

https://dspace.org/

Enhanced indexation of South African equities: a co-integration analysis

Modungwa, Dineo Ntshadi

2017

https://hdl.handle.net/10539/26908

Modungwa, Dineo Ntshadi (2017) Enhanced indexation of South African equities: a co-integration analysis, University of the Witwatersrand, Johannesburg, <http://hdl.handle.net/10539/26908>

Downloaded from a DSpace repository. https://dspace.org/

Enhanced indexation of South African equities: a co-integration analysis

A research submitted by

D Modungwa

Student Number: 765027

Supervisor:

Dr. Thanti Mthanti

Wits Business School

September 2017

ABSTRACT

This study presents an analysis of cointegration based tracking strategies; a classical index tracking strategy and a long-short equity market neutral tracking strategy. The first strategy attempted to track a reconstructed index, whilst the second strategy attempted to track an enhanced index. Quantitative portfolio managers have traditionally used the concept of correlation to analyse co-movements, which over years evolved into conditional correlation. In contrast, this study applied cointegration rather than correlation in portfolio optimisation. The concept of cointegration relies on the long-run relationship between time series i.e. stock prices and an index.

The data that was used was the price history of the JSE Top 40 market index and its constituent stocks for the period 02/11/2009 – 31/12/2015. The study found that it was only the designed tracking portfolio of 10 stocks that was sufficiently cointegrated with the index (JSE Top 40), albeit at higher volatility at various times in the out of sample period and performance that slightly lagged the index even before accounting for transaction costs. As far as the enhanced index tracking strategy based on cointegration, the study found no convincing empirical evidence for the South African equities market.

Further research should be done to test the robustness of the cointegration based index tracking strategies under different market conditions and its applicability within sophisticated trading strategies and stock selection methodologies.

DECLARATION

I, Dineo Modungwa, declare that this research report is my own work except as indicated in the references and acknowledgments. This thesis is submitted in fulfilment of the requirements for the degree of Master of Management in Finance and Investment with the University of Witwatersrand. I further declare that I have not submitted the work before for any degree or examination, either in its original or similar form, anywhere else.

Dineo Modungwa

Signed at ...Johannesburg....On the....07....Day of September 2017

ACKNOWLEDGEMENTS

I wish to extend my deep and sincere gratitude to:

- God, the Almighty, for guiding and strengthening me through this research process.
- Dr. Thanti Mthanti, my supervisor, for the guidance and Chris Holdsworth, for the additional technical guidance.
- My parents, siblings and my family at large for their prayers, love, patience and support.
- My spiritual leader and my family at the House of Prayer for their love, patience and prayers.
- My employer for the financial support.
- My friends and colleagues for their support.

TABLE OF CONTENTS

ABSTRACT i

DECLARATION ii

ACKNOWLEDGEMENTS iii

LIST OF TABLES vi

LIST OF FIGURES vi

1 CHAPTER 1: INTRODUCTION & BACKGROUND 1

1.1 Context of the study 1
1.1.1 Optimal Index Tracking 3
1.1.2 Cointegration vs. Correlation 4
1.1.3 Equity market neutral strategies for enhanced returns 5

1.2 Problem Statement 7

1.3 Research Questions 8

1.4 Purpose of Study 8

1.5 Data Collection 9

1.6 Overview of Research Methodology 9
1.6.1 The construction of portfolios 9
1.6.2 The rationale for constructing tracking portfolios based on cointegration relationships 9
1.6.3 Stock Selection 11
1.6.4 Tracking Portfolio Constituents' weights 11
1.6.5 Engel-Granger cointegration testing 12
1.6.6 Long-short Equity Market Neutral Strategy 12

2 CHAPTER 2: LITERATURE REVIEW 14

2.1 Enhanced Indexation 14

2.2 Cointegration in portfolio management 16

3 CHAPTER 3: RESEARCH METHODOLOGY AND INVESTMENT STRATEGIES 20

3.1 Concept of cointegration and its principles 20

3.2 The concept of cointegration justification 20
3.2.1 Cointegration equation estimation and testing 21

3.3 Background on the Engle-Granger Methodologies 21
3.3.1 Overview on Engel-Granger Approach to Estimation of Cointegration Systems 21
3.3.2 Overview on Engel-Granger Approach to testing for Cointegration 21

3.4 Unit Root Test 22

3.5 Portfolio Construction Strategies Based on Cointegration 22
3.5.1 Classic Index Tracking Strategy 22
3.5.2 Portfolio Construction 22
3.5.3 Back-Testing Procedure 24
3.5.4 Enhanced Indexation Strategy 28

4 CHAPTER 4: CO-INTEGRATION BASED PORTFOLIOS USING JSE TOP 40 EQUITY INDEX 31

4.1 Data 31

4.2 Unit Root Testing 32

4.3 Back-testing the Classic Index Tracking Strategy *34*
4.3.1 Rolling window EG co-integration test 34
4.3.2 Differential returns between the tracking strategy and the index 38
4.3.3 Sharpe Ratio and Information Ratio 39
4.3.4 Transaction costs 40
4.3.5 Volatility 40
4.3.6 Correlation 41
4.4 Back-testing the Enhanced Index Tracking Strategy *42*
4.4.1 Rolling window EG co-integration test 42
5 CHAPTER 5: CONCLUSIONS AND PROSPECTS FOR FUTURE RESEARCH 46
REFERENCES 48
APPENDIX 1: JSE TOP 40 CONSTITUENTS 52
APPENDIX 2 RETURNS 54

LIST OF TABLES

Table 1: Results of the ADF test....33
Table 2: Descriptive Statistics of Reconstructed Index and Tracking Portfolio....38

LIST OF FIGURES

Figure 1: EG Statistic of EG 10 Stocks....35
Figure 2: EG Statistic of EG 5 Stocks....36
Figure 3: EG Statistic of EG 20 Stocks....36
Figure 4: EG Statistic of EG 25 Stocks....36
Figure 5: EG Statistic of EG 30 Stocks....37
Figure 6: Investment Growth....39
Figure 7: Annualised Volatility....40
Figure 8: Graph showing correlations....41
Figure 9: EG Statistic of EG 10 Stocks portfolio ±5%....42
Figure 10: EG Statistic of EG 15 Stocks Portfolio ±5%....43
Figure 11: EG Statistic of EG Stocks Portfolio ±5%....43
Figure 12: EG Statistics of 25 EG Stocks Portfolio ±5%....44
Figure 13: EG Statistic of EG 30 Stocks Portfolio ±5%....45

1 CHAPTER 1: INTRODUCTION & BACKGROUND

1.1 Context of the study

The age-long debate about the best way to realise high returns while keeping down the cost of investments continues to occupy the investment industry worldwide. For some time now there have been two opposite extremes to this debate. The first argument refers to the considerable empirical evidence that active funds do not earn significant returns in excess of the comparable indices. For instance, Sharpe (1991) asserted that, on average, active fund managers cannot better the returns obtained from passive management strategies. The basis of the assertion is that the performance of the index equals the weighted average of both active and passive investors before expenses are deducted. He concluded that this showed that active management is merely a zero-sum game.

The second argument, holds that it is possible to find periods of dominance of active funds over index funds, depending on the time horizon of the data (Minor, 2001).

It is also important to understand the precepts informing each of these investment strategies. Active strategies are founded on the notion that markets are not efficient, thus through information advantage, investment strategies can be devised to generate excess profits. In contrast passive strategies firmly believe in market efficiency; a proposition that current stock prices reflect all available information about companies and thus no excess profits can be generated through information advantage. Evidence on market efficiency studies in developing markets have found the Johannesburg Stock Exchange to be weak-form efficient and the rest of the African Stock Markets to be inefficient (Mazviona & Nyangara, 2013). Implications of such findings suggest that South Africa is indeed a consideration candidate for passive investment options.

South Africa has also not escaped this debate. Despite this, there is still no firm conclusion on the superiority in performance over long periods of either strategy.

While this debate continues, the writer believes the growing attraction to passive investment strategies cannot be ignored. The PWC ETF 2020 report (2014) disclosed that the Exchange Traded Funds (ETFs) are holding assets of approximately $2,6 trillion under management

globally, with more than 5400 products listed in 60 exchanges. This rapid rise can be largely attributed to a growing acceptance of indexing.

The growth in passive investments has not only been a global phenomenon. South Africa's domestic passive market has, in line with global trends, also experienced exponential growth. According to etfSA The Home of Exchange Traded Funds (March 2015), there were approximately 70 ETPs listed on the JSE with a R126,6 billion market capitalisation (assets under management).

Passive strategies have been implemented through index tracking, a methodology summarised by Brink (2004, p. 24) as follows:

> Indexing is the structuring of a passively managed portfolio of stocks or bonds that seeks to replicate the returns of market indices such as the All Share Index. The index fund usually buys shares in securities included in a particular index, in proportion to the securities representation in that index. Index fund portfolios, provide a way for investors to closely match the performance of an index at a minimal cost. Although very low, all index funds underperform against the index they track by an amount equal to their annual expense charge.

Over the course of time, as equity indexation became increasingly widespread as a form of passive investing. A new challenge presented itself to the investment industry: investors discontentment with less than market equivalent returns offered through indexation. Investors are in constant pursuit of excess risk adjusted returns at minimal costs. This is not surprising, as it is considered to be human nature to always desire excess return.

The ever-evolving unrestrictive investment environment, which suggests no strict adherence to any one extreme of the broader pure passive or active investment strategies, saw the advent of enhanced indexation (also referred to as Smart Beta) as an extension to index tracking. Enhanced indexation is a hybrid strategy that combines active and passive investment, with the aim of outperforming an index by generating "excess" returns (diBartolomeo, 2000); Scowcroft and Sefton, 2003; Canakgoz and Beasley, 2008).

Empirical research has also further exacerbated the discussion by continuing to prove that active strategies on average only deliver marginal outperformance of the market over time.

Considering that we are also in an environment where money managers are increasingly coming under pressure and scrutiny to justify the fees they charge further fuels the active versus passive debate.

Transaction fees reduction thus emerges as another area requiring urgent concerted efforts by money managers. This is a challenge not just for South Africa but for the rest of the African Stock Markets that have been found to face a problem of high transaction costs per dollar invested (Mazviona & Nyangara, 2013). Efforts are underway to explore various methodologies to deliver excess market returns at minimal costs. Passive strategies such as Indexation remain one possible area of focus.

Brink (2004) also recognised that, the successful management of index funds relies on constant re-balancing to bring the constituent parts in line with the benchmark index. However, practical constraints, such as cash flows, transaction costs, liquidity differences among stocks and short-term market inefficiency, can inhibit a fund manager's ability to perfectly track an index through full replication. These disadvantages that have also been observed by Acosta-Gonzalez, Armas-Herrera and Fernandez-Rodriguez (2015), who suggest the disadvantages that are inherent in full replication can possibly be mitigated through investment in only a subset of the index constituents, referred to as partial replication. It is for these reasons that this study is particularly interested in pursuing a strategy that employs partial replication and modest rebalancing policies.

1.1.1 Optimal Index Tracking

An indexation strategy entails the construction of a portfolio, consisting of the index constituents, that tracks an index such that the tracking portfolio and the index move together in the long-run. It follows that the index tracking problem requires the determination and construction of a tracking portfolio such that there exists some co-movement relationship between the portfolio's price process and the index's price process. An index is after all merely a linear combination of its constituents' stock prices, and it is therefore justified to expect a replicating portfolio and an index to share a common stochastic trend.

Quantitative portfolio managers have traditionally used the concept of correlation to analyse co-movements, which over years evolved into conditional correlation. The works of pioneers

such as Markowitz (1959), Sharpe (1964), Lintner (1965), Black (1972) and the JP Morgan (1994) of RiskMetrics approaches in considering co-movements in financial instrument returns influenced this universal adoption of correlation analysis of asset returns, as a primary statistical tool for traditional portfolio optimisation. Correlation is however, time varying, which was affirmed by Taleb (2007) who posed that testing for correlation between the same time series over consecutive time periods may yield completely different results.

1.1.2 Cointegration vs. Correlation

The challenge with index tracking portfolios that are constructed on the basis of returns analysis such as correlation is the need for frequent rebalancing in order to retain co-movement of the portfolio with the index being tracked. This increases trading costs. However, if the portfolio allocations are designed such that the portfolio tracks an index based on cointegration in the short run the portfolio might deviate from the index but the two should still be tied together in the longer run. Optimal cointegration portfolios, as they rely on the long run trends between asset prices, should therefore not require as much rebalancing (Dunis & Ho, 2005).

Cointegration, initially introduced by Granger (1981), is an econometric concept which mimics the existence of a long-run equilibrium among economic time series (Ssekuma, 2011). When cointegration exits, stable co-relationships between groups of assets will remain stable for some time as a result of prevailing market factors.

Quantitative optimisation models based on correlation measures have been found to suffer several other drawbacks, especially when applied to a passive investment framework. Limitations are generated by the very nature of correlation as a measure of co-dependency; it is only applicable to stationary variables, such as stock returns. Return analysis is based on low autocorrelated market information, resulting in less stable portfolios having little long term predictive value, and limited trend information, while prices are highly autocorrelated, sharing long-term trends. As such correlation is only a short-term statistic and lacks stability. Depending on the model used to estimate it, correlation can be very sensitive to the presence of outliers and volatility clustering in data sample, which somehow limits the amount of historical data used and can thus lead to erroneous and subjective conclusions about the

nature of the long-term dependencies in the time series (Alexander & Dimitriu, 2002, 2004a). These short comings led to the pursuit of alternative time series analysis models that saw the displacement of the general acceptance of correlation.

In search of alternative quantitative models specifically designed to suit a passive investment framework, Alexandra (1999, 2003, 2004a) proposed a portfolio optimisation technique based on the cointegration of asset prices. An evident feature of Cointegration in these studies is that the price difference between the index and the tracking portfolio is, by construction stationary, proving mean reversion in price spreads; this model thus proved to produce stable portfolios that have strong relationships with the market index. Cointegration, as Maurer (2008) asserted, facilitates the determination of long term relationships between a tracking portfolio and the index, allowing for tracking portfolios stock weights to be determined on a long term equilibrium relationship with an index, rather than on short run association of returns. Due to portfolio weights that are based on the long run behaviour of stock prices, this result requires relatively less rebalancing, leading to reduced transaction costs (which are detractors to return). A key element of this study is the assessment of the potential benefits of less frequent portfolio rebalancing, as a result of the application of cointegration.

1.1.3 Equity market neutral strategies for enhanced returns

The Enhanced Indexation strategy with its prospect of excess risk adjusted returns at relatively reduced costs has certainly offered an expansion to investment options and a compelling alternative to practitioners and investors alike. Enhanced indexation through the application of cointegration will be the strategy of particular focus in this research.

Enhanced indexing essentially involves two basic goals in relation to both active and passive management: to outperform the index while maintaining some characteristics of the index. In order to outperform the index, a robust and stable alpha is required that is consistent over different investment cycles and sustainable over time. It must be noted that the second objective, risk containment, naturally restrains the first objective to a modest level of excess returns, especially when compared to traditional active strategies.

The attempt to achieve excess risk adjusted index returns has indeed been a much sought after end in investments. To this end, various approaches have been continuously tried and tested, some more complex than others, and among these are investment strategies seeking to profit from both increasing and decreasing asset prices in markets, while limiting general market exposure through a combination of long and short positions. These techniques, typically employed by hedge funds, are referred to as Market Neutral Strategies. Market neutral strategies have roots that date back to 1949, a time that marked the operation of the first hedged fund by Alfred Winslow Jones. A market neutral hedge strategy takes long and short positions in such a way that the impact of the overall market is minimized. Contrary to a long/short equity strategy, the market neutral strategy does not hold a long bias. It typically aims to neutralize the market impact by either employing a dollar neutral or beta neutral positioning (Furelid, 2015).

A long-short equity market neutral strategy consists of buying stocks (going long) that are expected to appreciate in value, and selling stocks expected to depreciate in value (going short). The goal is to neutralize exposure to the equity market and capture positive returns, regardless of the market's direction. The strategy strives to provide positive returns in both bull and bear markets. The spread between the performance of the longs and shorts, and the interest earned from short rebate, provides the value added of this investment strategy which seeks to provide a return in excess of the risk-free rate (AIMA, 2006).

In the market neutral strategy, portfolios will be constructed which are proven to have a long-run equilibrium relationship with an enhanced index, similar to Alexander and Dimitriu (2005a,b). Cointegration will be used to produce portfolios that track "plus" and "minus" benchmarks (i.e. enhanced index tracking), and then construct a self-financing strategy by being short (the portfolio tracking the "minus" benchmark) and being long (the portfolio tracking the "plus" benchmark).

Although cointegration has been applied to exploit long-run equilibrium relationships between stock prices and indexes in several markets; like the Dow Jones Industrial Average, Dow Jones EUROStoxx50, FTSE 100, CAC 40, S&P 100 and Bovespa indexes, no such studies have been done in the South African context. Considering the cited notable benefits and attractive features presented by cointegration as a method of portfolio construction, it is in the interest of this research to explore cointegration within portfolio management in the

South African equities space. This research seeks to empirically test the technique in the South African emerging market context, in order to detect and characterize the co-movement and long run equilibrium relationships (Maurer, 2008) between a JSE Top40 market index and constituent stock prices, for the purposes of devising and constructing classical and enhanced indexation portfolios. The outcome of the study will validate the applicability of cointegration to enhanced indexation, as a portfolio optimization tool in asset management in South Africa.

1.2 Problem Statement

As Blasdell and Zingone (2010, p. 1) so appositely states, "In a great debate between advocates of active and passive investment strategies, there is in fact, some middle ground; enhanced indexation. These strategies target the "sweet spot" of portfolio return/risk efficiency - between purely passive strategies (which strive to replicate, but not exceed benchmark performance) and fully active strategies which target higher, and generally more volatile above benchmark returns."

The rather ambitious quest for market plus premium returns at minimal costs still haunts the investment industry. The question arises: beyond comparing the performance of index-based funds to actively managed funds, is there a possibility to further enhance the index-based strategy to deliver excess risk-adjusted returns at reduced costs? This is to be achieved while fairly remunerating money managers for their expertise.

Roman, Mitra and Zverovich (2013) noted that although enhanced indexation was formulated as early as 2000, a review of the works of Canakgoz and Beasley (2008) revealed that only a few enhanced indexation methods were proposed in the research community. To date there is still no generally accepted portfolio construction method in this field. In addition, the study acknowledged that enhanced indexation methods researched were found to be mostly concentrated on overcoming the computational difficulty. This was raised by restricting the cardinality of the portfolios without establishing whether these portfolios actually obtained returns in excess of the index.

This research is based on studies by Alexander (1999), Alexander and Dimitriu (2002, 2004a), and Maurer (2008), proposing and analysing the applicability of cointegration techniques beyond index tracking to enhanced indexation strategies. Related empirical

studies in this area have not managed to fully demonstrate cohesive and convincing evidence of the impact and applicability of such an approach, specific to enhanced indexation. This is thus an area that still requires further justification. Further research is therefore required.

In order to further contribute to literature that validates the applicability of the cointegration technique to enhanced indexation, this study also aimed to establish and explain why and how cointegration works, if at all, when applied to enhanced indexation within the South African emerging market equity context. The study sought to further demonstrate that enhanced indexation techniques (at least those based on cointegration) can indeed be used to obtain returns in excess of the index, at reduced trading costs.

1.3 Research Questions

- How and when does cointegration work in the construction of enhanced indexation portfolios?
- Can tracking portfolios still truly replicate a market index by investing only in a subset of the constituents of the market index, both in the short term & long term?
- How do enhanced indexation portfolios based on cointegration behave under different market conditions within the South African context?
- Can excess return after transaction costs truly be generated for South African investors through enhanced indexation?
- Can a cointegrated enhanced indexation portfolio generate excess risk adjusted returns at reduced costs?

1.4 Purpose of Study

This paper attempted to contribute to empirical research by examining the long-run and short-run dynamic relationships between the market index and stock prices within the South African stock market, and how these can be explored to achieve the following:

- Constructing well diversified tracking portfolios comprising only a subset of JSE Top 40 Index constituents, using cointegration,
- Test and analyse whether:

- the constructed enhanced indexation portfolio offers market index excess risk adjusted returns, and

- this portfolio optimization approach results in reduced number of rebalancing and reduced trading costs within index tracking investment strategies

The results of the study attempted to propose a possible effective investment solution to the dilemma of delivering market index excess returns, at reduced costs to the seemingly discontent South African investors. This will possibly result in an additional portfolio construction tool for practitioners in meeting their client needs.

Even when an enhanced passive strategy is an adopted investment strategy of choice, money managers still need to identify well-tested alternative quantitative approaches that yield high probability of achieving their set investment objectives.

This study also sought to examine and reveal the short-run adjustment behaviour of the long-run relationships of stock price systems and market index, as well as when possible shocks occur.

1.5 Data Collection

The data used for purposes of this study covered 6 years' time series price history of the JSE Top 40 market index and its constituent stock prices. This data covered the period 2009 - 2015. This information can be obtained from Bloomberg.

1.6 Overview of Research Methodology

1.6.1 The construction of portfolios

Cointegration portfolios were constructed to track both a benchmark and an enhanced benchmark that are artificially created.

1.6.2 The rationale for constructing tracking portfolios based on cointegration relationships

This study draws from the works of Alexander (1999), Alexander and Dimitriu (2002, 2004a), Maurer (2008) and Dunis and Ho (2005). The focus of this paper is the construction of classical indexing as well as enhanced indexation portfolios based on a cointegration model, which allows the replication of various benchmarks, with different numbers of stocks,

an approach pioneered by Alexander (1999). If two or more series are cointegrated, in intuitive terms this implies that they have a long run equilibrium relationship that they may deviate from in the short run, but which will always be returned to in the long-run. Cointegration can be used to successfully model both the long and short term dynamics of a system.

Ssekuma (2011) cited that the advantage of cointegration, in distinction to correlation as a measure of dependency, is that it is a possible solution to non-stationarity found in many economic time series (i.e. equity prices). The challenge with nonstationary time series is that the assumptions upon which ordinary least squares (OLS) estimation rest, are violated, rendering its application inappropriate for analysis of such time series. Cointegration analysis allows nonstationary data such as stock prices to be used so that spurious and misleading results are avoided and provides finance practitioners and academics an effective formal framework for testing and estimating long-run models from actual time series data (Utkulu, 2012).

Alexandra (1999) explained that the rationale for constructing portfolios based on a cointegration relationship with the market index rather than correlation rests on the following features of cointegrated systems:

- The price difference between the index and the portfolio is by construction stationary, which implies that the tracking portfolio will be "tied" to the index in the long run.
- The stock weights, being based on the history of prices rather than returns, have enhanced stability.
- There is a full use of the information contained in level variables such as stock prices, avoids the loss of valuable information in analysis.

Despite its limitations, correlation in the portfolio construction process should not be completely discarded. A valid point, also raised by Alexander (2002, p. 4), was that, the limitations of correlation analysis for long-run inferences do not impede in any way the use of correlation for short-term relationship analysis. Correlation may be used to complement cointegration; "for instance, short-run correlation may be used as a stock selection technique, which is followed by portfolio optimization using cointegration."

There are several econometric approaches to estimate and test for cointegration, among them are the Engle-Granger procedure and Johansen multivariate framework, introduced by Engle and Granger (1987) and Johansen (1991) respectively. This study will apply the single equation methodology called Engle-Granger's (1987) two step estimation approach. Although this method only allows finding one cointegration vector even though many more may exist, it will suffice for the purposes of this study. The main advantage of the Engle-Granger approach is its simplicity and its intuitive interpretability. In a similar view to Gabriel (2002), this study argues that for portfolio management purposes the variance minimizing criterion is far more important than Johnsen's (1991) added criterion of maximising stationarity.

For both our index and enhanced index tracking processes this study followed two steps; first, the selection of stocks to be included in the tracking portfolio and second, determining the portfolio weights in each stock based on a cointegrated optimisation technique.

1.6.3 Stock Selection

For the purposes of this study, a simple stock selection criteria available based on the price ranking of the stocks in the index at the moment of portfolio construction was applied, as applied in previous studies by Alexander and Dimitriu (2002, 2004b), Maurer (2008) and Dunis and Ho (2005).

1.6.4 Tracking Portfolio Constituents' weights

The weights of the above selected stocks within the tracking portfolios will be estimated based on ordinary least squares (OLS) coefficients of a cointegration equation that regresses the index log price on tracking portfolio stocks log prices as independent variables, over a specified sample (calibration) period. The cointegration index tracking regression is given by:

$$\ln(I)_t = C_1 + \sum_{k=1}^{n} C_{K+1} \ln(P_{k,t}) + E_t \qquad (1)$$

Where I_t is the JSE Top 40 Index, C_1 is a constant, C_{K+1} is the approximate weight of stock k in the tracking portfolio, $P_{k,t}$ is the log price of the constituent stock P_k at time t, and E_t is the stationary spread between the index and the tracking strategy. If the series I_t and $P_{k,t}$ are cointegrated, it therefore means that the residual series i.e. the tracking error, E_t is stationary.

The application of OLS to non-stationary variables as in the above model is however, only valid in the special case of the existence of a cointegration relationship, that is if and only if the residuals in the above model are stationary. Testing for cointegration, is therefore an essential part in constructing cointegration-optimal tracking portfolios.

1.6.5 Engel-Granger cointegration testing

To test for cointegration, this study subjected the residuals of the above cointegration equation to the augmented Dickey-Fuller (ADF) (1997) test. The regression model estimated on the residuals of the cointegration regression (1) is:

$$\Delta\widehat{\varepsilon_t} = \gamma\widehat{\varepsilon_{t-1}} + \sum_{i=1}^{p} \alpha_i \Delta\widehat{\varepsilon_{t-i}} + u_t \qquad (2)$$

Where Δ denotes the difference operator, $\widehat{\varepsilon_t}$ is the estimated residuals from the cointegration regression 1, γ is the coefficient of $\widehat{\varepsilon_{t-1}}$, p is the lag order of the autoregressive process, $\Delta\widehat{\varepsilon_t}$ = $\widehat{\varepsilon_t}$ - $\widehat{\varepsilon_{t-1}}$ are first differences of $\widehat{\varepsilon_t}$, $\widehat{\varepsilon_{t-1}}$ are lagged values of order one of $\widehat{\varepsilon_t}$, $\Delta\widehat{\varepsilon_{t-i}}$ are changes in lagged values, and u_t is the white noise.

The null hypothesis tested is of no cointegration, i.e. γ =0, against a one-sided alternative of $\gamma < 0$. If the null hypothesis of a unit root in the DF test regression residuals is not rejected, it will be concluded that a stationary combination of the non-stationary variables has not been found and thus there is no cointegration. If the null is rejected, it will be concluded that a stationary combination of the non-stationary variables has been found and thus that the variables are cointegrated.

1.6.6 Long-short Equity Market Neutral Strategy

In the market neutral strategy, portfolios proven to have a long-run equilibrium relationship with an enhanced index were constructed, similar to Alexander and Dimitriu (2005a,b). This study used cointegration to produce portfolios that track "plus" and "minus" benchmarks (i.e. enhanced index tracking), and then construct a self-financing strategy by being short (the portfolio tracking the "minus" benchmark) and being long (the portfolio tracking the "plus" benchmark).

The 2 cointegration equations in this instance will be given by:

$$\ln(I_plus)_t = a_1 + \sum_{k=1}^{n} a_{K+1} * \ln(P_{k,t}) + E_t \quad (3)$$

$$\ln(I_minus)_t = b_1 + \sum_{k=1}^{n} b_{K+1} * \ln(P_{k,t}) + E_t \quad (4)$$

where I_t_plus is the JSE Top 40 "plus" index devised as a benchmark for the "long" portfolio and $P_{k,t}$ is the k-th stock price at time t, the series I_plus_t and $P_{k,t}$ are cointegrated and therefore the residuals series E_t is stationary. Similarly, for the "minus" I_t_minus is the JSE Top 40 "minus" index devised as a benchmark for the "short" portfolio and $P_{k,t}$ is the k-th stock price at time t.

Enhanced benchmarks were created by adding an evenly distributed annual premium to the actual index returns. It follows that the expectation is that the cointegration portfolios that successfully track the enhanced benchmarks will essentially consistently deliver a premium above market index return.

2 CHAPTER 2: LITERATURE REVIEW

The Literature on cointegration and enhanced indexation has been gradually expanding over time. Below, an overview of literature covering the field of Enhanced Indexation and Cointegration is presented.

2.1 Enhanced Indexation

An account of some of the enhanced indexation approaches and studies is given below. Since its emergence several theories and approaches to Enhanced Indexation have been proposed:

Two Broad Enhanced Indexation Strategies are framed as follows:

Schoenfeld and Yang (2003) segmented sources of properly defined enhanced indexation strategies under two broad categories: securities-based strategies and derivatives-based strategies. Each method exploits a different form of capital market imperfection and inefficiency. Securities-based strategies rely on a diversified portfolio of stocks to replicate the risk characteristics of the index and they exploit alpha through relative value mispricings or event-driven opportunities. Constructing a securities-based enhanced index portfolio involves identifying the appropriate alpha sources and choosing a technique to control tracking error relative to the benchmark. Most of the portfolio is constructed to replicate the index. A portion is then devoted to achieving alpha by overweighting or underweighting the sources of alpha. The core replication can be achieved through a full replication of the index. Alternatively, index-like returns can be achieved through an optimization method with a tilt toward the alpha source. Portfolio risk can be contained through diversifying the sources of alpha or neutralizing the overweights and underweights by index risk characteristics. When defined in the framework by Schoenfeld and Yang (2003), the securities-based enhanced indexing is based on capital market inefficiencies and not generally a fundamental alpha-target for each security. The authors called this index alpha, in that much of the outperformance of the index strategy is based on a more risk-seeking profile around trading index changes. Complexities with maintaining a securities-based portfolio increase with the complexity of the index itself, the asset class, and the number of multiple alpha sources.

A derivative-based approach to enhanced indexation replicates index performance through synthetic vehicles derived from the underlying index such as futures, options, swaps, and other exchange-traded vehicles. It should be noted that ETFs are not derivatives, but the creation/redemptions mechanisms can be used to capture mispricing - in fact, this mechanism

is what keeps ETFs trading in line with their underlying index. Derivative based approaches can simplify the complexities involved in maintaining securities-based strategies. An index arbitrage strategy is based on the arbitrage relationship between the index derivative and the underlying constituents of the index. Switching between futures and the securities in the underlying index typically achieves this process. To achieve optimal index outperformance, these strategies must cover the costs of the derivative positions as well as any tracking error and exceed the implied interest rate used in calculating the fair value of the derivative. The key to the successful implementation of options strategies for enhancement is to have diverse techniques and a deep experience set to determine when one approach would be more likely to succeed. Volatility-dampening strategies can also be used to transform the investment objective of a fund – for example, targeting a lower beta relationship with the market. Derivative-based strategies usually increase portfolio risk, introduce credit risk, and may lower the information ratio, and also contribute to additional operational complexities. The challenge for asset managers of the future will increasingly be that of finding the right blend of alpha and beta, and the beta component will likely come from index and/or enhanced index strategies.

diBartolomeo (2000) in a study underwritten by Invesco Management and Research, examined the properties of enhanced index funds as an alternative to index funds, as well as to combinations of index funds and active management. The study found that enhanced index funds offered no less likelihood of mean-variance efficiency than to passive funds invested in specified market indices. Investing in enhanced index funds was also found to be more efficient than equally risky combinations of index funds and active management. As compared to combination strategies, enhanced index funds were observed to reduce transaction costs, avoid capitalization biases, and provide better utilization of manager forecasting skills. In addition, enhanced index funds were found to lower estimation risks, allowing for more precise allocations of capital across asset classes and managers.

Alexander and Dimitriu (2004b, 2005a) executed Enhanced Indexation following a statistical arbitrage strategy based on a cointegration. In their model, they constructed portfolios that tracked two artificial indexes: the index plus a constant (index-plus) and the index minus the same constant (index-minus). They attempted to generate excess returns by selling the plus tracking portfolio and by purchasing the minus tracking portfolio. Their methodology was tested on 30 stocks of the Dow Jones Industrial Average (DJIA) over 13 years (Jan. 1990-Dec. 2003), using daily closing prices and considering as a benchmark a reconstructed DJIA

index. The same authors reapplied the same methodology to FTSE100, CAC, and S&P100, in another study, and found evidence of excess returns (called abnormal) under certain market conditions from enhanced indexation.

2.2 Cointegration in portfolio management

A number of studies have been undertaken to examine the application of cointegration to various trading strategies, however, limited studies exist which have thoroughly researched its applicability and impact in enhanced indexation. The studies conducted in this area have also mainly focused in developed stock markets, and not much has been done in emerging markets that have different dynamics.

In search for "cointegration alpha", Alexander and Dimitriu (2002) presented two applications of cointegration based trading strategies: a classic index tracking strategy and a long-short equity market neutral strategy. These two strategies were applied to the Dow Jones Industrial Average. Accurate replication of a benchmark in terms of returns and volatility was the aim of the first strategy whiles the strategy of the other, was to minimize volatility and generate steady returns under all market circumstances. The comprehensive set of back-tested results in this study demonstrated how cointegration works and validated its applicability to the construction of trading strategies; ranging from index and enhanced index tracking to long-short market neutral and alpha transfer techniques. Findings of the study in the case of the classic index tracking strategy were essentially that, provided a minimum number of stocks are used to construct the tracking portfolios and an appropriate calibration period is used, appropriate replica portfolios can be produced for the market index as well as for artificial benchmarks. Their enhanced stability resulted in a low amount of rebalancing and consequently reduced transaction costs. What was also interesting was that these same tracking portfolios produced abnormal positive returns under certain market conditions, even net of transaction costs. Statistical arbitrage portfolios which achieved returns in line with the spreads between the benchmarks tracked, also have lower volatility, low correlation and near to normal-return distributions. The successful performance depends on the number of stocks in the portfolio, the stock selection method, the spread between benchmarks tracked and the calibration period. By also using portable alpha techniques, the study by Alexander and Dimitriu (2002) showed how the characteristic of both strategies could be significantly

improved by combining them to create strongly market-neutral or enhanced index-tracking portfolios.

Conventional optimization models for passive investments were largely constructed based on tracking error and correlation measures, these models however, have been found to have significant drawbacks which are well documented in Alexandra and Dimitriu (2004), which pose limitations in terms of their applicability to a passive investments framework. To this end, Alexander and Dimitriu (2004a) proposed two portfolio construction models which they believed are designed to best suit passive investments framework: cointegration-based index tracking (Alexander, 1999) and common-trend replication portfolios (based on principal components) (Alexander & Dimitriu, 2003). The cointegration-based index tracking exercise showed that for as long as stock prices remain oscillating around the past equilibrium levels, accurate replicas of the benchmark are produced. However, in volatile markets, where returns are low and prices are moving towards new levels, this strategy yields consistent excess returns. Particularly different in approach in this study relative to others is that principal-component analysis is used for asset allocation rather than stock selection. A portfolio replicating the first principal component is constructed directly from the normalized eigenvectors of the variance-covariance matrix of stock returns. A portfolio constructed in this manner captures the largest proportion of the variation in the stock returns and is naturally suited for a passive investment framework - it requires a full replication tracking portfolio, but involves minimal rebalancing because it captures the major common trend in stock returns. The study concluded for both models that without any specific stock selection or explicit timing attempts, solely through smart optimization, the benchmark performance net of transaction costs can be significantly enhanced. These models can be applied to any type of value or capitalization weighted benchmark.

The same authors extended the study to investigate the statistical properties of the enhanced return (referred to in the study as the abnormal return) and identified a leading indicator for the abnormal return being a measure of stock price dispersion. By applying a Markov switching approach to the non-linear relationship between the abnormal return and lagged dispersion, it revealed the existence of two stock market regimes associated with the abnormal return. They found that the abnormal return was associated with the regime characterized by higher volatility and lower benchmark returns and that this regime is likely to occur when credit spreads are particularly high. Hence, investors that seek to diversify the risk of passive indexation when equity returns are low but wish to avoid the credit risk of

high yield bonds, should find a suitable alternative investment in this type of enhanced indexation.

They posited that the enhanced indexation, based on a cointegration relationship between the portfolio and the benchmark, can be interpreted as a relative pricing model with an implicit market timing element that pays off if the market switches from the regime with stable returns and low volatility to the high volatility regime when the benchmark returns are low.

The study showed that by tracking artificial indexes, such as index plus and index minus, returns that have low correlation with the index could be achieved. These returns were found to be slightly more volatile than the market index, but had significantly better Sharpe ratios. The study could still be refined to achieve more convincing evidence, as their results were based on rudimentary stock selection and rebalancing rules as well as conservative transaction and repo costs.

Maurer (2008) compared and tested two different index tracking strategies – one based on cointegration analysis of price processes of assets (CIT) and the other on market equilibrium and continuous time portfolio optimization approach (MIT), concentrating on the British and US developed markets. The first methodology tried to find long run equilibrium relations in price processes, while the second approach considers only asset returns. It was found that both strategies are able to track an index accurately. Of particular interest to this study were the findings of the study specific to enhanced index tracking, as this is the focus of this study. The study also showed that in a simple enhanced indexation approach, where one tries to track an artificial, enhanced index, the cointegration strategy was not valid at all or performed poorly. By simply enhancing an index while adding to each day's return a small constant return and then attempting to track the new enhanced index, the CIT strategy was not able to track any enhanced index. Maurer however contested that Alexander and Dimitriu (2002, 2004) in their above-mentioned study involving the analysis of some enhanced indexation strategies, made use of inaccurate critical values for the EG cointegration test analysis, and consequently ended up with incorrect conclusions. The research work concluded that for the enhanced indexation strategy proposed by Alexander, Giblin and Weddington (2002), and Alexander and Dimitriu (2002), there was no convincing empirical evidence that can be found for their findings.

Thomaidis (2013) expressed that the ultimate purpose of enhanced indexation, is being able to generate excess returns relative to the market while being largely unaffected by index movements. Similar to Alexandra and Dimitriu (2005), he employed a long-short investment framework, as one way of achieving a high alpha; low beta investment objective. This study attempted to extend Alexander and Dimitriu's (2005) methodology by presenting an easy to implement technique for deriving optimal allocation rules in a long-short investment context, using an assembled data set composed of the Dow Jones Composite Average (DJCA) stocks. The procedure applied combined careful stock picking with optimal portfolio allocations using an iterative restricted regression approach. In the study, trading positions were optimised, subject to cardinality constraints, upper and lower limits were placed on individual portfolio weights. The results showed that long-short portfolio tracking enhanced benchmarks were superior to a passive buy-and-hold strategy in terms of both reducing trading risk and generating positive income.

In their view that the quality of a tracking portfolio depends highly on the stock selection procedure, Acosta-Gonzalez, et al. (2015) proposed the use of a model selection technique based on optimising the cointegration level of the tracking portfolio and the benchmark index instead of selecting, since in prior studies the assets were based on technical analysis or other ad-hoc decisions. The heuristic optimization method applied was based on an adapted Genetic Algorithm version employed by Acosta-Gonzalez and Fernandez-Rodriguez (2007). The study combined the cointegration principles for describing the common stochastic trends between assets with new computational techniques of evolutionary heuristics devoted to finding the assets which constitute a tracking portfolio. This method permitted them to obtain tracking portfolios (with very few stock) that were able to track the benchmark index properly. Several advantages were also found when applying the methodology to Enhanced indexation and Long-short strategies previously applied by Alexander and Dimitriu (2005a,b). From the empirical results in the study it was evident that the optimisation of the augmented Dickey-Fuller (ADF) t statistic as a measure of cointegration between stocks and the index (including enhanced index) provided tracking portfolios that achieved the set return goals, despite using only five stocks in the portfolios.

Definition: A Genetic Algorithm is a class of optimization techniques, based on principles of natural evolution developed by Holland (1975), which tries to overcome noted problems of traditional optimization algorithms. Such noted problems are an absence of continuity or differentiability in the loss function (Acosta-Gonzalez, et al., 2015).

3 CHAPTER 3: RESEARCH METHODOLOGY AND INVESTMENT STRATEGIES

3.1 Concept of cointegration and its principles

This study drew from the works of Alexander (1999), Alexander and Dimitriu (2002, 2004a,b), Maurer (2008), Dunis and Ho (2005) and Dunis, Giorgioni, Laws and Rudy (2010). The focus of this paper is the construction of both classical index tracking and enhanced indexation portfolios based on a cointegration model, which allows the replication of various benchmarks, with different numbers of stocks, an approach pioneered by Alexander (1999). If two or more series are cointegrated, in intuitive terms, this implies that they have a long-run equilibrium relationship that they may deviate from in the short-run, but which will always be returned to in the long-run. A stationary linear combination of a stock prices and a market index, identified through cointegration, implies mean reversion in price spreads. The finding that the spread in a system of prices is mean reverting reveals that despite the size of the spread, the prices will remain tied together in the long term. As Maurer (2008) stated, through cointegration analysis, it is not only long term interrelations that may be described, but also disequilibrium situations that may appear in the short-term and subsequent mean-reversion tendencies towards the system's "steady state".

3.2 The concept of cointegration justification

Alexandra (1999) explained that the rationale for constructing portfolios based on a cointegration relationship with the market index rather than correlation rests on the following features of cointegrated systems:

- The price difference between the index and the portfolio is by construction stationary, which implies that the tracking portfolio will be "tied" to the index in the long run.
- The stock weights, being based on the history of prices rather than returns, have enhanced stability.
- There is a full use of the information contained in level variables such as stock prices.

3.2.1 Cointegration equation estimation and testing

There are several econometric approaches to estimate and test for cointegration, among them, the Engle-Granger procedure in bivariate cointegration analysis and the Johansen multivariate framework, introduced by Engle and Granger (1987), Engle and Yoo (1987) and Johansen (1991) respectively. Even though the Engel-Granger test was the first formulated, the Johansen test is considered to be the more superior and widely used test, as it allows for the estimation of more than one cointegration relationship, in the case of a data set that contains two or more time series. The Engle-Granger (1987) methodology is preferred in certain instances due to its intuitive interpretability and simplicity in application, an argument also put forth by Alexander and Dimitriu (2002). The Engel-Granger test was therefore applied in this research paper.

3.3 Background on the Engle-Granger Methodologies

3.3.1 Overview on Engel-Granger Approach to Estimation of Cointegration Systems

The Engle-Granger estimation methodology that relies on a two-step estimator is performed as follows; the first step consists of estimating a long-run equilibrium relationship and the second is the estimation of the dynamic error-correction relationship (ECM) using lagged residuals, in both steps using simple ordinary least squares estimation (LSE). The estimated cointegrated vector from the first step describes the long-run equilibrium relationship in the system, while the ECM from the second step characterizes the short-run disequilibrium situation and subsequent mean-reversion movements. Although this method only allows for finding one cointegration vector even though many more may exist, it will suffice for purposes of this study. For purposes of this study, only single cointegration relationship between the tracking portfolio and index is required. In similar view to Gabriel (2002), this study argues that for portfolio management purposes the error variance minimizing criterion is far more important for portfolio risk management purposes than Johnsen's (1991) added criterion of maximising stationarity.

3.3.2 Overview on Engel-Granger Approach to testing for Cointegration

A Cointegration test is essential a test for the existence of a long-run relationship between variables. The Engel-Granger cointegration test's validity is conditional on the existence of some cointegration relationship (cointegration vector) among variables. To test for

cointegration; in the case of known coefficients of a cointegration relationship, a unit root test is applied to the residuals of the cointegrated equation. However, when the cointegrated vector is unknown, the first step needs to be the estimation of a cointegration regression.

An illustration of how this is applied in the index tracking portfolio construction will be covered in the following paragraphs.

3.4 Unit Root Test

Cointegration requires that two variables be integrated of the same order. If the integrated variables are found to be integrated to the same order, then a cointegration test must be performed on them (Johansen, 1988). In order to test for stationarity as well as to determine the order of integration, the variables of the time series are subject to a unit root test. The candidate stock and index price series for non-stationarity using the augmented Dickey-Fuller (ADF) (1997) unit root test were tested using Eviews.

3.5 Portfolio Construction Strategies Based on Cointegration

3.5.1 Classic Index Tracking Strategy

Even though the specific focus of this study is on enhanced indexation, it is possible that the study would not be comprehensive enough, if it did not cover the basic index tracking problem. It is for this reason that, the study began by looking at the classic index tracking strategy which aims to replicate a specific benchmark in terms of returns and volatility using cointegration relationship rather than correlation, in order to benefit from the entire set of information contained in the stock prices. These replicating portfolios are expected to match the return and volatility as well as have high correlation with the index, albeit using only a subset of the index constituents.

3.5.2 Portfolio Construction

The main steps involved in the index tracking process are stock selection and weight allocations. These steps in more detail are covered in the paragraphs that follow.

3.5.2.1 Stock Selection

In portfolio construction, stock selection is usually executed through fundamental or technical analysis and/or proprietary selection models. It is worth noting that the success and quality of index tracking is highly dependent on the stock selection output, as such stock selection is a critical consideration in portfolio construction. Identifying the most successful stock selection methods does not constitute the focus of this study. For purposes of this study, a simple stock selection criteria available based on the price ranking of the stocks in the index at the moment of portfolio construction was applied, as has been applied in previous studies by Alexander and Dimitriu (2002, 2004b), Maurer (2008) and Dunis and Ho (2005). Alexander and Dimitriu (2002) found that to ensure cointegration, a minimum number of 20 stocks (in their case representing 0.67% of the index basket) and a minimum calibration period of 3 years are required. We will therefore construct portfolios consisting of the first 10, 20, 25 and 30 constituents stocks, ordered according to their prices from highest to lowest.

3.5.2.2 Constituents' weights

The weights of the above selected stocks within the tracking portfolio was estimated based on ordinary least squares (OLS) coefficients of a cointegration equation that regresses the index log price on tracking portfolio stocks log prices as independent variables, over a specified in sample (calibration) period. The cointegration index tracking regression is given by:

$$\ln(Index)_t = C_1 + \sum_{k=1}^{n} C_{K+1} * \ln(P_{k,t}) + E_t \qquad (1)$$

Where I_t is the JSE Top 40 Index, C_1 is a constant, $C_{K+1}$is the approximate weight of stock k in the tracking portfolio, $P_{k,t}$ is the log price of the constituent stock P_k at time t, and E_t is the stationary spread between the index and the tracking strategy. If the series I_t and $P_{k,t}$ are cointegrated, it therefore means that the residual series i.e. the tracking error, E_t is stationary.

The use of log prices has the advantage that the tracking error E_t is in return format. Provided that the levels variables are cointegrated, so will their logarithms. Henry and Juselius (2000) reported that if a variable had a unit root in its original units of measure, it would become deterministic over time if it had a constant error variance. The estimated coefficients however need to be normalised to sum up to one in order to provide the percentage weight of each selected stock in the index tracking portfolio.

The application of OLS to non-stationary variables as in model (1) is however only valid in the special case of the existence of a cointegration relationship, that is if and only if the residuals in (1) are stationary. On the basis of the theory that integrated variables of order one, I (1), may have a long run relationship, it is crucial to test for the existence of such a relationship. Testing for cointegration is therefore an essential part in constructing cointegration based tracking portfolios.

3.5.3 Back-Testing Procedure

For purposes of this analysis, the simplest stock selection criterion was used, i.e. the weight of the stocks in the index at the moment of the initial portfolio construction. Portfolios containing 10, 20, 25 and 30 constituents were constructed, ordered descending according to their weights in the index as at 31 December 2015. Only relative weights are subsequently modified Dunis and Ho (2005)

An initial in-sample portfolio (P0) was constructed for the period from November 2009 to December 2012, and it was progressively expanded monthly until December 2015; the initial portfolio was constructed over the period from November 2009 to December 2012, and simulated out of sample in January 2013 as the first tracking portfolio, the second tracking portfolio was then constructed over the period November 2009 to January 2013 and simulated out of sample in February 2013, the third tracking portfolio was constructed using data from November 2009 to February 2013 and simulated out of sample in March 2013, and so on (Dunis and Ho, 2005)

The initial portfolio (P0) was based on 3 years of daily data and the coefficients of the cointegration regression are subsequently re-estimated monthly using the Engel-Grager procedure. The first cointegration tracking portfolio was simulated from 2 to 30 December 2013, using estimation data from 2 November 2009 to 31 December 2012 to determine portfolio weights. The last tracking portfolio was simulated from 1 December 2015 to 30 December 2015.

Once the tracking portfolio weights were obtained, the sequence was then back-tested. The back-testing was run over a period following the calibration period. A number of measures were applied to assess the performance of the model:

3.5.3.1 Engel-Granger cointegration test

The expectation was for the above constructed tracking portfolios and reconstructed index to have a long run relationship, this was based on theory. The residuals of each cointegration regression were tested for stationarity following the Engel-Granger methodology for cointegration testing. Section 3.2, has shown that the second step of the Engle-Granger's method is an ADF test applied to the residuals of the long-run equation. To test for cointegration, we subject the residuals to the cointegration augmented Dickey-Fuller (ADF) (1997) unit root test at 95% confidence level. The cointegrating ADF regression estimated on the residuals of the cointegration regression is:

$$\Delta\widehat{\varepsilon_t} = \gamma\widehat{\varepsilon_{t-1}} + \sum_{i=1}^{p} \alpha_i \Delta\widehat{\varepsilon_{t-i}} + u_t \quad (2)$$

where Δ denotes the difference operator, $\widehat{\varepsilon_t}$ is the estimated residuals from the cointegration regression 1, γ is the coefficient of $\widehat{\varepsilon_{t-1}}$, p is the lag order of the autoregressive process, $\Delta\widehat{\varepsilon_t}$ = $\widehat{\varepsilon_t}$ - $\widehat{\varepsilon_{t-1}}$ are first differences of $\widehat{\varepsilon_t}$, $\widehat{\varepsilon_{t-1}}$ are lagged values of order one of $\widehat{\varepsilon_t}$, $\Delta\widehat{\varepsilon_{t-i}}$ are changes in lagged values, and u_t is the white noise.

The null hypothesis tested is that of no cointegration, i.e. γ =0, against a one-sided alternative of $\gamma < 0$. If the null hypothesis of a unit root in the DF test regression residuals is not rejected, it will be concluded that a stationary combination of the non-stationary variables has not been found and thus there is no cointegration. If the null is rejected, it will be concluded that a stationary combination of the non-stationary variables has been found and thus that the variables are cointegrated.

3.5.3.2 Returns on the tracking strategy

To ensure that the portfolios were indeed constructed on a cointegration relationship, the daily prices for each of the portfolios monitored was estimated. At consecutive re-estimations of the cointegration regression, the number of shares in each portfolio was kept constant.

Ideally tracking the index based on the cointegration coefficients would imply keeping the weights constant and not the number of stocks in the portfolio. This would mean daily (or

even intra-day) rebalancing of the portfolio to account for the effect of price changes on the stock weights. This would result in a rather active strategy, and while it would likely reduce the tracking error, it would also generate high transaction costs and counter one of the objectives of this study. The tracking error was based on a no trading for a period of a month between consecutive re-estimations of the cointegration coefficients and the number of stocks (instead of the weights) was kept constant.

Assuming that the portfolio weights $w_{k,T}$ are estimated at time T, the price of the portfolio $\pi_{(T+x,T-1)}$ at time T+x and T-1, can be computed based on the prices $P_{k,T}$ and $P_{k,T+x}$ of the n stocks in the portfolio as follows, with :

$$\pi_{T+x} = \pi_{T-1} \sum_{k=1}^{n} \frac{W_{k,T}}{P_{k,T}} P_{k,T+x} \qquad (5)$$

3.5.3.3 Transaction costs

To account for the impact of the price spread and the brokerage fees on the portfolio returns, a fixed amount of 20 basis points transaction costs on each turnover trade value was assumed. This is in line with previous studies such as Alexander and Dimitiu (2002). The JSE Top 40 constituents are the 40 largest companies by market capitalisation on the JSE and therefore known to be very liquid and therefore should have lower trading costs.

Transaction costs are simply defined as:

$$TC_t = 0.002 \sum_{k=1}^{N} abs\,(W_{k,t} - W_{k,t-1}) \qquad (6)$$

where TC denotes the transaction costs relative to the overall portfolio value at time t and $W_{k,t}$ is the portfolio weight of stock k at time t.

3.5.3.4 Volatility of strategy's returns

For each of the tracking portfolios constructed, the annual volatility of the excess returns was computed based on 252 trading days. An equally weighted approach was used to compute the unconditional volatility over the entire data sample and an exponentially weighted method to analyse the short-term volatility behaviour.

3.5.3.5 Correlation of the tracking portfolio returns with the index returns

For each replicating portfolio, the following was computed and reported:

- The correlation of its returns with the market returns, and
- The correlation of the excess returns (i.e. tracking error) with the market returns.

As estimation methods similar to the volatility calculations, an equally weighted approach to compute the unconditional correlation over the entire data sample and an exponentially weighted method to analyse the short-term correlation were used.

3.5.3.6 Skewness and excess kurtosis of tracking portfolio excess returns

For each of the replicating portfolios constructed the skewness and excess kurtosis was computed to determine the distributional properties of the tracking errors.

3.5.3.7 Sharp Ratio and information Ratio

The Sharpe Ratio and Information ratio of each tracking portfolio and where appropriate that of the benchmark was calculated. The Stefi for the Risk Free Rate over the sample period as provided by Morningstar was used.

3.5.4 Enhanced Indexation Strategy

3.5.4.1 Long-short Equity Market Neutral Strategy

In this next strategy, portfolios that are insulated from the upwards and downwards movements of the market and can deliver positive returns regardless of market direction were constructed. This was done through holding both long and short positions in equity portfolios. The analysis in the long short equity market neutral strategy was also based on cointegration, however this time the cointegrated portfolio prices were constructed on the basis of a long run equilibrium relationship with an enhanced index. A self-financing long-short strategy was then setup through holding portfolios tracking different "plus" and "minus" benchmarks, an approach similarly applied by Alexander and Dimitriu (2005). The spread between the performance of the long and short portfolios should result in a return in excess of the risk-free rate.

Alexander and Dimitriu (2005b) argued that, provided that each tracking portfolio in the strategy is a suitable replica of its "plus" or "minus" benchmark, the long-short market neutral equity strategy can be expected to generate returns according to the spread between the "plus" and the "minus" benchmarks.

As an extension for exploiting the tracking potential of cointegrated portfolios, replicating "plus" and "minus" benchmarks by creating long and short portfolios from the selected stocks was attempted. Cointegration portfolios were constructed to track enhanced benchmarks that were artificially created. Such benchmarks were created by adding and subtracting an evenly distributed annual premium to the actual index returns. It follows that the expectation is that the cointegration portfolios that successfully tracks the enhanced benchmarks will essentially consistently deliver a premium above market index return.

In practice, the construction of both long and short portfolios derives from the index tracking strategy: only this time two cointegrating portfolios tracking 2 benchmarks were aimed for, a benchmark "plus" and a benchmark "minus". For example, to construct JSE Top 40 + 5% (artificial index), having assumed a 252 day trading year 0.01984%, i.e. 5%/252 was added to the JSE Top40 daily returns. This was then followed by finding a cointegration relationship between this artificial benchmark and some of the constituents of the index.

The choice of the annual excess return to construct the "long and "short" cointegration portfolios is critical. While there is a good reason to expect *a priori* that a sufficiently large basket of component equities will be cointegrated with the reference market index, this may not be true in the case of artificial benchmarks, such as those created for the "long" and "short" portfolios.

3.5.4.2 Stock selection

As in the classical index tracking analysis, a simple stock selection criteria available based on the price ranking of the stocks in the index at the moment of portfolio construction will be applied.

3.5.4.3 Constituents' weights

To determine the portfolio constituents' weightings in the Enhanced Indexation problem ordinary least squares (OLS) will be employed. To construct portfolios that closely tracks the artificial indexes using ordinary least squares for the cointegration equation, the two cointegration equations are:

$$\ln(I_plus)_t = a_1 + \sum_{k=1}^{n} a_{K+1} * \ln(P_{k,t}) + E_t \qquad (3)$$

$$\ln(I_minus)_t = b_1 + \sum_{k=1}^{n} b_{K+1} * \ln(P_{k,t}) + E_t \qquad (4)$$

where I_t_plus is the JSE Top 40 "plus" index devised as a benchmark for the "long" portfolio and $P_{k,t}$ is the k-th stock price at time t, the series I_plus_t and $P_{k,t}$ are cointegrated and therefore the residuals series E_t is stationary. Similarly, for the "minus" I_t_minus is the JSE Top 40 "minus" index devised as a benchmark for the "short" portfolio and $P_{k,t}$ is the k-th stock price at time t.

Note: The stock weights are not restricted to be positive weights only in the tracking portfolios above, in fact it is likely that some short positions in the portfolios tracking both "plus" and "minus" benchmarks will be taken.

The stock holdings in the long-short strategy were obtained by netting their individual weights in the "plus" and "minus" portfolios. The transaction costs on the long-short strategy were less than the sum of the costs incurred by trading the "plus" and "minus" portfolios individually. Specifically, the transaction costs were determined as follows:

$$TC_t = 0.002 \sum_{k=1}^{n} abs\left((wplus_{k,t} - wminus_{k,t}) - (wplus_{k,t-1} - wminus_{k,t-1})\right) \quad (7)$$

where TC denotes the transaction costs relative to the overall portfolio value at time t and $W_{plus,minus,t,t-1}$, is the portfolio weight of stock k at time t and t-1.

In order to analyse the correlation of the long-short strategy returns with the market returns, the returns on the "plus" and "minus" portfolios can be written separately as functions of benchmarks as follows:

$$R_{+,t} = \rho + \frac{\delta_+}{\delta_{B+}} R_{B+,t} + \epsilon_{+,t} \quad (8)$$

$$R_{-,t} = \rho - \frac{\delta_-}{\delta_{B-}} R_{B-,t} + \epsilon_{-,t} \quad (9)$$

where $R_{+,/-}$ is return of the tracking portfolio "long" plus and tracking portfolio "short" minus, $\rho +/-$ are the correlation coefficients between the "plus" and "minus" portfolio returns with the plus /minus benchmark returns, $\delta_+/\ \delta_{B+}$ and δ_-/δ_{B-} are relative volatilities of the "plus" respectively minus portfolio returns to the "plus"/"minus" benchmark returns, and ϵ_+ and ϵ_- are the tracking errors of the "plus" and "minus" portfolios.

The "plus"/ "minus" benchmarks returns can be written in terms of market returns +/- the fixed annual spreads, denoted by μ_+ and μ_-, while the return on the long-short strategy will be the difference between the returns on the individual "plus" and "minus" portfolios:

$$R_{+/-,t} = R_{+,t} - R_{-,t} = R_{M,t}\left(\rho_+ \frac{\sigma_+}{\sigma_{B+}} - \rho_- \frac{\sigma_-}{\sigma_{B+}}\right) + \mu_+\rho_+ \frac{\sigma_+}{\sigma_{B+}} + \mu_-\rho_- \frac{\sigma_-}{\sigma_{B-}} + \epsilon_{+,t} - \epsilon_{+,t} \quad (10)$$

Again, the above explained procedures are in line with Alexander and Dimitriu (2002).

4 CHAPTER 4: CO-INTEGRATION BASED PORTFOLIOS USING JSE TOP 40 EQUITY INDEX

4.1 Data

For the historical cointegration analysis of price equilibrium the JSE Top 40 Index constituents' daily historical prices for the period concerned was required. The stock market data was obtained from Bloomberg and Morningstar. When working with long term data one encounters survivorship bias. In order to overcome this issue this study could only go back as far as 2009, where in all the index constituents as at 31/12/2015 daily price history was available. This research therefore utilised daily data series for the period 02/11/2009 – 31/12/2015, which contains 1608 observations. The six-year data is divided into two sub-periods: in-sample period called the calibration period (first 3 years) and out-of-sample period being the testing period (the last 3 years).

Similar to Alexander and Dimitriu (2002), in order to carry out the analysis, an artificial reconstructed JSE Top 40 Index was computed from daily close price of the stocks currently included in the JSE Top 40 basket, using available weights as at 31/12/2015. The value of the reconstructed index for one particular day in the sample was computed as a proportionally weighted sum of all constituents' stock prices.

The main difference between the actual index and the reconstructed one is the constituent stocks, which change periodically in the actual index but not in the reconstructed index. The use of a reconstructed index instead of the actual one is justified by this study's interest in the current structure of the index: that is, this study compared the performance of portfolios comprising the stocks currently included in the JSE Top 40 Index with a market index constructed from the same stocks.

Any missing observations were replaced with the last available closing price for that particular stock. The statistical properties of the log price series are reported in Appendix 1.

The above applies to both the classical indexation and enhanced indexation strategies.

4.2 Unit Root Testing

The Augmented Dickey-Fuller Test was used in order to test for unit root for all log price processes to check whether that are I(1), and thus cointegration analysis may be considered at all for further analysis. In the case of the log price data series, this study found that the null of a unit root can never be rejected on a 1% significance level and yet in the case of the first differences of the same series, the same null is always rejected with a confidence of 99.99%. Table 1 below, contains the results of the ADF test.

Table 1: Results of the ADF test

	LOG(AGL_SJ_EQUITY)	LOG(AMS_SJ_EQUITY)	LOG(ANG_SJ_EQUITY)	LOG(APN_SJ_EQUITY)	LOG(BAT_SJ_EQUITY)	LOG(BGA_SJ_EQUITY)	LOG(BIL_SJ_EQUITY)	LOG(BTI_SJ_EQUITY)	LOG(BVT_SJ_EQUITY)	LOG(CFP_SJ_EQUITY)	LOG(CPI_SJ_EQUITY)	LOG(DSY_SJ_EQUITY)	LOG(FFA_SJ_EQUITY)	LOG(FFB_SJ_EQUITY)	LOG(FSR_SJ_EQUITY)	LOG(GPT_SJ_EQUITY)	LOG(INL_SJ_EQUITY)	LOG(INP_SJ_EQUITY)	LOG(ITU_SJ_EQ
Mean	10.13736	10.74705	9.959002	9.695981	8.194022	9.603812	10.10956	10.67402	9.941062	8.710738	9.910200	8.724203	7.190391	6.525787	7.952369	7.674569	8.811457	8.802542	8.419135
Median	10.20176	10.74585	10.17502	9.663007	8.102432	9.564442	10.09950	10.70899	9.955748	8.801318	9.878144	8.638228	7.247437	6.545048	7.958577	7.774436	8.716044	8.699515	8.348445
Maximum	10.58583	11.32176	10.57132	10.69874	9.743612	9.900984	10.46569	11.37511	10.48690	9.397484	11.01863	9.646464	7.484369	8.366138	8.665613	8.022569	9.408371	9.412955	8.918249
Minimum	9.718600	9.711600	8.905987	8.743851	7.364351	9.392662	9.655731	10.00915	9.358760	7.670895	8.706917	7.971217	6.834550	4.861516	7.252189	7.207860	8.326275	8.323608	8.050697
Std. Dev.	0.283589	0.338742	0.441331	0.605545	0.681028	0.118448	0.143309	0.380036	0.324783	0.511712	0.499199	0.486250	0.171863	0.676255	0.401116	0.215145	0.296661	0.305932	0.222391
Skewness	-1.975160	-0.394726	-0.484443	0.081088	0.630662	0.545364	0.121899	-0.172621	-0.057751	-0.268311	0.215322	0.187891	-0.457804	0.207116	0.150139	-0.520830	0.483835	0.508456	0.556970
Kurtosis	8.468650	2.668795	1.774933	1.462825	2.230773	2.218261	2.963400	1.716338	1.559556	1.609952	3.309775	1.666795	1.963482	2.230930	1.760614	2.010098	1.883547	1.869774	1.993810
ADF Test Statistic Level	-0.613367	0.008008	-0.738331	-1.123393	1.190502	-2.502103	-2.026926	0.010205	-1.104503	-1.606193	-1.306171	-0.437573	-1.843581	-0.562354	-1.344273	-2.092965	-0.326680	-0.325647	-0.227985
ADF Test Statistic 1st diff	-41.93834	-30.06323	-39.86336	-40.49531	-39.68880	-32.36821	-39.87847	-41.09605	-44.34470	-39.51119	-38.05432	-31.52615	-34.23654	-47.09101	-32.26554	-42.85824	-43.95036	-42.31995	-43.21251

	LOG(MNP_SJ_EQUITY)	LOG(MRP_SJ_EQUITY)	LOG(MTN_SJ_EQUITY)	LOG(NED_SJ_EQUITY)	LOG(NPN_SJ_EQUITY)	LOG(NTC_SJ_EQUITY)	LOG(OML_SJ_EQUITY	LOG(PDF_SJ_EQUITY)	LOG(PEI_SJ_EQUITY)	LOG(PEM_SJ_EQUITY)	LOG(PMH_SJ_EQUITY	LOG(SAB_SJ_EQUITY)	LOG(SBK_SJ_EQUITY)	LOG(SHP_SJ_EQUITY)	LOG(SLM_SJ_EQUITY)	LOG(SNH_SJ_EQUITY)	LOG(SOL_SJ_EQUITY)	LOG(TBS_SJ_EQUITY)	LOG(VOD_SJ_E
Mean	9.259972	9.295511	9.68574	9.776439	11.09182	7.615008	7.798428	6.784187	7.44244	9.635518	8.21013	10.55992	9.375111	9.468796	6.297343	8.065317	10.59045	10.13427	9.
Median	9.131838	9.429114	9.666657	9.799154	10.90571	7.539694	7.796671	6.827087	7.386471	9.612299	8.256552	10.57128	9.336342	9.595194	8.317521	7.853651	10.55958	10.21464	9.
Maximum	10.42112	10.24885	10.16624	10.20736	12.31627	8.38936	8.427487	7.145199	8.072779	10.24572	8.894122	11.47765	9.777074	9.934841	8.999143	9.06184	11.07456	10.59476	9.
Minimum	8.261471	8.062118	9.19238	9.323669	10.12663	7.041412	7.193152	6.46386	6.96319	9.000724	7.422089	9.893437	9.108197	8.705994	7.635304	7.331008	10.19616	9.642123	8.
Std. Dev.	0.628861	0.592577	0.24527	0.236214	0.660132	0.400632	0.373271	0.167793	0.331001	0.390581	0.415363	0.434674	0.15472	0.301392	0.395281	0.517036	0.217101	0.236116	0.
Skewness	0.215343	-0.312568	0.061067	-0.05467	0.35148	0.468072	0.02001	0.069292	0.231194	0.008121	-0.053029	0.032089	0.664849	-0.841752	0.073483	0.563195	0.397527	-0.400651	-0.
Kurtosis	1.62988	2.019113	1.766494	1.677201	1.633467	1.779654	1.493151	1.821637	1.596381	1.530934	1.755445	1.616184	2.491815	2.564465	1.579971	1.879498	2.508143	1.862962	2.
ADF Test Statistic Level	-0.01007	-1.741167	-1.595702	-1.654402	0.032541	-0.756452	-0.615594	-1.807707	0.263047	-0.927838	-1.48289	0.187482	-1.893226	-2.479655	-1.333945	-0.06997	-1.781214	-1.90839	-1.
ADF Test Statistic 1st diff	-40.93018	-41.25843	-43.70301	-32.16326	-43.02759	-44.35105	-41.6209	-31.59626	-44.77903	-32.7175	-32.76006	-42.54602	-32.70557	-40.12067	-41.92143	-41.25021	-31.12554	-40.27318	-4

MacKinnon critical values:

1% level	-3.434214
5% level	-2.863134
10% level	-2.567666

The outcome of the ADF test suggests that the given data series are indeed all I(1), and thus cointegration analysis can apply.

Table 1 also contains some descriptive statistics of the data.

The average return of the stocks ranges between 11% and 6.5%, which is also aligned to the stocks in the tracking portfolio with 10 stocks having average returns ranging between 11% and 8%. Only four stocks have an average return below 8%, which are Fortress Income Fund Ltd (FFA & FFB), Reinet Investments SCA (REI) and Redefine Properties Ltd (RDF). All stocks appear to fall within one standard deviation of their mean. This seems to suggest that the stock returns are normally distributed.

For some stocks the mean is slightly less than the median, suggesting negatively skewed distribution due to some lower returns shifting the mean to the left. While some of the stocks like Standard Bank Group Ltd (SBK) and Steinhoff International Holdings NV (SNH) appear to be positively skewed distribution with resultant skewness of 0.66 and 0.56 respectively.

The data also shows a wide spread of the stock returns with a minimum value as low as 5 for Fortress Income Fund Ltd B (FFB) and a maximum value as high as 12 for Naspers Ltd. (NPN). It is also interesting that the spread of returns for each stock are however, not that wide, which has been confirmed by the finding that all the stocks fall within one standard deviation.

This data has positive kurtosis values that range between 1.46 and 8.47, indicating a distribution that has heavier tails and sharper peaks than the normal distribution. Most stocks' kurtosis value is below 2.97, with only one stock, Anglo American PLC (AGL), having a distribution that is much deviated from normal at a kurtosis value of 8.

4.3 Back-testing the Classic Index Tracking Strategy

4.3.1 Rolling window EG co-integration test

4.3.1.1 Engel-Granger cointegration test

This study tested the residuals from each EG cointegration regression for stationarity by applying the Engel-Granger methodology for cointegration testing.

It was mentioned earlier that Maurer (2008) contested the findings of Alexander and Dimitriu (2002, 2004) and other authors on the basis that they made use of inaccurate critical values to

perform the Engle-Granger cointegration test. Since this is a test on residuals of an estimated model, then the critical values applicable are different compared to those used for a DF or an ADF test on a series of raw data. Engle and Granger, and Engle and Yoo (1987) tabulated a new set of critical values, that are appropriate for the EG Test. The critical values for the t-statistic have thus been obtained using interpolation on published critical values calculated by Maurer (2008).

Based on the Engel-Granger test results as per Figure 1 the constructed tracking portfolios proved to not be sufficiently cointegrated with the reconstructed market index, except for the Tracking Portfolio consisting of 10 stocks. Its only in the case of a portfolio consisting of 10 stocks where cointegration was mainly found to exist at a 99.99% confidence level, as per figure 1. Albeit there are times when for short periods, the portfolio breaks the cointegration at this level but still remains cointegrated at both 95% and 90% confidence level. The null hypothesis of no cointegration could not be rejected for most of the calibration period for the portfolio with 30 stocks. Both the 25 and 20 stock portfolios go through periods of cointegration and no cointegration mainly in the out of sample period at a 90% confidence level. The 15 and 5 stock portfolios also proved to be less cointegrated with the index, with the 5 stocks portfolio depicting the least cointegration.

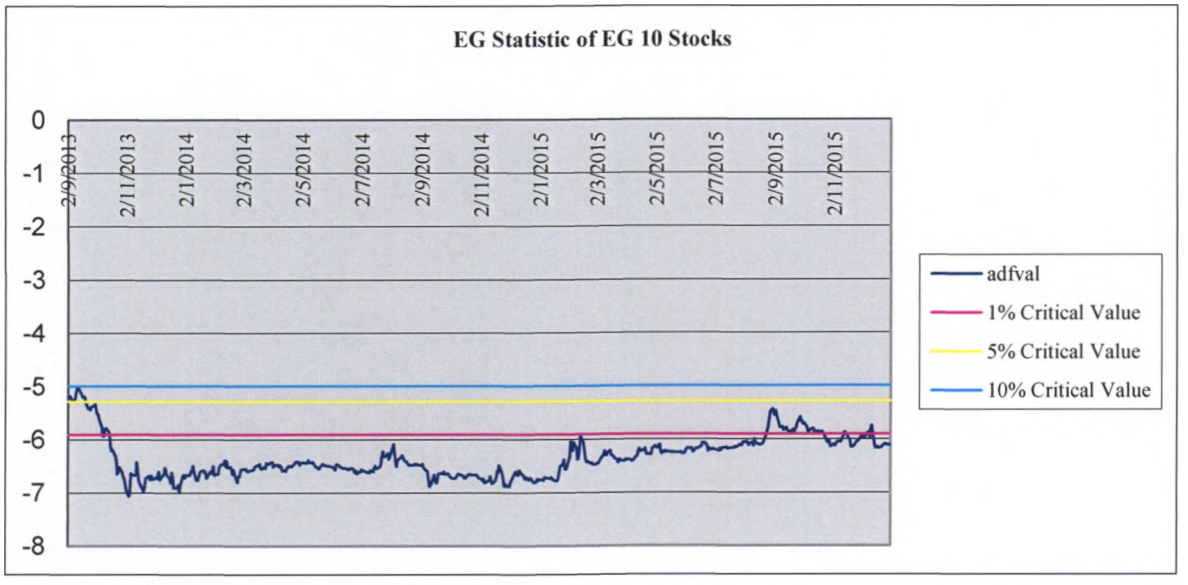


Figure 1: EG Statistic of EG 10 Stocks

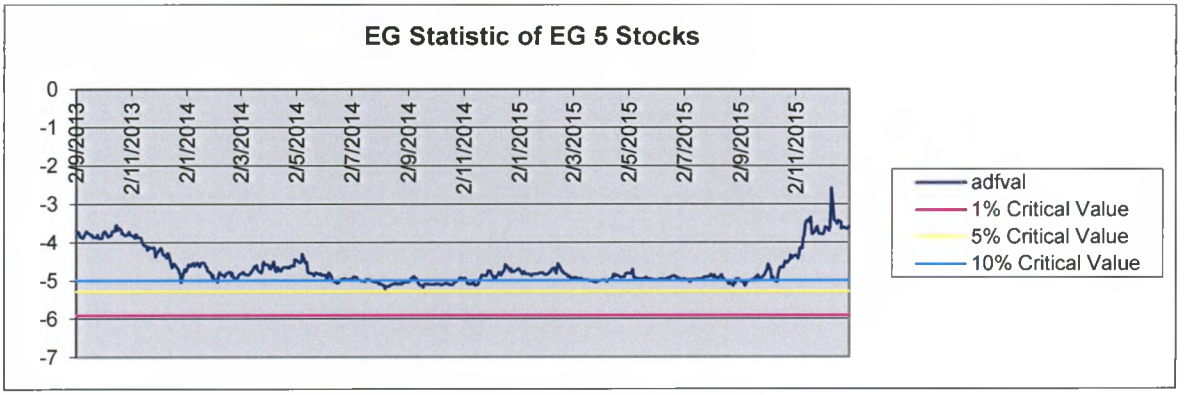


Figure 2: EG Statistic of EG 5 Stocks

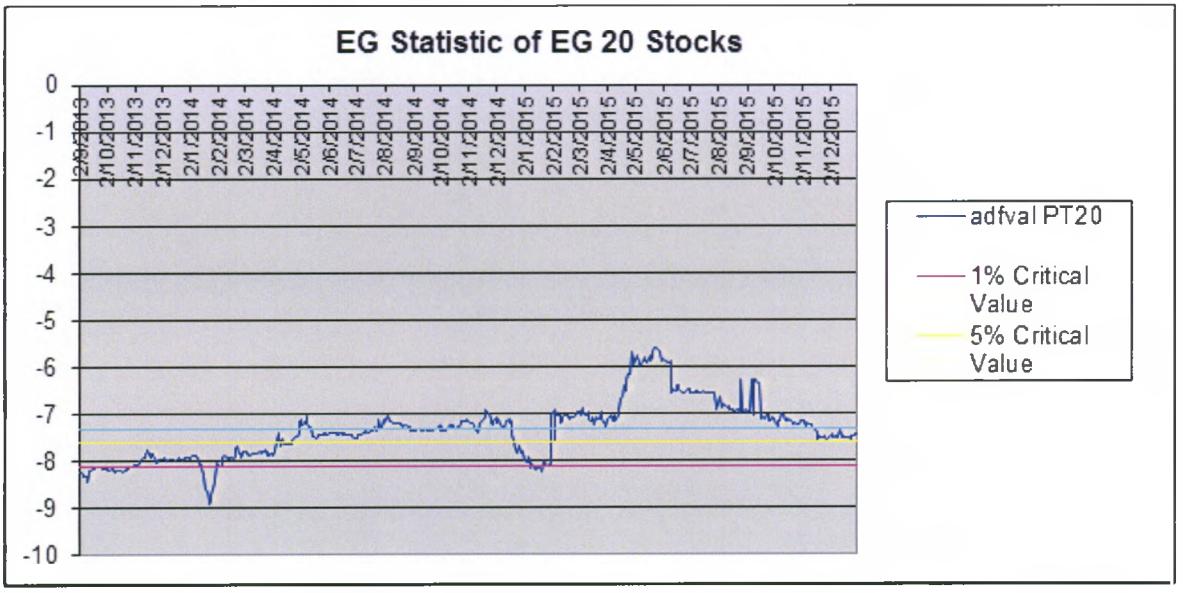


Figure 3: EG Statistic of EG 20 Stocks

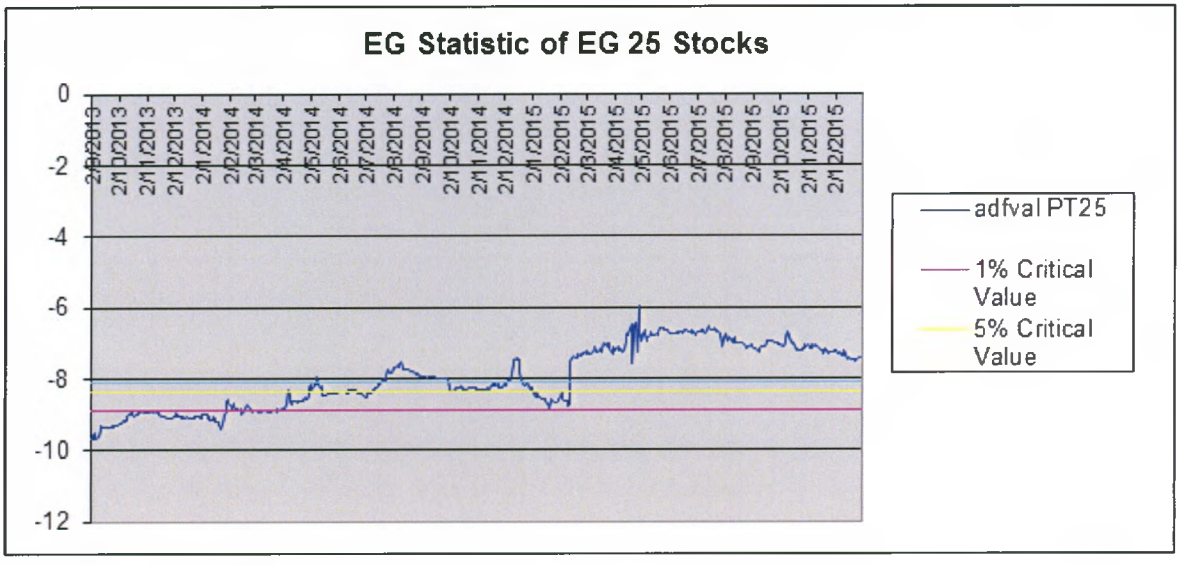


Figure 4: EG Statistic of EG 25 Stocks

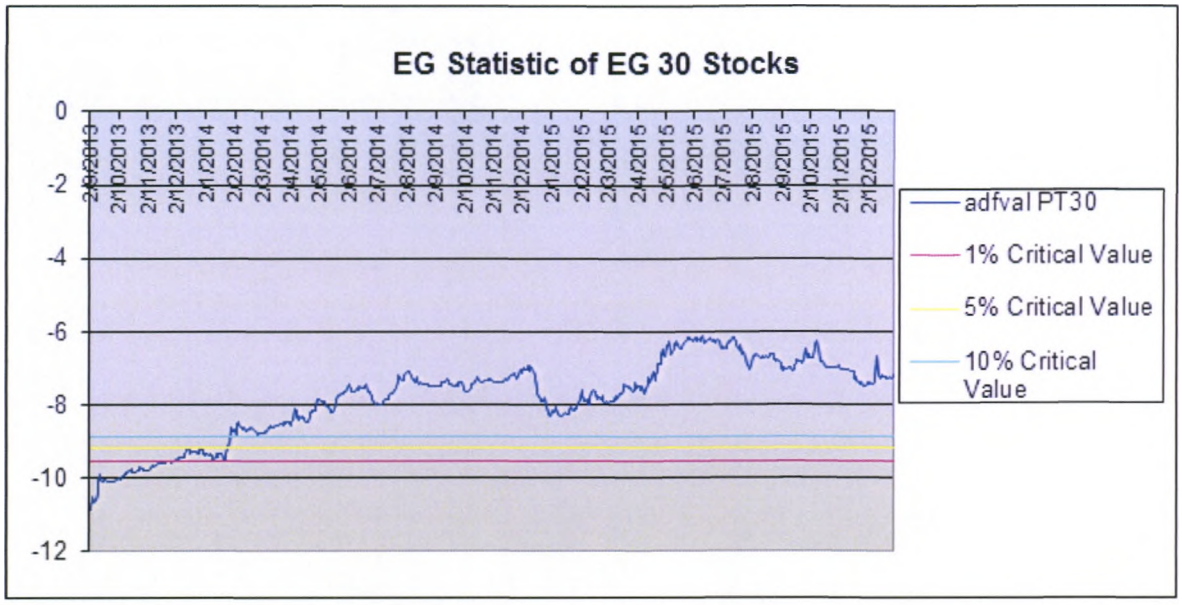


Figure 5: EG Statistic of EG 30 Stocks

Surprisingly and contrary to other studies as well as intuition in the case of South African Equities, the degree of cointegration appears to decrease as the number of stocks in the tracking portfolio increases. Based on the above results it can be concluded that portfolios consisting of 5, 15, 20, 25 and 30 stocks do not allow the construction of portfolios sufficiently cointegrated with the index. This study therefore only proceeded with tracking portfolios with 10 stocks in any further analysis.

This finding is not entirely surprising given that the South African equity market relative to developed equity markets is highly concentrated; with the largest stock in the SWIX 40 Index holding a weight of above 10%, while the top ten stocks in the index represent more than 40%. This concentration problem poses a challenge for the construction of well diversified portfolios, a fundamental principle of good portfolio management.

The estimated Error Correction Model was as follows:

Estimation Equation:

D(LNINDEX) = C(1) + C(2)*D(LNNPN) + C(3)*D(LNSAB) + C(4)*D(LNCFR) + C(5)*D(LNBIL) + C(6)*D(LNBTI) + C(7)*D(LNSOL) + C(8)*D(LNSNH) + C(9)*D(LNMTN) + C(10)*D(LNOML) + C(11)*D(LNAGL) + C(12)*U(-1)

Substituted Coefficients:

D(LNINDEX) = 0.0001195 + 0.1919650*D(LNNPN) + 0.1678463*D(LNSAB) + 0.0846344*D(LNCFR) + 0.0815268*D(LNBIL) + 0.0542859*D(LNBTI) + 0.0580172*D(LNSOL) +0.0737897*D(LNSNH) + 0.0723338*D(LNMTN) + 0.0883251*D(LNOML) + 0.0302058*D(LNAGL) - 0.0201771*U(-1)

where D denotes the first difference of the index and various stocks as dependent and independent variables, 0.1919650 is the intercept and U(-1) named the error correction term, is the one period lag of the residuals from the cointegration regression model.

The above equation outlines the short run coefficients of the various stocks in the tracking portfolio with 10 stocks. All the above short-run coefficients proved to be significant variables with p-values of 0.0000 being less than 5% in explaining the index movement. The error correction term corrects the disequilibrium of the system at a speed of adjustment rate of -2.02% daily and is a significant variable. The negative and significant error correction term validates that there is indeed a long run equilibrium relationship between the index and the tracking portfolio.

4.3.2 Differential returns between the tracking strategy and the index

On a cumulative-since-inception basis the Tracking Portfolio underperformed the reconstructed index by -46%, whereas on an annual basis the underperformance is on average -2.55%. Refer to Table 1 and Figure 2 below. The difference in performance between the Tracking Portfolio and the Index widens in the out of sample period. More information on performance is available in Appendix 3.

Table 2: Descriptive Statistics of Reconstructed Index and Tracking Portfolio

	Return	Std Dev	Average Drawdown	Information Ratio (arith)	Tracking Error	Sharpe Ratio	Skewness	Kurtosis	Correlation
Reconstructed Index	22.55	13.00	-5.07	—	0.00	1.69	0.46	-0.24	1.00
Tracking Portfolio 10	20.00	19.25	-11.49	-0.93	18.21	0.75	-0.02	0.06	0.48



Figure 6: Investment Growth

Closer examination of the comparative performance reveals that 2013 was the year with the most significant tracking error between the two of 41%, followed by 17% in 2014. The Tracking Portfolio is highly concentrated, consisting of only 10 stocks out of the 40 stocks in the Index. Further investigation into the third quarter of 2013, also revealed that the huge disconnect between the Tracking Portfolio and the Index was largely attributable to a combination of short to low weight positions in some good performing stocks, and high weight positions to some lower performing stocks. This caused the Tracking Portfolio to lag that of the Index. In addition, when looking at a rolling 1 year, 1 month shift the underperformance during 2013 and 2014 reverses beyond 07/2014 with the Tracking Portfolio subsequently starting to outperform the Index.

4.3.3 Sharpe Ratio and Information Ratio

To assess the risk adjusted performance, this study calculated both the Sharpe Ratio and Information Ratios. Over the period of 02/11/2009 to 31/12/2015 as per Table 1 above, the calculated Sharpe Ratio and Information Ratio for the Index are 1.69 and N/A that of the Tracking Portfolio are 0.75 and -0.93 respectively. While both Sharpe Ratios numbers are

positive, the Index appears to have better rewarded investors for every additional unit of risk taken than the Tracking Portfolio. This is further confirmed by the negative information ratio in the case of the Tracking Portfolio.

4.3.4 Transaction costs

Transaction costs were found to be negligible ranging between 0.0197 and 4.527 in Rand terms per transaction value, thus having minimal impact on the returns of the Tracking Portfolio. These negligible amounts are largely attributable to there being really minor changes in the monthly weights of the 10 stocks within the Tracking Portfolio, as well as the fact that it is a highly-concentrated portfolio. In addition, these 10 stocks represent the most liquid stocks in the South African Equity market within the JSE Top 40 Index. The finding that transaction costs can be minimal with this classical index tracking strategy is a positive.

4.3.5 Volatility

This study also computed the unconditional (historical) 21 day volatilities, see Figure 7.

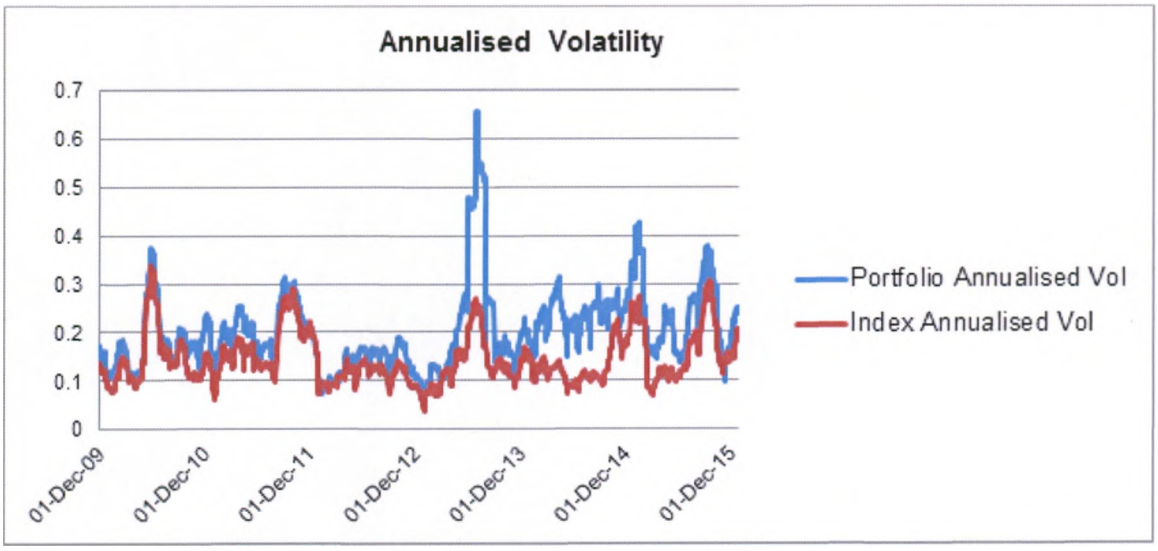


Figure 7: Annualised Volatility

The Tracking Portfolio displays the same annualised volatility pattern as the index in sample, this is despite that the tracking portfolio is only composed of 10 constituents of the index. Out of sample the tracking portfolio shoots up to a volatility level above 0.6 and subsequently adjusts lower although still remaining at slightly higher volatility levels than the index. The exacerbated volatility of the Tracking Portfolio relative to the Index in the out of sample period, is due to the Tracking Portfolio having higher weightings to stocks that individually exhibited higher volatility than the index in the period between mid 2013 to mid 2015, a

consequence of concentration risk. Some of the stocks that actually doubled in volatility during this time period were the likes of Cie Financiere Richemont SA, Steinhoff International Holdings NV, Naspers Ltd, Anglo American PLC and SABMiller PLC. From a stock specific weightings; an example of a notable difference in weight is the Cie Financiere Richemont SA stock that ranges between 11% and 22% in the tracking portfolio over time, vs ±7.4% in the Reconstructed Index. In order to mitigate concentration risk, more stocks (to the extent that the Tracking Portfolio remains cointegrated to the Index) can be incorporated into the tracking portfolio for diversification purposes. The annualised unconditional volatility of the Tracking Portfolio ranges from 0.65 to 0.1.

4.3.6 Correlation

Figure 4 below depicts the correlation between the Tracking Portfolio and the Index. It can be concluded that the Tracking Portfolio and the Index are generally positively correlated, with the correlation ranging between 0.2 and 1. The out of sample periods are characterised by periods of low correlation, again mainly attributable to the factors explained in the performance divergence and volatility discussion above.

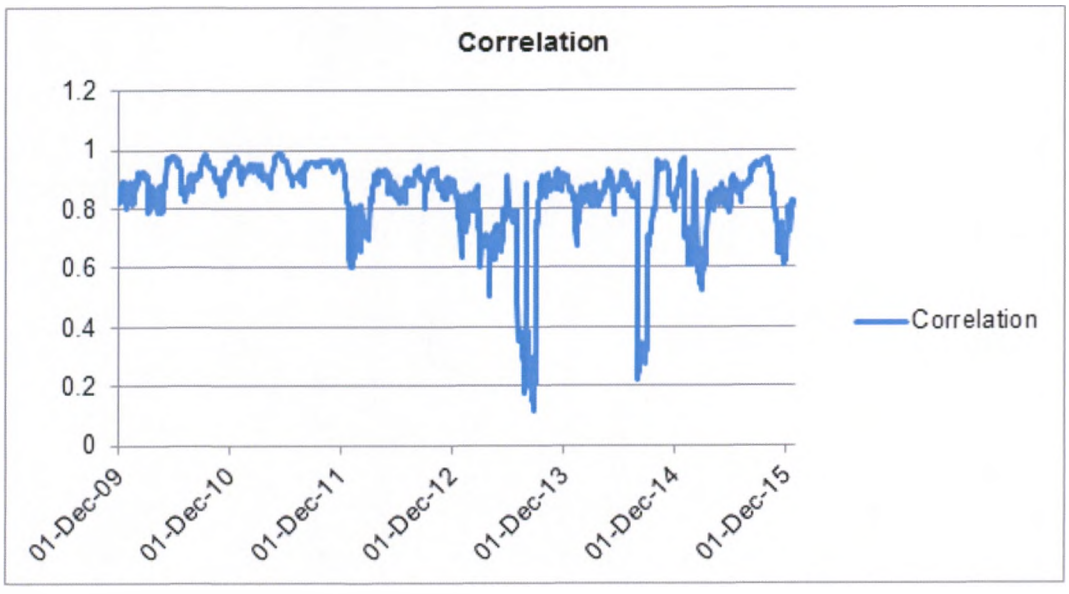


Figure 8: Graph showing correlations

4.4 Back-testing the Enhanced Index Tracking Strategy

4.4.1 Rolling window EG co-integration test

4.4.1.1 Engel-Granger cointegration test

In order to analyse the long-short market neutral strategy outlined in 2.3.2 above the same principle used for the classical index tracking analysis, were applied. This time cointegration was tested for using portfolios consisting of 10, 15, 20, 25 and 30 stocks, even though it was only the portfolio with 10 stocks that proved to be sufficiently cointegrated with the Index under the analysis in section 4.3.1. These portfolios would track an enhanced index of benchmark +/-5% annual return. A calibration period of 3 years was used.

The residuals from each EG cointegration regression were tested for stationarity applying the Engel-Granger methodology for cointegration testing. The analysis revealed that the Long/short portfolio consisting of 10 stocks tracking an artificially reconstructed index of index -/5% remained cointegrated mainly at an unsatisfactory 10% confidence level. The portfolio remained cointegrated 95% and 82% of the time, however only at the 10% and 5% significance levels respectively. This is not statistically significant throughout the period under review. The outcome of the rolling EG cointegration test is depicted in Figure 9.

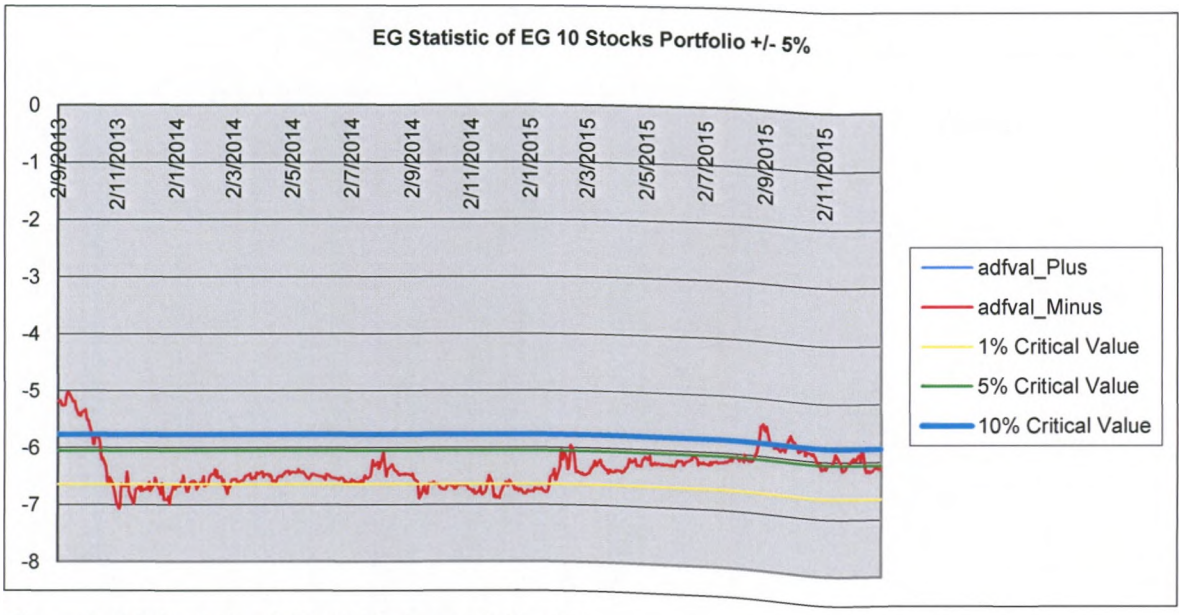


Figure 9: EG Statistic of EG 10 Stocks portfolio ±5%

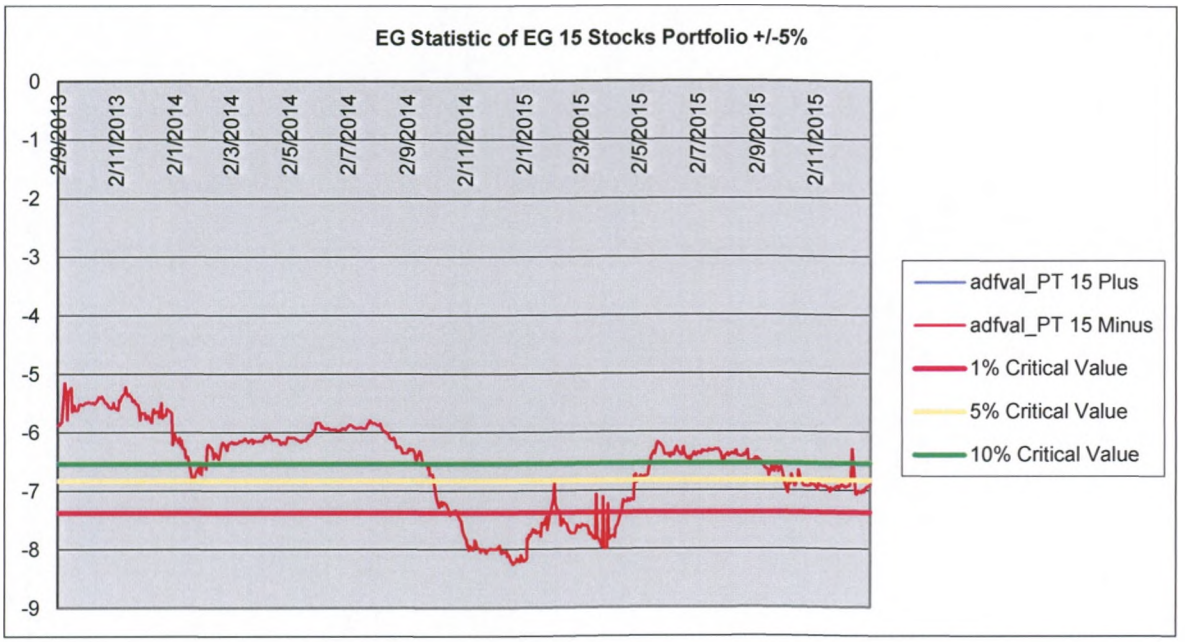


Figure 10: EG Statistic of EG 15 Stocks Portfolio ±5%

The 15 stock portfolio is interestingly contrary to the rest of the portfolios, appearing not to be cointegrated in the in sample period and then goes through periods of cointegration mainly at a 90% confidence level and no cointegration in the out of sample period.

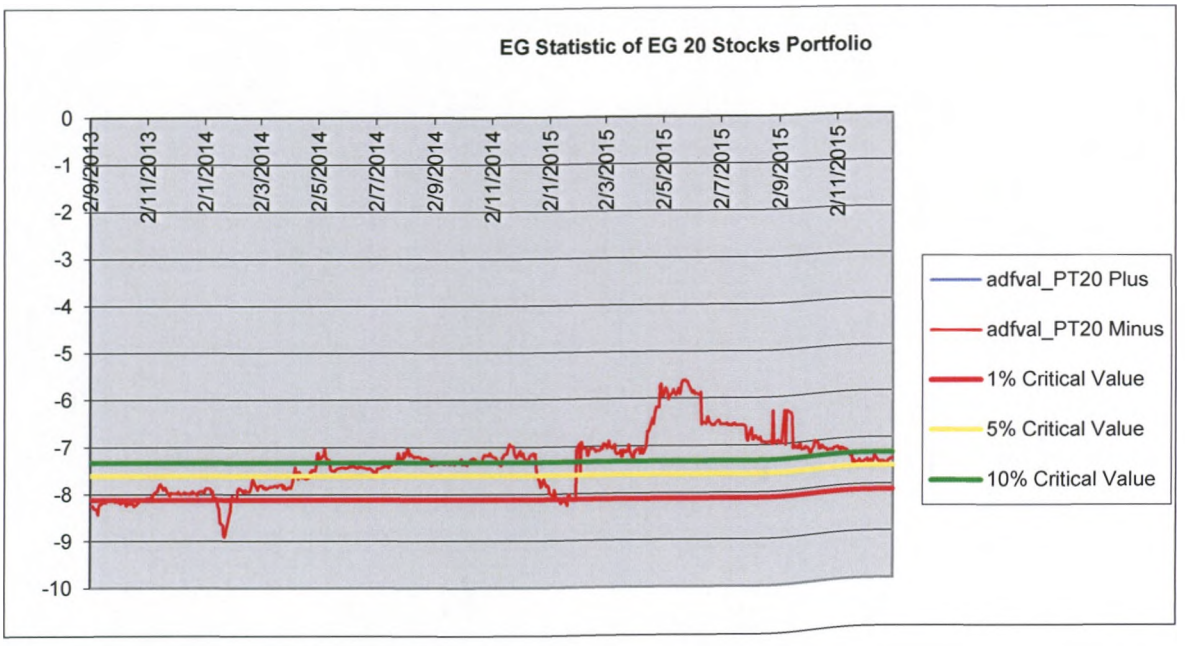


Figure 11: EG Statistic of EG 20 Stocks Portfolio ±5%

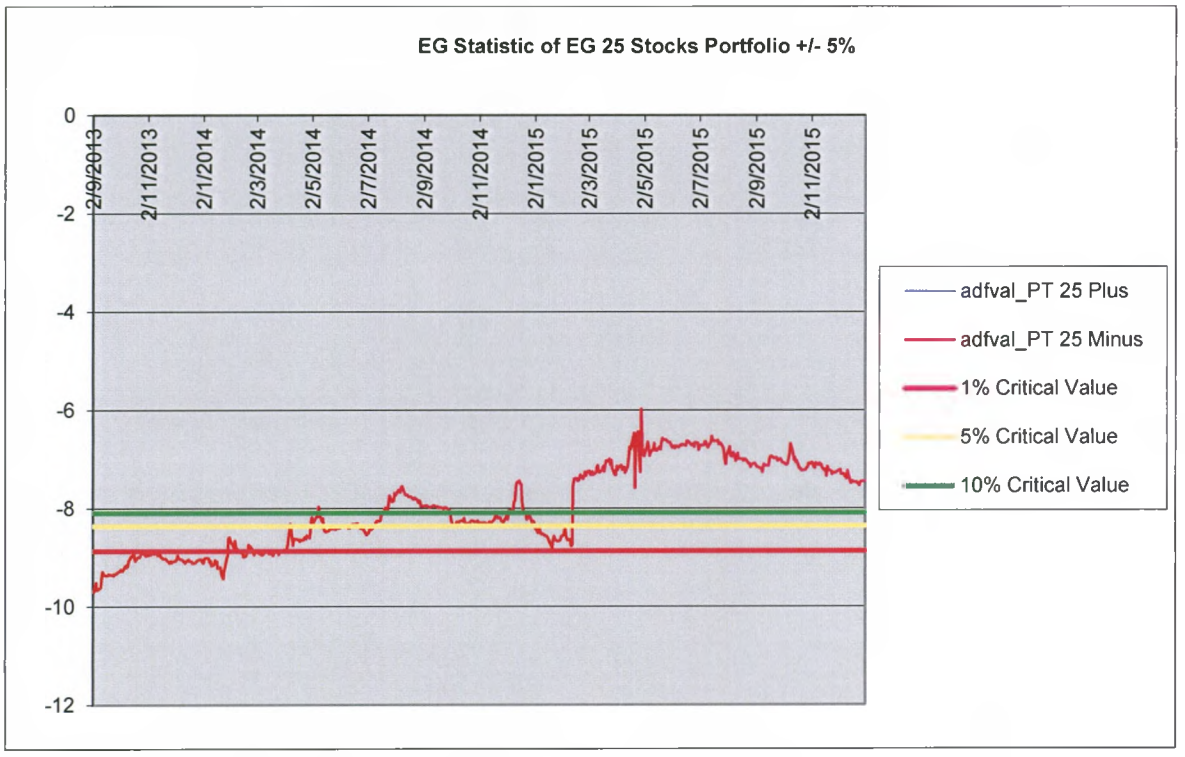


Figure 12: EG Statistics of 25 EG Stocks Portfolio ±5%

Observing figure 11 and figure 12, both the 20 and 25 stock portfolios yield almost exact opposite results to the 15 stocks portfolio, depicting some cointegration however mainly at unsatisfactory 90% and 95% levels in the in sample period and then goes through periods of no cointegration in the out of sample period, with the 20 stocks portfolio again moving towards slight cointegration towards the end of the calibration period.

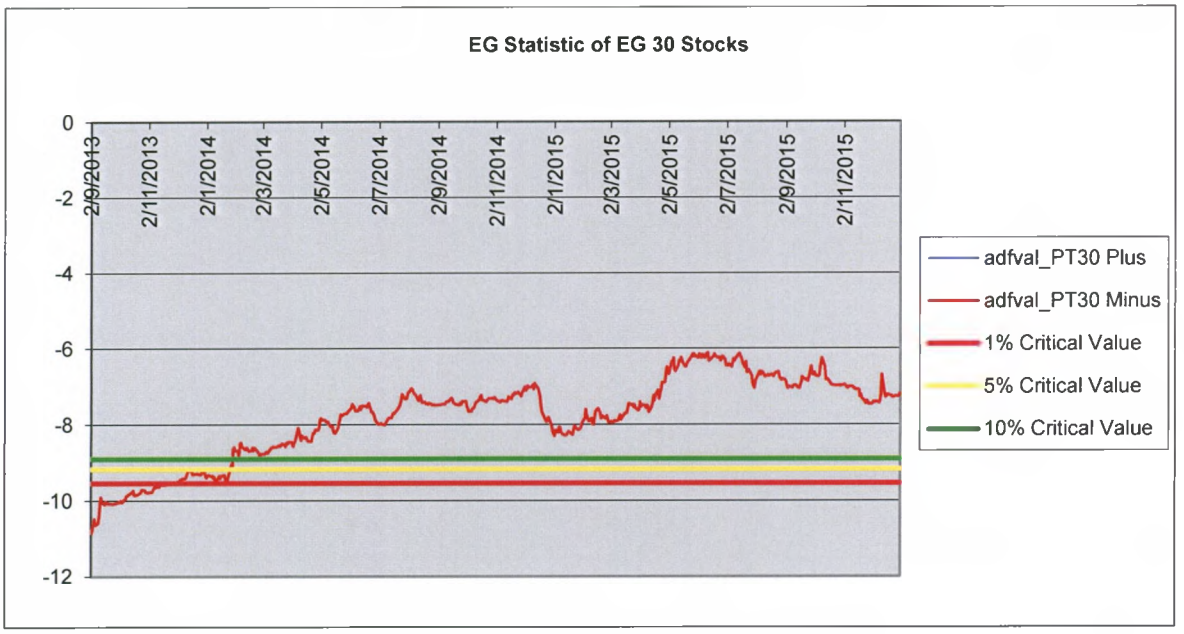


Figure 13: EG Statistic of EG 30 Stocks Portfolio ±5%

As per figure 13, the null hypothesis of no cointegration could not be rejected for most of the calibration period for the portfolio with 30 stocks.

Based on the above results it can be concluded that no arbitrage portfolio was found to be sufficiently cointegrated with an enhanced index. In fact, it appears that as the number of stocks increase in the tracking portfolio, the less cointegration there is with the enhanced index.

Additionally, this study has found that the portfolio weights for the portfolios tracking an enhanced index did not differ much to the weights of the portfolios in the classical index tracking scenario. This resulted in there being no difference in the prices and returns of the Enhanced Tracking Portfolio relative to the Classical Tracking Portfolio.

Given that this study has not been able to find a tracking portfolio that is sufficiently cointegrated with an enhanced index, further analysis will not be attempted.

5 CHAPTER 5: CONCLUSIONS AND PROSPECTS FOR FUTURE RESEARCH

This paper set out to analyse the extent to which the cointegration technique based on long term relationships of variables which may applied for portfolio construction in the South African Equites market. To this end, this study attempted to implement and test two types of strategies: Classical Index Tracking and Enhanced Index Tracking.

By applying cointegration techniques this study was able to find that, within the South African Equities market, a Tracking Portfolio consisting of only 10 constituents of the JSE Top 40 Index could sufficiently remain cointegrated with the index. In practice this would result in a highly-concentrated portfolio, thus going against the principles of good portfolio construction such as diversifying well for risk management. As is expected from such a concentrated portfolio the volatility was at times higher than that of the index. The Tracking Portfolio was also at times found to be lowly correlated with the Index. The challenge though with the determined Tracking Portfolio was that the returns were found to underperform the benchmark returns before transaction costs, an issue already currently faced by traditional classical index tracking strategies in practice.

Monthly rebalancing was applied and it was found that negligible transaction costs were incurred, which would suggest that with quarterly rebalancing the transaction costs would possibly reduce further as a result of less frequent rebalancing. This is a benefit of tracking strategies based on long run equilibrium relationship presented by cointegration.The outcome of minimal transaction costs calculated was positive for the classical index tracking strategy given that the Tracking Portfolio returns already lagged the Index returns before accounting for transaction costs.

The second strategy applied was Enhanced Index Tracking, with the aim of producing a positive return above the index using long and short portfolios. An index was enhanced by adding and subtracting to each day's return a constant return and then tried to track the artificial index similar to the classical index tracking analysis.This study was not able to find a tracking portfolio that was sufficiently cointegrated with an enhanced artificial index, and

thus could not carry on any further analysis of this strategy. Contrary to Alexander and Dimitriu (2002, 2004b) and Dunis and Ho (2005), but similar to Maurer (2008) for enhanced indexation strategies based on cointegration, no convincing empirical evidence was found for the South African equities market.

Some limitations of the study as a result of adopting simplistic approaches are discussed. The degree of cointegration, and consequently the tracking performance will highly dependant on the stock selection approach employed. As stock selection was not the focus of this study, a simplistic stock selection method was applied to construct the two tracking strategies by ranking the stocks according to their weight in the index at moment of the portfolio construction. This is not necessarily an optimal selection approach and therefore future studies should involve employing more sophisticated stock selection approaches. Other simplified approaches adopted in the study was the pre selection of 10, 15, 20, 25 and 30 of the stocks in the JSE Top 40 index, with the stocks remaining the same but only the weights being adjusted at points of rebalancing; a review of the tracking portfolio constituents and not just of the relative weights at rebalancing points would likely have yielded better results for the index tracking application.

Further research should be done to test the robustness of the cointegration based index tracking strategies under different market conditions and its applicability within sophisticated trading strategies and stock selection methodologies.

The conclusion of this study is that, it seems possible to track an index, albeit lagging the performance of such index using only a subset of its coefficients. Construction of tracking portfolios on the basis of cointegration can add economic value for South African investors and investment managers, and thus offer a worthwhile alternative for portfolio construction.

REFERENCES

Acosta-Gonzalez, E., Armas-Herrera, R., & Fernadez-Rodriguez, F. (2015).On The Index Tracking and The Statistical Arbitrage Choosing The Stocks By Means of Cointegration: The Role of Stock Picking. *Quantitative Finance, 15*(6), 1075-1091.

Alexander, C. (1999). Optimal Hedging Using Cointegration. *Philosophical Transactions of The Royal Society of London A:Mathematical, Physical and Engineering Sciences,* 357(1758), 2039-2058.

Alexander, C., & Dimitriu, A. (2002). The Cointegration Alpha: Enhanced Index Tracking and Long-Short Equity Market Neutral Strategies. *ISMA Discussion Papers in Finance*.

Alexander, C., & Dimitriu, A. (2004a). Equity indexing: Optimize your passive investments. *Quantitative Finance, 4*(3), 30-33.

Alexander, C., & Dimitriu, A. (2004b). A Comparison of Cointegration and Tracking Error Models for Mutual Funds and Hedge Funds. *ISMA Centre Discussion Papers in Finance, 4*, 1-26.

Alexander, C., & Dimitriu, A. (2005a). Indexing and Statistical Arbitrage: Tracking Error or Cointegration. *Journal of Portfolio Management, 31*(2), 50-63.

Alexander, C., & Dimitriu, A. (2005b). Indexation, cointegration and equity market regimes. *International Journal of Finance and Economics,* 213-231.

Alexander, C., Giblin, I., & Weddington, W. (2002). Cointegration and Asset Allocation: A New Active Hedge Fund Strategy. *Financial Risk and Financial Risk Management*, 16, 65-89.

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *Journal of Business*, 45, 444-454.

Blasdell, S., & Zingone, R. (2010, March). *Enhanced Indexation Targeting the Return/Risk "Sweets Spot".* USA: Insights. Accessed from https://careers.jpmorganchase.com/jpmpdf/1158630201479.pdf

Brink, M.E. (2004). *The Performance Of South African unit trusts for the period 1984 to 2003*, Doctoral Dissertation, Stellenbosch University, South Africa.

Canakgoz, N. A, Beasley, J. E. (2008). Mixed-integer Programming Approaches For Index Tracking and Enhanced Indexation. *European Journal of Operational Research*, 196(1), 384-399.

Dunis, C. L., & Ho, R. (2005). Cointegration portfolios of european equities for index tracking and market neutral strategies. *Journal of Asset Management, 6*(1), 33-52.

Dunis, C. L., Giorgioni, G., Laws, J., & Rudy, J. (2010). *Statistical arbitrage and high-frequency data with an application to Eurostoxx 50 equities.* Liverpool Business School, Working paper.

DiBartolomeo, D. (2000). *The enhanced index fund as an alternative to indexed equity management.* Boston: Northfield Information Services.

Engel, R. F., & Granger, C. W. J. (1987). Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica: Journal of the Econometric Society*, 55(2), 251–276.

Engel, R. F., & Yoo, B. S. (1987). Forecasting and Testing in Co-integrated Systems. *Journal of Econometrics*, 35(1), 143–159.

etfSA. (2015, May). *The Home of Exchange Traded Funds*. Retrieved from http://www.etfsa.co.za

Furelid, J. (2015). *The Market Neutral Strategy – An Introduction. Hedgenordic.* Retrieved from http://hedgenordic.com.

Gabriel, I. (2002). *Portfolio Construction Strategies Using Cointegration.* Masters Dissertation, Academy of Economic Studies, Doctoral School of Finance and Banking, Bucharest, Romania.

Granger, C. W. J. (1981). Some Properties of Time Series Data and Their Use in Econometric Model Specification. *Journal of Econometrics*, 16, 121-30.

Hendry, D. F., & Juselius, K. (2000). Explaining Cointegration Analysis: Part 2. The *Energy Journal*, 22(1), 75-120.

Holland, J. H. (1975). Adaptation in Natural and Artificial Systems. An introductory Analysis with Application to Biology, Control, and Artificial Intelligence. *Ann Arbor, MI: University of Michigan Press.*

Johansen, S. (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control*, 12(2-3), 231-254.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica: Journal of the Econometric Society,* 59, 1551-1581.

JP Morgan. (1994). *RiskMetrics Technical Document*, Morgan Guaranty Trust Company, New York.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risk Investments in Stock Portfolios and Capital Budgets. *Review of Economics and Statistics*, 4, 13-37.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments.* John Wiley&Sons, New York.

Maurer, T. A. (2008). *Cointegration in Finance: An Application to Index Tracking*, http://www.thomasmaurer.net/thomas_maurer_net_homepage/cointegration_in finance.pdf, 29 - 30

Mazviona, B.W, & Nyangara, D. (2013). A Test Of The Weak Form Efficiency Of The Zimbabwe Stock Exchange After Currency Reform. *International Journal of Business, Economics and Law,* 2(2), 43-48.

Minor, D. (2001). Beware of Index Fund Fundamentalists. *Journal of Portfolio Management*, *27*(4), 45-50.

PWC ETF 2020: Preparing for a new horizon, (2015, January 25). Accessed from: http://www.pwc.com/gx/en/industries/financial-services/asset-management/publications/pdfs/etf-2020-exchange-traded-funds-pwc.pdf

Roman, D, Mitra, G & Zverovich, V. (2013). Enhanced Indexation based on second-order stochastic dominance. *European Journal of Operational Research*, *228*(1), 273-281.

Sharpe, W. (1991). The Arithmetic of Active Management. *Financial Analyst Journal*, *47*(1), 7-9.

Sharpe,W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, 19(3), 425-442.

Schoenfeld, S. & Yang, J. (2003, October 01). Enhanced or Active? Precisely Defining Enhanced Indexing and Risk-Controlled Active Management. Accessed from : http://www.etf.com/publications/journalofindexes/joi-articles/1594.html?nopaging=1.

Scowcroft, A. & Sefton, J. (2003). Enhanced Indexation. In S. Satchel and A. Scowcroft (Eds.), *Advances in Portfolio Construction and Implementation*, 95 – 124. Elsevier, Butterworth-Heinemann Finance.

Ssekuma, R. (2011). A Study of Cointegrating Models with Applications. Doctoral Dissertation, University of South Africa, South Africa.

Taleb, N. N. (2007). The Black Swan: The Impact of the Highly Improbable. The Random House Publishing Group, 132-156

Thomaidis, N. S. (2013). On The Application of Cointegration Analysis in Enhanced Indexing. *Applied Economic Letters*, 20(4), 391-396.

Utkulu, U. (2012). How to estimate long-run relationship in economics: An Overview of Recent Developments.

APPENDIX 1: JSE TOP 40 CONSTITUENTS

Ticker	Name
NPN SJ Equity	Naspers Ltd
SAB SJ Equity	SABMiller PLC
CFR SJ Equity	Cie Financiere Richemont SA
BIL SJ Equity	BHP Billiton PLC
BTI SJ Equity	British American Tobacco PLC
SOL SJ Equity	Sasol Ltd
SNH SJ Equity	Steinhoff International Holdings NV
MTN SJ Equity	MTN Group Ltd
OML SJ Equity	Old Mutual PLC
AGL SJ Equity	Anglo American PLC
SBK SJ Equity	Standard Bank Group Ltd
FSR SJ Equity	FirstRand Ltd
REM SJ Equity	Remgro Ltd
MNP SJ Equity	Mondi PLC
SLM SJ Equity	Sanlam Ltd
APN SJ Equity	Aspen Pharmacare Holdings Ltd
BID SJ Equity	Bid Corp Ltd
ANG SJ Equity	AngloGold Ashanti Ltd
WHL SJ Equity	Woolworths Holdings Ltd/South Africa
MEI SJ Equity	Mediclinic International PLC
SHP SJ Equity	Shoprite Holdings Ltd
ITU SJ Equity	Intu Properties PLC
INP SJ Equity	Investec PLC
BGA SJ Equity	Barclays Africa Group Ltd
GRT SJ Equity	Growthpoint Properties Ltd
TBS SJ Equity	Tiger Brands Ltd
VOD SJ Equity	Vodacom Group Ltd
REI SJ Equity	Reinet Investments SCA
BAT SJ Equity	Brait SE
RDF SJ Equity	Redefine Properties Ltd
NTC SJ Equity	Netcare Ltd
MRP SJ Equity	Mr Price Group Ltd
BVT SJ Equity	Bidvest Group Ltd/The
DSY SJ Equity	Discovery Ltd
MND SJ Equity	Mondi Ltd
RMH SJ Equity	RMB Holdings Ltd
NED SJ Equity	Nedbank Group Ltd
CPI SJ Equity	Capitec Bank Holdings Ltd
RMI SJ Equity	Rand Merchant Investment Holdings Ltd
INL SJ Equity	Investec Ltd
CCO SJ Equity	Capital & Counties Properties PLC

FFB SJ Equity	Fortress Income Fund Ltd
AMS SJ Equity	Anglo American Platinum Ltd
FFA SJ Equity	Fortress Income Fund Ltd

APPENDIX 2 RETURNS

Figure 1: YTD, 1,3 and 5 years Returns

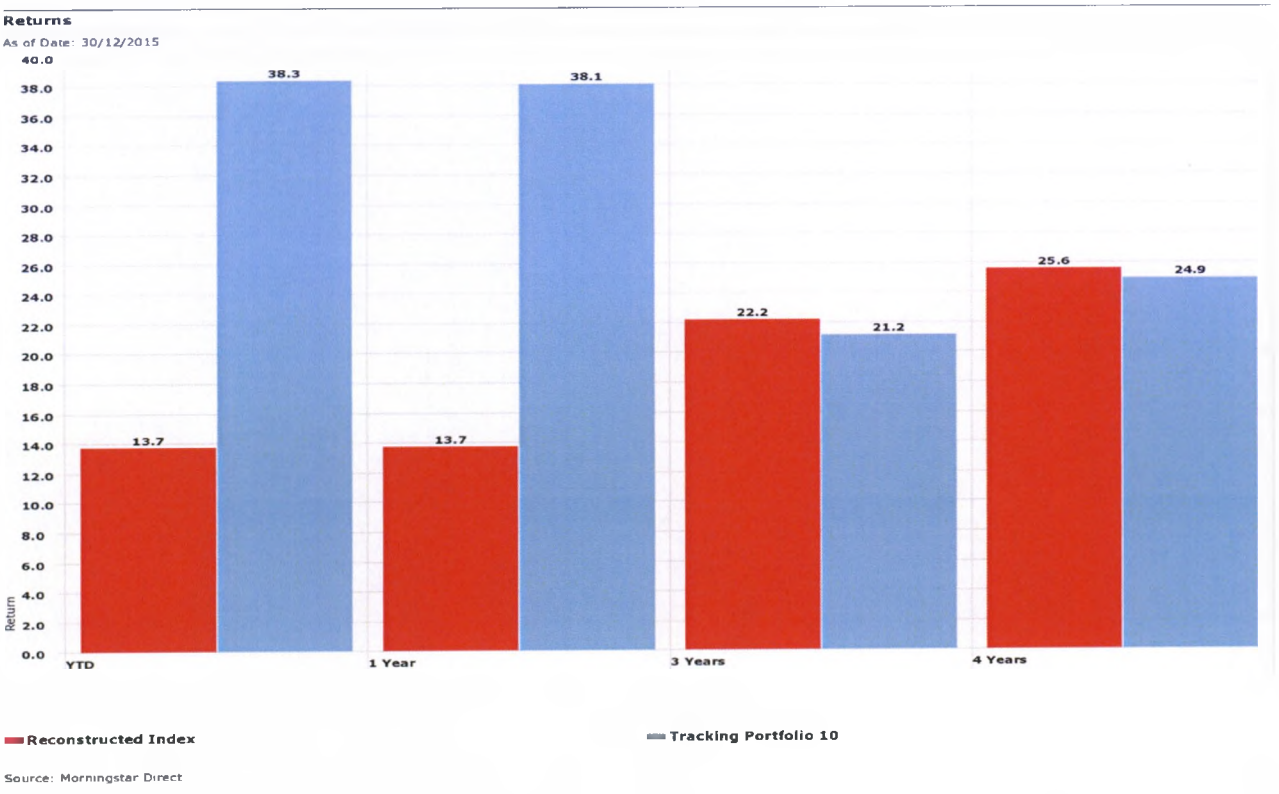
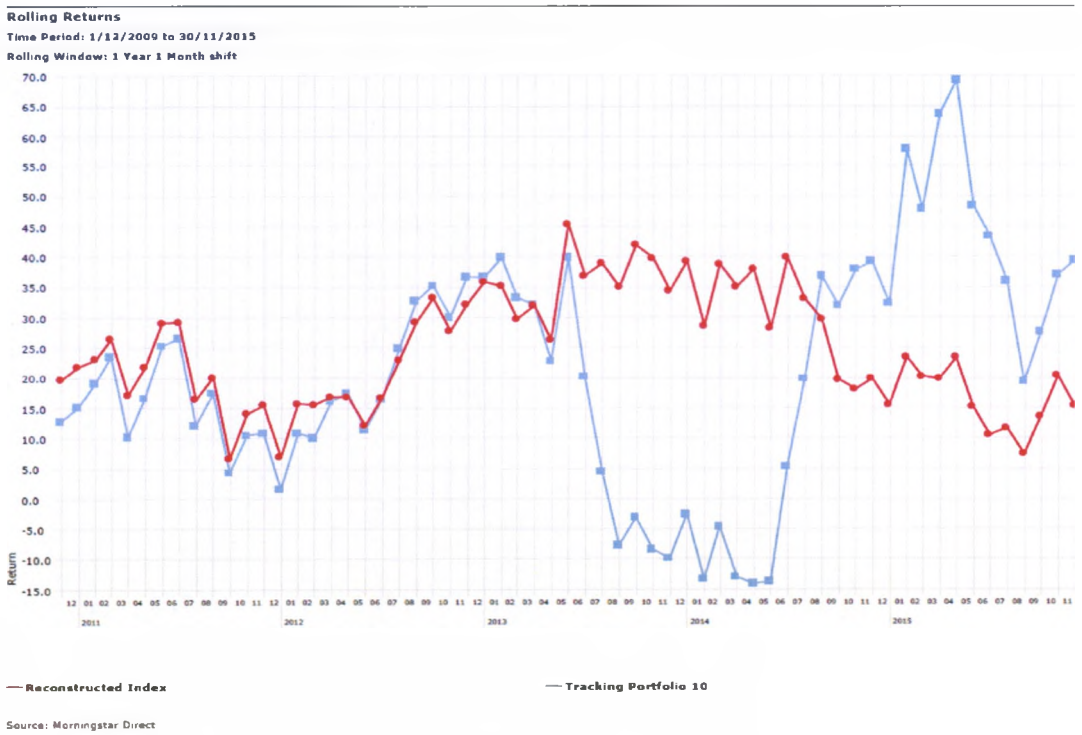


Figure 2: Rolling Window Returns



Source: Morningstar Direct

Table 2: Monthly Returns

	Tracking Portfolio 10	Reconstructed Index	Difference
2011-01	-2.43	-1.87	-0.56
2011-02	4.81	3.69	1.13
2011-03	-3.91	-0.56	-3.35
2011-04	4.73	3.43	1.30
2011-05	1.83	1.09	0.74
2011-06	-4.32	-2.70	-1.62
2011-07	-2.31	-1.34	-0.97
2011-08	0.09	0.03	0.06
2011-09	-1.50	-2.34	0.85
2011-10	9.92	9.78	0.13
2011-11	-1.97	0.15	-2.12
2011-12	-2.41	-2.01	-0.40
2012-01	6.49	6.23	0.26
2012-02	4.02	3.52	0.50
2012-03	1.41	0.54	0.88
2012-04	5.91	3.47	2.44
2012-05	-3.44	-3.03	-0.41
2012-06	0.07	1.25	-1.18
2012-07	4.71	3.93	0.78
2012-08	6.48	5.23	1.25
2012-09	0.35	0.63	-0.28
2012-10	5.69	5.30	0.40
2012-11	3.03	3.62	-0.59
2012-12	-2.49	0.73	-3.22
2013-01	9.03	5.68	3.35
2013-02	-0.88	-0.65	-0.23
2013-03	0.50	2.28	-1.78
2013-04	-1.52	-0.98	-0.54
2013-05	10.04	11.68	-1.63
2013-06	-14.11	-4.69	-9.41

	Tracking Portfolio 10	Reconstructed Index	Difference
2013-07	-8.91	5.45	-14.36
2013-08	-5.97	2.29	-8.26
2013-09	5.35	5.77	-0.42
2013-10	-0.11	3.70	-3.81
2013-11	1.36	-0.35	1.71
2013-12	5.51	4.37	1.15
2014-01	-3.05	-2.40	-0.66
2014-02	9.07	7.18	1.89
2014-03	-8.27	-0.42	-7.85
2014-04	-2.75	1.10	-3.86
2014-05	10.47	3.85	6.62
2014-06	4.90	3.94	0.96
2014-07	3.60	0.33	3.28
2014-08	7.29	-0.33	7.62
2014-09	1.70	-2.34	4.04
2014-10	4.33	2.28	2.05
2014-11	2.32	1.10	1.22
2014-12	0.27	0.64	-0.37
2015-01	15.64	4.30	11.34
2015-02	2.26	4.36	-2.10
2015-03	1.46	-0.79	2.25
2015-04	0.63	4.10	-3.47
2015-05	-3.12	-3.04	-0.08
2015-06	1.31	-0.27	1.59
2015-07	-1.75	1.37	-3.12
2015-08	-5.88	-4.14	-1.74
2015-09	8.76	3.23	5.52
2015-10	12.06	8.41	3.65
2015-11	4.06	-3.10	7.15