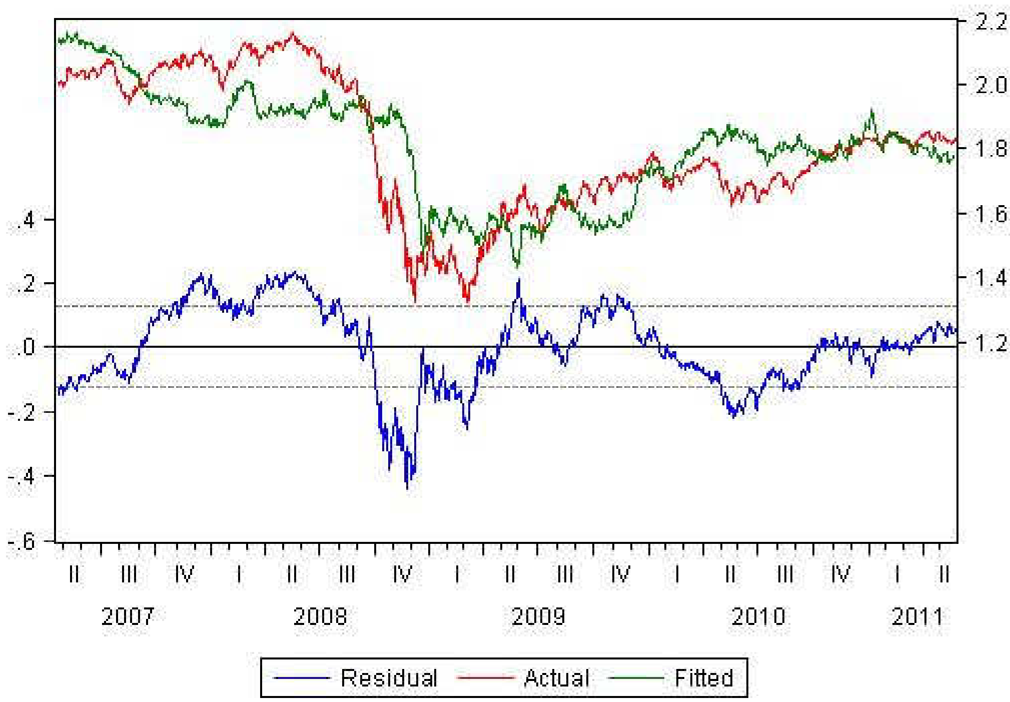
LOG(G17Y) = C + LOG(ETF) + u



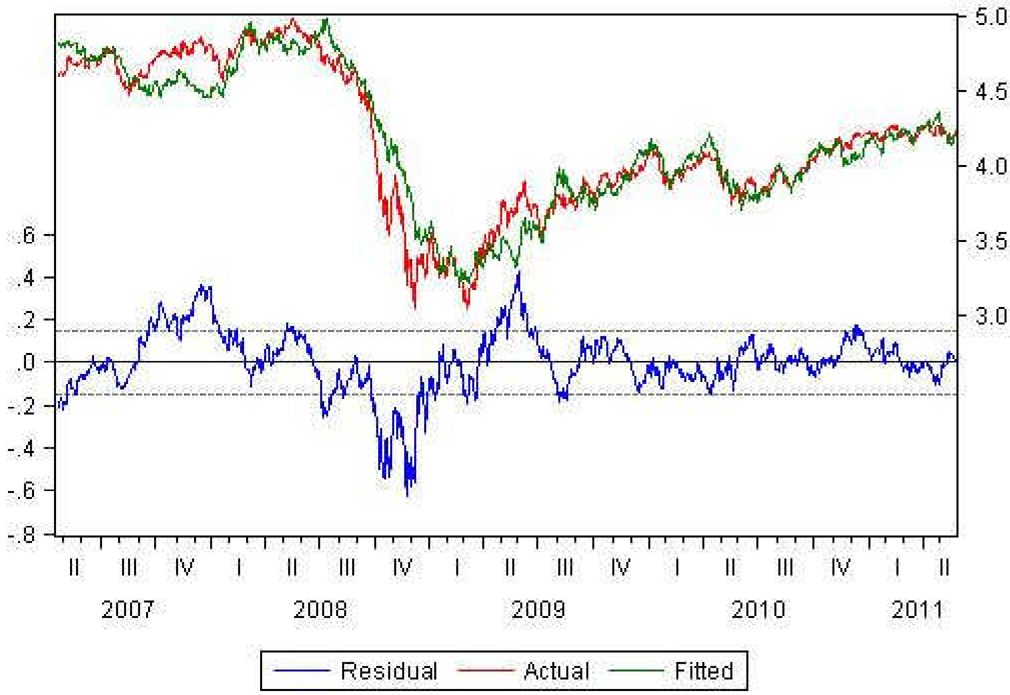
LOG(G17Y) = C + LOG(PORTFOLIO) + u



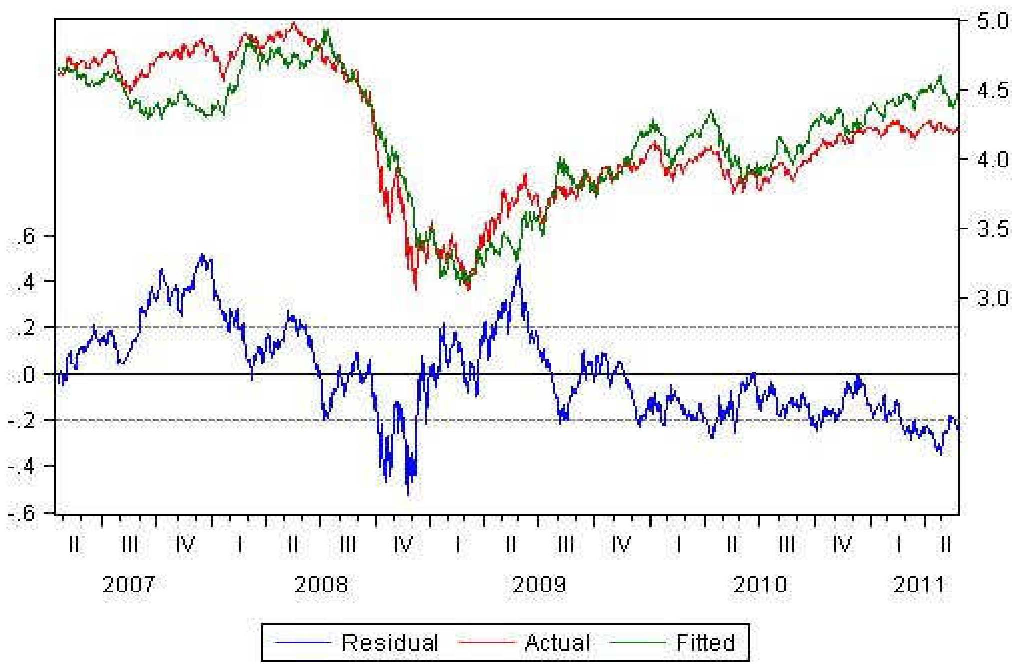
LOG(PORTFOLIO) = C + LOG(ETF) + u



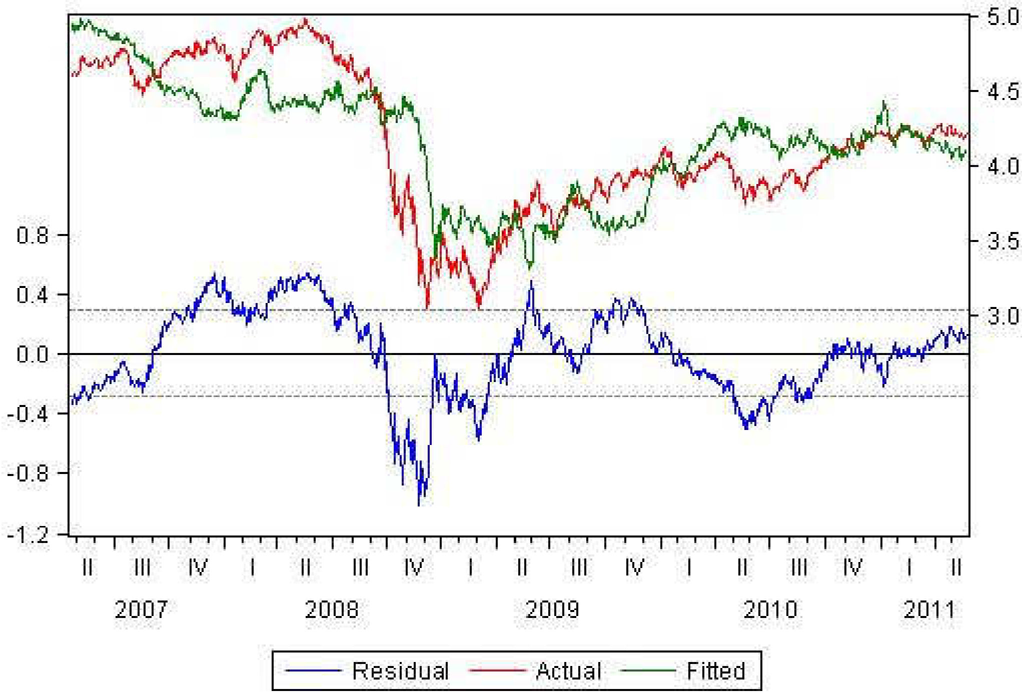
LOG(PCA) = C + LOG(G17Y) + u



LOG(PCA) = C + LOG(3MFUT) + u



LOG(PCA) = C + LOG(ETF) + u



Appendix K - Johansen Cointegration test - deterministic components

VAR-Model	AIC-Model 1	AIC-Model 2	AIC-Model 3	AIC-Model 4	AIC-Model 5	Rank	Lags
AA CENX	-10.39921	-10.40831	-10.41659*	-10.41037	-10.40896	1	2
AA RIO	-10.98246	-10.98246*	-10.97917	-10.97917	-10.97593	0	2
AA KALU	-11.66228	-11.66228*	-11.6586	-11.6586	-11.65581	0	2
CENX KALU	-10.37778	-10.37778*	-10.374	-10.374	-10.37123	0	2
CENX RIO	-9.824767	-9.824767*	-9.821202	-9.821202	-9.817594	0	2
RIO KALU	-11.10908	-11.10853*	-11.10661	-11.10566	-11.10469	1	2
AA CENX KALU RIO	-22.55847	-22.55656*	-22.55127	-22.55407	-22.54947	1	3
ETF G17Y	-14.07347	-14.07347*	-14.07017	-14.07017	-14.06823	0	2
ETF 3MFUT	-14.07628	-14.07628*	-14.07314	-14.07314	-14.07121	0	2
3MFUT G17Y	-18.99416	-18.99545	-18.99449	-19.00404*	-19.00355	0	2
PORTFOLIO ETF	-12.36349	-12.36378*	-12.3618	-12.3607	-12.35904	0	2
PORTFOLIO 3MFUT	-12.59674	-12.60275	-12.60085	-12.61885*	-12.61721	1	2
PORTFOLIO G17Y	-12.59401	-12.61694*	-12.61497	-12.61379	-12.61211	1	2
PCA 3MFUT	-12.30597	-12.32093*	-12.31892	-12.31963	-12.31777	1	2
PCA G17Y	-12.29536	-12.31372*	-12.31171	-12.31203	-12.31015	1	2
PCA ETF	-9.027364	-9.027652*	-9.025666	-9.024571	-9.022913	0	2
ETF 3MFUT G17Y	-26.16156	-26.16146	-26.16729	-26.17464*	-26.17151	0	2
3MFUT G17Y ETF PORTFOLIO	-31.88122	-31.87927	-31.89854*	-31.89697	-31.89152	1	2

*** Refers to the selected model based on AIC**

Appendix L - Johansen Cointegration Equation with Normalized Parameter Estimation

VAR	AA	CENX	KALU	RIO	PORTFOLIO	PCA	3MFUT	G17Y	ETF	Constant	Trend
AA CENX KALU RIO	1	-0.001	-3.562*	1.703*						1.592**	
		(0.005)	(5.181)	(-3.239)						(-2.065)	
PORTFOLIO 3MFUT					1		-1.875*				0.000*
							(15.478)				(-6.195)
PORTFOLIO G17Y					1			-1.878*		1.942*	
								(18.450)		(-9.526)	
PCA 3MFUT						1	-1.702*			-0.135*	
							(11.109)			(6.168)	
PCA G17Y						1		-1.465*		-0.196*	
								(10.118)		(6.985)	
PORTFOLIO 3MFUT G17Y ETF					1		0.387	-2.066*	-0.167		
							(-0.931)	(4.970)	(0.646)		

*Rejection of the null hyptothesis at 1% significance level

**Rejection of the null hyptothesis at 5% significance level

***Rejection of the null hyptothesis at 10% significance level

- Standard error in paranteses

Appendix M - Results Granger Causality Test

PORTFOLIO - 3MFUT		
Lags: 1	**F-Statistic**	**Prob.**
L3MFUT does not Granger Cause LPORTFOLIO	0.384	0.536
LPORTFOLIO does not Granger Cause L3MFUT	7.933	0.005
Lags: 2		
L3MFUT does not Granger Cause LPORTFOLIO	0.629	0.533
LPORTFOLIO does not Granger Cause L3MFUT	8.729	0.000
Lags: 3		
L3MFUT does not Granger Cause LPORTFOLIO	1.048	0.370
LPORTFOLIO does not Granger Cause L3MFUT	6.179	0.000

Null Hypothesis: Security 1 does not Granger Cause Security 2.

PCA - 3MFUT		
Lags: 1	**F-Statistic**	**Prob.**
L3MFUT does not Granger Cause LPCA	0.319	0.572
LPCA does not Granger Cause L3MFUT	23.881	0.000
Lags: 2		
L3MFUT does not Granger Cause LPCA	0.753	0.471
LPCA does not Granger Cause L3MFUT	13.859	0.000
Lags: 3		
L3MFUT does not Granger Cause LPCA	0.980	0.402
LPCA does not Granger Cause L3MFUT	9.297	0.000

Null Hypothesis: Security 1 does not Granger Cause Security 2.

PORTFOLIO - G17Y		
Lags: 1	F-Statistic	Prob.
LG17Y does not Granger Cause LPORTFOLIO	1.252	0.263
LPORTFOLIO does not Granger Cause LG17Y	29.356	0.000
Lags: 2		
LG17Y does not Granger Cause LPORTFOLIO	1.164	0.313
LPORTFOLIO does not Granger Cause LG17Y	18.945	0.000
Lags: 3		
LG17Y does not Granger Cause LPORTFOLIO	1.360	0.254
LPORTFOLIO does not Granger Cause LG17Y	13.324	0.000

Null Hypothesis: Security 1 does not Granger Cause Security 2.

PCA - G17Y		
Lags: 1	F-Statistic	Prob.
LG17Y does not Granger Cause LPCA	0.289	0.591
LPCA does not Granger Cause LG17Y	23.488	0.000
Lags: 2		
LG17Y does not Granger Cause LPCA	0.754	0.471
LPCA does not Granger Cause LG17Y	14.389	0.000
Lags: 3		
LG17Y does not Granger Cause LPCA	0.890	0.446
LPCA does not Granger Cause LG17Y	9.580	0.000

Null Hypothesis: Security 1 does not Granger Cause Security 2.

Appendix N – Return Equal Weight Modified Bollinger Band Strategies

Modified Bollinger Band Strategy #1

PCA-3MFUT	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
100-200	-22.87%	-22.87%	-22.87%	-22.87%	-20.50%	-18.44%	-21.74%
200-300	1.80%	2.91%	2.91%	2.91%	2.91%	2.91%	2.72%
300-400	-40.51%	-36.56%	-24.69%	-13.28%	0.00%	0.00%	-19.17%
400-500	33.04%	33.04%	33.04%	33.04%	0.00%	0.00%	22.03%
500-600	6.07%	22.78%	9.50%	9.50%	0.00%	0.00%	7.97%
600-700	-0.08%	-0.08%	0.00%	0.00%	0.00%	0.00%	-0.03%
700-800	2.06%	2.06%	0.00%	0.00%	0.00%	0.00%	0.69%
800-900	13.04%	5.05%	0.00%	0.00%	0.00%	0.00%	3.01%
900-997	2.17%	0.00%	0.00%	0.00%	0.00%	0.00%	0.36%
Total Return	-24.01%	-11.88%	-12.92%	0.27%	-18.19%	-16.06%	-13.80%
SD of returns	19.64%	19.73%	16.07%	14.45%	6.62%	5.98%	13.75%
SharpeRatio	-1.222	-0.602	-0.804	0.019	-2.748	-2.688	-1.341

PCA–G17Y	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
100-200	-23.63%	-23.63%	-23.63%	-23.63%	-23.63%	-18.93%	-22.85%
200-300	1.24%	1.24%	1.24%	1.24%	1.24%	1.24%	1.24%
300-400	-36.32%	-35.22%	-23.86%	-7.52%	0.00%	0.00%	-17.15%
400-500	33.66%	33.66%	33.66%	33.66%	0.00%	0.00%	22.44%
500-600	11.94%	26.68%	9.07%	9.07%	0.00%	0.00%	9.46%
600-700	15.09%	8.35%	0.00%	0.00%	0.00%	0.00%	3.91%
700-800	1.73%	0.00%	0.00%	0.00%	0.00%	0.00%	0.29%
800-900	-1.97%	6.89%	0.00%	0.00%	0.00%	0.00%	0.82%
900-997	-0.26%	-0.26%	0.00%	0.00%	0.00%	0.00%	-0.09%
Total Return	-15.44%	-1.77%	-14.18%	4.24%	-22.68%	-17.92%	-11.29%
SD of returns	19.41%	20.35%	16.14%	14.20%	7.49%	6.01%	13.93%
Sharpe Ratio	-0.7959	-0.087	-0.8786	0.29864	-3.03	-2.9816	-1.246

Modified Bollinger Band Strategy #1

PORT-3MFUT	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
100-200	8.01%	-1.41%	-0.07%	4.00%	8.04%	0.00%	3.10%
200-300	-7.84%	3.22%	12.60%	14.63%	0.00%	0.00%	3.77%
300-400	-34.25%	-34.25%	-34.25%	-34.25%	-13.22%	-13.22%	-27.24%
400-500	9.68%	9.68%	9.68%	9.68%	9.68%	9.68%	9.68%
500-600	-3.49%	-3.49%	-3.49%	-3.49%	-3.49%	-3.49%	-3.49%
600-700	-9.27%	-9.27%	-9.27%	-9.27%	-9.27%	-9.27%	-9.27%
700-800	-3.23%	-3.23%	-3.23%	-3.23%	-3.23%	-3.23%	-3.23%
800-900	6.69%	6.69%	6.69%	6.69%	6.69%	6.69%	6.69%
900-997	-7.89%	-7.89%	-7.89%	-7.89%	-7.89%	-7.89%	-7.89%
Total Return	-35.10%	-33.66%	-26.64%	-22.27%	-7.04%	-13.96%	-23.11%
SD of returns	12.54%	12.10%	13.11%	13.55%	7.68%	6.96%	10.99%
Sharpe Ratio	-2.7988	-2.7822	-2.0323	-1.6436	-0.9169	-2.0046	-2.0297

PORT-G17Y	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
100-200	1.60%	-2.55%	-0.72%	2.29%	7.09%	0.00%	1.28%
200-300	-4.57%	16.30%	12.69%	15.90%	0.00%	0.00%	6.72%
300-400	-33.00%	-33.00%	-33.00%	-25.06%	-13.04%	-7.70%	-24.13%
400-500	11.64%	11.64%	11.64%	11.64%	11.64%	11.64%	11.64%
500-600	0.78%	0.78%	0.78%	0.78%	0.78%	0.78%	0.78%
600-700	-4.09%	-4.09%	-4.09%	-4.09%	-4.09%	-4.09%	-4.09%
700-800	-1.17%	-1.17%	-1.17%	-1.17%	-1.17%	-1.17%	-1.17%
800-900	9.50%	9.50%	9.50%	9.50%	9.50%	9.50%	9.50%
900-997	-3.74%	-3.74%	-3.74%	-3.74%	-3.74%	-3.74%	-3.74%
Total Return	-24.13%	-11.31%	-12.46%	3.75%	8.75%	7.79%	-4.60%
SD of returns	12.07%	13.40%	12.92%	11.26%	7.25%	5.91%	10.47%
Sharpe Ratio	-1.9991	-0.8445	-0.9651	0.3334	1.2067	1.3180	-0.1584

Modified Bollinger Band Strategy #2

PCA-FUT	0.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
100-200	-12.13%	-10.62%	-21.62%	-21.49%	-20.24%	-26.52%	-18.77%
200-300	-16.85%	3.82%	5.25%	13.45%	13.45%	13.45%	5.43%
300-400	-15.78%	-25.08%	-20.07%	-41.34%	0.00%	0.00%	-17.04%
400-500	-7.17%	2.43%	10.89%	33.04%	0.00%	0.00%	6.53%
500-600	11.08%	25.11%	13.64%	13.64%	0.00%	0.00%	10.58%
600-700	-14.68%	-0.08%	0.00%	0.00%	0.00%	0.00%	-2.46%
700-800	-9.70%	2.06%	0.00%	0.00%	0.00%	0.00%	-1.27%
800-900	11.08%	7.91%	0.00%	0.00%	0.00%	0.00%	3.16%
900-997	3.22%	0.00%	0.00%	0.00%	0.00%	0.00%	0.54%
Total Return	-45.69%	-1.96%	-16.90%	-21.01%	-9.52%	-16.64%	-18.62%
SD of returns	10.59%	12.72%	11.48%	20.08%	8.07%	9.80%	12.12%
Sharpe Ratio	-4.3159	-0.1544	-1.4718	-1.0464	-1.1799	-1.6968	-1.6442

PCA-G17Y	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
100-200	-4.08%	-10.39%	-22.32%	-21.89%	-24.10%	-24.88%	-17.94%
200-300	-14.15%	8.08%	1.38%	11.18%	11.18%	11.18%	4.81%
300-400	-9.84%	-17.08%	-32.06%	-15.35%	0.00%	0.00%	-12.39%
400-500	-4.79%	6.56%	19.20%	33.66%	0.00%	0.00%	9.10%
500-600	13.87%	27.84%	13.11%	13.11%	0.00%	0.00%	11.32%
600-700	11.09%	9.23%	0.00%	0.00%	0.00%	0.00%	3.39%
700-800	-6.80%	0.00%	0.00%	0.00%	0.00%	0.00%	-1.13%
800-900	-5.17%	3.37%	0.00%	0.00%	0.00%	0.00%	-0.30%
900-997	-8.46%	-0.26%	0.00%	0.00%	0.00%	0.00%	-1.45%
Total Return	-20.97%	23.53%	-27.86%	11.14%	-15.61%	-16.48%	-7.71%
SD of returns	8.87%	12.00%	14.99%	15.27%	8.74%	8.96%	11.47%
Sharpe Ratio	-2.3644	1.9604	-1.8593	0.7297	-1.7865	-1.8383	-0.8597

Modified Bollinger Band Strategy #2

PORT-3MFUT	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
100-200	13.21%	-4.32%	-10.47%	4.77%	7.92%	0.00%	1.85%
200-300	-3.42%	12.35%	13.03%	14.51%	0.00%	0.00%	6.08%
300-400	0.00%	-11.26%	-9.26%	-24.37%	-14.16%	2.51%	-9.42%
400-500	0.00%	-6.53%	-5.42%	9.68%	9.68%	9.68%	2.85%
500-600	0.00%	-11.09%	-3.49%	-3.49%	-3.49%	-3.49%	-4.18%
600-700	0.00%	-22.46%	-9.27%	-9.27%	-9.27%	-9.27%	-9.92%
700-800	0.00%	-24.31%	-3.23%	-3.23%	-3.23%	-3.23%	-6.20%
800-900	-6.54%	0.39%	6.69%	6.69%	6.69%	6.69%	3.43%
900-997	0.00%	-11.64%	-7.89%	-7.89%	-7.89%	-7.89%	-7.20%
Total Return	2.18%	-53.28%	-21.48%	-10.03%	-8.15%	1.64%	-14.85%
SD of returns	5.03%	10.56%	7.56%	11.15%	7.83%	5.89%	8.00%
Sharpe Ratio	0.4332	-5.0462	-2.8412	-0.8996	-1.0410	0.2789	-1.5193

PORT-G17Y	.5SDs	1SDs	1.5SDs	2SDs	2.5SDs	3SDs	Average
100-200	2.49%	-2.93%	-11.53%	-3.79%	6.94%	0.00%	-1.47%
200-300	-1.59%	1.57%	11.10%	13.14%	0.00%	0.00%	4.04%
300-400	-3.47%	-9.49%	-20.60%	-3.98%	-17.01%	1.88%	-8.78%
400-500	0.00%	-2.01%	-2.37%	11.64%	11.64%	11.64%	5.09%
500-600	0.00%	-0.12%	0.78%	0.78%	0.78%	0.78%	0.50%
600-700	0.00%	-4.49%	-4.09%	-4.09%	-4.09%	-4.09%	-3.47%
700-800	0.00%	-1.17%	-1.17%	-1.17%	-1.17%	-1.17%	-0.98%
800-900	0.39%	9.50%	9.50%	9.50%	9.50%	9.50%	7.98%
900-997	-9.59%	-3.74%	-3.74%	-3.74%	-3.74%	-3.74%	-4.72%
Total Return	-2.27%	-9.72%	-20.30%	22.06%	3.65%	18.98%	2.07%
SD of returns	3.29%	4.86%	9.16%	6.87%	8.13%	5.14%	6.24%
Sharpe Ratio	-0.6880	-2.0008	-2.2169	3.2125	0.4483	3.6930	0.4080

Appendix O – Number of Trades and transaction cost

MBB#1		**0.5SDs**	**1SDs**	**1.5SDs**	**2SDs**	**2.5SDs**	**3SDs**
PCA-3MFUT							
	#Trades	6	3	1	0	0	0
	Transaction Cost	0.62%	0.31%	0.10%	0.00%	0.00%	0.00%
PCA-G17Y							
	#Trades	4	2	2	0	0	0
	Transaction Cost	0.42%	0.21%	0.21%	0.00%	0.00%	0.00%
PORT-3MFUT							
	#Trades	3	2	3	2	0	0
	Transaction Cost	0.31%	0.31%	0.31%	0.21%	0.00%	0.00%
PORT-G17Y							
	#Trades	8	3	1	1	0	0
	Transaction Cost	0.81%	0.31%	0.10%	0.10%	0.00%	0.00%

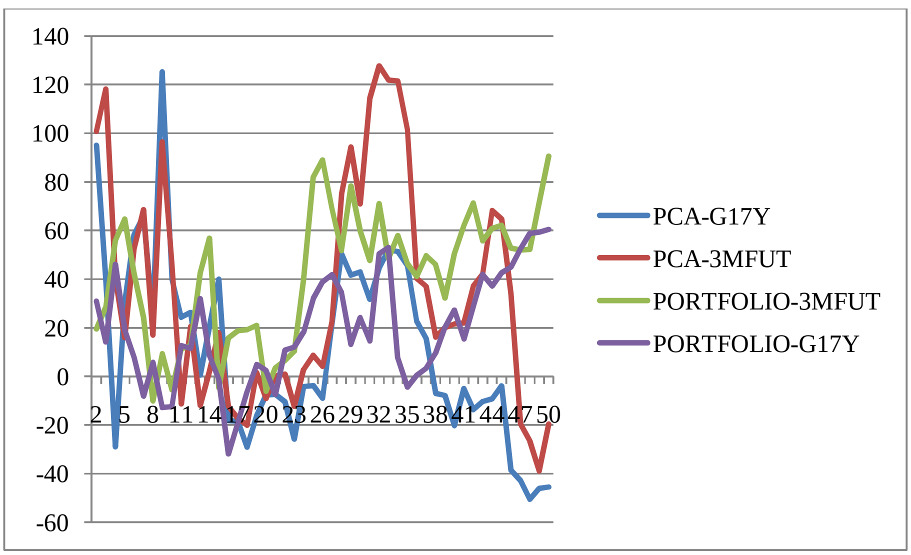
MBB#2		**0.5SDs**	**1SDs**	**1.5SDs**	**2SDs**	**2.5SDs**	**3SDs**
PCA-3MFUT							
	#Trades	13	3	1	1	1	0
	Transaction Cost	1.32%	0.31%	0.10%	0.10%	0.10%	0.00%
PCA-G17Y							
	#Trades	22	7	2	0	0	0
	Transaction Cost	2.22%	0.83%	0.21%	0.00%	0.00%	0.00%
PORT-3MFUT							
	#Trades	10	17	9	2	0	0
	Transaction Cost	1.02%	1.73%	0.92%	0.21%	0.00%	0.00%
PORT-G17Y							
	#Trades	27	2	1	1	0	0
	Transaction Cost	2.73%	0.21%	0.10%	0.10%	0.00%	0.00%

*SD – Standard Deviation

*One transaction costs 0.05% of the invested amount

Appendix P – Moving Average Trading Strategies Equal Weight

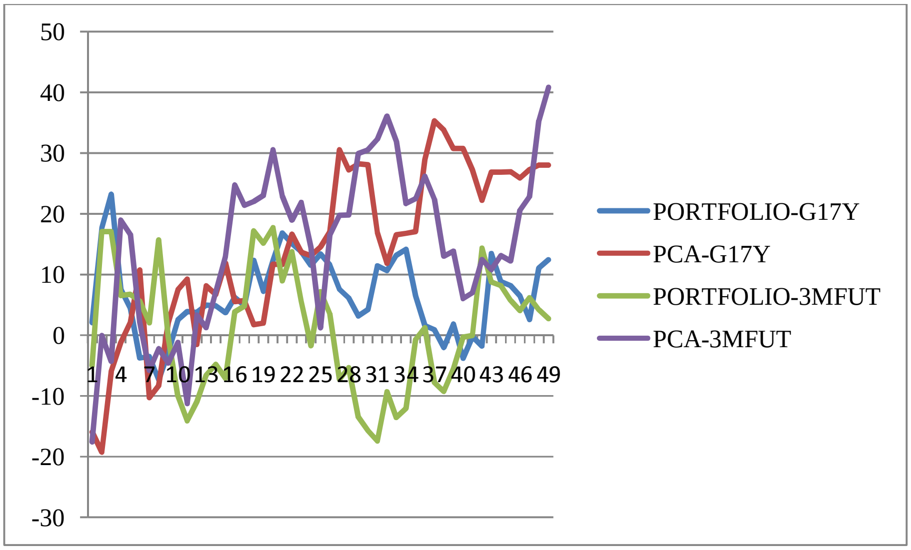
In-Sample	PORT-G17Y	PORT-3MFUT	PCA-G17Y	PCA-3MFUT
0-100	-11.57%	-11.26%	-7.32%	-6.15%
100-200	8.88%	12.80%	10.07%	11.48%
200-300	2.83%	1.46%	-18.32%	-20.06%
300-400	-11.82%	-12.02%	23.63%	23.19%
400-500	-2.99%	-1.18%	3.17%	5.89%
500-600	-10.62%	-12.81%	-19.79%	-23.38%
600-700	11.68%	15.07%	6.03%	16.16%
700-800	17.13%	27.17%	2.10%	7.14%
800-900	11.95%	20.38%	10.13%	21.25%
900-997	13.14%	9.83%	11.63%	10.47%
R/period	5.83%	11.79%	2.21%	8.92%
Total Return	19.95%	41.39%	10.81%	34.12%
SD	15.06%	17.87%	17.81%	20.70%
S-ratio	0.37579	0.66260	0.02473	0.38160



Returns for each value of MA for each pair

Appendix Q – Moving Average Trading Strategies Engle Granger Weighted

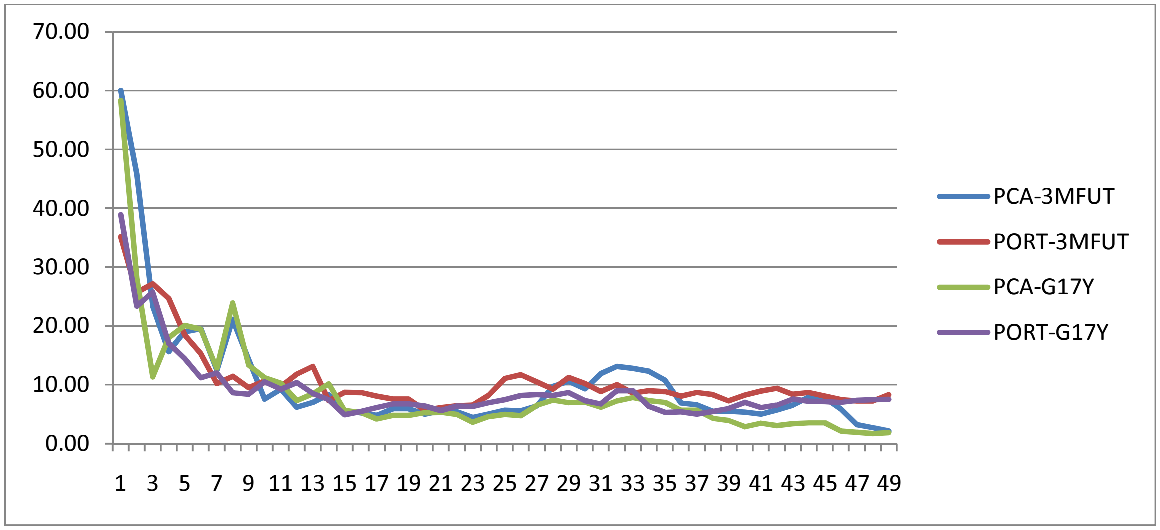
Out-of-Sample	**PORT-G17Y**	**PORT-3MFUT**	**PCA-G17Y**	**PCA-3MFUT**
500-600	-5.68%	-4.58%	-12.49%	-12.61%
600-700	4.35%	1.66%	3.24%	1.23%
700-800	-0.59%	-3.44%	13.67%	12.02%
800-900	4.89%	4.84%	6.55%	10.26%
900-997	4.08%	3.20%	4.20%	5.22%
R/period	3.20%	0.35%	6.72%	7.05%
Total Return	6.59%	0.93%	14.31%	14.98%
St.Dev.	5.32%	6.51%	8.58%	9.88%
Sharpe-Ratio	0.61002	0.01518	0.68271	0.66123



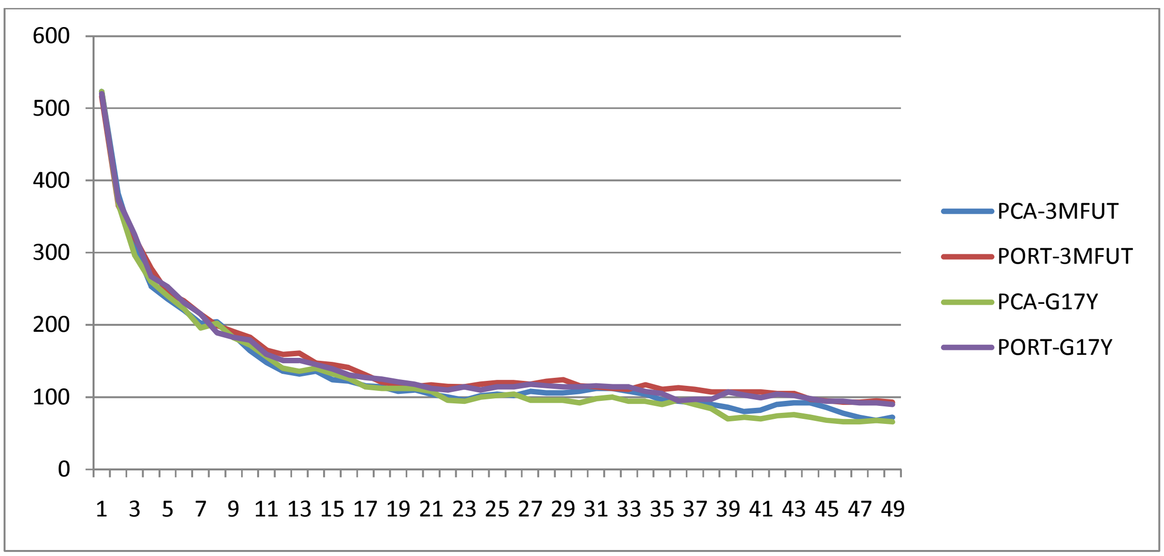
Returns for each value of MA for each pair

Appendix R – Transaction Cost Moving Averages Equal Weighted

Transaction Costs in terms of reduced return

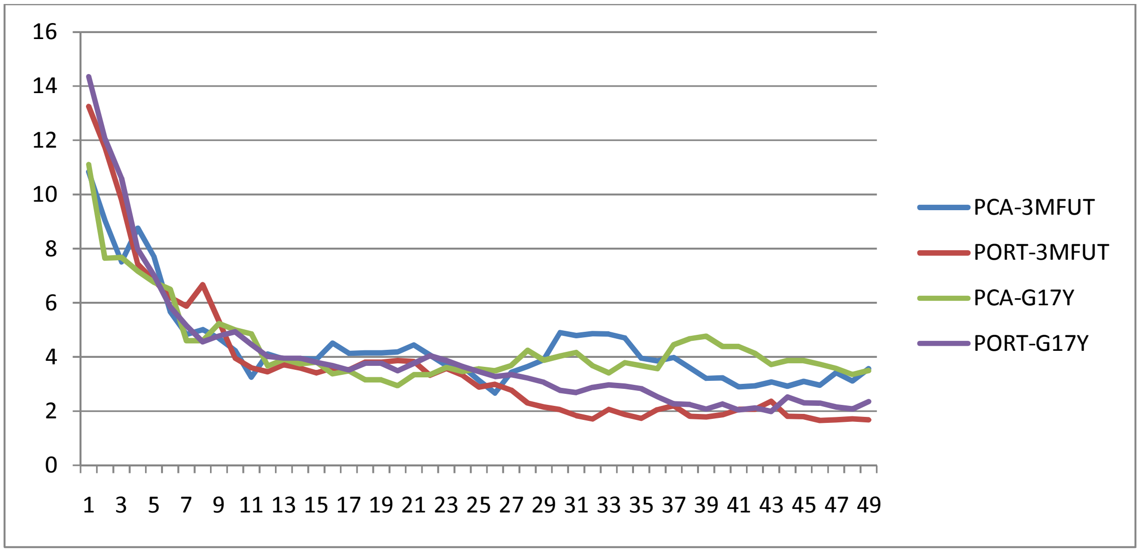


Number of transactions



Appendix S – Transaction Cost Moving Averages Engle Granger Weighted

Transaction Costs in terms of reduced return



Number of transactions