

A Work Project, presented as part of the requirements for the Award of a Master's degree in Finance from the Nova School of Business and Economics.

Commodity Terms-of-Trade as a Foreign Exchange Pairs Trading Signal

Ciaran Jack Callaghan

Work project carried out under the supervision of:

Nicholas H. Hirschey

20-12-2023

Abstract: Group Part

This study examines multi-asset portfolio construction using commodities, foreign exchange, and three equity ESG strategies in the Nordic market. Different allocation methods, including the Maximum Sharpe Ratio and asset weighted strategies, were analysed to understand how non-financial factors such as ESG and Commodity Terms of Trade Signals can impact portfolio performance. The Asset-Weighted strategy, with its strategic diversification across multiple asset classes, delivered consistently high returns and maintained elevated Sharpe Ratios, adeptly balancing risk and return in line with the shifting dynamics of the market.

Abstract: Individual Part

This paper analyses the usefulness of commodity terms of trade as a trading signal in a universe of advanced and emerging market currencies. Three variations of the strategy were tested; one which traded long and short positions, one only long positions and one only short positions. All were found to generate statistically significant alpha in out-of-sample testing. This paper contributes to the literature as it provides evidence that the uncovered interest rate parity condition does not always hold, as well as providing examples of how commodity terms of trade can be used as a trading signal to exploit these deviations.

Keywords: *Finance, Financial Markets, Data Analysis, Equity, Factor-Based Investing, ESG, Commodity Terms of Trade, Foreign Exchange, Currency, Commodities, Systematic Trading Strategy, Currency Pairs, Long Short, Sustainability, Sustainable Investment, Cointegration, , Futures, Quantitative Strategies, Pairs Trading, ESG Momentum, Nordic Stock Market, Sustainable Investment, ESG Rating, Machine Learning, Gradient Boosting Tree, SHAP*

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209)

1 Introduction

Constructing a multi-asset portfolio can be a valuable method of building a successful strategy, providing benefits from diversification and a balanced approach to managing risk and return. The key to this approach is leveraging the strengths of different asset classes to improve portfolio performance and reduce risk. This paper delves into the complexities of constructing a multi-asset portfolio, particularly evaluating various investment strategies.

Firstly, a literature review and economic rationale of individual strategies are provided, including an overview of Environmental, Social, and Governance investing (ESG), equity investment strategies, and strategies focusing on commodities and FX.

After providing a comprehensive explanation of each tactic, the techniques and fundamental principles of each approach are elucidated. Including ESG criteria in various strategies conforms to the rising inclination towards ethical investment and underscores its mounting significance in contemporary portfolio management.

To evaluate the effectiveness of the combined strategies in a multi-asset portfolio context, a detailed analysis of the portfolios' performance and characteristics is provided in the subsequent chapters. This assessment pertains to both the in-sample and out-of-sample performance of the portfolios.

To maximise investment returns, portfolio managers often turn to innovative and multi-faceted approaches. This study investigates the potential of combining five different investment strategies - Gradient Boosting Tree, ESG Momentum, Best-in-Class ESG, Dynamic Commodity Pairs and Commodity Terms of Trade FX - to improve portfolio performance. This research aims to investigate whether the combination of these individual strategies can outperform the individual strategies and determine the most effective allocation between them.

Equity investments are investments in companies and offer the potential for high returns, although with relatively high exposure to systematic risk. Company-specific factors and broader economic trends influence their performance.

Foreign exchange trading involves global currencies and is influenced by international economic factors, interest rates and geopolitical events. The inclusion of foreign exchange in an equity-oriented portfolio can offer diversification benefits due to the different market dynamics that influence these asset classes (Kroencke et al., 2011).

The value of commodities such as precious metals, oil and agricultural products are primarily determined by supply and demand dynamics, geopolitical factors and inflation trends (Ng & Pirrong, 1994) Investments in commodities can serve as a hedge against inflation and are often inversely related to equities (Gorton and Rouwenhorst, 2004).

The focus of this study is to examine potential methods of combining traditional equity investments with alternative assets, specifically, foreign exchange (FX) and commodities. This approach aims to broaden the scope of portfolio diversification and risk management beyond traditional equity and fixed-income portfolios. The weighting methodologies used in this study include risk-parity, asset-weighted, maximising Sharpe ratio, and equal-weighted strategies.

Overall, the aim of analysing various allocation strategies is to understand how different combinations and weightings can impact the risk-return characteristics of a multi-asset portfolio.

2 Literature Review and Economic Reasoning of the Individual Strategies

This section provides an in-depth analysis of five distinct investment strategies: commodities, foreign exchange, and three unique approaches to ESG strategies in the Nordic market. These ESG approaches include Momentum, Best-in-Class, and Gradient Boosting Tree (GBT). The discussion will cover the economic rationale behind each strategy, detailing the construction of the respective models and evaluating their performance. Initially, the economic motivations for the three ESG strategies will be addressed from a unified perspective, followed by an individual examination of each strategy.

2.1 ESG Overview

In the field of sustainable investing, the integration and examination of Environmental, Social, and Governance factors play a central role. First recognised in the 2006 United Nations Principles for Responsible Investing, ESG elements are integral to discussions surrounding their impact on achieving sustainable development goals and influencing financial performance.

Extensive financial research has been undertaken to unravel the complexities inherent in ESG metrics, yielding results that indicate an encouraging upward trajectory in their effectiveness and impact. Research such as that of Friede et al. (2015), who conducted a meta-analysis of over 2000 studies, revealed a positive correlation between high ESG ratings and superior financial performance, suggesting that companies with strong ESG practices often outperform those with weaker commitments. Further studies such as NYU Sterns (2020) also establish the connection between ESG and financial performance, ESG performance is positively correlated with various metrics such as operational efficiency, stock performance, and lower cost of capital. Significantly, the analysis underscored that the financial benefits of ESG become more apparent over longer time horizons. In this study, long-term focused research was 76% more

likely to observe positive or neutral results, reinforcing the notion that the impact of ESG on financial performance may be more profound and evident over extended periods.

Whether named corporate social responsibility or ESG, the wider concept has demonstrated its staying power and significance in company analysis, indicating its critical role in shaping future business evaluations. Three distinct ESG strategies were developed to harness this evolving trend: a momentum-based approach, a best-in-class method, and an innovative machine learning strategy utilising Gradient Boosting Trees. Each of these strategies will be thoroughly examined in the following sections, providing detailed insights and justifications, reflecting their contributions to the field of ESG investing.

2.2 Gradient Boosting Tree ESG Strategy

As the Environmental, Social, and Governance (ESG) domain expands, the variability in ESG scoring methods has created a murky landscape for investors. A poignant example of this can be seen in the case highlighted by ESG Clarity (2020), where divergent ratings led to the exclusion of Tesla from the S&P 500 ESG Index while retaining companies such as ExxonMobil, notoriously known for their bad ESG practices. This decision, driven by different methodologies and ratings scales employed by ESG agencies, underscores the challenge of achieving a consistent approach to ESG assessment (Hallez, 2022).

Such ambiguity between ESG methodologies and their applications, although it poses a challenge, also underscores ESG's importance in investment decisions. One of the main challenges in ESG investing is this very ambiguity in ESG scoring, where different agencies provide vastly different scores for the same companies, often based on opaque models (Esma, 2021). The lack of transparency in these scores creates a significant obstacle for investors, complicating the process of determining the most accurate and reliable ESG assessments. With

over 600 agencies employing ESG scores, the landscape is further muddled, presenting an opportunity to create a model that addresses these inconsistencies.

Integrating machine learning into modelling ESG metrics on stock performance offers a powerful solution to these challenges. Machine learning's ability to analyse large and complex datasets makes it an ideal tool for revealing the intricate relationships between ESG practices and financial performance. This is particularly relevant in the Nordic markets, where companies are often at the forefront of ESG innovation. The advanced ESG practices in this region provide a wealth of diverse data, ideal for machine learning algorithms to analyse and extract meaningful insights.

Furthermore, applying machine learning in this context aligns with economic reasoning. The Nordic market, recognized for its high ESG standards, presents a unique case where the integration of ESG is closely linked to a company's financial health. Machine learning models can detect patterns that demonstrate how ESG metrics influence operational efficiency, risk mitigation, and long-term profitability, especially in these markets.

Modelling ESG metrics on financial performance using machine learning is not just a strategy but a response to the need for more transparency and accuracy in ESG investing. This approach is particularly beneficial in the Nordic context, where ESG leadership is synonymous with innovation and sustainable financial performance. It allows investors to navigate through the inconsistency of ESG ratings and directly assess the impact of ESG practices on a company's financial health and return potential.

2.3 ESG Momentum Strategy

The rationale for ESG momentum is rooted in the idea that companies with improving ESG scores are increasingly seen as a predictor of positive future stock performance. Similar to stock momentum, the ESG momentum strategy is predicated on the belief that firms with historical

improvements in their ESG profiles are likely to experience positive future stock returns, while those failing to address ESG concerns may face declining stock performance.

There are several studies investigating the relationship between ESG scores and corporate financial performance (Albuquerque et al., 2019; Galema et al., 2008; Kempf and Osthoff, 2007; Von Wallis and Klein, 2015). However, only a few studies focus on the changes in the ESG ratings in relation to financial improvement. A critical aspect of ESG momentum research is highlighted in Nagy et al.'s studies (2013; 2016), which contrast various approaches to ESG investing, including the impact of ESG ratings changes on market performance in the US market. Their findings indicate that ESG momentum strategies tend to deliver better risk-adjusted performance and active returns compared to strategies based solely on high ESG scores. Ouaknine et al. (2010) found similar results for the European stock market, where ESG momentum outperformed the positive screening ESG strategy by over 14%.

Extensive research, notably by Khan, Serafeim, and Yoon (2016), found a significant positive relationship between improvements in ESG scores and future stock performance. This relationship indicates that companies with strong and improving ESG practices often exhibit better financial health, making them attractive investment targets. While Kaiser (2020) reveals that improvements in ESG ratings can predict future stock performance, this relationship shows mixed results. Particularly when focusing on strategies based solely on environmental factors improvements generally correlated with negative stock returns. On the other hand, Nagy et al. (2016) demonstrate that trading on positive ESG momentum yielded significant alpha during 2007 and 2015 with an annual excess return of 2.2%. Interestingly, the study found that the majority of these excess returns can be attributed to idiosyncratic risk, which might be related to ESG signals. This implies that the market may not fully price in the nuances of ESG improvements, creating opportunities for informed investors. (Nagy et al., 2016). This lag of markets not yet fully integrating ESG information into stock prices creates opportunities for

investors to benefit from the market's slow response to ESG-related improvements (Gloßner, 2017). With its particular emphasis on sustainability, the Nordic market presents a fertile ground for exploiting these market inefficiencies. This trend is supported by evidence that early adopters of corporate policies focusing on environmental and social issues often achieve higher stock returns (Eccles et al., 2014). Similarly, Dimson et al. (2015) found that successful shareholder engagement in ESG issues leading to changes in business practices resulted in higher abnormal returns over one year.

While ESG momentum has been less extensively studied compared to traditional ESG strategies, the available research, including works by Berg et al. (2022) and Galema and Gerritsen (2022), indicates its potential to leverage the financial gains associated with continuous improvements in ESG scores over mid to long-term periods.

2.4 Best-in-Class ESG Strategy

Various studies have examined methodologies similar to the best-in-class approach for ESG investing in different global regions. These methodologies may be labelled differently, with terms like best-in-sector also being used. Despite their differences in aspects, such as the proportion of stocks selected for investment, these approaches share a common focus on prioritising the best ESG performers. This paper specifically references influential studies in the field of ESG investment strategies that are frequently cited to ensure a strong analytical basis.

Research has suggested that companies with strong ESG performance may outperform their peers over the long term. This is illustrated by Bennani, Guenedal, and Lepetit's findings that between 2014 and 2017, best-in-class stocks in the Eurozone achieved an annualised excess return of 6.6% compared to their worst-in-class counterparts (Bennani et al., 2018).

Investing in companies with strong ESG performance can help mitigate risks such as reputational, regulatory, and operational risks, which can contribute to more stable and resilient investment portfolios. This is a crucial factor in the increasing focus on ESG integration in sustainable investing strategies, as Eurosif's 2018 report emphasised.

The Best-in-Class ESG Strategy is also key to driving ESG improvements across all industries. By directing investment towards companies that excel in ESG factors, it incentivises other companies to improve their ESG performance, promoting positive changes in corporate behaviour and industry standards (Eccles & Viviers, 2011)

The study by Kempf and Osthoff (2007) is a remarkable example of the effectiveness of the best-in-class approach in the area of socially responsible investing. They developed a trading strategy that involved buying stocks with high socially responsible ratings and selling stocks with low ratings. This method resulted in significant abnormal returns, reaching up to 8.7% per year, highlighting the potential financial benefits of this investment strategy.

It is essential to acknowledge that there is no consistency in the findings of studies on this subject. For instance, Statman and Glushkov (2008) conducted a study from 1992 to 2007, which arrived at a varied conclusion. Although portfolios that follow socially responsible investing initially displayed an advantage in returns, the particular approach used, such as excluding certain stocks from undesirable companies, could potentially nullify these benefits. This outcome supports using the best-in-class technique, which concentrates on socially responsible criteria without eliminating so-called "sinful" companies (Statman and Glushkov, 2008).

In the best-in-class approach, all sectors offered by Refinitiv are considered for investment but are exclusively based on companies with high ESG ratings. The strategy involves taking only long positions and not excluding any sector.

2.5 Commodity Pairs Strategy

The incorporation of commodities into an investment portfolio is rooted in their distinct behaviour relative to traditional asset classes, such as stocks and bonds. Commodities, characterised by tangible and finite resources, not only exhibit low correlation with equities but also lack significant correlation among themselves (Erb and Harvey, 2006). This inherent diversification potential provides investors with a means to mitigate overall portfolio risk, aligning with the fundamental principles of modern portfolio theory as advocated by Harry Markowitz (Markowitz, 1952).

Moreover, commodities offer unique benefits, particularly during inflationary periods, showcasing a positive historical relationship with rising prices (Gorton and Rouwenhorst, 2004). Including commodities in a diversified portfolio serves as a safeguard against the erosive effects of inflation on the real value of investments. Tang and Xiong's research (2012) further bolsters the economic rationale for incorporating commodities by emphasising their role in enhancing long-term risk-adjusted returns. Their findings indicate that commodities contribute positively to portfolio performance over extended time horizons.

The mean-reverting nature inherent in commodities, as found by Bessembinder et al. (1995), suggests a tendency to revert to historical averages or equilibrium levels over time. When commodity prices deviate significantly from their historical averages due to external shocks or price pressure, an opportunity exists to take positions anticipating a mean reversion. This leads to a situation where being both long and short in commodity futures could prove beneficial. This approach aligns with the broader concept of statistical arbitrage, where investors seek to profit from temporary price dislocations.

Furthermore, economic motivation extends beyond risk mitigation, inflation hedging and long-short strategy potential. Commodities, positioned as an asset class, provide exposure to global

economic cycles and serve as a barometer of economic health (Erten and Ocampo, 2013). The associated diversification benefits, coupled with the potential for capital appreciation during economic expansions, underscore commodities as a valuable addition to a diversified portfolio.

2.6 Commodity Terms of Trade FX Strategy

The concept of using macroeconomic variables to predict changes in asset prices is far from new. Many of the most common currency trading strategies, such as the carry-trade strategy, the dollar-carry trade strategy and the momentum strategies, attempt to exploit exchange rate movements which deviate from the uncovered interest rate parity condition, which suggests that the difference in interest rates between two countries should be equal to the expected percentage change in the exchange rate between those two currencies. Bilson (1987), Fama (1984), and Froot and Thaler (1990) have all found empirical evidence for deviations from the condition, which imply that interest rates are not the optimal predictor of future exchange rates (Filippou & Taylor, 2017). These deviations imply that other variables contribute to predicting the future exchange rate.

Multiple prior studies have attempted to model real exchange rates using a variety of methodologies, Meese and Rogoff (1983) famously found that a range of models exchange rate models based on macroeconomic fundamentals were unable to predict nominal and real exchange rates more accurately than simple random walk model, this finding was reiterated by Cheung, Chinn and Pasual (2005). Much of the empirical evidence which refutes a quantifiable link between exchange rates was conducted on developed economies, however, more recent studies which focus on commodity-exporting, developing economies claim to overturn these findings, such as Kohlscheen, Avalos and Schrimpf (2017), who found (pseudo) out-of-sample predictability of exchange rates using commodity prices in a sample of commodity-exporting countries, or Cashin, Céspedes and Sahay (2004) who found evidence of a long run relationship

between real exchange rate and real commodity prices. Indeed, if exchange rates do follow a random walk, then changes to the exchange rates would persist and vary randomly, as opposed to returning to a long-run equilibrium as proposed by purchasing power parity, which posits that exchange rates should move towards the level that the prices of a basket of goods and services in different countries (Terborgh, 1926).

A macroeconomic variable which theoretically could influence exchange rates is commodity terms of trade. In general, a country's terms of trade, the ratio of their export prices to their import prices, can be interpreted as the quantity of import goods an economy can purchase per unit of their export goods. As one might expect, 'commodity' terms of trade follow the same definition but consider only the ratio of commodity export and import prices.

Ceteris paribus, improvement in terms of trade should be a positive sign for an economy (Collier & Goderis, 2012), imports become cheaper compared with exports and demand for the country's exports increases, which in turn increases demand for the country's currency, causing it to appreciate. Given this theoretical link, it is not hard to imagine that accurate and timely terms of trade data could bring a rich vein of insight to the appropriate market participant. The most common way in which terms of trade changes are measured is through commodity terms of trade indices. These indices weigh the effect of price changes in goods and services based on the contribution of those goods and services to the country in question's economy to determine the net effect of a price change on the country's terms of trade. The problem is that official terms of trade figures are released with a lag as trade data is not reported instantaneously, the figures are entirely compiled in hindsight, using only actual prices paid for completed trade. To solve the problem of data lag, market participants have come to focus more on commodity terms of trade indices. These indices consider a basket of relevant primary commodities, while not encompassing all of a country's trade, their prices tend to vary more than those of more processed goods and services and thus tend to have a larger influence on

short-term macroeconomic dynamics (Arezki et al., 2012). Additionally, prices of such commodities are available point-in-time from markets. This strategy will build rolling weight CToT indices for the selected 36 currencies following a methodology similar to that of Gruss and Kebhaj (2019), with key differences implemented to develop indices which can predict the effect of commodity price changes on a country's terms of trade, as well as ensuring they are representative of a country's current trade flows, providing real-time data for use as a signal for trading foreign currencies both long and short against the US dollar in a dollar-denominated portfolio. To the best of the author's knowledge, this strategy has not been tested and published on any public forum.

3 Description of the individual strategies

In this chapter, different investment strategies that are suited to particular market conditions and goals are discussed. These strategies vary from utilising advanced machine learning techniques in ESG investing to employing dynamic approaches in the commodity and foreign exchange markets, giving a thorough outline of various investment methodologies.

3.1 Gradient Boosting Tree ESG Strategy

A well-known model in the field of machine learning is XGBoost, a refined version of Gradient Boosting Trees, known for its ability to efficiently process complex and incomplete datasets, a characteristic often found in ESG data. The model's effectiveness is underscored by its predictive accuracy with unseen datasets, especially structured datasets, which has established it as a favoured tool in machine learning competitions at data platforms such as Kaggle, which indicates its suitability for the task.

The effectiveness of XGBoost was empirically tested across various datasets in the foundational paper by Tianqi Chen and Carlos Guestrin (2016). Key results include the algorithm's robust performance across different computational settings, its efficient handling

of missing data, and the use of advanced techniques like exact greedy and approximate algorithms to determine the best way to split the data at each node of the decision tree. These features have been shown to improve both the speed and accuracy of the model, indicating the suitability for complex predictive tasks such as modelling stock returns based on ESG metrics.

Despite the extensive use of machine learning in various competitions, the direct application of these techniques, including XGBoost, in forecasting stock returns based on ESG metrics, remains a relatively untapped field. Limited studies, such as Bloomberg (2020) point towards the potential of such approaches, where they managed to outperform both the Russel 1000 and S&P 500 in the selected period.

This strategy, therefore, aims to explore this gap, utilizing XGBoost's proven strengths to provide new insights into the relationship between ESG metrics and stock market performance. Further, SHapley Additive exPlanations (SHAP) introduced by Lundberg and Lee (2017), will be utilized in selecting the top 20 performing ESG metrics. SHAP is an approach based on game theory used to explain the actual predictions of the model as it is of great interest to understand what factors are guiding the predictions.

To evaluate the model's precision, a grid search on optimizing the model's mean squared errors (MSE) for the training period outliers was conducted. A grid search optimizes the Hyperparameters of the model and selects the one which would achieve the best results during the training period.

To summarise, this strategy offers a comprehensive approach to understanding and leveraging ESG metrics in investment decisions. By integrating advanced machine learning techniques like XGBoost and SHAP, this strategy goes beyond traditional analysis to capture the complex relationship between ESG metrics and financial returns. It addresses the need for greater clarity

and objectivity in ESG investing, Utilizing one of the more prominent markets – the Nordic stock market.

3.2 ESG Momentum Strategy

While price momentum is a stable part of financial markets as established by Jegadeesh and Titman in 2001 and evident across various markets (Moskowitz, Ooi, and Pedersen, 2012; Hartley, 2020), ESG scores are emerging increasingly. The concept of winners minus losers has extended to socially responsible investing, as Vojtko and Padysak (2019) found that ESG scores can be effectively used in various investment strategies, including negative screening and momentum approaches.

The ESG momentum strategy, a novel approach to sustainable investing, diverges from traditional long-term investment strategies by focusing on the dynamic nature of ESG scores. Central to this strategy is the belief that a company's future stock performance is closely tied to changes in its ESG quality. It operates on the insight that improvements in ESG scores signal a company's better capability to mitigate ESG-related risks, which market participants quickly recognize and factor into the company's share price. Unlike strategies that focus on companies with the highest ESG scores, the ESG momentum strategy targets those exhibiting significant positive changes in their ESG performance.

The cornerstone of this strategy is the computation of the ESG score change, calculated as a 12-month differential, reflecting the annual nature of ESG score reporting. This methodology facilitates the identification of corporations demonstrating significant positive shifts in their ESG profiles, aligning with the strategy's core thesis: firms exhibiting substantial ESG improvements are poised to outperform their counterparts.

Papers like Nagy (2016) showed that the execution of this strategy is distinct from any ESG tilt strategy, which leans towards companies with consistently high ESG ratings. The ESG

momentum strategy instead prioritizes companies that have demonstrated a notable increase in their ESG ratings over the past year. This focus does not directly aim to raise the overall ESG profile of the portfolio but identifies potential short-term gains from positive ESG developments.

Moreover, Nofsinger and Varma (2013) posited that higher ESG ratings enhance a company's resilience against market downturns, legal challenges, and regulatory fines. Their findings are bolstered by subsequent studies (e.g., Nagy et al., 2013; Verheyden et al., 2016; Pollard et al., 2018), which collectively underscore the profitability potential of the ESG momentum strategy.

3.3 Best-in-Class ESG Strategy

The Best-in-class approach, deeply rooted in ESG investment principles, focuses on selecting the top 10% of companies based on their ESG performance within each OMX Nordic Large Cap Index sector. This methodological choice is underpinned by reducing sector bias, a concern highlighted by Eurosif (2018), which noted the growing importance of ESG integration and best-in-class strategies in sustainable investing. After determining which stocks are eligible for investment, they are then weighted in the portfolio based on their respective ESG scores to ensure that those with a higher ESG performance have a proportionally greater impact on the investment strategy.

In implementing this approach, the aim is to achieve a balanced and diversified portfolio that accurately reflects the leading sustainability practices in the Nordic market. Through this methodical selection and weighting of the best ESG performers from different sectors, the strategy ensures a comprehensive representation of sustainable excellence in the region.

A range of ten sectors, including Industrials, Financials, Consumer Cyclicals, Healthcare, Technology, Real Estate, Basic Materials, Consumer Non-Cyclicals, Energy and Utilities, are considered in this approach.

It is important to acknowledge that certain industries, such as Consumer Non-Cyclicals, Real Estate, Energy, and Utilities, are combined into a single category called "Other." This decision was made because insufficient data is available for these sectors, especially in the first year of data observation. The purpose of this grouping is to ensure that the assessment and selection process is fair and unbiased across all industries, regardless of any limitations in data or differences in ESG performance that may be inherent to specific sectors.

This strategy follows a trading cycle that spans one year, during which ESG scores are used as the primary criterion for decision-making in January. This annual cycle aims to ensure consistency in the assessment of ESG performance and the composition of the portfolio. It is important to note that companies that are newly included in a given year are only considered for inclusion in the portfolio in the following year's assessment.

3.4 Dynamic Commodity Pairs Strategy

This commodity pairs strategy introduces a dynamic cointegration framework, departing from conventional static models with fixed formation and trading periods. The innovation lies in enhancing the strategy's adaptability to changing market conditions. By incorporating a signal with a rolling cointegration window, the approach provides a more nuanced understanding of specific pairs relationships, enabling investors to capitalise on short-term deviations from their historical means. Moreover, this dynamic framework expands the spectrum of tradable pairs, fostering increased diversification effects and, ideally, creating more profitable trading opportunities.

The approach involves purchasing an underpriced commodity future and concurrently selling an overpriced counterpart relative to each other when the z-score of the pairs spread surpasses a predefined threshold. The determination of the number of contracts for each future is guided by a hedge ratio, ensuring the formation of a zero-cost portfolio. Pairs trading offers a

compelling avenue to capitalise on profit opportunities arising from market overreactions or underreactions to new information.

The pairs strategy presented in this paper leverages 17 distinct continuous front-end commodity future contracts, all traded on US exchanges, facilitating a pool of 136 tradable pairs. In conjunction with surpassing the z-score threshold, a signal to initiate a position is contingent upon the pairs demonstrating a cointegration relationship over the preceding 200 trading days, a criterion that must persist for the subsequent 5 days, thus limiting pairs cointegrated by mere chance ability to enter the tradable pool.

Cointegration fundamentally encapsulates the idea of mean reversion in asset prices. When two assets exhibit cointegration, it indicates the presence of an equilibrium relationship between them. In pairs trading, regardless of the chosen selection method, the objective is to leverage deviations from this equilibrium within a nonstationary common factor asset-pricing framework (Nicolas Huck and Komivi Afawubo, 2014).

The cointegration approach is specifically adopted due to its demonstrated econometric reliability in identifying comoving securities. In the literature, articulated by Krauss (2016), there exist several different methodologies within pairs trading literature. While each approach has its merits, the cointegration approach stands out for its ability to establish a more robust equilibrium relationship. Empirical studies, exemplified by the work of Gatev et al. (2006) and subsequently reinforced by Do and Faff (2010), emphasise the effectiveness of cointegration tests in identifying pairs with persistent relationships.

The research conducted contributes to the literature by offering a fresh outlook on commodity pairs trading through the application of a dynamic cointegration approach. Going beyond conventional methodologies, this novel approach creates opportunities for pairs trading on commodities even in the absence of apparent fundamental relationships. The introduction of

dynamic cointegration aims to capture evolving interdependencies among commodity pairs, enabling adaptability to changing market conditions. Moreover, the inclusion of commodities lacking evident fundamental connections adds an innovative dimension to the study, challenging conventional notions and expanding the potential scope of trading opportunities.

3.5 Commodity Terms of Trade FX Strategy

The commodity terms of trade strategies were tested on a universe of 36 developed and emerging market currencies, including the US dollar as the base currency, against which the remaining 35 currencies were traded. The decision was made to trade foreign currencies against the US dollar to ensure the most liquid foreign exchange market for each currency pair, aside from a few select European countries whose most liquid market is with the Euro but are traded in sufficient volumes with the US dollar that liquidity was not a concern.

The strategy is unique in its approach to currency trading as it uses the commodity terms of trade of each currency union to determine the weight and direction of trades based on the change relative to the change in the commodity terms of trade of the US. The logic behind this signal is that if a foreign country or currency union's terms of trade have improved more than that of the US, or indeed if the US commodity terms of trade have fallen, then there should be an appreciation in the foreign currency relative to the US dollar as more of the foreign currency is demanded, and relatively less of the US dollar is demanded. In this scenario, the strategy took a long position in the foreign currency against the US dollar in anticipation of a subsequent appreciation. In the case of a decrease in the foreign currency unions commodity terms of trade relative to the US dollar, a short position was taken in the foreign currency as by the same logic the US dollar would be expected to appreciate relative to the foreign currency.

The size of the position as well as the weighting given to the trade in the combined portfolio of currency pairs was determined by the magnitude of the deviation relative to the historical

mean, with deviations of higher magnitude being allocated greater weight in the overall portfolio. An advantage of using the historical mean is that as the strategy operates for longer, the mean will become more stable and representative of the ‘true mean’, in theory allowing the strategy to improve in performance as time progresses.

Deviations were measured over the previous four weeks to ensure rapid or short-lived commodity price fluctuations do not lead to positions taken without sufficient economic reason, while weekly rebalancing still allows the strategy to be dynamic enough to capture significant terms of trade dislocations and leverage them to make profitable trades. Statistical techniques were also employed to ensure that excessively large positions in single markets were not taken, which could compromise the diversification effect of holding a larger basket of currency pairs.

As the commodity dependence of different currency unions varies significantly, the ability of commodity terms of trade to predict changes in each union’s exchange rate with the US is also likely to vary. To address this issue, the strategy was first tested on an in-sample testing period before excluding the currency pairs, which did not contribute to the risk-adjusted return of the strategy.

4 Data and Research Methodology

In this chapter, we delineate the methodologies adopted for data sourcing and preparation across the spectrum of strategies under review. Ensuring analytical coherence, particularly in the combined strategy analysis, necessitates a standardisation process wherein the returns from all strategies have been methodically converted to a daily return format. All equity strategies use the OMX Nordic Large Cap, further referred to as OMXNLC, as a benchmark.

4.1 Data and Research Methodology of the individual strategies

4.1.1 Gradient Boosting Tree ESG Strategy

This section describes the dataset selected from the Refinitiv Platform for the years 2007 to 2022, focusing on the OMXNLC stocks return and respective ESG metrics. The periods are divided into training (2007-2013), in-sample(2014-2017), and out-sample(2018-2022) periods. The predictions from 2013 onwards use rolling windows, meaning that the model trains extensively on all available previous data. This is incorporated to make sure that all available data is utilized and helps in back-testing the model. The Nordic market was chosen for its leadership in ESG reporting and the dataset initially contained 729 ESG metrics, reduced to 353 after excluding those with a 100% missing rate or irrelevancy.

The methodology incorporates a Gradient Boosting Tree approach using XGBoost, tailored for handling ESG data complexities like missing values and non-linearity, as outlined in the strategy profile. In the selection of stocks, further exclusion had to be made as OMXNLC in some cases includes A & B stocks, where they are linked to the same ESG metrics. The final dataset consists of 198 stocks for 2022, while the first dataset for 2007 consisted of 105 different stocks. Stocks are evaluated one year after they are introduced to the market, as there is always a lag in the reporting of ESG metrics. The signal used for trading is created by utilizing XGBoost to model ESG metrics to predict a stock's active return, which measures its performance against a benchmark. This strategy is grounded in the belief that stocks excelling in ESG criteria are likely to surpass their market peers. Therefore, the justification for focusing on Active Return as the dependent variable, rather than mere stock return, is to target stocks that achieve outperformance in their respective sectors, attributing success to strong ESG practices rather than general market movements.

To evaluate the performance of the model, a grid search was conducted in the training period, optimizing the mean squared error for top & bottom 15% performers. As it is of interest to understand which features affect the decision-making in the trees, Shapley Additive exPlanations (SHAP) will be employed, helping to interpret the contributions of each feature to the model's predictions. SHAP, based on Shapley values from cooperative game theory, provides interpretable explanations for the predictions of any machine learning model by quantifying the contribution of each feature to individual predictions. a new model will be trained based on the top 20 features selected based on their SHAP values, where features not showing ambiguity are prioritized. This means that a feature that had a positive effect on top values in one year but a negative effect in the next year is excluded.

4.1.2 ESG Momentum Strategy

The data for the ESG momentum strategy was retrieved from Refinitiv Eikon, ranging from 01/01/2007 to 31/12/2022. The data sample consisted of a historical composite of OMXNLC-listed stocks with daily returns and market capitalization, as well as the corresponding stock-specific ESG scores. For each individual year, the actual composition of the OMXNLC was utilized and then merged to ensure the correct compositions and to avoid survivorship bias.

The methodology for constructing ESG momentum portfolios implemented findings from previous studies highlighting the impact of stock selection thresholds on portfolio performance. The approach was rooted in the findings of Jegadeesh and Titman in 1993, who demonstrated that using stocks only in the top and bottom deciles for a long-short portfolio yielded superior results compared to a lower threshold encompassing all stocks in the sample universe. Similarly, Bird et al. (2017) found that including more stocks in momentum portfolios with a lower threshold negated statistically significant returns.

In this study, three distinct ESG momentum portfolios were constructed using a data set based on the evolution of ESG scores: a long-short portfolio, a long-only portfolio, and a short-only portfolio. Each year, companies were categorized as winners or losers based on their ESG score development. Winners were those with the most positive ESG score development over the year, with losers being those with the worst ESG score development. The portfolios were annually reconstructed based on the change in ESG scores from the previous year. For instance, the 2008 portfolios were based on ESG score changes during 2007–2008, and the selected stocks were held for a year before the portfolio was reconstituted. The positive ESG momentum strategy based solely on equities demonstrating ESG score enhancements, generated the strongest returns during the in-sample analysis, thereby serving as the referencing model for the out-of-sample application.

After careful evaluation of various thresholds for equity inclusion in the ESG momentum portfolios, a strategic decision was made to raise the inclusion threshold by dividing the stock universe into quintiles according to ESG score development. The positive ESG momentum portfolio takes a long position in the top 20% (winners), effectively narrowing down the number of stocks in the portfolios. An essential consideration in this methodology is the company-specific risk, which increases significantly when only the top and/or bottom deciles of stocks are included due to the drastic reduction in the number of companies in the portfolios. By focusing only on the top 20%, the strategy seeks returns at the extreme ends of the stock universe and excludes companies with minor changes in their ESG scores. The portfolio was built by building an equal-weight scheme where the chosen stocks are scaled according to their market capitalization.

4.1.3 Best-in-Class ESG Strategy

ESG investing using the Best-in-Class approach aims to create a portfolio that prioritizes responsible investing through a value-weighted selection of stocks. This method focuses on choosing stocks with ESG scores that rank in the top 10% within their sectors. The strategy is specifically designed for the Nordic region and uses the OMXNLC as its primary source of data. This index is an appropriate choice due to the strategy's geographic emphasis.

Data was collected from 2007 to 2022 using Refinitiv Eikon and its Eikon API, with Python used to ensure efficient and accurate data retrieval. This collection included daily close prices, sectors, and ESG scores for 231 unique instruments. The time frame for analysis was split into two parts: an in-sample period from 2007 to 2017, and an out-of-sample period from 2018 to 2022. This division allowed for a thorough analysis and validation of the investing approach.

Every year, the portfolio is created and adjusted to reflect changes in ESG scores, which are updated in December. However, the ESG score used for portfolio creation is applied in January each year to ensure consistency in the annual investment cycle. It's crucial to understand that newly listed companies are only evaluated for inclusion in the following year's portfolio, allowing for a full year's worth of data to be analysed.

To ensure a fair assessment process, sectors with lower data availability, such as Consumer Non-Cyclicals, Real Estate, Energy, and Utilities, were grouped as "Other." This grouping was necessary as a result of varying data availability levels, especially in the earlier years. Despite the differences in ESG performance or inherent data limitations, the decision to combine these sectors is made to maintain balance and fairness.

This methodology represents a thoughtful and data-driven approach to ESG investing in the Nordic region. By focusing on high-performing ESG stocks within their sectors and adjusting

for annual changes in ESG scores, the Best-in-Class portfolio aims to be an investment model for responsible investing.

4.1.4 Dynamic Commodity Pairs Strategy

For the empirical analysis of the strategy, the data consists of liquid commodities with concurrent data accessibility from Refinitiv. The dataset that was compiled spans from the first trading day of 2007 through the conclusion of 2022. This dataset is partitioned into an in-sample period, covering 2007 until the end of 2017, and an out-of-sample period encompassing the remainder of the dataset from 2018 to the end of 2022. The data was imported into Python using the Eikon API. The selected commodities were categorised into three main groups: Energy (Crude WTI, Natural Gas, Heating Oil, Gasoline), Metals (Gold, Platinum, Silver), and Agriculture (Wheat, Soybeans, Coffee, Sugar, Corn, Soybean Oil, Lean Hogs, Cotton, Feeder Cattle, Oats). Resulting in the creation of 136 unique pairs possible to trade in the strategy. All the contracts corresponding to the individual commodities were continuous front-end contracts denominated in US dollars.

Cointegration calculations are performed on a 200-day rolling basis for each commodity pair. The choice of a 200-day window was made deliberately, striking a balance between allowing sufficient time for the time series to exhibit dependence and maintaining adaptability to changing market conditions. This window size demonstrated robustness during the in-sample testing phase. Unlike traditional approaches involving a forming period followed by a trading period, cointegration was employed in this paper to function as a trading signal. The Statsmodels Python library's `coint()` function, which is based on the augmented Engle-Granger two-step cointegration test, is employed as a measure of cointegration in the analysis. This test is designed to evaluate the null hypothesis of no cointegration between variables, assuming that the variables in y_0 and y_1 are integrated of order 1, $I(1)$.

A pair was considered tradeable if it had been cointegrated for 200 days for at least five consecutive days. This requirement aimed to mitigate sensitivity to signals arising by chance, ensuring that mostly meaningful signals were considered for trading. A synthetic portfolio was formed of each pair, involving the opening of both a long and a short position. In addition to cointegration, the pair's z-score had to deviate from the rolling mean by at least 1.5 standard deviations for a position to be initiated. This strategy effectively involves buying the underpriced future and selling the overpriced one relative to each other. Positions were maintained until the spread surpassed the equilibrium by ± 0.75 standard deviations. This approach was designed to enable the strategy to capture momentum in the mean-reverting spreads.

To manage unexpected behaviour, positions were closed if a pair did not exhibit the anticipated mean reversion and remained between 1.5 and 2.5 standard deviations from the mean for more than 60 trading days. To safeguard against significant downside risk, the strategy incorporated a parameter whereby positions were automatically closed if the z-score exceeded a threshold of 3.5 or -3.5. This precaution was introduced to account for the volatile nature of the underlying contracts and to prevent the strategy from being adversely affected by temporary large deviations due to external shocks, which did not adhere to the assumptions of the strategy. Lastly, during the in-sample period, the analysis involved a pre-selection phase followed by a post-selection phase. In the post-selection, the strategy discarded the bottom half of pairs from the trading pool based on risk-adjusted returns obtained in the pre-selection phase. This selection criterion is grounded in the rationale that pairs demonstrating ineffectiveness within the strategy framework during the pre-selection period are deemed unsuitable for further trading.

4.1.5 Commodity Terms of Trade Strategy

Given the systematic nature of the strategy, all computations and executions were carried out through a model with predefined parameters in Python. After constructing the CToT indices for each currency, the initial step involved calculating the z-score for each currency's CToT index on every date. To prevent forward-looking bias, only past CToT scores were considered when calculating the mean. Z-scores were then winsorised at two standard deviations to address outliers and mitigate excessive risk-taking in any particular market.

These z-scores played a pivotal role in determining both the weight of each currency in the portfolio and the trade direction (long or short against USD). Currency weights were calculated based on the magnitude of the terms of trade change relative to the USA's terms of trade over the last four weeks. The larger the change, the greater the weight. The trade direction was determined by the sign of the terms of trade change, with a long position taken if the change relative to the USA was positive and vice versa. Normalisation of weights ensured that the sum of their absolute values always equalled one, with weekly rebalancing.

Returns from the strategy encompass two components: the fluctuation in the exchange rate of the currency (exchange rate effect) and the foreign risk-free rate earned in long positions or paid in short positions (interest rate effect). Strategy transaction costs were estimated following Bollerlev et al. (2016), using half the median of the bid/ask spread of each currency over the last nine months. The weighted returns of each foreign currency/USD pair were then aggregated to derive the portfolio return.

The strategy underwent an initial in-sample testing phase from the beginning of 2007 until the end of 2017. Subsequently, all currencies that performed with a negative Sharpe Ratio under the strategy weighting scheme were excluded. The strategy was then assessed during an out-of-sample period, both individually and as part of the combined group portfolio.

4.2 Data and Methodology for the Combined Strategies

The analysis examined a combined investment strategy consisting of five different variations: an equal-weighted approach, an equal-weighted approach based on the asset class, a target volatility approach, a risk-parity approach, and a maximum Sharpe Ratio approach. The investigation covered two distinct periods, namely an in-sample period spanning from 2014 to 2017 and an out-of-sample period spanning from 2018 to 2022. The strategy weights were applied to the returns daily, with the weighting scheme updating annually.

4.2.1 Equal-Weighted Strategy

The equally weighted strategy holds an equally distributed amount of all the individual strategies over the periods. Every strategy is included in the portfolio with a weight of 20% until the end of the sample. This allocation remains unchanged over time, regardless of the fluctuating performance of the individual asset classes.

4.2.2 Asset-weighted Strategy

The asset-weighted portfolio strategy is characterized by its simplicity and diversification approach. This method differs from more traditional weightings, such as equally weighted portfolios, by emphasizing capital allocation based on different asset classes.

In this context, the asset-weighted strategy is applied specifically to three broad asset classes. The first three strategies - Machine Learning Approach, ESG Momentum and Best in Class ESG Strategy - are grouped as one asset class, reflecting their common focus on equity-based investments. The portfolio is then expanded into more asset classes through the inclusion of commodities and foreign exchange, making it a multi-asset portfolio with each class weighted at one-third. Within this framework, the equity investments are equally weighted so that each equity strategy has a weighting of one-sixth.

4.2.3 Risk Parity Strategy

Incorporating insights from the S&P Risk Parity Indices, the risk parity approach for a multi-asset portfolio of equities, commodities, and FX aimed for a balanced risk distribution across these diverse asset classes. A Risk-Parity Portfolio is based on the principle that each asset within the portfolio contributes equally to the overall risk. Unlike traditional portfolios where asset allocation is based on capital, in risk-parity portfolios, allocation is based on the risk contribution of each asset. This means that riskier assets will have a smaller capital allocation, while less risky assets will have a larger capital allocation. The objective is to achieve a balanced risk distribution across various asset classes, leading to more stable and potentially improved risk-adjusted returns over time. (Brznek et al., 2020)

In this strategy, the individual strategies were clustered into the asset classes; equities, commodities, and FX to ensure each contributes equally to the portfolio's overall risk profile, echoing the principles observed in the S&P Risk Parity Indices. The approach requires annual rebalancing, to distribute risk equally and adjust allocations dynamically to maintain a predefined volatility level, thereby aiming for consistent performance across various market conditions. (Brznek et al., 2020)

4.2.4 Maximized Sharpe Ratio Strategy

The Sharpe Ratio is a measure of the excess return (return above the risk-free rate) per unit of risk in an investment. A maximum Sharpe Ratio strategy seeks to maximize this ratio for the given period. It's a specific application of the mean-variance optimization that not only considers the balance between risk and return but also incorporates the risk-free rate in its calculation. The goal is to find the portfolio that offers the highest excess return for each unit of risk taken. The weighting scheme is optimized in the in-sample period, selecting the returns

yielding the best results, where the same weighting is later applied to the out-sample period to test any differences.

5 Performance and characteristics of the portfolios

This section will provide a comprehensive overview of the in-sample performance for each individual strategy, spanning from the inception date of each strategy to the conclusion of 2017. Additionally, it will examine the performance of various combined portfolio strategies employing multiple weighting schemes. The evaluation will cover the period from February 2014, marking the earliest date when all strategies could generate returns, through the end of 2017.

5.1 In-Sample

The table below contains selected summary statistics of each individual strategy, to which the performance descriptions in the following section will refer.

Table 1: Individual Strategies In-Sample Performance Statistics

	Return	Volatility	Skewness	Kurtosis	Sharpe Ratio	Max Drawdown	Transaction Costs
GBT-Model	17.66%	18.60%	-0.11	5.45	0.94	-19.18%	0.00385%
ESG Momentum	5.79%	20.45%	0.09	10.10	0.23	-57.64%	NA
Best in Class ESG	8.99%	16.12%	-0.31	4.87	0.54	-50.92%	NA
Commodity Pairs	26.81%	22.53%	0.90	10.61	1.18	-19.39%	0.0033%
Currency Pairs	1.67%	4.19%	0.44	9.73	0.083	-12.00%	0.0041%

Note: All statistics are reported as annualised figures except transaction costs which are the average % of notional trade value

5.1.1 Gradient Boosting Tree Strategy

The In-sample period for the GBT model, incorporating SHAP values analysis spanned from the beginning of 2014 to the end of 2017 with annual rebalancing of the portfolio. The model employed rolling windows, meaning each training period incorporated all available previous data, leveraging cumulative historical information. During the in-sample period, the portfolio demonstrated solid performance, achieving an annualised return of 17.66% and an annualised

volatility of 18.60%, resulting in a Sharpe Ratio of 0.94. The relatively low Max Drawdown indicates that the portfolio experiences minimal fluctuations. Additionally, the portfolio's Kurtosis of 5.45, which exceeds the standard kurtosis of 3 for a normal distribution, suggests a distribution with more pronounced tails, implying a higher likelihood of extreme returns. The negative skewness of -0.11, while slight, indicates a marginal asymmetry in the distribution, with a subtle inclination towards negative values.

5.1.2 ESG Momentum Strategy

The in-sample period for the ESG momentum strategy ranges from 01/01/2008 until 31/12/2017, with the entirety of 2007 utilized to generate the ESG momentum signal. The sample period is thus one year shorter, due to the 12-month lag needed to construct the signal. Referring to the results presented in Table 1, the positive ESG momentum portfolio generated an annualised return of 5.79%, paired with a volatility of 20.45%, reflecting a strategy with balanced growth prospects tempered by substantial risk exposure. The strategy's Sharpe Ratio stood relatively low at 0.23, indicating positive risk-adjusted returns. Notably, the maximum drawdown was -57.64%, reflecting substantial but not extreme declines, although this decline is a reaction to the global financial crisis where equity markets around the world went down. The skewness of 0.09 indicates that returns are relatively symmetrically distributed, with a slight inclination towards positive outcomes. The higher kurtosis at 10 implies a likelihood of occasional outsized returns, which could mean infrequent but significant gains or losses. The strategy portrays moderate returns with the potential for both stable gains and significant fluctuations.

5.1.3 Best-in-Class ESG Strategy

From 2007 to 2017, during the in-sample period, the Best-in-Class strategy, to invest in the top 10% of ESG Score in each sector, achieved an annualised return of 8.34% and a remarkable

annualised volatility of 17.85%. During this period, a maximum drawdown of -50.8% was recorded and the strategy generated a Sharpe Ratio of 0.47, indicating that the return is obtained with significant risk. It is important to note that these performance figures are calculated excluding transaction costs.

5.1.4 Dynamic Commodity Pairs Strategy

The in-sample period for the commodity pairs strategy spans from the beginning of 2007 through 2017. The strategy incorporates a 1-year rolling window, initiating live signals in early 2008. From 2014 until the end of the sample period, the strategy selects the top-performing 50% of pairs based on their risk-adjusted return contribution to the strategy. These pairs will then be the only pairs constituting the trading pool from 2014 until 2018. After this selection, the strategy delivered a respectable annual return of 17.8% with an annual volatility of 18.04%. The resulting Sharpe Ratio for the in-sample period stands at 0.99. Despite a drawdown of -13.95%, the strategy demonstrates resilience within its market context. Notably, the strategy exhibits a substantial kurtosis of 10.95, suggesting quite a high degree of tail risk in the return distribution. Additionally, the skew of 0.84 indicates that the strategy has a longer right tail and the potential for positive returns is higher than the potential for negative returns.

5.1.5 Commodity Terms of Trade FX Strategy

The in-sample testing period for this strategy spanned the beginning of 2007 until the end of 2017, the strategy requires 4 weeks (29 days) to generate trading signals and as a result, has zero returns until the first trading day at least 29 days from the beginning of the year. As can be seen from **Table 1** above the CToT strategy provided an annualised return of 1.67% as well as a positive risk-adjusted return, exhibiting a Sharpe Ratio of 0.02. During this period the strategy also experienced relatively low annualised volatility and moderate smaller maximum

drawdown and positive skewness indicating a higher chance of extreme positive than negative returns.

After the in-sample testing period, the performance of each currency pair under the weighting scheme was evaluated in risk-adjusted terms, with currency pairs which contributed negatively by this metric (Sharpe Ratio) being excluded from out-of-sample trading, allowing greater weight to be allocated to the currency pairs which contributed positively. 4 currency pairs remained in the portfolio, the Brazilian Real, Israeli New Shekel, Nigerian Naira and Russian Ruble (all paired with the US dollar).

5.1.6 Combined Strategies

5.1.6.1 Hypothesis

As mentioned in the Introduction, it is assumed that the asset-weighted strategy will be the best performing. The strategy allocates the capital equally among the asset classes, leveraging the benefits of diversification. By assigning equal risk, to equities, commodities, and FX, rather than to the strategies, equities are not overweighted as in the equal-weighted strategy. In this way, the strategy aims to mitigate the idiosyncratic risk of any single asset class dominating the portfolio's performance. This equal distribution of risk across varied asset classes is intended to harness the diversification effect, which can potentially lead to more stable and improved risk-adjusted returns over time. It does not rely on complex rebalancing algorithms or predictions about future volatility, which can be error-prone. By equally dividing risk among varied asset classes, it naturally adjusts to market changes, potentially offering a more robust performance in differing market conditions without the need for frequent adjustment. Furthermore, unlike the equal-weighted strategy, which disproportionately allocates 60% towards equities given its three individual equity strategies compared to one each for FX and commodities, the asset-weighted strategy offers a more balanced risk profile.

5.1.6.2 *In-sample results*

In shaping an effective portfolio allocation strategy, consideration of the individual strategies and their interactions within the broader market is crucial. The correlation among them underscores the potential benefits of prudent allocation, leveraging the diversification effects inherent in assets with low correlations. Notably, the presence of three individual strategies focused on ESG criteria in the Nordic market introduces a nuanced balance—offering diversification benefits while also necessitating a cautious approach to avoid overexposure to specific stocks. The inclusion of two alternative market strategies, FX (foreign exchange) and commodities, in our portfolio offers a diversification benefit. This is evident from their very low correlation of -0.0212 with each other, and the highest correlation with equities being only 0.0526. Such low correlation figures indicate that these assets move relatively independently of each other, presenting an opportunity to spread portfolio risk across different market segments. As expected, the correlation between the different Equity strategies, drawing returns from the same market, is quite high.

Table 2 - Correlation Matrix - In sample period: 2014-2017

	FX returns	ESG BIC	COM return	ESG MOM	ESG ML
FX returns	1.0000	0.0527	-0.0212	0.0478	-0.0052
ESG BIC	0.0526	1.000	0.0059	0.8925	0.7485
COM return	-0.0212	0.0059	1.0000	0.0126	-0.0171
ESG MOM	0.0478	0.8925	0.0126	1.0000	0.8369
ESG ML	-0.0052	0.7485	-0.0171	0.8369	1.0000

The subsequent section quantifies these considerations by examining both active and passive allocation approaches. It aims to offer a nuanced comprehension of the optimal weighting scheme for the diversified portfolio. Refer to Table 3 for the relevant statistics corresponding to each of the allocation strategies.

Table 3: Combined Strategy in Sample Performance Statistics

	Return	Volatility	Sharpe Ratio	Kurtosis	Max Drawdown	Skewness
Equal-Weighted	13.97%	10.48%	1.31	6.41	-11.19%	-0.23
Asset-Weighted	14.25%	9.16%	1.53	7.31	-7.57%	0.47
Risk-Parity	8.79%	6.57%	1.30	7.58	-6.14%	-0.19
Sharpe Optimizer	12.13%	7.24%	1.64	5.85	-5.68%	0.33
Benchmark	5.86%	5.03%	1.11	4.53	-6.84%	-0.19

Note: All statistics are reported as annualised figures

5.1.6.3 Equal-Weighted

In assessing the equal-weighted allocation strategy, it's clear that the various sub-strategies exhibit diversification even without optimization. While the strategy boasts a respectable Sharpe Ratio of 1.31 and annual returns of 13.97%, its performance is relatively mediocre in most metrics compared to the alternative allocations. As far as the distribution characteristics of the returns are concerned, a kurtosis of 6.41 and a skewness at -0.23 were determined. The maximum drawdown was calculated at -11.19 %.

5.1.6.4 Asset-Weighted

The asset-weighted approach distributes a larger allocation in the alternative markets, constituting 66% of the portfolio. Considering their negative correlations and different risk-return characteristics, the allocation performs significantly better than risk parity in most metrics. Limiting only 11% each to seemingly profitable equities strategies leads to a flatter and less volatile return profile when compared to the equity-weighted strategy. The Sharpe Ratio also has a higher value than the equal-weighted strategy.

5.1.6.5 Risk-Parity

As evidenced in Table 2, the risk-parity allocation strategy performs with mediocrity in terms of risk-adjusted returns when compared to the alternative allocations. It also has significantly lower annual returns. The predominant factor contributing to the resulting performance is the

overweight position in the FX strategy, exceeding 53% of the total portfolio throughout the entire sample period. The remaining allocation was distributed relatively evenly among the other strategies for most of the period. This allocation was influenced by the FX strategy's low volatility during the in-sample period. However, it is crucial to recognize that this low volatility is accompanied by relatively subdued returns. Risk parity, relying on risk-based capital allocation, also encounters challenges in responding effectively to shifts in market sensitivity.

Despite being less favourable in terms of annual return, the risk-parity strategy still demonstrates a noteworthy Sharpe Ratio of 1.30. This Sharpe Ratio surpasses that of any individual strategy by quite some margin, underscoring the relative effectiveness of the risk-parity approach and the diversification effects of combining the different strategies.

5.1.6.6 Maximized Sharpe Ratio

Utilizing the Sharpe Optimizer approach, which seeks to enhance the Sharpe Ratio following principles similar to those advocated by Harry Markowitz, the strategy exhibits commendable performance. The strategy expectedly achieved the highest Sharpe Ratio of 1.64 after the optimization, indicative of considerable risk-adjusted returns, while also exhibiting lower volatility of 7.24% compared to the Equal-Weighted and Asset-Weighted strategies. This suggests a more stable performance with fewer fluctuations. The maximum drawdown is also the lowest at -5.68%, implying less probability of large downside fluctuations in portfolio value. Moreover, the positive skewness of 0.33 hints at a bias towards more frequent positive returns as opposed to negative ones. These statistics, collectively, underscore the effectiveness that could be obtained if one leverages the allocations to their strengths.

Interestingly, the weighting scheme adopted based on the in-sample period fully excludes 2 strategies – ESG BIC and ESG MOM, resulting in investments only being made into 3 of the strategies.

Table 6: Combined Strategy Out of Sample Performance Statistics

	Return	Volatility	Sharpe Ratio	Kurtosis	Max Drawdown	Skewness
Equal-Weighted	14.27%	12.82%	1.02	7.22	-21.26%	-0.59
Asset-Weighted	15.61%	10.63%	1.36	5.07	-11.29%	-0.15
Risk-Parity	15.07%	9.83%	1.42	5.29	-10.57%	-0.17
Sharpe Optimizer	17.38%	10.08%	1.61	4.95	-13.05%	-0.07
Benchmark	3.93%	7.45%	0.37	9.61	-16.65%	-0.22

Note: All statistics are reported as annualised figures

Table 4 - Maximized Sharpe Ratio - weights

FX	ESG BIC	COM	ESG MOM	ESG ML
52%	0%	24%	0%	24%

5.2 Out-of-Sample

The out-of-sample results present a comprehensive narrative about the efficacy of the in-sample tested portfolio optimization strategies. The correlation in the out-of-sample period between the individual strategies is different in certain aspects compared to the in-sample period. The ESG MOM portfolio showcases quite a different correlation to the COM and FX portfolios, with absolute differences of 0.0619 and 0.1221, demonstrating differences in the returns. The Same effect is noticeable for ESG BIC, with an absolute difference of 0.0559 and 0.1053 respectively. The ESG ML Showcases small differences in correlation, but the effect is not nearly as prevalent. Interestingly enough, The ESG strategies have a negative correlation towards the COM and FX portfolios, indicating even further diversifications than in the In-sample period, where only the ESG ML had a negative correlation.

Table 5 - Correlation Matrix - Out-of-sample: 2018-2022

	FX returns	ESG BIC	COM return	ESG MOM	ESG ML
FX returns	1.000	-0.0526	-0.0032	-0.0744	-0.0601
ESG BIC	-0.0526	1.0000	-0.0499	0.8988	0.6920
COM return	-0.0032	-0.0499	1.0000	-0.0493	-0.0381
ESG MOM	-0.0744	0.8988	-0.0493	1.0000	0.7897
ESG ML	-0.0601	0.6920	-0.0381	0.7897	1.0000

Table 5 provides a comprehensive overview of the performance metrics of the different investment strategies over the specified out-of-sample period. The equal-weighted strategy in the out-of-sample period had an annualised return of 14.27%, slightly higher than its in-sample return of 13.97%. However, the volatility also increased from 10.48% to 12.82%, and the Sharpe ratio fell from 1.31 to 1.02. The maximum drawdown of -21.26% was significantly higher than in the sample (-11.19%) and the skewness changed more negatively to -0.59.

This asset-weighted strategy performed better out-of-sample with an annualised return of 15.61%, compared to 14.25% in-sample. Volatility increased from 9.16% to 10.63%, and the Sharpe ratio worsened from 1.53 to 1.36. Maximum drawdown was less pronounced out-of-sample (-11.29% vs. -7.57%), and skewness decreased slightly to -0.15. The improvement in both returns and Sharpe ratio in the out-of-sample period suggests that this strategy is more adaptable to changing market conditions and provides a better balance between return and risk.

The annualised return for the risk-parity strategy was 15.07% out-of-sample, significantly higher than the 8.79% for the in-sample period. Volatility was higher at 9.83% compared to 6.57%, and the Sharpe ratio increased from 1.30 to 1.42. The maximum drawdown improved from -6.14% to -10.57%, with skewness remaining negative at -0.17. An increase in the return can also be observed here.

Out-of-sample, the Sharpe Optimizer achieved the highest annualised return of 17.38%, compared to 12.13% in the sample. Volatility was slightly higher at 10.08% compared to 7.24%, and the Sharpe ratio decreased from 1.64 to 1.61. The maximum drawdown was less pronounced at -13.05% compared to -5.68%, and skewness was slightly negative at -0.07.

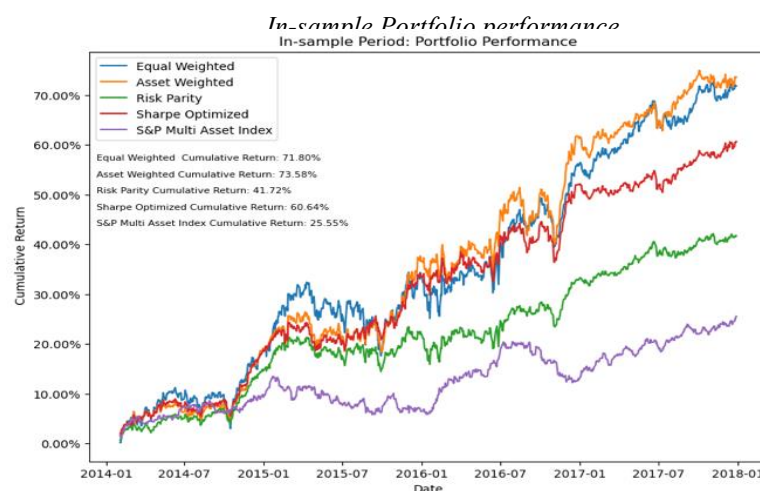
The equal-weighted strategy recorded marginal gains in returns at the expense of increased risk, while the asset-weighted strategy improved both returns and risk management. The risk

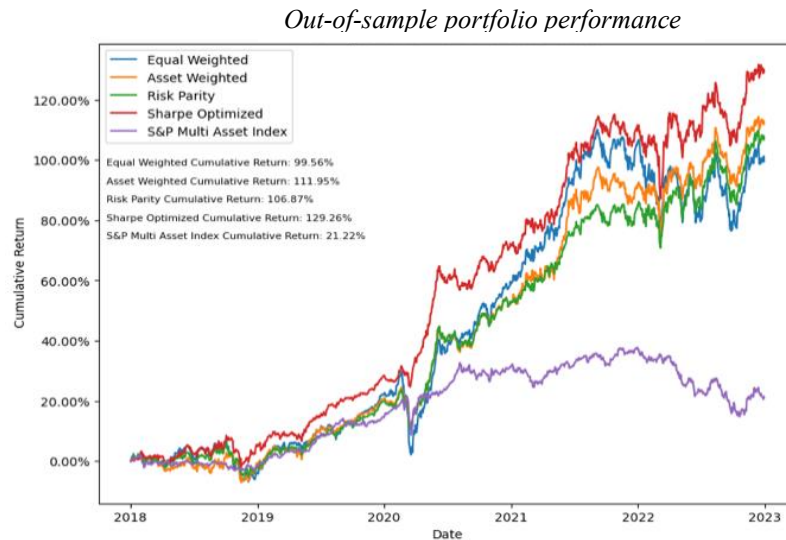
parity strategy recorded a significant increase in returns and a balanced improvement in risk ratios. The Sharpe Optimiser strategy in particular saw a significant increase in returns and efficient risk-adjusted performance.

5.3 Discussion of Performance Results

The varied outcomes from the portfolio strategies offer several notable points for discussion. The hypothesis that the asset-weighted portfolio will be the most optimal way is rooted in the equal distribution of risk across diverse asset classes like equities, commodities, and FX, and its diversification benefits. This diversification is firstly underscored by the correlation matrix showing the high correlation between the three equity strategies and the extremely low correlation between equities, FX and commodities around 0.05. While almost all portfolios increased in performance during the out-of-sample exciding the benchmark by far, the Sharpe optimizer stands out as the best performer due to its superior risk-adjusted returns and resilience as evidenced by its Sharpe Ratio and maximum drawdown metrics during in-sample and out-of-sample. It demonstrates that maximizing the Sharpe Ratio can lead to a strategy that not only performs well in-sample but also out-of-sample, a testament to its robustness. Nevertheless, in-sample and out-of-sample the second best-performing strategy was the asset-weighted portfolio, which also showed strong improvements, highlighting the effectiveness of diversification and balanced risk management. It stood out due to its superior Sharpe Ratio compared to the Equal-Weighted and Risk-Parity strategies, while also limiting volatility through a balanced allocation across various asset classes. It showed notable improvements from the in-sample period, achieving an annualized return of 15.61% and an improved Sharpe ratio of 1.36. This strategy demonstrated adaptability and an enhanced balance between return and risk, suggesting it was better suited to the changing market conditions of the out-of-sample period.

Some notable events occurred during the investment periods, where the market conditions varied greatly from the in-sample period to the out-sample period. As indicated by the graph below, the in-sample period was relatively stable with not a lot of short crashes and a relatively stable increase over time. However, the out-sample period graph tells a different case, where extreme spikes happen during notable times. After 2020 there was a large dip in performance, mostly explained by the volatile period of COVID-19. However, The Sharpe Optimized portfolio does not see a similar spike as the others. This can be explained by the Weighting scheme of the Sharpe Optimized portfolio, where 52% of the funds are allocated to the FX strategy. This indicates that the FX-strategy performs well during times of market uncertainty, and further reinforces the diversification benefits. The same phenomenon is however not present in 2022, another year of high market uncertainties, where the first increases in risk-free rate started. The Graph also highlights the volatile nature of the year 2022, where the returns fluctuate a lot, presumably because of the war in Ukraine, which had dramatic effects on the market. The chosen Benchmark does offer a much safer investment strategy than anticipated and seems to consistently offer safer but lower returns.





Comparing the change from in-sample and out-of-sample performance, the Risk-Parity strategy showed the highest improvement. It had a marked increase in its annualized return, going from 8.79% in-sample to 15.07% out-of-sample. Additionally, its Sharpe Ratio improved from 1.30 to 1.42, reflecting better risk-adjusted performance despite increased volatility and a higher maximum drawdown in the out-of-sample period.

In conclusion, the asset-weighted strategy, with its balanced risk distribution and more realistic adaptability, seems to offer the most compelling case for a robust portfolio management approach. Besides, the Sharpe Optimizer strategy also stands out for its efficient risk-adjusted returns and stability but is trickier to integrate into a retail portfolio.

5.4 Regression

This section presents a regression analysis performed on the investment strategies against a specified benchmark, Fama French 5 Factors + Momentum (FF5FM). As our portfolio is quite diversified and covers different areas of the market, the chosen Benchmark was the S&P U.S. Balanced Multi-Asset Index. This benchmark is reflective of the asset classes targeted by the strategies: equities, FX, and commodities, but also includes fixed income. The benchmark was chosen as it is a fair representation of the returns an investor could generate passively by investing in a multi-asset index.

For this analysis, monthly returns data for each strategy are adjusted by the risk-free rate to compute excess returns, which are then subjected to a regression against the excess returns of the benchmark and FF5FM. This methodology facilitates a thorough investigation of the alpha and beta coefficients of each strategy. The alpha coefficient offers insight into the performance of the strategy beyond the market's influence, while the beta coefficient quantifies the sensitivity of the strategy's returns to market fluctuations, serving as an indicator of systematic risk.

Table 7: Regression Results - Fama French Factors + Momentum

In-Sample: 2014-2017				
	Equal weight	Asset Weighted	Risk parity	Max Sharpe
Alpha (a)	14.65%** (0.000)	16.15%** (0.000)	8.85%** (0.000)	12.56%** (0.000)
Market (b)	0.26 (0.270)	0.14 (0.585)	0.16 (0.302)	0.11 (0.550)
Size	0.28** (0.000)	0.33** (0.000)	0.22** (0.000)	0.27** (0.000)
Value	-0.35** (0.000)	-0.34** (0.000)	-0.19** (0.000)	-0.17** (0.000)
Profitability	-0.0844 (0.160)	-0.09 (0.121)	0.02 (0.623)	0.09 (0.053)
Investment	0.13 (0.069)	0.62** (0.000)	0.046 (0.312)	0.41** (0.000)
Momentum	-0.0587 (0.028)	0.13** (0.000)	-0.01 (0.780)	0.16** (0.000)
R^2	0.138	0.179	0.131	0.170
Information Ratio	1.81	1.89	1.69	1.93
Tracking Error	0.081	0.085	0.052	0.065

Notes: p-values in (); **, * denote statistical significance at 1 and 5% level, respectively. Alpha, Information ratio and Tracking error are given in annual terms.

Within the in-sample period, the Asset Weighted portfolio has the highest alpha of 16.15%, with an Information ratio of 1.89 and a Tracking error of 0.085, showcasing a strong performance on key metrics, while also having the highest R-squared, 0.180, indicating the most correlation with FF5F+Momentum. The Sharpe Optimized Portfolio, not surprisingly, performs the highest in terms of Information Ratio at 1.93 while having an alpha of 12.56%. all the portfolios have a p-value below 0.05 for the Alpha, indicating that the results are statistically significant. In terms of the other factors, Size & Value have all a p-value below 0.05 across the board and are hence statistically significant. The Size factors are positive for

the portfolios, which indicates that the portfolios reflect a tilt or preference towards small-cap companies or stocks. Conversely, the negative relationship with the value factor indicates that the portfolios favour growth over value stocks. The asset-weighted portfolio and Sharpe Ratio-optimized portfolio have significantly positive values for the Momentum factor, indicating that during the sample period, those strategies capitalized on a continuation of existing market trends. The same relationship could be seen with the Investment factor, implying a preference for or better performance with firms that are actively investing (high-investment firms). In terms of the Market factor, none of the P-values are statistically significant and hence no conclusions can be drawn.

Table 8 : Regression Results - Fama French Factors + Momentum

	Out-of-Sample: 2018-2022			
	Equal weight	Asset Weighted	Risk parity	Max Sharpe
Alpha (a)	10.74%* (0.026)	12.97%** (0.008)	12.05%** (0.005)	14.19%** (0.002)
Market (b)	0.99** (0.000)	0.43** (0.036)	0.50** (0.006)	0.45** (0.019)
Size	0.43 (0.012)	0.25 (0.133)	0.2341 (0.115)	0.16 (0.305)
Value	0.06 (0.662)	-0.05 (0.750)	-0.0041 (0.974)	-0.15 (0.264)
Profitability	0.25 (0.205)	0.07 (0.711)	0.1819 (0.291)	0.04 (0.808)
Investment	-0.17 (0.405)	0.04 (0.849)	-0.0512 (0.771)	0.07 (0.708)
Momentum	-0.07 (0.563)	-0.04 (0.733)	-0.0522 (0.611)	-0.02 (0.894)
R^2	0.499	0.161	0.265	0.188
Information Ratio	1.16	1.39	1.47	1.62
Tracking Error	0.093	0.093	0.082	0.088

Notes: p-values in (); **, * denote statistical significance at 1 and 5% level, respectively. Alpha, Information Ratio and Tracking error are given in annual terms.

During the out-of-sample period, the regression analysis exhibits similarities in Factor values but demonstrates notable differences, particularly in terms of p-values. The Sharpe-Optimized portfolio had the highest alpha of 14.19%, followed by the Asset Weighted portfolio of 12.97%. Interestingly, the Equal Weighted portfolio shows a much higher correlation with an R-square of 0.499 and the same trend is also noticeable for the rest of the portfolios. The alpha is, similarly to the in-sample period, statistically significant. An interesting change is that all

portfolios now showcase a p-value below 0.05 for the market factor, showcasing a statistically significant relationship. This means that, as the Market factor is positive across the board, the portfolios tend to move in the same direction as the Benchmark. For the equal-weighted portfolio, The Factor is 0.9928, meaning that the returns of the strategy are very closely aligned with the Benchmark. Most of the other factors, Size, Value, Profitability and Momentum, do not have statistically significant P-values for the out-of-sample period, barring the Size Factor for the Equal Portfolio. Hence, no real conclusions from these factors can be drawn for the Out-of-sample period, regardless of their factor value.

In summary, all of the portfolios showcase a high positive alpha towards the Benchmark, both during the In-sample-period and Out-of-sample period. The high Information Ratio also tells us that our portfolios consistently outperform the given Benchmark, While the relatively High Tracking error tells us that the returns of the portfolio's returns are not closely aligned with the Benchmark. As described in the previous section, that result is not surprising, as the chosen Benchmark has, over both the in-sample and out-of-sample period, a much lower Annualised Volatility and Annualised Returns. As the FF5 + Momentum factors p-value changes in between the periods, no real relationship can be consistently drawn across the entire investment period, and further analysis combining or analysing different periods needs to be conducted to further validate the claims.

6 Limitations

A key challenge investors encounter stems from the costs of operating an active strategy. Transaction costs vary based on factors such as monthly trade volumes and whether the investor is categorised as retail or professional, with the latter often benefiting from lower fees. The individual strategies GBT, CToT and Commodity Pairs Trading consider different transaction costs based on their asset class and the nature of their trades aside from ESG

Momentum and ESG Best-in-Class, which may lead to an overestimation of actual returns for these two strategies. Further, the strategy is for simplicity's sake rebalancing each weighting scheme to the weight of that year. This means that winners get punished and could skew the results negatively. It is however also possible that strategies that fluctuate highly in a certain year do not get punished for it, as the fluctuation is not fully encapsulated by the weighting scheme. Overall, the effect should presumably skew returns upwards, as Transaction costs for these trades are not included. Besides explicit costs, investors must account for implicit costs such as the bid-ask spread and slippage/latency, typically more significant for professionals due to the larger size of their trades. We use the end-of-day/week/month closing price in our analysis. Still, it is important to note that this price might not always be achievable in real trading scenarios, where buying and selling occur at ask and bid prices, respectively. High market volatility, like that observed during the beginning of the COVID-19 pandemic, can widen spreads, especially for illiquid assets, and even minimal spreads can accumulate significantly over time, impacting performance.

Additionally, liquidity can pose a major challenge in the commodity futures strategy, particularly with distant contracts that are most liquid during monthly rollovers. Investors needing immediate liquidity to act on trading signals might find this strategy underestimates the impact of illiquidity. The commodity futures market is also evolving rapidly, with its financialization detailed in a separate report. As this market expands, it is unclear how the risk premium will be affected. Similarly, in the commodity terms of trade strategy, the tendency to trade more in emerging, commodity-dependent currencies, which can be less liquid and have higher bid-ask spreads, can significantly increase transaction costs as well as have the potential to limit the scalability of the strategy in the case of, for example, a large institution trading high volumes in currencies with lower market liquidity.

Furthermore, the commodity strategy mandates the maintenance of a margin account to facilitate the use of leverage. While the strategy allows for leveraging futures contracts, it demands a substantial capital commitment due to the sizable notional value associated with certain contracts. For example, strategies employing a long-short approach require a margin account for stock borrowing, making them challenging for retail investors. This difficulty arises not only from the risks and financial savvy needed for managing gross exposures exceeding 100% but also due to the same substantial capital requirements. Portfolio rebalancing, essential for achieving optimal asset allocation, involves buying and selling stocks and futures. While stock prices typically range from \$10 to \$1,000, dealing with commodity futures, where a single contract can be worth up to \$100,000, complicates rebalancing. Therefore, achieving precise asset allocation is mostly feasible for institutional-sized portfolios.

Moreover, the adage “past performance is not indicative of future returns” is particularly relevant in finance and more applicable than ever. The analysis period began post-global financial crisis and post-Eurozone debt crisis, encompassing a notably strong decade for equities and the post-pandemic surge linked with the ECB’s unconventional monetary policies, including quantitative easing and negative interest rates.

Expanding on this point, portfolios like Risk-parity and Maximzed Sharpe are constructed using historical returns, leading to a bias that favors portfolios with exceptional past performance. Consequently, this bias results in certain weightings that exclude entire portfolios, undermining the diversification advantages. For instance, in the Maximum Sharpe portfolio, no investments are allocated to ESG BIC and ESG MOM, resulting in only 3 strategies being invested in.

In addition, an overfitting bias in the in-sample period is present for the Maximized Sharpe Ratio strategy as the weights of the strategies were fitted to the historical data on which they

were tested, resulting in a model that is intricate and specific to the idiosyncrasies of this particular data set. However, this bias is not present in the out-of-sample period, as the weighting scheme is the same as the one optimized for the in-sample period.

The forward-looking bias in our ESG scores is another notable limitation. This bias emerges from the reliance on contemporary ESG data that may not have been accessible or fully developed during the historical periods under examination, which can lead to artificially deflated or inflated performance results, painting an overly optimistic picture of the investment strategies based on this data. As a result, the backtesting may inadvertently incorporate insights and assessments based on information that, at the time, was not available to investors or the public.

Survivorship bias is an important factor to consider in our composite strategy, particularly for the equity strategies that utilize the OMX Nordic Large Cap index. Such Indices undergo periodic revisions to their constituent stocks, typically adding (subtracting) stocks due to high (low) performance or due to higher (lower) market capitalization. This process can induce survivorship bias, presenting an inflated historical performance by only reflecting currently successful stocks and omitting those eliminated due to poor performance. In financial analysis, failure to acknowledge this bias can result in a distorted perception of performance, leading to potentially flawed investment decision-making predicated on partial or biased data.

The Equity strategies also pose a certain risk in terms of currency fluctuations, as the chosen Market incorporates 4 different types of Currencies – The Swedish Crown, The Norwegian Crown, The Danish Crown and The Euro. Typically, one would assume a cost of Hedging these types of risks, but due to the complication of implementing it, it has been excluded, which means that the returns, in a realistic scenario, would further be decreased.

7 Results/ Conclusions

This work was a comprehensive attempt to evaluate the performance and profitability of combining different investment strategies within a multi-asset portfolio, particularly in the Nordic market. Incorporating commodities, foreign exchange, and the three novel ESG approaches, Gradient Boosting Tree, ESG Momentum, and Best-in-Class, was noteworthy. Each of these strategies brought distinct features and performance trends to the overall portfolio, highlighting the intricacies of multi-asset investment. The analysis uncovered that although conventional asset categories like equities and commodities are fundamental to portfolio building, integrating ESG factors as non-financial components could enhance risk-adjusted returns.

For each strategy, the empirical analysis yielded different results. The asset-weighted approach, which heavily invested in alternative markets, outperformed the risk parity method across most metrics. This indicates the advantages of diversifying across asset classes. The Maximum Sharpe Ratio strategy demonstrated its effectiveness by delivering the highest risk-adjusted returns. This confirms the notion that a well-balanced multi-asset strategy can generate above-average portfolio performance. However, strategies such as risk parity provided stable but lower total returns. This highlights the trade-off between managing risk and maximizing returns.

The research confirmed the original hypothesis: diversification across commodities and currencies and the inclusion of ESG strategies can improve a portfolio's risk-adjusted returns. This finding is central to understanding the construction of multi-asset portfolios in today's markets, particularly given the increasing importance of ESG factors in investment decisions.

To advance the study of multi-asset portfolio management, a wider and more diverse range of research is necessary. Examining various global markets, not limited to the Nordic region

studied here, could yield valuable insights into the adaptability and variability of these strategies in different economic contexts. Additionally, exploring the inclusion of emerging market dynamics into these portfolios presents an interesting opportunity, given the potential for growth and unique risk-return profiles of these markets. Such research could investigate how emerging market assets interact with traditional assets like bonds, equities, and commodities, potentially revealing novel strategies for risk management and diversification. Moreover, incorporating bonds into multi-asset frameworks would provide a more comprehensive view of portfolio construction and a deeper understanding of how different asset classes can synergize to enhance portfolio resilience and performance in varying market conditions.

To summarise, this paper demonstrates the significant benefits of a multi-asset approach to portfolio management, with a focus on the integration of ESG factors. The findings provide investors with clear guidance on the benefits of diversified, risk-adjusted strategies. These findings highlight the effectiveness of multi-asset portfolios in achieving balanced performance in different market conditions and emphasize their relevance in the broader context of global investment strategies.

8 References

- Albuquerque, Rui, Yrjö Koskinen, and Chendi Zhang. 2019. "Corporate Social Responsibility and Firm Risk: Theory and Empirical Evidence." *Management Science* 65 (10): 4451–69. doi:10.1287/mnsc.2018.3043.
- Arezki, M.R., Pattillo, M.C.A., Quintyn, M.M.G., and Zhu, M. 2012. "Commodity Price Volatility and Inclusive Growth in Low-Income Countries." International Monetary Fund.
- Bennani, Leila, Théo Le Guenedal, Frédéric Lepetit, Lai Ly, Vincent Mortier, Thierry Roncalli, and Takaya Sekine. 2018. "How ESG Investing Has Impacted the Asset Pricing in the Equity Market." *Social Science Research Network*, January. doi:10.2139/ssrn.3316862.
- Berg, Florian, Julian F Kölbel, and Roberto Rigobón. 2022. "Aggregate Confusion: The Divergence of ESG Ratings." *Review of Finance* 26 (6): 1315–44. doi:10.1093/rof/rfac033.
- Bilson, John F. O., and David A. Hsieh. 1987. "The Profitability of Currency Speculation." *International Journal of Forecasting* 3 (1): 115–30. doi:10.1016/0169-2070(87)90082-3.
- Bollerslev, Tim, Benjamin Hood, John Huss, and Lasse Heje Pedersen. 2018. "Risk Everywhere: Modeling and Managing Volatility." *The Review of Financial Studies* 31 (7): 2729–73. doi:10.1093/rfs/hhy041.
- Brzenk, Phillip, Liu, Berlinda and Tianyin Chen. 2020. "Indexing Risk Parity Strategies." S&P Dow Jones Indices, October 14.
- Cashin, Paul, Luis Felipe Céspedes, and Ratna Sahay. 2004. "Commodity Currencies and the Real Exchange Rate." *Journal of Development Economics* 75 (1): 239–68. doi:10.1016/j.jdeveco.2003.08.005.

- Chen, Tianqi, and Guestrin, Carlos. 2016. "*XGBoost: A Scalable Tree Boosting System*." In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16, 785–794. doi:10.1145/2939672.2939785
- Chen, Zhiwu, and Peter J. Knez. 1995. "Measurement of Market Integration and Arbitrage". *Review of Financial Studies*, 8: 287-325.
- Cheung, Yin-Wong, Menzie David Chinn, and Antonio García Pascual. 2005. "Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?" *Journal of International Money and Finance* 24 (7): 1150–75. doi:10.1016/j.jimonfin.2005.08.002.
- Collier, Paul, and Benedikt Goderis. 2012. "Commodity Prices and Growth: An Empirical Investigation." *European Economic Review* 56 (6): 1241–60. doi:10.1016/j.euroecorev.2012.04.002.
- Dimson, Elroy, Oguzhan Karakas, and Xi Li. 2015. "Active Ownership." *Review of Financial Studies* 28 (12): 3225-3268.
- Eccles, Neil Stuart, and Suzette Viviers. 2011. "The Origins and Meanings of Names Describing Investment Practices That Integrate a Consideration of ESG Issues in the Academic Literature." *Journal of Business Ethics* 104 (3): 389–402. doi:10.1007/s10551-011-0917-7.
- Eccles, Robert G., Ioannis Ioannou, and George Serafeim. 2014. "The Impact of Corporate Sustainability on Organizational Processes and Performance." *Management Science* 60 (11): 2835–57. doi:10.1287/mnsc.2014.1984.
- Erb, Claude B. and Harvey, Campbell R. 2006. "The Tactical and Strategic Value of Commodity Futures". SSRN Electronic Journal.

- Erten, Bilge, and José Antonio Ocampo. 2013. "Super Cycles of Commodity Prices since the Mid-Nineteenth Century." *World Development*, 44 (April): 14–30. doi:10.1016/j.worlddev.2012.11.013.
- European Securities and Markets Authority (ESMA). 2021. "Report on Trends, Risks, and Vulnerabilities." European Securities and Markets Authority.
- Eurosif. 2018. "European SRI 2018 Study." <https://www.eurosif.org/wp-content/uploads/2021/10/European-SRI-2018-Study.pdf>.
- Fama, Eugene F. 1984. "Forward and Spot Exchange Rates." *Journal of Monetary Economics* 14 (3): 319–38. doi:10.1016/0304-3932(84)90046-1.
- Filippou, Ilias, and Mark P. Taylor. 2017. "Common Macro Factors and Currency Premia." *Journal of Financial and Quantitative Analysis* 52 (4): 1731–63. doi:10.1017/s0022109017000424.
- Friede, Gunnar, Timo Busch, and Alexander Bassen. 2015. "ESG and Financial Performance: Aggregated Evidence from More than 2000 Empirical Studies." *Journal of Sustainable Finance & Investment* 5 (4): 210–33. doi:10.1080/20430795.2015.1118917.
- Froot, Kenneth A., and Richard H. Thaler. 1990. "Anomalies: Foreign Exchange." *Journal of Economic Perspectives* 4 (3): 179–92. doi:10.1257/jep.4.3.179.
- Galema, Rients, and Dirk Gerritsen. 2022. "ESG Rating Score Revisions and Stock Returns." *Social Science Research Network*, January. doi:10.2139/ssrn.4218969.
- Galema, Rients, Auke Plantinga, and Bert Scholtens. 2008. "The Stocks at Stake: Return and Risk in Socially Responsible Investment." *Journal of Banking and Finance* 32 (12): 2646–54. doi:10.1016/j.jbankfin.2008.06.002.

- Gloßner, Simon. 2017. "ESG Risks and the Cross-Section of Stock Returns." *Social Science Research Network*, January. doi:10.2139/ssrn.2980917.
- Gorton, Gary, and K. Geert Rouwenhorst. 2006. "Facts and Fantasies about Commodity Futures." *Financial Analysts Journal* 62 (2): 47–68. doi:10.2469/faj.v62.n2.4083.
- Hallez, Emile. 2022. "Variability in ESG Ratings Makes a Murky Picture for Investors." *ESG Clarity*. August 19. <https://esgclarity.com/variability-in-esg-ratings-makes-a-murky-picture-for-investors/>.
- Hartley, Jonathan. 2020. "Interest Rate Momentum Everywhere across Global Yield Curves." *Social Science Research Network*, January. doi:10.2139/ssrn.3542813.
- Harvey, Campbell R., Edward Hoyle, Russell Korgaonkar, Sandy Rattray, Matthew Sargaison, and Otto Van Hemert. 2018. "The Impact of Volatility Targeting." *Social Science Research Network*, January. doi:10.2139/ssrn.3175538.
- Huck, Nicolas, and Komivi Afawubo. 2014. "Pairs Trading and Selection Methods: Is Cointegration Superior?" *Applied Economics*, 47 (6): 599–613. doi:10.1080/00036846.2014.975417.
- IMF. 2023. "Groups and Aggregates." *World Economic Outlook*. <https://www.imf.org/en/Publications/WEO/weo-database/2023/April/groups-and-aggregates>.
- Kaiser, Lars. 2020. "ESG Integration: Value, Growth and Momentum." *Journal of Asset Management* 21 (1): 32–51. doi:10.1057/s41260-019-00148-y.
- Kempf, Alexander, and Peer Osthoff. 2007. "The Effect of Socially Responsible Investing on Portfolio Performance." *European Financial Management* 13 (5): 908–22. <https://doi.org/10.1111/j.1468-036x.2007.00402.x>.

- Khan, Mozaffar, George Serafeim, and Aaron Yoon. 2016. "Corporate Sustainability: First Evidence on Materiality." *The Accounting Review* 91 (6): 1697-1724.
- Krauß, Christopher. 2016. "STATISTICAL ARBITRAGE PAIRS TRADING STRATEGIES: REVIEW AND OUTLOOK." *Journal of Economic Surveys*, 31 (2): 513–45. doi:10.1111/joes.12153.
- Kroencke, Tim Alexander, Felix Schindler, and Andreas Schrimpf. 2011. "International Diversification Benefits with Foreign Exchange Investment Styles." *Social Science Research Network*, January. doi:10.2139/ssrn.1790764.
- Lundberg, S. (2017, May 22). A Unified Approach to Interpreting Model Predictions. arXiv.org. <https://arxiv.org/abs/1705.07874v2>
- Markowitz, Harry M. 1952. "Portfolio Selection." *The Journal of Finance*, 7 (1): 77. doi:10.2307/2975974.
- Moskowitz, Tobias J., Yao Hua Ooi, and Lasse Heje Pedersen. 2011. "Time Series Momentum." *Social Science Research Network*, January. doi:10.2139/ssrn.2089463.
- Nagy, Zoltán T., Douglas G. Cogan, and Dan Sinnreich. 2013. "Optimizing Environmental, Social and Governance Factors in Portfolio Construction: Analysis of Three ESG-Tilted Strategies." *Social Science Research Network*, January. doi:10.2139/ssrn.2221524.
- Nagy, Zoltan, Kasam, Atlaf and Linda-Eling Lee. 2016. "Can ESG Add Alpha?" SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3175538>.
- Ng, Victor, and Stephen Craig Pirrong. 1994. "Fundamentals and Volatility: Storage, Spreads, and the Dynamics of Metals Prices." *The Journal of Business* 67 (2): 203. doi:10.1086/296630.

NYU Stern. Uncovering the Relationship by Aggregating Evidence from 1,000 Plus Studies

Published between 2015 – 2020. New York University Stern School of Business 2021,

URL:<https://www.stern.nyu.edu/sites/default/files/assets/documents/ESG%20Paper%20Aug%202021.pdf>

Ouaknine, Yannick, Nimit Agarwal, Aiswarya Sankar, Lorna Lyett, and Virgile Haddad. 2019.

"ESG Rating and Momentum: Do Stocks with Improving ESG Ratings Represent Potential Upside for Investors?" Societe Generale Cross Asset Research, Equity Research, March 5.

TangKe, and XiongWei. 2012. "Index Investment and the Financialization of Commodities."

Financial Analysts Journal, 68 (6): 54–74. doi:10.2469/faj.v68.n6.5.

Vidyamurthy, Ganapathy. 2004. Pairs Trading: Quantitative Methods and Analysis. New

Jersey: John Wiley & Sons, Inc. doi:10.1007/s40685-014-0015-7.

Vojtko, Radovan, and Matúš Padyšák. 2019. "Quant's Look on ESG Investing

Strategies." *Social Science Research Network*, January. doi:10.2139/ssrn.3504767.

Von Wallis, Miriam, and Christian Klein. 2014. "Ethical Requirement and Financial Interest:

A Literature Review on Socially Responsible Investing." *Business Research* 8 (1): 61–98.

Bossaerts, Peter, Richard C. Green. 1989. "A General Equilibrium Model of Changing Risk Premia: Theory and Tests". *The Review of Financial Studies*, 2(4): 467–493.

Ouyang, Liwen. 2020. "Seeking Signals from ESG Data." Bloomberg Enterprise Quants. 2020,

URL:

https://data.bloomberglp.com/professional/sites/10/798791_SeekingSignalsFromESGData.pdf

1 Introduction: An Overview of Past Literature

This paper analyses the application of ‘Commodity Terms of Trade’ (CToT) as a trading signal in a universe of 36 different emerging and advanced market currencies. It is hypothesised that currencies with higher commodity dependence scores will perform better under the strategy as their currencies should, in theory, be more affected by changes in commodity prices.

Using macroeconomic variables to predict changes in asset prices is far from a novel concept. Many common currency trading strategies attempt to exploit exchange rate movements which deviate from the uncovered interest rate parity condition, which suggests that two countries' interest rate differential is equivalent to the expected change in the exchange rate between those two currencies. Bilson (1987), Fama (1984) and Froot and Thaler (1990) have all found empirical evidence for deviations from the condition which imply interest rates are not the optimal predictor of future exchange rates (Filippou & Taylor, 2017). These deviations imply that other variables could be used to predict future movements. This paper will attempt to use CToT to this end in a dollar-denominated portfolio trading long and short in foreign currencies.

In general, an economy's terms of trade, the ratio of its export prices to its import prices, can be interpreted as the quantity of import goods an economy can purchase per unit of its export goods. As one might expect, ‘commodity’ terms of trade follow a similar definition, but consider only the ratio of commodity export and import prices. Multiple prior studies have attempted to model real exchange rates using a variety of methodologies, Meese and Rogoff (1983) famously found that a range of exchange rate models based on macroeconomic fundamentals were unable to predict nominal and real exchange rates more accurately than a simple random walk model, a finding reiterated by Cheung, Chinn and Pasual (2005). Much of the empirical evidence which refutes the predictability of exchange rates was conducted on developed market currencies, however, more recent studies which focus on commodity-exporting, developing economies claim to overturn these findings, such as Kohlscheen, Avalos

and Schrimpf (2017), who found (pseudo) out-of-sample predictability of exchange rates using commodity prices in a sample of commodity-exporting countries, or Cashin, Céspedes and Sahay (2004) who found evidence of a long run relationship between real exchange rate and real commodity prices. Indeed, if exchange rates do follow a random walk then changes to the exchange rates should persist and vary randomly, as opposed to returning to a long-run equilibrium as proposed by the theory of purchasing power parity, which posits that exchange rates should move towards the level that the prices of a basket of goods and services in different countries (Terborgh, 1926). This paper will build rolling weight CToT indices for the selected 36 currencies following the methodology of Gruss and Kebhaj (2019), before using these indices as a trading signal for each currency against the US dollar (USD).

2 Economic Reasoning For The Strategy

Ceteris paribus, improvement in terms of trade should be a positive sign for an economy (Collier & Goderis, 2012), imports become relatively cheaper compared with exports and demand for the country's exports increases, which in turn increases demand for the country's currency, causing it to appreciate. Given this theoretical, it is not hard to imagine that accurate and timely terms of trade data could bring a rich vein of market insight to market participants.

The most common way in which terms of trade changes are measured is through indices which weigh the effect of price changes in goods and services based on the contribution to the economy, to determine the net effect of a price change on terms of trade. However, official figures are released with a lag, as trade data isn't reported instantaneously, figures are entirely compiled in hindsight, using only actual prices paid for completed trade. To address this issue, market participants have come to focus more on CToT indices. These indices consider a basket of relevant primary commodities, while not encompassing all of a country's trade, their prices tend to vary more than those of more processed goods and services and thus tend to have a larger influence on short-term macroeconomic dynamics (Arezki et al., 2012). Additionally,

prices of such commodities are available point-in-time from markets. The most comprehensive is the monthly IMF database following the methodology of Gruss and Kebhaj (2019), which covers 182 economies and 45 different commodities. However, when it comes to use in financial markets, the space has been dominated by Citibank's indices which are updated daily or weekly, depending on the economy. Being a for-profit institution, the methodology used to create these indices is not publicly available, however, analysis of these indices compared with others of known methodology such as the IMF indices seems to indicate they use a fixed weighting scheme. That is one which determines the weights of each commodity based on a fixed period, meaning they are static over time and may not accurately represent the changing significance of different commodities to specific economies. This paper will aim to build upon existing literature, to develop predictive indices, as well as ensuring they are representative of a country's current trade flows, providing real-time data for use as a signal for currency trading.

3 Data and Research Methodology

3.1 Data Review – Commodity Terms of Trade Indices

Daily historical price data for the 34 primary commodities considered was extracted from Bloomberg, details of each commodity and ticker can be found in *Appendix 1*.

The World Bank's Manufactured Goods Unit Value Index (MUV) was also taken from the 'Commodity Markets' data centre to deflate historical commodity prices, as well as annual trade data for each traded economy, details of which can be found in *Appendix 2*, from the UN CommTrade international trade database.

This database is compiled using voluntary reporting and as such is vulnerable to missing data and inaccurate reporting. Gaps in trade flows reported by countries were surprisingly prevalent, with some countries lacking any reporting whatsoever for multiple years. However, as countries

report both their export and import flows, as well as trade partners, it was possible to fill any gap years by analysing the data for which the problem country had been cited as a trade partner.

Some commodities such as natural gas can be linked with one specific heading (321) but most are associated with multiple. Commodities were grouped by STIC codes which refer to the unprocessed commodity as well as products derived from the commodity, whose price is expected to closely follow the primary commodity. A full table of the primary commodities and corresponding STIC codes can be found in *Appendix 3*.

3.2 Data Review – Financial Data Used in Return Calculation and Analysis

The earliest point in time from which each required dataset was available was the beginning of 2004, for this reason, each dataset was collected from then until the end of 2022, the last full calendar year.

Historical daily spot rates and bid/ask spreads to calculate transaction costs were collected for each currency pair from Refinitiv, as well as the historical central bank policy rate of each currency union from the respective central bank's website, to serve as a risk-free rate in calculations. Each currency was also classified as either advanced or emerging according to the IMF's World Economic Outlook classification (2023), and given a commodity dependence score according to the United Nations State of Commodity Dependence Report (2021). In total, the sample of currencies comprises 17 and 19 advanced and emerging economies respectively, with an average commodity dependence score of 47.98%. A full table of the economies, their currency, currency classification, and commodity dependence score can be found in *Appendix 2*. It should be noted that these classifications and scores are from 2023 and 2021, respectively, therefore they may not provide an accurate picture of economies on all dates covered in the testing period, however, this does not affect the construction or performance of the strategy, it is simply a limitation of the subsequent analysis.

To provide a point of comparison, the J.P. Morgan USD Spot Index (JPMQUSD Index) was used as a benchmark. The index tracks the performance of the USD against a basket of 10 global

currencies and was chosen as it provides a fair representation of the returns an investor could generate passively through an index of currencies trading against the USD. Details of the index can be found in *Appendix 4*. Emerging market factors were taken from the Dartmouth Ken French factor database to analyse the returns exposure to known factors via OLS regression under the Fama French 5-factor model. The risk-free rate used to calculate excess returns was the return of a one-month US treasury bill, as the portfolio is dollar-denominated.

3.3 Methodology – Construction of CToT indices

This paper follows a similar methodology to Gruss and Kebhaj (2019), constructing a country-specific price index for the chosen commodities, with adjustments which allow them to be forecast instead of historical and daily as opposed to monthly. The indices were constructed according to the following equation:

$$\Delta \text{Log}(\text{Index})_{i,t} = \sum_{j=1}^J \Delta P_{j,t} \Omega_{i,j,t} \quad (1)$$

$P_{j,t}$ denotes logged real price of commodity j at time t , Δ denotes first differences, $\Omega_{i,j,t}$ denotes the country-specific time-varying weights.

Calculation of Index Weights

Since the indices will be used to predict the effect of real-time changes in commodity prices on exchange rates, time-varying weights were chosen over fixed weights, ensuring that the weights more closely track the significance of different commodities to the country in question's economy. The weights were calculated as the average trade flow over the previous three-year period (Gruss & Kebhaj, 2019) (see equation 2):

$$\Omega_{i,j,t} = 1/3 \sum_{s=1}^3 \omega_{i,j,T-s} \quad (2) \quad \omega_{i,j,T} = \frac{x_{i,j,T} - m_{i,j,T}}{TT_{i,T}} \quad (3)$$

The weight of each individual commodity $\omega_{i,j,T}$, was determined by weighting commodity j by the contribution of that commodity's net exports to total trade flow in the country $TT_{i,T}$ (see

equation 3). $x_{i,j,T}$ and $m_{i,j,T}$, respectively, denote the exported and imported value of commodity j for country i in year T . As trade data is reported annually, the opportunity to rebalance weights was limited, however as weights are calculated as three-year averages, weights were able to be rebalanced monthly, using a 36-month rolling window. Moreover, commodity price data is available daily, allowing the capture of the effect of price changes on terms of trade daily. This, as well as the fact that weights for the current period were predetermined by the previous periods, allows for the price effect on CToT to be isolated from the supply effect of a price change (e.g. commodity price increases so a country increases export of that commodity), while still incorporating the increased significance of the commodity in the index, assuming an increase or decrease in price is sustained. The use of net exports to weight commodities is preferred as it allows for the effect of variations in the price of commodities which are net imports to be taken into account. For example, many countries are net exporters of non-energy commodities but are net importers of energy commodities, thus general increases in commodity prices may result in an improvement in their CToT.

4 Strategy Framework and Definition

The strategy was tested on a universe of 36 developed and emerging market currencies, including the US dollar as the base currency against which the remaining 35 currencies are traded. The decision was made to trade foreign currencies against the USD to ensure the most liquid foreign exchange market for each currency pair, aside from a few select European countries whose most liquid market is with the Euro but are traded in sufficient volumes with the USD that liquidity was not a concern. As the strategy is systematic, all calculation and implementation was done via a model with pre-defined parameters in Python.

The strategy is unique in its approach to currency trading as it uses the CToT of each currency union to determine the weight and direction of trades based on the change relative to the change in the CToT of the US. The logic behind this signal is that if a foreign country or

currency union's terms of trade have improved more than that of the US, or indeed if the US CToT has fallen, then there should be an appreciation in the foreign currency relative to the USD as more of the foreign currency is demanded, and relatively less of the USD is demanded. In this scenario, the strategy took a long position in the foreign currency against the USD in anticipation of a subsequent appreciation. In the case of a decrease in the foreign currency unions CToT relative to the USD, a short position was taken in the foreign currency as by the same logic, the USD was expected to appreciate relative to the foreign currency. The size of the position as well as the weighting given to the trade in the combined portfolio of currency pairs was determined by the magnitude, measured in z-scores, of the deviation relative to the historical mean, with deviations of higher magnitude being allocated greater weight in the overall portfolio. Using the historical mean prevents forward-looking bias and means that as the strategy operates for longer, the mean becomes more stable and representative of the 'true mean', in theory allowing the strategy to improve in performance as time progresses. The weights were rebalanced weekly, with the trading day being Friday, and were constructed according to the following equations:

$$(4) \quad \Omega_{i,t} = \frac{\Delta Z_{i,t} - \Delta Z_{USA}}{\sum_{s=1}^{35} \Delta Z_{i,t} - \Delta Z_{USA,t}} \quad (5) \quad \Delta Z_{i,t} = Z_{i,t} - Z_{i,t-4}$$

Where $\Omega_{i,t}$ and $Z_{i,t}$, respectively represent the weight and z-score of currency i at time t . Following their calculation the weights were then normalised to ensure the absolute value of their sum was one. Deviations were measured over the previous 4 weeks to ensure rapid or short-lived commodity price fluctuations did not lead to positions taken without sufficient economic reasoning, while weekly rebalancing still allowed the strategy to be dynamic enough to capture significant terms of trade dislocations and leverage them to make profitable trades. Z-scores were winzorised at 2 standard deviations to mitigate the effect of data outliers and prevent excessive risk-taking in any single market, which could unnecessarily compromise the diversification effect of holding a larger basket of currency pairs.

Returns to the strategy were comprised of two components; the change in the exchange rate of the currency (exchange rate effect) and the foreign risk-free rate earned in long positions or paid in short positions (interest rate effect). The transaction costs of the strategy were estimated according to Bollerslev et al. (2016), using half the median of the bid/ask spread of each currency over the last nine months. The so-called ‘half-spread’ is used as the methodology assumes strategic trading would reduce the likelihood of paying the full spread, as well as the fact there was a ‘netting effect’ in cases where entire positions were not required to be bought or sold upon rebalancing. The weighted returns of each foreign currency/USD pair were then summed to give the portfolio return.

As the commodity dependence of different currency unions varies significantly, the ability of CToT to predict changes in each union’s exchange rate is also likely to vary. To determine which currencies operated best under the weighting scheme, the strategy was first tested in-sample before excluding the currency pairs which did not contribute to the risk-adjusted return of the strategy, allowing greater weight to be allocated to the remaining currencies. Three variations of the above methodology were tested, ‘long/short’, ‘long only’ and ‘short only’. As the strategy portfolio was dollar-denominated, the long-only strategy took only positions long a foreign currency against the USD and vice versa.

5 Performance and characteristics of the portfolio

To properly evaluate the relative performance of the CToT strategies, performance statistics are compared to the aforementioned JPMQUSD Index, hereafter referred to as ‘the benchmark’.

5.1 In-Sample Testing

The performance of the strategies trading 35 different currency pairs (36 currencies) during the in-sample period, spanning 01/01/2004 to 2017 end is shown in *Table 1* below. All figures are annualised, transaction costs represent the total expense ratio, the same is true of *Table 2*.

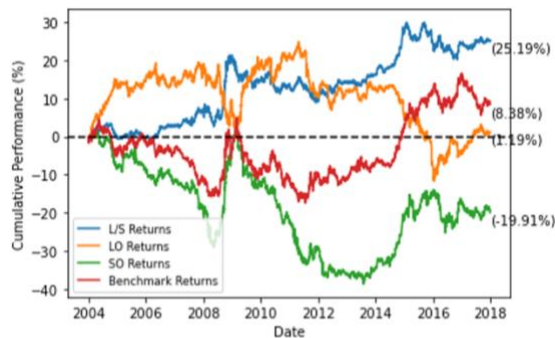
Table 1: In-Sample Performance Statistics

	Annualized Return	Annualized Volatility	Skewness	Kurtosis	Sharpe Ratio	Max Drawdown	Transaction Costs
Long/Short Strategy	1.67%	4.19%	0.44	9.73	0.083	12.00%	0.39%
<i>Exchange Rate Effect</i>	<i>(1.82%)</i>						
<i>Interest Rate Effect</i>	<i>(-0.15%)</i>						
Long only Strategy	-0.11%	6.13%	-1.09	14.96	-0.178	31.23%	0.44%
<i>Exchange Rate Effect</i>	<i>(3.45%)</i>						
<i>Interest Rate Effect</i>	<i>(-3.56%)</i>						
Short only Strategy	-1.56%	6.12%	0.23	6.75	-0.466	33.95%	0.56%
<i>Exchange Rate Effect</i>	<i>(-4.83%)</i>						
<i>Interest Rate Effect</i>	<i>(3.27%)</i>						
Benchmark	0.10%	5.72%	-0.06	3.69	-0.184	22.67%	

Note: Strategy returns are decomposed into the exchange rate effect and the interest rate effect in ()

During the in-sample period, **Table 1** illustrates that the benchmark yielded an annualized return of 0.1%. However, the risk-adjusted return and skewness were negative, with a Sharpe ratio of -0.184, indicating a higher chance of extreme negative returns. In contrast, the long/short strategy emerged as the top performer, delivering an annualised return of 1.67% and a positive risk-adjusted return with a Sharpe ratio of 0.083, the only positive return and ratio of the strategies. Throughout this period, the long/short strategy also exhibited the lowest volatility, the smallest maximum drawdown, and the most positive skewness, suggesting a heightened likelihood of extreme positive returns. These observations are visually depicted in **Figure 1**, where the line representing the long/short strategy returns demonstrates a more consistent and upward trajectory, with instances of volatility tending towards the positive direction. In

Figure 1: Cumulative Returns In-Sample **Figure 2: Cumulative Returns Out-of-Sample**



contrast, the benchmark, long-only, and short-only trends appear less consistent, characterized by more frequent periods of downside volatility.

Given the results of the in-sample tests, the long/short strategy was chosen as the ‘focal’ strategy, the other two were still analysed out-of-sample, but only to provide a comparison to the focal strategy. After the in-sample testing period, the performance of each currency under the weighting scheme was evaluated using the aforementioned selection criteria.

5.2 Out-of-Sample Testing

A full table detailing the currencies which were excluded/kept in each strategy can be found in *Appendix 5*. The performance statistics for the strategies with the selected currencies are shown in *Table 2* below.

Table 2: Out-of-Sample Performance Statistics

	Annualized Return	Annualized Volatility	Skewness	Kurtosis	Sharpe Ratio	Max Drawdown	Transaction Costs
Long/Short Strategy	8.84%	11.40%	1.17	15.51	0.624	21.22%	1.11%
<i>Exchange Rate Effect</i>	<i>(7.49%)</i>						
<i>Interest Rate Effect</i>	<i>(1.35%)</i>						
Long only Strategy	4.52%	6.62%	-0.28	5.24	0.473	9.35%	0.65%
<i>Exchange Rate Effect</i>	<i>(3.65%)</i>						
<i>Interest Rate Effect</i>	<i>(0.87%)</i>						
Short only Strategy	19.05%	11.61%	2.21	26.08	1.452	12.06%	1.14%%
<i>Exchange Rate Effect</i>	<i>(13.64%)</i>						
<i>Interest Rate Effect</i>	<i>(5.41%)</i>						
Benchmark	1.67%	5.40%	0.34	5.31	0.094	12.85%	
Note: Strategy returns are decomposed into the exchange rate effect and the interest rate effect in ()							

During this period the benchmark produced positive annualised returns, both nominally and risk-adjusted, exhibiting a Sharpe ratio of 0.094, while maintaining a relatively consistent volatility. It also increased in skewness to a positive level of 0.34, which makes sense given the increased returns and similar volatility, as well as the maximum drawdown reducing to 12.85%.

In this sample, all three of the strategies outperformed the benchmark with the short-only strategy showing the strongest performance, with an impressive annualised return of 19.05% and Sharpe ratio of 1.452. The focal, long/short strategy increased nominal and risk-adjusted returns as well as skewness from in-sample testing levels to 8.84%, 0.624 and 1.17, respectively. The volatility and Maximum drawdown of the strategy however did increase, although this increase in risk was justified in risk-adjusted terms given the increase in the Sharpe ratio. These effects were also expected as the number of currencies being traded by the strategy significantly reduced and therefore the overall portfolio was less diversified and more exposed to fluctuations in the returns of the individual currencies.

Given the focal strategy's outperformance of the benchmark in nominal and risk-adjusted returns both in-sample and out-of-sample and the outperformance of both the long-only and short-only strategy out-of-sample, simple linear regressions were run to determine the beta with the benchmark and if there was a statistically significant excess return which the Fama-French 5 factor model could not explain. The abbreviated results are shown in **Table 3** below, for the full table, including coefficients of factors in the models, refer to **Appendix 6**.

Table 3: Fama French 5 Factor Model Regression Results

	In-Sample			Out-of-Sample		
	L/S	LO	SO	L/S	LO	SO
Alpha (α)	0.12% (0.927)	-0.96% (0.483)	-3.12%* (0.022)	11.16%* (0.024)	4.08%* (0.040)	19.8%* (0.022)
Market Risk (β)	0.107 (0.136)	-0.514** (0.00)	0.701** (0.00)	0.352 (0.155)	-0.8662** (0.000)	0.835* (0.014)
R^2	0.072	0.524	0.587	0.213	0.600	0.158
Information Ratio	0.008	-0.062	-0.202	0.327	0.296	0.434
Tracking Error	0.013	0.013	0.013	0.028	0.012	0.038
Note: p-values in (); **, * denote statistical significance at 1 and 5% level, respectively; α shown as annual %						

In-Sample, only the Long-short strategy generated a positive alpha but this was not statistically significant at the 95% confidence level. However, out-of-sample, all strategies produced a positive alpha, with the short-only strategy understandably producing the largest at 19.8%

annually. The focal long/short strategy produced the second largest at 11.16%, with the long-only producing 4.08% per year. The market betas varied highly across strategies, with the largest coming from the short-only strategy.

Given the results, we can say that CToT strategies can produce statistically significant alpha, yielding returns that exposure to the market and established emerging market factors cannot explain. We cannot conclusively say this in the case of the long/short strategy, however, that is not to say that the strategy does not produce alpha, only that in using the Fama-French model and the chosen benchmark we fail to reject the null hypothesis that the observed alpha is simply a result of exposure to the market and emerging market factors.

6 Discussion

Clear evidence for the hypothesis that emerging market currencies and those with greater commodity dependence would perform better under the strategy was found. In the in-sample period, 75%, 67% and 63% of currencies which contributed to the risk-adjusted return for the long/short, long-only and short-only strategies respectively, were emerging market currencies. Moreover, as **Table 4** below depicts, the average commodity dependence of the strategy constituents increased significantly.

Table 4: Average Commodity Dependence of Strategy Constituents

	Long/Short	Long Only	Short Only
In-Sample	47.98%	47.98%	47.98%
Out-of-Sample	67.00%	68.04%	56.39%

A further finding which follows the expectation of emerging market currencies performing well under the strategy is that the transaction costs of the strategy increased between testing periods for each of the three strategies. This can be explained by the fact that, as previously mentioned, the resulting portfolios after the out-of-sample currency selection criteria were made up of a

much higher percentage of emerging market currencies. These currencies tend to be less liquid and have higher bid/ask spreads than those of highly traded, advanced economy currencies.

It can be seen that the short-only strategy performed significantly better in the out-of-sample period despite posting negative returns in the in-sample period. This makes sense as the benchmark also performed considerably better out-of-sample therefore it is understandable that the strategy would perform well when the USD is appreciating against other currencies.

Despite lower risk-adjusted returns out-of-sample than the short-only strategy, the decision to select the long-short strategy as the focal strategy remained valid upon reflection at the end of the out-of-sample period. The long/short strategy provided positive risk-adjusted returns, consistently outperforming the benchmark in both sample testing periods even before the currency selection criteria were applied. The short-only and long-only strategy did outperform the benchmark out-of-sample, despite the long-only strategy taking only short positions in USD during a time when the USD performed overall positively, but, like the long-only strategy, it performed negatively in the in-sample period when the USD performed less strongly. Consequently, the ability of the long/short strategy to dynamically take both long and short positions against the USD appears to have allowed it to generate positive risk-adjusted returns in the sample testing periods regardless of the performance of the USD against other currencies.

7 Implementation obstacles

One of the main sources of implementation obstacles when building a systematic trading strategy stems from data. This concern is exacerbated by the fact the CToT strategies tend to trade emerging and more commodity-dependent currencies. Unfortunately, it is these countries that tend to have less reliable reporting standards and consistency when it comes to trade data used to build the CToT indices. Although there are techniques to tackle this issue such as using data from trade partners, instances can occur where there is trade between two countries' economies which neither reports, and therefore there is no reliable and accurate account of the

trade flows. This could have the potential to distort CToT indices weights, reducing the efficacy of the trading signal. Moreover, the fact that trade data is available only annually means that the cyclical patterns of international trade throughout the year could be understated. A further potential issue associated with trading emerging and more commodity-dependent currencies is the reduced liquidity that comes with trading such currencies, as well as the higher transaction costs. This could reduce the scalability of the strategy in the case of a large investor.

8 Conclusion

In conclusion, this paper contributes to the literature by detailing a method of building predictive, rolling weight commodity terms of trade indices, which take into account daily commodity price changes for a range of 36 emerging and advanced currency unions, as well as three systematic trading strategies which use these commodity terms of trade indices as a signal. Analysis shows that after the in-sample selection criteria were applied, all three of the long/short, long only and short-only strategies outperformed the benchmark both in absolute and risk-adjusted terms, while the short-only and long-only provided statistically significant alpha when regressed against the Fama-French 5 Factor model, during a period where the US dollar performed strongly. These findings lend support to the findings of Filippou and Taylor (2017), implying that macroeconomic variables other than interest rate can be used to predict changes in exchange rates and further that interest rate is not the sole optimal predictor as the uncovered interest rate parity condition would imply.

The long/short strategy was the only one to outperform the benchmark in the in-sample period, a time when the US dollar performed more modestly, however, the strategies had not been tested and currency selection applied at this time. This leaves two obvious areas for research in the future; firstly, to determine how would the three strategies perform in a time of modest USD performance such as the in-sample period if the selection criteria had already been applied; and secondly, to determine performance during a period of USD underperformance.

9 References

- Arezki, Mr Rabah, Ms Catherine A. Pattillo, Mr Marc G. Quintyn, and Min Zhu. *Commodity price volatility and inclusive growth in low-income countries*. International Monetary Fund, 2012.
- Bilson, John FO, and David A. Hsieh. "The profitability of currency speculation." *International Journal of Forecasting* 3, no. 1 (1987): 115-130.
- Bollerslev, Tim, Benjamin Hood, John Huss, and Lasse Heje Pedersen. "Risk everywhere: Modeling and managing volatility." *The Review of Financial Studies* 31, no. 7 (2018): 2729-2773.
- Cashin, Paul, Luis F. Céspedes, and Ratna Sahay. "Commodity currencies and the real exchange rate." *Journal of Development Economics* 75, no. 1 (2004): 239-268.
- Cheung, Yin-Wong, Menzie D. Chinn, and Antonio Garcia Pascual. "Empirical exchange rate models of the nineties: Are any fit to survive?." *Journal of international money and finance* 24, no. 7 (2005): 1150-1175.
- Collier, Paul, and Benedikt Goderis. "Commodity prices and growth: An empirical investigation." *European Economic Review* 56, no. 6 (2012): 1241-1260.
- Fama, Eugene F. "Forward and spot exchange rates." *Journal of monetary economics* 14, no. 3 (1984): 319-338.
- Filippou, Ilias, and Mark P. Taylor. "Common macro factors and currency premia." *Journal of Financial and Quantitative Analysis* 52, no. 4 (2017): 1731-1763.
- Froot, Kenneth A., and Richard H. Thaler. "Anomalies: foreign exchange." *Journal of economic perspectives* 4, no. 3 (1990): 179-192.
- IMF. 2023. Groups and aggregates. *World Economic Outlook*,. Available at: <https://www.imf.org/en/Publications/WEO/weo-database/2023/April/groups-and-aggregates>
- Meese, R.A. and Rogoff, K., 1983. Empirical exchange rate models of the seventies: Do they fit out of sample?. *Journal of international economics*, 14(1-2), pp.3-24.
- Terborgh, George W. "The purchasing-power parity theory." *Journal of Political Economy* 34, no. 2 (1926): 197-208.

United Nations Conference on Trade and Development. 2021. The State of Commodity Dependence. *The State of Commodity Dependence Report 2021*. Available at: <https://unctad.org/publication/state-commodity-dependence-2021>

10 Appendix

Appendix 1 – Primary Commodities and their tickers

Commodity	Bloomberg Ticker
Coal	XA1 Comdty
Crude Oil	CO1 Comdty
Natural Gas	NG1 Comdty
Aluminium	LMAHDY Comdty
Copper	LOCADY Comdty
Gold	GOLDLNPM Index
Silver	SI1 COMB Comdty
Platinum	PL1 COMB Comdty
Iron Ore	ISIX62IU Index
Lead	LOPBDY Comdty
Nickel	LONIDY Comdty
Tin	LOSNDY LME Comdty
Zinc	LOZNDY LME Comdty
Bananas	WRCOBANA Index
Barley	WA1 Comdty
Beef	LC1 Comdty
Cocoa	COCOCBUS Index
Coffee	KC1 Comdty
Corn (Maize)	C 1 Comdty
Fish	CSA1 Comdty
Fish Meal	BROLFM Index
Olive Oil	OLO8SSM1 Index
Oranges	JO1 Comdty
Palm Oil	KO1 Comdty
Poultry	WPFCFRCH Index
Rice	RR1 Comdty
Soybean Oil	BO1 Comdty
Soybeans	S 1 Comdty
Sugar	QW1 Comdty
Swine Meat	LH1 Comdty
Tea	MTEASOUT Index
Wheat	W 1 Comdty
Cotton	CT1 Comdty

Appendix 2 - Country/Currency Union Information

Country/ Currency Union	Currency	Symbol	Emerging/Advanced Economy	Commodity Dependence Score
Argentina	Argentine Peso	ARS	Emerging	82.1%
Australia	Australian Dollar	AUD	Advanced	89%
Brazil	Brazilian Real	BRL	Emerging	73.4%
Canada	Canadian Dollar	CAD	Advanced	52.1%
Chile	Chilean Peso	CLP	Emerging	88.3%
China	Chinese Yuan	CNY	Emerging	5.8%
Hong Kong	Hong Kong Dollar	HKD	Advanced	NA
Colombia	Colombian Peso	COP	Emerging	78.6%
Eurozone	Euro	EUR	Advanced	NA
Israel	Israeli New Shekel	ILS	Advanced	22.1%
Japan	Japanese Yen	JPY	Advanced	8%
Malaysia	Malaysian Ringgit	MYR	Emerging	28.8%
Mexico	Mexican Peso	MXN	Emerging	17.5%
New Zealand	New Zealand Dollar	NZD	Advanced	81.3%
Peru	Peruvian Nuevo Sol	PEN	Emerging	90.9%
Philippines	Philippine Peso	PHP	Emerging	19.8%
Poland	Polish Zloty	PLN	Emerging	20.3%
Rep. of Korea	South Korean Won	KRW	Advanced	12%
Singapore	Singapore Dollar	SGD	Advanced	21%
Sweden	Swedish Krona	SEK	Advanced	23.3%
Switzerland	Swiss Franc	CHF	Advanced	28.6%
Taiwan	Taiwan Dollar	TWD	Advanced	NA
Thailand	Thai Baht	THB	Emerging	28.3%
Turkey	Turkish Lira	TRY	Emerging	21.4%
USA	United States Dollar	USD	Advanced	31.9%
United Kingdom	British Pound	GBP	Advanced	28.6%
Norway	Norwegian Krone	NOK	Advanced	81.6%
Iceland	Icelandic Krona	ISK	Advanced	87.2%
Hungary	Hungarian Forint	HUF	Emerging	13%
Romania	Romanian Leu	RON	Emerging	18.5%
Nigeria	Nigerian Naira	NGN	Emerging	95%
Czechia	Czech Koruna	CZK	Advanced	9.3%
Congo	Congolese Franc	CDF	Emerging	94%
Niger	West African CFA Franc	XOF	Emerging	90.4%
Russian Federation	Russian Ruble	RUB	Emerging	77.5%
South Africa	South African Rand	ZAR	Emerging	63.7%

Appendix 3 – Aggregated Primary Commodities and their STIC Codes

Commodity	STIC Codes						
Coal	321						
Crude Oil	333						
Natural Gas	343						
Aluminium	684	2851	28404	28823			
Copper	682						
Gold	971						
Silver	6811	28911					
Platinum	6812	89741					
Iron Ore	281	2821					
Lead	685	2834	2874	28406			
Nickel	683	28403	2832	2841	2842	2872	28822
Tin	687	28409	2876	28826	51368		
Zinc	686	2835	28407	28408	2875	28825	
Bananas	0513	0573					
Barley	043	0430					
Beef	0011						
Cocoa	072	073					
Coffee	071						
Corn (Maize)	044	04711					
Fish	034	037	03502	03503	03504		
Fish Meal	03501	08142					
Olive Oil	42141	42142					
Oranges	0511	0571					
Palm Oil	4222						
Poultry	0014	0114	01181	0123			
Rice	042						
Soybean Oil	4211						
Soybeans	2222	2214					
Sugar	0611	0612					
Swine Meat	0122						
Tea	0741						
Wheat	041						
Cotton	263	652					

Appendix 4 – J.P. Morgan USD Spot Index Details

Currency	Weight
Euro	29.14%
Japanese Yen	12.82%
Canadian Dollar	11.55%
British Pound	10.16%
Mexican Peso	9.31%
Chinese Yen	7%
Swiss Franc	4.81%
Austrian Dollar	4.35%
South Korean Won	3.27%
Indian Rupee	2.73%
Singapore Dollar	2.61%
Taiwan Dollar	2.25%

Note: Weights are determined annually based on their share of international trade and FX liquidity, weights above are indicative purely indicative and correct as of 2023, Bloomberg ticker JPMQUSD Index

Appendix 5 – Currency Pairs Which Remained In Each Strategy After In-Sample Selection Criteria

Long/Short	Long Only	Short Only
Brazil	Argentina	Chile
Israel	Brazil	Israel
Nigeria	Mexico	Rep. of Korea
Russian Federation	New Zealand	Singapore
	Turkey	Thailand
	Norway	Nigeria
	Iceland	Congo
	Nigeria	Niger
	Russian Federation	

Appendix 6 – Full Results of Regressions

	In-Sample			Out-of-Sample		
	L/S	LO	SO	L/S	LO	SO
Alpha (α)	0.12% (0.927)	-0.96% (0.483)	-3.12%* (0.022)	11.16%* (0.024)	4.08%* (0.040)	19.8%* (0.022)
Market Risk (β)	0.107 (0.136)	-0.514** (0.00)	0.701** (0.00)	0.352 (0.155)	-0.8662** (0.000)	0.835* (0.014)
Size	-0.078 (0.236)	-0.094 (0.152)	0.058 (0.353)	-0.436 (0.093)	0.054 (0.603)	-0.405 (0.241)
Value	0.017 (0.840)	0.168* (0.044)	-0.008 (0.918)	-0.617* (0.014)	0.007 (0.940)	-0.337 (0.303)
Profitability	0.019 (0.860)	-0.196 (0.065)	0.148 (0.144)	-0.483 (0.129)	0.061 (0.632)	-0.253 (0.550)
Investment	0.1718* (0.040)	-0.2812** (0.001)	0.218** (0.007)	0.430 (0.150)	0.008 (0.949)	0.068 (0.864)
R^2	0.072	0.524	0.587	0.213	0.600	0.158
Information Ratio	0.008	-0.062	-0.202	0.327	0.296	0.434
Tracking Error	0.013	0.013	0.013	0.028	0.012	0.038
Note: p-values in (); **, * denote statistical significance at 1 and 5% level, respectively; α shown as annual %						