Investment Analysts Journal

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/riaj20

Statistical arbitrage on the JSE based on partial co-integration

A. J. Hoffman

To cite this article: A. J. Hoffman (2021) Statistical arbitrage on the JSE based on partial co-integration, Investment Analysts Journal, 50:2, 110-132, DOI: 10.1080/10293523.2021.1886723

To link to this article: https://doi.org/10.1080/10293523.2021.1886723

INVESTMENT ANALYSTS JOURNAL
2021, VOL. 50, NO. 2, 110–132
https://doi.org/10.1080/10293523.2021.1886723

Check for updates

Statistical arbitrage on the JSE based on partial co-integration

A. J. Hoffman

School of Electrical, Electronic and Computer Engineering, North-West University, Potchefstroom, South Africa

ABSTRACT

Early forms of statistical arbitrage exploited the mean reversion of a model error extracted from pairs of instruments with a tendency to move together. Pairs trading was extended by Engle and Granger and by Johansen to include several co-integrated instruments. Partial co-integration was proposed by Clegg and Krauss to allow for model errors that contain both random walk and mean-reverting components. In this paper we implement a modified version of partial co-integration using a Kalman filter approach that allows the behaviour of the mean-reverting error component to be optimised. Co-integrated sets of shares are compiled over the period from January 1990 to November 2020 based on membership of sectors on the Johannesburg Stock Exchange. We demonstrate that optimal selection of the Kalman filter gain enables the improvement of risk-adjusted returns generated by the partial co-integration strategy. We optimise the parameters that define the partial co-integration trading strategy and find that it significantly outperforms market returns and a strategy based on normal co-integration. We observe higher returns during bear cycles compared with bull cycles, making statistical arbitrage based on partial co-integration an attractive option to combine with trading strategies that perform well during bull markets.

ARTICLE HISTORY
Received 10 July 2020
Accepted 31 January 2021

KEYWORDS
statistical arbitrage; partial co-integration; Kalman filter; sector-based

1. Introduction

Most investment and trading strategies are based on the extraction of historical relationships that are expected to be at least partly continued into the future. Statistical arbitrage is a variation on this theme, as it assumes that well-established relationships between two or more financial instruments will tend to be restored once they have deviated from the historical norm. An early version of statistical arbitrage is pairs trading that identifies pairs of instruments that normally move together (Zhou & Kang, 1984). This concept was extended by Engle and Granger (Co-Integration and Error Correction: Representation, Estimation and Testing, 1987) and later by Johansen (Likelihood-based Inference in Cointegrated Vector Autoregressive Models, 1995) to include not only two but a number of related instruments. They defined methods to test the presence of co-integration: if a set of instruments are co-integrated, the residual of a linear regression relationship between them will be quasi-stationary. As a stationary time-series will also tend to be mean reverting, it can be exploited for trading purposes by assuming that significant deviations from the long-term mean are likely to be followed by movements back towards this mean.

A weakness in statistical arbitrage models is the fact that co-integrated relationships extracted from historical behaviour are not always sustained into the future (Gatev, Goetzmann, & Rouwenhorst, 2006). The implication is that the model error, that is assumed to be stationary, do not always revert to its long term mean but may start behaving like a random walk. This will result not only in

CONTACT A. J. Hoffman alwyn.hoffman@nwu.ac.za
Investment Analysts Journal is co-published by NISC Pty (Ltd) and Informa Limited (trading as Taylor & Francis Group)

infrequent trading opportunities but can also lead to large losses if no stop loss is implemented. This phenomenon was addressed by Clegg and Krauss (Pairs trading with partial cointegration, 2018) through a concept called partial co-integration. Their revised model allows for both a random walk and a mean reverting component in the residual model error and formulate co-integration as a state space model that can be solved using Kalman filter techniques (Kalman, 1960). By trading on the mean reverting component rather than on the total model error, a higher frequency of trades will be possible that may result in higher returns and the avoidance of large losses. A problem associated with partial co-integration is the fact that no profit is guaranteed when the mean reverting component of the error returns to zero, as the random walk component also contributes to the effective gain for that trade. The method proposed by Clegg and Krauss (Pairs trading with partial cointegration, 2018) does not allow any control over the size of the mean reverting compared to the random walk components of the co-integration error; the only method to control the strategy is by selecting the threshold level at which trades are entered. Their results were limited to the selection of pairs of instruments, rather than allowing any number of instruments to form part of a co-integrated set. This limits the number of arbitrage opportunities in markets that do not have as many listed instruments as the NYSE. They furthermore limited their approach to using a 12-month formation and 6-month trading period for extracted models. There is no theoretical or empirical evidence that this is optimal; the fact that most markets have been shown to change behaviour over time indicates that more frequent updating of trading models could be beneficial.

In our paper we address the fundamental question whether it is possible to create artificial investment instruments from real instruments displaying behaviour that is sufficiently close to stationary to allow profitable exploitation through statistical arbitrage. For this purpose, we implement a modified version of partial co-integration which allows us to control the fraction of the error signal residing in the random walk and mean reverting components respectively. This is achieved by selecting the values of the mean-reverting autoregressive coefficient and the Kalman filter gain before solving the set of Kalman filter equations. We demonstrate that optimisation of these model parameters results in superior risk-adjusted returns.

In line with earlier research we also investigate the performance of trading strategies based on co-integration and partial co-integration during different cycles of the market and find that both strategies achieve higher returns during bear cycles compared to bull cycles of the market. A unique contribution of this paper is that it is the first report on the application of partial co-integration based statistical arbitrage that allows more than two instruments per co-integrated set. We furthermore investigated the influence of all of the important parameters that must be selected to define the trading strategy.

Against the above background we formulate the following research questions to be addressed by this paper:

1. Does partial co-integration produce superior results compared to normal co-integration when applied to JSE shares, and does it outperform market returns on a risk-adjusted basis?
2. What is the impact of bear and bull market cycles on the profitability of a statistical arbitrage trading strategy?
3. Which parameters forming part of the definition of a partial co-integration trading strategy have a significant influence on trading returns, and what are optimal ranges for their values?
4. Are the co-integration relationships between sets of shares on which statistical arbitrage is based sufficiently stable to produce profitable results over future time periods when using models extracted from past time periods?

The rest of the paper is organised as follows. Section 2 provides an overview of relevant literature, while section 3 describes the dataset and methodology that were used. In section 4 we provide the results and discuss them. Section 5 concludes and makes recommendations for future research.

2. Previous research

Arbitrage as a trading strategy is theoretically intriguing as it can provide risk-free profit at zero cost. In practice, however, risks do exist in arbitrage such as the devaluation of an instrument that is being traded in (Sharpe & Alexander, 1990).

2.1. Literature on statistical arbitrage, pairs trading and partial co-integration

Many statistical arbitrage strategies are focused on the concept of mean-reversion of security prices. Some forms of statistical arbitrage are pairs trading and long/short strategies (Thorp, A Perspective on Quantitative Finance: Models for Beating the Market, 2003). Statistical arbitrage is very popular in the hedge fund industry, using market neutral strategies or long/short strategies to produce low-volatility investment strategies that inherently take advantage of diversification across assets (Thorp, A Mathematician on Wall Street, 2000).

The concept of mean-reversion trading is built on the assumption that a security's high and low prices are only temporary and that the series will revert to a certain mean value over time (Bhave & Libertini, 2013). Mean-reversion strategies are however only effective when price series are stationary. The price series of most securities are not stationary since prices are subject to drifts such as those caused by trends and momentum. Even though single price series are seldom stationary, a stationary price series can be obtained by creating a linear (weighted) combination of securities that exhibit a co-integrated relation.

Order of integration d denoted $I(d)$, is a summary statistic that reports the minimum number of differences that is required to obtain a covariance stationary series. If the returns of two or more series are individually integrated and the order of integration between the series differ, the series are said to be co-integrated (Engle & Granger, 1987). Some co-integration testing techniques include the Engle-Granger two-step method (Sjo, 2008), the Johansen test (Johansen, 1995) and the Phillips-Ouliaris test (Phillips & Ouliaris, 1990). In contrast to the Engle-Granger method and Phillips-Ouliaris test, the Johansen test can be used to test multiple time series for co-integration.

Pairs trading was pioneered as a popular market-neutral trading strategy to exploit co-integrated relations that exist in the market between pairs of instruments (Zhou & Kang, 1984). Vidyamurthy (Pairs Trading: Quantitative Methods and Analysis, 2004) suggested that arbitrage pricing theory (APT), which was first suggested by Ross (The arbitrage theory of capital asset pricing, 1976) to determine asset prices, can be used to detect tradeable pairs. Gatev (Pairs Trading: Performance of a Relative Value Arbitrage Rule, 2006) proposed the minimum distance method to identify pairs to which statistical arbitrage could be applied.

In the case of pairs trading, a certain hedge ratio must be determined after obtaining securities that have a co-integrated relation. Different approaches to calculating the optimal hedge ratio have been investigated in the past. Previous studies (Rudy, 2011), (Chan, 2013) provide promising results employing the Kalman filter to determine the hedge ratio dynamically.

The application of statistical arbitrage trading on US energy futures markets was investigated by Kanamura, Rachev, and Fabozzi (The application of pairs trading to energy futures markets, 2008). A simple pairs trading strategy was implemented using a mean-reverting process of the futures price spread of specifically WTI crude oil, heating oil and natural gas.

Harlacher (Cointegration Based Statistical Arbitrage, 2012) proposed an approach based on an error correction model (ECM) to identify pairs. In an ECM, the dynamics of one time series at a certain time point is a correction of the last period's deviation from the equilibrium with the addition of possible lag dynamics.

Avellaneda and Lee (Statistical Arbitrage in the U.S. Equities Market, 2008) studied model-driven statistical arbitrage strategies in the US equity market over the period of 1997 to 2007. They focused on the use of principle component analysis (PCA) applied to sector ETFs by modelling the residuals and idiosyncratic components of stock returns as a mean-reverting process. PCA-based strategies achieved an average annual Sharpe ratio of 1.44 over the period.

In a study on the São Paulo stock exchange Caldeira and Moura (Selection of a Portfolio of Pairs Based on Cointegration: A Statistical Arbitrage Strategy, 2013) applied a cointegration-based statistical arbitrage strategy for the period of January 2005 to October 2012. Their empirical analysis focused on estimating long-term equilibrium and modelling the resulting residuals. Their model obtained excess returns of 16.38% per annum with an average Sharpe ratio of 1.34.

Jarrow and Li (Exploring statistical arbitrage opportunities in the term structure of CDS spreads, 2016) estimated an affine model for the term structure of credit default swap (CDS) spreads on North American companies and identified mis-valued CDS contracts along the credit curve. Their trading rules were contrarian by betting that mis-valuations will disappear over time. In some studies of statistical arbitrage, such as by Miao (High Frequency and Dynamic Pairs Trading Based on Statistical, 2014), a correlation matrix is calculated for the entire universe. Filters are then applied by only selecting highly correlated securities to form pairs for trading. Frey and Dueck (Clustering by passing messages between data points, 2007) provide arguments for using non-parametric clustering methods with the introduction of a clustering algorithm called affinity propagation.

Focardi et al. (A new approach to statistical arbitrage: Strategies based on dynamic factor models of prices and their performance, 2016) introduced a new statistical arbitrage strategy based on dynamic factor models of prices. Their empirical analysis applied to the stock of companies included in the S&P 500 index statistically tested the relative forecasting performance using the Diebold–Mariano framework and performing the test for statistical arbitrage proposed by Hogan et al. (Testing market efficiency using statistical arbitrage with applications to momentum and value strategies, 2004). Their results show that statistical arbitrage applied to share prices allow for significantly more accurate forecasts than when applied to share returns.

Baviera and Baldi (Stop-loss and leverage in optimal statistical arbitrage with an application to energy market, 2019) developed a statistical arbitrage trading strategy with two key elements in high-frequency trading: stop-loss and leverage. They considered a mean-reverting process for the security price with proportional transaction costs and used stop-loss and leverage in an optimal trading strategy. Sánchez-Granero et al. (Testing the efficient market hypothesis in Latin American stock markets, 2020) used the evolution of the Hurst Exponent of a pair to show how in emerging markets a pairs trading strategy can still be profitable though it is not profitable any more in developed markets, in accordance with the weak form of market efficiency.

A potential weakness of co-integration is the fact that the error signal is often not completely stationary but also contains random walk behaviour. Partial co-integration as defined by Clegg & Krauss (Pairs trading with partial cointegration, 2018) is a weakening of the co-integration condition that allows for the co-integrating residual to contain a random walk and a mean-reverting component. They derived its representation as a state space model that can be solved using Kalman filter techniques, providing a maximum likelihood-based estimation routine, and a suitable likelihood ratio test.

2.2. Literature on the application of statistical arbitrage to JSE stocks

Mashele, Terblanche, and Venter (Pairs trading on the Johannesburg Stock Exchange, 2013) implemented a pairs trading strategy using bank shares listed on the JSE over the period January 2005 to January 2012, selecting pairs from stocks in economically meaningful groups. They also specifically investigated the impact of trading costs and trader skills and found this to have a significant impact on the profitability of the pairs trading strategy.

Appelbaum (Does Pairs Trading Work On The Johannesburg Stock Exchange?, 2015) performed a study based on data of the top 80 shares on the JSE, selected based on market capitalisation, for the period 1998–2012. They tested strategies including both unrestricted pairs (where any two shares can be paired) and restricted (only pairs within the same sectors could be paired). They found that most of their strategies outperformed the ALSI after trading costs, and that pairs trading resulted in lower standard deviations of returns than the ALSI. They also observed a decline in pairs trading performance over time and found that pairs trading outperformed the market during periods of increased market stress, e.g., around 2008.

Visagie and Hoffman (Comparison of Statistical Arbitrage in Developed and Emerging Markets, 2017) investigated the application of statistical arbitrage based on co-integration in developed and emerging markets, including the JSE, over the period of 2006 to 2016. Their results indicate that statistical arbitrage systems produce higher excess returns during non-trending markets and that these systems generally provide higher returns in emerging markets.

3. Data and methodology

3.1. Data

JSE share data was sourced from IRESS for the period from January 1990 to November 2020. The closing price data was adjusted for share splits and similar corporate actions. Dividend data was used to adjust daily share returns on the day that dividends accrued to shareholders. The data set included shares that were delisted during the period, making the data free from survivorship bias.

It is known that a statistical arbitrage strategy involves frequent trading and can thus be expected to be sensitive to trading costs. We therefore excluded shares that may be subject to high effective trading costs. Shares that are not liquid tend to suffer from much higher effective trading costs than official brokerage fees due to among other reasons large bid-ask spreads and small trading lots. As microcap shares are not well-behaved in either of these respects, we excluded all microcap shares from our selected portfolios. This is in line with the strategies employed by other researchers (Krauss, Do, & Huck, 2017), (Stübinger, 2019) who only considered shares forming part of the index for the respective exchanges.

3.2. Compiling groups of shares from which to extract cointegrated sets

Our objective was to use the techniques developed by Engle and Granger (Co-Integration and Error Correction: Representation, Estimation and Testing, 1987) and Johansen (Likelihood-based Inference in Cointegrated Vector Autoregressive Models, 1995) to identify groups of shares that in combination display co-integration, rather than pairs. This was expected to result in a larger number of trading opportunities, which could enable the strategy to be always invested and to include more diversification, by investing in more than one co-integrated group of shares at any point in time.

One of the potential weaknesses of statistical arbitrage is that the co-integrated relationships may not be sustained outside of the period from which it was extracted. To reduce this risk, we decided to only search for sets of co-integrated shares within the same sector or industry. For this purpose, we used information obtained from the JSE that describe the membership of all firms to industries, super-sectors, sectors and subsectors. The JSE categorises shares in this way to group together companies with similar operations and that are exposed to similar external factors. 'Industries' is the widest definition used for this purpose; this definition becomes increasingly more specific for super-sectors, sectors and subsectors, e.g., Awethu Breweries belongs to Industry: Consumer Goods, Super-sector: Food & Beverages, Sector: Beverages and Subsector: Distillers & Vintners.

As the size of industries and sectors varies significantly, we set a minimum and a maximum size for the group of shares to be identified, in this case 2 and 30 respectively. We also set a maximum limit of 30% on the fraction that one group of shares may represent of the total average market capitalisation of the exchange. Using these rules we started off at the level of industries (each of which including one or more super-sectors) and worked our way down to the level of subsectors, until we identified a combination of share groupings at either industry, super-sector, sector or subsector level that satisfied the above size and market capitalisation constraints. For model sectors with more than 12 shares we selected the 12 with the highest market capitalisation, as the Johansen technique cannot handle more than 12 instruments at one time. Following this approach we arrived at 18 so-called model sectors with size varying from 2 to 12 shares and that in combination included all non-microcap shares, as displayed in Table 3.

3.3. Pairs trading, statistical arbitrage and partial co-integration

Co-integration is defined when the error term in regression modelling is stationary. In mathematical terms, if two variables x_t and y_t are co-integrated, a linear combination of them must be stationary such that:

$$x_t - \beta y_t = u_t \tag{1}$$

where u_t is a stationary process. The Augmented Dickey-Fuller (ADF) test is used to test a time series for stationarity. If a time series passes the test and is indeed stationary, it is expected that a dependency exists between historic values and future values of the time series. If a previous value was above the mean it is expected that the upcoming value will tend to move down towards the mean, and *vice versa*. These expectations have a strict probability after stationarity has been confirmed by the ADF test (Fuller, 1976).

Johansen (Johansen, 1995) considers a general p dimensional vector autoregressive ($VAR(p)$) model for k time series, integrated of order d such that $\{x\}_t \sim I(d)$:

$$X_t = \mu + \Phi D_t + \Pi_p X_{t-p} + \ldots + \Pi_1 X_{t-1} + \varepsilon_t \qquad , \ t = 1, \ \ldots, \ T \tag{2}$$

where μ is a $k \times 1$ vector of constants, ΦD_t represents deterministic trends, X_{t-p} is the p^{th} lag of X and ε_t is a $k \times 1$ vector of error terms. As with a unit root test, it is possible that either a constant term (μ), a trend term (D_t), both terms or neither terms may be present in the model.

The method to identify co-integrated share sets consists of two distinguishable periods: a formation period and a trading period. If two security's price series are both integrated of order d and there is a linear combination of the two price series that creates a series which is integrated of order $d - b$ where ($b > 0$), then the two series are considered to be co-integrated ($CI(d, b)$). In the framework of statistical arbitrage, interest is placed on the situations where $d - b = 0$, such that there exists a stationary time series for the spread. As many price series are integrated of order 1 ($I(1)$) focus is placed on the situation where $b = d = 1$. A very advantageous part of two co-integrated price series X_t and Y_t is that these series can be represented in an error correction model (ECM). In an ECM, the dynamics of one time series at a certain time point is a correction of the last period's deviation from the equilibrium with the addition of possible lag dynamics. Harlacher (Cointegration Based Statistical Arbitrage, 2012) expresses this relation mathematically as:

$$\Delta y_t = \psi_0 - \gamma_y(y_{t-1} - \alpha - \beta x_{t-1}) + \sum_{i=1}^{K} \psi_{x,i}\Delta x_{t-i} + \sum_{i=1}^{L} \psi_{y,i}\Delta y_{t-i} + \epsilon_{y,t} \tag{3}$$

and similarly

$$\Delta x_t = \xi_0 + \gamma_x(y_{t-1} - \alpha - \beta x_{t-1}) + \sum_{i=1}^{K} \xi_{y,i}\Delta y_{t-i} + \sum_{i=1}^{L} \xi_{x,i}\Delta x_{t-i} + \epsilon_{x,t} \tag{4}$$

where $\epsilon_{y,t}$ and $\epsilon_{x,t}$ represents white noise and the terms Δy_t and Δx_t represents one period differences in y_t and x_t respectively. If there is no deterministic trend in the series, the constants ψ_0 and ξ_0 are zero. The advantage of using an ECM is that active predictions can be simply done by using past information.

Table 3. List of model sectors and their trading characteristics.

Nr	Sector Name	Number shares per sector	Fraction periods no co-integration	Number of transactions	Fraction positions profitable	Fraction transaction profitable	Accumulative return	Sharpe ratio	Ave number share positions
1	Forestry & Paper	2	0.99	0	0.00	0.00	0.0%	0.00	0.00
2	Industrial Metals	6	0.78	97	0.56	0.67	39.6%	−0.08	0.27
3	General Mining	5	0.67	133	0.51	0.70	52.1%	0.03	0.37
4	Gold Mining	10	0.83	81	0.47	0.59	−10.2%	−0.49	0.38
5	Platinum & Precious Metals	5	0.62	202	0.54	0.76	244.2%	0.43	0.34
6	Basic Materials C	3	0.90	39	0.56	0.69	18.2%	−0.40	0.07
7	Food & Beverages	11	0.70	159	0.52	0.67	24.3%	−0.05	0.97
8	Consumer Services	12	0.83	105	0.54	0.67	2.0%	−1.35	0.79
9	Banks	8	0.59	255	0.56	0.67	14.5%	−0.15	0.90
10	Financial Services	12	0.37	380	0.57	0.55	34.4%	−0.15	1.49
11	Insurance	10	0.70	115	0.56	0.70	112.0%	0.13	0.38
12	Investment Instruments	5	0.65	154	0.54	0.78	176.8%	0.17	0.29
13	Real Estate	12	0.78	103	0.49	0.60	−4.2%	−1.01	0.64
14	Health Care	4	0.87	63	0.57	0.73	4.6%	−0.32	0.10
15	Industrials	12	0.70	168	0.52	0.64	3.0%	−0.61	1.26
16	Technology	3	0.96	14	0.48	0.29	−26.7%	−1.08	0.03
17	Telecomms	5	0.91	33	0.47	0.52	−19.5%	−0.78	0.10
18	Unclassified	12	0.83	86	0.50	0.62	32.3%	−0.84	0.67

From equations (3) and (4) it is clear that the part that represents the deviation from the long-run equilibrium is $(y_{t-1} - \alpha - \beta x_{t-1})$. For error-correcting behaviour to be present this term must be weakly stationary and the two coefficients γ_y and γ_x must have opposite algebraic signs. To test for co-integration the procedure that was proposed by Engle and Granger (Engle & Granger, 1987) must be applied. This procedure consists of two steps. First, a linear regression is run of the one series on the other:

$$y_t = \alpha + bx_t + \varepsilon_t \tag{5}$$

where ε_t is the error term. If the two variables are integrated of order one, the error term of the regression must be weakly stationary ($I(0)$). If this is the case, x_t and y_t are co-integrated. To test for stationarity in the error term, the augmented Dickey-Fuller test (Fuller, 1976) can be used. Since the series to be tested is based on estimated coefficients, it is necessary to search for the presence of a unit root in the residuals by using the critical values that are provided by Phillips and Ouliaris (Asymptotic properties of residual based tests for, 1990).

The composition of the co-integration portfolio is dictated by the regression coefficients of the model, as described in equation (6) below, which was derived from equation (5) above:

$$\varepsilon_t = y_t - \alpha - bx_t \tag{6}$$

In this case b and x_t are vectors, as there may be more than two shares in the co-integrated set. We define the coefficients for the shares, that will also determine their relative weights in the co-integrated portfolio, as follows:

$$\beta = (1, \ -b) \tag{7}$$

The size of the position taken in the share that is modelled y_t therefore has a relative value of one, while the relative sizes of the positions taken in the other shares are determined by the components of b. By investing into each share from the model sector in proportion to the size of the regression coefficient, a resulting portfolio is obtained for which the combined value at any point in time corresponds to the model error or residual signal.

The next step is to select a trading strategy for opening and closing positions. During trading the residual signal of each co-integrated set is checked dynamically by calculating a z-score with regards to the moving average μ_ε and standard deviation σ_ε of the error that were determined over the formation period:

$$z_t = \frac{\varepsilon_t - \mu_\varepsilon}{\sigma_\varepsilon} \tag{8}$$

Trading signals are generated when z_t assumes predetermined values, e.g., when $z_t = -2$ a long position is taken in the portfolio, while a short position is taken when $z_t = 2$.

The solution proposed by Clegg and Krauss (Pairs trading with partial cointegration, 2018) extends the model in equation (5) above as used by Engle and Granger to provide for an error that consist of a random walk element, a mean reverting autoregressive element, as well as a stochastic component. Equation (5) can then be rewritten as follows:

$$y_t = c_0 + bx_t + \varepsilon_t = c_0 + bx_t + \varepsilon_{s,t} + \varepsilon_{mr,t} + \varepsilon_{rw,t} \tag{9}$$

where $\varepsilon_s(t)$ is the stochastic, $\varepsilon_{mr}(t)$ the mean reverting and $\varepsilon_{rw}(t)$ the random walk components of the error signal. As the different error signal components and the true values of the shares are not directly observable, Clegg and Krauss (Pairs trading with partial cointegration, 2018) restated the model in state space; this allows the unknown variables to be estimated using a Kalman filter approach. A Kalman filter effectively estimates a more accurate value for a stochastic variable that cannot be measured accurately by using a weighted average between a value estimated from previous observations and the value of the current observation, which is assumed to be uncertain

(Kalman, 1960). The relative weights for the two contributions are based on the levels of uncertainty that exists about the two contributions, derived from noise figures.

In this case the different states to be estimated from observations are the true share values x_t, the mean reverting error component $\varepsilon_{mr,t}$ and the random walk error component $\varepsilon_{rw,t}$. The observation equation and state equation can respectively be written as follows, similar to the equations used by Clegg and Krauss (Pairs trading with partial cointegration, 2018):

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{vmatrix} b & 1 & 1 \\ 1 & 0 & 0 \end{vmatrix} \begin{vmatrix} x_t \\ \varepsilon_{mr,t} \\ \varepsilon_{rw,t} \end{vmatrix} \tag{10}$$

$$\begin{vmatrix} x_t \\ \varepsilon_{mr,t} \\ \varepsilon_{rw,t} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & \rho & 0 \\ 0 & 0 & 1 \end{vmatrix} \begin{vmatrix} x_{t-1} \\ \varepsilon_{mr,t-1} \\ \varepsilon_{rw,t-1} \end{vmatrix} \begin{vmatrix} \varepsilon_{xs,t-1} \\ \varepsilon_{mrs,t-1} \\ \varepsilon_{rws,t-1} \end{vmatrix} \tag{11}$$

where $\varepsilon_{xs,t-1}$, $\varepsilon_{mrs,t-1}$ and $\varepsilon_{rws,t-1}$ are the innovations of the state variables $\bar{x}_t$, $\varepsilon_{mr,t}$ and $\varepsilon_{rw,t-1}$. By solving these equations it is then possible to estimate the values of the mean reverting error and random walk error. Instead of trading on the total error, only the mean reverting error component is used to trigger trades; this increases the trading frequency and prevents situations where indefinite losses can be incurred due to long term drift in the random walk component.

3.4. Co-integration trading strategy

The statistical processes to test instruments for co-integration require a significant time period of share availability in order to extract reliable relationships. Other researchers extracted the model over a training period (12 months) and then tested it over a subsequent period (6 months) (Gatev, Goetzmann, & Rouwenhorst, 2006) (Clegg & Krauss, 2018). As no proof was provided that this approach was optimal, we rather found an optimal value for the model training period by varying the length between 250 days (approximately one year) and 2000 days (approximately 8 years). For a share to be considered for inclusion in the cointegrated set, it had to be available for the complete cointegration period.

A new model was extracted for each model sector on a daily basis and tested for a sufficient level of co-integration using the h-values of the Engle-Granger or Johansen tests. The Dickey-Fuller test was then applied to the resulting model error to test for stationarity. Both tests had to be passed before a model was accepted to compile a co-integrated portfolio. If a co-integrated relationship was found, the same model was used for that sector until the conditions for exiting the trade were satisfied. After exiting a co-integrated portfolio, the daily compilation of new models was resumed until a new co-integrated relationship was found. The standard deviation multiplier or z-value of equation (8) based on which trades were entered was varied as part of the optimisation process. A position was exited once the model error once again reached a zero value.

For short positions the effective positions in the shares were the negative values of the model coefficients. As model coefficients can have any sign, in most cases each portfolio consisted of both short and long positions in the various shares. This meant that the effective investment position was mostly close to zero, as long positions were partly offset by short positions in other shares. A typical trade that is entered and exited is illustrated in Figure 1. The first graph shows the closing prices of the shares forming part of the co-integrated set. The 2nd graph shows the error signal and the high and low thresholds to trigger a trade, which can be seen to vary as the level of volatility of the error signal changes over time. The blue star shows the position where a long trade is entered, once the error signal broke through the lower threshold, and the red star the point where the trade exits once the error signal broke upwards through zero. The points where the trades occurred are also shown in the closing price graph, with an indication of the returns produced by each share. The bottom graph shows the net return of the trade as function of time – it can be seen that in this case a

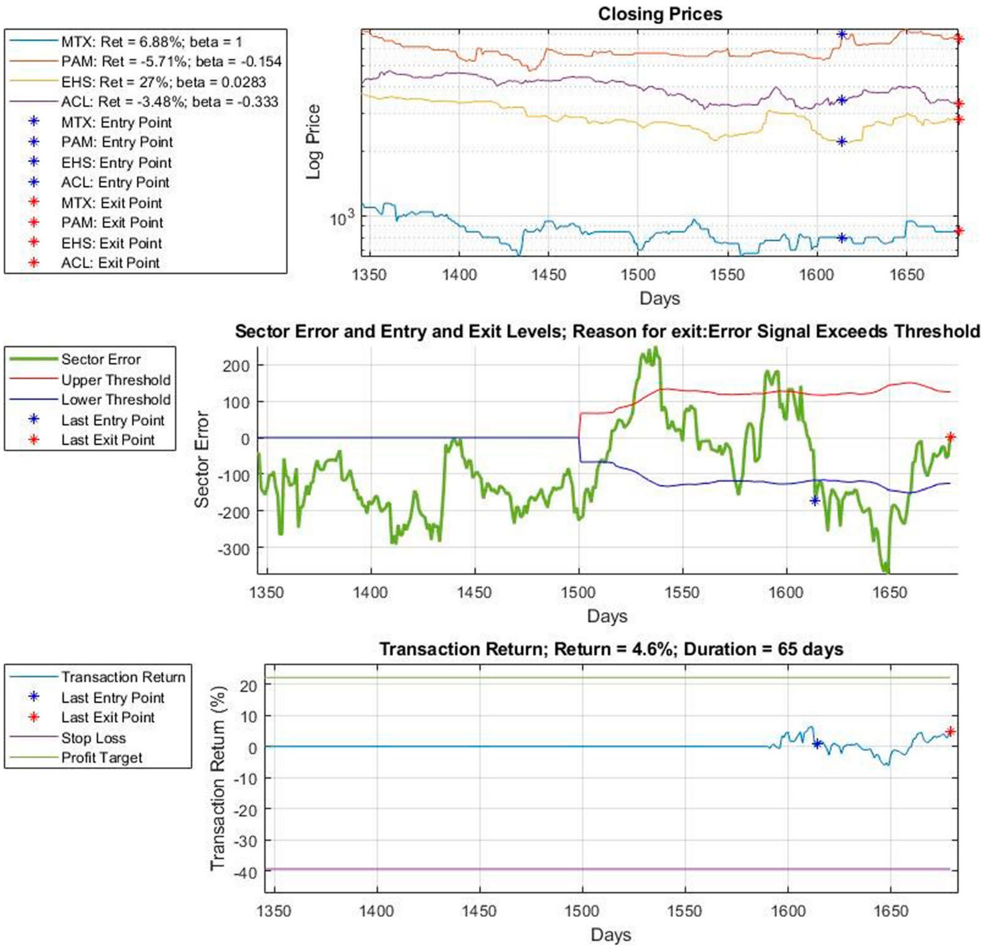


Figure 1. Illustration of a typical co-integration strategy trade. The top graph displays the closing prices of a group of co-integrated shares. The middle graph displays the error signal and the trading thresholds, with entry and exit point indicated by *. The bottom graph displays the net return, stop loss and profit target.

loss is first registered as the error signal moves further below the threshold, but that this turned into a profit by the time that the trade exits. The return graph also displays the stop loss and profit target levels.

If the above trading strategy is followed using the entire model error to trigger trades, a guaranteed profit is made (excluding trading costs) each time that a position is taken and exited once the model error reaches a value of zero. This is however only the case if the model error indeed reverts to zero within an acceptable time period. Should the co-integrated relationship between the shares break down after a position was taken, it could mean that the model error could move away from zero indefinitely, as is typical for random walks. In such a case an unlimited loss can be incurred on the position that was assumed; for this reason, it was found to be necessary to also implement stop losses. Even when using stop losses, a model error behaving like a random walk could result in a situation where the co-integration strategy repeatedly enters loss-making trades during the same long-term trend of the error signal. To prevent such a situation an additional condition for entering a trade was defined: if a stop loss was hit for a specific sign of the error signal, a trade in the same direction could only be entered once the error signal has again crossed the zero threshold.

It could similarly happen that a profit is shown before the zero-error level is reached, but that this turns into a loss before the zero-crossing condition for exiting the position was satisfied. In order to take such profits a profit target was set that provides an alternative condition to exit the strategy on a profitable basis. This is specifically of relevance as the size of the error at the point of entering a position depends on the standard deviation of the error at that stage. Due to the heteroscedasticity that is often observed among financial instruments, the volatility of the underlying shares and thus of the model error is changing all the time, thus leading to widely varying standard deviations.

It may also happen that a co-integrated set of shares is identified, but that while the trade is still active one or more of the underlying shares become delisted. If such situations are avoided by not considering shares that are not available for the entire testing period, an element of survivorship bias would be built into the strategy. We therefore added a rule that forced the portfolio to be exited once any of the underlying shares became unavailable. Similarly, the trade was exited when the co-integration between the selected set of shares was lost, based on the Engle-Granger, Johansen and Dickey-Fuller tests.

3.5. Partial co-integration trading strategy

We found that in most cases the model error for a set of shares that satisfied the conditions for co-integration at the time of model extraction, tends to display behaviour that is a combination of a random walk and mean-reversion. The result of the random walk component of the error signal is to reduce the frequency of mean reversion, and in some cases to prevent the error signal from returning to zero at all, resulting in large trading losses.

The state equations for the mean reverting and random walk error components, that form part of equation (11) above, can respectively be rewritten as follows:

$$\varepsilon_{mr,t} = \rho\varepsilon_{mr,t-1} + \varepsilon_{mrs,t} \tag{12}$$

$$\varepsilon_{rw,t} = \varepsilon_{rw,t-1} + \varepsilon_{rws,t} \tag{13}$$

where each state variable is assumed to have its own zero mean innovation process $\varepsilon_{mrs,t}$ (with variance σ^2_{mr}) and $\varepsilon_{rws,t}$ (with variance σ^2_{rw}) respectively. By regarding the closing prices and the mean-reverting and random walk error signals as system states, we formulated the equations as a Kalman filter problem, in line with (Clegg & Krauss, 2018). Each consecutive observed closing price $P_{i,t}$ is viewed as a noisy measurement of the true value of the share $x_{i,t}$; this value is combined with a previous estimated value to obtain an updated estimation. Application of Kalman filter theory (Kalman, 1960) results in an estimated value $x_{i,t}$ of each share at time t in terms of its estimation $x_{i,t-1}$ at time *t-1* and the latest share price observation $P_{i,t}$:

$$x_{i,t} = (1 - K_k)x_{i,t-1} + K_k P_{i,t} + \varepsilon_{xs,t}, \varepsilon_{xs,t} \sim N(0, \sigma^2_x) \tag{14}$$

The parameter K_k is the Kalman filter gain (Kalman, 1960). By solving the above set of equations as a Kalman filter problem, the relative sizes of the random walk and mean reverting components can be determined. Clegg and Krauss (Pairs trading with partial cointegration, 2018) solved for ρ, K_k and the coefficients of β through a maximum likelihood technique. We followed a somewhat different approach: we extracted the value of β by using the techniques of Engle-Granger (Co-Integration and Error Correction: Representation, Estimation and Testing, 1987) or Johansen (Likelihood-based Inference in Cointegrated Vector Autoregressive Models, 1995), and then optimised the trading strategy profit in terms of the values of ρ and K_k. While our solution does not yield maximum likelihood values for the unknown parameters, we observed that the selection of specific optimal values for ρ and K_k still resulted in random walk and mean reverting error components, and enabled us to significantly improve net trading returns (see section 4.4 below), as it allowed the fraction of the error signal residing in the mean-reverting components to be controlled.

Instead of trading based on the total error as in the case of normal co-integration, for partial co-integration trades are based on the mean-reverting element of the error only, with the benefit of more frequent trades and exits from a position before large losses are incurred. The value of the mean reverting coefficient ρ in effect determines the fraction of the total co-integration error that resides in the random walk and the mean reverting components respectively. The impact of ρ and K_k on the profitability of the trading strategy is discussed in section 3.6 below. In section 4.4 we describe how we optimised the values of these parameters as part of strategy optimisation.

Figure 2 displays a typical trade based on partial co-integration. The main difference with the co-integration case in Figure 1 is the fact that in the second graph the error signal now consists of a mean-reverting and a random walk component. The trade is entered as the mean-reverting component breaks upward through the upper threshold, and exits when this component returns to zero. The net return is however a result of the movements of both the mean-reverting and the random walk components. In this case the random walk component performed a large movement contrary to the movement of the mean-reverting component; as a result a loss is generated rather than a profit.

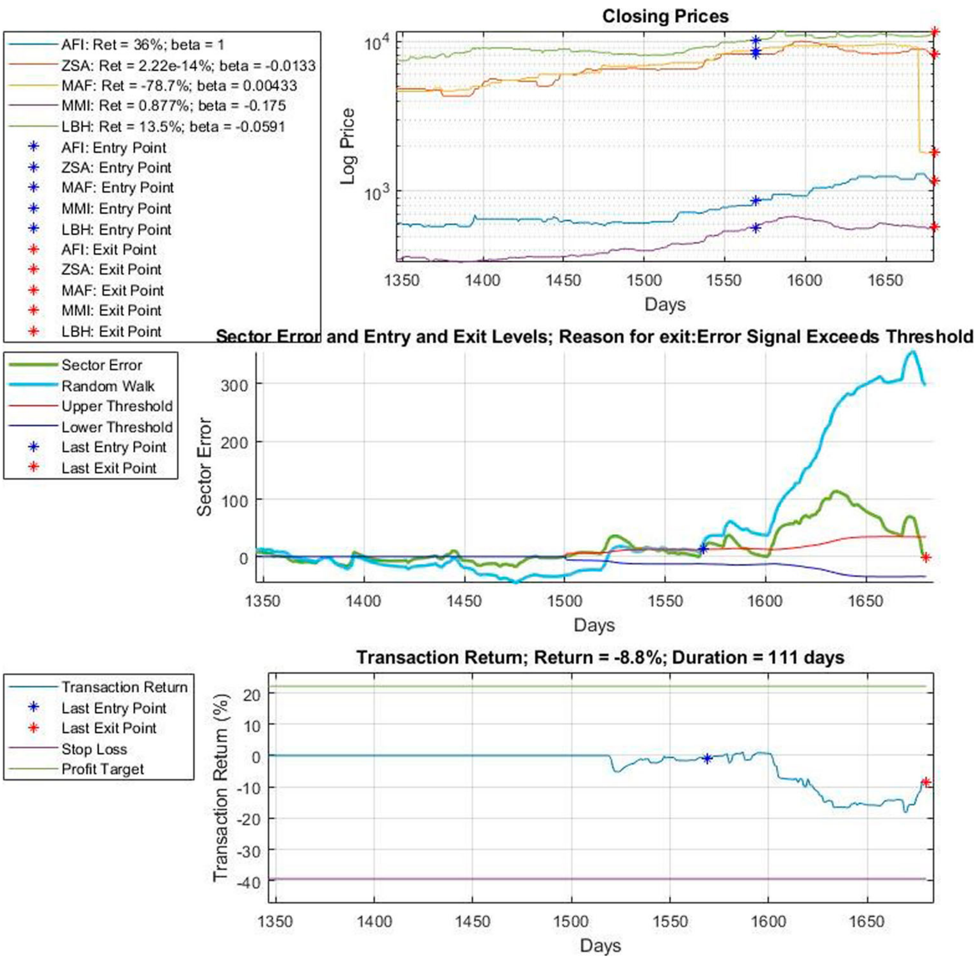


Figure 2. Illustration of a typical partial co-integration strategy trade. The top graph displays the closing prices of a group of co-integrated shares. The middle graph displays the mean reverting and random walk error signals and the trading thresholds, with entry and exit point indicated by X. The bottom graph displays the net return, stop loss and profit target.

The trading strategy was implemented by repeatedly extracting co-integrating share sets for each model sector for consecutive time periods. After each time period where trading positions were assumed, the available funds were divided equally between available trades from each of the 18 model sector strategies. This allowed maximum diversification between all the available shares. It was observed that for some model sectors very few trades were generated, while other sectors provided almost continuous trading opportunities.

In line with Mashele et al. (Pairs trading on the Johannesburg Stock Exchange, 2013) we used a trading cost equal to 0.2% of the size of the trade upon entry and exit, thus resulting in a 0.4% total trading cost per trade. As some share positions are short, we assumed that security had to be deposited equal to the value of the shares that were sold short; our results for net return on investment are therefore fairly conservative, as in practice a much smaller net amount needs to be deposited to effect such trades.

After implementing the trading strategy for the entire available date range, the annual profits (or losses) generated by the strategy were calculated, as well as the average annual returns over the entire period. We also calculated the risk adjusted profits, using the Sharpe ratio (Sharpe & Alexander, 1990). It has been reported before (Clegg & Krauss, 2018) that statistical arbitrage tends to perform better during bear compared to bull markets. For this reason, we also identified bear and bull markets for the available time period, and separately evaluated the performance of the co-integration and partial co-integration strategies for bear and bull markets.

3.6. Optimisation of trading parameters

We observed through practical experiments that several of the parameters forming part of the partial co-integration trading strategy had a significant impact on the profitability of the strategy. Prior research made little mention of this fact, creating the somewhat misleading impression that statistical arbitrage is almost guaranteed to generate profits if trading costs can be limited. We found that the following set of parameters should be optimised from time to time as the underlying behaviour of the shares and the nature of their relationships change over time:

1. The threshold at which trades are entered, defined as a multiple of error standard deviations.
2. The period over which this standard deviation is measured.
3. Time window over which the co-integration relationships are extracted.
4. The level α of statistical significance at which the co-integration tests are performed.
5. The level of the stop loss.
6. The profit target level.
7. Mean reversion coefficient ρ.
8. Kalman filter gain.
9. Mean Reversion Random Walk Ratio Limit: this is the maximum allowed value for the ratio between the standard deviation of the mean-reverting error signal and the random walk error signal.

In practice a balance must be struck between sub-optimal parameter selection, over-fitting and computational limitations. We iteratively tested the profitability of the strategy while changing one of the above parameters over a range of values that proved to provide profitable returns. The range of parameter values that maximised returns was noted in each case. This was repeated separately for a training set (the first 50% of the data) and a test set (the last 50% of the data). It was determined in each case if the optimal range of values as extracted for the training set would provide close to optimal results for the test set. This prevented overfitting of the model and served to verify how repeatable the model results would be once optimal parameter value were selected.

4. Results

4.1. Risk-adjusted performance of the partial co-integration trading strategy

The accumulated returns generated over the period under investigation by the partial co-integration strategy are compared against the market index and the co-integration trading strategy in Figure 3. For market returns we do not use the published ALSI index returns for the JSE, but rather calculate weighted average annual returns for all the shares that were considered for inclusion in the trading strategy portfolios, using market capitalisation as weights. Our version of market returns is therefore calculated in exactly the same way as our portfolio returns, to ensure that we measure the ability of the trading strategy to correctly select shares rather than the performance of specific share sets. It is clear than not only did the partial co-integration strategy far outperform the market returns and normal co-integration returns, but the returns it generated appear to be more stable over time compared to market returns.

This is confirmed by the results displayed in Table 1. The overall annualised returns of the partial co-integration strategy were 22.3% compared to market returns of 4.96%, while the Sharpe ratio, as measure of risk-adjusted returns, was 0.87 compared to 0.15. Partial co-integration therefore did not only drastically improve average returns, but also reduced the volatility of returns over a period of more than 25 years. The normal co-integration strategy could not outperform the market either in terms of average returns or risk-adjusted returns. Based on these summarised results we can affirmatively answer the first research question, whether partial co-integration produces superior results compared to normal co-integration when applied to JSE shares, and whether it outperform market returns on a risk-adjusted basis.

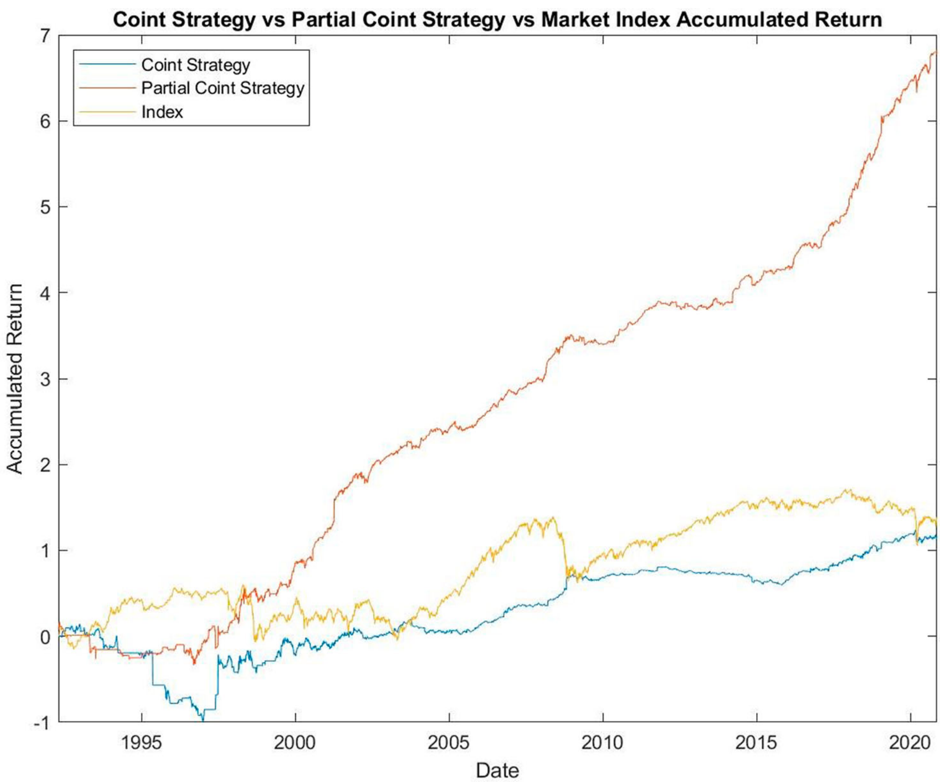


Figure 3. Comparison of the accumulated returns of the market index with the co-integration and partial co-integration strategies.

Table 1. Average annual returns during different parts of the market cycle.

	Market	Co-integration	Partial co-integration
All	4.96%	3.81%	22.32%
Low market return	−10.85%	5.78%	27.87%
High market return	19.71%	1.84%	17.15%
Bear cycles	−13.26%	12.54%	29.79%
Bull cycles	17.82%	−1.84%	17.05%
Sharpe ratio	0.15	0.09	0.87

Table 2 displays the reasons for exiting both co-integration and partial co-integration trading positions. In most cases the partial co-integration strategy exits a position because the mean-reversion error exceeded its threshold value; this shows that the trading strategy behaves as intended. In about 25% of the cases positions are exited because the profit target has been reached. For normal co-integration trades are seldomly exited because the model error exceeded the threshold, as the model error in this case tends to be a random walk rather than a mean-reverting signal. The total number of trades is also much higher for partial co-integration compared to normal co-integration.

4.2. Performance of partial co-integration during bear and bull cycles

More insight into the potential benefits of a partial co-integration trading strategy is obtained by observing the values of average annual returns over bear and bull cycles of the market respectively. To allow us to investigate this aspect Table 1 separately displays the average annual results

- for all market cycles,
- for the lower and upper halves of trading years sorted based on annual returns, and
- for bear and bull cycles.

We defined bear and bull cycles as years where market returns were respectively below or above a risk-free rate of return set at 2%. The results in this table provide a clear answer to the second research question. The partial co-integration strategy performs similarly to the market during bull and above average years but outperforms the market by approximately 40% in terms of annualised returns during bear cycles and below average years. The normal co-integration strategy outperforms the market during bear cycles and below average years but performs poorly during bull cycles and above average years.

4.3. Sector-specific behaviour of partial co-integration

Table 3 shows the list of model sectors that were constructed, the number of shares in each, the fraction of the overall trading period that each of these model sectors produced a co-integrated relationship, the number of trades over this period, the average annual profits and Sharpe ratio generated by trades in each sector and the average number of shares selected for each model sector, taking into account the fraction of the total time period that a transaction was active. While some sectors produced very high annualised returns, the Sharpe ratios of most sectors were

Table 2. Reason for exiting partial co-integration trading positions.

Number of exits / reason	No co-integration	Reached profit target	Reached stop loss	Same shares not available	Error signal exceeds threshold
Co-integration	0	395	34	13	17
Partial co-integration	0	527	10	7	1636

lower than that of the market. This is confirmed by Figure 4 that displays the accumulated returns for the individual model sectors. The individual sector arbitrage portfolio returns were quite volatile, indicating that the good results of the overall partial co-integration strategy are only made possible by the degree of diversification that is achieved by simultaneously exploiting arbitrage opportunities available in all sectors.

The number of shares included in the co-integrated sets for the various model sectors ranged from 2 to 12. Combining the available co-integrated portfolios of all sectors allowed the total portfolio to on average be invested into 12.5 shares. An arbitrage trading strategy based on pairs trading will struggle to provide the same level of diversification, as each trade will provide exposure to two shares only.

4.4. Influence of trading strategy parameters

In this section we address the 3rd research question by finding optimal values for the various parameters that define the partial co-integration trading strategy. The performance criteria that we used is the average annualised return of the strategy over the entire period for which data was available. As secondary criteria we use the average annualised returns over bear market cycles, as a statistical arbitrage strategy is expected to be of most value during such periods. Due to space

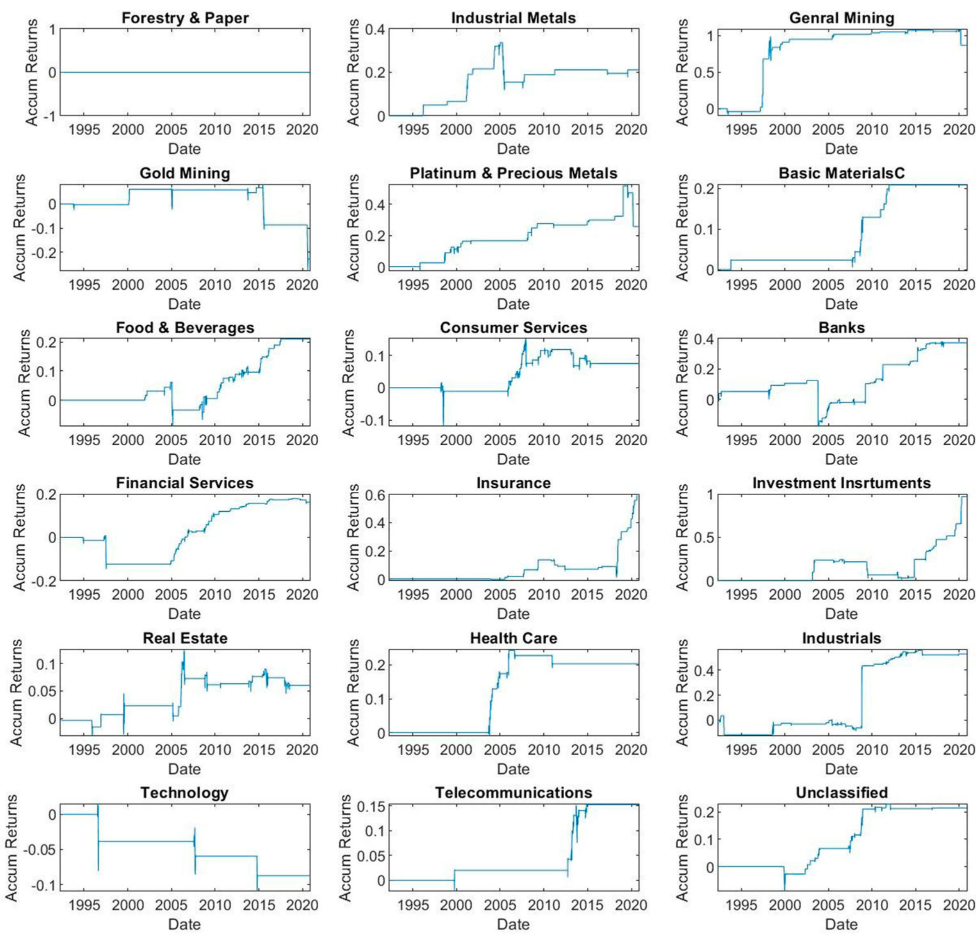
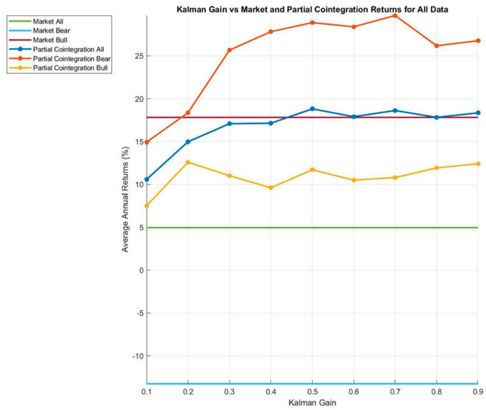
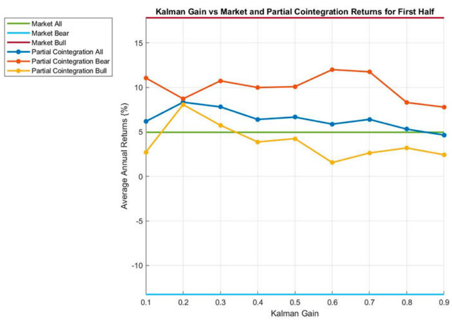


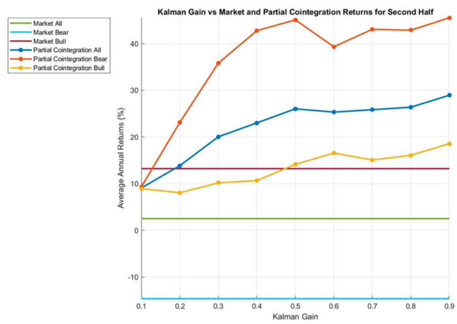
Figure 4. Accumulated returns of the partial co-integration strategy produced by different model sectors.



(a) Return measured over total time period



(b)Return measured over 1st half of data set



(c) Return measured over 2nd half of data set

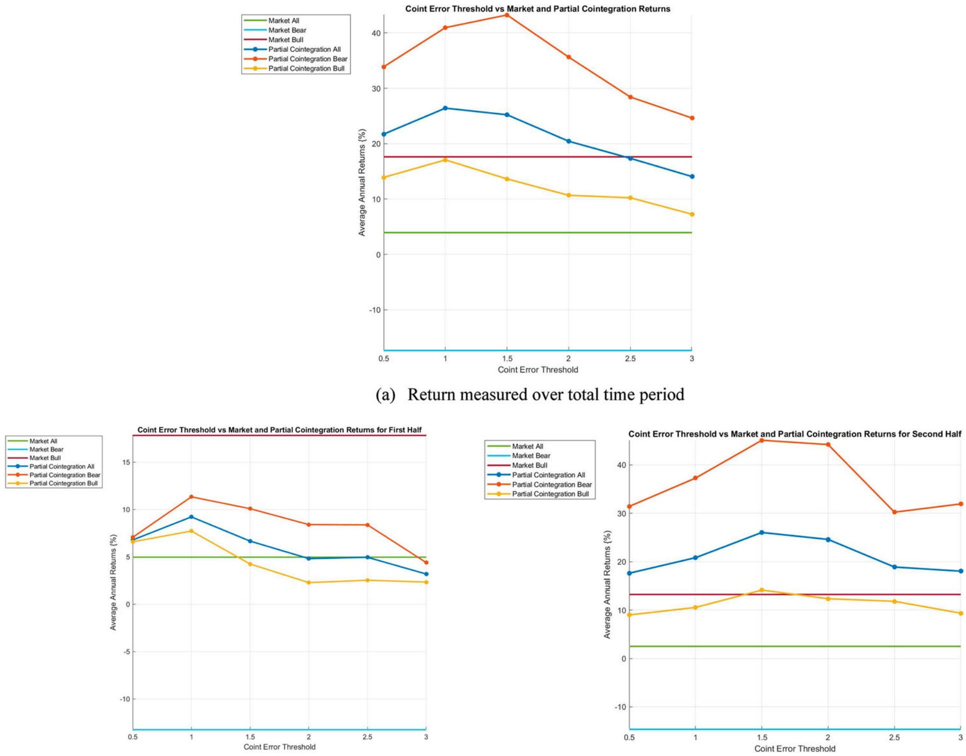
Figure 5. Partial co-integration trading strategy average annual profits over the period 1996–2020 for different values of Kalman filter gain

constraints we only show graphs of performance as function of parameter value for those parameters that had the biggest impact on trading performance.

In Figure 5 (a) it can be seen that the Kalman filter gain has a more significant impact on trading profitability than ρ, with a value of 0.7 being close to optimal for bear markets. Within the context of Kalman filtering, this means that the measured observation of closing prices must be significantly smoothed to obtain optimal trading results. In Figure 5 (b) and (c) we display the impact of Kalman filter gain on returns as extracted from the first half and second half of the data respectively. While the relationships for the two periods are not identical, application of the optimal Kalman filter gain as extracted from the first half will still produce close to optimal results when applied to the second half of the data. Model application is therefore effective out-of-sample.

The mean reversion coefficient ρ controls the fraction of the total model error that is represented by the mean reverting component vs the random walk component. We found values between 0.2 and 0.5 to be close to optimal.

As the trading strategy is dictated by the co-integration error signal levels where trades are initiated, we next investigated the impact of trading threshold on performance. The trading parameter in this case is the multiplier for mean-reverting error signal standard deviation, i.e., the value of z as reflected in equation (8) above. From the graph in Figure 6 (a) it can be seen that a value of 1.5 was close to optimal. For larger values a smaller number of trading opportunities exists, while for smaller values the full profit potential of mean-reverting swings of the error signal is not exploited. We repeated the calculations for the first and second halves of the available data set respectively; the results displayed in Figure 6 (b) and (c) shows that the optimal values for the trading threshold are very similar for both data sets.



(a) Return measured over total time period

(b) Return measured over 1st half of data set

(c) Return measured over 2nd half of data set

Figure 6. Partial co-integration trading strategy average annual profits over the period 1996–2020 for different values of trading threshold multiplier.

Figure 7 shows that the time period over which error signal standard deviation is measured is also important, with 300 days the optimal measurement period for our data set. If this period is too short the measurement becomes unreliable; if too long the measurement reflects behaviour that has already changed by the time that a trading decision is made.

Almost all previous research on co-integration-based trading used a 5% level of significance for the statistical tests used to verify co-integration of the share set. We found a level of 4% to produce slightly better results. For larger values the strategy will trade on error signals for share sets that are not reliably co-integrated, while for smaller values trading opportunities will be ignored for cases where co-integration is strong enough to provide reliable mean-reversion behaviour.

We also investigated the impact of stop loss levels. If the absolute value is too small, trades that would eventually produce a profit are too often exited after a small loss is incurred. For partial co-integration we however did not find large stop losses to have much impact, as the mean-reverting component mostly triggered trade exits before large losses were incurred.

The complete list of selectable parameter values that form part of the partial co-integration strategy are displayed in Table 4. For practical application of the technique these parameter values should be optimised from time to time for the available data set.

4.5. Repeatability of the partial co-integration trading results

The fourth research question addresses the issue whether the co-integration relationships between sets of shares on which statistical arbitrage is based are sufficiently stable to produce profitable results over future time periods when using models extracted from past time periods. Figure 3

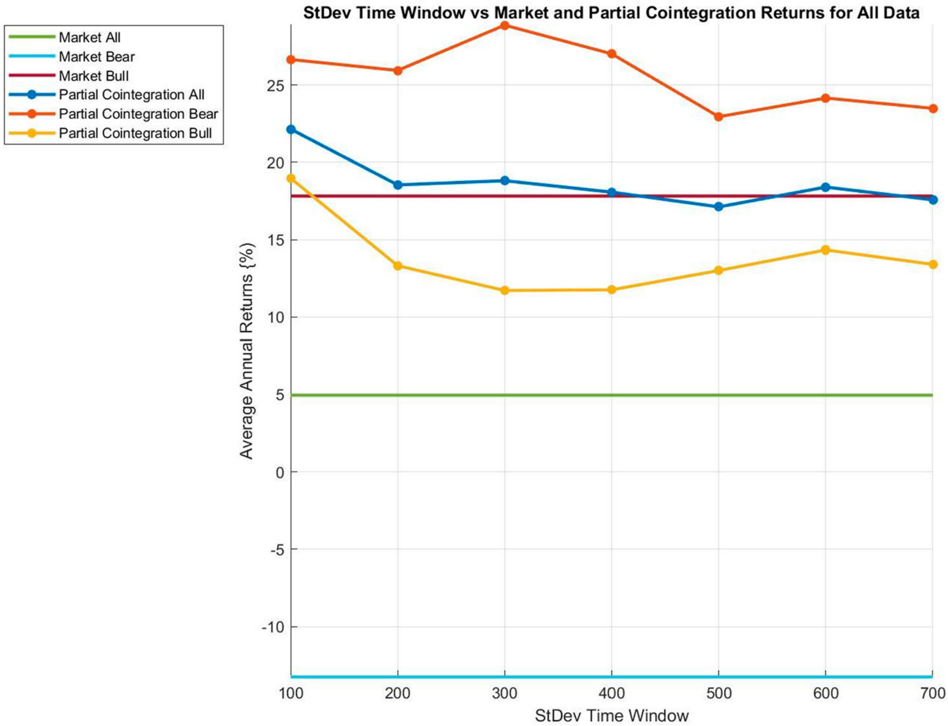


Figure 7. Partial co-integration trading strategy average annual profits over the period 1996–2020 for different values of length of time window over which error standard deviation is measured.

provides evidence that the partial co-integration trading strategy dominates market returns for most of the period under investigation, except for the first few years when the market was in a strong upward cycle. This indicates that the strong results produced by the partial co-integration trading strategy were not limited to a short period when specific conditions may have prevailed.

The possibility however still exists that the trading strategy may have been optimised for specific values of the trading parameters that are only valid for the period from which they were extracted. By studying Figure 5 (b) and (c) and Figure 6 (b) and (c) the optimal values for the two most prominent trading parameters, Kalman filter gain and Co-integration Error Threshold, can be compared for the first and second halves of the calendar period for which data was available. While the optimal values are not identical for both periods, it is evident that optimal values extracted from the first half of the data would have produced close to optimal results during the second period. We can therefore state that the relationships between the co-integrated sets of shares are sufficiently stable to ensure that the results that we obtained will be repeatable in future.

Table 4. Optimal values for parameters forming part of partial co-integration trading strategy.

Parameter	Value
Error threshold multiplier	1.5
Error standard deviation time window	300 days
Significance level of co-integration test	4%
Stop loss log return	Less than −0.2
Profit target log return	0.015
Co-integration model extraction time window	400 days
Mean reversion coefficient	0.2
Kalman gain	0.5
Mean reversion random walk ratio limit	15

4.6. Co-integration and partial co-integration trading statistics

In order to characterise the behaviour that is displayed by co-integration and partial co-integration trading strategies, we extracted a list of statistics from the results, including the following:

1. The fraction of times that trading strategies were exited for various reasons, including reaching the zero-error threshold, reaching a stop loss or profit target, loss of co-integration between the selected set of shares and exiting due to the delisting of one of the shares in a co-integrated set.
2. The number (and fraction) of long and short trades that were entered (considering that each of these may involve long and short positions in different shares).
3. The number (and fraction) of profitable and lossmaking trades that were entered, calculating these also separately for long and short positions.
4. The average profits (and losses) generated by all trades, by long and short trades and by profitable and lossmaking trades respectively.
5. The average duration of all, long, short, profitable and lossmaking trades.
6. The fraction of total trading time that one or more trades were active, and the fraction of days that the positions that were held were profitable or lossmaking.
7. The fraction of time that each set of shares (i.e., each model sector) did not produce a co-integrated relationship, thus preventing a trade from being entered.
8. The average number of shares involved in trades, and the average number of trades that the strategy is entered into at any point in time – this would provide an indication of the level of diversification offered by the strategy.

Table 3 indicates that the strategy took positions in all model sectors that had 2 or more shares, and therefore also in almost all the available shares at some stage. When using the optimal set of parameters as displayed in Table 4 each sector is involved in a trade for on average 20.6% of the time, and on average trades were active in 4.2 model sectors at a time. The average number of shares that the partial co-integration strategy invested into at any point in time was 12.48, which confirms that the strategy provides an acceptable level of risk diversification.

In Table 5 it can be seen that for normal co-integration, where trading is based on the complete error signal that contains large random walk components, the fraction of positive trades is high (in this case almost 88%) but the small fraction of loss-making trades have on average high negative return values (-15.6%) compared to the much smaller returns of the profitable trades (2.84%). As a result, the overall strategy has an average annual return of only 3.81%. The long average duration of the trades of 45 days also does not reflect the high frequency trading typical of statistical arbitrage based on truly stationary error signals.

As indicated in Table 6, for the partial co-integration strategy, where trading is based on the mean-reverting component only, the number of trades is much higher (2187 vs 465) and the average duration of the trades much shorter (4.57 days) as the mean-reverting error signal is truly stationary. While the fraction of profitable trades is much lower (65%) the average size of profitable

Table 5. Statistics extracted for a typical co-integration trading strategy.

	Tot	Profit	Loss	Long	Profit long	Loss long	Short	Profit short	Loss short
Ave ret transaction	0.64%	2.84%	−15.60%	0.69%	2.79%	−16.06%	0.59%	2.89%	−15.12%
Ave duration transaction	45.0	21.8	218.4	44.4	25.1	199.9	45.7	17.8	237.5
Number/fraction of transactions	465	0.877	0.118	0.544	0.482	0.060	0.456	0.396	0.058
Number/fraction of positions	21170	0.495	0.493	0.538	0.266	0.263	0.462	0.228	0.230

Table 6. Statistics extracted for a typical partial co-integration trading strategy.

	Tot	Profit	Loss	Long	Profit long	Loss long	Short	Profit short	Loss short
Ave ret transaction	0.32%	1.44%	−1.79%	0.26%	1.36%	−1.77%	0.38%	1.51%	−1.82%
Ave duration transaction	4.57	2.93	7.68	4.66	2.95	7.81	4.48	2.92	7.55
Number/fraction of transactions	2187	0.65	0.35	0.49	0.32	0.17	0.51	0.34	0.17
Number/fraction of positions	9999	0.53	0.46	0.50	0.26	0.23	0.50	0.27	0.22

Ave ret trans: average return per transaction
Ave dur tran prof: average duration of profitable transactions
Ave dur tran loss: average duration of loss-making transactions
Frac tran prof: fraction of profitable transactions
Ave annual ret all: average annualized return of the partial cointegration strategy for all market cycles
Ave annual ret bear: average annualized return of the partial cointegration strategy during bear cycles
Ave annual ret bull: average annualized return of the partial cointegration strategy during bull cycles

trades is now almost the same as the average size of loss-making trades (1.44% vs −1.79%); as a result, the average annual return is now an impressive 22.3%.

In Table 2 we showed the reasons for exiting a trading position and the number of times that each reason occurred when using the set of trading parameters in Table 4. Due to the short trading durations associated with partial co-integration, co-integration was almost always maintained for the period that the trades were active, and the stop loss was seldom reached. The fact that in most cases the trades exited because the zero-threshold value was reached by the error signal, provides evidence that the trading strategy functioned as intended and was close to optimal.

It can be appreciated that the optimal trading strategy, and thus parameter selection, will differ for different trading scenarios. For example, if the statistical arbitrage method is combined with a momentum strategy, then the optimal selection of parameters for statistical arbitrage should be based on the performance during bear cycles only, opposed to a situation where statistical arbitrage is implemented on its own, in which case parameter selection must be based on performance across all cycles. The effective trading cost (which is not the same for different traders) will also determine whether a larger number of transactions with smaller returns per transaction will produce higher returns compared to a smaller number of transactions with higher returns per transaction. Our results demonstrate that it is possible to optimise a statistical arbitrage strategy for such different scenarios.

5. Conclusions and recommendations

In this paper we applied a trading strategy based on partial co-integration to sector-based share sets on the JSE. We found that the well-known techniques developed by Engle and Granger (Engle & Granger, 1987) and by Johansen (Johansen, 1995) can be used to identify sets of co-integrated shares, and that the defined sectors of the JSE provide a suitable basis for selecting candidate sets. We however observed that the model error produced by standard co-integration models tends to display random walk behaviour most of the time. This results in infrequent trading, long average trade durations and a small fraction trades generating large losses that negatively impact on the performance of the trading strategy.

We modified the approach of Clegg and Krauss (Pairs trading with partial cointegration, 2018) to apply a trading strategy based on partial co-integration that uses sets of co-integrated shares, rather than pairs, selected from 18 defined model sectors. This is shown to increase the level of diversification as the selected portfolio on average includes approximately 12.5 shares at any point in time. Our approach allowed us to optimally divide the extracted model error between the mean-reverting and random walk components through selection of the mean reversion coefficient and Kalman filter gain. A partial co-integration approach with optimised parameters dramatically improves the profitability of trading returns compared to conventional co-integration-based trading. We found that partial co-integration based trading strategies, in contrast with conventional co-integration strategies, produce a larger number of trades that are shorter in duration, with a smaller fraction of profitable trades but much reduced average size of losses in loss-making trades, thus resulting in superior net performance. The optimal values for the critical model parameters proved to be sufficiently stable over time to ensure the repeatability of the results that were observed.

Our partial co-integration strategy outperformed the market by almost 20% on an annualised basis for the time period under consideration. It also outperformed the market on a risk-adjusted basis using the Sharpe ratio as performance measure. In line with previous research we found that the statistical arbitrage trading strategy performs better during bear cycles of the market than during bull cycles: while matching the market during bull cycles it outperformed the market by more than 40% during bear cycles. This makes statistical arbitrage an attractive option to combine with trading strategies like earnings momentum that are known to perform well during bull markets.

Future research will apply the partial co-integration trading strategy to share data from other exchanges and will combine it with other trading strategies that are suited to bull markets, to determine if a combined trading strategy can further improve the risk-adjusted return performance.

ORCID

A. J. Hoffman http://orcid.org/0000-0003-4909-1073

References

Appelbaum, M. (2015). *Does pairs trading work on the Johannesburg Stock Exchange?* Master's thesis, University of Cape Town, South Africa.

Avellaneda, M., & Lee, J.-H. 2008. Statistical arbitrage in the U.S. equities market. [Online]. Retrieved from https://www.math.nyu.edu/faculty/avellane/AvellanedaLeeStatArb071108.pdf

Baviera, R., & Baldi, T. (2019). Stop-loss and leverage in optimal statistical arbitrage with an application to energy market. *Energy Economics*, *79*, 130–143. https://doi.org/10.1016/j.eneco.2018.03.024

Bhana, N. (1995/96). The share market reaction to earnings announcements: A test of the efficiency of the Johannesburg Stock Exchange. *Investment Analysts Journal*, Summer, Issue 42, 45–57.

Bhave, A., & Libertini, N. (2013). *A study of short term mean reversion in equities.* [Online]. Retrieved from http://www.361capital.com/wp-content/uploads/2013/10/361_Capital_Study_Short_Term_Mean_Reversion_Equities.pdf

Caldeira, J., & Moura, G. (2013). Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. *Brazilian Review of Finance*, *11*(1), 49–80. https://doi.org/10.12660/rbfin.v11n1.2013.4785

Chan, E. P. (2013). Algorithmic trading: Winning strategies and their rationale. John Wiley & Sons. https://doi.org/10.1002/9781118676998

Clegg, M., & Krauss, C. (2018). Pairs trading with partial cointegration. *Quantitative Finance*, *18*(1), 121–138. https://doi.org/10.1080/14697688.2017.1370122

Frey, B. J., & Dueck, D. (2007). Clustering by passing messages between data points. *Science*, *315*(5814), 972–976. https://doi.org/10.1126/science.1136800 PMID:17218491

Engle, R. F., & Granger, C. (1987). Co-Integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, *55*(2), 251–276. https://doi.org/10.2307/1913236

Focardi, S., Fabozzi, F., & Mitov, I. (2016). A new approach to statistical arbitrage: Strategies based on dynamic factor models of prices and their performance. *Journal of Banking & Finance*, *65*, 134–155. https://doi.org/10.1016/j.jbankfin.2015.10.005

Fuller, W. A. (1976). Introduction to Statistical Time Series. John Wiley and Sons.

Gatev, E., Goetzmann, W., & Rouwenhorst, K. (2006). Pairs Trading: Performance of a Relative Value Arbitrage Rule. *Review of Financial Studies*, *19*(3), 797–827. https://doi.org/10.1093/rfs/hhj020

Harlacher, M. (2012). Cointegration based statistical arbitrage. [Online]. Retrieved from https://stat.ethz.ch/research/mas_theses/2012/harlacher.pdf

Hogan, S., Jarrow, R., Teo, M., & Warachka, M. (2004). Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial Economics*, *73*, 525–565. https://doi.org/10.1016/j.jfineco.2003.10.004

Jarrow, R. A., Li, H., & Ye, X. (2016). Exploring statistical arbitrage opportunities in the term structure of CDS spreads. [Online]. Retrieved from http://ssrn.com/abstract=2686284

Johansen, S. (1995). Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Clarendon Press. https://doi.org/10.1093/0198774508.001.0001

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, *82*(1), 35–45. https://doi.org/10.1115/1.3662552

Kanamura, T., Rachev, S. T., & Fabozzi, F. J. 2008. *The application of pairs trading to energy futures markets.* [Online]. Retrieved from http://statistik.ets.kit.edu/download/doc_secure1/ptem342v144techreport.pdf

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, *259*, 689–702. https://doi.org/10.1016/j.ejor.2016.10.031

Mashele, H., Terblanche, S., & Venter, J. (2013). Pairs trading on the Johannesburg Stock Exchange. *Investment Analysts Journal*, *42*(78), 13–26. https://doi.org/10.1080/10293523.2013.11082559

Miao, G. J. (2014). High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. *International Journal of Economics and Finance*, *6*(3), 96–110. https://doi.org/10.5539/ijef.v6n3p96

Phillips, P. & Ouliaris, S. (1990). Asymptotic properties of residual based tests for cointegration. *Econometrica*, *58*(1), 165–193.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economics*, *13*(3), 341–360.

Rudy, J. (2011). *Four essays in statistical arbitrage in equity markets.* [Online]. Retrieved from http://www.harvestalpha.com/wp-content/uploads/2013/08/phd.pdf

Sánchez-Granero, M., Balladares, K., Ramos-Requena, J., & Trinidad-Segovia, J. E. (2020). Testing the efficient market hypothesis in Latin American stock markets. *Physica A*, *540*, 123082. https://doi.org/10.1016/j.physa.2019.123082

Sharpe, W., & Alexander, G. (1990). Investments. Prentice Hall.

Sjo, B. (2008). Testing for unit roots and cointegration. [Online]. Retrieved from https://www.iei.liu.se/nek/ekonometrisk-teori-7-5-hp-730a07/labbar/1.233753/dfdistab7b.pdf

Stübinger, J. (2019). Statistical arbitrage with optimal causal paths on high-frequency data of the S&P 500. *Quantitative Finance*, *19*(6), 921–935. https://doi.org/10.1080/14697688.2018.1537503

Thorp, E. (2000). A Mathematician on Wall Street. [Online]. Retrieved from http://www.wilmott.com/pdfs/080617_thorp.pdf

Thorp, E. (2003). A Perspective on quantitative finance: Models for beating the market. [Online]. Retrieved from http://media.wiley.com/product_data/excerpt/11/04700235/0470023511.pdf

Vidyamurthy, G. (2004). *Pairs trading: Quantitative methods and analysis.* s.l. Wiley.

Visagie, G., & Hoffman, A. (2017). Comparison of statistical arbitrage in developed and emerging markets. *International Journal of Trade, Economics and Finance*, *8*(2), 67–72. 541-EF0012.pdf (ijtef.org)