Norwegian School of Economics
Bergen, Fall 2023

Exploring The Possibilities of Investing in Artificial Intelligence

A comprehensive analysis of NQROBO index performance

Bjørnar Kaldheim & Marius Senderud

Supervisor: Gernot Peter Doppelhofer

Master thesis, Economics and Business Administration

Major: Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

This thesis investigates the potential of beating the market index for an investor by investing in Artificial Intelligence (AI). We have analysed the performance of Nasdaq CTA Artificial Intelligence & Robotics (NQROBO) from January 2018 to August 2023, comparing it to the Nasdaq Composite (NASDAQ) and S&P 500. We have simulated the behaviour of an open-minded investor who uses simple prediction models to forecast returns. We have tried to make this simulation as realistic as possible using minimal hindsight. Our thesis is based on three analyses: a historical analysis evaluating NQROBO's performance, a pseudo-out-of-sample forecasting performance analysis exploring how an investor in real time utilising a forecasting tool would perform, and lastly, an optimal relative weighting analysis of NQROBO, based on the pseudo-out-of-sample analysis.

The historical analysis revealed that NQROBO outperformed the market from 2020 through 2022. It also uncovered that the Alpha was primarily positive from 2020 to early 2022, before turning negative in 2022. The Beta was lower than the market until 2022 before increasing sharply and stabilising at 1,1. Regarding the Fama French Factors, we identified the market as a consistent driver for returns. HML, RMW and CMA fluctuating greatly, being mostly negative, suggesting that NQROBO performs best when the market favours growth-oriented firms with an aggressive investment strategy. Indicating that the index has the potential of outperforming the market over certain periods if the market conditions are favourable. Furthermore, the pseudo-out-of-sample forecasting performance analysis showed that portfolios utilising Sharpe Ratio, RMSE and Hybrid RMSE weighting could outperform the market, if rebalancing daily. Suggesting that potential gains of investing in NQROBO is short lived. Lastly, our optimal relative weighting analysis of NQROBO's shows that a highly dynamic weight allocation that is rebalanced frequently is beneficial. Enabling the portfolio to capture short-term gains and beating the market index over the period. The findings suggest that investing in AI offer the potential of beating the market index if done flexibly.

Acknowledgement

This master thesis was written during the fall of 2023 as a part of our Master of Science in Economics and Business administration at NHH Norwegian School of Economics.

When choosing the topic of our thesis we wanted to apply some of the theoretical theory we have learnt from our courses at NHH, meanwhile, we also wanted our topic to be unique and relevant today and in years to come. The research process has been educational, challenging and rewarding. Resulting in a deep understanding of forecasting methodologies and the potential impact Artificial Intelligence technology may have on the financial landscape.

Lastly, we would like to express our gratitude to our supervisor, Professor Gernot Peter Doppelhofer, who has contributed with advice, guidance, and constructive input throughout this thesis. The prompts received, together with supportive feedback, has made our work educational and motivating. His guidance has undoubtedly enhanced the quality of our research significantly.

Norwegian School of Economics, December 2023

Bjørnar Kaldheim

Marius Senderud

Contents

1. Introduction

As technology keeps developing, becoming more accessible across the world, it has been increasingly integrated into our daily lives. The third Industrial Revolution marked a crucial turning point with the introduction of computing power. This introduction challenged prior conceptions and limitations of what has been deemed impossible. Artificial Intelligence stands at the frontline of innovation in an era that can be characterized as being rapidly developing, continuously introducing new technological advancements. The technology displays an incredible capacity to mimic human-like cognitive thinking and perform complex tasks resulting in an increased adoption across industry sectors. This development is reflected in the remarkable growth in Artificial Intelligence funding, which has grown by over 7x since 2015, reaching $93,5 billion in 2021 (Duarte, 2023). With no signs of slowing down, the global AI market is projected to compound at an annual growth rate of 37,3% until 2030, resulting in an expected market value of $1847,58 billion (Grand View Research, 2023).

As Artificial Intelligence continues its inevitable integration into the core of industries and economies, there is no denying that the technology is here to stay. However, as the road ahead is filled with uncertainties tied to regulations, ethical dilemmas, implementation, and future development, investors find themselves at a crossroads between opportunity and uncertainty. Stephen Hawking (2016) famously said:

"It will either be the best thing that's ever happened to us, or it will be the worst thing. If we're not careful, it very well may be the last thing."

This perfectly highlights how captivating and exciting the potential of Artificial Intelligence is and how it can introduce new possibilities and generate excess returns. However, Hawking also warns about the danger of the technology. More recently, there has been a rising concern regarding the development of the technology as tech leaders has pleaded to have a pause in the development. Currently, it exists no global framework regulating the technology. The dilemma between upside and downside challenges a broad spectrum of stakeholders, from huge institutional asset managers to individual investors.

The crossroad is a thrilling dilemma and has greatly inspired us when writing our master thesis.

1.1 Research Question

This thesis aims to examine the possibility of generating excess returns through investing in Artificial Intelligence technology, ultimately beating the market index. In our analyses, we have regarded excess returns as returns relative to risk-free rate. We define the S&P 500 index as the market and "beating the market" as achieving a higher final wealth by investing in an alternative portfolio rather than investing in the market. This is achieved by examining the development of the Nasdaq CTA Artificial Intelligence & Robotics Index (NQROBO). The index tracks the performance of companies engaged in the AI and robotics segment of the technological, industrial, medical, and other economic sectors. In this thesis, we want to explore the approach for simulating an investor's behaviour in real time through forecasting models, using as little hindsight as possible. By as little hindsight we mean only utilizing historical information that is available for the investor in real time. The thesis aims to provide an answer to our research question:
To what extent does investing in Artificial Intelligence technology, supported by a forecasting tool only using historical information, offer the potential of beating the market index for a risk-neutral investor with a short time horizon?

To answer this question, we have constructed the thesis into three different parts, each part is constructed to enrich our understanding of the research question.

1. **Historical analysis of NQROBO:** Evaluating the historical performance of the index in terms of excess returns, Alpha, Beta, and the Fama French Factors.
2. **Pseudo-out-of-sample forecasting performance analysis:** Constructing and assessing the performance of synthetic portfolios
3. **Optimal relative weighting analysis:** Analysing the optimal relative weightings in dynamic portfolios**.**

To answer the first part of our thesis, we have conducted a historical analysis of the performance of NQROBO. Here, we examine the performance of the NQROBO, the Nasdaq Composite (NASDAQ) and the S&P 500. We divided our findings into three sub-samples to gain deeper insights into the development. We also assessed the index's Alpha and Beta over time before analysing how the Fama French Factors affected NQROBO.

To answer the second part of our thesis, we retrieved financial data for all three indices and constructed different forecasting models. All forecasting models was trained over one year and subsequently, identifying and utilising the most accurate model for each of the three indices. After this, we constructed six synthetic portfolios, each weighted differently based on simple statistical weights, replicating real-world investment strategies (Pesaran & Timmermann, 1995). Ultimately, this leads to a ranking of all portfolios, providing a better understanding of how an investor could beat the market.

Lastly, we wish to examine the optimal weighting within class of models considered of the NQROBO and the other indices. We find it interesting to evaluate the development of the weighting in our dynamic portfolios over time to better understand when it would be most beneficial to invest in NQROBO. This is done by computing all weights for the portfolios and illustrating the development.

1.2 Thesis Structure

The thesis is organised into different chapters. Chapter 2 presents the reader with the basics of Artificial Intelligence technology. The chapter explains essential critical concepts within the technology and presents, current, and possible future development. Chapter 3 introduces the reader to fundamental relevant financial theories and performance measurement on which our analyses are based on. Chapter 4 outlines the methodology of our research. Chapters 5 and 6 present six synthetic portfolios and our data to the reader. In Chapter 7, we conduct a historical analysis of NQROBO's performance and analyse Alpha, Beta, and the Fama French Factors. In chapter 8, we present our empirical portfolio performance results. Chapter 9 consists of a discussion of the results presented in the prior chapter, limitations, robustness analysis and proposals for further research. Chapter 10 summarises the main findings from our thesis before we conclude.

2. Artificial Intelligence Technology

This chapter gives a short introduction to the fundamental principles of Artificial Intelligence (AI) and its multiple applications. We provide context for a deeper understanding of AI technology by discussing historical developments and future opportunities.

2.1 Nasdaq CTA Artificial Intelligence & Robotics

The NQROBO index was introduced on 18 December 2017, and tracks the performance of companies engaged in the AI and robotics segment of the technological, industrial, medical, and other economic sectors. NQROBO utilises modified equal weighting, rebalances quarterly, and reconstitutes in March and September (semi-annually). The index categorises the constitutents within “Enabler”, “Engager”, and “Enhancer”, based on the perceived degree of AI and robotics sector involvement. Companies that develop the building block components for AI and robotics are viewed as Enablers. Examples of components are advanced machinery, self-driving vehicles, and databases for machine learning. The second category, Engager, consists of companies that design, create, integrate, or deliver robotics in the form of products, software, or systems. Companies that provide value-added services within the Artificial Intelligence and robotics ecosystem, which are not core to their products or services, are categorised as Enhancers. The weights between the three categories are 25% in Enablers, 60% in Engagers, and 15% in Enhancers (NASDAQ, 2022). The index consists of 109 companies, whereas well-known technology companies such as Alphabet, Apple, Microsoft, NVIDIA, and Tesla are represented. The index is weighted relatively evenly distributed between the 109 companies with GENTEX holding the highest weight of 2,37%, and OMNICELL INC the lowest weight of 0,3%, per August 2023 (NASDAQ, 2023).

Figure 1: Cumulative excess returns of NQROBO, NASDAQ and S&P 500 over time

Note: Returns relative to risk-free T-Bill rate

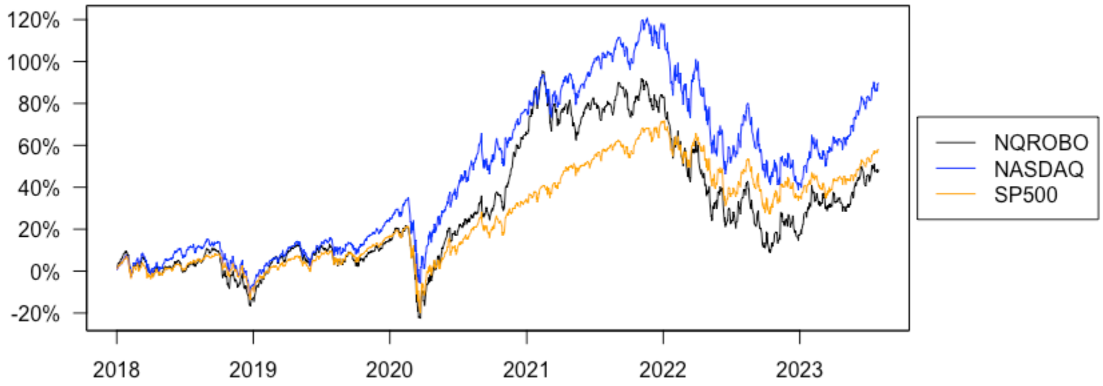


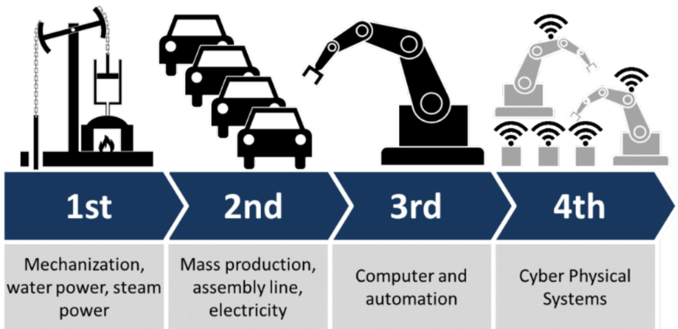
Figure 1 shows that all three indices followed similar pattern from 2018 to 2020. Due to Covid-19 pandemic, the graph had a clear downshift in early 2020, affecting all indices. From April 2020 to 2021, NQROBO outperformed both S&P 500 and NASDAQ before the technology "sell-off" in early 2022. Based on the realized cumulative ex post returns, NQROBO only outperforms the NASDAQ index on one occasion, early 2021. However, it outperformed the S&P 500 for a year and a half from May 2020 to 2022. With this in mind, it will be interesting to observe how different investment strategies will perform in "real time", which will later be analysed in Chapter 8.

There are three main benefits of investing in AI funds. The first is the opportunity to tap into a high-growth market. The second is diversification, enabling your portfolio to be resilient to the potential disruption unleashed by AI. Lastly, there are close to no "pure play" publicly traded AI companies. However, there are also certain risks of investing in AI funds, such as the valuation of AI firms may being influenced by shifts in investors' sentiments, as demonstrated by the decline observed over the past 18 months. During the Covid-19 pandemic, when the interest rate was zero, investors had little to no choice but to invest in growth-related assets because bonds and cash weren't providing any return. In the backlash of the pandemic and the sharply rising interest rates, the valuation of the firms was heavily affected. This results in investors having to differentiate between quality AI firms with long-term growth and firms that rode the AI wave of 2020-2022 (Schmidt & MarketBeat, 2023).

2.2 The Beginning and Evolution of Artificial Intelligence

Throughout history, humanity has witnessed the emergence of three essential industrial revolutions. The first revolution commenced in Britain as early as 1784 with the introduction of the first steam engine. Subsequently, the second revolution unfolded in 1870 with the introduction of electricity, which changed the production methods in a significant way. The third revolution marked the onset of the IT era in 1969. We are currently finding ourselves amid the fourth revolution centered around Artificial Intelligence (Schwab, 2017). This revolution is often characterised by the proliferation of extensive automation, big data, and an interconnected world, all driven by AI technology. The Fourth Industrial Revolution and the development made using AI technology build further upon the foundation laid during the Third Industrial Revolution and can be said to be a fusion of advancements that connects the borderlines between physical, biological, and digital domains (Skilton & Hovsepian, 2017).

Figure 2: The impact of the different Industrial Revolutions (Schultz-Bergin, 2021)



Even though the fourth revolution and AI-based development began relatively recently, the idea of intelligent machines or automatons can be traced back to old Greece when Homer introduced the Automata of the Greek god Hephaestos in the sixth century BCE (McCorduck, 2004). However, only recently have we been able to build and test such machines. The origins of the first newer development can be traced back to Alan Turning, a mathematician and code breaker during World War II, who grappled with whether “machines could exhibit thoughts”. Nevertheless, it was not until 1956 that the term “Artificial Intelligence” was officially coined by John McCarthy. The Oxford English Dictionary defines AI as follows (Lexology, 2017):

“Artificial intelligence is the theory and development of computer systems that can perform tasks normally requiring human intelligence such as visual perceptions, speech recognition, decision-making, and translation between languages.”

It is important to note that there is no clear consensus on the definition of AI. The reasoning for this disagreement stems from the technical complexity of the concept but also because AI is intertwined with a broader philosophical debate regarding what it means to be human (Singer, 2009). This has led to a significant ethical debate regarding the definition of intelligence and the boundaries and applications of AI.

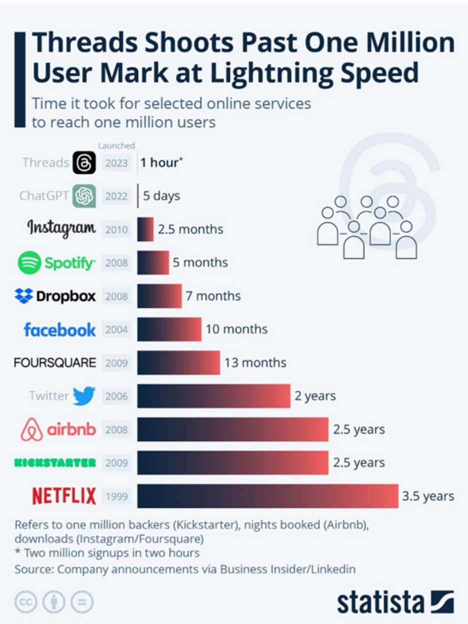
In this thesis, our primary focus is on the application of AI within the organisational sectors. However, it is important to note that the majority of the early-stage AI development was not directed towards utilisation within the organizational sectors. Instead, most of the efforts were dedicated to Bayesian statistics, which makes up the foundation for the machine learning techniques we utilise today. A primary challenge during this period was the limited availability of data power and storage capacity. Despite substantial financial investments put into AI research, the outcome wasn't particularly impressive. Consequently, the interest and funding for AI declined heavily in the 1970s, leading to what is now called the "AI winter". Fortunately, the "AI winter" was short-lived as the field of AI experienced a resurgence, mainly due to advancements in computer technology and renewed financial support (Lutkevich, 2022). Notably, significant investments from Japan, the UK and the US played a pivotal role in the revival. As a result of the heavy investment made, a significant milestone occurred in 1982 when James Simons established the quantitative investment firm "Renaissance Technologies". This marked the initial introduction of AI solutions into the financial sector. The firm later became famous for their innovative financial processing technologies, particularly in pattern recognition (Zuckerman, 2019).

As AI developed through the 1980s, the technology gained traction within fraud detection in the 1990s. In 1993, the Financial Crimes Enforcement Network (FinCEN) was the first significant implementation of AI technology within this sector with a system to detect money laundering, FinCEN Artificial Intelligence System (FIAS). The system reviewed and monitored over 200 000 transactions weekly, and within two years of operating, the system detected 400 money laundering attempts worth 1 billion dollars (Senator et al., 1995).

Due to the growth of computer processing power, more storage, and other technological developments, we have received new possibilities within deep learning, which has been a massive breakthrough in AI. Deep learning has allowed computational models with multiple processing to learn informational data with multiple abstractions. This has brought significant breakthroughs in processing images, videos, audio, and speech (LeCun et al., 2015).

AI is not a sudden, unforeseen technological innovation; instead, it is a continuously evolving technology that has constantly been improved and developed over the years. Nevertheless, no one could foresee ChatGPT's impact on the world. ChatGPT is an AI chatbot built on top of OpenAI's foundational large language model (LLM). This chatbot has set a new benchmark in AI, demonstrating how these machines can comprehend and learn the complexities of human language and interaction. The journey began in June 2018, with the release of the first GPT-1 (Generative pre-trained transformer) model by OpenAI (Marr, 2023). This first iteration consisted of 117 million parameters and set the foundational architecture for ChatGPT we know today. GPT-1 displayed the power of unsupervised learning in language understanding using books as training data to predict words. In February 2019, OpenAI released their new and improved GPT-2, representing a significant upgrade with over 1,5 billion parameters. The model showcased an essential upgrade in text generation capabilities and produced coherent multi-paragraph texts. However, perhaps the most significant milestone in AI language models was the introduction of GPT-3 in June 2020. This model was trained on 175 billion parameters and showcased advanced text-generating capabilities and human-level performance on language tasks. This led to use in various applications, such as writing articles, creating poetry, rephrasing text, and even drafting emails. OpenAI released an early demo of ChatGPT-3.5 on 30 November 2022, and the chatbot quickly went viral on numerous social media. Within five days, the chatbot had attracted over one million users.

Figure 3: Number of days to reach one million users (Buchholz, 2023)



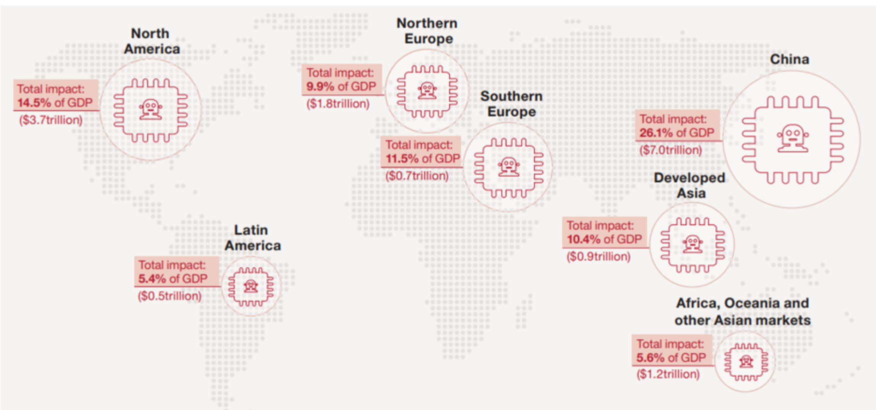
The newest iteration, ChatGPT-4, continues to improve the technology with several enhancements. Introducing the following improvements: better model alignment, reduced likelihood of generating offensive content, better factual accuracy, and real time internet connectivity, which enables real time information to be integrated. Each milestone brings us closer to a future where AI technology is seamlessly integrated into everyday routines, enhancing communication, productivity, and creativity.

The evolution and journey of ChatGPT by OpenAI exemplifies the rapid progression and development of Artificial Intelligence. This groundbreaking model has not only introduced the mainstream public to their first real meeting with AI but also propelled the progress in AI development as one could build further onto already existing AI frameworks. However, the journey is far from over. As we look ahead, we can anticipate ChatGPT and similar AI-powered devices to continue shaping our world.

2.3 Market Size and Growth Opportunities of Artificial Intelligence

The United States were the first country to invest heavily in AI. Between 2000 and 2016, the United States emerged as the leading country in AI technology, hosting and creating 3 033 startups, which constituted 37% of all AI startups globally (Buchanan & Cao, 2018). This accounted for around 72% of the total AI funding worldwide. To put into perspective, from 2012 to 2016, the UK and China invested $850 million and $2,6 billion, respectively, compared to the US, which invested $18,2 billion in this technology. However, in 2017, the US dropped their leading position to China, which began investing heavily in the technology. Below is an illustration showcasing which regions are predicted to capture the most from AI.

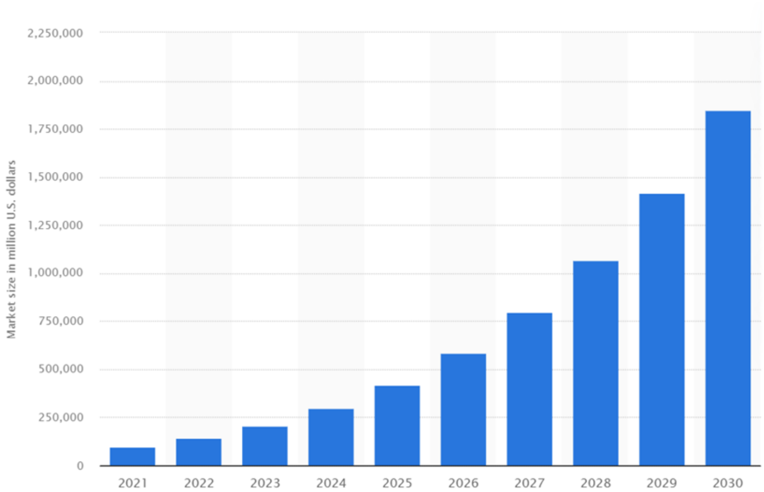
Figure 4: Which region will capture the most GDP from the AI technology in 2030 (PwC, 2016)



As depicted in Figure 4, all regions of the global economy will reap benefits from implementing AI. It is also clear that China and North America is expected to benefit the most prominent economic gains with a GDP improvement of respectively $7 trillion and $3,7 trillion, constituting 70% of the global economic impact (PwC, 2016). It is essential to note that North America is expected to experience the majority of AI faster than China. However, China could ultimately end up seeing a more significant impact on GDP in 2030, despite a slower initial uptake.

In 2022, the global AI market was valued at $136,55 billion, and it is projected to expand at a compounded annual growth rate of 37,3% from 2023 to 2030. This would accumulate to $1847,58 billion by 2030 (Grand View Research, 2023). This growth is mainly due to continuous research and innovation efforts led by tech giants, which drive the adoption of advanced technologies in industry verticals. The AI market covers many industries, including supply chains, analysis, research, and product marketing. These are among some of the fields that will adopt Artificial Intelligence within their business.

Figure 5: Forecasting Artificial Intelligence market size worldwide up to 2030 (PwC, 2016)



As illustrated above, we can see that the market for AI will grow significantly by 2030. According to PwC, industries such as healthcare, financial services and transport & logistics have the biggest potential for positive impact of the integration of AI (PwC, 2016). Historically the adoption of new technology tends to grow exponentially before leveling off, following a logistic pattern (Kucharavy & De Guio, 2015).

The market for AI technology is already quite substantial, however, what lies ahead in the future is a remarkable opportunity for further growth. The potential for expansion and integration in AI is immense, promising several advancements across several industries, introducing more customisable, cost- and time-effective solutions. However, with the logistic pattern described earlier by Kucharavy & De Guio in mind, it will be important for an investor to monitor the development as the adoption rate of the technology may level off.

2.4 Ethical Dilemmas and Possible Challenges With AI

Although AI offers numerous advantages for businesses and consumers, it also creates challenges and dilemmas. Compared to traditional technological products, the ethical dilemmas posed by AI appear to require more urgent consideration due to rapid growth and expanding capabilities. As outlined in the Norwegian government's "National Strategy for Artificial Intelligence", AI technology should be developed with ethical principles, digital security, and data privacy in mind (Regjeringen, 2020). In March 2023, more than 1 000 technology leaders and researchers, including Musk and Wozniak, urged to pause further development of AI systems more powerful than ChatGPT-4 in an open letter (Samuel, 2023). They addressed their concern and how AI systems can present a "profound risk to society and humanity". The reasoning behind the pause is to address concerns such as; "Machines flood our information channels with propaganda", "automate away all the jobs, including fulfilling ones", and "develop nonhuman minds that eventually outnumber, outsmart, obsolete and replace us" (Future Of Life, 2023).

Privacy Concerns

Privacy regards the right to manage personal information (DesJardins, 2014). This right is violated if personal data is collected or used without consent or knowledge. There are numerous forms of which privacy can be violated, for example, information collection and distribution by third parties. From the consumer's perspective, ensuring that data owners are not misusing their personal information is difficult. From the data owners' perspective, preserving the information without violating any informational regulations is challenging. Regarding AI technology, privacy concerns arise due to AI's usage of Big Data (Kaplan & Haenlein, 2019). The utilisation of AI technology relies heavily on Big Data, increasing the amount of personal information collected. Big Data utilises factors such as financials, socioeconomics, demographics, age, gender, and other transactional data to increase the precision of AI. As AI technology advances, the volume of consumer data being collected, utilised, and transmitted is growing at the same rate, presenting challenges for consumers' privacy protection.

Ethical Dilemmas

Multiple ethical dilemmas arise with the development of AI. The concern that most rapidly occurs is unemployment. As the automation of labour has gradually developed since the Third Industrial Revolution, the market forces have delegated and created more complex tasks, resulting in less physical labour and more strategic work-related tasks. However, with the introduction of AI, the workforce is concerned about further development. The question going forward is whether we must (1) regulate the development of AI, (2) let the market forces decide or (3) find other solutions. Another concern that must be addressed is how we will distribute the wealth created by machines amongst each other. This was first discussed at the World Economic Forum (Bossmann, 2016). The global economy currently operates on a compensation system, rewarding employees with wages for their contributions to the economy. With the introduction of AI, wealthy businesses can adopt AI and reduce the number of employees, which again will reduce wage costs and increase their overall profit. This results in a higher profit for fewer people increasing the economic inequality.

AI bias

With AI technology having superior autonomous capabilities and computing power, technology is increasingly utilised to mitigate individual decision-making. Therefore, consumer decisions such as insurance selection, loan application and even movie/series selection are often influenced by AI through recommendations. There is a common misconception that technology is more objective and, therefore, less prone to biases than humans. However, it has become evident that bias is a significant weakness of AI, impacting the quality of decision-making (Du & Xie, 2021). The reason for this bias is that AI-powered products often rely on machine learning, which utilises large training data sets. Biased and unbalanced dataset is the driver of AI bias, often stemming from imbalances related to variables such as education, geography, income, and gender (Du & Xie, 2021).

With the abovementioned ethical dilemmas and challenges, it is crucial for an investor to monitor the development of the technology closely if contemplating whether to invest in AI or not. These are just some of the challenges the technology faces in the near future, which could impact the development, and subsequently, the possibility of generating returns.

2.5 How AI Can Add Organisational Value

Organisations that advantageously implement AI are expected to attain added business value gains in the form of cost reduction, increased revenue, and business efficiency (Alsheibani et al., 2020). Up to 80% of organisations view AI as a strategic opportunity, and almost 85% of organisations see AI as a tool to achieve a competitive advantage (Ransbotham et al, 2017). Despite the growing organisational interest in AI, many firms fail to realise the true value of AI (Fountaine et al., 2019). The use of AI for corporate use can be categorised into automation and augmentation (Enholm et al., 2021). Automation refers to replacing human work, and augmentation refers to enhancing the human decision-making process by providing valuable insight. The implementation of AI is divided into two categories: first-order impact and second-order impact. First-order impact refers to how AI changes the process level of an organisation, while second-order impact refers to how AI affects the operations. In first-order effects, AI would typically enhance process efficiency, which is achieved by reducing human errors, improving productivity, and generating an overall greater precision. The second-order effects are when AI is used as a tool to improve human decision-making, this is achieved by allowing AI to access and analyse vast amounts of data.

2.6 Related Literature

When researching relevant literature for the thesis, we found that the existing literature exploring the AI sector as an investment strategy was relatively scarce. However, we identified two articles that were particularly relevant and significantly contributed to the foundation of our thesis.

Zhang et al. (2023) explores how forecasting different AI-related indices is important for financial market stability. The article argues that NQROBO is superior, compared to other AI and robotics indices. NQROBO reflects the overall stock price change and associated development in the AI industry comprehensively. We have chosen to utilise NQROBO as our index to track the performance of companies engaged in AI & robotics. This is also consistent with other studies that have solved it similarly, for instance, Tiwari et al. (2021) and Huynh et al. (2020). Pesaran & Timmermann (1995) has primarily laid the foundation for our forecasting methodology. The article explores how an open-minded investor who believes that stock returns could be predicted to a great extent, but does not know which model is correct, thus utilises multiple forecasting models that evolve. We will cover the forecasting methodology more comprehensively in Chapter 4.

3. Financial Theories and Performance Measurement

This chapter introduces important and relevant financial theories that make the foundation for the analyses we will conduct later. Simultaneously, we introduce our chosen financial performance measurement, enabling us to evaluate an index's performance.

3.1 Important Financial Theories

In the following section, we will introduce some fundamental financial theories essential for our analysis. This section may be viewed as redundant for experienced readers; however, we have chosen to include it as the theories presented are crucial for our upcoming analyses and discussions.

3.1.1 Efficient Market Hypothesis (EMH)

The Efficient Market Hypothesis (EMH) is an essential concept in financial economics. It states that markets are regarded as efficient when the price of an asset reflects all relevant information about said asset. "Efficiency", in this context, implies that investors cannot earn above-average returns without undertaking above-average risk. Fama divided the EMH into three degrees based on the informational level present in the market: weak form, semi-strong form, and strong form (Fama, 1965). If the informational level present in the market is weak, today's share price only reflects historical data and prices, making it impossible to receive any additional information about tomorrow's price. A semi strong market is considered to imply that all publicly available information, both historical and current, is reflected in the price. Lastly, if the market is in strong form, the EMH argues that all information, both public and private, inside information, is fully reflected in the prices. For our discussion, we adopt the semi-strong form of the EMH.

Because of this, investors cannot exploit mispricing to generate excess returns. Therefore, a investor who believes in the EMH would only hold the market portfolio as there is no excess return to make because all assets are correctly priced. Nevertheless, the EMH is often contested as opposing investors arguing that investors have consistently outperformed the market over a longer period, this would be impossible if EMH holds true.

3.1.2 The Capital Asset Pricing Model

The Capital Asset Pricing Model (CAPM) is a widely used financial model that evaluates the expected return on an investment relative to risk and overall market return (Sharpe, 1964). The model was first introduced by Treynor (1961), Sharpe (1964), Lintner (1965) and Mossin (1966). There are three key components of CAPM: risk-free rate, market risk premium and Beta. Whereas the risk-free rate is the theoretical return an investor can earn without risk of financial loss, market risk premium represents the additional return expected for taking on higher risk and investing in the market. Beta is a measurement of an investment's risk that cannot be eliminated through diversification. CAPM operates with the assumption that expected returns on all assets are linearly related to their systematic risk. Thus, the market only prices systematic risk. CAPM is expressed in the equation below:

Equation 1

$$E[r_i] = r_f + \beta_i(r_m - r_f)$$

Where, $E[r_i]$ is the expected return for an investment i, r_f is the risk-free rate, β_i is the Beta for investment i and $(r_m - r_f)$ is the market risk premium.

3.1.3 Beta

As mentioned above, Beta is a metric expressing the systematic risk in CAPM. It assesses the relationship between the return of an investment and those of the market and measures the volatility of an investment in relation to the overall market. Beta as a measurement helps investors assess an asset's systematic risk, which cannot be computed away through diversification and provides insights into how an investment is likely to perform. Beta can be computed through several different methods. It can be computed by regression and historical data, amongst others. Finding Beta using regression analysis involves using statistical techniques to compute the slope of the best-fit line between the return of an investment and the market. Beta computed through historical data can be expressed through the equation below:

Equation 2

$$\beta_i = \frac{Cov(r_i, r_m)}{\sigma^2(r_m)}$$

Where, β_i is the Beta of investment i, $Cov(r_i, r_m)$ is the covariance between the return of investment i and the market return and σ^2 is the variance market return

3.1.4 Alpha

Alpha is a financial metric used to describe an investment's excess return generated relative to a set benchmark or market. It is important to note the difference between active return and Alpha. Alpha refers to the risk-adjusted contribution of active management, and it is represented as the intercept in a regression analysis comparing active return to benchmark or other risk factors (Chen, 2023). Meanwhile, active return represents the difference between return and benchmark measuring the contribution of active management.

3.1.5 Fama French Factor Models

As mentioned before, the Capital Asset Pricing Model (CAPM) is a widely utilised framework in asset pricing and only prices systematic risk and not idiosyncratic risk. To enhance the model, Fama & French (1993) introduced the Three-Factor model. The factor model builds further onto the time series regression approach of Black, Jensen and Scholes (1972) and aims to explain the variation of returns of stocks and portfolios. By including factors such as market risk, size, and value, Fama & French attempts to capture various systematic risk factors coherently in the market. Consequently, any Alpha, as indicated by the intercept, can be attributed to idiosyncratic risk not accounted for by the model, implying that the Alpha (α) can be interpreted as index-specific risk. The regression of the Three-Factor model can be written as:

Equation 3

$$R_p^{\alpha} = \alpha + \beta_{MKT}MKT_t + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \varepsilon_t$$

The market return (**MKT**) factor represents the excess return of the entire stock market over the risk-free rate. Similar to the CAPM, it accounts for systematic market risk.

The size factor (**SMB**), small minus big, reflects the difference between returns on a diversified portfolio of small firms compared to big firms. This is done because small firms with a low market capitalisation tend to outperform large firms with a high market capitalisation.

The value factor (**HML**), high minus low, showcases the difference between a well-diversified portfolio of stocks with a high book-to-market ratio (value stocks) compared to a portfolio consisting of stocks with a low-book-to-market ratio (growth stocks). The results from the Fama & French study suggested that value stocks tend to outperform growth stocks. One of the reasons behind this is that firms with a high book-to-market ratio are significantly more exposed to financial distress and require more compensation for taking on risk.

Later Fama & French introduced two additional factors, profitability, and investment, in the Five-Factor model. One assumes that when applying the Five-Factor model, it captures the expected return of the factors. The regression for the Five-Factor model is denoted by Equation 4.

Equation 4

$$R_p^{\alpha} = \alpha + \beta_{MKT} MKT_t + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \beta_{RMW} RMW_t + \beta_{CMA} CMA_t + \varepsilon_t$$

The profitability factor (**RMW**), robust minus weak, is found by taking the average return of a robust firm minus the average return of a weak firm. A robust firm is categorised as a firm with high profitability and a weak firm with low profitability.

The investment factor (**CMA**), conservative minus aggressive, examines the historical excess return difference between firms with conservative investment practices (low asset growth) versus firms with aggressive investment practices (high asset growth).

3.2 Financial Performance Measurement

3.2.1 Sharpe Ratio

Sharpe Ratio (SR) is a reward-to-volatility measurement. The ratio quantifies how much excess return an investment generates for each unit of risk. It is commonly used when evaluating the performance of an asset and exhibits the average risk-adjusted excess return compared to a risk-free rate (Sharpe, 1966). A higher SR is generally preferred as it indicates better risk-adjusted performance. SR can be found through the following equation:

Equation 5

$$Sharpe\ Ratio = \frac{E[r_i] - r_f}{\sigma_i}$$

Where, $E[r_i]$ is the expected return, r_f is the risk-free rate and σ_i is the standard deviation of portfolio.

3.3 Transaction Cost

Transaction cost refers to anything that results in a trade having an additional cost. The EMH's assumptions overlook this cost, which is an unrealistic premise in the real world (Fama, 1965). In reality, we know that transaction costs exist in the market and vary among market participants. The cost would typically vary due to the market participant's size, varying time factors, location, and different funds. A large investor, such as the Government Pension Fund Global (GPFG), would have a lower transaction cost than a small investor. These are market frictions one must consider when operating within the market. Transaction cost would influence the excess returns negatively as it "eats" of the profit. In our case, the fictional investor would have to pay transaction cost both when buying and selling indices. As the transaction costs vary between different funds (NQROBO, NASDAQ, and S&P 500), it is difficult to choose one rate that would be correct for all funds. Because of the difficulty of selecting one appropriate rate and given the increased complexity it would entail, we have decided to disregard transaction costs in our analysis.

3.4 Portfolio Management

In the following section, we discuss the key differences that distinguish passive and active management. The differences will be important when constructing our portfolios later in the thesis.

3.4.1 Passive Management

A passive investor would typically have a "Hands-off" approach and try to replicate the market. According to Sharpe (1991), a passive investor can be defined as someone who believes in the EMH and, therefore holds the same securities and weights them accordingly to replicate the market. If the markets are efficient, passive investors would hold the market. This is because all securities are priced correctly, making it impossible to outperform the market consistently. The goal of a passive investor is to achieve returns equal to the market, and passive investors are characterised by low transaction frequencies and, consequently, low transaction costs.

3.4.2 Active Management

An active investor tries to outperform the market. This is achieved by trading based on information trying to buy undervalued securities and sell overvalued securities. Thus, an active investor does not believe that the market is 100% efficient in terms of EMH and therefore, tries to invest in securities that differ from the market (Sharpe, 1991). To outperform the market, an active investor believes in timing and tries to identify under- and overvalued securities. The goal of an active investor is, therefore, to outperform the market. Consequently, active investors tend to take on more risk and trade more frequently compared to passive investors implying a higher degree of transaction costs compared to a passive investor.

4. Forecasting Methodology

This chapter serves to explain the statistical models and methods applied within the thesis to answer the research question. Initially, we present an autoregressive model before building further onto this model and introduce an ARIMA model. These models are used when forecasting future values of the various indices. We will also describe our pseudo-out-of-sample forecasting and present a forecast performance measurement. Given the scope of this thesis, we exclude the analysis of stochastic volatility of returns, that a GARCH-type model could capture, as this would be beyond our work.

4.1 Autoregressive Models

An unvariate autoregressive model (AR) is a method utilised when forecasting future values of a chosen variable, and the predicted values are found by using the past values of the same variable (Hyndman & Athanasopoulos, 2018). As the term "autoregressive" indicates, the model regresses the dependent variable linearly on the lagged variables. The lagged variables are used as predictors, where the number of lags determines the model's order. For instance, an autoregressive model with a single lag can be referred to as a first-order AR model, also called AR(1). While an AR model with two lags can be referred to as a second-order AR model. The AR model is a simple statistical model because it only relies on the past values to explain future values. Despite its basic structure, the model captures random fluctuations through its error term, also known as "white noise". The white noise represents variations that the model's past values cannot explain.

The autoregressive process of order p can be expressed as follows:

Equation 6

$$y_t = c + \sum_{i=i}^{p} \phi_i y_{t-i} + \varepsilon_t$$

Where, y_t is the dependent variable, c is a constant, ϕ is the coefficient for lagged value i and ε_t is the error term.

One of the most basic yet most insightful forecasting models is the process of AR(1). In the model, the dependent variable is exclusively regressed on itself one period ago. The AR(1) model is crucial when conducting our forecasting, thus making it crucial when trying to answer our research question. The autoregressive process with a lag order of one can be expressed as follows, assuming $\phi_1 < 1$.

Equation 7

$$y_t = c + \phi_1 y_{t-1} + \varepsilon_t$$

4.2 Random Walk

A random walk is a stochastic process where each data point or step is determined by adding a random error or noise term to the previous data point. The error term is often characterised as white noise. The key feature of a random walk is that future values are unpredictable, and each step is independent of the previous ones (Smith, 2023). Random walk is not stationary, this is because the mean of a random walk is constant, however, the variance is not. The random walk theory states that, given a time series random movement, future values would be today's value plus the sum of residuals from now until future date T. Because future errors are not known today, the best predictor of future value is today's value. The process of random walk can be denoted in the following equation:

Equation 8

$$y_t = y_{t-1} + \varepsilon_t$$

4.3 Autoregressive Integrated Moving Average

Autoregressive Integrated Moving Average, known as ARIMA, models are often used in pseudo-out-of-sample forecasting to evaluate their ability to make accurate predictions on unseen data. It is a time series forecasting model that predicts future values based on past values and uses lagged moving averages to "smooth" the time series data. By conducting a pseudo-out-of-sample forecast, one can assess how well the ARIMA model performs when estimating beyond the estimation period. This helps to determine the model's accuracy and predictive properties (Shumway & Stoffer, 2017).

This model is a combination of autoregressive models and moving average models. In an autoregressive model, we forecast the variable using a linear combination of past values of the variable. On the other hand, a moving average model uses past forecast errors in a regression-like model (Hyndman & Athanasopoulos, 2018).

The equation for a general ARIMA model can be denoted as:

Equation 9

$$y'_t = \Delta^d y_t$$

Equation 10

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Where, y'_t is the differenced time series with d number of differentiations needed to ensure stationarity. The right-hand side of Equation 10 consists of both lagged values and lagged errors of y_t. The model can be written as $ARIMA(p,d,q)$, where p is the order of autoregressive parts, d is the degree of first differencing involved needed to ensure stationarity, and q is the order of the moving average part.

In our forecasting, we have estimated the most appropriate ARIMA model for each estimation window, hereby called ARIMA best fit. This means that the order of p, d, and q is adjusted for the individual time series. To obtain the ARIMA best fit model, we have used the Box-Jenkins methodology. This approach is divided into three phases: (1) test for stationarity in time series and identification of the model, (2) estimation of chosen model, and (3) check for normality of residuals and further improvements.

To support step 1 and 2 in our model selection procedure, we have used the auto.arima() function in the R package *forecast*. The function follows a systematic process to identify the most suitable combination of p, d, and q for each time series (Hyndman & Khandakar, 2008). The inclusion of this function limits the consequences of human error. For further details on the model selection and residuals results, see Appendix Chapter 4.

4.4 Pseudo-Out-of-Sample Forecasting

Pseudo-out-of-sample is a term commonly used within the field of financial modelling, it is often used when evaluating the performance of predictive models. Such predictive models are often utilse when forecasting financial markets and economic variables. The technique is often used as experience has demonstrated that forecasting a good-in-sample fit does not necessarily imply a good-out-of-sample performance. To mitigate this issue, the method often partitions the data into two parts: one is used for model estimation, also called training data, and the other is used for model evaluation.

4.4.1 Forecasting Description

We have chosen to analyse three different forecasting horizons: one day-, one week- and one month-ahead. Our forecasting period begins on 2 January 2019, and continues to August 2023. The one day ahead forecasts end on 1 August 2023, the one week ahead forecasts end on 7 August 2023, while the one month-ahead forecasts end on 29 August 2023.

The forecasting of excess return is based on data of the past 250 observations (days). For example, is the predicted daily excess return on 2 January 2019 based on daily data from 2 January 2018 to 31 December 2018. To forecast the weekly excess return, we forecast the excess returns five days ahead using dynamic forecasting. This means that we do not use information after 31 December 2018 to forecast the returns from 2 January 2019 to 8 January 2019. We use the same method to forecast the monthly excess return for the indices. After forecasting the daily, weekly, and monthly excess return at time t, we move the estimation window one day ahead and repeat the procedure. In total, this procedure was repeated 1152 times. An overview of the forecasting models used for the different indices in this thesis is given in Table 1.

Table 1: The different forecasting models utilised for the indices

Note: Nine different forecasting models were used to forecast NQROBO and NASDAQ and three different forecasting models were used to forecast S&P 500

	Index		
Forecasting Model	**NQROBO**	**NASDAQ**	**S&P 500**
Random Walk	X	X	X
AR(1)	X	X	X
ARIMA best fit	X	X	X
CAPM w/Random Walk	X	X	
CAPM w/ARIMA	X	X	
FF3 w/Random Walk	X	X	
FF3 w/ARIMA	X	X	
FF5 w/Random Walk	X	X	
FF5 w/ARIMA	X	X	

The three forecasting models that we have used for all indices was the Random Walk, AR(1) and ARIMA best fit. As mentioned in 4.3, we have estimated an individual ARIMA model for each estimation window. For the CAPM-model forecasting, we have used a linear regression model. The model takes the realized excess return of the relevant index as the dependent variable and the realized excess market return as the independent variable. As mentioned earlier, we consider the S&P 500 index as the market. We have included two versions of the CAPM model in our forecasting. The first model uses the Random Walk model to predict the market return. The second model employs an ARIMA best fit model to forecast future market returns.

We have adopted a similar methodology to the Fama French Three-Factor and Five-Factor model. To obtain the factors, we have downloaded them from their website, this is commented further in Chapter 6. The first Three-Factor and Five-Factor model utilize the market return and factor premiums from the previous period (t-1) as the input for forecasting. In our second Three-Factor and Five-Factor model, we have employed an ARIMA best fit model to predict the market return and used the same factor premiums as in the first model.

4.5 Forecasting Evaluation

To evaluate our forecasting models, we compute the root mean squared error (RMSE) for every forecast model and compare it to each other.

The RMSE measures the average difference of the errors between forecasted and actual values. The formula for RMSE can be written as:

Equation 11

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Where, y_i is the actual value, and $\hat{y_i}$ is the forecasted value. A lower RMSE indicates a better-fitting model, as the forecasted values are closer to the actual values.

5. Construction of The Synthetic Portfolios

This chapter introduces our constructed six portfolios, observing them over the chosen period. Each portfolio is constructed to replicate a possible real-life investment strategy. By this, we mean, to mimic an investor who wants to evaluate alternative investment strategies, not using future values.

When constructing our synthetic portfolios, we have used six different weighting approaches. The portfolios based on Sharpe Ratio and RMSE measurements is dynamically weighted, we have defined these as “dynamic portfolios”. Each portfolio is weighted differently resulting in a comprehensive understanding of the performance of the different portfolios.

Table 2: Overview of synthetic portfolios

Name	Description
Portfolio 1	Holding the market
Portfolio 2	Equally weighted portfolio
Portfolio 3	Sharpe Ratio performance based portfolio
Portfolio 4	RMSE performance based portfolio
Portfolio 5	Hybrid portfolio of equal weights and Sharpe Ratio performance
Portfolio 6	Hybrid portfolio of equal weights and RMSE performance

Portfolio 1 “Holding the market” showcases how a passive investor who has invested 100% into the market would perform. The other five portfolios are constructed as a combination of the three indices: NQROBO, NASDAQ and S&P 500. Our analysis spans from 2019 to August 2023, with daily updates for our forecasting. For our fictional investor, we have chosen a start investment of $100 000.

The investments are made daily, with holdings lasting for a day, a week, and a month before rebalancing. When an investor with a weekly perspective invests, will they invest an amount every trading day (Monday to Friday), and the holdings are maintained for a week before they rebalance and reinvest. When reinvesting, the weights invested into the three indices will differ amongst the six portfolios. When investing the first time, the amount will be divided equally on the 5 trading days. Below is a visual illustration of the weekly investment strategy.

Table 3: Visual illustration of a weekly investment strategy

Invest
Holding

	Week 1							Week 2				
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday
Investment 1	**Invest $20 000**							**Reinvest**				
Investment 2		**Invest $20 000**							**Reinvest**			
Investment 3			**Invest $20 000**							**Reinvest**		
Investment 4				**Invest $20 000**							**Reinvest**	
Investment 5					**Invest $20 000**							**Reinvest**

Likewise, an investor with a monthly perspective will divide the $100 000 on the 21 first trading days. The investor will further invest for every available trading day, holding the investment for a month before reinvesting. The investment weights of each index are denoted by λ (lambda), and the approach used when calculating λ differs between all five portfolios. For the simplicity of our thesis, we have chosen to limit λ to be between 0% and 100%, not enabling any form for shorting.

5.1 Portfolio 1 – Holding The Market

Our first synthetic portfolio is constructed to showcase how a passive investor who only holds the market would perform over the chosen period. The investor would invest $100 000 into the S&P 500 on day one and hold this to the end of our period.

$$\lambda_{S\&p\ 500} = 1$$

5.2 Portfolio 2 – Equally Weighted Portfolio

The second synthetic portfolio is also constructed as a benchmark portfolio, displaying how an equally weighted portfolio would perform without any rebalancing. It represents a simple and common investment strategy an active investor may consider. The portfolio is weighted as follows:

$$\lambda_{NQROBO} = \frac{1}{3}$$

$$\lambda_{NASDAQ} = \frac{1}{3}$$

$$\lambda_{S\&p\ 500} = \frac{1}{3}$$

Portfolios 1 and 2 will serve as a baseline comparison between the different portfolios when evaluating the performance of other more complex portfolios and ultimately help determine whether the more advanced portfolios add value.

5.3 Portfolio 3 – Sharpe Ratio Performance-Based Portfolio

The third portfolio is performance-based on the measurement of the Sharpe Ratio. The portfolio weight is determined through the predicted Sharpe Ratio of the respective indices. As addressed in the financial theory chapter, the Sharpe Ratio is a reward-to-volatility measurement where the ratio quantifies how much excess return an investment generates for each unit of risk. To calculate exact weights, we divide the relevant Sharpe Ratio on the sum of all Sharpe Ratios. Below is the equation used to calculate the weights.

Equation 12

$$\lambda_{Index}^{SR} = \frac{SR_{Index}}{\sum SR}$$

If the Sharpe Ratio is negative for all three indices, Portfolio 3 would weight 100% in the market. However, in certain scenarios, two indices have a positive Sharpe Ratio, and the last generates a negative Sharpe Ratio. This is solved by only investing in the two indices with a positive Sharpe Ratio and ignoring the last negative index. When only one index generates a positive Sharpe Ratio, the optimal strategy is to invest 100% into the positive index. Despite the Sharpe Ratio being a financial measurement utilising historical data, we have used this measurement when deciding future weights. The thought process behind this is that historical data has a certain "stickiness" to it, and yesterday's data impacts the data tomorrow. When simulating an investor making "real time" decisions using Sharpe Ratio to score alternative investments strategies they will allocate weights depending on SR at the time of investment up to the forecasting horizon and not beyond. If using future value, it will likely lead to overfitting and misleading scores.

5.4 Portfolio 4 – RMSE Performance-Based Portfolio

The fourth portfolio is a performance-based portfolio denoted by the RMSE. As presented in Chapter 4.5, RMSE is a forecasting evaluation measurement found by calculating the difference between the predicted value and the actual observed value. Consequently, this means that the lower values are preferred, indicating a better prediction. Because lower values indicate better fit, we utilise inverse RMSE when determining the weights of the different indices. Below is the equation for calculating the relevant weights if the expected return is positive.

Equation 13

$$\lambda_{Index}^{RMSE} = \frac{RMSE_{Index}^{-1}}{\sum RMSE^{-1}}$$

However, as RMSE is only a measurement displaying how accurate the forecasting is, this alone would not suffice to determine whether or not it is profitable to invest. We have solved this by only investing if the expected return is positive. If the expected return for all three indices is negative, the investor would hold the market, if two indices have a positive expected return and the last has a negative expected return, the investor would only invest in the two positive indices. This strategy may be conceived as conservative, however, in a typical bear market an investor would remain calm, holding the market rather than being eaten up by transaction costs (Pesaran & Timmermann, 1995). An alternative investment strategy could be shorting, however as we do not enable any form for shorting this will be disregarded. In Chapter 5 of the Appendix, Figure 18 illustrates the steps involved in the investment process of Portfolio 4.

5.5 Portfolio 5 – Hybrid Portfolio of Equal Weights and Sharpe Ratio Performance

Our fifth portfolio is a hybrid portfolio between the performance measurement Sharpe Ratio and equal weights. The hybrid portfolio is constructed so that 50% of the portfolio has fixed weights distributed equally to the three different indexes. As a result of these fixed weights, each asset class is guaranteed an investment weight of 16,67%. The remaining 50% is distributed through the Sharpe Ratio measurement method described in Equation 12. We have chosen to split the hybrid portfolio into 50/50 between equal and dynamic weighting as this is the most balanced. The formula for calculating weights is presented below.

Equation 14

$$\lambda_{Index}^{SR\ Hybrid} = \frac{SR_{Index}}{\sum SR} * 0{,}5 + \frac{1}{3} * 0{,}5$$

It would be interesting to see how the portfolio would perform with different weights, such as 30/70. We have done this as a part of our robustness analysis in Chapter 9.

5.6 Portfolio 6 – Hybrid Portfolio of Equal Weights and RMSE Performance

Our last portfolio is similar to the fifth, a hybrid portfolio between equal weights and the performance of RMSE. The construction is also similar, with the only difference being that the performance measurement is the RMSE. This includes utilizing inverse RMSE as lower values are preferred. We operate with the same assumptions from 5.4, only investing if the expected return is positive. The equation for calculating exact weights is presented below and is constructed as an add-on from Equation 13, including 50% fixed weights distributed equally on the three indices.

Equation 15

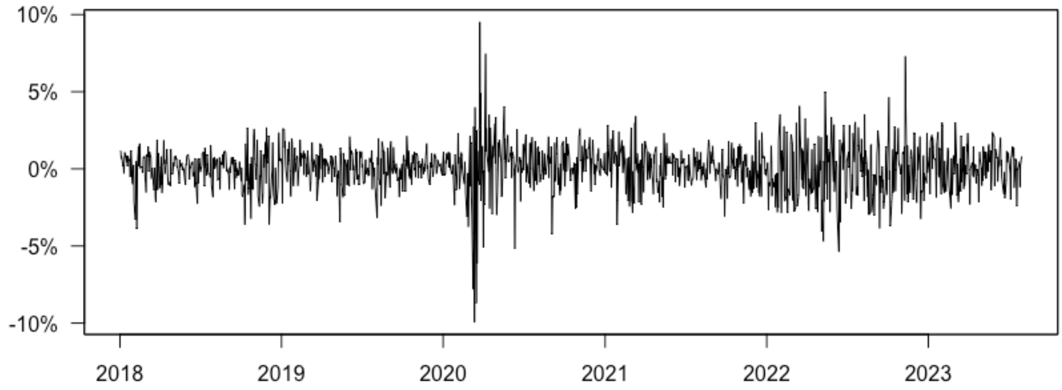
$$\lambda_{Index}^{RMSE\ Hybrid} = \frac{RMSE_{Index}^{-1}}{\sum RMSE^{-1}} * 0{,}5 + \frac{1}{3} * 0{,}5$$

6. Data

We have downloaded daily return data of the NQROBO index. The return data consists of 1423 daily observations from January 2018 to August 2023. Figure 6 presents the distribution of excess returns over the whole period.

Figure 6: Excess returns NQROBO

Note: Returns relative to risk-free T-Bill rate

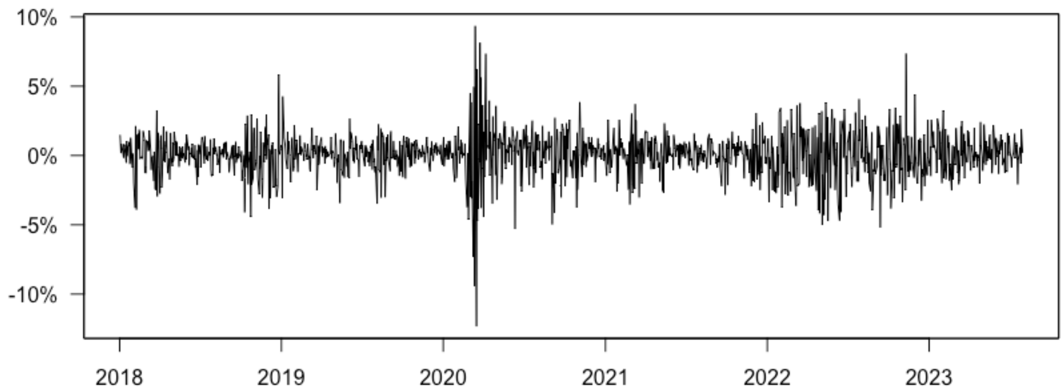


NASDAQ Composite Data

We have chosen to utilise the Nasdaq Composite index as one of the indices we use to compare the performance of NQROBO. NASDAQ measures all domestic (US-based) and international common-type stocks listed on the Nasdaq stock market. NASDAQ consists of 3490 securities, including well-known firms such as Apple, Microsoft and Tesla. It is known for being heavily weighted towards the technology sector, with a 55% exposure to technological industries (NASDAQ, 2023).

Figure 7: Excess returns NASDAQ COMP

Note: Returns relative to risk-free T-Bill rate

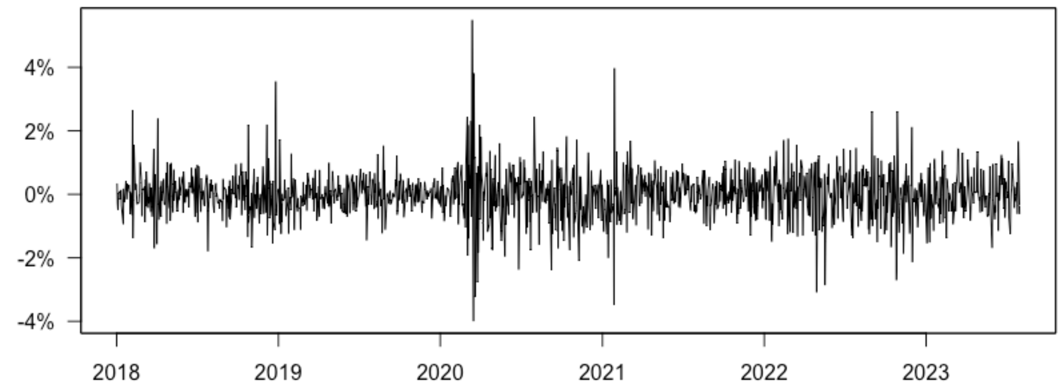


When retrieving the NASDAQ, we have chosen to regress NASDAQ on NQROBO in an attempt to “strip” away the AI part that is included in NASDAQ, resulting in “NASDAQ

Corrected". We have done this to ensure that AI isn't weighted twice in our analyses. The movements are quite similar when observing the differences between the original NASDAQ and NASDAQ Corrected in Figure 7 and 8. However, NASDAQ Corrected has significantly lower volatility, with excess returns fluctuating between 4% and –4% compared to NASDAQ fluctuating between 10% and –10%.

Figure 8: Excess returns NASDAQ Corrected

Note: Returns relative to risk-free T-Bill rate

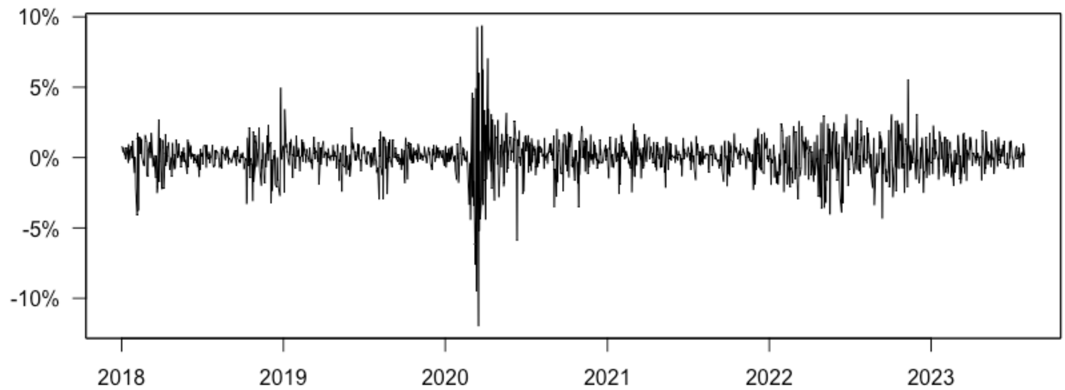


S&P 500 Data

As mentioned earlier, we have chosen the S&P 500 index as the market in our analysis. The data consists of daily observations from January 2018 to August 2023, displayed in Figure 9. The index measures the performance of 500 leading companies and is widely considered to be a proxy of the U.S. equity market, covering approximately 80% of available market capitalisation (S&P 500, 2023). The index utilises a float-adjusted market cap when weighting and is rebalancing quarterly. Unlike NASDAQ, the S&P 500 exhibits a more evenly spread exposure across various sectors. This is why we have chosen to utilise the S&P 500 as the market in our analyses.

Figure 9: Excess returns S&P 500

Note: Returns relative to risk-free T-Bill rate



The Fama French Factor Data

For Fama & French's Three-Factor model, we have downloaded daily factor returns from Kenneth R. French's Data Library (Kenneth R. French, 2023). The risk-free rate factor is the simple daily rate that, over the number of trading days in the month, compounds to a 1-month T-Bill rate.

Data Merging

When calculating excess return, we merged the retrieved datasets containing NQROBO's, NASDAQ's and S&P 500's daily returns. When the data was retrieved, it was first collected and sorted in Excel. Observing the datasets, we found discrepancies in the number of observations between the indices. To be able to evaluate the indices against each other, we had to match the observed returns. When there was a mismatch of observations, we excluded all observations for that particular day. NQROBO had 28 more observations than NASDAQ and S&P 500. There are several potential reasons for such a disparity, such as index rebalancing, data providers and calculation methodology.

Specific Events

With our observations beginning in 2018 and ending in August 2023, an important event has affected the economic landscape greatly and, consequently, our observation data. The Covid-19 pandemic, which began in March 2020, greatly influenced customer demand, policy rates, and investment opportunities. As global policy rates approached the zero lower bound, there was a notable increase in demand and investments. Subsequently, in the backlash of the pandemic, policy rates have been increased historically fast in an attempt to slow down inflation. This has impacted the financial landscape significantly and resulted in a financial cooldown. In our analysis, we have not treated this event any differently. However, it is essential to recognise the importance of this event when interpreting the results.

7. An Evaluation of NQROBO: Index Performance

In the following chapter, we will present a historical analysis of the NQROBO index, enabling us to answer the first part of the thesis. We utilise well-known financial performance measurements, such as excess return, standard deviation, and Sharpe Ratio, to gain valuable insights into the index's performance. The purpose of doing this is to better understand how the index has performed over the years. Enabling us to evaluate the index's capacity to deliver risk-adjusted returns over a specific period. We have decided to do multiple regressions for four different sample periods to understand the market fluctuations. The four periods are 2018-2020, 2020-2022, 2022-August 2023 and the whole period 2018-August 2023.

7.1 Risk and Return Historical Performance

This section aims to provide a comprehensive summary of the NQROBO index's historical performance. We have computed weekly and monthly arithmetic excess return, standard deviation and Sharpe Ratio for NQROBO, NASDAQ Corrected and S&P 500. The table below presents the key performance measurements for the abovementioned factors.

Table 4: Weekly and monthly excess returns for NQROBO, Nasdaq Corrected and S&P 500 over four different sample periods

Note: Returns relative to risk-free T-Bill rate

Weekly returns

Sample 1 (2018-2020) Weekly returns

	NQROBO	NASDAQ Corrected	S&P 500
Mean Ex.Weekly returns	0,13 %	-0,08 %	0,14 %
S.D. of return	2,54 %	0,97 %	2,08 %
Sharpe ratio	0,05	-0,08	0,07

Sample 2 (2020-2022) Weekly returns

	NQROBO	NASDAQ Corrected	S&P 500
Mean Ex.Weekly returns	0,53 %	0,01 %	0,42 %
S.D. of return	4,09 %	1,44 %	3,31 %
Sharpe ratio	0,13	0,01	0,13

Sample 3 (2022-2023) Weekly returns

	NQROBO	NASDAQ Corrected	S&P 500
Mean Ex.Weekly returns	-0,26 %	-0,02 %	-0,08 %
S.D. of return	3,85 %	1,51 %	2,73 %
Sharpe ratio	-0,07	-0,02	-0,03

Sample 4 (2018-2023) Weekly returns

	NQROBO	NASDAQ Corrected	S&P 500
Mean Ex.Weekly returns	0,16 %	-0,03 %	0,17 %
S.D. of return	3,54 %	1,31 %	2,75 %
Sharpe ratio	0,05	-0,02	0,06

Monthly returns

Sample 1 (2018-2020) Monthly returns

	NQROBO	NASDAQ Corrected	S&P 500
Mean Ex.Monthly returns	0,73 %	-0,25 %	0,71 %
S.D. of return	6,02 %	1,94 %	4,23 %
Sharpe ratio	0,12	-0,13	0,17

Sample 2 (2020-2022) Monthly returns

	NQROBO	NASDAQ Corrected	S&P 500
Mean Ex.Monthly returns	2,21 %	0,09 %	1,77 %
S.D. of return	7,30 %	3,10 %	5,66 %
Sharpe ratio	0,30	0,03	0,31

Sample 3 (2022-2023) Monthly returns

	NQROBO	NASDAQ Corrected	S&P 500
Mean Ex.Monthly returns	-1,16 %	-0,12 %	-0,34 %
S.D. of return	7,71 %	2,48 %	5,70 %
Sharpe ratio	-0,15	-0,05	-0,06

Sample 4 (2018-2023) Monthly returns

	NQROBO	NASDAQ Corrected	S&P 500
Mean Ex.Monthly returns	0,70 %	-0,09 %	0,77 %
S.D. of return	7,03 %	2,52 %	5,20 %
Sharpe ratio	0,10	-0,04	0,15

We have found that NQROBO has a weekly excess return of 0,13%, 0,53% and –0,26% and a monthly excess return of 0,73%, 2,21% and –1,16% for the respective three sub-periods. For the whole duration of the dataset, NQROBO has a weekly excess return of 0,16% and a monthly excess return of 0,70%. Compared to NASDAQ Corrected, which is NASDAQ "stripped" away of AI, NQROBO does outperform the index at Sample 1, 2 and 4 for both weekly and monthly excess returns. In Sample 3, we see that NQROBO has a greater negative excess return than NASDAQ Corrected. From the table above, we can observe that the S&P 500 outperforms NQROBO in terms of excess return for all samples except Sample 2 weekly and monthly and Sample 1 monthly.

Table 4 shows that NQROBO, NASDAQ Corrected, and the market performed best during Sample 2 (2020-2022), having high returns and a relatively low standard deviation. This is also reflected in the Sharpe Ratio being the highest during this sample. In Sample 3 (2022-2023), we can see that all indices have negative weekly and monthly returns, resulting in a negative Sharpe Ratio.

In this section, we have provided an overview of the performance of NQROBO compared to NASDAQ Corrected and S&P 500. The results are in line with what we presented in Figure 1, namely, a notable increase in weekly and monthly returns from 2020-2022 before the technology “sell-off" early 2022. It is also worth mentioning how NQROBO was outperformed by S&P 500 for both weekly and monthly returns and the Sharpe Ratio when looking at the whole period.

7.2 Evaluating Active Returns

This section further builds on the return and risk performance measurements presented in section 7.1. To better understand the importance of active returns and how these are affected by different risk factors. By active returns we mean, the difference between S&P 500 and the actual return of NQROBO. We first estimate and evaluate NQROBO’s Alpha before using the Fama French Three-Factor and Five-Factor model to identify and understand the impact of risk factors within the financial markets.

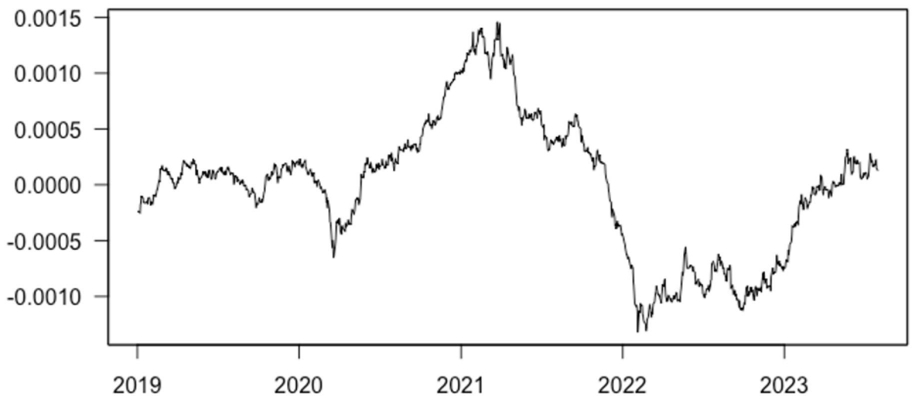
7.2.1 Alpha and Beta Estimation

We have calculated an estimation for Alpha and Beta by regressing the actual excess return of NQROBO on the market's actual excess return. The estimated intercept Alpha represents the average contribution of active returns after adjusting for risk.

In Figure 10 and 11, we have illustrated the development of NQROBO’s Alpha and Beta over time. The purpose of the illustrations is to showcase better how the Alpha and Beta of NQROBO have developed over the chosen period.

Figure 10: NQROBO’s Alpha over time

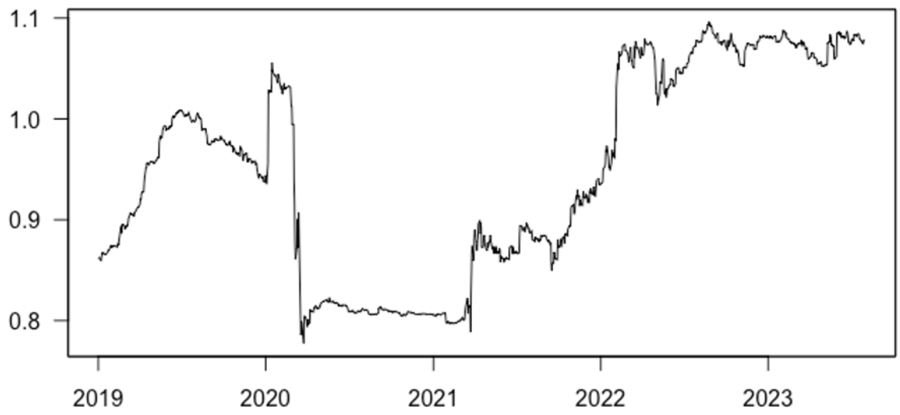
Note: Regressed NQROBO’s excess return over market’s (S&P 500) excess return



When observing Alpha over time, it begins slightly below zero. The fluctuations in 2019, resulting in the Alpha being both negative and positive, suggest that NQROBO both over- and underperformed compared to the market. Between early 2020 and early 2021 active return was positive and peaked close to 0,0015, indicating significant outperformance relative to market. In 2022, the Alpha plummeted to an all-time low and was underperforming compared to the market. There could be multiple reasons for this development. However, this is most likely due to a combination of increasing interest rates, uncertainties due to the Ukraine invasion and the technology "sell-off" in 2022. However, the latest movements of 2023 suggest that the Alpha is recovering and moving towards being positive, indicating an outperformance relative to the market.

Figure 11: NQROBO's Beta over time

Note: Regressed NQROBO's excess return over market's (S&P 500) excess return



Observing NQROBO's Beta over time, it fluctuates around 1, suggesting a volatility similar to the market. We see that it begins around 0,85 before an increase in early 2019. We can observe a sharp increase in early 2020 before a steep dip. Such sharp movements suggest that there are specific events that have affected the index's volatility, in this case it is due to Covid-19. After the dip Beta stables around 0,8, indicating that NQROBO was less volatile than the market between 2020 and 2021. This coincides with the period where NQROBO outperformed both NASDAQ Corrected and S&P 500 in terms of excess return. This period was greatly affected by interest rates close to 0%, which substantially increased the investment intensity in AI technology. This, combined with the forced increased usage of technology due to the pandemic, drove the development of the whole technological sector. From 2021 to mid-2022, the Beta increased to a level slightly above 1, indicating that NQROBO is more volatile than the market.

7.2.2 Fama French Three-Factor and Five-Factor Model Regression

By including the Fama French Three-Factor (FF3) model and Five-Factor (FF5) model, we can be more confident in our results of the estimated Alpha, this is because FF3 and FF5 capture structural trends affecting the market. When conducting a regression analysis including the different factors, we receive a more comprehensive historical assessment of the Alpha accounted for risk factors. Below is the graphical illustration of FF3 and FF5 regressions over time, providing a better understanding of the development.

Figure 12: NQROBO's FF3-coefficients

Note: Regressed NQROBO's excess return over the Three-Factor model (MKT, SMB, HML)

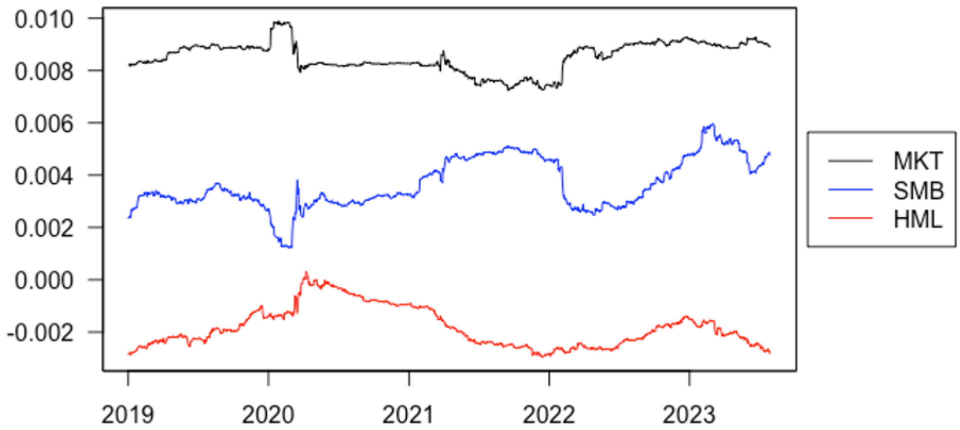
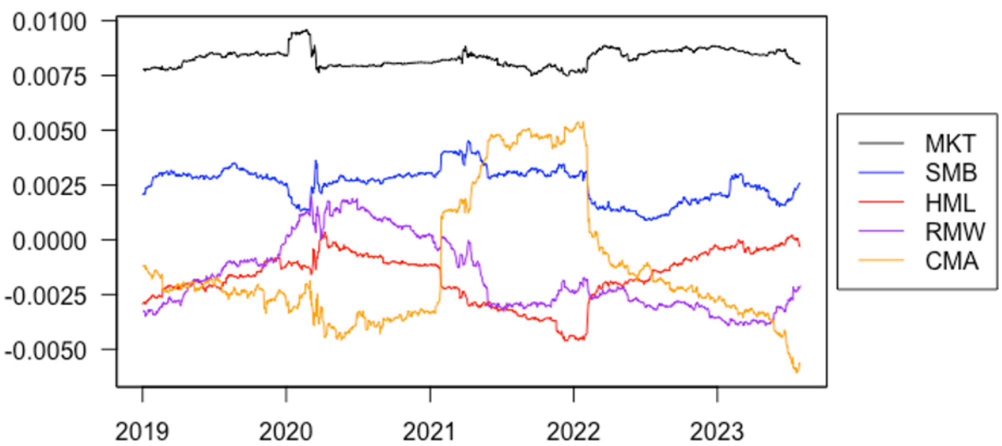


Figure 13: NQROBO's FF5-coefficients

Note: Regressed NQROBO's excess return over the Five-Factor model (MKT, SMB, HML, RMW, CMA)



We can see that for both FF3 and FF5 models, the **MKT** coefficient is positive and relatively stable across the time horizon. This suggests that the broad market risk is a consistent driver of returns for NQROBO.

The **SMB** coefficient fluctuates for both models. However, the coefficient is smaller in magnitude compared to MKT. This suggests that the size coefficient does not consistently capture the returns of NQROBO. Given that firms engaged in AI technology tend to vary significantly in size, from large well-established firms to start-ups, the influence of the size factor may change based on market dynamics.

The **HML** factor is primarily negative, however it seems to diminish over time, this is especially well illustrated in the FF5 model. When the coefficient is negative, this suggests that the NQROBO index may have an inverse relationship with the value factor, implying a growth stock orientation. This is unsurprising as tech firms tend to be growth-oriented rather than value-oriented.

We see that both **RMW** and **CMA** fluctuate significantly, with both coefficients mainly being negative. A negative RMW coefficient suggests that NQROBO tend to perform poorly when firms with high profitability outperform those with low profitability, implying that NQROBO has a high composition of weak profitability firms. This is unsurprising as the index is heavily weighted towards tech companies with high growth potential but currently low profitability. Similarly, the CMA coefficient fluctuates greatly but is primarily negative with an exception between 2021 and 2022. A negative CMA coefficient suggests that NQROBO tend to perform better when the market favours companies with an aggressive investment strategy. This is also no surprise, given that the index consists of tech firms that are dependent on investments to drive technology development.

Given the period of 2019-August 2023, the robotics and automation sector has seen significant developments. Due to the introduction of technological advancements, increased adoption across industries, and effects of external macroeconomic factors, such as the Covid-19 pandemic (March 2020 to January 2022) and long-term interest developments affecting the sector. All these factors should be considered when interpreting these coefficients, as they have greatly affected the analysis. When interpreting the performance of NQROBO relative to these factors, we have provided insights into how the index is aligned or deviates from the general movements in the market.

7.3 Summary of Historical Analysis

In this section, a comprehensive historical analysis has been conducted to provide deeper insights into the performance of the NQROBO index. In section 7.1, we provided a table displaying a detailed overview of NQROBO's weekly and monthly performance in terms of risk and return compared to the market. The analysis spans three sample periods as well as the entire duration of the dataset. Notably, NQROBO has generated a monthly return of 0,70%, over the entire period and was outperformed in terms of excess return by the market for all sample periods except 2020-2022.

When evaluating the active returns, we divided the section into two parts, the first part focuses on the development of Alpha and Beta over time, while the other part observes the Fama French Factors. When observing Alpha's development over time, we found that NQROBO generally displays a positive Alpha, indicating significant outperformance relative to the market, except between 2022-2023. In terms of Beta, NQROBO initially had a Beta at 0,85 before some spikes were identified, primarily due to the pandemic. After these spikes, the Beta continued to increase, ultimately ending close to 1,1. This suggests that NQROBO's volatility relative to the market has increased and remains more volatile than the market.

Furthermore, we conducted a regression including the FF3 and FF5 models based on the equations presented in section 3.1.5. The analysis revealed that the MKT coefficient is stable over time, suggesting that the broad market risk is a consistent driver of returns for the index. The SMB coefficient exhibited substantial fluctuations over time. This is not a surprise as tech firms typically vary in size, and therefore the size factor does not consistently contribute to returns for the index. The value factor, HML, is primarily negative, however, it seems to diminish over time, not making the value factor a driver of returns. This is aligned with our expectations as tech firms tend to be growth-oriented rather than value-oriented. Both CMA and RMW factors fluctuate greatly over time, mainly being negative, suggesting that NQROBO tend to do well when the market favours firms with an aggressive investment strategy and that NQROBO consists mainly of "weak" firms with high growth potential but low profitability.

8. Portfolio Performance Results

In this section, we will provide some graphical displays of our main portfolio performance results. Firstly, we will present the forecasted excess return for our portfolios. Further, we summarise the forecasting results and analyse the forecasting model selection in Portfolio 4. Then we present the trading results and a visualization of weight variations for Portfolio 3 and 4. We end this chapter by analyzing the portfolio performance over two different sub-periods.

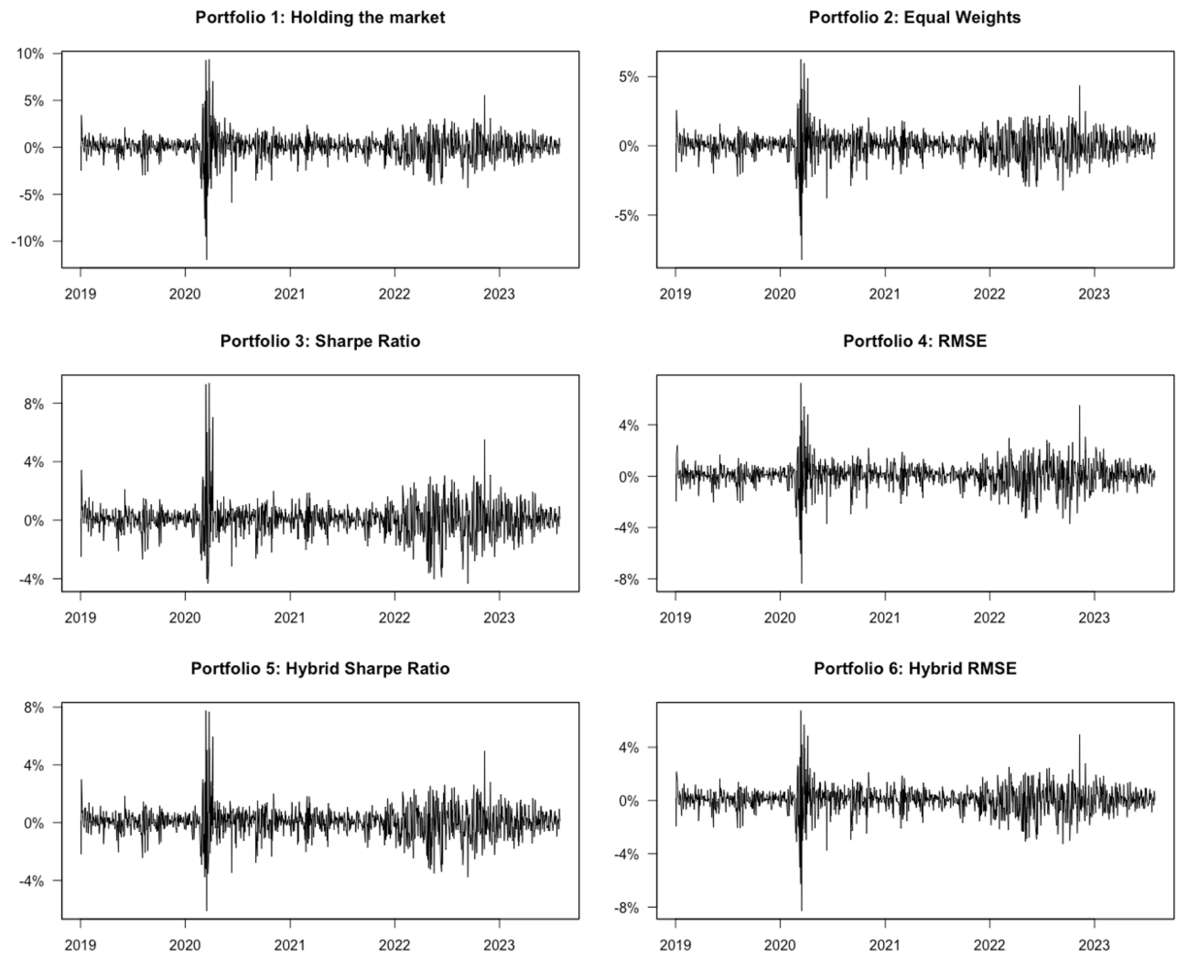
8.1 Excess Return of Our Portfolios

Figure 14 shows the daily excess returns, return relative to risk-free rate, of our six portfolios over the forecasting period. Portfolio 1 appears to have the highest volatility, with returns from +10% to -10%. The other portfolios also show considerable volatility but are less extreme than Portfolio 1.

The portfolios have similar patterns, and all show high volatility in March 2020. As mentioned earlier, this coincides with the Covid-19 pandemic and heightened uncertainty in the stock market. These findings also align with the one week and one month horizon, see Chapter 8 Figure 24 and 27 in Appendix.

Figure 14: Daily excess return for our synthetic portfolios

Note: Excess returns relative to risk-free T-Bill rate
P1 Holding the market, 100% S&P 500
P2 = Equal weights, 1/3 in all three indices
P3 = SR, weightings based on SR
P4 = RMSE, weightings based on RMSE
P5 = SR Hybrid weightings based on SR and Equal weighting
P6 = RMSE Hybrid = weightings based on RMSE and Equal weighting
See chapter 5 for further information about portfolio construction



8.2 Summary of The Forecasting Results

In this section, we will summarise the forecasting results. We used nine forecasting models for the NQROBO index and the NASDAQ at each point in time, see Table 1. The S&P 500 index was forecasted with three models at each point in time. For graphical displays, see Figures 19, 20, and 21 Chapter 8 in Appendix.

NQROBO and NASDAQ

The actual excess returns of the NQROBO and NASDAQ index exhibit high volatility and no clear pattern, which is typical for financial returns. Compared to the actual returns of the degree of volatility is much smaller for the predictions. There is also a varying degree of smoothing of excess returns. Generally, the ARIMA best fit models provide smoother forecasts which minimize the impact of short-term volatility. This indicates that these could be less accurate to sudden market movements. Because of that, the CAPM model and Fama French models with Random Walk seems to capture more of the market risk factor than the corresponding forecasts with ARIMA best fit.

S&P 500

The actual returns for the S&P 500 also exhibit a high degree of volatility and the AR(1) and ARIMA best fit models predicts less volatile returns. The ARIMA best fit forecasting are close to zero in 2019 and after 2022. This means that the time series do not exhibit strong patterns or trends in this period that can be used for forecasting.

8.3 RMSE Model Selection

Table 5 presents the frequency of when a model has the lowest RMSE given a positive predicted excess return.

Table 5: Utilization frequencies

Note: Highlighted numbers indicating the highest frequency of utilization for each index at each time horizon

NQROBO				Frequency					
Time Horizon	Random Walk	AR(1)	ARIMA best fit	CAPM w/RW	CAPM w/ARIMA	FF3 w/RW	FF3 w/ARIMA	FF5 w/RW	FF5 w/ARIMA
One day	112	170	101	100	**216**	81	100	52	99
One week	80	170	138	96	**226**	46	111	56	107
One month	51	196	165	76	**209**	28	140	50	113

NASDAQ	Random Walk	AR(1)	ARIMA best fit	CAPM w/RW	CAPM w/ARIMA	FF3 w/RW	FF3 w/ARIMA	FF5 w/RW	FF5 w/ARIMA
One day	81	**218**	75	103	172	57	156	69	131
One week	72	**260**	79	70	168	43	175	48	140
One month	52	**331**	69	46	163	42	179	43	127

S&P500	Random Walk	AR(1)	ARIMA best fit
One day	389	**458**	157
One week	278	**543**	181
One month	200	**610**	185

For the NQROBO forecasting, CAPM with ARIMA best fit tends to be the most used model across all time horizons. The AR(1)-model has the second highest frequency and is highly used in the one month horizon.

When predicting the NASDAQ excess return, the AR(1) model has the highest frequency for all three horizons. FF3 with ARIMA best fit and CAPM with ARIMA best fit are the two models that are the second and third most used.

The bottom panel of this table shows that the AR(1)-model is the most used in the RMSE selection for the S&P 500 weighting.

8.4 Empirical Portfolio Results

Table 6 displays the outcomes of our six synthetic portfolios. We have presented the mean excess return, standard deviation, Sharpe Ratio, and final wealth for each portfolio. These calculations assume that an investor initially invests $100 000 in January 2019.

Table 6: Portfolio performance results

Note: See Chapter 5 for further explanation of construction of portfolios

	Portfolio 1: Holding the market		
	Daily	Weekly	Monthly
Mean ex. return	0,053%	0,250%	1,070%
S.D. of return	1,37%	2,80%	5,34%
Sharpe ratio	0,04	0,09	0,20
Final wealth ($)	**167 698**	**167 698**	**167 698**

	Portfolio 2: Equal Weighting		
	Daily	Weekly	Monthly
Mean ex. return	0,036%	0,176%	0,675%
S.D. of return	1,01%	2,00%	4,06%
Sharpe ratio	0,04	0,09	0,17
Final wealth ($)	**143 343**	**143 124**	**138 142**

	Portfolio 3: Sharpe Ratio			**Portfolio 4: RMSE**		
	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Mean ex. return	0,055%	0,192%	0,542%	0,079%	0,153%	0,705%
S.D. of return	1,12%	2,20%	4,40%	1,03%	2,05%	4,30%
Sharpe ratio	0,05	0,09	0,12	0,08	0,07	0,16
Final wealth ($)	**176 309**	**147 841**	**128 032**	**232 641**	**135 437**	**141 612**

	Portfolio 5: Hybrid Sharpe Ratio			**Portfolio 6: Hybrid RMSE**		
	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Mean ex. return	0,046%	0,184%	0,608%	0,058%	0,164%	0,690%
S.D. of return	1,04%	2,03%	4,12%	1,00%	1,96%	4,06%
Sharpe ratio	0,04	0,09	0,15	0,06	0,08	0,17
Final wealth ($)	**159 493**	**145 853**	**133 231**	**183 004**	**139 601**	**139 800**

Daily Results

We will first consider the results for a one day investment horizon. The mean excess daily return on the S&P 500 index is 0,053% from 2019 to 1 August 2023. This is higher than the mean excess return of the Portfolio 2 and 5. Portfolio 3 has a marginally higher mean excess return compared to holding the market. The portfolios based on the RMSE selection criteria have a significantly higher mean excess return than the other portfolios.

These differences in mean excess return are reflected in the final wealth. Portfolio 4 pays almost $65 000 more than Portfolio 1. If we compare the RMSE portfolio with the Sharpe Ratio portfolio, RMSE performs better, generating approximately $56 000 more.

We also discover that the standard deviation of the returns of the weighted portfolios lies in a range from 1,0% to 1,12%, which is substantially lower than holding the market (1,37%). Adding the lower standard deviation with the higher mean excess return for Portfolios 3, 4 and 6 results in higher Sharpe Ratio values for these portfolios.

Weekly and Monthly Results

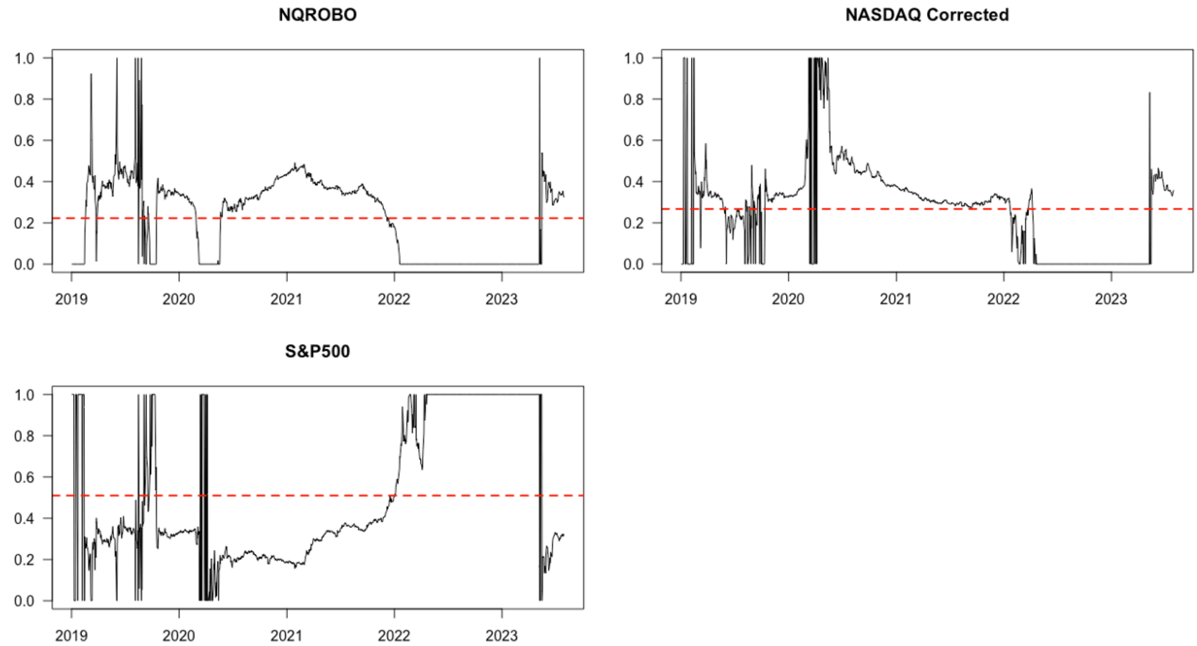
When we extend the investment horizon to one week or one month, we can see that holding the market gives us the highest mean excess return across all portfolios. Portfolio 1 also gives us the highest final wealth for these horizons. Furthermore, we observe that the standard deviations for Portfolio 2-6 are still lower than holding the market.

8.5 Portfolio Weights

This section introduces the variation in weighing for Portfolio 3 and 4. For each of the indices, Figure 15 displays the weights of the Sharpe Ratio selection portfolio over time.

Figure 15: Portfolio 3 weights in each index over time

Note: P3 = SR, weightings based on SR measurements
Red dotted lines visualize the average weights over the whole period



By looking at the variations in weights, we see that in large parts of 2022 and 2023, the weights of NQROBO and NASDAQ Corrected are zero. This is because of a negative Sharpe Ratio for all the indices in this period and we therefore choose to weight the S&P 500 index 100%. On average, the weights of the S&P 500 are 51%. For NQROBO and NASDAQ, the weights are on average 22% and 27%, respectively. Compared to the equal weighted portfolio, we can conclude that this portfolio overweighted the S&P 500 index and underweighted NQROBO and NASDAQ on average.

Figure 16: Portfolio 4 weights in each index over time

Note: P4 = RMSE, weightings based on RMSE measurements
Red dotted lines visualize the average weights over the whole period

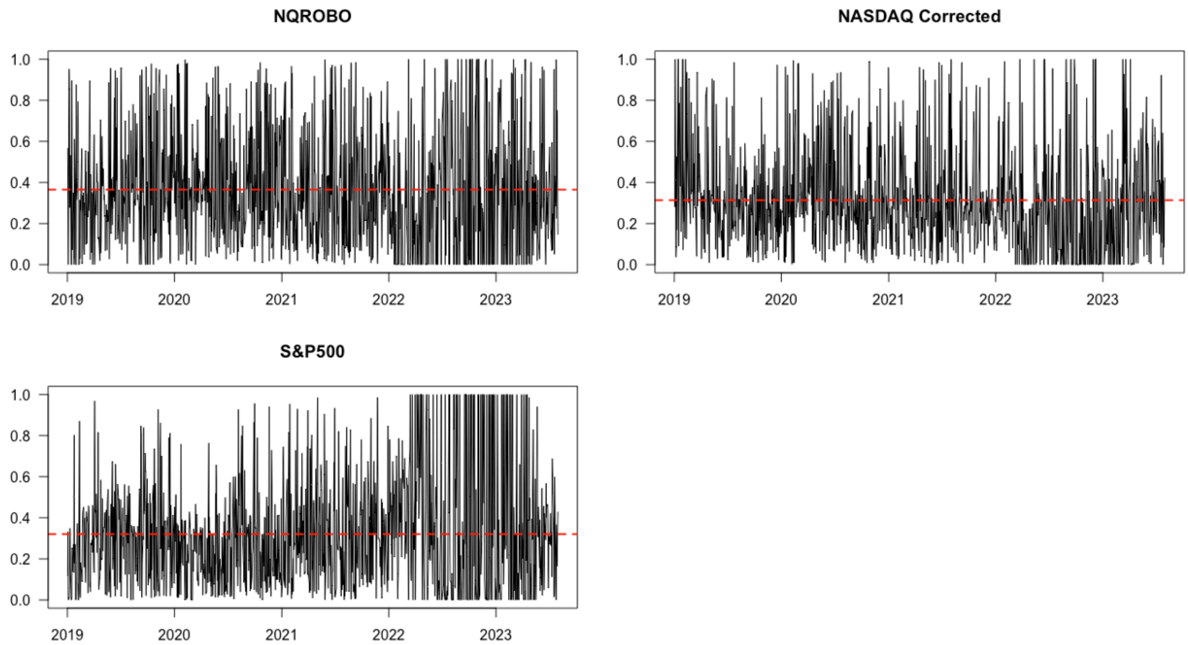


Figure 16 shows the weights for each index in the RMSE selection portfolio with a time horizon of one day. We can observe that this portfolio tends to do more frequent changes in the index weightings. In Portfolio 4, the average weight for NQROBO is 37%, 31% for NASDAQ Corrected, and 32% for S&P 500. This means that the portfolio is, on average, overweighted in NQROBO and underweighted in NASDAQ Corrected and S&P 500 compared to the equal weighted portfolio.

If we extend the horizon to weekly, the average weight in NQROBO decreases and NASDAQ Corrected and S&P 500 increase. For a monthly horizon, the weight on average in NQROBO is 35%, and 31% in NASDAQ. For graphical displays of weight variations for one week and one month, see the Appendix Figure 25 and 28.

8.6 Sub-Periods

We have also analysed the daily performance of the synthetic portfolios over two different sub-periods of our data sample. To better understand the development in performance, we have included cumulative return as a measurement. The first sub-period is from January 2019 to December 2020, and the second is from January 2021 to August 2023. The reason for choosing these two sub-periods is because we want to compare the portfolio performance under different market conditions.

Table 7: Daily portfolio results for subperiods (2019-2020 and 2021-August 2023)

	Portfolio 1: Holding the market		**Portfolio 2: Equal Weights**	
	2019-2020	2021-2023	2019-2020	2021-2023
Mean ex. return	0,089%	0,031%	0,070%	0,010%
Cum. return	46,1%	14,8%	38,2%	3,7%
S.D. of return	1,63%	1,16%	1,13%	0,91%
Sharpe ratio	0,05	0,03	0,06	0,01
Final wealth ($)	**146 107**	**167 698**	**138 221**	**143 343**

	Portfolio 3: Sharpe Ratio		**Portfolio 4: RMSE**	
	2019-2020	2021-2023	2019-2020	2021-2023
Mean ex. return	0,108%	0,014%	0,099%	0,063%
Cum. return	66,9%	5,7%	60,0%	45,4%
S.D. of return	1,15%	1,09%	1,10%	0,98%
Sharpe ratio	0,09	0,01	0,09	0,06
Final wealth ($)	**166 867**	**176 309**	**160 034**	**232 641**

	Portfolio 5: Hybrid Sharpe Ratio		**Portfolio 6: Hybrid RMSE**	
	2019-2020	2021-2023	2019-2020	2021-2023
Mean ex. return	0,089%	0,012%	0,085%	0,036%
Cum. return	52,3%	4,7%	48,8%	23,0%
S.D. of return	1,09%	0,99%	1,10%	0,92%
Sharpe ratio	0,08	0,01	0,08	0,04
Final wealth ($)	**152 279**	**159 493**	**148 834**	**183 004**
N	505	647	505	647

As shown in Table 7, Portfolio 3 has the best performance during the first sub-period, with the highest final wealth after two years when rebalancing daily. However, in the second sub-period the performance of this portfolio has a sharp decline, and a lower cumulative return than holding the market. Further, we find that Portfolio 4 and 6 has a higher cumulative return and final wealth than Portfolio 1 in both sub-periods. We also observe that Portfolio 5 outperforms Portfolio 1 in the first sub-period, but significantly underperformed in the second sub-period compared to holding the market. The results are aligned with the results presented in Table 6, indicating that daily rebalancing and reinvesting perform best. This is possibly due to frequent rebalancing, being able to capture short term gains in the market that don’t persist over a week or month.

9. Discussion

In this chapter, we will analyse and discuss our findings presented in Chapter 8 to answer our second and third part of our thesis. We want to examine to what extent investing in AI, supported by a forecasting tool using minimal hindsight, offers the potential of beating the market for a moderately risk-averse investor. Furthermore, we assess the optimal relative weighting of Portfolio 3 and 4.

9.1 Discussion of Forecasting Results

Our discussion of forecasting results is based on the findings represented in Table 6, section 8.4.

Daily

From Table 6, we can observe the performance of our six synthetic portfolios. Our results have identified that Portfolio 3, 4 and 6, respectively using Sharpe Ratio, RMSE and Hybrid RMSE for weighting, have outperformed holding the market in terms of mean excess return. When observing the other financial measurements for all portfolios, we find that Portfolio 1 has the shared lowest Sharpe Ratio. This is due to Portfolio 1 having both high return and the highest standard deviation. When assessing all portfolios, we find that Portfolio 3, 4 and 6 are the only ones outperforming holding the market regarding final wealth. Looking at both hybrid portfolios, we see that their performance is between the equal weighted portfolio and those strictly computed using Sharpe Ratio and RMSE. This seems reasonable, given the construction of the portfolios.

Weekly

When observing the weekly financial measurements, certain changes need to be addressed. In terms of mean excess return and final wealth, Portfolio 1 now outperforms all portfolios. Several reasons could explain this development. One of the most likely reasons is that the market is generally efficient over longer periods. This is due to daily fluctuations and noise being “smoothed” away. We also see that holding the market has the highest standard deviation of all the portfolios. Therefore, it is not surprising that holding the market also has a relatively high mean excess return as investors demand high returns when taking on higher risks.

Monthly

Observing monthly, we find that holding the market now also outperforms all other portfolios in terms of mean excess return, Sharpe Ratio and final wealth. One can interpret this as monthly rebalancing does not capture the short-term gains that daily rebalancing capture. Meaning that the potential gains of investing in AI must be captured quickly as profits are short-lived.

Sub-Periods

We chose to observe the synthetic portfolios under two sub-periods to see how they would perform under different market conditions. We found that the Portfolio 3 performed the best, having the highest cumulative return, during the first sub-period. However, the portfolio performed significantly worse in the second sub-period. This is surprising given that RMSE is performing relatively consistent across both periods. A potential reason for this decline could be the forecasting accuracy for SR being worse, resulting in bad investment decisions and economic loss during the second sub-period.

9.2 Discussion of Portfolio Weights

When analysing the portfolio weight, we look at the empirical results presented in section 8.5. The aim of analysing the weights is to observe and identify when and how much the portfolios over- and underweights the different indices compared to the equaled weighted portfolio. By conducting such an analysis, we hope to discover a pattern across all portfolios, identifying a period where potential gains by investing in NQROBO could be made.

By comparing the weights of Portfolio 3 and 4, we observe that the weights of Portfolio 4 are more volatile than the weights of Portfolio 3. This implies a more active portfolio management approach with the attempt to capitalize on short-term movements. On average, Portfolio 3 tends to overweight the S&P 500 index, while Portfolio 4 seems to overweight the NQROBO index compared to Portfolio 2. Considering the final wealth, we discover that a flexible weighting strategy with an average overweight in NQROBO delivers the highest return and wealth in this period.

9.3 Discussion of Risk Profile

In our thesis, we have not chosen to take an explicit stance on the risk aversion of our fictional investor. However, we have constructed the synthetic portfolios utilising RMSE and SR. Both measurements could be rationalised to be connected to a loss function. We argue that RMSE is reconciled with quadratic utility, implying certainty equivalence, as our investor is more interested in the expected return rather than the magnitude of the return. Likewise, utilising SR when deciding the weights of Portfolio 3 and 5, implies that our fictional investor is using risk-adjusted excess returns. We define our fictional investor as risk-neutral, considering both risk and return closely.

9.4 Limitations of The Analyses

Several assumptions and simplifications were made when conducting our analyses. This section aims to address the main limitations that directly affect our research question.

9.4.1 Quality of Data

One of the most apparent identified limitations of our analysis is regarding the data quality. Firstly, the return data consists of 1423 daily observations from January 2018 to August 2023. This is a relatively short period, with data observed over only five and a half years. Secondly, the quality of the data is debatable as our data was retrieved in a period with a significant specific event being the Covid-19 pandemic and the backlash of this. There is no denying that the pandemic had great implications on macroeconomic factors, investment intensity and the increased usage of new technologies. Therefore, it is essential to discuss to what extent our findings are realistic and applicable when in a “normal situation”. Conducting a more extensive study sometime in the future, during a “normal situation”, would enable a more thorough exploration of our research question.

We also have to address the missing observations. When retrieving our datasets containing daily returns from NQROBO, NASDAQ and S&P 500, we found discrepancies in the number of observations between the three indices. We solved this by disregarding all observations for that particular day for all three indices if an observation were missing. We found that this was the optimal solution, however, it is important to acknowledge the weaknesses it entails. In total, 28 trading days were disregarded because of deficient information.

9.4.2 Transaction Cost

Our thesis and the conducted analyses rest on several assumptions and simplifications that must be addressed. One of these simplifications is that transaction cost is disregarded.

In the real world, this assumption is unrealistic. However, the reasoning behind the choice is the difficulties in finding one appropriate rate for all three indices and the increased complexity it would entail. If transaction costs were accounted for, it would have a negative impact on the final wealth for all six synthetic portfolios, especially the last five portfolios with an active management. The higher trading frequency an investor has, the more they will incur transaction costs, consequently “eating” into their final wealth. So, while our approach simplifies things, it is important to recognise the importance of transaction cost and how this would significantly impact the actual outcome of investment strategies.

9.4.3 Nasdaq CTA Artificial Intelligence & Robotics

As mentioned in section 2.6, we have chosen to only use the NQROBO index to track the performance of companies engaged in Artificial Intelligence. By only using one index to track the whole development of AI, there are certain limitations that must be addressed. The main limitation regarding the use of only NQROBO is the limited representation. The index may not be able to capture the whole breadth of the AI landscape due to the fast-evolving nature of AI technology and the exclusion of non-public companies not being represented in the index. By not having a complete representation of the whole AI landscape due to the nature of the index, we do not possess the full information to evaluate the performance of AI-engaged firms entirely correct.

9.5 Robustness Analysis

In this section, we conduct a robustness analysis of our findings in Table 6. The purpose of this is to examine the validity of our findings. Our robustness analysis can be divided into two parts. The first part examines if our results hold true if the forecasting is now computed on a six-month window instead of the original twelve-month window. The second part will examine how different reference weights in our hybrid portfolios affect final wealth.

9.5.1 Forecasting Based on Six Months of Training Data

The first part of our robustness test examines how our results differ if we change the duration of the training data before forecasting. Initially, we have chosen to base our training data on 1-year observations for all three indices. In Table 8, we have displayed the portfolio performance based on a six-month estimation window.

Table 8: Portfolio performance based six-months of training data

	Portfolio 3: Sharpe Ratio			**Portfolio 4: RMSE**		
	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Mean ex. return	0,045%	0,194%	0,637%	0,081%	0,199%	0,713%
S.D. of return	1,2%	2,4%	4,49%	1,04%	2,05%	4,25%
Sharpe ratio	0,04	0,08	0,14	0,08	0,10	0,17
Final wealth ($)	**153 857**	**146 793**	**134 431**	**238 014**	**151 887**	**141 976**

	Portfolio 5: Hybrid Sharpe Ratio			**Portfolio 6: Hybrid RMSE**		
	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Mean ex. return	0,041%	0,188%	0,671%	0,059%	0,191%	0,709%
S.D. of return	1,1%	2,1%	4,17%	1,01%	1,94%	4,01%
Sharpe ratio	0,04	0,09	0,16	0,06	0,10	0,18
Final wealth ($)	**150 227**	**146 521**	**137 672**	**186 668**	**148 879**	**141 280**

We have identified some key differences between Table 6 (twelve-months) and Table 8 (six-months). The most prominent difference is that when utilising six-months of training data, all portfolios do better, in terms of final wealth, when rebalancing and reinvesting monthly. One possible reason for this is that when the model is trained on twelve months of data, it may be too generalised and does not capture trends or adapt to changes in the market, compared to the model trained on a shorter period.

The other key difference is that when rebalancing daily and weekly, Portfolio 4 and Portfolio 6 perform better when trained on six months of data, while Portfolio 3 and 5 perform worse. This could be due to the nature of the portfolio constructions. RMSE-based portfolios may be more responsive to market volatility, better capturing the underlying market structure and seasonality trends.

Another reason may be that the SR portfolios overfit the data, meaning that the model is too closely tailored to the characteristics in the training data. A possible reason for overfitting may be due to nature of the construction of the portfolios, relying too much on historical data.

To summarise, Portfolio 4 and 6, still outperforms holding the market. They still generate a higher excess return after changing the duration of the training data, leading to similar results as in Table 6. They key differences found is that when training the forecasting on a shorter period the RMSE based portfolios perform better when rebalancing daily and weekly, while the SR based portfolios perform better monthly.

9.5.2 Changes in Reference Weights

This part of the robustness test is largely motivated by the Government Pension Fund Global (GPFG), which is weighted 70% in stocks and 30% in bonds (NBIM, 2023). We want to examine if our findings hold true if changes in the reference weights in our hybrid portfolios (Portfolio 5 and 6) are made. With inspiration from the GPFG, we have changed the weights to be 70% dynamic weighted and 30% fixed weights. See Appendix Chapter 5, Equation 20 and 21, to understand how the new weights of the portfolios are found. Table 9 presents a summary of the second part of our robustness analysis.

Table 9: Changes in Portfolio 5 and 6 when the reference weights change

	Porfolio 5: Hybrid Sharpe Ratio			**Portfolio 6: Hybrid RMSE**		
	Daily	Weekly	Monthly	Daily	Weekly	Monthly
Mean ex. return	0,040%	0,207%	0,765%	0,066%	0,160%	0,696%
S.D. of return	1,14%	2,3%	4,41%	1,01%	1,98%	4,13%
Sharpe ratio	0,04	0,09	0,17	0,07	0,08	0,17
Final wealth ($)	**147 713**	**152 215**	**144 246**	**201 546**	**138 009**	**140 508**

Comparing the Table 6 and 9, we see some differences in the performance measurements due to changes in the reference weights.

Looking at the Hybrid Sharpe Ratio we can see that the final wealth has increased for weekly and monthly rebalancing. However, it has decreased daily when the dynamic weighting has been increased to 70%. This suggests that dynamic weighting is more effective on longer rebalancing intervals. Observing the standard deviation, we see that it has increased for all rebalancing durations for both portfolios. All portfolios have equal or higher Sharpe Ratio when increasing the weights of dynamic allocation to 70%.

The Hybrid RMSE has increased the final wealth for daily and monthly rebalancing durations when the dynamic allocation has increased to 70%. Suggesting that the RMSE weighting is particularly good at forecasting daily and monthly market changes. However, the model may experience noise, resulting in the model reacting to false trends that do not persist weekly, resulting in a decrease in final wealth when the amount of dynamic allocation increases.

To summarise, the results presented in Table 6 still holds true as we receive similar results in Table 9 when changing the reference weights. Portfolio 6 is still the only hybrid portfolio outperforming holding the market with daily rebalancing in terms of final wealth. The key difference to highlight is that a higher dynamic allocation is beneficial for the Hybrid RMSE when rebalancing daily, while it is beneficial for the Hybrid Sharpe Ratio when rebalancing weekly and monthly.

9.6 Further Research

This thesis examines to what extent investing in AI, supported by a forecasting tool, offers the potential of beating the market. When creating our portfolios, we have excluded the possibility of shorting the three indices based on the intricacies this would entail. Given the nature of the investment period, it would be interesting to see how the portfolios would perform if shorting were allowed. By shorting we mean, selling indices that we expect, based on our measurements, to generate negative returns. If allowing for shorting one also have to allow for analysis of negative Sharpe Ratios. The period is greatly influenced by significant fluctuations in returns and high volatility due to the pandemic, thus making the correct decision whether to invest or short highly lucrative. This would be especially lucrative for an investor in 2022 when all indices had negative returns.

Another area of interest for future research, could be the analysis of stochastic volatility. For example, estimating a GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model and examining the effects found. The primary purpose of conducting such a model is to forecast and examine future volatility. This is done given the idea behind GARCH is that past volatility has a lasting impact on present and future volatility, resulting in volatility clustering.

Lastly, it would be interesting to examine the inclusion of transaction costs further and observe how this would affect the final wealth of the different portfolios.

10. Conclusion

In this thesis, we have assessed the performance of the Nasdaq CTA Artificial Intelligence & Robotics index and explored the extent to which investing in AI offers the possibility of beating the market. We define "beating the market" as achieving a higher final wealth by investing in an alternative portfolio rather than investing in the market. Our thesis aims to address the following research question and is structured accordingly.

To what extent does investing in Artificial Intelligence technology, supported by a forecasting tool only using historical information, offer the potential of beating the market index for a risk-neutral investor with a short time horizon?

To answer this question, we divided the thesis into three parts, each part supported by an analysis. A historical analysis evaluating NQROBO's performance, a pseudo-out-of-sample forecasting performance analysis exploring how an investor in real time utilising a forecasting tool would perform, and lastly, an optimal relative weighting analysis of NQROBO, based on the pseudo-out-of-sample analysis.

The **historical analysis** identified that NQROBO outperformed the market between January 2020 and December 2021, in terms of weekly and monthly returns. For the whole period, NQROBO was slightly outperformed by the market. Observing NQROBO's Alpha and Beta over time, we discovered that the Alpha was primarily positive from 2020 to early 2022 before turning negative later in 2022. The Beta began slightly below 1, before decreasing to 0,8 in 2020, then sharply rising in 2022, stabilising at 1,1. When looking at how the Fama French Factors affect NQROBO, we found that the market was a consistent driver for their return. SMB fluctuated greatly, being both negative and positive, suggesting that the size factor is not a consistent driver. The remaining factors, HML, RMW and CMA, all fluctuates, being mostly negative. This suggest that NQROBO is performing best when the market favours growth-oriented companies with an aggressive investment strategy. To summarize, the index displays potential of outperforming the market for certain periods if the market conditions are favourable.

To answer the second part of the thesis, we proposed an approach for simulating an investor's behaviour in real time using as little hindsight as possible. Assessing the **pseudo-out-of-sample forecasting performance analysis** indicates an opportunity to generate a higher final wealth by investing in AI. Different forecasting models were trained over a year, utilising dynamic portfolios, and only using optimal forecasting. We found that dynamic weighting based on Sharpe Ratio, RMSE and Hybrid RMSE outperform holding the market in terms of final wealth when rebalancing and reinvesting daily. At a weekly and monthly basis, holding the market outperformed all portfolios suggesting that potential gains present are short-term.

To answer our third and last part of the thesis, we observe the NQROBO's weighting over time in our **optimal relative weighting analysis**. Our analysis shows that the RMSE weighting portfolio, that is more volatile and dynamically weighted, outperforms the other portfolios. Implying that a portfolio with a high degree of dynamic weighing may be able to capture and capitalize on short-term movements in the market. These findings are also supported by the dynamic portfolios performing better when rebalancing and reinvesting daily, compared to weekly and monthly.

This thesis concludes that investing in Artificial Intelligence exhibits a clear opportunity of beating the market for a risk-neutral investor. However, it is important for an investor to utilise dynamic weighting frequently to be able to capture short-term gains present in the market and the market conditions must be favourable.

References

Alsheibani, S. A., Cheung, Y., Messom, C., & Alhosni, M. (2020). *Winning AI Strategy: Six-Steps to Create Value from Artificial Intelligence*. Americas Conference on Information Systems. https://aisel.aisnet.org/amcis2020/adv_info_systems_research/adv_info_systems_research/1/

Bossmann, J. (2016). *Top 9 ethical issues in artificial intelligence*. World Economic Forum. https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994) *Time Series Analysis: Forecasting & Control* (3rd edition). Pearson Education India.

Box, G.E.P. et al. (2015) *Time Series analysis: Forecasting and Control* (5th edition). John Wiley & Sons.

Buchanan, B., & Cao, C. X. (2018). *Quo vadis?: A comparison of the fintech revolution in China and the West.* Social Science Research Network. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3264146

Buchholz, K. (2023). *Threads shoots past one million user mark at lightning speed*. Statista Daily Data. https://www.statista.com/chart/29174/time-to-one-million-users/

Chen, J. (2023). *Alpha: What it means in investing, with examples*. Investopedia. https://www.investopedia.com/terms/a/alpha.asp

DesJardins, J. R. (2013). *An introduction to business ethics* (5th edition). McGraw Hill.

Duarte, F. (2023). How many AI Companies are there? *Exploding topics.*

Du, S., & Xie, C. (2021). Paradoxes of artificial intelligence in consumer markets: Ethical challenges and opportunities. *Journal of Business Research*, 129, 961–974. https://doi.org/10.1016/j.jbusres.2020.08.024

Enholm, I. M., Papagiannidis, E., Mikalef, P., & Krogstie, J. (2021). Artificial Intelligence and Business Value: a Literature Review. *Information Systems Frontiers*, 24(5), 1709–1734. https://doi.org/10.1007/s10796-021-10186-w

Fama, E. F. (1965). The behavior of Stock-Market prices. *The Journal of Business*, 38(1), 34. https://doi.org/10.1086/294743

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56. https://doi.org/10.1016/0304-405x(93)90023-5

Fountaine, T. (2019). *Building the AI-Powered organization*. Harvard Business Review. https://hbr.org/2019/07/building-the-ai-powered-organization

Future of Life Institute. (2023). *Pause Giant AI Experiments: An Open Letter*. Retrived from: https://futureoflife.org/open-letter/pause-giant-ai-experiments/

Grand View Research. (2023). *Artificial Intelligence Market Size, Share & Trends Analysis Report by solution, by technology (Deep learning, machine Learning), by end-use, by region, and segment Forecasts, 2023 - 2030*. Retrived from: https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market

Hawking, S. (2016). *The best or worst thing to happen to humanity*. University of Cambridge. https://www.cam.ac.uk/research/news/the-best-or-worst-thing-to-happen-to-humanity-stephen-hawking-launches-centre-for-the-future-of

Huynh, T. L. D., Hille, E., & Nasir, M. (2020). Diversification in the age of the 4th industrial revolution: The role of artificial intelligence, green bonds and cryptocurrencies. *Technological Forecasting and Social Change*, 159, 120188. https://doi.org/10.1016/j.techfore.2020.120188

Hyndman, R.J. and Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for R. *Journal of statistical software*, *27*, 1-22.

Jensen, M. C., Black, F., & Scholes, M. S. (1972). The capital asset pricing model: Some empirical tests. *Studies in the theory of capital markets*.

LeCun, Y., Bengio, Y., & Hinton, G. E. (2015). Deep learning. *Nature*, 521(7553), 436–444. https://doi.org/10.1038/nature14539

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Kucharavy, D., & De Guio, R. (2015). Application of logistic growth curve. *Procedia engineering*, *131*, 280-290.

Lutkevich, B. (2022). *AI winter*. Enterprise AI. https://www.techtarget.com/searchenterpriseai/definition/AI-winter

Marr, B. (2023). *A short history of ChatGPT: How we got to where we are today*. Forbes. https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=3232a077674f

McCorduck, P., & Cfe, C. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. CRC Press.

NASDAQ. (2022). *Nasdaq CTA Artificial Intelligence and Robotics index*. Retrived from: https://indexes.nasdaqomx.com/docs/methodology_NQROBO.pdf

NASDAQ. (2023). *Weighting values for DRBG*. Retrived from: https://www.nasdaq.com/NQROBO

NASDAQ. (2023). *Nasdaq Composite*. Nasdaq Indexes. Retrived from: https://indexes.nasdaqomx.com/docs/FS_COMP.pdf

Norges Bank Investment Management. (2023). *Investeringsstrategien.* NBIM. https://www.nbim.no/no/oljefondet/slik-er-fondet-investert/investeringsstrategien/

Pesaran, M. H., & Timmermann, A. (1995). Predictability of stock returns: Robustness and economic significance. *The Journal of Finance*, *50*(4), 1201–1228. https://doi.org/10.1111/j.1540-6261.1995.tb04055.x

PwC. (2016). *Sizing the prize: What's the real value of AI for your business and how can you capitalize?* Retrived from: https://www.pwc.no/no/publikasjoner/Digitalisering/pwc-ai-analysis-sizing-the-prize-report.pdf

Ransbotham, S., Kiron D., and Reeves M. (2017). *Reshaping business with artificial intelligence*. MIT Sloan Management Review. https://sloanreview.mit.edu/projects/reshaping-business-with-artificial-intelligence/

Regjeringen. (2020). *Nasjonal strategi for kunstig intelligens*. Retrived from: https://www.regjeringen.no/contentassets/1febbbb2c4fd4b7d92c67ddd353b6ae8/no/pdfs/ki-strategi.pdf

S&P Dow Jones Indices. (2023). *S&P 500*. Retrived from: https://www.spglobal.com/spdji/en/indices/equity/sp-500/#overview

Samuel, J. (2023). Response to the March 2023 "Pause Giant AI Experiments: An Open Letter" by Yoshua Bengio, signed by Stuart Russell, Elon Musk, Steve Wozniak, Yuval Noah Harari and others. *Social Science Research Network*. https://doi.org/10.2139/ssrn.4412516

Schwab, K. (2017). *The Fourth Industrial Revolution*. Currency.

Schmidt, D. & MarketBeat (2023). *Why Do Tech Stocks Go Down When Interest Rates Rise?* Nasdaq. https://www.nasdaq.com/articles/why-do-tech-stocks-go-down-when-interest-rates-rise

Schultz-Bergin, M. (2021, May 1). *The 4th Industrial Revolution*. The Primacy of the Public. https://pressbooks.ulib.csuohio.edu/principles-of-engineering-ethics/chapter/the-4th-industrial-revolution/

Senator, T. E., Goldberg, H. G., Wooton, J., Cottini, M. A., Khan, A. F. U., Klinger, C. D., Llamas, W. M., Marrone, M. P., & Wong, R. W. H. (1995). Financial Crimes Enforcement Network AI System (FAIS) Identifying Potential Money Laundering from Reports of Large Cash Transactions. *Ai Magazine*, *16*(4), 21–39. https://doi.org/10.1609/aimag.v16i4.1169

Sharpe, W.F. (1964). Capital Asset Prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3), 425-442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, 39(1), 119-138. https://doi.org/10.1086/294846

Sharpe, W. F. (1991). The arithmetic of active management. *Financial Analysts Journal*, 47(1), 7-9. https://doi.org/10.2469/faj.v47.n1.7

Shumway, R. H., Stoffer, D. S. (2017). *Time series analysis and its applications* (4th edition). Springer International Publishing AG.

Singer, Peter W. (2009). *Wired for War: The Robotics Revolution and Conflict in the Twentyfirst Century*. Penguin Books.

Skilton, M., Hovsepian, F. (2017). *The 4th Industrial Revolution: Responding to the Impact of Artificial Intelligence on Business*. Palgrave Macmillan.

Smith, T. (2023). *Random walk theory: Definition, how it's used, and example*. Investopedia. https://www.investopedia.com/terms/r/randomwalktheory.asp

Tiwari, A. K., Abakah, E. J. A., Le, T. N. L., & La Hiz, D. I. L. (2021). Markov-switching dependence between artificial intelligence and carbon price: The role of policy uncertainty in the era of the 4th industrial revolution and the effect of COVID-19 pandemic. *Technological Forecasting and Social Change*, *163*, 120434. https://doi.org/10.1016/j.techfore.2020.120434

Wooldridge, J.M. (2019). *Introductory Econometrics: A Modern Approach*. Cengage Learning.

Wooldridge, J. M. (2015). Introductory econometrics: A modern approach. Cengage Learning.

Zhang, Y. J., Zhang, H., & Gupta, R. (2023). A new hybrid method with data-characteristic-driven analysis for artificial intelligence and robotics index return forecasting. *Financial Innovation*, 9(1), 75. https://doi.org/10.1186/s40854-023-00483-5

Zuckerman, G. (2019). *The man who solved the market: How Jim Simons launched the quant revolution*. Penguin.

Appendix

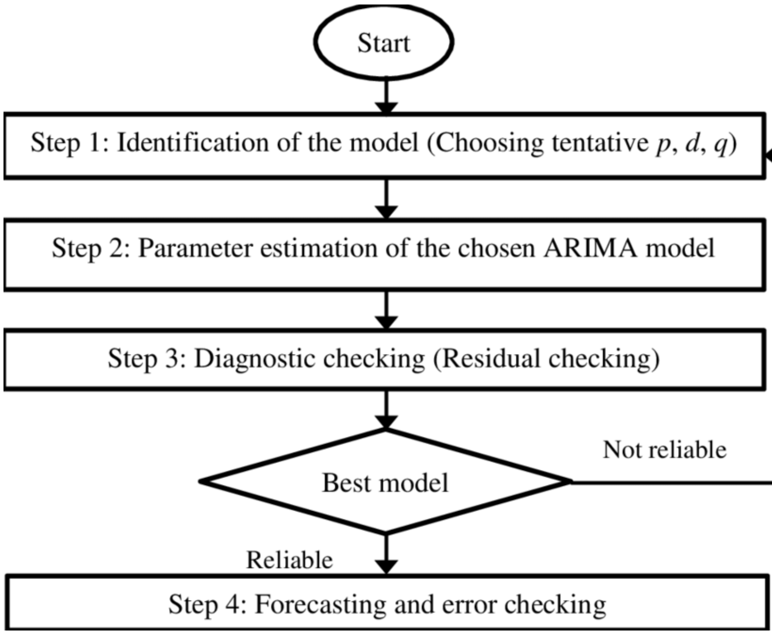
The appendix follows the same structure as the thesis.

4. Forecasting Methodology

The Box-Jenkins Approach for Model Selection

We used the Box-Jenkins methodology to determine the most appropriate ARIMA model. This approach is a systematic iterative method for identifying, estimating, testing, and applying ARIMA (Box and Jenkins, 1994). As demonstrated in Figure 17, the method is separated into three phases.

Figure 17: The Box-Jenkins Approach



In the identification phase, we check for stationarity and select the parameters of an ARIMA model that best summarizes the time series most accurate. In the next phase, estimation, we use the data to train the parameters of the ARIMA model. The last step of the process is to evaluate the fitted model and check for improvements (Box and Jenkins, 1994).

Phase 1: Identification

Stationarity

The first step is to test the time series for stationarity. By stationary data, we mean that the properties of the time series do not change across time. This means that time series with trends or with seasonality are not stationary. In other words, a stationary time series does not have any long-term predictable patterns (Hyndman & Athanasopoulos, 2018). Three conditions need to be fulfilled to consider a time series stationary; (1) constant mean over time, (2) constant variance over time, and (3) constant autocovariance. If the time series is not stationary, it could lead to spurious forecasting results and it would be necessary to differentiate the data (Hyndman and Athanasopoulos, 2018). By removing changes in the level of a time series, differences can help stabilize the meaning of a time series, therefore eliminating seasonality and trend. When the time series is stationary, the component, d, will be equal to zero.

Augmented Dickey-Fuller test

The ADF test is a statistical test where one investigates for stationarity in the data, by checking for unit root (Wooldridge, 2015). The null hypothesis of the ADF test is that there is a unit root in the time series, which implies that the data is not stationary. The alternative hypothesis is that the data is stationary. If the p-value is lower than 0,05, we can reject the null hypothesis and conclude that the time series is stationary (Wooldridge, 2019).

Autoregressive and moving average order

The next step in the identification process is to identify the autoregressive order (p) and the moving average order (q). To find an appropriate value for these parameters, one could analyze the autocorrelation function and the partial autocorrelation function.

The autocorrelation function is the correlation between the time series with a lagged version of itself. The second function that expresses information useful in determining the order of an ARIMA model is the partial autocorrelation function. This function calculates the partial correlation between the values of two time periods and adjusts out the influence of intermediate lags. The estimated value of p is determined by the last lag with a large value of the appropriately differenced series. If the partial autocorrelation function does not have a cut-off after a few lags, the alternative is to either have a moving average model (p=0) or an ARIMA model with $p > 0$ and $q > 0$. The estimated value of q is determined by the last lag with a large value from the autocorrelation function. If the function does not have a cut-off after a few lags, the alternative is to either have an autoregressive model (q=0) or an ARIMA model with $p > 0$ and $q > 0$ (Hyndman and Athanasopoulos, 2018).

Phase 2: Model estimation

The Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are useful techniques to select the parameters of the ARIMA model. One obtains good models, by minimising AIC and BIC. The AIC offers a good balance between giving biased estimates when the model's order is too low, and there is a potential for increasing the variance when an excessive number of regressors are included.

The AIC can be written as:

Equation 16

$$AIC = -2\log(L) + 2(p + q + k + 1)$$

The BIC equation is outlined below:

Equation 17

$$BIC = AIC + [\log(T) - 2](p + q + k + 1)$$

The AIC tend to prefer larger models, with more p and q lags, compared to BIC which penalizes model complexity more. There is some ambiguity in the selection criteria between AIC and BIC, because there is no statistical test to compare these two estimators. This means that we have to do a subjective comparison of the two criteria.

Auto-arima()

As mentioned in Chapter 4, we have supported our identification and model estimation with *auto.arima()*, in order to remove possible human error. The *auto.arima()* function from the package *forecast* in R, uses a Hyndman-Khandakar algorithm to estimate the most suitable ARIMA model. This algorithm combines unit root tests and minimization of the AIC, BIC, and AICc (Hyndman & Athanasopoulos, 2008).

The first step of the function is to determine whether the time series is stationary. This is done using a ADF test. If the time series is not stationary, *auto.arima()* determines the order of differencing (d) required to achieve stationarity. The next step in the function, is to calculate the information criteria for each model combination. In this thesis, we have set the information criteria to AIC, BIC and AICc. The default information criterion in the function is AICc, which means that in cases with different criteria values, this criterion is deciding. The output of the *auto.arima()* is the best ARIMA model that describes the relevant data series, based on the information criterion.

Phase 3: Model diagnostics

The last phase of the Box-Jenkins approach is to ensure that the ARIMA model is a good fit for the data. Two useful methods of checking a model are overfitting and residual errors.

By checking for overfitting, we check if the model is more complex than it needs to be and captures random noise. This is a problem because it impacts the ability of the model to generalize negatively, and leads to poor forecast performance (Box et al., 2015).

From an ideal model, we should observe that the errors follow white noise and are normally distributed. We use the Ljung-Box test to test whether or not the errors are independent and identically distributed (IID) random variables. The null hypothesis of the test is that the model does not show a lack of fit. The alternative hypothesis is that the model does show a lack of fit (Box et al., 2015).

To calculate the test statistics Q for the Ljung-Box test, the following equation is used:

Equation 18

$$Q(m) = n(n+2)\sum_{j=1}^{m}\frac{r_j^2}{n-j},$$

Where, r_j is the accumulated sample autocorrelation and m is the time lag. We reject the null hypothesis and conclude that the model shows a lack of fit if:

Equation 19

$$Q > \chi^2_{1-\alpha,h}$$

Where, $\chi^2_{1-\alpha,h}$ is the value in the chi-square distribution table for α significance level and h degree of freedom.

Our results from the Ljung-Box test shows that three ARIMA best fit models for NQROBO has a lower p-value than 0,05. There was zero ARIMA best fit models for NASDAQ and one ARIMA best fit model for S&P 500 with a p-value below 0,05. This means that we cannot reject the null hypothesis of the residuals being white noise for almost every ARIMA best fit model.

5. Construction of The Synthetic Portfolios

Below you will find a visual representation of the investment process for the RMSE portfolio. We have included this example to help the reader to better understand the different steps involved when investing in the portfolio.

Figure 18: The steps involved in the investment process for Portfolio 4

Step 1: Forecast excess returns for all indices

↓

Step 2: Choose the model with the lowest RMSE and positive predicted return

↓

Step 3: Weight the portfolio based on RMSE

↓

Step 4: Invest in the portfolio

Finding the weights of portfolio 5 with new reference weights – 70% and 30%

Equation 20

$$\lambda_{Index}^{SR\ Hybrid} = \frac{SR_{Index}}{\sum SR} * 0{,}7 + \frac{1}{3} * 0{,}3$$

Finding the weights of portfolio 6 with new reference weights – 70% and 30%

Equation 21

$$\lambda_{Index}^{RMSE\ Hybrid} = \frac{RMSE_{Index}^{-1}}{\sum RMSE^{-1}} * 0{,}7 + \frac{1}{3} * 0{,}3$$

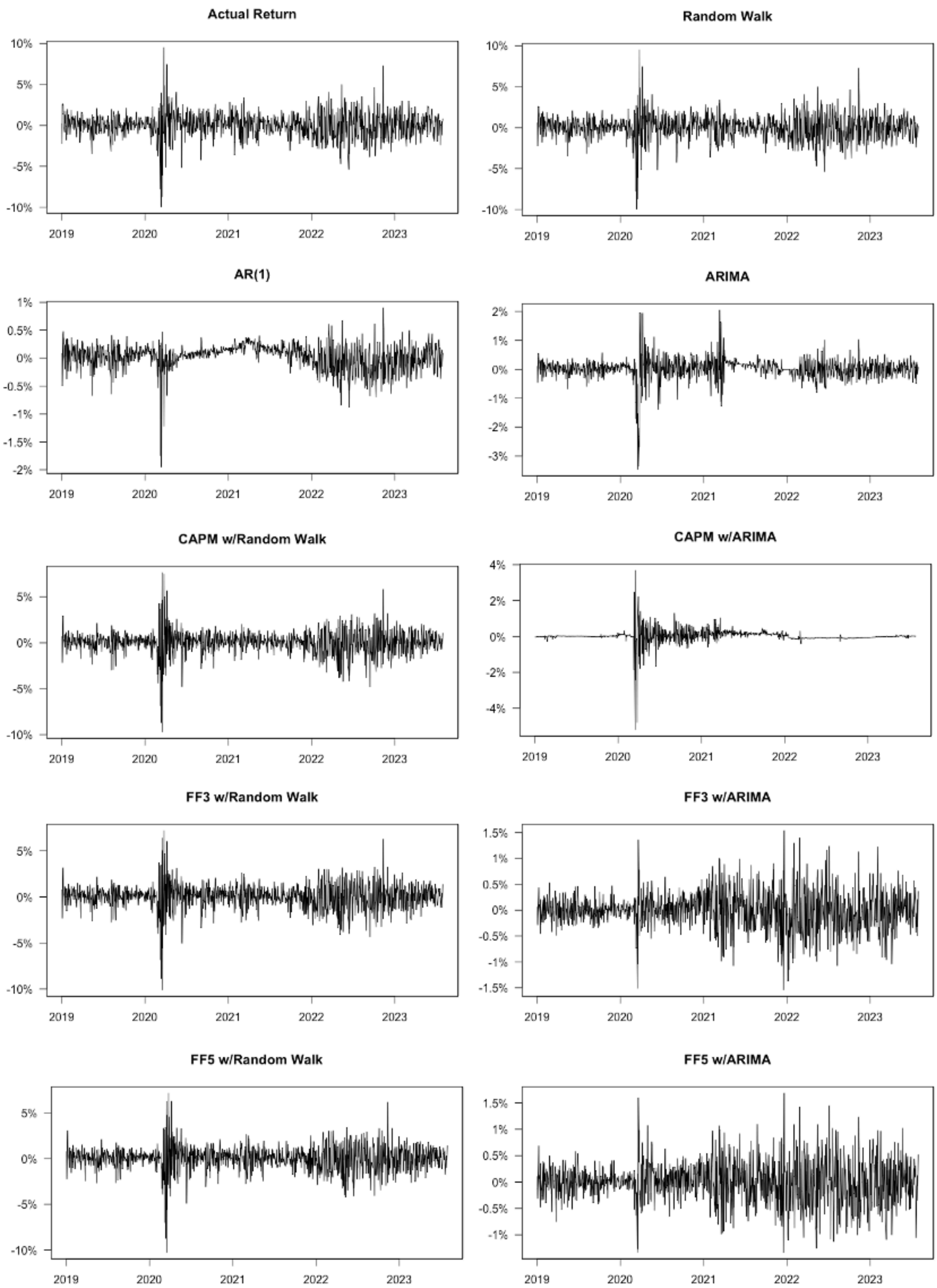
8. Forecasting Results

NQROBO Forecasting

Figure 19 presents the daily forecasting results for the NQROBO index. The actual daily excess returns of the NQROBO index over the forecast period are shown at the top left. Compared to the actual returns, the volatility of the predictions is much smaller.

Figure 19: NQROBO forecasting of daily excess returns

Note: Excess returns relative to risk-free T-Bill rate



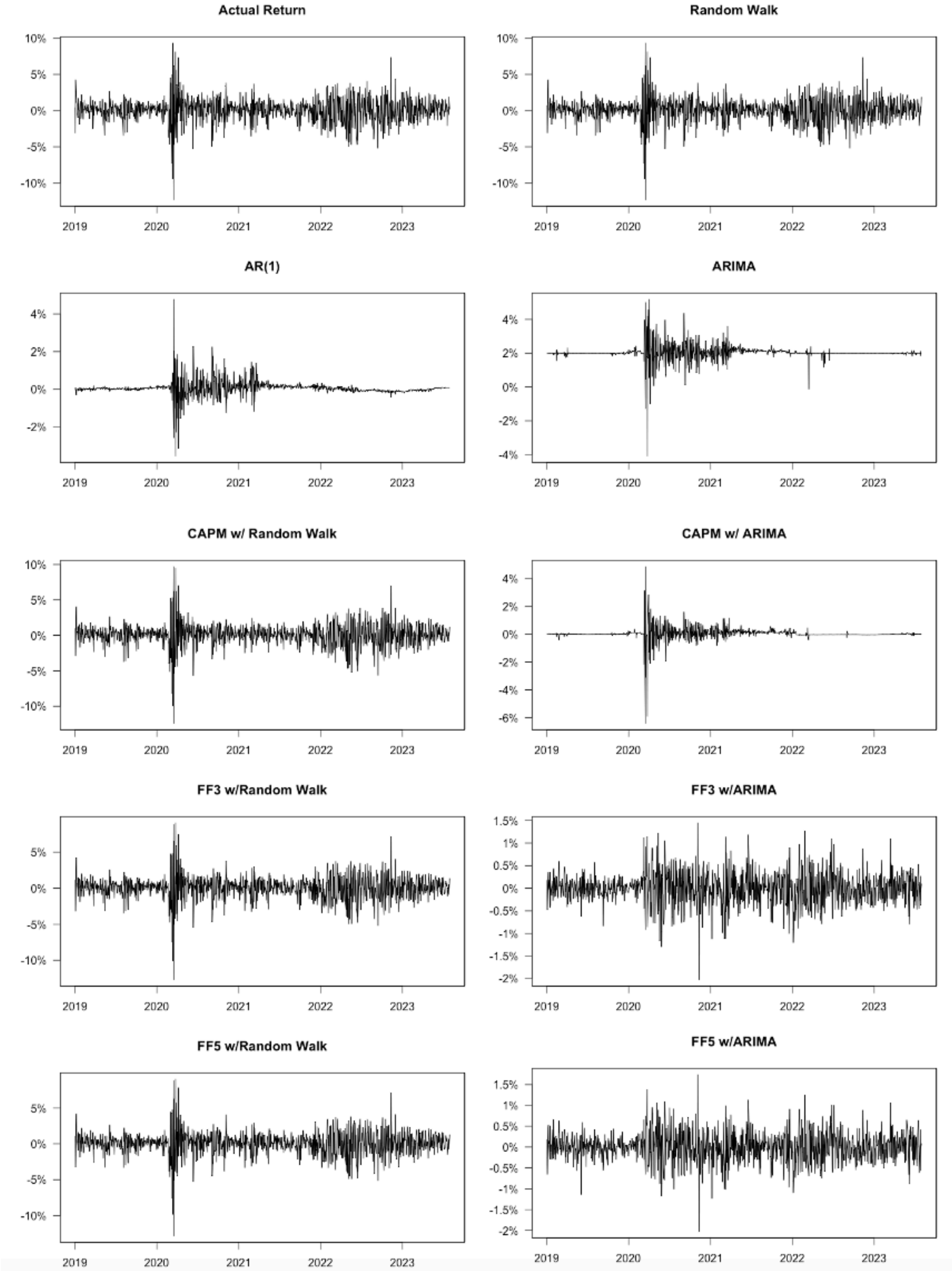
The middle figures show the CAPM model with ARIMA best fit market forecasting for a one day horizon. We can see that because the market forecasts are almost always zero from the beginning to March 2020, and from 2022 and onwards, the estimations are only based on the Alpha in this part of the forecasting period.

NASDAQ Forecasting

The NASDAQ index's forecasting estimation and actual excess return is shown in Figure 20. The volatility of the returns was higher in 2020 and 2022, compared to 2019, 2021 and 2023.

Figure 20: NASDAQ forecasting of daily excess returns

Note: Excess returns relative to risk-free T-Bill rate



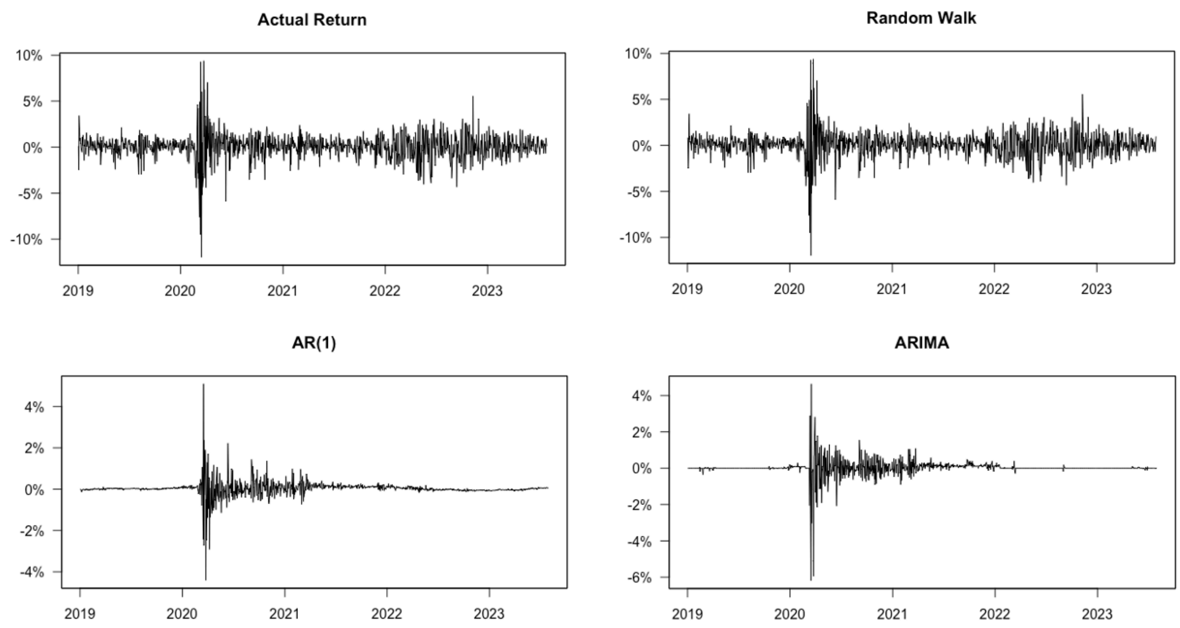
It is interesting to compare the AR(1)-model forecasts for NASDAQ with the corresponding forecasts for the NQROBO. We can see that the volatility was similar for the first year, but after March 2020, the volatility was much higher for the NASDAQ predictions. This difference continues until the end of 2020 and then returns to being similar.
If we look at the ARIMA best fit forecasts, it is the opposite. The volatility of NASDAQ predictions is lower than NQROBO predictions in 2019, 2021, 2022 and 2023. In 2020, the volatility is similar. The CAPM forecasts have similar patterns to the corresponding NQROBO forecasts, but the CAPM with Random Walk market forecasts tends to estimate higher NASDAQ returns. For CAPM with ARIMA best fit market forecast, the same problem with zero market estimations is present. There are some differences between the FF3/FF5 models with Random Walk predictions and FF3/FF5 models with ARIMA best fit predictions regarding trends and volatility.

S&P 500 Forecasting

The actual daily excess returns and the predicted returns of the S&P 500 are shown in Figure 21.

Figure 21: S&P 500 forecasting of the daily excess returns

Note: Excess returns relative to risk-free T-Bill rate



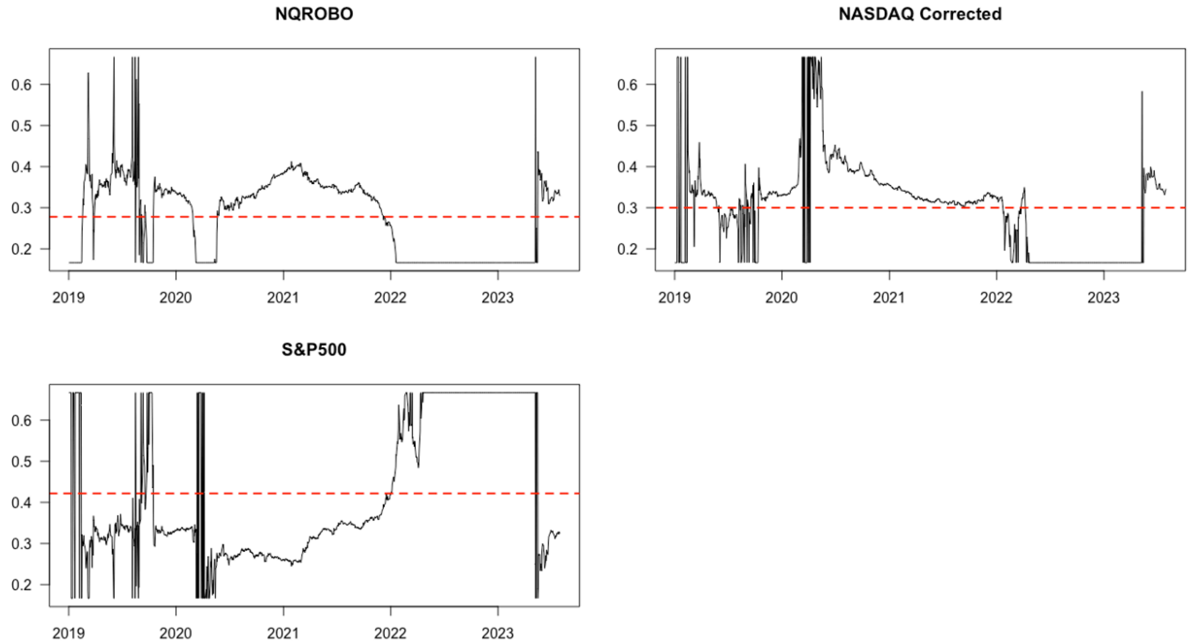
The volatility of the AR(1) forecasting and the ARIMA best fit forecasting of the S&P 500 index has similar patterns. In 2019 and after 2022, the estimated excess return is close to zero and in 2020-2021 the volatility has a downward trend.

One Day Investment Horizon

The weights of Portfolio 5 and 6 range from 17% to 67%. Figure 22 present the weighting variations in Portfolio 5 for all investment horizons. On average, the weight of NQROBO is 28%, 30% in NASDAQ, and 42% in S&P 500.

Figure 22: Portfolio 5 weighting in the indices over time, all horizons

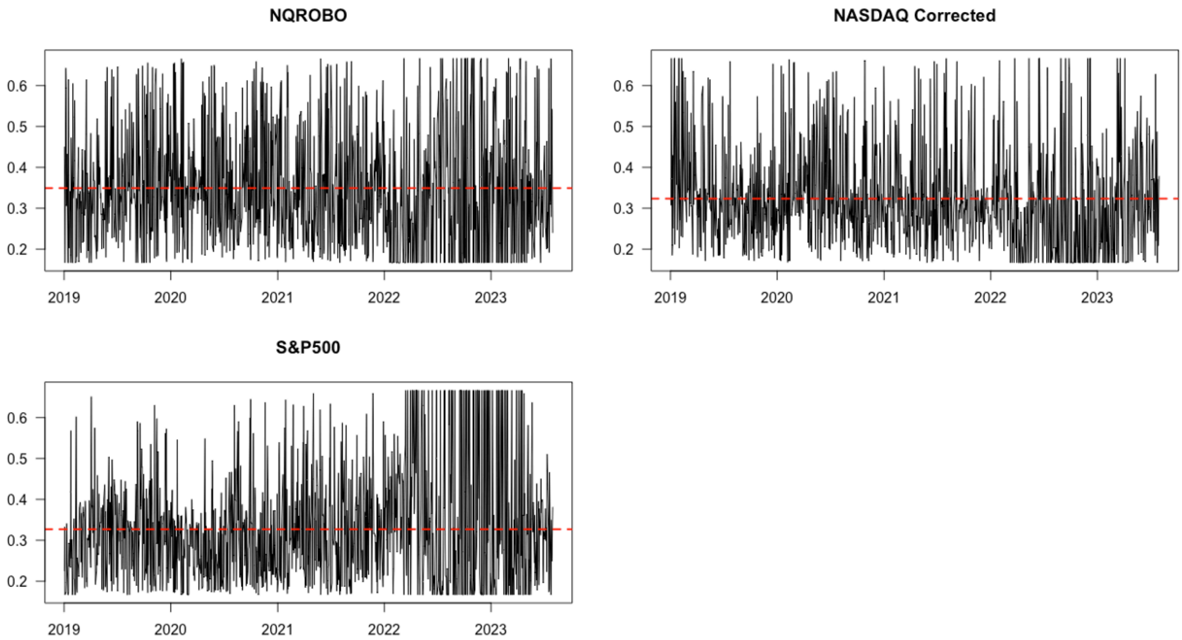
Note: P5 = Hybrid Sharpe Ratio
Red dotted lines visualize the average weights over the whole period



In the following figure, the weights of Portfolio 6 with one day investment horizon are shown. The average weight for the NQROBO index is 35%, 32% for the NASDAQ Corrected, and 33% for the S&P 500.

Figure 23: Portfolio 6 weighting in the indices over time, 1-day

Note: P6 = Hybrid RMSE
Red dotted lines visualize the average weights over the whole period

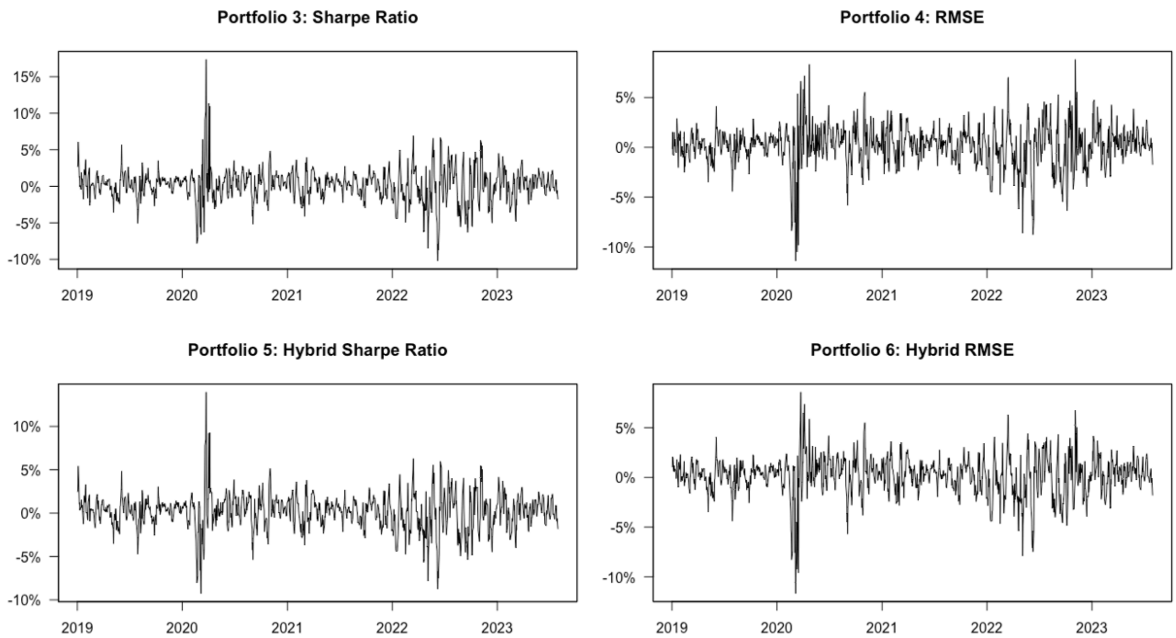


One Week Investment Horizon

In this section we present the empirical results from our analysis with a one week investment horizon. Figure 24 shows the weekly returns of the dynamic portfolios. The fluctuations are greater compared to the daily returns.

Figure 24: Weekly excess return for our dynamic portfolios

Note: Returns relative to risk-free T-Bill rate



The weights for Portfolio 4 and Portfolio 6 are presented in Figure 25 and 26. For the NQROBO index and the S&P 500 index, the average weight in this period is 33% for both portfolios. The average weight in NASDAQ Corrected in both portfolios is slightly higher, 34%.

Figure 25: Portfolio 4 weighting in the indices over time, weekly

Note: P4 = RMSE
Red dotted lines visualize the average weights over the whole period

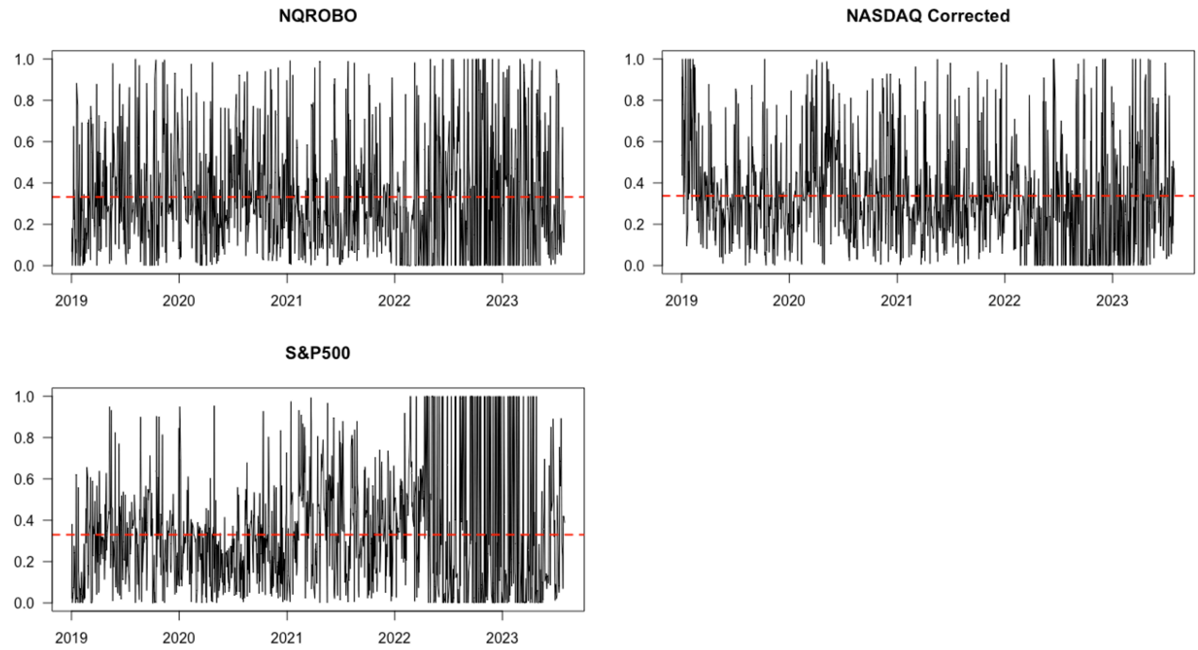
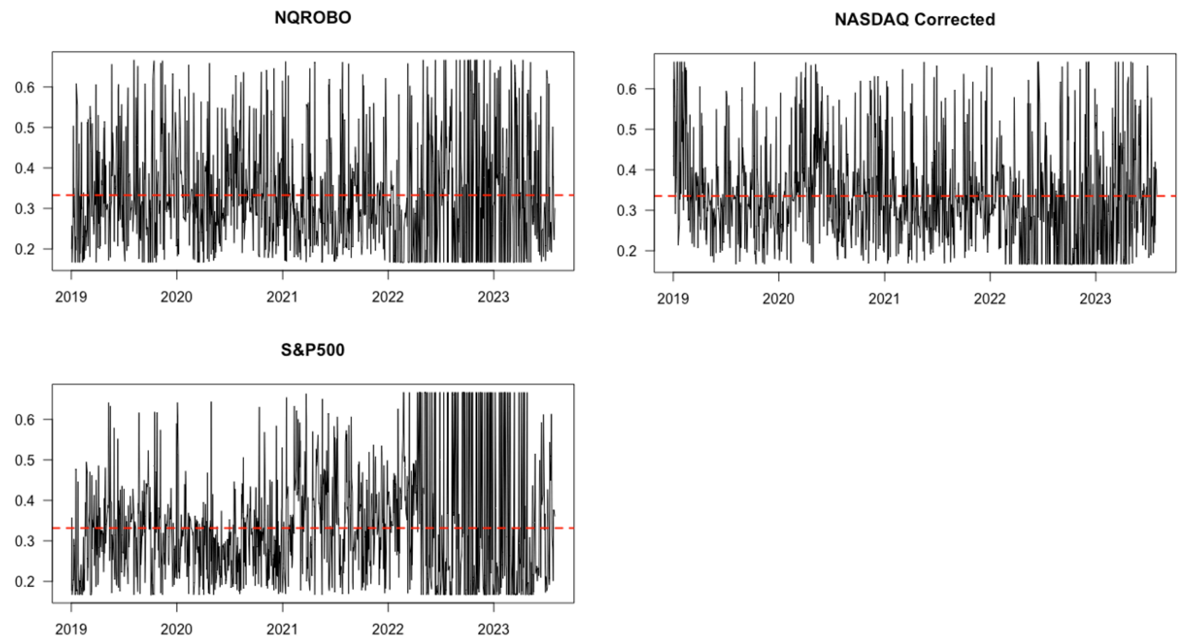


Figure 26: Portfolio 6 weighting in the indices over time, weekly

Note: P6 = Hybrid RMSE
Red dotted lines visualize the average weights over the whole period



One Month Investment Horizon

In this section we present the empirical results from our analysis with a one month investment horizon. Figure 27 shows the excess returns of the dynamic portfolios. We can observe a higher degree of volatility compared to the daily and weekly returns.

Figure 27: Monthly excess return for our dynamic portfolios

Note: Returns relative to risk-free T-Bill rate

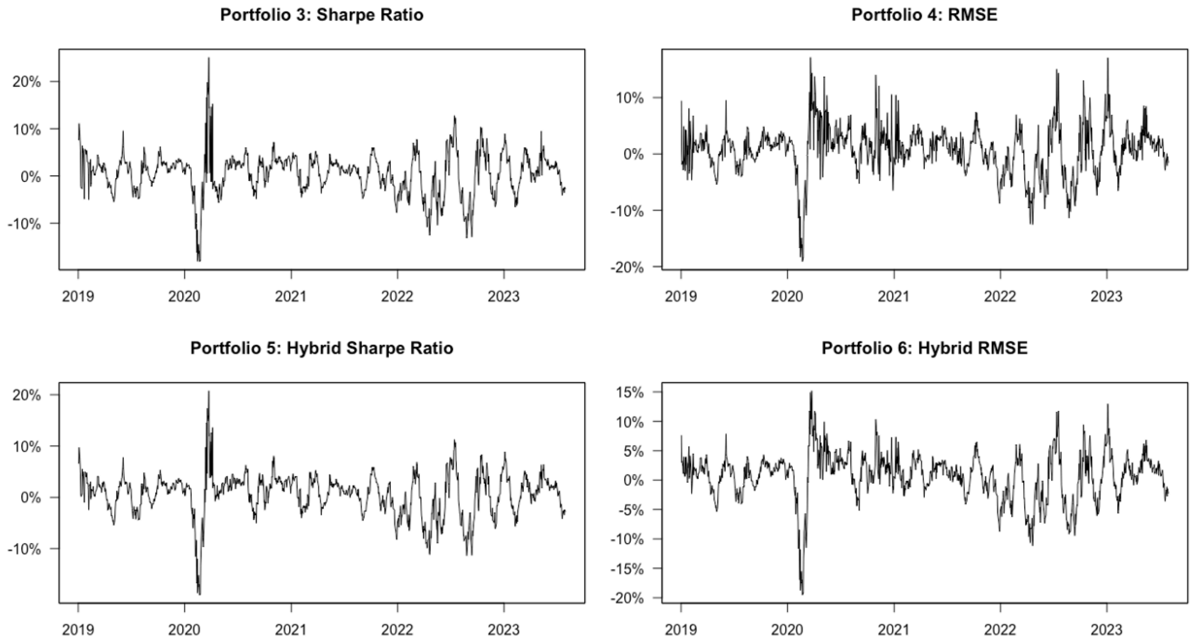


Figure 28 and 29 shows that the average weight of NQROBO in Portfolio 4 and 6 is 34%, while the average weight for NASDAQ Corrected and S&P 500 is 33%.

Figure 28: Portfolio 4 weighting in the indices over time, monthly

Note: P4 = RMSE
Red dotted lines visualize the average weights over the whole period

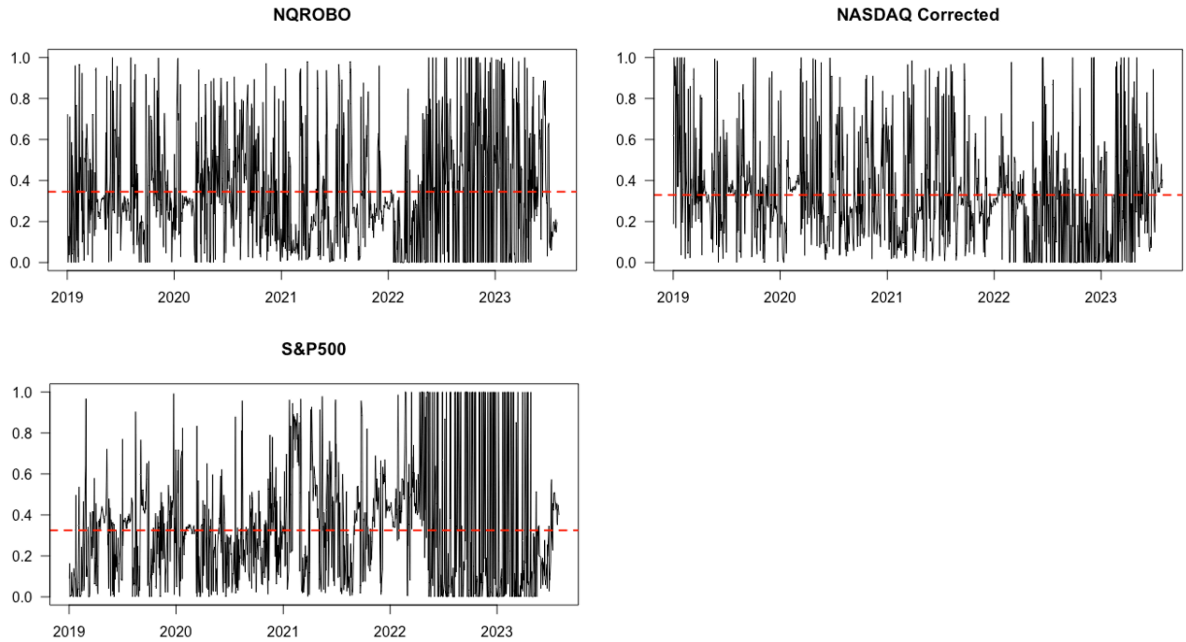


Figure 29: Portfolio 6 weighting in the indices over time, monthly

Note: P6 = RMSE Hybrid
Red dotted lines visualize the average weights over the whole period