Performance replication of the Spot Energy Index with optimal equity portfolio selection: evidence from the UK, US and Brazilian markets.

Kostas Andriosopoulos[1]* and **Nikos Nomikos**[2]

[1]ESCP Europe Business School, London, UK

[2]Cass Business School, City University, London, UK

December, 2011

Abstract

This paper reproduces the performance of a geometric average Spot Energy Index by investing only in a subset of stocks from the Dow Jones Composite Average, the FTSE 100 and Bovespa Composite indexes, and in two pools including only stocks of the energy sector from the US and the UK respectively. Daily data are used and the index-tracking problem for passive investment is addressed with two innovative evolutionary algorithms; the differential evolution algorithm and the genetic algorithm, respectively. The performance of the suggested investment strategy is tested under three different scenarios: buy-and-hold, quarterly, and monthly rebalancing; accounting for transaction costs where necessary.

JEL Classifications: C6, C7, F37, F21, G11, G15, Q4

Keywords: Energy markets, International equities, International investment, Index Tracking, Evolutionary Algorithms.

***Corresponding Author:** Kostas Andriosopoulos, ESCP Europe Business School, 527 Finchley Road, London NW3 7BG, UK, Tel: +44 (0) 207 443 8809, Fax: +44 (0) 207 443 8874, E-mail: kandriosopoulos@escpeurope.eu.

†We would like to thank the participants of the 33rd International Conference of the International Association for Energy Economics (IAEE) in Rio de Janeiro, Brazil and Juliana Caicedo-Liano, Duncan Shand of Black Rock, Thomas Chiang and other participants of the 3rd Emerging Markets Finance Conference of the Emerging Markets Group at Cass Business School, London, UK for their helpful comments. In addition, we would like to thank Michael Doumpos for kindly providing some of the routines used.

1. Introduction

It is well documented in the literature that investors can benefit by getting exposure in commodities as part of their long-term asset allocation plan. Over the past decade impressive gains have been witnessed in commodity prices, with this pattern accelerating in the last few years. The aforementioned trend, along with empirical evidence supporting the idea that passive strategies are better than active ones (see Konno and Hatagi, 2005; Frino and Gallagher, 2001; Barber and Odean, 2000; Sorenson et al., 1998; Malkiel, 1995; Sharpe, 1991 among others), especially in the longer term, made passive strategies increasingly popular. One of the most popular forms of passive trading strategies is index tracking (Beasley et al., 2003); a method that attempts to replicate/ reproduce the performance of an index. This has attracted investors' attention and led to an impressive growth of index investing in the commodity markets. In general there are three major ways of investing in a commodity index; first, by choosing an index and replicating it by following the related Rule Book; second, by investing in a fund that replicates the chosen index; finally, the most popular approach lately is by buying shares of an Exchange Traded Fund (ETF) that its strategy is to follow the respective commodity index. This trend has been recognised by investors and prompted them to set-up the first commodity ETF in November 2004[1]. As of January 2010 the market capitalization of that first commodity ETF was exceeding 39 billion US dollars, competing with numerous other commodity-related ETFs established since then. Many other ETFs investing in physical commodities, futures, and commodity-related equities, have followed since then.

Generally, commodities are seen as a hedge against inflation (Bodie, 1983; Gorton and Rouwenhorst, 2006). Though currently inflation is relatively low and stable, mounting worries about potential inflation pressures moving forward can be enticing more investors to the commodities market. In addition, since most energy commodities and especially crude oil are quoted in US dollars, any weakening of the USD against an international basket of major currencies and especially the euro, leads to an appreciation of the energy commodities in dollar terms. This happens on the one hand because demand is global, taking place in an international market scene, reflecting global currency prices, and on the other hand because these energy

[1] The first listed commodity ETF was the streetTRACKS Gold Shares ETF, with its sole assets being gold bullion and from time to time cash.

commodities are used by investors as a hedge against further US dollar weakness and other floating currencies. Moreover, the long lead times to bring additional capacity to satisfy the newly created excess demand for energy commodities, driven by the billions of people entering the global consumer economy, will attract even more investors to the energy commodity markets going forward.

There are many papers applying various momentum and market timing strategies to commodity futures markets, with the findings in the literature suggesting that there is mixed evidence on their performance (see for example, Miffre and Rallis, 2007; Alizadeh et al., 2008; Marshall et al., 2008; Szakmary et al., 2010). In addition, there is a plethora of studies focusing on the effects of oil price changes on the economy (Hamilton, 2003), on whether oil price risk is priced in stock markets (Jones and Kaul, 1996), and whether oil prices forecast future stock market returns (Driesprong et al., 2008). However, the question whether returns of equity portfolios can be used to replicate the performance of physical energy price returns, aggregated in a portfolio and proxied by a spot index, has received almost no attention in the existing literature.

The aim of this paper is to replicate the unique price/ return behaviour of direct energy commodity investment using equities. The proposed approach is based on previous research findings that in the case of equally weighted long-only portfolios of commodity futures, with a changing composition over the studied period, their statistically significant returns are similar to those of stocks (Bodie and Rosansky, 1980; Fama and French, 1987; Gorton and Rouwenhorst, 2006). In addition, it is documented in the literature that after the 2000s, commodities have gone through a financialization process, exposing them to the wider financial shocks (Tang and Xiong, 2010). The goal is accomplished by applying two very efficient in terms of tracking error strategies, the Differential Evolution Algorithm (DE) and the Genetic Algorithm (GA), to solve the index tracking problem in the energy markets as represented by the constructed Spot Energy Index (hereafter named SEI). Low tracking error strategies provide several advantages to investors; they result in better diversified portfolios, make the long-only constraint of a fund manager less binding, and in general tend to provide higher returns for various equity strategies. As of 2005, more than 50% of the trading volume on NYSE was performed using some form of program trading strategies (Lamle and Martell, 2005).

More specifically, the performance of the SEI is reproduced by investing in a small basket of stocks picked either from the stocks comprising three well known financial indexes, or from two pools of energy-only related stocks. In particular, the cases of the US, UK and Brazilian investors are considered under the assumption that they want to invest in the SEI and prefer to access only their local stock markets due to cost savings and/or better knowledge of the respective markets. They represent two developed and one developing stock market, with the latter having its unique energy significance in the global scene. The recent reforms and regulations that took place in Brazil brought transparency, sophistication and additional liquidity to its financial markets. It is this reliability in the Brazilian stock market data that led to the selection of this market for testing and implementing the proposed investment strategy. The lack of transparency and liquidity in other emerging stock markets, which have a large number of commodity related firms listed, as for example in Russia, can be questionable as it could lead to obscure datasets. In addition, while recently many developed countries have sputtered amid weak economic growth, Brazil has continued to thrive, given its rich reserve of natural resources and growing middle class, becoming the fifth-largest economy in the world.

In addition, it is well documented in the literature that energy prices affect national economies and have a different impact on the various business sectors. As Hammoudeh et al. (2004) point out in their study, the oil related industries are amongst the most affected sectors, with higher oil prices having a positive impact on most companies. Oil, and in effect energy prices, affect companies' earnings and their bottom lines, thus having an immediate effect on their stock prices. Hence, based on intuition and previous research findings, the two pools of energy-only related stocks used in the analysis should perform very well in tracking the SEI. Moreover, the three non-energy specific stock pools are used as a relative performance measure, as there is a possibility that the stocks of various companies operating in other, non-energy related industries to be directly affected by the movements in energy prices, thus making them a good selection for constructing the portfolios that track the SEI. The methodology implemented can track the SEI or any other benchmark index by investing in a basket of stocks that each of the evolutionary algorithms will determine. Baskets of maximum 10, 15 and 20 stocks are selected from the following stock pools: Dow Jones Composite Average, FTSE 100, Bovespa Composite, and two unique pools of energy-only related stocks from the US and the UK stock markets respectively. The proposed methodology allows investors to be more comfortable with their investment selection since this is drawn out of a stock market that they are more familiar with.

Hence, the first contribution of this paper in the literature is that the index tracking problem in the energy commodities market is addressed and both the DE and GA are applied. Second, investors are provided with the opportunity to invest in the energy spot markets by choosing stocks from a specific domestic equity market which could appeal more to their investing criteria/ preferences. Third, by tracking the performance of the energy sector with stocks selected by two innovative evolutionary algorithms, a cost effective implementation and true investability is promoted for the popular segment of energy style investors. Barberis and Sheleifer (2003) argue that style investing is attractive mostly because of the fact that institutional investors act as fiduciaries and thus they must follow systematic rules of portfolio allocation, and because of its simplified performance evaluation process. However, there are many funds that cannot invest in commodities directly as in the case of pension funds, where governments in their effort to protect peoples' savings strictly regulate the industry by placing stringent restrictions on the types of assets held. Usually futures contracts and other derivative products in alternative investments such as commodities are excluded from their portfolios (Nijman and Swinkels, 2003). Nevertheless, by following the proposed investment strategy and investing in stock portfolios selected by the evolutionary algorithms used in this paper, these funds could now participate in the energy markets by investing in an ETF that would track the performance of the SEI.

Fourth, given the importance of equities in a multi-asset class portfolio, by choosing those stocks that can track the SEI, the selected equity portfolios are indirectly insulated from inflation; a key consideration among investors and fund managers in an uncertain economic environment. In their investigation over the period 1972-2001, Nijman and Swinkels (2003) find that investors with liabilities indexed to the interest rate and inflation, such as insurance companies and pension funds, can significantly increase their risk-return trade-off through commodity investment because of the positive relation of commodities with inflation. Fifth, it is the first time that a broad energy index incorporates in its calculation electricity market prices, thus reflecting the full spectrum of energy commodities and their by-products besides the commonly used crude oil and its refined fuels. Finally, this paper contributes to the existing literature by investigating three different investment strategies during the three year out-of-sample period, buy-and-hold, quarterly, and monthly rebalancing; accounting for transaction costs where necessary.

The findings of this paper have several positive implications for investors. Although the SEI represents the economic importance of the energy group of commodities to the global economy it primarily serves as a performance benchmark, given the limited ability for a direct investment. However, the proposed approach provides investors with an option to track the performance of this Spot Energy Index using a basket of equities that are liquid and fully investable. This allows investors to get closer to the underlying commodity market price trends, something they cannot achieve using a futures price index. Historically, futures index returns have lagged price index returns, with this decoupling of performance being a constant frustration for index investors. For comparison reasons the performance of two well established energy excess return indexes is reported, namely the Dow Jones–UBS Energy Sub-Index and the Roger’s Energy Commodity Index, against the performance of the constructed SEI and the selected portfolios.

Adding to the aforementioned, the proposed investment strategy provides a low cost – compared to actively managed funds – means of accessing the energy spot markets. In particular, sector rotation investment managers can benefit from the findings of this paper. By tactically shifting assets, they can over- or under-weigh specific sectors according to their due diligence, economic outlook or market objective. Diversification is another important implication. Instead of taking concentrated risks by purchasing individual stocks, the investors can own our proposed baskets and at the same time avoid the diligent attention that individual stocks require. Furthermore, investors who on the one hand want to participate in the performance of the volatile spot energy sector, but on the other hand do not want the high risk exposure of holding the individual energy commodity, can invest in the selected stock baskets that exhibit substantially lower volatility. Finally, investors that cannot physically hold the energy commodities can benefit from the selected equity baskets that allow for both long and short position to be taken. Most commodity trading advisors and commodity pool operators use investment strategies that can be long-only or systematic long/short, using leverage to take the short positions. The latter strategy assumes that investors take opposite positions than those taken by commercial hedgers (Jaeger et al., 2002). So an effective index tracking strategy, as the one proposed in this paper, should allow for both the replication of the performance of the benchmark index, and the implementation of this long/short strategy that can significantly improve the risk/ return profile of traditional asset portfolios.

The structure of this paper is as follows. Section 2 presents a literature review on energy commodity investing; the various energy indexes in existence and the relation between commodities and equities. Section 3 gives an explanation of the constructed spot energy index and the data used in the analysis. In section 4, the DE and GA evolutionary algorithms are explained, with the problem formulation also being described. Section 5 offers the empirical results of the study and, finally, section 6 concludes the paper.

2. Energy commodity investing

2.1. Energy indexes

There are two ways of investing in energy commodities. The first is the direct physical investment that includes all relevant costs for maintaining and managing the inventory. The second is the indirect investment via equity or debt ownership of energy companies and utilities, engaged in oil exploration, production, refining, marketing etc. However, in recent years there has been an increasing number of direct energy commodity-based products available to investors such as the respective energy futures contracts that require constant active management, and the energy commodity indexes. There is a large number of mutual funds, hedge funds, Exchange Traded Funds (ETFs), Exchange Traded Notes (ETNs) and OTC return swaps that follow the energy sector through index investing. In fact, in the US alone, assets allocated to commodity index strategies via futures contracts has risen from $13 billion in 2003 to $260 billion as of March 2008, with an estimated 70 percent of these funds invested in the energy sector (Hamilton, 2009b). From the total of commodity index investing in the US exchanges, about 42% is conducted by institutional investors (pension and endowment funds), 25% by retail investors (ETFs, ETNs and similar exchange-traded products), 24% by index funds (a client/

counterparty with a fiduciary obligation to match or track the performance of a commodity index), and 9% by Sovereign wealth funds (CFTC, 2008).

Commodity indexes attempt to replicate the returns equivalent to holding long positions in various commodities markets without having to actively manage the positions. Being uncorrelated with the returns of traditional assets such as stocks and bonds, commodity index investments' returns provide a significant opportunity to reduce the risk of traditional investment portfolios; thus explaining the economic rationale for including a commodity index investment in institutional portfolios such as those of pension funds and university endowments. Currently there are more than ten publicly available futures' indexes, with different risk and return profiles, offering exposure to commodity markets; each of these indexes also offers specific exposure to certain commodity sectors via their traded sub-indexes. The variations in commodity index performance across indexes and during different market conditions lie with the differences in the construction methodology of each index. The main differentiations relate to the index sectors' composition, constituent commodities selection, rolling and rebalancing strategy, which are both crucial and apply only for futures indexes, and the methodology used for calculating the constituents' respective weights. The later has been an important determinant of the indexes' performance, especially with the recently large weight allocations towards the energy sector across all indexes (AIA, 2008). This remark strengthens the approach of this paper that focuses only on the energy sector which has recently drawn the most activity in index investing. Another issue that complicates the historical analysis of commodity futures index returns is the lack of a universal way to define their composition, because commodities cannot have a market capitalization-based portfolio weighting scheme; because at any time, the value of all open long futures contracts is offset by the value of the open short futures contracts (Black, 1976).

There are several risks and disadvantages associated with futures' based commodity indexes. In the case of a futures index, unlike a passive equity portfolio which entitles the holder to a continuing stake in a company, commodity futures contracts specify a certain date for the delivery of the physical commodity. In order to avoid the delivery process and maintain a long futures position, a passive futures portfolio requires regular transactions; nearby contracts must be sold and contracts with later deliveries must be purchased. This process is referred to as “rolling”. The difference between the prices of the two contracts, the nearby and the more distant delivery one, is called the “roll yield”. Even though the term structure of commodity prices has historically been an important driver of realised commodity futures’ excess returns, there is no guarantee that the term structure will remain the same in the future. Also, there is a possibility that the futures term structure of an individual commodity be, on average, in backwardation, yet the particular contract that an index mechanically rolls into might be in contango. When commodity markets are in contango this could result in negative roll yields that would adversely affect the value of the futures index. These negative roll yields can significantly decrease the value of the futures index over time when the nearby contracts or spot prices of the underlying commodities are stable or increasing. Also, in the opposite scenario of decreasing spot prices, the value of the futures index can significantly decrease when some or all of the constituent commodities are in backwardation.

Furthermore, although most of the energy commodities have liquid futures contracts with expiration every month, there are some that expire less frequently, thus rolling forward can be more costly and vulnerable to longer duration and smaller liquidity. Moreover, Gorton and Rouwenhorst (2006) find that commodity futures contracts become illiquid in the delivery month as most traders avoid delivery of the physical commodities. In addition, the explicit rolling procedure that needs to be used when tracking a commodity futures index is another major disadvantage. Any transparent commodity futures index publishes the specific rules of rebalancing making them available to all market participants. This means that other traders and speculators can take advantage of these known future transactions mandated by those rules. Under the prevailing trend of these index funds to constantly grow in size, they will only become more vulnerable to such trading exploitation.

In addition, external market and macroeconomic factors can have a major impact on a futures index. The market prices of the index’s components may rapidly fluctuate due to changes in supply and demand relationships, and due to other numerous factors such as weather, major political and economic events, technological developments, fiscal and monetary programs. Recently, even the performance of the equities markets has become a significant factor affecting the performance of commodity indexes, especially when the index holds large positions of illiquid contracts or maturities. It has been observed that during periods of steep equity market movements there is a tendency of aggressive buying or selling of commodity indexes (Tang and Xiong, 2010). Investors tend to rebalance the mix of their portfolios between equities and commodities, either for hedging or speculating purposes, or because of their view of the market being short- or long-term. Kyle and Xiong (2001), argue that investors with a short term strategy trade more aggressively against noise trading than those with a long term strategy. All these factors can affect the spot prices of the physical commodities, the underlying of the futures contracts, causing the prices and the volatilities of the components of the index to fluctuate in inconsistent directions and at inconsistent rates. This could quickly lead specific trades against the investor, resulting in a loss of the initial deposit required before being able to close the position.

Moreover, suspension or disruptions of market trading in the commodities futures markets could adversely affect the value of a futures index. Such events that disrupt the functionality of the futures markets, like lack of liquidity, replacement or delisting of a futures contract, changes in the quality specifications of the underlying physical commodities, increased participation of speculators, governmental regulation and intervention, adversely affect a futures commodity index. In fact, the recent increase in volume on the buy side of the futures contracts, in its major part to support index investing, is argued that has an apparent effect on commodity prices drifting them away from their fundamental value and creating a speculative price bubble; a conclusion that can lead to increased government regulation on futures markets. Hamilton (2009a) suggests that speculative investing in oil futures contracts contributed to the oil shock of 2007-08. The steep decline in short-term interest rates in 2008 resulted in negative real interest rates that in turn attracted a great deal of investment in physical commodities, and thus fuelled commodity speculation, especially for crude oil and other energy products (Frankel, 2008).

One can argue that this financialization of commodities introduced a speculative bubble in the price of physical energy commodities, especially crude oil, which subsequently burst. Moreover, in the case of pension funds where governments in their effort to protect people's savings strictly regulate the industry, there are stringent restrictions on the types of assets held by a fund. Usually, futures contracts and other derivative products in alternative investments such as commodities are excluded from their portfolios (Nijman and Swinkels, 2003). Speculation in the commodities markets has been in the centre of a heated debate in the past few years amongst industry and policy circles, on whether it is the driver of excessive increases and the resulted excessive price volatility in the energy and food markets. Following these debates, there have been increasing calls for a more stringent supervision of the energy markets, and in particular for their paper markets, from both the industry's bodies as well as international governments.

The abovementioned risks and disruptions can be avoided when following the investment strategy proposed, by using as a performance benchmark for the energy markets the SEI which allows investors to get closer to the underlying commodity price trends, and by investing in the selected equity portfolios. Using the evolutionary algorithms and the methodology suggested in this paper, stock investors can optimally select their portfolios for tracking the SEI without spending time, effort, and money, trying to identify which stocks can simultaneously act as a profitable investment and a good commodity play.

2.2. Commodities and their relation to equities

Kilian (2009) finds that all major real oil price increases since the mid-1970s can be attributed to increases in global aggregate and/or oil-specific demand, and much less to disruptions of crude oil production. Even when political events affect the oil prices, like the Persian Gulf War, it is mostly the increased sudden demand for oil, triggered by fears for the future oil supply, which drives oil prices and not the actual disruptions in oil supply. In the same lines, Hamilton (2009a) finds that the run-up in oil prices of 2007-08 should be attributed to the strong demand for crude oil in combination with a stagnating world production. From an asset-only perspective, previous research suggests that depending on investors risk tolerance, commodities as proxied by cash-collateralized commodity futures, should be about a quarter of investors' portfolios in their strategic, long-term, asset allocation (Anson, 1999; Jensen et al., 2000).

In addition, Hong et al. (2007) argue that the returns of a number of industry stock portfolios, including that of petroleum, which are informative about macroeconomic fundamentals, can forecast the returns of the aggregate stock market with a lead of up to two months. They also find that high returns for some industries, including that of petroleum, mean bad news for future economic activity and the aggregate stock market. In addition, Driesprong et al. (2008) find that a rise in oil prices significantly lowers future stock market returns, especially for the markets of those countries classified as net energy importers, and the world market index. They also suggest that investors tend to underestimate the direct economic effect of oil price changes on the economy and thus act with a delay. Their conclusion is strengthened by the fact that this under-reaction is less pronounced in the oil-related equity sectors, where market players are more informed and aware of the economic consequences of oil price changes.

Findings by Erb and Harvey (2006) suggest that portfolios of commodity futures can have equity-like returns if a high enough diversification return can be achieved, or if the portfolio exposures are skewed toward contracts that are more likely to have positive roll or spot returns in the future[2]. Gorton and Rouwenhorst (2006) construct a fully-collateralized commodity futures index and conclude that historically, between 1959 and 2004, their index has a similar risk/ return performance to equities, using the S&P500 as a proxy. They also find that correlation between the returns of stocks and bonds and those of the commodity futures is negative; a conclusion that can be attributed to the different behaviour that the various asset classes exhibit over the business cycle. In contrast, Schneeweis and Spurgin (1997) conclude that over the period January 1987 to February 1995, commodity and managed futures indexes have sources of risk and return that are distinct from indexes of traditional assets such as stocks and bonds. Nonetheless, they also find that the unique construction methodology of each index results in differential return correlation with alternative assets, making each index very useful as a performance benchmark for unique portfolios.

Research evidence suggests that before the 2000s commodity indexes had negative correlation with equities, e.g., Greer (2000), Gorton and Rouwenhorst (2006), and Erb and Harvey (2006). However, after the 2000s, commodities were heavily promoted as a new asset class, with various instruments based on commodity indexes attracting billions of dollars from wealthy individuals and institutions, resulting in a financialization process that exposed commodities to the wider shocks of financial markets, as shown in Tang and Xiong (2010). The latter authors also find that

[2] The diversification return is defined as the synergistic benefit of combining two or more assets to reduce variance, enhanced when the portfolio is rebalanced. Roll returns can originate from an upward- or downward-sloping term structure of the individual futures prices.

this exposure gradually increased, especially after 2004, with the spill-over effects of the recent financial crisis contributing to the subsequent large increase of commodity price volatility. Equities and other financial assets mainly derive their value from future cash flows, whereas commodities, being real assets, derive their value from physical supply and demand conditions. Despite this fundamental difference between equities and commodities, the need of commodity producers and consumers to share price risk with the broader investment community was the main driver of the resulted integration of commodities and financial markets.

Why, especially in recent years, are commodities expected to behave more like financial assets? This question can be answered with the following arguments: First, taking into consideration that commodity index investors have a big impact into commodities prices it can be assumed that the remaining participants, such as commercial hedgers and speculators, cannot fully absorb the price impact (Tang and Xiong, 2010). Second, it is known that any shocks affecting the market-wide risk premium, subsequently, affect all financial assets to a varying degree (e.g., Cambell and Cochrane, 1999). It is thus valid to argue that, as commodities become more and more integrated with the financial markets, they should also be affected. Third, when price shocks in one asset occur, by rebalancing his/ her portfolio, the shocks spill-over to the other assets that the marginal investor holds (Kyle and Xiong, 2001). Hence, commodity index investors that usually hold additionally large positions in stocks are exposed to stock market shocks when they reallocate their funds between commodities and stocks. Fourth, Barberis and Shleifer (2003) find that each asset of a certain class is exposed to shock spillovers from other assets in the same class. Therefore, according to Tang and Xiong (2010), individual commodities' prices are exposed to both the shocks to those commodities that participate in the indexes held by index investors, and, to a certain degree, the shocks to off-index commodities. Finally, all non-US commodity index investors are also exposed to exchange rate shocks, as all commodity indexes are denominated in US dollars.

When making portfolio allocation decisions, most investors categorize assets into broad categories called styles (Barberis and Sheleifer, 2003). Stocks within a particular country, index or industry, value stocks or growth stocks, can all be considered as style examples. While some styles persist over the years, such as government bonds, financial innovation guarantees the appearance of new styles, as is the case for instance with mortgage-backed securities. Simplification and performance evaluation are the two main reasons that individual and institutional investors follow style investing[3]. The former makes the processing of vast amounts of information relatively easy and efficient, whereas the latter can help evaluate money managers relative to a performance benchmark specific to their style (Sharpe, 1992). Energy commodity investing could be considered as a new style investment, with a plethora of funds and ETFs that track passive benchmarks of commodity and energy sector equity indexes. The work of this paper could motivate investors, private and institutional, to follow the international energy industry, a sector that deserves sole attention. The potential benefits of commodity investments for institutions date at least back to Bodie (1980), and especially in the case of insurance companies and pension funds these benefits are recently pointed out in Nijman and Swinkels (2003). Many new energy commodity ETFs and ETNs[4] have come to the market, making it easier for a retail investor to obtain exposure to commodities. There are various types of these Energy Index Funds either based on the construction type of the fund (single- or multi-contract, long-only or bearish[5]), or based on the energy sector they track (broad energy or sector specific).

These tracking funds have a number of advantages over traditional debt instruments (notes, bonds, certificates). They offer less expensive and less risky investment products, while at the same time providing protection against inflation. Also, they can provide easy access to a broad range of investors, a simple way to manage accounting and disclosure procedures, and can lead to fewer taxes since in many countries index fund returns are treated as capital gains and not as income. An energy ETF can be used by the energy industry market players to complete parts of their existing portfolio or to perform tactical strategies. They can be used for hedging energy investment risk, portfolio diversification, or as a control measure of inflation exposure. To that end, the proposed methodology offers an effective, and at the same time inexpensive way to operate such a fund, giving the full flexibility of any investment style, long or short, that equities can provide.

[3] Style investing is particularly attractive to institutional investors because acting as fiduciaries they must follow systematic rules of portfolio allocation (Barberis and Shleifer, 2003).
[4] An ETN, although it is structured similar to an ETF, exposes the investor to counterparty risk making it a much riskier investment.
[5] Bearish Energy Index Funds have the same structure as bullish (long-only) funds with the major difference that investors are not only allowed to buy the fund, but also to put on a short position (sell the fund).

Investors that want exposure to spot commodities returns, usually cannot invest in the actual physical products, besides the case of precious metals, and thus seek alternative approaches such as commodity futures and commodity-related equities. However, although commodity futures provide exposure to their respective underlying commodity, as their prices converge to the spot prices on a monthly basis, the link between long-term commodities futures and spot returns is distorted because of the effect it has on the term structure the prevailing backwardation or contango. This effect has been more profound in recent years, since 2004, when contango started prevailing in the energy markets. Commodity equities on the other hand overcome these term structure effects, with relevant research showing a direct and powerful link between the returns of commodity-related equities and their business-related spot commodity prices.

On that note, empirical evidence shows that commodity-market returns are very similar to equity-market returns in terms of magnitude, with equity-like risk (Bodie and Rosansky, 1980; Nash, 2001). The latter finding has recently increased the interest from institutional investors to integrate commodities in their strategic asset allocation and to develop tactical asset allocation strategies. Nijman and Swinkels (2003) test a tactical switching strategy between commodities and stocks and they find that commodity investments can be beneficial to pension funds within a mean-variance framework. Vrugt et al. (2004) use a market timing strategy based on a dynamic multi-factor approach, to forecast monthly commodity returns with a broad range of indicators related to the business cycle, the monetary environment, and the general market sentiment; they find that investors can have superior returns when following their timing asset allocation strategy. It is evident in the literature that up until the early 2000s, commodities and commodity funds perform well during a financial market downturn, while having at the same time a lower correlation to equities (Chow et al., 1999; Edwards and Caglayan, 2001), with energy commodities in specific being consistently negatively correlated to equities. As Till and Eagleeye (2003) conclude, whenever a commodity investment is intended to act as a diversifier for equities it needs to be heavily weighted in energy markets, as it is the energy complex that exhibits a persistent negative correlation to equities.

Investors generally expect that futures indexes are a good proxy for a spot index, because of the high correlation between spot and futures prices. However, this is not entirely true as according to Chada (2010) the Spot Commodity Index used in his paper outpaces the respective Commodity Futures Index by over 5.6 percent per year, even though their correlation is exceeding 99 percent. The correlation measure is not the most important factor for determining which is the best investment alternative, as it only measures the degree to which two variables are likely to move together. It does not provide an adequate measure of the magnitude of the moves, and it also fails to capture the overall trend of the variables' returns over time, especially as those returns compound. A risk-adjusted return measure, as the Information Ratio, is a better and more appropriate performance measure. In addition, long-only futures commodity indexes have little protection against any sudden and large in magnitude downward price spikes, as they have no ability to sell short, they have inherent limitations based on the state of the futures curve (backwardation or contango), and most of them rebalance only once a year. Furthermore, investing in a broad commodity futures index does not reflect any short-term, tactical response to prices, in either the individual constituents or the aggregate commodity market, which can be better captured by investing in a specific segment of the commodities markets, such as the energy sector.

Investing in commodity-related[6] equities is considered to be the best alternative for avoiding some of the inefficiencies of futures returns, outlined previously, as they can play a crucial part in providing exposure to the commodity markets. Some argue that investing in commodities equities is primarily an investment in equities, which does not significantly help to reduce the overall volatility of the portfolio, or improve its risk-adjusted returns. The main concern of the advocates of this argument is that commodities equities are subject to the actions of their company's management in the same manner as for all other equities, which implies that they can destroy shareholder value or break the link between these stocks and the underlying commodities' price movements. Although the aforementioned argument can be valid in some instances, it is generally accepted that commodity equities are not too far removed from the actual commodity, as the value of a commodities company is directly tied to the value of the

[6] Commodity-related equities are the securities of those companies that are mainly engaged in the production and distribution of commodities and commodities-related products, the so-called pure-play companies.

commodities it produces/ trades. The latter could be justified by the fact that the equity markets of Russia, Brazil and other emerging market countries that their economies depend heavily on commodities, and more specific on energy commodities, and thus have a large number of commodity related listed stocks, have witnessed a thriving performance during every recent commodities boom. Moreover, there are plenty of strategies, and their related opportunities, connected to energy production, distribution, and trade finance that are not directly available to futures investors, irrespectively of their approach, passive or active. These opportunities can only be available to investors via the equities markets, as part of the respective companies' valuation.

In general, any increase in the underlying commodity price should result in an increase in the company's earnings, leading into an increase in shareholder value, which in turn is reflected in the share price. Chada (2010) constructs an equally-weighted portfolio of the eight largest energy stocks as of December 2009, and then maps the aggregated changes in revenues and earnings of these stocks with changes in the WTI spot oil price. He concludes that earnings of oil companies tend to generally relate to the spot price of oil, tracking it closely both in up and down markets. Building on the aforementioned, it is believed that tracking the performance of spot energy prices, as proxied by the proposed in this paper Spot Energy Index (SEI), can be best achieved by optimally selecting portfolios of stocks, and most probably from energy-related stock pools. With such an investment approach, commodity investors can have all the means at their disposal to protect against any sudden downward price movements, that investing in the selected equities portfolios can deliver, and thus can capture all the alpha opportunities that a passive futures index would miss.

3. Benchmark energy index, spot and equity data

Because centralized trading lacks for many commodities, the most reliable spot prices are for those that trade active and liquid futures contracts, since these are typically used as a pricing benchmark. In the case of the energy commodities, the NYMEX is the world's largest futures exchange. Initially, a spot price energy index is constructed, constituted by daily prices of the following six energy commodities that also trade futures contracts on the NYMEX[7]:

[7] The main reason for selecting these energy commodities that trade futures contracts on the NYMEX is that since most energy commodity futures markets are denominated in US dollars, the indexes constituted mostly by local US

1. Heating Oil, New York Harbour No.2 Fuel Oil, quoted in US Dollar Cents/Gallon[8] (US C/Gal); hereafter named as "HO";
2. Crude Oil, West Texas Intermediate (WTI) Spot Cushing, quoted in US Dollars/Barrel (US$/BBL); hereafter named as "WTI";
3. Gasoline, New York Harbour Reformulated Blendstock for Oxygen Blending (RBOB), quoted in US C/Gal; hereafter named as "Gasoline";
4. Natural Gas, Henry Hub, quoted in US Dollars/Milion British Thermal Units (US$/MMBTU); hereafter named as "NG";
5. Propane, Mont Belvieu Texas, quoted in US C/Gal; hereafter named as "Propane";
6. PJM, Interconnection Electricity Firm On Peak Price Index, quoted in US Dollars/Megawatt hour (US $/Mwh); hereafter named as "PJM".

All six energy commodities that are included in the index, as a result of large volume daily trading of standardization qualities, serve as indicators of impeding changes in business activity as they are sensitive to factors affecting both current and future economic conditions. The Spot Energy Index (SEI) is constructed as an un-weighted geometric average of the individual commodity ratios of current prices to the base period prices, set at January 31, 2006 until February 1, 2010. The base date for the SEI is the same date that the equity sample is obtained. Considering that the boom in commodity index investing is a relatively new phenomenon, recent data are utilized to test the proposed investment strategy. The index's construction methodology is similar to that of the world-renowned CRB Spot Commodity Index. The SEI is designed to offer a timely and accurate representation of a long-only investment in energy commodities using a transparent and disciplined calculation.

Geometric averaging provides a broad-based exposure to the six energy commodities, since no single commodity dominates the index. It also helps increase the index diversification by giving even to the smallest commodity within the basket a reasonably significant weight. Gordon (2006)

commodities will have a smaller currency exposure when the commodity is produced and delivered in the US. In the case that the marginal buyer of the underlying commodity is outside the US, then the return to holding that commodity has a large currency exposure. Additional reasons for the commodities' selection are the following: 1) Quality standardization so that uniform price quotations can be obtained, 2) High trading volume in an open market, 3) Sensitive to changing market conditions.

[8] Fuel Oil and Gasoline spot prices that are quoted in US C/gallon are converted into US $/Barrel, taking into account that there are 42 gallons in one barrel and 100 cents per dollar.

finds that a geometrically weighted index is preferred to alternative weighting schemes, because the daily rebalancing allows the index not to become over- or, under-weighted. This avoids the risks that other types of indexes are subject to, like potential errors in data sources for production, consumption, liquidity, or other errors that could affect the component weights of the index. Furthermore, through geometric averaging the SEI is continuously rebalanced which means that the index constantly decreases (increases) its exposure to the commodity markets that gain (decline) in value, thus avoiding the domination of extreme price movements of individual commodities. As Erb and Harvey (2006) point out, the indexes that rebalance annually eventually become trend followers because commodity prices movements constantly change the weightings, whereas those that rebalance daily stay closer to the original intent of the index. In addition, Nathan (2004) shows that the indexes that use geometric rebalancing, and thus rebalance their weightings daily, generally exhibit lower volatility.

The mathematical specification used to calculate the geometric average Spot Energy Index (SEI) is the following:

$$\mathrm{SEI}_t = \left(\prod_{i=1}^{n} \frac{P_t^n}{P_0^n} \right)^{\frac{1}{n}} \times 100 = \sqrt[n]{\frac{P_t^1}{P_0^1} \times \frac{P_t^2}{P_0^2} \times \ldots \frac{P_t^n}{P_0^n}} \times 100,\ \mathrm{n} = 1, 2, ..6. \tag{1}$$

where, SEI_t is the index for any given day, n represents each one of the six commodities comprising the index, P_t^n is the price of each commodity for any given day, and P_0^n is the average (geometric) price of each commodity in the base period.

The SEI provides a stable benchmark so that end-users can be confident that historical performance data is based on a structure that resembles to both the current and future composition of the index; making SEI suitable for institutional investment strategies. The stable composition of the index is an important element, because when the composition of an index changes over time, the average return of the index does not equal the return of the average index constituent, especially when indexes are equally weighted. The latter makes historical index performance a bad proxy to prospective index returns, thus distorting the information that investors seek (Erb and Harvey, 2006). Moreover, it is a better means for evaluating the movement in energy commodity prices because it is based on spot prices and not on highly volatile prices for future delivery which are subject to contango and backwardation. The SEI is the best indicator of the activity and the trend prevailing in the energy markets, and thus by default provides a gauge of world growth and any potential inflationary pressures. Both private and institutional investors can use the SEI to track its performance, or as a benchmark for actively or passively managed portfolios. In addition, there could be numerous other ways to invest in the SEI such as OTC swaps, structured notes or products offered by third-party asset managers that provide energy commodity exposure benchmarked on the index.

As far as the equity data sample is concerned, it includes daily prices for stocks that are picked from the Dow Jones Composite Average, FTSE 100 and Bovespa Composite indexes; representing two developed, and one developing stock market with a distinct significance in the global energy scene. The index is also tracked with portfolios that include stocks from a unique pool of energy related stocks from the US and the UK stock markets, respectively. These two energy related equity pools are used because according to Scholtens and Wang (2008) oil related firms' earnings are more likely to be affected by changes in oil prices, as explained by the highly significant estimated coefficients of the earnings-to-price factor returns for their total oil firms' sample. After employing a multi-factor APT model, Al-Mudhaf and Goodwin (1993) find that oil price changes in a period surrounding the 1973 oil shock can explain the return differences in 29 US oil companies that they examine. In addition, Boyer and Filion (2007) with their APT model also find that stock returns of Canadian oil and gas companies have a significant relationship with oil price changes. The selection of the equities included in the two pools is being made according to the Industry Classification Benchmark (ICB) jointly developed by Dow Jones and FTSE (see appendix 1). In the sample used, the two filtered pools include all stocks from the US and UK stock markets that are engaged in the various phases of energy production and processing, listed in the following four sectors: 1) Oil and Gas Producers, 2) Oil Equipment, Services and Distribution, 3) Alternative Energy, and 4) Electricity. After applying the filtering procedure to the US and UK stock markets, two energy-related stock pools are constructed hereafter named US Filter and UK Filter, respectively.

Hence, to test the proposed heuristic approach and the efficiency of both the DE and the GA as index-tracking methodologies, five data sets are selected. All stock prices are closing prices adjusted for capital gains according to the annualised dividend yield, and they are all obtained on daily basis for the period January 31, 2006 to February 1, 2010 from Thomson Financial Datastream. All stock prices are in US dollars thus reflecting the local currency exchange rate against the USD at every point in time for the period examined. Should a company cease trading due to an event (merger, bankruptcy etc.), within the test period, it is dropped from the sample; that is why the total number of stocks in the FTSE 100 and Bovespa pools is less than the total number of stocks included in each index. Moreover, after adjusting for all US and UK Bank Holidays, 1,008 observations are sorted to calculate daily returns for each stock. Considering 252 trading days in a calendar year, the heuristic approach is tested under various assumptions by selecting the first year as the in-sample period and the last three years as the out-of-sample period. The final five data sets have the following number of stocks: N=41 (UK Filter), N=53 (Bovespa Composite), N=65 (Dow Jones Composite Average), N=77 (US Filter), and N=97 (FTSE 100 Index). See appendix 2 for a detailed list of all stocks used in each pool.

4. Methodology

4.1. Evolutionary Algorithms

EAs have been applied to numerous optimization problems in business, engineering, cognitive and applied sciences (Goldberg, 1989). More specifically, since the 1980s, a rapid expansion of their practical and theoretical financial applications has been witnessed. Some of the applications include portfolio optimization (Lorashi and Tettamanzi, 1996; Beasley et al., 2003; Chang et al., 2009), insurance risk assessment (Hughes, 1990), technical trading rules and market timing strategies (Bauer, 1994; Neely et al., 1997; Allen and Karjalainen, 1999), time series forecasting and econometric estimation (Marimon et al., 1990; Dorsey and Mayer, 1995; Leinweber and Arnott, 1995; Mahfoud et al., 1997). Primarily, there are four paradigms that can be identified as different techniques that belong to the family of EAs. These are the Genetic Algorithms (Holland, 1962, 1975), Genetic Programming (Koza, 1992, 1994), Evolutionary Strategies (Recheuberg, 1973), and Evolutionary Programming (Fogel et al., 1996).

Evolutionary Algorithms (EAs) are widely used in the operational research literature for solving multi-objective optimization problems (Coello Coello, 1999; Deb, 2001), and have many advantages over traditional operational research techniques (Zitzler and Thiele, 1999). Issues regarding the convexity, concavity, and continuity or multiple local optima of the objective functions do not need to be taken into consideration. The main feature that differentiates an evolutionary search algorithm from other traditional search algorithms such as random sampling (e.g. random walk) and heuristic sampling (e.g. gradient descent), is that it is population based. Evolutionary algorithms use a population of points to search the space rather than a single point making them superior to random search. They also have the advantage of avoiding the hill-climbing behaviours of gradient-based search algorithms (Sivanandam and Deepa, 2007). Traditional optimization techniques, such as the gradient methods, break down due to their inability to handle the constraint that restricts the number of assets included in the tracking portfolio.

In general, an EA generates a population of potential solutions and evaluates the quality of each one based on a problem-specific fitness function that defines the evolution environment. Because it is this cost function that guides the search, no supplementary knowledge is needed. In addition EAs use probabilistic transition rules rather than deterministic ones, and an encoding of the search space rather than a single point (Kingdon and Feldman, 1995). Using various operators, new solutions are generated by selecting the relatively fit population members and then these are recombined, performing an efficient direct search and thus reducing the uncertainty about the search space. However, EAs do have some limitations like the fact that the user cannot easily incorporate problem-specific information, making them less efficient than special purpose algorithms in well understood domains. Another weakness is that in differentiable problems an EA could prematurely converge, or converge to a non-zero gradient point if there is limited genetic variation left in the population.

Nevertheless, for most real world financial problems, a number of unknown factors affect the multi-objective target functions of large search spaces. These are complex problems characterized by irregular features such as multiple optima, nonlinearities, and discontinuities of the objective function. Many option pricing, trading rules and constrained portfolio optimization problems for which a closed form solution is not available, serve as examples. The ability of the EAs to handle the solutions of these types of problems, and to find the global optimum relatively fast, strengthens the conclusion that they are a powerful and robust optimization technique.

4.2. Genetic Algorithm (GA) and Differential Evolution Algorithm (DE)

The most popular technique in evolutionary computation research is the Genetic Algorithm (GA). One of the most important steps of the GA is the selection of the individuals used to produce the successive generations. Any single individual in the population has a chance of being selected at least once in order to be reproduced into the next generation. There are many different schemes and their variations that can be used for the selection process such as the roulette wheel selection, which was the first scheme introduced, the tournament and ranking selection, scaling techniques and elitist models (Goldberg, 1989; Michalewicz, 1994). The genetic algorithm used in this paper applies the tournament selection scheme that requires only the evaluation function to map the solutions to a partially ordered set, allowing for minimization and negativity. It is used in this paper, because unlike other more conventional schemes, it does not assign any probabilities. Under this scheme, k individuals are randomly selected from the population, with replacement, with the best individual being selected to participate in the new population; each individual represents a vector of prices. This process is repeated until N individuals are selected.

The next most important step in the GA is to select the scheme of the genetic operators used to provide the building block of the search mechanism. The two basic operators are the mutation and the crossover. In the GA variation applied in this paper, real valued representations are used for both operators as developed by Michalewicz (1994), the uniform mutation and the arithmetic crossover. Let for every variable j, a_j and b_j be the lower and upper bounds, respectively. Next, the uniform mutation selects a random variable j^* which is set equal to a uniform random number, i.e.:

$$x'_{ij} = \begin{cases} U_j \sim unif\left(a_{j^*}, b_{j^*}\right), & \text{if } j = j^* \\ x_{ij}, & \text{otherwise} \end{cases} \tag{2}$$

Under the arithmetic crossover scheme, two complimentary linear combinations of the parents are generated based on the random number r drawn from a uniform distribution $U_i \sim unif\left(0,1\right)$. The two new individuals $\bar{X}'$ and $\bar{Y}'$ are created based on the following equations:

$$\bar{X}' = r\bar{X} + \left(1-r\right)\bar{Y} \tag{3}$$

$$\bar{Y}' = \left(1-r\right)\bar{X} + r\bar{Y} \tag{4}$$

For each new solution to be reproduced, a pair of "parent" solutions, $\bar{X}'$ and $\bar{Y}'$, is selected from breeding from the pool selected previously. Hence, by producing a "child" solution using the abovementioned methods of crossover and mutation, a new solution is created which generally shares many of the characteristics of its "parents". Finally, the GA moves from one generation to the next, selecting and reproducing parent solutions until a termination criterion is met. For the purposes of this paper the process is repeated until either the population converges to the global optimum (i.e. the optimum solution that satisfies the criteria set) or the pre-specified maximum number of generations is reached. A more extensive discussion on the genetic algorithms' functionalities, extensions and applications, can be found in Holland (1975), Goldberg (1989), Davis (1991) and Michalewicz (1994).

DE, on the other hand, is one of the latest heuristic approaches which also belongs to the family of Evolutionary Algorithms (EAs) and has been developed by Storn and Price (1995) for solving nonlinear and non-differentiable continuous space functions. DE is a stochastic optimization method which can minimize a function capable for modelling the problem's objectives, while at the same time incorporate all necessary solution constraints. More specifically, DE has the following advantages over rival approaches; fast convergence, use of few control parameters, ability to find the true global minimum irrespective of the initial parameter values, robustness, and ease of use (Storn and Price, 1997). What is more, DE's claimed advantages are apparent when applied to the index tracking problem. Maringer and Oyewumi (2007) show evidence for the latter from the Dow Jones Industrial Average by analysing the financial implication of cardinality constraints for tracking portfolios when using a subset of its components. DE does not use binary encoding or a probability density function to self-adapt its parameters as a simple EA; there are, however, modified GAs that use real number representation, similar to the one used in this paper. The DE algorithm has also been used in other recent studies using hybrid and multi-objective schemes (Krink et al., 2009; Krink and Paterlini, 2011), as well as in the context of loss aversion (Maringer, 2008) and mutual fund replication (Zhang and Maringer, 2010). Other recently proposed algorithmic procedures include immune systems (Li et al., 2011), hybrid algorithms (Ruiz-Torrubiano and Suárez, 2009), robust optimization (Chen and Kwon, 2012), and mixed-integer programming formulations (Canakgoz and Beasley, 2008; Stoyan and Kwon, 2010).

Furthermore, the main difference between the GA and the DE lies on the schemes used for the selection process, the mutation and the crossover operators. In the GA, two parents are selected for crossover and the child is a recombination of the parents, whereas in DE three parents are selected for crossover and the child is a perturbation of one of them (Sarker and Abbass, 2004). The DE is a self adaptive algorithm, with all possible solutions having the same chance of being selected as parents with no dependence on their fitness value, and at the same time it is also a "greedy" algorithm, whereas only the best new solution and its parent are kept. Comparisons on various benchmark problems show that DE performs better when compared to other evolutionary algorithms (Sarker et. al. 2002, Sarker and Abbass, 2004). DE's proven past performance is the reason why it is used to solve the index tracking problem in this paper, serving as a comparison methodology next to the modified GA.

There are various approaches with respect to the way mutation is computed and to the type of the recombination operator used to solve the global optimization problem. The general notation, for the variant schemes/ strategies for the DE algorithm as introduced by Storn and Price (1997), is the following: DE/x/y/z where, "DE" stands for Differential Evolution, "x" specifies the methodology used to choose the population vector to be mutated, "y" is the total number of vector differences that contributes to the differential, and "z" indicates the crossover scheme used. In the optimization problem presented in this paper the following notation is used, with x = rand-to-best, y = 1 and z = exp, identifying the "DE/rand-to-best/1/exp" variant as the most suitable. "Rand-to-best" indicates that the population vectors are selected to compute the mutation values that lie on the line defined by the randomly generated and the best-so-far vectors; "1" is the number of pairs of solutions chosen (how many vector differences contribute to the differential); and finally, "exp" means that an exponential crossover scheme is used. Compared to the basic version of the DE, the aforementioned scheme is used in this paper because it enhances the greediness of the algorithm by incorporating the current best vector into the scheme.

Definition 1: Let $\mathbf{u}_{ji,G+1}$ be the trial vector, $\mathbf{v}_{ji,G+1}$ the mutant vector, $\mathbf{x}_{ji,G}$ the parent solution from the current generation G, $x_{jr_1,G}$, $x_{jr_2,G}$ and $x_{jr_3,G}$ three randomly chosen integer indexes which are mutually different and also different from the running index i. Define,

Mutation: $v_{ji,G+1} = x_{jr_1,G} + F\left(x_{jbest,G} - x_{jr_1,G}\right) + F\left(x_{jr_2,G} - x_{jr_3,G}\right)$ (5)

Crossover: $\mathbf{u}_{ji,G+1} = \begin{cases} \mathbf{v}_{ji,G+1}; & \mathrm{u}_j \le CR \text{ or } j = j_{rand} \\ \mathbf{x}_{ji,G}; & \text{otherwise} \end{cases}$ (6)

Selection: $\mathbf{x}_{i,G+1} = \begin{cases} \mathbf{u}_{i,G+1}; & f\left(\mathbf{u}_{i,G+1}\right) \le f\left(\mathbf{x}_{i,G}\right) \\ \mathbf{x}_{i,G}; & \text{otherwise} \end{cases}$ (7)

$i = 1, 2, \ldots NP;\ r_1, r_2, r_3 \in \{1, 2, \ldots NP\}$
$r_1 \ne r_2 \ne r_3 \ne i;\ NP \ge 4$
$j = 1, 2, \ldots D;\ \mathrm{u}_j \square\ unif\,[0,1]$
$G = 1, 2, ..G_{\max}$
$CR \in [0,1]$
$F \in [0,2]$

where NP is the total number of D-dimensional parameter vectors that represent the population of the available decision variables for each generation, which also remains constant during the minimization process. Also, $\mathbf{x}_{jbest,G}$ is the best solution of the population, CR is the crossover probability that controls the fraction of parameter values that are copied from the mutant, and F is a real and constant factor that controls for the magnitude of the differential variations $\left(x_{jbest,G} - x_{jr_1,G}\right)$ and $\left(x_{jr_2,G} - x_{jr_3,G}\right)$, respectively.

The steps of the DE that describe Definition 1 are the following: The first step is the population structure where a random sample of solution vectors is generated, after both the upper and lower bounds for each parameter are specified. A uniform probability distribution for all random solutions is assumed. Then, for every target vector $\mathbf{x}_{i,G}$ a mutant vector $\mathbf{v}_{i,G+1}$ is generated (eq. 5), which combines other randomly selected population vectors. Compared to the basic version of the DE, the control variable F is introduced twice to enhance the greediness of the algorithm by incorporating the current best vector $\mathbf{x}_{best,G}$ into the scheme. This step is known as "mutation".

Then as a third step, an index j that contains randomly chosen numbers $\mathbf{u}_j$ from the uniform distribution$[0,1]$, ensures that $\mathbf{u}_{i,G+1}$gets at least one parameter from$\mathbf{v}_{i,G+1}$. If $\mathbf{u}_j$ is less than or equal to the crossover probability CR, then the mutant vector $\mathbf{v}_{ji,G+1}$is being mixed with the parameters of another predetermined vector, the solution-parent$\mathbf{x}_{ji,G}$, to produce the so-called trial vector$\mathbf{u}_{ji,G+1}$(eq. 6); otherwise, the parameter is copied from the target vector$\mathbf{x}_{ji,G}$. Moreover, the trial parameter with the randomly chosen index, j_{rand}, is taken from the mutant vector to ensure that the trial vector does not duplicate$\mathbf{x}_{ji,G}$. This step is known as "crossover". Finally, during the selection process, to decide whether or not to keep the trial vector $\mathbf{u}_{i,G+1}$as a member of the generation$G+1$, its cost function is compared with the target vector$x_{i,G}$using the greedy criterion. If the objective function value of the trial vector $\mathbf{u}_{i,G+1}$is less or equal to that of the target vector$\mathbf{x}_{i,G}$, then it replaces the target vector in the subsequent generation (eq. 7); otherwise, the parent solution $\mathbf{x}_{i,G}$ is retained. This final step is known as "selection".

As mentioned earlier, in order to use the DE algorithm, it needs to be fine-tuned using just three control parameters; the crossover constant (CR); the weighting factor (F); and the number of parents (NP). The CR parameter is responsible for controlling the influence of the parent on the generation of the offspring, with higher values having a reduced effect. The F parameter controls the influence of the pair of solutions that calculate the mutation value (for the variant specification used in this paper that includes only one pair[9]). For most optimization problems, as a rule of thumb, F and CR should both be set in the range of [0.5, 1], while NP should be between 5*D and 10*D, where D equals the number of decision variables (in the present case this is the number of available stocks) (Price et al., 2005; Storn and Price, 1997). Based on the aforementioned, the combination of F, CR and NP that is used for the optimization problem solved in this paper is 0.7, 0.5 and 10*D, respectively. The following table summarizes the parameters used as inputs for both the GA and the DE.

Table 1: Parameters used as inputs in the algorithms.

Genetic Algorithm (GA)	
Solution representation	Binary with 10 digits
Selection	Tournament - stochastic with replacement
Crossover	Arithmetic - 2 individuals
Crossover probability	0.8
Mutation	Uniform
Mutation probability	0.001
Population size	100N
Number of generations	200
Differential Evolution Algorithm (DE)	
Solution representation	Space vector R^N
Crossover	Exponential
Crossover probability	0.5
Mutation	DE/rand-to-best/1
Mutation constant	0.7
Population size	10N
Number of generations	100

4.3. Formulating the objective function and its constraints

To test the performance of the proposed heuristic three different scenarios are examined. In the first one, both algorithms are tested without rebalancing the tracking portfolios for the out-of-sample period; in the second scenario the portfolios are rebalanced quarterly; and finally, in the third scenario, the portfolios are rebalanced on a monthly basis. In both cases of rebalancing, transaction costs are taken into consideration. The main purpose of testing the algorithms under

[9] Increasing either the population size or the number of pairs of solutions, in order to compute the mutation values, will increase the diversity of possible movements; hence a balance should be kept to make the algorithm more efficient (Feoktistov and Janaqi, 2004).

these three scenarios is to examine whether by including additional information in the index-tracking algorithm – by regular rebalancing of the portfolio - is more rewarding than buying the initial selected portfolio and holding it throughout the test period.

For each case examined, N number of stocks are held within the in-sample time period [1,2..,T] and the price of the index tracked. The goal is to create tracking portfolios consisting of maximum K stocks (K<N), and replicate the tracked index during the out-of-sample period [T, T+Δt]. The tracking portfolios are created based on the stocks that the algorithms choose, using every time the available data from the in-sample period. To decide which stocks will form the tracking portfolio two main objectives are employed: the tracking error and the excess return.

The tracking error (TE) is defined by the p-norm as:

$$TE = \frac{1}{T}\left\| r_t - R_t \right\|_p = \left(\sum_{t=1}^{T} \left| r_t - R_t \right|^p \right)^{\frac{1}{p}}; \ \mathrm{p} > 0, \tag{8}$$

where r_t and R_t are the returns for the tracking portfolio and the index respectively. Portfolios' returns are adjusted for transaction costs when rebalancing occurs; 0.5% per transaction. For p = 2, the p-norm is equal to the Euclidean norm which represents the Root Mean Squared Error (RMSE) as expressed by the following equation:

$$TE = RMSE = \sqrt{\sum_{t=1}^{T} \left(r_t - R_t \right)^2 / T}. \tag{9}$$

The tracking error is measured with the RMSE criterion, which according to Beasley et al. (2003) is one of the most effective measurements for addressing this type of index tracking problems. Using only the variance of $\left\{ \left(r_t - R_t \right) \middle| t = 1, \ldots, T \right\}$ as a tracking error measure (see Franks, 1992; Pope and Yadav, 1994; Connor and Leland, 1995; Buckley and Korn, 1998; Larsen and Resnick, 1998; Rohweder, 1998; Wang, 1999), could potentially lead to erroneous results, as the tracking portfolios would constantly underperform the index because they would ignore the bias proportion $\left(r_t - R_t\right)$. For example, let $M > 0$ be a constant, when $r_t = R_t - M \; \forall t$ the tracking portfolio has a zero tracking error, but will always underperform the benchmark index.

The mean Excess Return (ER) over that of the benchmark index is given by the following equation:

$$ER = \sum_{t=1}^{T} \left(r_t - R_t\right) / T. \tag{10}$$

Excess return gives a competitive advantage to any index fund that can historically show returns over and above the index, even at the cost of a higher degree of tracking error. It can be a measurement for distinguishing between competing funds besides the amount they charge for participation. The complete formulation of the objectives and constraints used to solve the index tracking problem is the following:

Minimize: $\lambda \times RMSE - \left(1-\lambda\right) \times ER$ (11)

Under the constraints: $\sum_{i=1}^{N} P_{iT} x_i = C$ (12)

$$z_i \varepsilon C \le P_{iT} x_i \le z_i C \qquad \forall i = 1,...,N \text{ ; } \varepsilon \ge 0.05\text{*C} \tag{13}$$

$$\sum_{i=1}^{N} z_i \le K \tag{14}$$

$$x_i \ge 0, \qquad z_i \in \left\{0,1\right\} \qquad \forall i = 1,...,N$$

where λ $\left(0 \le \lambda \le 1\right)$ is the generalised minimization objective for the index tracking problem; a metric controlling for the trade-off between tracking error and excess return. In case λ = 1, the tracking portfolio has as its main objective to minimize the tracking error (pure index tracking), whereas when λ = 0, the portfolio's main goal is to maximize the excess return. The first constraint ensures that the value of the portfolio at the end of the in-sample period will be equal to the available capital to the investor, C. Using the rolling window method, the same rule applies for every rebalancing period. In addition, P_{iT} is the price of stock i at time T, whereas x_i is the weight of each stock that participates in the tracking portfolio. The last two constraints relate to the weights and total number of each participating stock in the portfolio; variable ε represents the minimum weight of each stock set at 5% of the available capital, and variable z is a decision variable which takes the value one (zero) when a stock is (is not) included in the basket. Finally it is assumed that all portfolios are long-only and also fully invested.

5. Empirical results

5.1. Tracking the Spot Energy Index

After developing an investable model for seeking returns comparable to the Spot Energy Index, the performance characteristics of the proposed strategy are examined. This section presents the empirical evidence on index tracking in the energy commodity markets using equity portfolios. The size of the five test problems ranges from N = 41 (UK Filter) to N = 97 (FTSE 100 Index); in the case of the Bovespa Composite N = 53, for the Dow Jones Composite Average N = 65, and for the US Filter N = 77. The stocks picked by both the DE and the GA from the aforementioned stock pools are used to track the performance of the SEI. The initial capital of the investment portfolio is set equal to C = $100,000. Figures 1 and 2 show the convergence of both the DE and the GA during the in-sample period, of the Spot Energy Index with the Bovespa, DJIA, FTSE 100, UK Filter and US Filter baskets respectively. The case considered in the two graphs is for monthly rebalancing, with λ=0.6 and portfolios of maximum 15 stocks. In the empirical analysis, tracking portfolios consisting of maximum K stocks are used with K = 10, 15, and 20. This aligns with the findings of Chang et al. (2009) that investors should include in their tracking portfolios about one third of the total assets included in the search space, since those tracking portfolios that included more assets constantly underperformed. In another study, Maringer and Oyewumi (2007) show that including roughly 50% of the available assets is satisfactory enough to get the desirable properties in the tracking portfolios. Different attitudes corresponding to three different trade-offs between tracking error and excess return are also considered, with λ = 0.6, 0.8, and 1; thus, moving from maximising excess return to minimising tracking error. Then, the heuristic is repeated ten times with the same set of parameters per run, from which the best solution is chosen.

Figure 1: DE convergence, during the in-sample period, of the Spot Energy Index with the Bovespa, DJIA, FTSE 100, UK Filter and US Filter baskets, respectively; λ=0.6, with maximum 15 stocks in the basket, rebalanced monthly.

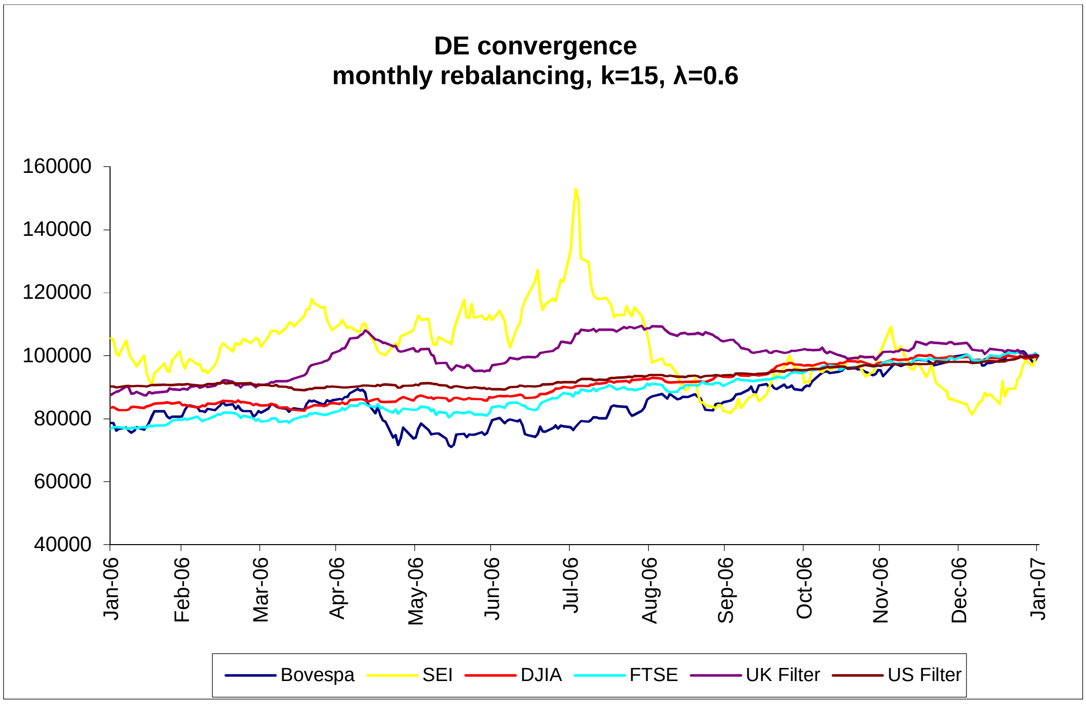


Figure 2: GA convergence, during the in-sample period, of the Spot Energy Index with the Bovespa, DJIA, FTSE 100, UK Filter and US Filter baskets, respectively; λ=0.6, with maximum 15 stocks in the basket, rebalanced monthly.

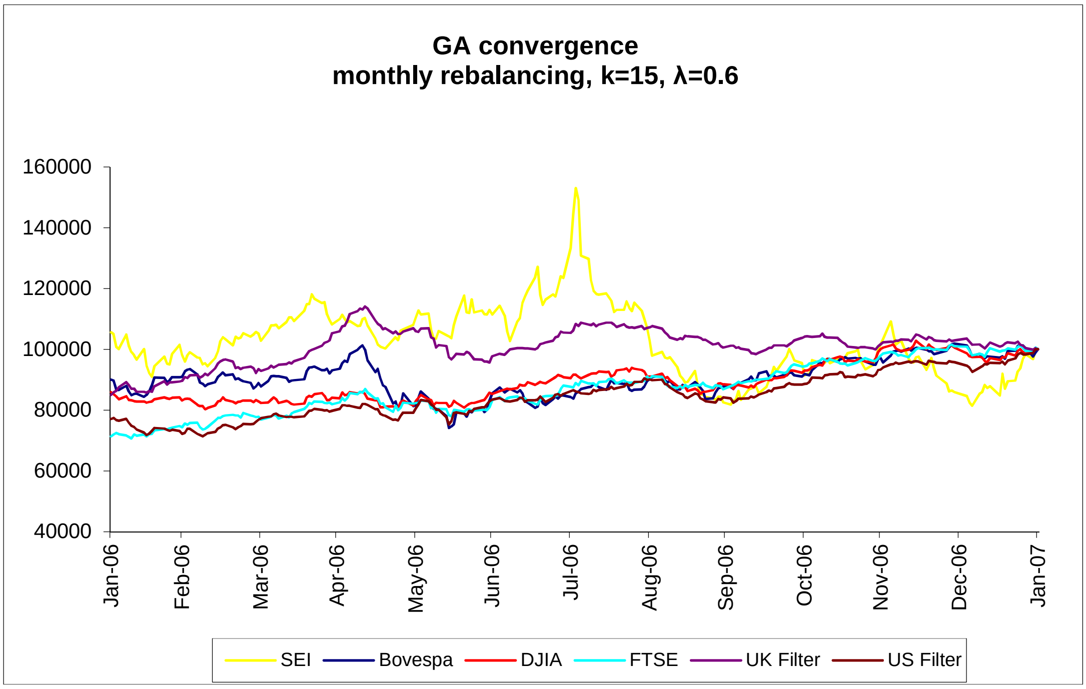


Figure 3 presents the performance of a $100K portfolio fully invested in three energy commodity indexes; the SEI, the Dow Jones-UBS Energy Index, and the Rogers Energy Commodity Index. The former represents the return available to the holder of the basket of the physical energy commodities comprising the SEI[10], and the latter total return indexes reflect the return on fully collateralized futures positions. The Dow Jones–UBS Energy Sub-Index and the Roger's Energy Commodity Index are selected for comparison reasons against the constructed SEI and the selected portfolios, as they are two of the most established indexes in the market; besides, the correlation between the energy sub-indexes of other well-known commodity indexes, such as the S&P GSCI, is extremely high. From figure 3 it is also observed that for most of the out-of-sample period, the SEI and Rogers Energy have performed better than the DJ UBS-Energy. However, especially during the last year, SEI has outperformed both futures based indexes. This

[10] The constructed Spot Energy Index tracks the evolution of the relevant commodities' spot prices setting an upper bound on the return available to an investor, since it ignores any costs associated with the holding of the physical commodities like storage, insurance etc.

confirms the fact that futures' based indexes underestimate the underlying commodity market price trends in relation to a spot index.

Figure 3: Three-year out-of-sample performance comparison of long-only portfolios invested in the SEI, Rogers Energy and DJ UBS Energy Indexes.

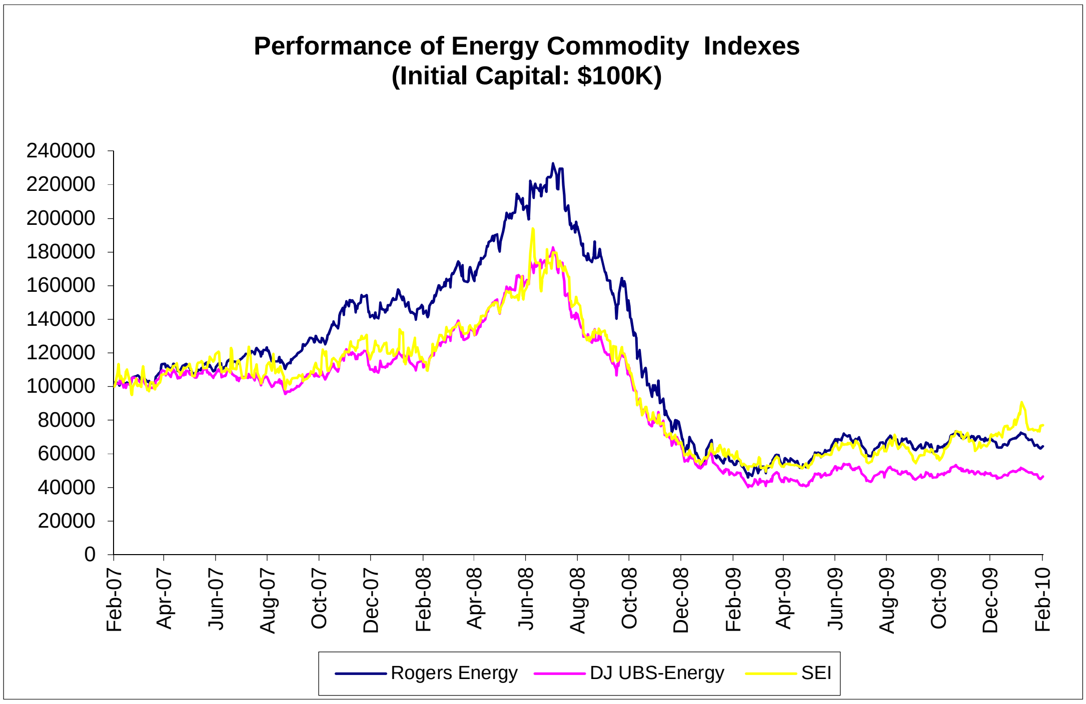
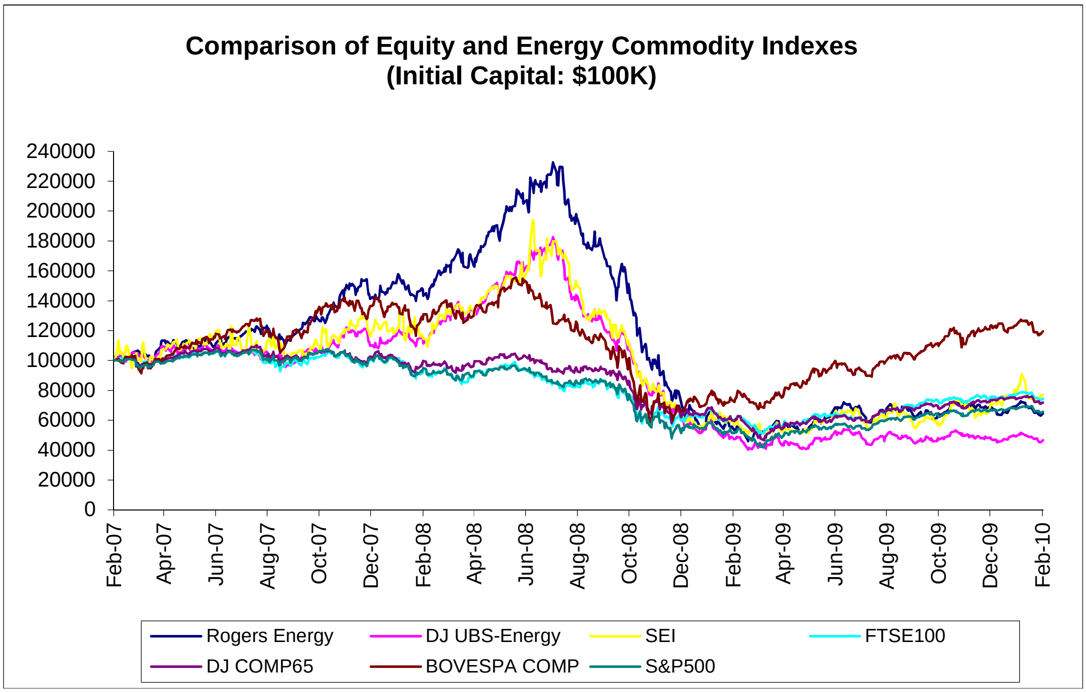


Figure 4 shows the relative performance over the out-of-sample period of the three aforementioned commodity indexes next to four financial indexes, the S&P 500 Composite, the Dow Jones Composite Average, FTSE 100, and Bovespa Composite. When global markets entered the recent global economic recession towards the end of 2007, a big price correction in both equities and commodities markets followed. It is observed that energy commodities delivered higher returns for about one year, until the end of 2008, proving to be a better investment during the recession period. This finding aligns with Weiser (2003) who concludes that commodity futures, during the period of 1970-2003, perform well in the early stages of a recession when usually stocks tend to disappoint. Gorton and Rouwenhorst (2006), as well as Vrugt et al. (2004) also find that during late expansion and early recession periods of the business cycle, commodity returns are generally above their average, outperforming stocks and bonds that generally are below their average. The aforementioned prove that there is huge potential for various timing and index tracking strategies, as the one proposed in this paper, to be applied to energy commodities markets and deliver superior returns to investors. From figure 4 it can also be seen that the indexes from the US and UK equity markets are not capable to follow the upward trend of energy commodities, except the Bovespa index that follows rather closely the high commodities' returns during the recession period, having a faster rebounding during the last year, outperforming all other equity and commodity indexes. This reflects the unique energy significance of Brazil to the global scene, and thus justifies the inclusion in this paper of stocks from the Bovespa pool to track the performance of the SEI.

Figure 4: Three-year out-of-sample performance comparison of long-only portfolios invested in the three Energy Commodity Indexes, SEI, Rogers Energy and DJ UBS Energy, and in the four benchmark Stock Indexes, FTSE 100, S&P 500, DJ Comp65 and Bovespa Comp.



Next, figures 5 and 6 display the SEI against quarterly rebalanced portfolios selected from the DE and GA respectively. The portfolios consist of maximum 15 stocks and these are the FTSE 100, DJIA, Bovespa, UK Filter and US Filter, respectively; results are shown for $\lambda = 1$. Looking at the figures it is observed that during and towards the end of the recession period, the benchmark index can be better tracked with the Bovespa baskets followed by the UK Filter baskets; whereas during the last year it is the US Filter and DJIA baskets that perform better. The portfolios comprising of optimally selected energy related stocks can successfully track the SEI, generating similar returns for most of the out-of-sample period. This is in line with Hammoudeh et al. (2004) who conclude that WTI spot prices and their respective NYMEX future prices explain the stock price movement of oil related firms, with the spot and futures prices volatility having a volatility-echoing effect on the respective stock prices. However, there are contradictory views in the literature as Schneeweis and Spurgin (1997) conclude that direct stock and bond investment cannot provide consistent risk/ return attributes similar to various commodity and managed futures indexes. In this study, the US Filter and UK Filter results verify that when energy related stocks are selected, they can better replicate the risk and return trade-off of the SEI. The same applies for the Bovespa baskets since the Brazilian stock exchange has a large number of energy and commodity related listed companies that would closely follow any developments in the international energy markets. In addition, between the DE and GA selected portfolios, from the graphs it seems that the latter ones can follow more closely the performance of the SEI, achieving highest excess returns for the final out-of-sample year.

Figure 5: Out-of-sample tracking of the Spot Energy Index with the Bovespa, DJIA, FTSE 100, UK Filter and US Filter baskets, respectively; λ=0.8, with maximum 15 stocks in the basket, rebalanced quarterly using the DE.

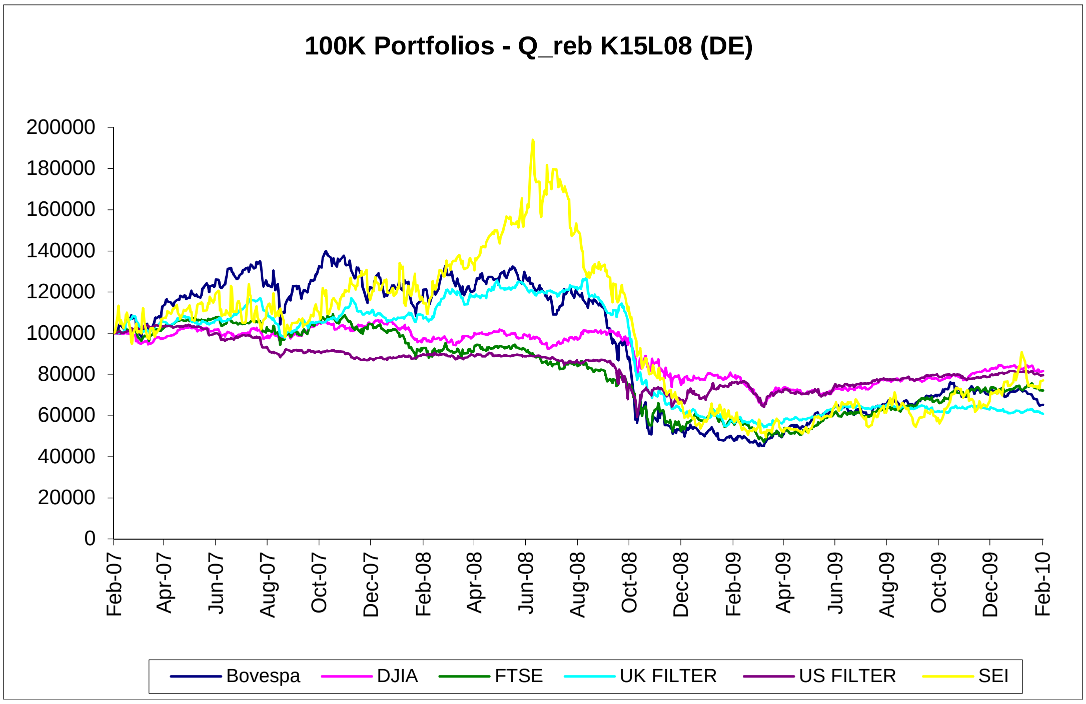


Figure 6: Out-of-sample tracking of the Spot Energy Index with the Bovespa, DJIA, FTSE 100, UK Filter and US Filter baskets, respectively; λ=0.8, with maximum 15 stocks in the basket, rebalanced quarterly using the GA.

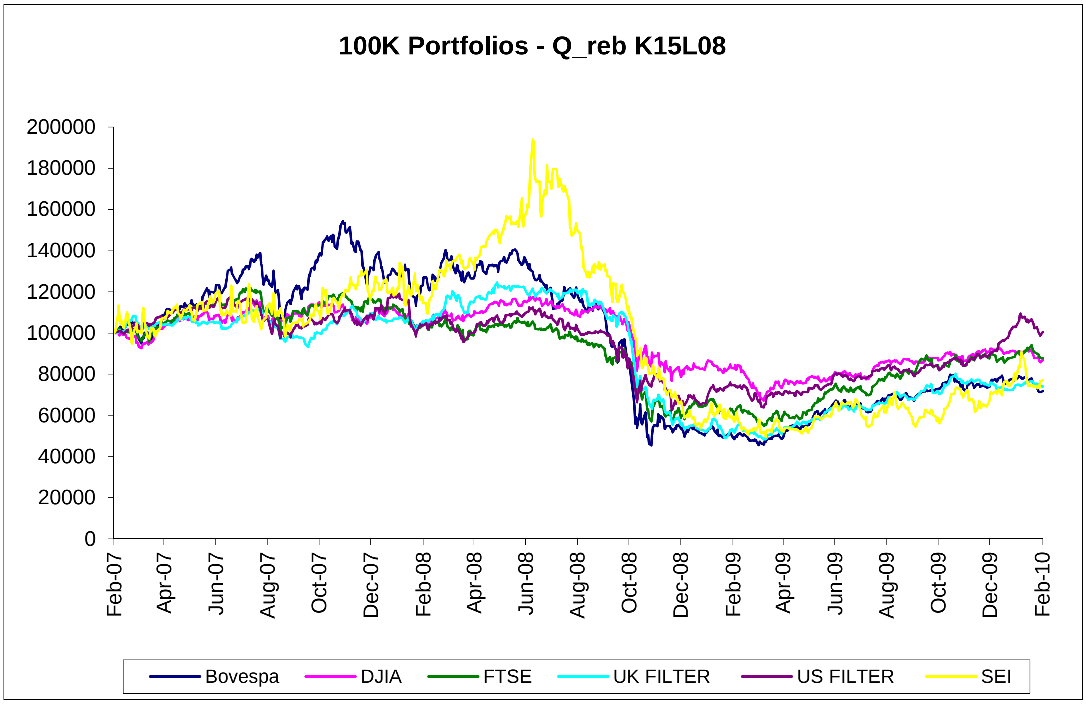


Table 2 presents the root mean squared errors and the mean excess returns of both the Genetic and Differential Evolution algorithms employed, under all three rebalancing strategies; buy-and-hold, monthly, and quarterly rebalancing. Using formal statistical evaluation criteria, the better tracking performance of the UK Filter and US Filter baskets is also confirmed. In terms of the competing portfolios' RMSEs, the DE is more consistent across the various portfolios, whereas the GA selects portfolios that exhibit larger differences between the worst and best performing ones. Additionally, in general GA tends to select portfolios that have a lower tracking error and thus track better the benchmark index when compared to the ones selected from the DE. Another interesting observation is that, although the RMSEs are improved when rebalancing occurs, increasing the frequency from quarterly to monthly has only a marginal effect. These results are more profound for the portfolios selected by the DE and align with Dunis and Ho (2005) who find that when comparing alternative rebalancing frequencies, a quarterly portfolio update is preferable to monthly, semi-annual or annual reallocations. In terms of their excess returns, in most cases, the portfolios selected by the GA tend to outperform the ones selected by the DE.

The UK Filter and US Filter baskets, that also have the lowest tracking errors (see panels D and E), have excess returns that in some cases are positive, indicating that the selected portfolios, on average, over the out-of-sample period, over-perform the SEI. In the case of the US Filter baskets selected by the GA, the index is constantly outperformed in terms of excess returns (8.10% for K=20 and λ=0.6 under monthly rebalancing, and 6.14% for K=15 and λ=0.6 under quarterly rebalancing); there is only one exception for both rebalancing frequencies when λ=1 and K=10 where the portfolios under-perform the index. This is an indication that the trade-off criterion does work, and leads to portfolios that compromise any excess return over a better tracking performance as expressed by the smaller RMSEs. Thus, taking into account the fact that commodity indexes performed better compared to the financial indexes over the three-year out-of-sample period (except the Bovespa Composite, see figure 4), with the methodology employed the performance of the SEI is closely replicated, and in the case of the energy related stock portfolios the benchmark index is even outperformed.

Table 1: Index tracking performance of selected portfolios.

Our sample spans from February 15, 2006 to February 18, 2009. The first two years are used as the estimation period whereas the last year is our test period. The tracking portfolios are created based on the stocks that the Differential Evolution and Genetic Algorithms choose. To decide which stocks will be included in the tracking portfolio, we use two main objectives, the tracking error and the excess return. K is the maximum number of stocks allowed to be included in the selected baskets. λ is the generalised minimization objective for the index tracking problem; in the case that λ takes the value of 1, the tracking portfolio has as its main objective to minimize the tracking error, whereas, when λ equals 0 the portfolio's main goal is to maximize the excess return. Our tracking portfolios include stocks picked each time from the Dow, FTSE 100, Bovespa, UK Filter and US Filter stock pools which contain N = 65, 97, 53, 41, and 77 stocks, respectively. Panels A, B, C, D and E report the out-of-sample daily Root Mean Squared Errors (RMSE) and mean daily percentage (%) Excess Returns, as defined in equations (5.8) and (5.9), respectively. We also report the results for monthly and quarterly rebalancing. Under both rebalancing strategies the weights of the tracking portfolios are estimated based on the available data in the rolling window in-sample period (one year), every month and quarter, respectively. Portfolios' returns are adjusted for transaction costs of 0.5% for each transaction.

		No Rebalance				**Monthly Rebalance**				**Quarterly Rebalance**			
		RMSE		**Mean ER (%)**		**RMSE**		**Mean ER (%)**		**RMSE**		**Mean ER (%)**	
(K)	**(λ)**	**DE**	**GA**	**DE**	**GA**	**DE**	**GA**	**DE**	**GA**	**DE**	**GA**	**DE**	**GA**
Panel A: Bovespa													
10	0.6	0.0346	0.0344	0.0136	0.0324	0.0331	0.0329	-0.0432	-0.0104	0.0333	0.0332	-0.0389	0.0134
	0.8	0.0343	0.0359	0.0176	0.0347	0.0330	0.0326	-0.0480	-0.0471	0.0332	0.0329	-0.0438	-0.0416
	1	0.0343	0.0362	0.0189	0.0133	0.0330	0.0327	-0.0545	-0.0689	0.0333	0.0332	-0.0472	-0.0236
15	0.6	0.0345	0.0359	0.0161	0.0239	0.0331	0.0327	-0.0427	-0.0063	0.0333	0.0332	-0.0411	-0.0148
	0.8	0.0343	0.0361	0.0181	0.0334	0.0330	0.0327	-0.0487	-0.0298	0.0332	0.0331	-0.0431	-0.0280
	1	0.0343	0.0356	0.0180	0.0238	0.0330	0.0327	-0.0533	-0.0418	0.0332	0.0333	-0.0442	-0.0312
20	0.6	0.0345	0.0354	0.0148	0.0233	0.0331	0.0331	-0.0436	0.0094	0.0333	0.0335	-0.0417	0.0209
	0.8	0.0343	0.0358	0.0186	0.0329	0.0330	0.0327	-0.0488	-0.0052	0.0332	0.0333	-0.0427	0.0000
	1	0.0343	0.0357	0.0164	0.0284	0.0330	0.0328	-0.0541	-0.0346	0.0333	0.0334	-0.0461	-0.0210
Panel B: DJIA													

10	0.6	0.0319	0.0328	-0.0232	-0.0257	0.0318	0.0315	-0.0479	-0.0115	0.0319	0.0319	-0.0302	-0.0243
	0.8	0.0319	0.0330	-0.0238	-0.0210	0.0318	0.0316	-0.0511	-0.0312	0.0318	0.0318	-0.0323	-0.0273
	1	0.0319	0.0330	-0.0249	-0.0218	0.0318	0.0313	-0.0522	-0.0274	0.0319	0.0317	-0.0314	-0.0172
15	0.6	0.0320	0.0329	-0.0244	-0.0200	0.0319	0.0315	-0.0503	-0.0332	0.0319	0.0318	-0.0297	-0.0172
	0.8	0.0319	0.0330	-0.0240	-0.0250	0.0318	0.0314	-0.0515	-0.0244	0.0319	0.0319	-0.0311	-0.0192
	1	0.0319	0.0328	-0.0246	-0.0239	0.0318	0.0313	-0.0515	-0.0410	0.0319	0.0319	-0.0314	-0.0283
20	0.6	0.0319	0.0328	-0.0228	-0.0251	0.0319	0.0315	-0.0514	-0.0239	0.0319	0.0319	-0.0313	-0.0005
	0.8	0.0319	0.0329	-0.0235	-0.0289	0.0318	0.0315	-0.0529	-0.0300	0.0319	0.0318	-0.0301	-0.0332
	1	0.0319	0.0328	-0.0253	-0.0323	0.0318	0.0313	-0.0505	-0.0344	0.0319	0.0317	-0.0308	-0.0051
Panel C: FTSE 100													
10	0.6	0.0315	0.0318	-0.0450	-0.0359	0.0309	0.0299	-0.0597	-0.0260	0.0308	0.0303	-0.0438	0.0106
	0.8	0.0317	0.0316	-0.0469	-0.0246	0.0309	0.0302	-0.0701	-0.0416	0.0309	0.0305	-0.0475	-0.0255
	1	0.0316	0.0314	-0.0495	-0.0193	0.0310	0.0300	-0.0735	-0.0635	0.0310	0.0307	-0.0461	-0.0334
15	0.6	0.0315	0.0318	-0.0512	-0.0253	0.0309	0.0303	-0.0674	-0.0327	0.0308	0.0303	-0.0468	-0.0180
	0.8	0.0316	0.0313	-0.0477	-0.0220	0.0309	0.0302	-0.0634	-0.0449	0.0309	0.0306	-0.0416	-0.0127
	1	0.0316	0.0312	-0.0490	-0.0175	0.0310	0.0303	-0.0699	-0.0682	0.0310	0.0306	-0.0456	-0.0349
20	0.6	0.0315	0.0317	-0.0507	-0.0271	0.0309	0.0303	-0.0705	-0.0311	0.0308	0.0305	-0.0442	-0.0092
	0.8	0.0316	0.0313	-0.0484	-0.0297	0.0310	0.0303	-0.0681	-0.0656	0.0309	0.0305	-0.0445	-0.0145
	1	0.0316	0.0313	-0.0492	-0.0245	0.0310	0.0301	-0.0679	-0.0600	0.0310	0.0306	-0.0449	-0.0208
Panel D: UK Filter													
10	0.6	0.0318	0.0309	-0.0900	-0.0834	0.0299	0.0294	-0.0712	0.0019	0.0300	0.0296	-0.0681	-0.0032
	0.8	0.0315	0.0312	-0.0818	-0.0834	0.0300	0.0290	-0.0680	-0.0725	0.0301	0.0296	-0.0611	-0.0412
	1	0.0317	0.0307	-0.0809	-0.0751	0.0300	0.0292	-0.0713	-0.1371	0.0301	0.0297	-0.0632	-0.1049
15	0.6	0.0312	0.0309	-0.0825	-0.0519	0.0299	0.0294	-0.0782	-0.0427	0.0300	0.0298	-0.0711	-0.0341
	0.8	0.0313	0.0309	-0.0847	-0.0408	0.0300	0.0293	-0.0720	-0.0501	0.0300	0.0296	-0.0707	-0.0410
	1	0.0313	0.0308	-0.0846	-0.0531	0.0300	0.0293	-0.0782	-0.1083	0.0301	0.0297	-0.0601	-0.0459
20	0.6	0.0311	0.0305	-0.0796	-0.0586	0.0299	0.0297	-0.0764	-0.0508	0.0300	0.0299	-0.0717	-0.0446
	0.8	0.0311	0.0303	-0.0858	-0.0451	0.0299	0.0294	-0.0752	-0.0790	0.0300	0.0298	-0.0697	-0.0391
	1	0.0311	0.0304	-0.0763	-0.0516	0.0300	0.0295	-0.0747	-0.0794	0.0301	0.0296	-0.0676	-0.0494
Panel E: US Filter													
10	0.6	0.0307	0.0329	-0.0258	-0.0442	0.0306	0.0297	-0.0449	0.0710	0.0309	0.0307	-0.0364	0.0249
	0.8	0.0308	0.0321	-0.0265	-0.0780	0.0309	0.0295	-0.0603	0.0607	0.0310	0.0300	-0.0345	0.0240
	1	0.0309	0.0318	-0.0234	-0.0314	0.0310	0.0294	-0.0688	-0.0278	0.0310	0.0298	-0.0367	-0.0172
15	0.6	0.0307	0.0321	-0.0246	-0.0581	0.0309	0.0306	-0.0497	0.1241	0.0310	0.0308	-0.0322	0.0614
	0.8	0.0308	0.0327	-0.0244	-0.0511	0.0309	0.0296	-0.0575	0.0212	0.0310	0.0301	-0.0336	0.0016
	1	0.0308	0.0322	-0.0254	-0.0566	0.0309	0.0295	-0.0648	-0.0027	0.0310	0.0302	-0.0342	0.0204
20	0.6	0.0307	0.0327	-0.0261	-0.0668	0.0309	0.0301	-0.0510	0.0810	0.0310	0.0308	-0.0274	0.0345
	0.8	0.0308	0.0319	-0.0251	-0.0320	0.0309	0.0296	-0.0603	0.0210	0.0310	0.0303	-0.0329	0.0369
	1	0.0307	0.0311	-0.0226	-0.0649	0.0309	0.0294	-0.0662	0.0071	0.0310	0.0301	-0.0352	0.0126

Now in terms of the risk/ return trade-off (λ), it is observed that results are very similar between portfolios where λ=0.8 and 1. In most cases, the risk/ return trade-off criterion tends to perform well, selecting portfolios with higher returns and also relatively higher RMSEs. Moreover, the portfolios selected by the GA tend to be more consistent when the risk/ return trade-off rule is applied, compared to the ones selected by the DE. Overall, when considering both the tracking performance and the excess returns of the various portfolios, those with λ=0.8 should be preferred. As far as the maximum number of stocks criterion is concerned, in all three rebalancing scenarios, portfolios with K=10 tend to perform worst in terms of RMSEs but they do slightly better in terms of excess returns, for both the DE and GA selected portfolios. This is also an indication that the more stocks are included in the portfolio, the higher the transaction costs when a rebalancing occurs. Overall, it is suggested that portfolios with a maximum of 15 stocks should be selected, as there still seems to be a valuable compensation for the additional information and diversification when rebalancing, against the extra rebalancing costs.

According to the results, for both algorithms, monthly rebalancing is overall the best option in terms of RMSEs, closely followed by quarterly rebalancing; whereas when looking at excess returns, quarterly rebalancing appears to improve portfolio performance. This last observation can be confirmed by figures 8 and 10 where the UK Filter baskets selected by the DE and GA, respectively, are plotted, with K=20 and λ=1, for all three rebalancing frequencies. Also, from figures 7 and 9 it is clearly seen that for the Bovespa baskets, the buy-and-hold strategy performs better than both the quarterly and monthly rebalancing. The return of a buy and hold portfolio may be higher than that of a rebalanced portfolio when transaction costs are considered, but it is important to determine the source of the higher return; whether it is greater capital efficiency as expressed by a higher Sharp or Information ratio, or greater risk. Plaxco and Arnott (2002) showed that rebalanced portfolios typically have higher Sharpe ratios than buy-and-hold portfolios; a finding that suggests that the possible outperformance of a buy-and-hold portfolio may be the result of greater risk. Results are more apparent for the GA portfolios, as for the DE portfolios the difference between monthly and quarterly rebalancing is only marginal. In the case of the UK Filter basket, picked by the GA, there is an obvious difference in performance when rebalancing quarterly, against a monthly rebalancing. A more in depth analysis comparing the portfolios' information ratios is presented in the following section. On average, based on the results from table 2, K=15 and λ=0.8 is the most desirable combination providing the best results for most tracking portfolios. Although it is up to the investors' risk/ return appetite to decide whether rebalancing their portfolio quarterly, which comes with an extra cost, it is better than no rebalancing at all. The same applies and as to whether λ=0.8 should be used compared to a more risky trade-off when λ=0.6.

Figure 7: Out-of-sample performance of the Bovespa portfolio; λ=1, with maximum 20 stocks in the basket, under the three rebalancing frequencies as selected by the DE.

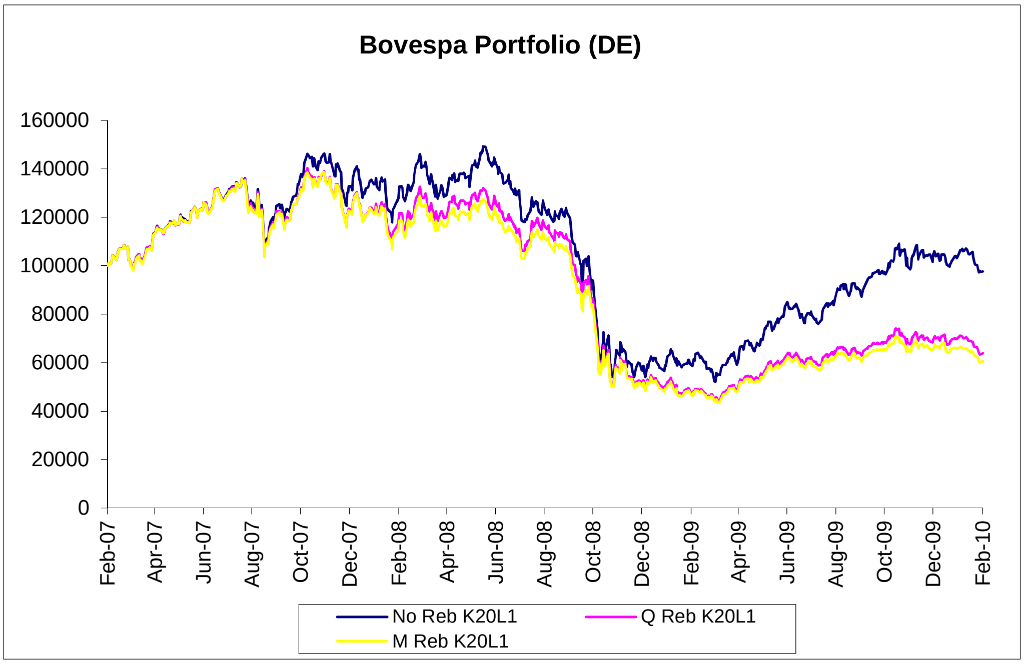


Figure 8: Out-of-sample performance of the UK Filter portfolio; λ=1, with maximum 20 stocks in the basket, under the three rebalancing frequencies as selected by the DE.

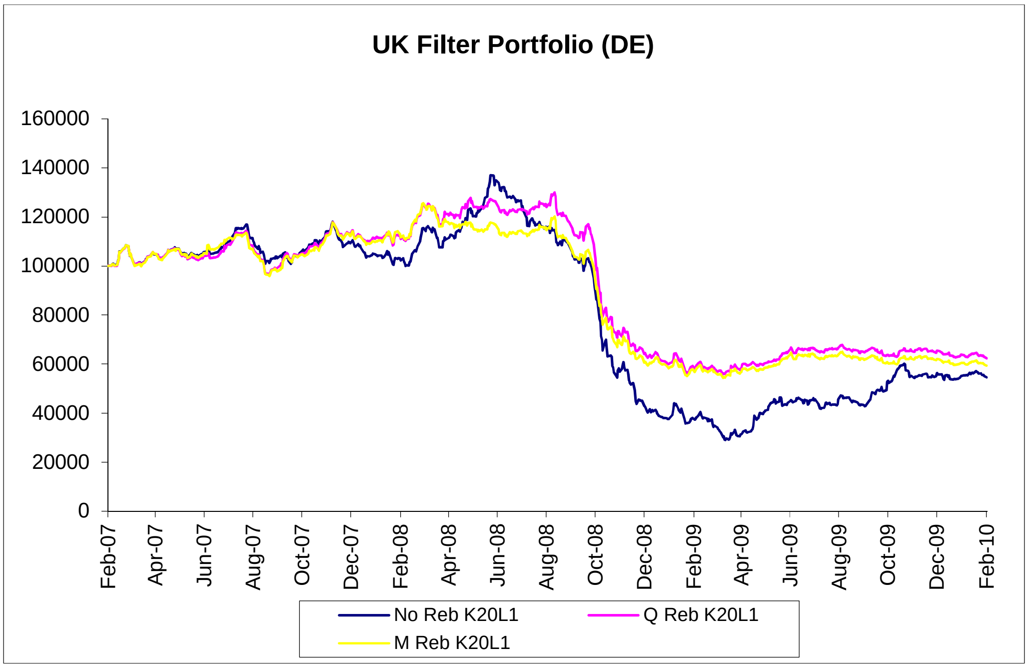


Figure 9: Out-of-sample performance of the Bovespa portfolio; λ=1, with maximum 20 stocks in the basket, under the three rebalancing frequencies as selected by the GA.

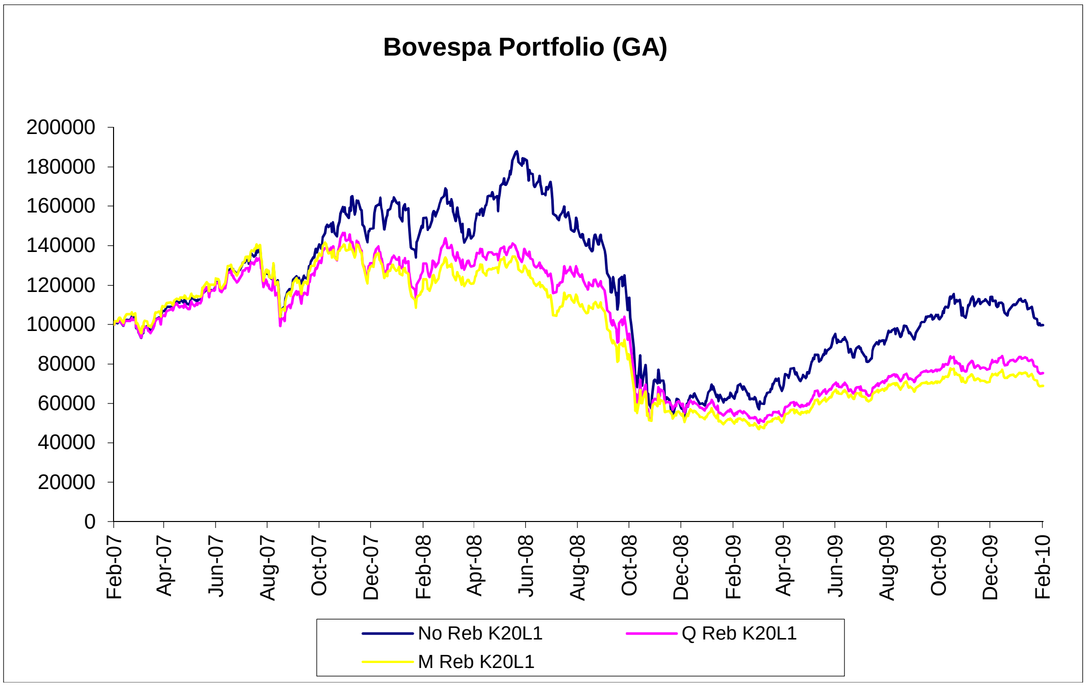
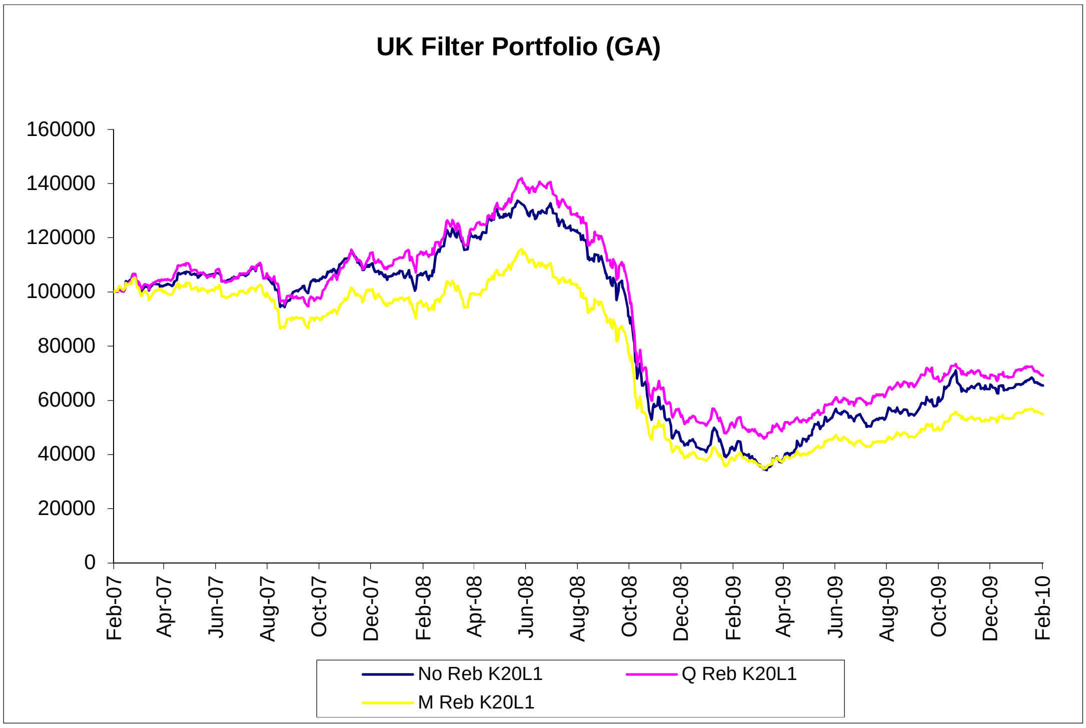


Figure 10: Out-of-sample performance of the UK Filter portfolio; λ=1, with maximum 20 stocks in the basket, under the three rebalancing frequencies as selected by the GA.



5.2. Statistical properties of selected portfolios

Tables 3, 4 and 5 present some distributional statistics of the selected portfolios' returns under the buy-and-hold, monthly and quarterly rebalancing respectively. Also, in panel F of each aforementioned table, the statistics and relevant performance measures for the following indexes are reported for comparison reasons: two Total Return Energy Commodity Indexes, the DJ UBS-Energy and Rogers Energy Commodity, the three stock indexes used to draw stocks from to construct the tracking portfolios, Bovespa, DJIA and FTSE 100, and finally the most commonly used benchmark in the finance industry, the S&P 500. According to the historical annualised volatilities for the out-of-sample period, the SEI is more volatile than the DJ UBS-Energy and Rogers Energy Commodity Indexes; 48.40% as compared to 36.21% and 41.11% respectively. The respective volatility of the equity indexes is in the range of 27% to 38%. However, when comparing the information ratios, only the Bovespa index is able to generate a better risk-return performance compared to the SEI.

Table 2: Distributional statistics of portfolios' daily returns.

This table presents the annualised returns and volatilities of the tracking portfolios, the skewness and kurtosis, the correlation coefficient between the returns of the benchmark index and the portfolio that is used each time to replicate this benchmark, and the Information Ratio, under the No Rebalancing strategy. The Information Ratio (IR) is the ratio of each portfolio's return above the return of the benchmark index to the volatility of those returns. It measures the ability of the portfolio to generate excess returns relative to the benchmark index, and at the same time suggests consistency of performance. The IR can be expressed as the following ratio: IR = (Mean Excess Return of the Portfolio) / (Excess Returns' Volatility). Panels A, B, C, D and E represent the portfolios that include stocks picked each time from the Dow, FTSE 100, Bovespa, UK Filter and US Filter stock pools. Panel F presents, for comparison reasons, the relevant performance measures for two Total Return Energy Commodity Indexes, the DJ UBS-Energy and Rogers Energy Commodity, for the three stock indexes used to draw stocks from in order to construct the tracking portfolios, Bovespa, DJIA and FTSE 100, and finally the most commonly used benchmark in the finance industry, the S&P 500.

		No Rebalancing											
		An. Ret (%)		**An. Vol. (%)**		**Skewness**		**Ex. Kurtosis**		**Correl. (%)**		**Info Ratio**	
(K)	**(λ)**	**DE**	**GA**	**DE**	**GA**	**DE**	**GA**	**DE**	**GA**	**DE**	**GA**	**DE**	**GA**
Panel A: Bovespa													
10	0.6	6.44	11.16	40.16	41.03	-0.282	-0.389	7.582	6.609	24.19	26.22	0.062	0.149
	0.8	7.44	11.76	39.22	45.37	-0.316	-0.325	7.813	5.933	24.01	26.19	0.081	0.153
	1	7.76	6.37	39.34	47.10	-0.320	-0.304	7.658	4.825	24.17	27.68	0.087	0.059
15	0.6	7.06	9.03	39.85	44.92	-0.272	-0.299	7.748	6.313	23.89	25.48	0.074	0.106
	0.8	7.56	11.42	39.27	47.02	-0.311	-0.359	7.732	4.550	23.95	27.63	0.083	0.147
	1	7.55	9.00	39.46	45.61	-0.327	-0.374	7.560	5.105	24.25	27.70	0.083	0.106
20	0.6	6.73	8.86	39.96	44.26	-0.275	-0.260	7.633	5.512	24.10	26.69	0.068	0.104
	0.8	7.68	11.29	39.54	45.01	-0.307	-0.350	7.608	5.942	24.40	26.04	0.086	0.146
	1	7.14	10.16	39.72	45.09	-0.324	-0.337	7.389	5.609	24.74	26.69	0.076	0.126
Panel B: DJIA													
10	0.6	-2.85	-3.46	22.18	31.14	0.571	0.406	11.674	11.823	12.33	19.84	-0.116	-0.124
	0.8	-2.98	-2.28	21.50	32.93	0.490	0.390	11.175	11.229	11.55	21.07	-0.118	-0.101

	1	-3.28	-2.48	21.44	31.48	0.366	0.547	10.852	12.525	11.10	19.31	-0.124	-0.105
15	0.6	-3.14	-2.03	22.69	31.68	0.563	0.546	12.006	12.512	12.66	19.97	-0.121	-0.096
	0.8	-3.05	-3.31	22.02	32.20	0.489	0.240	11.446	10.909	12.04	20.23	-0.119	-0.120
	1	-3.20	-3.01	21.86	32.17	0.426	0.394	10.942	11.654	11.76	21.26	-0.122	-0.115
20	0.6	-2.73	-3.33	22.55	32.03	0.515	0.220	11.418	10.750	12.83	20.96	-0.113	-0.122
	0.8	-2.91	-4.27	22.18	32.85	0.463	0.130	10.919	10.488	12.18	21.96	-0.117	-0.139
	1	-3.38	-5.13	21.65	31.66	0.403	0.250	10.538	10.939	11.57	20.61	-0.126	-0.156
Panel C: FTSE 100													
10	0.6	-8.34	-6.04	28.22	31.64	-0.059	-0.231	6.344	6.944	23.50	25.81	-0.227	-0.179
	0.8	-8.82	-3.18	28.84	30.89	-0.080	0.013	6.418	7.273	23.06	25.95	-0.235	-0.123
	1	-9.47	-1.87	29.44	30.66	-0.104	0.021	5.995	7.300	23.98	27.03	-0.248	-0.098
15	0.6	-9.90	-3.37	28.64	30.53	-0.110	-0.108	6.347	6.971	24.03	24.67	-0.258	-0.126
	0.8	-9.01	-2.54	28.99	30.12	-0.077	-0.044	6.360	7.176	23.51	26.91	-0.239	-0.112
	1	-9.33	-1.41	29.16	30.44	-0.080	0.041	6.207	6.916	23.71	27.35	-0.246	-0.089
20	0.6	-9.76	-3.83	28.49	30.41	-0.091	-0.183	6.393	6.922	23.82	24.94	-0.256	-0.136
	0.8	-9.20	-4.48	28.84	32.12	-0.063	0.021	6.499	6.589	23.75	29.06	-0.244	-0.151
	1	-9.38	-3.18	29.11	32.57	-0.080	-0.001	6.133	6.136	23.78	29.67	-0.247	-0.125
Panel D: UK Filter													
10	0.6	-19.68	-18.02	30.55	29.32	-0.006	-0.250	10.129	5.788	24.55	28.13	-0.449	-0.429
	0.8	-17.60	-18.01	29.29	30.23	-0.109	-0.114	9.151	5.918	24.65	27.38	-0.412	-0.424
	1	-17.37	-15.93	29.84	29.62	0.020	-0.404	10.024	4.821	24.47	29.50	-0.405	-0.389
15	0.6	-17.78	-10.08	29.25	31.89	-0.336	-0.712	7.537	4.866	26.08	30.75	-0.419	-0.266
	0.8	-18.35	-7.27	29.06	31.87	-0.241	-0.628	8.014	5.012	25.46	30.85	-0.430	-0.209
	1	-18.31	-10.37	29.00	30.59	-0.235	-0.658	8.539	4.740	25.51	29.82	-0.429	-0.273
20	0.6	-17.05	-11.75	28.76	30.28	-0.361	-0.703	7.774	4.804	26.08	30.94	-0.406	-0.304
	0.8	-18.61	-8.36	28.68	28.51	-0.323	-0.723	7.597	4.314	26.13	30.49	-0.438	-0.236
	1	-16.23	-9.99	28.20	28.48	-0.362	-0.808	7.526	5.115	25.88	29.77	-0.390	-0.269
Panel E: US Filter													
10	0.6	-3.49	-8.14	18.71	36.75	0.378	-0.125	16.744	7.485	17.50	27.28	-0.133	-0.213
	0.8	-3.68	-16.65	18.87	30.59	0.487	-0.031	19.319	6.308	16.69	22.75	-0.137	-0.385
	1	-2.89	-4.91	18.82	31.54	0.344	0.182	19.821	11.993	16.14	25.83	-0.120	-0.157
15	0.6	-3.21	-11.63	18.93	32.85	0.531	0.528	18.389	16.749	17.68	25.64	-0.127	-0.287
	0.8	-3.14	-9.86	18.98	35.24	0.467	0.240	20.067	12.534	16.81	26.05	-0.126	-0.248
	1	-3.39	-11.26	18.96	34.46	0.617	-0.104	21.177	8.574	17.00	27.26	-0.131	-0.279
20	0.6	-3.56	-13.83	19.05	33.98	0.526	0.374	17.797	15.279	17.95	24.28	-0.135	-0.324
	0.8	-3.32	-5.06	19.06	33.69	0.611	-0.091	20.461	7.872	16.95	27.94	-0.129	-0.159
	1	-2.69	-13.35	18.98	26.94	0.474	-0.361	21.563	9.317	17.35	24.35	-0.117	-0.332

Panel F: Indexes	**An.Ret. (%)**	**An.Vol. (%)**	**Skewn.**	**Ex. Kurt.**	**Correl. (%)**	**Info Ratio**
SEI	3.01	48.40	0.094	2.283	-	-
Bovespa	13.21	38.04	0.026	4.875	20.09	0.185
DJIA	-7.07	28.03	-0.053	4.636	12.90	-0.191
FTSE 100	-6.01	27.42	-0.009	5.374	24.34	-0.182
S&P500	-9.46	30.07	-0.162	5.999	14.51	-0.235
DJ UBS Energy-TR	-18.94	36.21	-0.166	1.102	43.83	-0.477
Rogers Energy Commodity-TR	-6.15	41.11	-0.189	2.099	44.02	-0.192

Table 3: Distributional statistics of portfolios' daily returns.

For further details, see notes in previous table.

		Monthly Rebalancing											
		An. Ret (%)		An. Vol. (%)		Skewness		Ex. Kurtosis		Correl.		Info Ratio	
(K)	(λ)	DE	GA	DE	GA	DE	GA	DE	GA	DE	GA	DE	GA
Panel A: Bovespa													
10	0.6	-7.88	0.39	35.05	37.73	-0.685	-0.618	7.390	5.738	23.75	28.52	-0.207	-0.050
	0.8	-9.09	-8.87	34.67	36.78	-0.670	-0.653	7.242	6.686	23.79	28.43	-0.231	-0.229
	1	-10.74	-14.35	34.77	36.32	-0.651	-0.648	7.485	6.393	23.86	27.61	-0.262	-0.335
15	0.6	-7.77	1.41	35.05	37.50	-0.693	-0.384	7.545	6.491	23.76	29.05	-0.205	-0.031
	0.8	-9.27	-4.51	34.81	36.94	-0.667	-0.571	7.549	7.385	23.80	28.40	-0.234	-0.145
	1	-10.42	-7.52	34.78	36.19	-0.634	-0.405	7.463	7.580	23.87	27.44	-0.256	-0.203
20	0.6	-7.99	5.39	35.04	37.63	-0.689	-0.646	7.536	6.739	23.71	27.51	-0.209	0.045
	0.8	-9.30	1.69	34.81	36.13	-0.657	-0.598	7.467	5.913	23.79	27.05	-0.235	-0.025
	1	-10.62	-5.71	34.77	36.57	-0.647	-0.520	7.503	7.536	23.77	27.23	-0.260	-0.167
Panel B: DJIA													
10	0.6	-9.06	0.10	19.45	22.79	0.572	0.165	12.589	7.598	8.91	16.14	-0.239	-0.058
	0.8	-9.88	-4.85	19.62	22.96	0.554	0.422	13.159	10.442	9.21	15.63	-0.255	-0.156
	1	-10.14	-3.89	19.63	23.24	0.562	0.424	13.418	10.173	9.08	18.05	-0.260	-0.139
15	0.6	-9.68	-5.37	19.63	21.72	0.546	0.304	12.686	7.835	8.80	15.10	-0.251	-0.168
	0.8	-9.98	-3.15	19.61	22.41	0.573	0.270	13.150	8.099	9.17	16.58	-0.257	-0.123
	1	-9.96	-7.32	19.61	23.45	0.576	0.571	13.430	12.986	9.11	18.53	-0.257	-0.208
20	0.6	-9.96	-3.02	19.57	23.26	0.577	0.386	12.735	9.342	8.75	16.62	-0.256	-0.120
	0.8	-10.32	-4.56	19.63	22.81	0.567	0.174	13.190	8.242	8.93	16.41	-0.264	-0.151
	1	-9.73	-5.66	19.51	22.86	0.577	0.358	13.330	9.564	9.01	17.53	-0.252	-0.174
Panel C: FTSE 100													
10	0.6	-12.05	-3.54	26.26	28.79	0.005	0.008	6.062	8.298	24.46	32.96	-0.307	-0.138
	0.8	-14.67	-7.47	26.39	29.71	-0.016	0.097	5.871	6.692	24.45	32.29	-0.360	-0.219
	1	-15.51	-13.00	26.15	29.76	-0.029	0.100	5.730	8.650	24.13	33.24	-0.377	-0.336
15	0.6	-13.99	-5.23	26.23	29.08	-0.002	-0.179	6.251	6.001	24.45	30.85	-0.346	-0.171
	0.8	-12.96	-8.31	26.04	29.65	-0.080	0.059	6.050	8.000	24.08	32.12	-0.325	-0.236
	1	-14.61	-14.17	26.38	29.44	-0.058	-0.223	6.116	6.708	23.85	31.60	-0.357	-0.357
20	0.6	-14.77	-4.84	26.26	29.22	0.002	-0.311	6.298	6.610	24.27	31.24	-0.362	-0.163
	0.8	-14.16	-13.54	26.35	29.42	-0.011	-0.002	6.168	7.674	24.05	31.29	-0.349	-0.344
	1	-14.10	-12.11	26.43	29.22	0.019	0.015	6.227	6.851	23.96	32.02	-0.347	-0.316
Panel D: UK Filter													
10	0.6	-14.94	3.47	17.80	23.13	-1.134	-0.707	6.977	4.513	23.12	31.42	-0.377	0.010
	0.8	-14.14	-15.26	17.61	22.86	-1.060	-1.535	6.811	9.672	22.44	33.72	-0.360	-0.397
	1	-14.97	-31.53	17.68	23.42	-1.050	-0.925	6.879	5.862	22.50	33.09	-0.377	-0.746
15	0.6	-16.70	-7.75	17.72	23.56	-1.175	-0.839	7.074	4.375	23.16	31.61	-0.415	-0.231
	0.8	-15.13	-9.62	17.72	23.55	-1.145	-0.929	7.070	5.514	22.89	32.41	-0.381	-0.272
	1	-16.69	-24.28	17.69	23.71	-1.112	-1.054	6.971	6.130	22.79	32.60	-0.414	-0.587
20	0.6	-16.25	-9.78	17.71	24.10	-1.167	-0.912	6.890	4.482	23.02	30.10	-0.405	-0.271
	0.8	-15.94	-16.90	17.67	24.16	-1.140	-0.819	6.983	4.836	22.99	32.15	-0.398	-0.427
	1	-15.83	-16.99	17.64	24.40	-1.105	-0.867	6.832	5.252	22.65	31.68	-0.395	-0.427

Panel E: US Filter													
10	0.6	-8.31	20.89	19.22	26.62	-0.755	-0.140	19.511	10.044	18.65	31.80	-0.233	0.379
	0.8	-12.19	18.31	20.26	25.75	-0.742	-0.373	24.991	11.898	17.41	32.28	-0.309	0.326
	1	-14.33	-4.00	20.48	24.96	-0.954	-0.260	26.671	11.698	16.91	32.58	-0.352	-0.150
15	0.6	-9.52	34.28	20.19	27.25	-0.831	0.012	25.504	10.244	17.46	27.74	-0.255	0.645
	0.8	-11.48	8.34	20.25	26.54	-0.773	-0.118	24.625	16.027	17.59	32.31	-0.295	0.113
	1	-13.33	2.33	20.17	26.56	-0.870	-0.170	25.108	12.796	17.60	33.33	-0.333	-0.014
20	0.6	-9.84	23.41	20.28	27.26	-0.859	-0.280	25.937	7.271	17.37	30.34	-0.262	0.427
	0.8	-12.20	8.29	20.19	25.34	-0.853	0.180	24.818	9.723	17.25	31.49	-0.310	0.112
	1	-13.67	4.81	20.32	25.50	-0.836	-0.367	26.336	12.638	17.48	32.73	-0.340	0.039

Panel F: Indexes	**An.Ret. (%)**	**An.Vol. (%)**	**Skewn.**	**Ex. Kurt.**	**Correl. (%)**	**Info Ratio**
SEI	3.01	48.40	0.094	2.283	-	-
Bovespa	13.21	38.04	0.026	4.875	20.09	0.185
DJIA	-7.07	28.03	-0.053	4.636	12.90	-0.191
FTSE 100	-6.01	27.42	-0.009	5.374	24.34	-0.182
S&P500	-9.46	30.07	-0.162	5.999	14.51	-0.235
DJ UBS Energy-TR	-18.94	36.21	-0.166	1.102	43.83	-0.477
Rogers Energy Commodity-TR	-6.15	41.11	-0.189	2.099	44.02	-0.192

Table 4: Distributional statistics of portfolios' daily returns.

For further details, see notes in previous table.

		Quarterly Rebalancing											
		An. Ret (%)		An. Vol. (%)		Skewness		Ex. Kurtosis		Correl.		Info Ratio	
(K)	(λ)	DE	GA	DE	GA	DE	GA	DE	GA	DE	GA	DE	GA
Panel A: Bovespa													
10	0.6	-6.79	6.38	35.68	38.32	-0.572	-0.588	7.688	7.146	23.76	27.67	-0.185	0.064
	0.8	-8.04	-7.48	35.39	36.15	-0.541	-0.499	7.696	7.198	23.72	26.04	-0.209	-0.200
	1	-8.88	-2.94	35.49	37.28	-0.537	-0.565	7.846	7.791	23.62	26.46	-0.225	-0.113
15	0.6	-7.36	-0.73	35.72	38.38	-0.578	-0.516	7.699	7.113	23.84	28.06	-0.196	-0.071
	0.8	-7.86	-4.05	35.49	37.33	-0.548	-0.620	7.910	7.932	23.79	26.89	-0.206	-0.134
	1	-8.14	-4.87	35.45	36.76	-0.532	-0.461	7.734	7.889	23.65	25.36	-0.211	-0.149
20	0.6	-7.49	8.27	35.73	38.45	-0.570	-0.494	7.661	7.896	23.95	26.36	-0.199	0.099
	0.8	-7.77	3.01	35.42	37.53	-0.544	-0.481	7.675	7.498	23.57	26.21	-0.204	0.000
	1	-8.62	-2.29	35.50	37.69	-0.534	-0.485	7.801	8.467	23.64	25.94	-0.220	-0.100
Panel B: DJIA													
10	0.6	-4.61	-3.13	19.76	22.72	0.543	0.329	12.944	9.405	8.96	13.36	-0.151	-0.121
	0.8	-5.14	-3.87	19.79	22.40	0.563	0.444	13.201	9.707	9.13	13.44	-0.161	-0.136
	1	-4.90	-1.33	19.76	22.87	0.630	0.437	13.884	10.343	8.97	14.63	-0.156	-0.086
15	0.6	-4.48	-1.33	19.85	22.44	0.536	0.405	12.659	10.195	9.01	13.63	-0.148	-0.086
	0.8	-4.83	-1.83	19.80	23.63	0.563	0.210	13.169	8.742	9.04	14.64	-0.155	-0.095
	1	-4.91	-4.12	19.87	24.36	0.600	0.475	13.712	12.793	8.97	15.65	-0.156	-0.141
20	0.6	-4.87	2.88	19.84	22.41	0.543	0.335	12.801	7.553	9.00	12.49	-0.156	-0.002
	0.8	-4.58	-5.36	19.83	24.40	0.542	0.355	13.054	9.969	9.07	16.10	-0.150	-0.165
	1	-4.75	1.72	19.86	23.42	0.587	0.526	13.684	10.842	8.93	15.57	-0.153	-0.026
Panel C: FTSE 100													
10	0.6	-8.03	5.68	25.87	28.61	0.040	-0.010	5.981	6.623	24.57	30.30	-0.225	0.056
	0.8	-8.96	-3.41	25.82	29.42	-0.019	0.082	5.743	8.084	24.11	30.01	-0.244	-0.132
	1	-8.62	-5.42	26.14	28.74	0.039	0.018	6.319	8.876	24.07	28.52	-0.236	-0.173
15	0.6	-8.78	-1.54	26.18	29.32	0.006	0.060	6.170	7.373	25.07	31.08	-0.241	-0.094
	0.8	-7.49	-0.19	26.03	28.89	0.004	-0.026	6.140	7.309	24.12	29.36	-0.214	-0.066
	1	-8.47	-5.78	26.26	30.48	-0.016	-0.106	6.310	7.594	24.01	30.57	-0.233	-0.180
20	0.6	-8.12	0.68	26.12	29.30	0.033	0.091	6.108	7.646	25.00	29.88	-0.228	-0.048
	0.8	-8.22	-0.64	26.12	29.07	-0.023	0.076	6.140	7.321	24.23	30.02	-0.229	-0.075
	1	-8.32	-2.24	26.17	29.43	-0.037	0.068	6.138	7.613	23.62	29.92	-0.230	-0.108
Panel D: UK Filter													
10	0.6	-14.16	2.21	18.43	23.56	-1.545	-0.908	11.532	5.806	22.81	29.94	-0.360	-0.017
	0.8	-12.40	-7.37	18.40	23.53	-1.540	-1.322	11.741	8.974	22.59	30.14	-0.323	-0.221
	1	-12.91	-23.42	18.47	22.11	-1.506	-1.353	11.692	9.453	22.38	28.14	-0.333	-0.560
15	0.6	-14.91	-5.58	18.45	23.98	-1.556	-0.908	11.403	4.967	23.06	28.91	-0.376	-0.181
	0.8	-14.81	-7.32	18.57	23.19	-1.602	-1.126	12.077	6.813	22.93	30.08	-0.373	-0.220
	1	-12.13	-8.57	18.59	24.84	-1.560	-0.947	11.759	5.099	22.40	30.45	-0.317	-0.245
20	0.6	-15.06	-8.22	18.38	24.71	-1.595	-1.115	11.618	6.180	22.97	29.35	-0.379	-0.237
	0.8	-14.55	-6.86	18.38	24.85	-1.600	-0.995	11.910	5.192	22.74	30.23	-0.368	-0.209
	1	-14.03	-9.44	18.48	23.93	-1.611	-1.037	11.846	6.240	22.36	30.48	-0.357	-0.265

Panel E: US Filter													
10	0.6	-6.16	9.29	20.51	26.77	-0.303	0.650	28.721	16.322	17.51	26.39	-0.187	0.129
	0.8	-5.70	9.06	20.64	24.56	-0.246	0.018	27.642	6.268	16.95	28.30	-0.177	0.127
	1	-6.23	-1.33	20.68	24.22	-0.289	-0.217	29.105	7.641	17.55	29.06	-0.188	-0.091
15	0.6	-5.12	18.48	20.57	26.87	-0.252	-0.104	28.952	5.516	17.48	25.97	-0.165	0.317
	0.8	-5.47	3.41	20.63	25.42	-0.200	-0.165	28.577	8.188	17.38	28.33	-0.172	0.008
	1	-5.62	8.15	20.73	24.86	-0.194	0.000	28.466	6.699	17.42	27.10	-0.175	0.107
20	0.6	-3.91	11.69	20.58	27.18	-0.289	-0.154	28.874	5.360	17.46	26.41	-0.141	0.178
	0.8	-5.27	12.30	20.65	26.32	-0.206	0.287	28.549	7.590	17.31	27.99	-0.168	0.193
	1	-5.87	6.19	20.84	26.44	-0.235	0.371	28.229	11.545	17.32	29.28	-0.180	0.067

Panel F: Indexes	**An.Ret. (%)**	**An.Vol. (%)**	**Skewn.**	**Ex. Kurt.**	**Correl. (%)**	**Info Ratio**
SEI	3.01	48.40	0.094	2.283	-	-
Bovespa	13.21	38.04	0.026	4.875	20.09	0.185
DJIA	-7.07	28.03	-0.053	4.636	12.90	-0.191
FTSE 100	-6.01	27.42	-0.009	5.374	24.34	-0.182
S&P500	-9.46	30.07	-0.162	5.999	14.51	-0.235
DJ UBS Energy-TR	-18.94	36.21	-0.166	1.102	43.83	-0.477
Rogers Energy Commodity-TR	-6.15	41.11	-0.189	2.099	44.02	-0.192

Furthermore, moving from no rebalancing to monthly rebalancing, the information ratios tend to go down in all cases, except in the case of the US Filter baskets for GA, and that of the UK Filter baskets for both DE and GA. This can be explained by the higher transaction costs which have a greater impact on the portfolios' returns, especially during falling markets. It can be argued that when rebalancing, the additional information available from the latest price data does make a difference on reducing the portfolios' volatility, but the small return improvement coupled with the rebalancing costs out-weighs the volatility benefits. Results are consistent for all cases for the risk-return trade-off λ. Among monthly and quarterly rebalancing the differences are relatively small, but the information ratios are in all cases higher for the monthly rebalanced portfolios, with only one exception for the FTSE selected baskets. This is an indication that greater capital efficiency can be achieved with the more frequent rebalancing. Under the buy-and-hold scenario, the best performance in terms of information ratios is reported for the Bovespa portfolios, and under both monthly and quarterly rebalancing it is reported for the US Filter portfolios. In most cases, negative information ratios are reported, indicating that these portfolios over the out-of-sample period under-perform the benchmark as they are associated with the lowest excess returns[11]. This observation can be explained by the fact that energy markets, as represented by the SEI, have been resistant to the recent economic recession, even though they have experienced one of their most severe up- and down-trends in their history.

Historically it has been shown that commodities have had an equity-like risk/ return profile, while at the same time being negatively correlated with stocks. Moreover, financial activity in commodity markets during the past decade has grown too much in size relative to physical production, leading to non-commercial net long positions to be less influenced by the commodities' diversification benefits observed in the past (Domanski and Heath, 2007). Looking at tables 3, 4, and 5, it can be seen that when switching from quarterly to monthly rebalancing, correlations tend to marginally improve, with results being more profound for the baskets selected by the GA. The relatively low correlations of the selected equity portfolios with the SEI (between 9% and 33%) suggest that investors who want to participate in the energy sector can still benefit from the addition of the selected baskets to a well diversified portfolio of assets. This observation aligns with the findings of Buyuksahin et al. (2010) that the correlation between equity and commodity returns is not often greater than 30%, besides some noticeable fluctuation that occurs over time. Also, correlation is not the most appropriate performance measure, as it only measures the degree to which the selected equity baskets and the SEI move in tandem, and does not capture the magnitude of the returns and their trajectories over time. Moreover, as it is well documented in the literature and also verified in the results presented in this paper, equity returns, represented by the financial indexes and the selected portfolios, deviate from a normal distribution displaying skewness and fat tails. The same is true for the returns of the SEI which exhibit positive skewness and relatively high excess kurtosis. Both futures commodity indexes have excess kurtosis similar to the SEI, with their skewness however being negative. Most equity portfolios selected by both the DE and GA exhibit negative skewness, indicating that the equity portfolios have more weight in the left tail of the distribution in contrast with the SEI that has more weight in the right tail.

[11] Note that investors who would have taken short positions on these baskets would realise the highest excess returns.

Moreover, looking at table 6 it can be concluded that the strategy and methodology used in this paper is much more efficient than a “naïve” strategy of randomly selected stocks, forming equally weighted portfolios constituted of 10, 15, and 20 stocks respectively. The evidence concur that this happens for both, achieving a good tracking performance (low RMSEs), and good returns relative to the SEI (positive or very small negative ERs). Under the “naïve” strategy there is a large dispersion of outcomes and no consistency, e.g. for the UK Filter portfolios with 10, 15 and 20 randomly selected socks, the respective information ratios are -0.62%, 0.09% and -0.12%.

Table 5: Performance of randomly selected portfolios.

This table presents a "Naïve" investment strategy of randomly selected stocks forming equally weighted portfolios consisting in each case by 10, 15 and 20 stocks, respectively. The stocks are selected from the same five equity pools used by the EAs, from a uniform distribution, thus giving equal probability for all stocks to be chosen.

	No Stocks	RMSE	ER (%)	An. Ret (%)	An. Vol. (%)	Skewness	Ex. Kurtosis	Correl. (%)	Info Ratio
Bovespa	10	0.04	-0.01	1.32	45.20	-0.20	6.10	21.44	-0.03
	15	0.04	0.03	9.73	45.31	-0.41	6.41	22.37	0.12
	20	0.04	0.02	7.80	42.79	-0.30	6.64	21.35	0.08
DJIA	10	0.04	-0.06	-12.05	35.64	-0.07	2.84	5.62	-0.26
	15	0.03	-0.02	-2.80	28.90	-0.19	4.03	12.56	-0.11
	20	0.03	-0.03	-3.62	30.57	-0.14	3.14	10.69	-0.12
FTSE 100	10	0.03	-0.04	-6.30	28.22	0.30	7.78	23.98	-0.19
	15	0.04	-0.09	-19.96	43.62	-0.02	4.35	25.15	-0.41
	20	0.03	-0.03	-5.80	41.27	-0.20	3.73	29.49	-0.16
UK_FILTER	10	0.04	-0.14	-31.62	39.15	-2.00	20.65	18.78	-0.62
	15	0.03	0.02	7.90	35.80	-0.54	4.71	26.38	0.09
	20	0.03	-0.02	-3.00	26.57	-0.48	3.52	24.72	-0.12
US_FILTER	10	0.03	-0.06	-10.97	38.48	-0.76	7.52	23.03	-0.26
	15	0.03	-0.04	-6.00	33.87	0.10	10.88	27.64	-0.18
	20	0.03	-0.04	-7.97	40.37	-0.44	7.40	29.29	-0.21

In addition, looking at the no rebalancing strategy in table 7 it can be observed that both algorithms in most cases do not utilise the maximum number of stocks allowed to select. The case is stronger for the GA selected portfolios. For instance, for all λ scenarios and for K=20, the maximum number of stocks selected in the case of the Bovespa, DJIA, and FTSE 100 stock pools is 8, 7, and 10 respectively. A general observation that can be made is that the algorithms tend to utilise almost the maximum number of available stocks when choosing from the UK Filter and US Filter pools. This can be justified by the fact that because only energy related stocks are included in the pools, there can be more stock combinations identified for inclusion in the selected portfolios, capable of tracking the performance of the SEI. Moreover, between the two evolutionary algorithms, the DE tends to use more stocks in the various selected portfolios, reaching the maximum number allowed most of the times. Finally, the DE is more stable in the number of stocks picked between the various cases of the risk/ return trade-off, whereas the GA tends to select portfolios quite different in terms of their composition. This can be confirmed by the much higher total number of stocks selected during all rebalancing frequencies, for both quarterly and monthly rebalancing strategies. For example, under monthly rebalancing and K=15, irrespectively of λ, the maximum total number of stocks that the DE selects is 49 and 45 for the FTSE 100 and US Filter baskets, while the GA selects 70 and 65 stocks respectively.

Table 6: Statistics of Portfolios (number of stocks used from algorithms).

Over the whole out-of sample period, "No Reb", "Q Reb" and "M Reb" shows the total number of stocks selected in each tracking portfolio i.e. under No rebalancing, Quarterly rebalancing and Monthly rebalancing, respectively. Note that "No Reb" is also the initial number of selected stocks for both "Q Reb" and "M Reb" because at t0=0 the estimation period is the same for all three rebalancing frequencies; hence, the number of stocks involved is identical. For further details, see also table 5-2.

		No Reb		**Q Reb**		**M Reb**	
(K)	**(λ)**	**DE**	**GA**	**DE**	**GA**	**DE**	**GA**
Panel A: Bovespa							
10	0.6	10	7	19	22	22	38
	0.8	10	5	19	25	25	34
	1	10	6	22	20	23	32
15	0.6	10	5	20	23	24	39
	0.8	11	6	20	24	25	36
	1	10	3	20	23	25	34
20	0.6	11	8	20	36	25	47
	0.8	10	8	21	30	25	42
	1	10	7	22	30	24	44
Panel B: DJIA							
10	0.6	10	5	24	23	31	30
	0.8	10	3	23	23	29	34
	1	10	3	23	27	27	38
15	0.6	15	4	31	28	35	37
	0.8	15	3	29	30	32	38
	1	15	2	29	27	32	38
20	0.6	17	6	31	36	36	42
	0.8	20	5	32	32	33	39
	1	19	7	33	35	32	43
Panel C: FTSE 100							
10	0.6	10	9	33	41	41	58
	0.8	10	4	32	43	40	61
	1	10	2	34	41	42	62
15	0.6	15	9	43	46	49	70

	0.8	15	7	40	47	46	66
	1	15	8	39	48	47	60
20	0.6	16	10	44	51	48	64
	0.8	17	10	42	50	48	63
	1	16	6	38	50	48	64
Panel D: UK Filter							
10	0.6	10	10	28	31	30	37
	0.8	10	5	26	24	29	37
	1	10	10	26	28	28	36
15	0.6	15	14	31	35	34	39
	0.8	15	15	30	37	33	40
	1	15	15	30	39	32	40
20	0.6	16	20	33	39	36	40
	0.8	17	20	30	40	34	41
	1	18	19	31	39	33	41
Panel E: US Filter							
10	0.6	10	10	25	43	38	54
	0.8	10	10	25	40	33	56
	1	10	10	29	45	34	64
15	0.6	15	11	34	44	44	61
	0.8	15	12	33	42	45	65
	1	15	15	35	51	40	64
20	0.6	16	12	35	50	43	65
	0.8	16	10	34	56	44	69
	1	16	19	34	58	39	72

6. Conclusions

In this paper, a Geometric Average Spot Energy Index is constructed and then its performance is being reproduced with stock portfolios. This is achieved by investing in small baskets of equities, selected from five stock pools, the Dow Jones, FTSE 100, Bovespa Composite, and the UK and US Filters. The investment methodology used employs two advanced EAs, the GA and the DE. Both algorithms are self-adaptive stochastic optimization methods, superior to other rival approaches when applied to the index tracking problem. To test the performance of the tracking baskets three different rebalancing scenarios are examined, also taking transaction costs into consideration: a) buy-and-hold, b) monthly rebalancing, and c) quarterly rebalancing. For comparison reasons the performance of a “naïve” investment strategy of randomly selected stocks forming equally weighted portfolios is also reported.

It is found that energy commodities, as proxied by the SEI, can have equity-like returns, since they can be effectively tracked with stock portfolios selected by the investment methodology followed in this paper. Overall, during the three-year period examined, which reflects a period before, during and towards the end of the recent global economic recession, an investor would realise positive returns by investing in commodities, as the SEI returns suggest. With the methodology employed that performance is closely replicated, and in the case of the energy related stock portfolios and those selected from the Bovespa equity pool, the benchmark index is even outperformed. In most cases there seem to be no major differences between the DE and GA selected portfolios, though the GA tends to select portfolios that have a lower tracking error. Both algorithms, in most cases, do not utilise the maximum number of stocks allowed to select, with the DE being more stable in the number of stocks picked between the various cases of the risk/ return trade-off; the GA tends to select portfolios quite different in terms of their composition.

On average, based on the results of this paper, portfolios with 15 stocks and a risk-return trade-off value of 0.8 are the most desirable combination providing the best results for most tracking portfolios. Also, it is found that when rebalancing, the additional information available from the latest price data does make a difference on reducing the portfolios' volatility; the resulting return deterioration however, out-weighs the volatility benefits leading to smaller information ratios. Moving from the Buy and Hold strategy to Quarterly Rebalancing and then to the more frequent Monthly Rebalancing strategy, returns tend to deteriorate for most selected portfolios, by both the DE and the GA. Nonetheless, the same holds for the portfolios' volatilities that also tends to go down when moving from no rebalancing to the more frequent one. Between monthly and quarterly rebalancing the differences are relatively small in terms of the portfolios' return and volatility performance; however the information ratios are in almost all cases higher for the quarterly rebalanced portfolios. The only exception is for the US Filter in the case of the baskets selected by the GA. Thus, it is concluded that greater capital efficiency can be achieved with rebalancing, preferably every quarter, compared to the buy-and-hold strategy.

The investment approach proposed in this paper, for tracking the performance of the energy sector with stocks selected by two innovative evolutionary algorithms, promotes a cost effective implementation and true investability. While most mutual funds cannot invest in commodities directly, they can track the performance of the SEI by investing in the stocks selected by the evolutionary algorithms used in this paper. There are many investment houses around the globe that use evolutionary algorithms for tactical asset management[12]. The work and findings presented in this paper can encourage asset and fund managers to recognise the importance of the energy sector and prompt them to set-up similar funds that will track the constructed Spot Energy Index. To that end, the proposed methodology suggests an effective, and at the same time, least expensive way to operate such a fund, giving the full flexibility of any investment style, long or short, that equities can provide.

[12] First Quadrant a US based investment firm started using EAs in 1993 to manage its investments, at the time $5 billion USD allocated across 17 countries around the globe, claiming that have made substantial profits (Kieran, 1994).

7. References

AIA Research Report, 2008. Comparing commodity indices: multiple approaches to return, Alternative Investment Analytics LLC, Amherst, MA.

Alizadeh, A., Nomikos, N., and Pouliasis, P., 2008. A Markov regime switching approach for hedging energy commodities. Journal of Banking & Finance 32, 1970-1983.

Al-Mudhaf, A. Goodwin, T.H., 1993. Oil shocks and oil stocks: Evidence from the 1970s. Applied Economics 25, 181-190.

Allen, F., Karjalainen, R., 1999. Using genetic algorithms to find technical trading rules. Journal of Financial Economics 51, 245-271.

Alexander, C., Dimitriu, A., 2002. The cointegration alpha: enhanced index tracking and long-short equity market neutral strategies. ISMA Discussion Papers in Finance, 2002-08, ISMA Centre, University of Reading.

Anson, M. J.P., 1999. Maximizing utility with commodity futures diversification. Journal of Portfolio Management 25 (4), 86-94.

Bauer, R.J. Jr., 1994. Genetic Algorithms and Investment Strategies. Wiley, New York.

Barber, B.M., Odean, T., 2000. Trading is hazardous to your wealth: the common stock investment performance of individual investors. The Journal of Finance 55(2), 773-806.

Barberis, N., Shleifer, A., 2003. Style investing. Journal of Financial Economics 68, 161-199.

Beasley, J.E., Meade, N., Chang, T.-J., 2003. An evolutionary heuristic for the index tracking problem. European Journal of Operational Research 148(3), 621-643.

Black, F., 1976. The pricing of commodity contracts. Journal of Financial Economics 3, 167-179.

Bodie, Z., 1980. An innovation for stable real retirement income. Journal of Portfolio Management 7 (1), 5-13.

Bodie, Z., Rosansky, V., 1980. Risk and return in commodity futures. Financial Analyst Journal 36 (3), 27-39.

Boyer, M.M., Filion, D., 2007. Common and fundamental factors in stock returns of Canadian oil and gas companies. Energy Economics 29, 428-453.

Buckley, I.R.C., Korn, R., 1998. Optimal index tracking under transaction costs and impulse control. International Journal of Theoretical and Applied Finance 1 (3), 315-330.

Cambell, J., Cochrane, J., 1999. By force of habit: A consumption-based explanation of aggregate stock market behaviour. Journal of Political Economy 107, 205-251.

Canakgoz, N.A., & Beasley, J.E., 2008. Mixed-integer programming approaches for index tracking and enhanced indexation. European Journal of Operational Research 196, 384-399.

Chang, T.J., Yang, S.C., Chang, K.J., 2009. Portfolio optimization problems in different risk measures using genetic algorithm. Expert Systems with Applications 36, 10529-10537.

Chen, C., & Kwon, R.H., 2012. Robust portfolio selection for index tracking. Computers & Operations Research 39, 829-837.

Coello Coello, C.A., 1999. A comprehensive survey of evolutionary-based multi-objective optimization techniques. Knowledge and Information Systems 1 (3), 269-308.

Commodity Futures Trading Commission, 2008. Staff report on commodity swap dealers and index traders with commission recommendations (September), 1-70.

Connor, G., Leland, H., 1995. Cash management for index tracking. Financial Analysts Journal 51 (6), 75-80.

Deb, K., 2001. Multi-objective Optimization using Evolutionary Algorithms. John Wiley & Sons, New York.

Domanski, D., Heath, A., 2007. Financial investors and commodity markets. Bank for International Settlements, Quarterly Review, March.

Dorsey, R.E., Mayer, W.J., 1995. Genetic algorithms for estimation problems with multiple optima, non-differentiability, and other irregular features. Journal of Business and Economic Statistics 13, 53-66.
Dose, C., Cincotti, S., 2005. Clustering of financial time series with application to index and enhanced index tracking portfolio. Physica A 355, 145-151.
Driesprong, G., Jacobsen, B., and Maat, B., 2008. Striking oil: Another puzzle? Journal of Financial Economics 89, 307-327.
Dunis, C.L., Ho, R., 2005. Cointegration portfolios of European equities for index tracking and market neutral stategies. Journal of Asset Management 6(1), 33-52.
Erb, C.B., Harvey, C.R., 2006. The strategic and tactical value of commodity futures. Financial Analyst Journal 62 (2), 69-97.
Famma, E. F., French, K. R., 1987. Commodity futures prices: some evidence on forecast power, premiums and the theory of storage. Journal of Business 60 (1), 55-73.
Feoktistov, V., Janaqi, S., 2004. Generalization of the strategies in differential evolution. In: The 18th International Parallel and Distributed Processing Symposium, Santa Fe, New Mexico, USA..
Frankel, J.A., 2008. The effect of monetary policy on real commodity prices. In John Cambell (ed.). Asset prices and Monetary policy. University of Chicago Press, 291-327.
Franks, E.C., 1992. Targeting excess-of-benchmark returns. The Journal of Portfolio Management 18 (4), 6-12.
Frino, A., Gallagher, D.R., 2001. Tracking S&P 500 index funds. The Journal of Portfolio Management 28(1) 44-55.
Gaivoronski, A.A., Krylov, S., Van der Wijst, N., 2005. Optimal portfolio selection and dynamic benchmark tracking. European Journal of Operational Research 163, 115-131.
Gilli, M., Kellezi, E., 2001. Threshold accepting for index tracking. Working Paper, Department of Econometrics, University of Geneva.
Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization & Machine Learning, Addison-Wesley.
Gorton, G., Rouwenhorst, G., 2006. Facts and fantasies about commodity futures. Financial Analyst Journal 62 (2), 47-68.
Hammoudeh, S., Dibooglu, S., Aleisa, E., 2004. Relationships among US oil prices and oil industry equity indices. International Review of Economics and Finance 13, 427-453.
Hamilton, J.D., 2009. Causes and consequences of the oil shock of 2007-08. Working Paper, University of California San Diego (March).
Hamilton, J.D., 2008. Understanding Crude Oil Prices. NBER Working Paper Series, No. 14492, November.
Hamilton, J.D., 2003. What is an oil shock? Journal of Econometrics 113 (2), 363-398.
Hong, H., Torous, W., and Valkanov, R., 2007. Do industries lead stock markets? Journal of Financial Economics 83, 367-396.
Holland, J. H., 1975. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor.
Holland, J. H., 1962. Outline for a logical theory of adaptive systems. Journal of the Association for Computing machinery 3, 297-314.
Hughes, M., 1990. Improving products and processes – nature's way. Industrial Management and Data Systems 90 (6), 22-25.
Jaeger, L., Cittadini, P., Jacquemai, M., 2002. The sGFI futures index. Journal of Alternative Investments 5 (10), 73-80.
Jensen, G.R., Johnson, R.R., and Mercer, J.M., 2000. Efficient use of commodity futures in diversified portfolios. Journal of Futures Markets 20 (5), 489-506.
Jones, C.M., Kaul, G., 1996. Oil and the stock markets. The Journal of Finance 51 (2), 463-491.

Kieran, V., 1994. Growing money from algorithms. New Scientist, Dec. 94, No. 1954.

Killian, L., 2006. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. Working Paper, University of Michigan (September).

Kingdon, J., Feldman, K., 1995. Genetic algorithms and applications to finance. Applied Mathematical Finance 2, 89-116.

Krink, T., & Paterlini, S., 2009. Multiobjective optimization using differential evolution for real-world portfolio optimization. Computational Management Science 8, 157-179.

Krink, T., Mittnik, S., & Paterlini, S., 2009. Differential evolution and combinatorial search for constrained index-tracking. Annals of Operations Research 172, 153-176.

Konno, H., Hatagi, T., 2005. Index-plus-alpha tracking under concave transaction cost. Journal of Industrial and Management Optimisation 1(1), 87-98.

Kyle, A., Xiong, W., 2001. Contangion as a wealth effect. Journal of Finance 56, 1401-1440.

Lamle, H.R., Martell, T.F., 2005. A new era for commodity investments. Journal of Financial Transformation 15, 1-6.

Larsen-Jr, G.A., Resnick, B.G., 1998. Empirical insights on indexing. The Journal of Portfolio Management 25(1), 51-60.

Leinweber, D., Arnott, R., 1995. Quantitative and Computational Innovation in Investment Management. Journal of Portfolio Management 21 (2), 8-15.

Li, Q., Sun, L., & Bao, L., 2011. Enhanced index tracking based on multi-objective immune algorithm. Expert Systems with Applications 38, 6101-6106.

Loraschi, A., Tettamanzi, A., 1996. An Evolutionary Algorithm for Portfolio Selection Within a Downside Risk Framework. In Dunis, C. (ed.), Forecasting Financial Markets, John Wiley, Chichester, pp. 275-285.

Mahfound, S., Mani, G., Reigel, S., 1997. Nonlinear versus linear techniques for selecting individual stocks. In Weigend, A.S., Abu-Mostafa, Y., Refenes, A-P. (eds.), Decision Technologies for Financial Engineering, World Scientific, Singapore, pp. 65-75.

Markowitz, H., 1952. Portfolio selection. The Journal of Finance 7, 77-91.

Marimon, R., McGrattan, E., Sargent, T.J., 1990. Money as a medium of exchange in an economy with artificial intelligent agents. Journal of Economic Dynamics and Control 14, 329-373.

Maringer, D., 2008. Constrained index tracking under loss aversion using differential evolution. In: Brabazon, A. & O'Neil, M. (eds.), Natural Computing in Computational Finance, Studies in Computational Intelligence 100, Springer, Berlin, 7-24.

Maringer, D., Oyewumi, O. 2007. Index tracking constrained portfolios. Intelligent Systems in Accounting, Finance and Management 15, 57-71.

Malkiel, B., 1995. Returns from investing in equity mutual funds 1971 to 1991. The Journal of Finance 50(2), 549-572.

Marshall, B.R., Cahan, R.H., and Cahan, J.M., 2008. Can commodity futures be profitably traded with quantitative market timing strategies? Journal of Banking & Finance 32, 1810-1819.

Meade, N., Beasley, J.E., 2004. An evaluation of passive strategies to beat the index. Working Paper Series, Tanaka Business School.

Meade, N., Salkin, G.R., 1989. Index funds – construction and performance measurement. Journal of the Operational Research Society 40, 871-879.

Miffre, J., Rallis, G., 2007. Momentum strategies in commodity futures markets. Journal of Banking & Finance 31, 1863-1886.

Neely, C.J., Weller, P., Dittmar, R., 1997. Is technical analysis in the foreign exchange market profitable? A genetic programming approach. Journal of Financial Quantitative Analysis 32 (4), 405-426.

Nijman, T., Swinkels, L., 2003. Strategic and tactical allocation to commodities for retirement savings schemes. Discussion Paper, Tilburg University.

Plaxco, L.M., Arnott, R.D., 2002. Rebalancing a global policy benchmark. Journal of Portfolio Management 28 (2), 9-22.

Politis, D. N., Romano, J. P., 1994. The stationary bootstrap. Journal of The American Statistical Association 89 (428), 1303–1313.
Pope, P.F., Yadav, P.K., 1994. Discovering errors in tracking error. The Journal of Portfolio Management 20 (2), 27-32.
Price, K.V., Storn, R.M., Lampinen, J.A., 2005. Differential Evolution: A Practical Approach to Global Optimization. Springer, Heidelberg.
Rohweder, H.C., 1998. Implementing stock selection ideas: does tracking error optimization do any good? The Journal of Portfolio Management 24 (3), 49-59.
Ruiz-Torrubiano, R., & Suárez, A., 2009. A hybrid optimization approach to index tracking. Annals of Operations Research 166. 57-71.
Sarker, R., Liang, K., Newton, C., 2002. A new multi-objective evolutionary algorithm. European Journal of Operational Research 140 (1), 12–23.
Sarker, R., Abbass, H., 2004. Differential evolution for solving multi-objective optimization problems. Asia-Pacific Journal of Operations Research 21(2), 225-240.
Schneeweis, T., Spurgin, R., 1997. Comparisons of commodity and managed futures benchmark indexes. The Journal of Derivatives 4 (4), 33-50.
Scholtens, B, Wang, L., 2008. Oil risk in oil stocks. The Energy Journal 29 (1), 89-112.
Sharpe, W., 1991. The arithmetic of active management. The Financial Analyst Journal 47 (1), 7-9.
Sharpe, W., 1992. Asset allocation: management style and perfomance measurement. Journal of Portfolio Management 18, 7-19.
Sorenson, E.H., Miller, K.L., Samak, V., 1998. Allocating between active and passive management. Financial Analyst Journal 54(5), 18-31.
Storn, R., Price, K., 1995. Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. Technical Report TR-95-012, International Computer Science Institute, Berkeley.
Storn, R., Price, K., 1997. Differential evolution, a simple and efficient heuristic strategy for global optimization over continuous spaces. Journal of Global Optimization 11, 341-359.
Sivanandam, S.N., Deepa, S.N., 2007. Introduction to Genetic Algorithms. Springer.
Stoyan, S.J., & Kwon, R.H., 2010. A two-stage stochastic mixed-integer programming approach to the index tracking problem. Optimization and Engineering 11(2), 247-275.
Sullivan, R., Timmermann, A., White, H., 1999. Data-snooping, technical trading rule performance, and the bootstrap. The Journal of Finance 54, 1647-1691.
Szakmary, A.C., Shen, Q., and Sharma, S.C., 2010. Trend-following trading strategies in commodity futures: A re-examination. Journal of Banking & Finance 34, 409-426.
Tabata, Y., Takeda, E., 1995. Bicriteria optimization problem of designing an index fund. Journal of the Operational Research Society 46, 1023-1032.
Tang, K., Xiong, W., 2009. Index and the financialization of commodities. Working Paper (September).
Wang, M.Y., 1999. Multiple-benchmark and multiple-portfolio optimization. Financial Analysts Journal 55 (1), 63-72.
Weiser, S., 2003. The strategic case of commodities in portfolio diversification. Commodities Now, 7-11.
White, H., 2000. A reality check for data snooping. Econometrica 68 (5), 1097–1126.
Wilmott, P., 1998. Derivatives: The Theory and Practice of Financial Engineering. Wiley, New York.
Zhang, J., & Maringer, D., 2010. Index mutual fund replication. In: Brabazon, A., O’Neil, M., & Maringer, D. (eds.), Natural Computing in Computational Finance, Studies in Computational Intelligence 293, Springer, Berlin, 109-130.
Zitzler, E., Thiele, L., 1999. Multi-objective evolutionary algorithms: a comparative case study and the strength pareto approach. IEEE Transactions on Evolutionary Computation 3 (4), 257–271.

Appendix 1: Industry Classification Benchmark (ICB)

The Industry Classification Benchmark (ICB) is a company classification system developed jointly by Dow Jones and FTSE. It is used to segregate markets into a number of sectors within the macro-economy. The ICB uses a system of 10 industries, partitioned into 19 super sectors, which are further divided into 41 sectors, which then contain 114 subsectors.

The principal aim of the ICB is to categorize individual companies into subsectors based primarily on a company's source of revenue or where it constitutes the majority of revenue. If a company is equally divided amongst several distinct subsectors, the judging panel from both Dow Jones and FTSE makes a final decision. Firms may appeal their classification at any time.

The ICB is used globally (though not universally) to divide the market into increasingly specific categories, allowing investors to compare industry trends between well-defined subsectors. The ICB replaced the old classification systems used previously by Dow Jones and FTSE on 3 January, 2006, and is used today by the NASDAQ, NYSE and several other markets around the globe. All ICB sectors are represented on the New York Stock Exchange except Equity Investment Instruments (8980) and Non-equity Investment Instruments (8990).

Table 8-1 below presents the ICB codes used for filtering all US and UK stock markets, creating the two energy-related stock pools named US Filter and UK Filter, respectively.

Table 8: Industry Classification Benchmark (ICB) codes

Industry	**Super-sector**	**Sector**	**Sub-sector**
0001 Oil & Gas	0500 Oil & Gas	0530 Oil & Gas Producers	0533 Exploration & Production
			0537 Integrated Oil & Gas
		0570 Oil Equipment, Services & Distribution	0573 Oil Equipment & Services
			0577 Pipelines
		0580 Alternative Energy	0583 Renewable Energy Equipment
			0587 Alternative Fuels
7000 Utilities	7500 Utilities	7530 Electricity	7535 Conventional Electricity
			7537 Alternative Electricity

Appendix 2: Stocks used in all five equity pools

The table below includes all stocks used in the five equity pools from which the final stock portfolios were selected by the two algorithms, GA and DE, respectively.

Table 9: List of all stocks used in each pool for the selection of the tracking stock portfolios.

FTSE 100 (98 stocks in total)	DJIA 65 (65 stocks in total)	Bovespa (56 stocks in total)	UK Energy Filter (54 stocks in total)	US Energy Filter (89 stocks in total)
3I GROUP	3M	ALL AMER LAT UNT	AFREN	ALON USA ENERGY
ADMIRAL GROUP	AES	AMBEV PN	ALKANE ENERGY	AMERICAN OIL & GAS
ALLIANCE TRUST	ALCOA	ARACRUZ PNB	ANDES ENERGIA	ARENA RES.
AMEC	ALEX.& BALDWIN	BANCO BRASIL ON	ASCENT RESOURCES	ATLAS AMERICA
ANGLO AMERICAN	AMER.ELEC.PWR.	BRADESCO PN	BALTIC OIL TERMINALS	ATP OIL&GAS
ANTOFAGASTA	AMERICAN EXPRESS	BRADESPAR PN	BORDERS & SOUTHERN PTL.	BASIC ENERGY SVS.
ASSOCIATED BRIT.FOODS	AMR	BRASIL TELCOM PARTP.PN	BOWLEVEN	BGE CAPITAL TST.II
ASTRAZENECA	AT&T	BRASIL TELECOM PN	CDS OIL & GAS GROUP	BILL BARRETT
AUTONOMY CORP.	BANK OF AMERICA	BRASKEM PNA	CERES POWER HOLDINGS	BOARDWALK PIPELINE PTNS.
AVIVA	BOEING	BRF FOODS ON	CIRCLE OIL	BRONCO DRILLING
BAE SYSTEMS	BURL.NTHN.SANTA FE C	CCR RODOVIAS ON	CLIPPER WINDPOWER (REGS)	CANO PETROLEUM
BALFOUR BEATTY	CATERPILLAR	CELESC PNB	D1 OILS	CHINA NTH.ET.PTL.HDG.
BARCLAYS	CENTERPOINT EN.	CEMIG PN	DRAX GROUP	CIMAREX EN.
BG GROUP	CH ROBINSON WWD.	COMGAS PNA	EGDON RESOURCES	CNX GAS
BHP BILLITON	CHEVRON	COMPANHIA BRASL.DISTB. PNA	EMPYREAN ENERGY	COMPLETE PRDN.SVS.
BP	CISCO SYSTEMS	COPEL PNB	ENCORE OIL	COPANO ENERGY
BRITISH AIRWAYS	COCA COLA	COSAN ON	EUROPA OIL & GAS (HDG.)	CROSSTEX EN.
BRITISH	CON-WAY	CPFL	FALKLAND OIL	CROSSTEX

AMERICAN TOBACCO		ENERGIA ON	& GAS	EN.SHBI
BRITISH LAND	CONSOLIDATED EDISON	CYRELA REALT ON	FAROE PETROLEUM	CUBIC ENERGY
BRITISH SKY BCAST.GROUP	CONT.AIRL.B	DURATEX PN	FORUM ENERGY	DAYSTAR TECHS.
BT GROUP	CSX	ELETROBRAS ON	FRONTERA RESOURCES	DCP MIDSTREAM PTNS.
BUNZL	DOMINION RES.	ELETROBRAS PNB	GETECH GROUP	DELEK US HOLDINGS
CABLE & WIRELESS	DUKE ENERGY	EMBRAER ON	GLOBAL ENERGY DEV.	DRESSER-RAND GROUP
CADBURY	E I DU PONT DE NEMOURS	GAFISA ON	GOOD ENERGY GROUP	DTE EN.TST.II GTD TOPRS
CAIRN ENERGY	EDISON INTL.	GERDAU PN	GULFSANDS PETROLEUM	DUNE ENERGY
CAPITA GROUP	EXELON	GOL PN	HALLIN MAR.SUBSEA INTL.	ENBRIDGE EN.MAN.
CARNIVAL	EXPEDITOR INTL.OF WASH.	ITAUSA PN	HARDY OIL & GAS	ENCORE ACQ.
CENTRICA	EXXON MOBIL	ITAUUNIBANCO PN	HYDRODEC GROUP	ENDEAVOUR INTL.
COBHAM	FEDEX	KLABIN SA PN	INDEPENDENT RESOURCES	ENERGY TRANSFER EQ.
COMPASS GROUP	FIRSTENERGY	LIGHT ON	IPSA GROUP	ENTERGY MS.6% 1ST.MGE. BDS.
DIAGEO	FPL GROUP	LOJAS AMERIC PN	ISLAND OIL AND GAS	ENTERPRISE GROUP HDG.
FOREIGN & COLONIAL	GATX	LOJAS RENNER ON	ITM POWER	EVERGREEN SOLAR
FRIENDS PROVIDENT GROUP	GENERAL ELECTRIC	METALURGICA GERDAU PN	LANSDOWNE OIL & GAS	EXCO RESOURCES
G4S	HEWLETT-PACKARD	NATURA ON	MAX PETROLEUM	FMC TECHNOLOGIES
GLAXOSMITHKLINE	HOME DEPOT	NET PN	MEDITERRANEAN OIL & GAS	GASCO EN.
HAMMERSON	HUNT JB TRANSPORT SVS.	PETROBRAS ON	MERIDIAN PETROLEUM	GEOPETRO RESOURCES
HOME RETAIL GROUP	INTEL	PETROBRAS PN	NAUTICAL PETROLEUM	GLOBAL ENERGY HDG.GP.
HSBC HDG. (ORD $0.50)	INTERNATIONAL BUS.MCHS.	ROSSI RESID ON	NOVERA ENERGY (LON)	GLOBAL PARTNERS UNITS

ICAP	JETBLUE AIRWAYS	SABESP ON	OFFS.HYDROCA RBON MAPPING	GMX RES.
ICTL.HTLS.GP.	JOHNSON & JOHNSON	SADIA PN	PANTHEON RESOURCES	GRAN TIERRA ENERGY
IMPERIAL TOBACCO GP.	JP MORGAN CHASE & CO.	SIDER.NACIO NAL ON	PETROFAC	GREEN PLAINS RENEW.EN.
INMARSAT	KRAFT FOODS	SOUZA CRUZ ON	PETROLATINA ENERGY	HECO CAPITAL TST.III 6.5%
INTERNATIONAL POWER	LANDSTAR SYSTEM	TAM PN	PLEXUS HOLDINGS	HERCULES OFFSHORE
INTERTEK GROUP	MCDONALDS	TELE NRLES.PARTP .ON	REGAL PETROLEUM	HILAND PARTNERS
INVENSYS	MERCK & CO.	TELE NRLES.PARTP .PN	RENEWABLE ENERGY GNRTN.	HOKU SCIENTIFIC
JOHNSON MATTHEY	MICROSOFT	TELEMAR NRLES.PNA	RENEWABLE ENERGY HDG.	HOLLY ENERGY PTNS.
KAZAKHMYS	NISOURCE	TELESP PN	RHEOCHEM	HORNBECK OFFS.SVS.
KINGFISHER	NORFOLK SOUTHERN	TIM PART ON	ROCKHOPPER EXPLORATION	HOUSTON AMERICAN EN.
LAND SECURITIES GROUP	OVERSEAS SHIPHLDG.GP.	TIM PART PN	RURELEC	ITC HOLDINGS
LEGAL & GENERAL	PFIZER	TRAN PAULIST PN	SERICA ENERGY (LON)	KINDER MORGAN MAN.
LIBERTY INTL.	PG&E	ULTRAPAR PARTP.PN	SOVEREIGN OILFIELD GP.	LINN ENERGY
LLOYDS BANKING GROUP	PROCTER & GAMBLE	USIMINAS ON	VENTURE PRODUCTION	MAGELLAN MIDSTREAM HDG.
LONDON STOCK EX.GROUP	PUB.SER.ENTE R.GP.	USIMINAS PNA	VICTORIA OIL & GAS	MAGELLAN MIDSTREAM PTNS. UTS.
LONMIN	RYDER SYSTEM	VALE ON	WOOD GROUP (JOHN)	MARINER ENERGY
MAN GROUP	SOUTHERN	VALE PNA		MARTIN MIDSTREAM PTNS.
MARKS & SPENCER GROUP	SOUTHWEST AIRLINES	VIVO PN		MIRANT
MORRISON(WM)SP MKTS.	TRAVELERS COS.			MMC ENERGY
NATIONAL GRID	UNION PACIFIC			NATURAL GAS SVS.GP.
NEXT	UNITED PARCEL SER.			NEW GNRTN.BIFL.H DG.
OLD MUTUAL	UNITED TECHNOLOGIE			NORTHWESTE RN

	S			
PEARSON	VERIZON COMMUNICATIONS			NRG ENERGY
PENNON GROUP	WAL MART STORES			NUSTAR ENERGY LP
PETROFAC	WALT DISNEY			OCEAN POWER TECHS.
PRUDENTIAL	WILLIAMS COS.			OIL STS.INTL.
RANDGOLD RESOURCES	YRC WORLDWIDE			OILSANDS QUEST
RECKITT BENCKISER GROUP				ORMAT TECHS.
REED ELSEVIER				PLAINS EXP.& PRDN.
REXAM				PORTLAND GEN.ELEC.
RIO TINTO				RAM ENERGY RESOURCES
ROLLS-ROYCE GROUP				RASER TECHS.
ROYAL BANK OF SCTL.GP.				REGENCY ENERGY PTNS.
ROYAL DUTCH SHELL A(LON)				RIO VISTA EN.PTNS.LP.
ROYAL DUTCH SHELL B				ROSETTA RESOURCES
RSA INSURANCE GROUP				RRI ENERGY
SABMILLER				SOUTH TEXAS OIL
SAGE GROUP				SUNOCO LOGIST.PTNS.LP
SAINSBURY (J)				SUNPOWER 'A'
SCHRODERS				SUPERIOR WELL SVS.
SCHRODERS NV				TEEKAY LNG PARTNERS
SCOT.& SOUTHERN ENERGY				TETON ENERGY
SERCO GROUP				TRANSMONTAIGNE PTNS.
SEVERN TRENT				TRICO MARINE SVS.
SHIRE				ULTRA PTL.
SMITH & NEPHEW				UNION DRILLING

SMITHS GROUP				W&T OFFSHORE
STANDARD CHARTERED				WARREN RESOURCES
TESCO				WESTERN REFINING
THOMAS COOK GROUP				WHITING PTL.
TUI TRAVEL				WILLIAMS PARTNERS
TULLOW OIL				
UNILEVER (UK)				
UNITED UTILITIES GROUP				
VEDANTA RESOURCES				
VODAFONE GROUP				
WOLSELEY				
WPP				
XSTRATA				