

Bachelor Thesis in Finance Spring 2024

Does Cybersecurity Risk Impact Stock Returns of Nordic Large Cap Firms?

Cedric Bratt (25295)
Axel Söderberg (25222)
June 2, 2024

Does Cybersecurity Risk Impact Stock Returns of Nordic Large Cap Firms?

Abstract :

In this study, we investigate the impact of cybersecurity risk on the stock returns of Nordic large-cap firms by employing textual analysis on firms' annual reports. Inspired by methodologies used in prior research on U.S. firms, we develop a cybersecurity risk score tailored to the Nordic market. We then create three cybersecurity risk-sorted portfolios and a long-short portfolio that goes long in low-risk cybersecurity stocks and short in high-risk cybersecurity stocks. Our findings indicate that cybersecurity risk does not impact Nordic large-cap stock returns, with any differences in the generated returns of our cybersecurity risk-sorted portfolios being explained by market risk, size, and value factors. Additionally, we provide descriptive evidence that larger firms with high price-to-book ratios are more likely to be subject to cyberattacks. Furthermore, we evaluate the cybersecurity risks of specific industries and how these risks have developed over time. Our study contributes to the growing field of research on cybersecurity's impact on financial assets and provides first-of-its-kind evidence in the Nordic market.

Keywords :

Cybersecurity, Textual Analysis, Efficient Market Hypothesis, Quantitative Research

Authors :

Cedric Bratt (25295), Axel Söderberg (25222)

Tutor :

Ye Zhang, Department of Finance

Examiner :

Adrien d'Avernas, Department of Finance

Bachelor Thesis
Bachelor Program in Business and Economics
Stockholm School of Economics
© Cedric Bratt and Axel Söderberg, 2024

Contents

1 Introduction **3**

2 Theoretical Background **4**

2.1 Jaccard Similarity Coefficient 5

2.2 Cosine Similarity 5

2.3 Efficient Market Hypothesis 6

2.4 Capital Asset Pricing Model 6

2.5 Fama and French Three Factor Model 7

2.6 Related Literature 7

2.6.1 Cybersecurity Risk Related 7

2.6.2 Textual Analysis Related 9

3 Data **9**

4 Methodology **10**

4.1 Construction of Cybersecurity Risk Score 10

4.2 Analyzing Financial Returns 11

5 Results **12**

5.1 Cybersecurity Risk Over Time and across industries 12

5.2 Financial Performance of Cybersecurity Risk Sorted Portfolios 14

6 Summary and interpretation **18**

6.1 Interpretation of Descriptive Evidence 18

6.2 Interpretation of Cybersecurity Risk Impact on Stock Returns 18

6.3 Study Limitations 19

6.3.1 Geographical Restriction and Small Sample Size 19

6.3.2 Lack of Cybersecurity Risk Disclosures in Nordic firms 20

6.3.3 Possible Inaccuracies of our Cybersecurity Risk Score 20

6.4 Further Research and Possible Improvements 21

6.4.1 Further Research 21

6.4.2 Possible Improvements of Our Cybersecurity Risk Score 21

7 Concluding remarks **22**

8 AI Transparency Statement **26**

9 Appendix **26**

1 Introduction

Cyberattacks have more than tripled in number over the last decade and doubled since the onset of the COVID-19 pandemic globally and cybersecurity is raised as a key financial vulnerability in 2024 by the IMF (2024). These attacks are not only increasing in frequency but also in severity, with potential damages reaching up to U.S.D 2.5 billion according to IMF (2024). Scandinavia is not spared with the Swedish government (2024) prioritizing strengthening cybersecurity in 2024 and the EU implementing new measures that puts higher demand on how organization act and protect themselves against digital threats (Finansdepartementet, 2024). Notably, the number of cyberattacks in Sweden in 2023, targeting critical infrastructure and organizations, reached approximately 100 - equivalent to the total number of cyberattacks recorded in 2021 and 2022 combined according to MSB (2024). Moreover, Scandinavia has witnessed significant cyber incidents during the last decade; for instance, the 2017 attack on the global shipping giant AP Moller-Maersk resulted in over U.S.D 300 million in losses and widespread disruptions to shipping operations worldwide according to Leovy (2017). More recently, Tieto Evry reported an attack that severely impacted healthcare services across Sweden, forcing several regions into emergency mode and some to revert temporarily to paper records to maintain patient care (Lindh, Sundström, 2024). These events underscore the vast socioeconomic impacts of cyberattacks, affecting not only organizations but also society as a whole. Given the sharp increase in cybersecurity risks and the limited existing literature of the impact of cybersecurity risks on the financial market, further investigation into this area is both timely and essential.

Due to the relatively recent escalation of cybersecurity threats to organizations, it remains uncertain whether the stock market can accurately price in cybersecurity risk. Organizations not only handle these threats differently but also vary widely in how they communicate them. Additionally, the regulatory landscape is diverse globally. The U.S. has taken a leading role by requiring public companies to disclose both material cybersecurity incidents they experience and, on an annual basis, provide substantial information regarding their cybersecurity risk management, strategies, and governance, according to SEC (2023). In contrast, the EU is noticeably behind, with no clear explanation given that threat levels are presumed similar across regions. This is particularly striking when compared to the EU's Corporate Sustainability Reporting Directive ("CSRD"), which mandates all public companies to report risks arising from social and environmental issues and was enforced starting in 2024 (Jakobsson, 2024). The absence of a similar mandate for cybersecurity threat disclosures could leave EU investors less informed about the cybersecurity risks public companies face, potentially impairing their ability to price these risks accurately. This gap in regulatory requirements and information disclosure might threaten the Efficient Market Hypothesis, by keeping critical risk information from investor making them unable to price in cybersecurity risks.

The literature on the impact of cyberattacks on the financial performance of public companies remains limited, although it is a fast growing field due to the recognition of the devastating effects cyberattacks can have on a company and society. However, to the best of our knowledge there has been no study investigating the impact of cybersecurity risk on the financial returns of public Nordic companies. The studies in this subject have been published during the last three to four years and almost entierly focused on the U.S. market. Given the increasing awareness and cybersecurity threats, we believe it is highly relevant to attempt to quantify the potential financial impacts that cybersecurity risks can have on Nordic companies. Even though quantifying cybersecurity risks in organizations presents challenges due to their complex and technical nature, we adopt a textual analysis approach, which has gained popularity in recent years and proven useful when analyzing hard-to-estimate factors or variables. After assigning public companies different cybersecurity risk scores it will be possible to evaluate how high risk compared

to low risk companies perform financially over time. Furthermore, we also aim to develop insights of the shared characteristics of Nordic high cybersecurity risk firms. Finally, we hope to form an indicative conclusion regarding how cybersecurity risk impacts the financial performance of public Nordic firms.

In this study, we investigate cybersecurity risk within the Nordic large-cap market, aiming to quantify firm-level exposure and discern its influence on stock returns. Our methodology draws inspiration from Florackis, Louca, Michaely and Weber (2023) which utilizes a textual analysis approach to quantify public U.S. firms' cybersecurity risk and how it correlates with financial returns. By examining firms' which have been subject to a cyberattack annual reports and comparing them to other firms which have not been attacked, we construct a measure of cybersecurity risk of Nordic large cap firms. The measure builds on the idea that firms which have been attacked have increased cybersecurity risk and express this increased risk in their annual report. Moreover, firms with similar cybersecurity risk levels utilize similar wording to firms which have been subject to a cyberattack - thus allowing us to construct a cybersecurity risk score.

By sorting stocks based on their cybersecurity-risk scores and constructing three portfolios associated with varying levels of cybersecurity risk exposure we explore the pricing of cybersecurity risk in the cross-section of stock returns within the Nordic large-cap market. We also test the statistical robustness of our results by employing Fama-French Three Factor Model (1993) and the Capital Asset Pricing Model (CAPM) (1962) regressions in order to control for known risk factors which can impact firms' stock returns. Our study is considered to be of interest to investors and policymakers that ponder about the financial impact of cybersecurity risk. In addition, our study also provides information about what type of firms that are most likely to be attacked and should therefore also be of interest to companies wanting to estimate their cybersecurity risk. It also contributes to the growing field of textual analysis in important but hard to estimate subjects.

Our results indicate that a portfolio strategy that includes going long in stocks with lower cybersecurity risk score and short in stocks with higher cybersecurity risk score outperforms the market, yielding an alpha of 0.01 under the CAPM and an alpha of 0.002 per year under the Fama-French Three-Factor Model, although the latter alpha was non-significant. Furthermore, we observed that cybersecurity risk scores has increased over time during our research period. Among the sectors analyzed, industrial firms experienced the most attacks, while the energy sector exhibited the highest cybersecurity risk scores. Additionally, we found that firms with higher price-to-book ratios and higher market capitalizations generally have a higher cybersecurity risk score than their lower counterparts. These findings highlight the importance of cybersecurity risk management, offering useful insights for investors and decision-makers in the Nordic large-cap market.

2 Theoretical Background

In this section we present the most relevant concepts and literature for conducting the study which also serves to add understanding in explaining the results. The concepts are the Jaccard Similarity Coefficient, Cosine Similarity, Efficient Market Hypothesis, Capital Asset Pricing Model (CAPM), and the Fama and French three factor model. Furthermore, this study also draws heavily on the paper from Florackis et al. (2023), which also will be discussed briefly.

2.1 Jaccard Similarity Coefficient

The Jaccard Similarity Coefficient, also referred to as the intersection over the union, was introduced by Paul Jaccard in 1901 as a similarity metric. As the name implies it calculates the size of the intersection between two vectors A and B divided by the size of the union of these two vectors. The Jaccard measure is binary which means that each item word in the vector is counted only once. The resulting score ranges from 0 to 1, where 0 signifies no similarity between the vectors, and 1 represents that the two vectors are equal as the intersection is equal to the union. The equation for the Jaccard similarity measure is displayed below:

$$\text{Jaccard Similarity}(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Jaccard similarity is commonly used in textual analysis due to its effectiveness in comparing binary vectors. In this context, vectors A and B represent embedded word vectors, indicated by 1s and 0s. This binary representation ensures that a word present in both vectors contributes to both the intersection and the union, whereas a word present in only one vector contributes only to the union. The binary nature of the Jaccard measure makes it ideal for cases where the presence of words is more significant than their frequency, thereby aligning with binary data comparisons.

Furthermore, Jaccard similarity has been employed in the financial sector for textual analysis purposes. In the study by Cohen, Malloy, and Nguyen (2023) the Jaccard similarity coefficient is utilized to demonstrate that U.S. firms that make changes to their 10-K filings are more likely to experience declines in earnings, reduced profitability, and a higher risk of bankruptcy. Conversely, firms that maintain consistency in their filings tend to achieve positive abnormal returns (Cohen, 2023). This highlights the potential and suitability of utilizing the Jaccard similarity to identify meaningful patterns in textual analysis of corporate communications and the financial industry.

2.2 Cosine Similarity

Cosine similarity, unlike the Jaccard similarity coefficient, quantifies similarity between two vectors A and B by taking the dot product between the two vectors normalized by their vector lengths, represented by ||A|| and ||B|| which are the Euclidean norms of the vectors. Geometrically, the dot product of two vectors A and B is the product of the magnitudes of the vectors and the cosine of the angle θ between them. It is defined as follows:

$$A \cdot B = |A||B| \cos\theta$$

This operation results in a scalar quantity that has magnitude but no direction. In other words, the formula successfully measures the the cosine of the angle between two vectors and determines whether two vectors are pointing roughly in the same direction within an inner product space. The resulting score ranges from 0 to 1 with positively correlated similarity scores. In contrast to Jaccard similarity, cosine similarity is not binary and puts emphasis on the frequency of words and not only the presence of them. The formula for cosine similarity is:

$$\text{Cosine Similarity}(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|}$$

Cosine similarity is commonly applied in information processing and textual analysis. It can be used as a standalone metric within the intersection of finance and textual analysis as in Hanley and Hoberg (2010) when it is used to show that greater informative content in an IPO prospectus is likely to result in more accurate offer prices and less under pricing. It can also be used together with the Jaccard similarity score as the the combination of the two scores allows for both a binary and non-binary interpretation when

calculating the similarity between two vectors, as is done in Florackis et al. (2023).

2.3 Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH) translates into that in a fully efficient market, asset prices accurately reflect all available information, leaving no room for investors to consistently outperform the market by exploiting undervalued or overvalued securities (Fama, 1970). Under the EMH, stock returns are proportional to the level of risk associated with each stock, meaning that investors can expect higher returns only if they accept higher levels of risk (Fama, 1970).

Incorporating the Efficient Market Hypothesis into this study is important for framing our analysis of how cybersecurity risk affects asset prices. By examining the validity of EMH in this context, we aim to determine whether the market efficiently integrates information related to cybersecurity into stock prices, or if discrepancies exist that could lead to market inefficiencies. This examination is especially important given the findings of Florackis et al. (2023), which indicated that the EMH holds and that cybersecurity risk is priced into stock returns. By investigating the alignment between EMH and the real-world behavior of Nordic large cap firms' stock prices in response to cybersecurity risks, this study seeks to uncover potential existing market dynamics.

2.4 Capital Asset Pricing Model

The Capital Asset Pricing model (CAPM), introduced by Treynor (1962), Sharpe (1964), Lintner (1965) and Mossin (1966), provides a method for estimating expected returns by considering the systematic risk of individual securities, captured by their beta coefficients, and the alpha which indicates the over or underperformance of a given portfolio. In equilibrium, the CAPM suggests that the market portfolio, represents the most efficient portfolio in terms of maximizing returns relative to risk. Consequently, the CAPM implies that there are no portfolios offering higher risk-adjusted returns than the market portfolio, and that beta is the sole determinant of expected returns for individual securities. The CAPM is captured by the following formula:

$$E(R_i) = R_f + \beta_i \times (E(R_m) - R_f)$$

Where:

- $E(R_i)$ represents the expected return of the investment.
- R_f represents the risk-free rate of return.
- β_i represents the beta of the investment, indicating its volatility relative to the market.
- $E(R_m)$ represents the expected return of the market.
- $(E(R_m) - R_f)$ is known as the market risk premium, which is the additional return investors demand for taking on risk when investing in stocks compared to a risk-free asset.

Including CAPM in this study allows us to control for market risk and measure the expected returns on assets given their cybersecurity risk aligned with the notion that higher risk should be compensated by higher returns. In other words, high cybersecurity risk stocks should have higher expected returns if a cybersecurity risk premium exists. If it does not exist Beta should be sufficient in explaining a portfolio's expected return.

2.5 Fama and French Three Factor Model

The Fama and French three-factor model, developed by Fama and French (1993), was designed to enhance the Capital Asset Pricing Model (CAPM). They recognized that returns are influenced by multiple dimensions of risk beyond just market beta and introduced two additional factors: size and value. Their model was based on earlier studies, such as Banz (1981) and Rosenberg, Lanstein, and Reid (1985), which identified the "size effect" and the positive correlation between stock returns and Book-To-Market (BTM) equity ratios, respectively. They observed that smaller stocks and stocks with higher BTM ratios systematically yielded higher returns than what CAPM would predict. To account for these dimensions, they introduced the "small-minus-big" (SMB) and "high-minus-low" (HML) factors, demonstrating that these additional factors could better predict expected returns. In our thesis, we aim to use the Fama-French model to control for these known risk factors and investigate whether cybersecurity risk is priced into stock prices, providing more robust results than just utilizing a single factor model.

$$E(R_i) = R_f + \beta_{\text{mkt}}(E(R_m) - R_f) + \beta_{\text{smb}}\text{SMB} + \beta_{\text{hml}}\text{HML}$$

Where:

- $E(R_i)$ represents the expected return of the investment.
- R_f represents the risk-free rate.
- β_{mkt} represents the sensitivity of the expected return to the expected excess market return.
- $E(R_m) - R_f$ represents the market risk premium.
- β_{smb} represents the sensitivity of the returns to the size premium.
- SMB (Small Minus Big) represents the size premium, calculated by subtracting the average return of the low market-cap stocks from the average return of the high market-cap stocks.
- β_{hml} represents the sensitivity of the returns to the value premium.
- HML (High Minus Low) represents the value premium, calculated by subtracting the average return of the low Book-To-Market stocks from the average return of the high Book-To-Market stocks.

2.6 Related Literature

2.6.1 Cybersecurity Risk Related

Our thesis contributes to the small but fast growing field of studies investigating cybersecurity's impact on firms' stock returns which includes Florackis et al. (2023), Jiang, Khanna, Yang and Zhou (2020), Helene, Rustam and Tahoun (2023), and Jungmin, Milidonis, Stulz, Kamiya and Jun-Koo (2021).

We primarily utilized the study from Florackis et al. (2023) as a methodological outline in constructing and applying our cybersecurity risk score. The paper utilizes a training set of firms which have been subject to a cyberattack and information in cybersecurity risk disclosures in annual 10-K filings of U.S. listed firms to construct cybersecurity risk proxies. Florackis et al. (2023) showed that firms with high cybersecurity risk tend to outperform those with low cybersecurity risk in both one-factor and three-factor models - indicating that investors that purchase high cybersecurity risk stocks demand higher returns. Thus, their results proving that the Efficient Market Hypothesis still holds and that cybersecurity risk is priced in stock returns. Furthermore, their findings indicate that high risk cybersecurity firms perform poorly in times of heightened cybersecurity risk and better in low risk periods, suggesting that investors

demand a premium for holding high risk cybersecurity stocks. Florackis et al. (2023) also observed that larger firms tend to be targeted more frequently, being assigned higher cybersecurity risk scores. In addition, their average cybersecurity risk score has increased over time in line with rising number of observed cyberattacks suggesting that cyberthreats have increased.

Jiang et al. (2020) employs a machine learning approach that leverages a wide range of firm characteristics to determine the cybersecurity risk of a given firm. Utilizing data from July 2008 to June 2019, they conduct Fama and MacBeth (1973) cross-sectional regressions to analyze the impact of cyber risk on stock returns. Their findings indicate that a cybersecurity risk premium exists and is factored into market pricing. Notably, during periods of heightened cybersecurity risk, as indicated by New York University's Index of Cyber Security (ICS), stocks with high cybersecurity risk tend to underperform the market. This underperformance underscores the market's sensitivity to fluctuations in perceived cyber threats.

In contrast to previous studies, Tahoun et al. (2021) utilizes quarterly earnings' call data spanning from 2002 to 2021 comprimising more than 13,000 firms in order to construct a firm-level cybersecurity risk score. Firstly, their result provide an approximate calculation that cybersecurity risk exposure amounts to $226 billion in lost net income per year globally highlighting the importance of this topic. Secondly, they show that cybersecurity risk is priced in the option market. Thirdly, the results indicate that firm-level cybersecurity risk can have systemic implications as there is a negative spill over effect from the company subject to the cyberattack to other companies in the same industry and country. In conclusion, their study highlights the scalability of utilizing textual analysis to analyze hard-to-value risks and the impact cybersecurity risk has on stock returns and the potential systemic effects.

In Kamiya's et al. (2021) study, which involves firms with a mean market value of approximately $58.93 billion, they discover that cyberattacks involving financial information loss lead to an average value loss of over $1.06 billion per attacked firm, compared to those without such losses. Their findings also reveal total out-of-pocket for their sample firms amount to $1.72 billion, with mean and median out-of-pocket costs of $82.01 million and $16 million, respectively. Thus, Kamiya et al. (2021) conclude that out of pocket expenses are not responsible for shareholder value losses. However, their analysis also showed that stock prices do not, on average, react strongly to cyber incidents with the mean negative market movement of -0.23% when there is no financial information loss, with the corresponding number being -1.09% when there is a loss of financial information. Furthermore, the finance, consumer industries, and manufacturing sectors emerge as the most at-risk, indicating that firms experiencing cyberattacks are generally larger, older, and more frequently listed among the Fortune 500 compared to those not attacked. These targeted firms also exhibit higher profitability (higher return on assets), lower risk (lower stock return volatility), greater future growth opportunities (higher Tobin's q), higher leverage, and higher asset intangibility but invest less in capital expenditures and research and development (Kimaya et al. 2021).

Another study also investigating the characteristics of breached firms is Lending, Minnick and Schorno (2018). They found that firms with higher standards of governance and social responsibility — specifically those with smaller, more financially expert boards, and higher rankings in environmental and product safety — are less likely to suffer data breaches. They do this by utilizing data from 271 U.S. firms that experienced breaches between 2004 and 2012 and financial records from the Privacy Rights Clearinghouse to assess the impacts of data breaches on firm performance and governance changes. Moreover, Lending et al. (2018) found that breached firms who improve their social and governance standards are less likely to be attacked compared to firms that do nothing. Furthermore, Lending et al. (2018) indicate that breached firms show an immediate stock return decline of -3.5% and longer-term operational performance setbacks,

such as significant declines in deposits for banks and in sales for non-banking entities, highlighting the significant negative effects of a data breach.

In contrast to the above litterature, Hilary, Segal and Zhang (2016) examines the effects of cyberattacks on publicly listed U.S. firms. Using the Chronology of Data Breaches database, which lists privacy data breaches from 2005 to 2014, they observe that the market reaction to such breaches is minimal, with a median three-day period market-adjusted abnormal return of approximately -0.5%. This muted response suggests that the market does not view these events as significantly damaging to firm value. Interestingly, despite an increase in disclosure following a breach, this is consistent with a general trend of increased disclosures across all firms, indicating no unique response to cyber incidents. Therefore, Hilary et al. (2016) points out that it does not seem like firms' subject to a data breach increase their cybersecurity risk disclosures. In addition, their analysis extends to long-term returns, finding no significant long-term effects post-breach, suggesting that the market neither overreacts nor underreacts to such incidents and that cybersecurity risk is accurately priced into stock returns.

In addition, the study by Jing et al. (2023) found that firms experiencing a data breach tend to increase their amount of cybersecurity risk factor disclosures compared to matched firms without a data breach. Interestingly, while increased disclosures in breached firms' subsequent annual reports do not lead to significant market reactions, a significant negative market reaction occurs if breached firms do not change/decrease their cybersecurity risk factor disclosures, regardless of the breach's severity. This suggests that the market expects increased disclosures following data breaches and reacts negatively when expectations are not met.

2.6.2 Textual Analysis Related

Our methodology of determining cybersecurity risk scores for each firm relies on the use of textual analysis. In the Florackis et al. (2023) study, a comparable technique was employed to scan textual documents for the frequency of predetermined cybersecurity keywords. To further assess the suitability of such an approach for our purposes, we examined other studies. Loughran and McDonald (2011) utilized various textual analysis techniques to test the predictive power of textual measures on financial outcomes such as stock returns and volatility. Their findings indicated that textual analysis of 10-K filings, engineered to search for finance-related keywords, can provide insights that predict financial outcomes.

In addition to Loughran and McDonald (2011), the paper by Gentzkow, Kelly, and Taddy (2019) explores textual analysis using more advanced methodologies, including deep learning and neural networks. Their research demonstrates how textual analysis can uncover insights into human behavior and decision-making, extending beyond financial applications to political economy and labor economics.

Although our methodology differs from that of Loughran and McDonald (2011) and Gentzkow et al. (2019), their work highlights that textual analysis is a powerful tool for extracting valuable, but hard to measure, information from complex, qualitative data like annual reports.

3 Data

Our thesis covers the five year period from 2018 to 2022 and focuses on the 248 companies listed on the Nasdaq Nordic Large Cap list. We obtained annual reports for each firm from their respective websites in English. However, due to varying reporting standards (e.g., the report only being available in the firm's local language) and some being IPO:d during the time period, a few Nordic large cap firms had to be excluded from our dataset.

For each company, we collected stock price, market capitalization, Book-To-Market ratio, and geographical location of firms' headquarters data from Factset during the study's time period. The stock prices retrieved from Factset represent the price of the stock on the last trading day of each month. To determine the risk-free rate of return, we obtained data from the Swedish central bank (Riksbanken) (2024) during the study's time period and used it as the risk-free rate, also known in Swedish as "styrränta".

Furthermore, in order to construct our training set for our cybersecurity risk score we conducted a desktop search between the years 2017-2021 on every firm present in the Nordic large cap index to investigate whether they have been subject to a publicly disclosed cyberattack or not. Since only publicly disclosed cyberattacks can have a severe impact on stock prices we believed this was a suitable approach. Notably, there were no publicly disclosed cybersecurity breaches in any of the Nordic large cap firms during 2018. Consequently, there is no training set available for that year, resulting in the absence of Jaccard similarity coefficients and Cosine similarity scores for the year 2019.

In our textual analysis of the annual reports, we searched for specific keywords associated with cyberattacks. The list of 451 words used was extracted from the "Vocabulary" list provided by the U.S. governmental agency National Initiative for Cybersecurity Careers and Studies (NICCS) (2024).

We also utilized google search trend data on the word "cyberattack" in Sweden to determine periods when the public displayed heightened cybersecurity risk awarness which was generated by trends.google.com.

4 Methodology

4.1 Construction of Cybersecurity Risk Score

We construct our cybersecurity risk score by firstly downloading the annual reports for each firm belonging to the Nordic large cap index. We then transform the PDF files into analyzable txt files before embarking on our textual analysis.

In our textual analysis, we utilize several packages in R, including tidytext, tm, lsa, and wordcloud, to process and examine the text data. The first step involves transforming the input text into a corpus, a structured set of text documents. This transformation is important as it prepares the data for subsequent preprocessing tasks. Following the creation of the corpus, we embark on a series of text preprocessing steps aimed at cleaning and normalizing the text. This process includes converting all text to lowercase to ensure uniformity, removing punctuation and numbers to isolate words effectively, and eliminating common stopwords from the English language. Stopwords, which are frequently occurring words that carry minimal meaning on their own, are removed to focus the analysis on more significant words. Additionally, we remove any extra spaces from the text to standardize its format.

After preprocessing, the corpus is converted into a term-document matrix (TDM). This matrix quantifies the frequency of terms across the collection of documents, with rows representing terms and columns representing documents. To facilitate analysis, we convert this matrix into a tidy format, a data frame where each row corresponds to a term-document pair, including information on the term, its document, and the frequency of occurrence.

The next phase of our analysis focuses on identifying the frequency of specific cybersecurity-related keywords within the text. We are using a list of 451 cybersecurity keywords provided by NICCS (2024). By filtering the tidy data for these predefined keywords, we can isolate and analyze the occurrences of terms relevant to our study. Additionally, we annotate the data with the document's name or identifier.

We then perform research on news outlets to identify every Nordic large-cap firms that have experienced publicly disclosed cyberattacks between 2017 and 2021. This investigation stems from our hypothesis that only cyberattacks known to the public can influence a company's stock prices, thereby affecting their financial returns. To build our training set, we pinpoint firms that have been publicly recognized as victims of cyberattacks in year t_{N-1}. These firms are subsequently used as our training sample for the following year, t_N. For each year, t_N, we assess the similarity of each firm to those in our training set, utilizing both cosine and Jaccard similarity metrics. The cybersecurity risk score for each firm, for each year, is then calculated as the average of these cosine and Jaccard similarity scores. Thus the higher similarity to our training sample the higher cybersecurity risk score. The equations for calculating the cosine and jaccard cybersecurity risk scores are displayed below:

$$\text{Cybersecurity Risk}_{i,t} = \frac{\sum_{n=1}^{N} CS_{i,n,t}}{N_{t-1}}$$

$$\text{Cybersecurity Risk}_{\text{Jaccard},i,t} = \frac{\sum_{n=1}^{N} JS_{i,n,t}}{N_{t-1}}$$

4.2 Analyzing Financial Returns

We sort companies based on their median cybersecurity risk score over the time period to construct three equally sized portfolios from high to low cybersecurity risk. Then for each stock the monthly return was calculated by taking the price at month T divided by the price at month T_{-1}. Moreover, a long-short portfolio is created by taking portfolio 3's returns minus portfolio 1's returns.

We then calculate the monthly market return by taking the average return of all Nordic large cap companies. The market risk premium is then calculated by taking the average market return minus the risk free rate. We then calculate the monthly excess return for each portfolio by taking the stock returns of the portfolio minus the risk free rate.

We then run CAPM regressions on each portfolio to estimate β and α by utilizing the lm function in R which fits a linear regression. β is how well the portfolio correlates with the market risk premium and α is the excess return of each portfolio.

Furthermore, we then go on to estimate the Small minus Big (SMB) and High minus Low (HML) factors in the fama french model. We utilize monthly market cap data and create a function to classify companies into "Big" or "Small" based on their market capitalization relative to the median market cap for each month. This classification is achieved by calculating the median market capitalization for each date and then assigning a label ("Big" if the market cap is above the median, "Small" otherwise) using the ifelse function. The returns for each category are then isolated using a numeric mask—1 for "Big" and 0 for "Small"—and multiplied by the firm's returns to filter out the other category. Average returns for each category are computed monthly, and the SMB factor is determined by subtracting the average returns of "Big" firms from "Small" firms.

The HML factor is created by categorizing firms based on their Book-To-Market (BTM) ratios into three groups—low, medium, and high—using tertiles. This categorization process starts by removing any missing values to ensure accuracy in tertile calculation. Tertiles are then computed for the BTM ratios, and firms are assigned a category of 1 (low), 2 (medium), or 3 (high) based on these tertiles. To manage any missing periods, values from previous periods are replicated forward. For the return calculations, numeric masks are created for the high and low BTM groups (high is 3, low is 1), and these

masks are used to filter the firms' returns data accordingly. Returns for each group are calculated by multiplying the returns matrix by the respective BTM group mask, converting zeros to NA for accurate mean calculations. Monthly averages of returns for both high and low BTM groups are then computed. Finally, the HML factor is determined by subtracting the average returns of the low BTM group from the high BTM group for each period, providing a measure of the value effect in asset pricing, which captures the excess returns of high Book-To-Market stocks over low Book-To-Market stocks.

Finally, we run the Fama-French three-factor regression model (1993) with the market risk premium, size (SMB), and value (HML) factors on the excess returns of the four portfolios. Using R, the regression models are set up with the lm() function. Robustness against potential heteroscedasticity is ensured by employing the vcovHC() function from the sandwich package to calculate standard errors, specifically using the "HC1" type. The regression coefficients, along with these robust standard errors, are then evaluated using the coeftest() function from the lmtest package, which facilitates statistical testing of the factors' significance in explaining the portfolios' returns.

We also conduct an analysis of how portfolios perform under periods of heightened cybersecurity risk. Normalized Google search trends for the word "cyberattack" was obtained for Scandinavia during our timeperiod 2018-2022. We then categorize all months in which the normalized google search trends was higher than our cutoff value of 10 as a period of high cybersecurity risk. We then calculate our different portfolios' annual returns during periods with heightened cybersecurity risk compared to periods with normal cybersecurity risk. Furthermore, statistical t testing was conducted between the two periods for each portfolio to test the statistical robustness of the analysis.

5 Results

5.1 Cybersecurity Risk Over Time and across industries

Over the five-year span from 2018 to 2022, our longitudinal study reveals a steadily climbing trend in the average cybersecurity risk scores among Nordic large-cap index firms. Notably, there was a significant increase in our cybersecurity risk score from 2021 to 2022, which aligns with the rise in the number of identified cyberattacks from 4 to 9, as demonstrated in Figure 1. This correlation suggests that our constructed cybersecurity risk score reflects real-time cybersecurity threats. Furthermore, Figure 1 highlights an era of escalating cybersecurity challenges, evidenced by both the higher cybersecurity risk scores and the increased frequency of cyberattacks. The ascending trajectory of both the number of attacks and the risk scores indicates that cyberattacks are not only becoming more frequent but also that companies are increasingly aware of the risks, as evidenced by more frequent mentions of cybersecurity keywords in their annual reports resulting in higher cybersecurity scores.

The cybersecurity risk scores distributed across the various industries as defined by Nordic Nasdaq reveal that the Energy sector bears the highest risk, followed by Utilities and Consumer Services. In contrast, the Financial sector exhibits the lowest risk score. The Industrials industry, which includes companies such as AP Moeller Maersk and Boliden, and the Healthcare industry, featuring companies like AstraZeneca and Novo Nordisk, have experienced the highest number of cyberattacks (8 and 4 respectively) compared to other industries during the study period. These industries, could potentially be more reliant on information technology, which then makes them inherently more vulnerable to cyberattacks. For example, the Health Care industry heavily utilizes digital systems to store sensitive patient data and data from clinical trials, making it a prime target for cyberattacks. Moreover, these industries consist of a larger number of companies compared to others in our dataset; the Industrials industry includes 47 companies,

Figure 1: Cybersecurity Risk Over Time

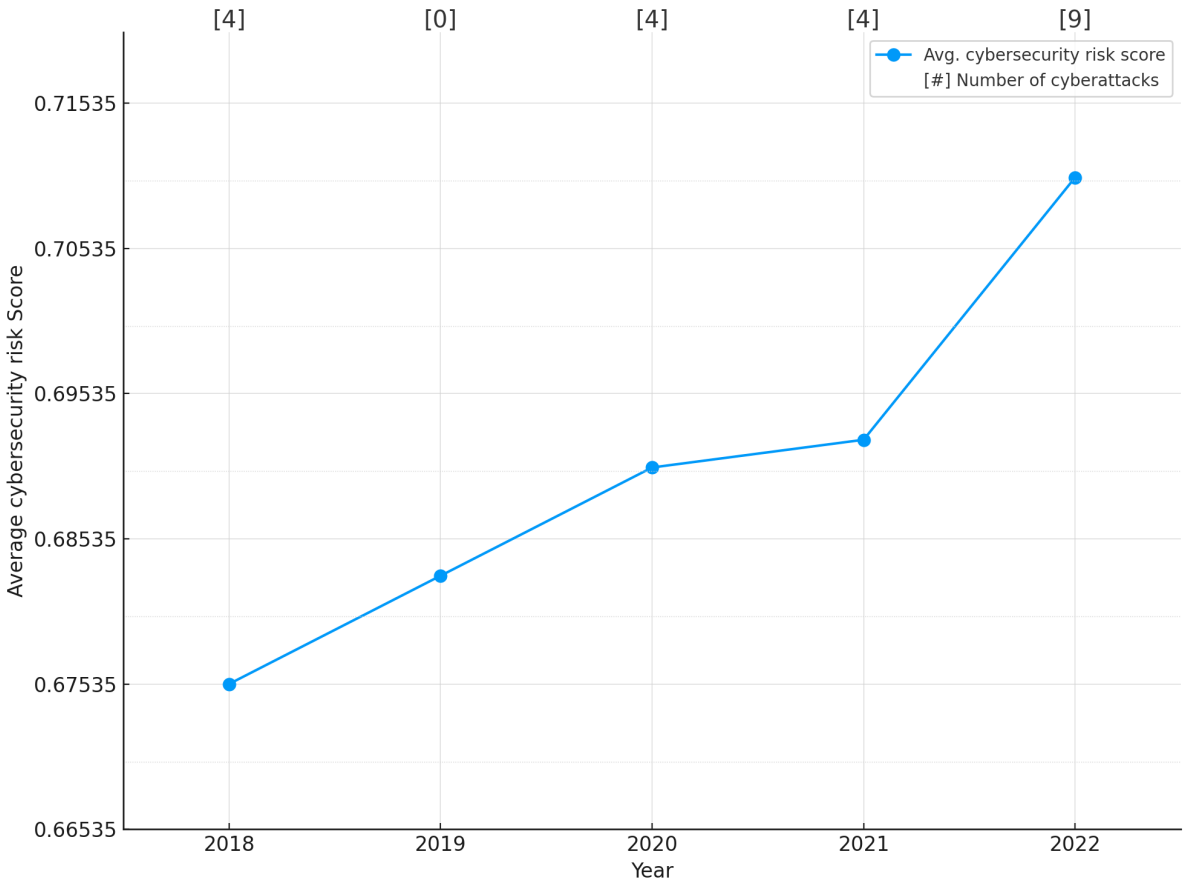
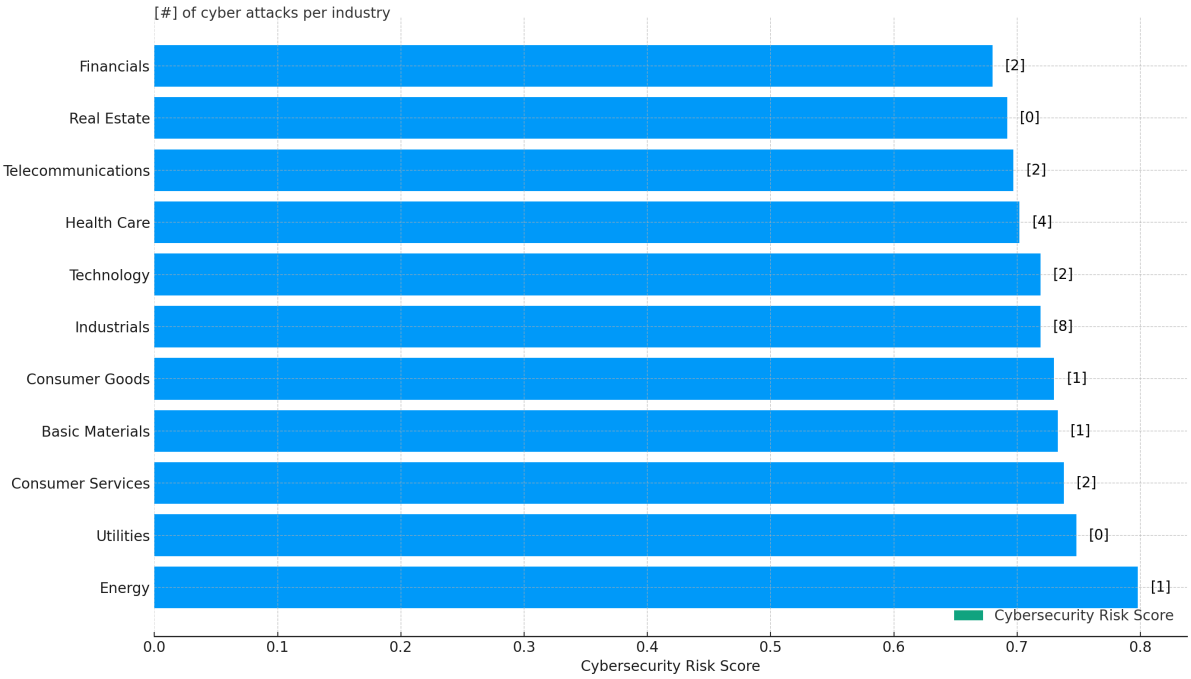


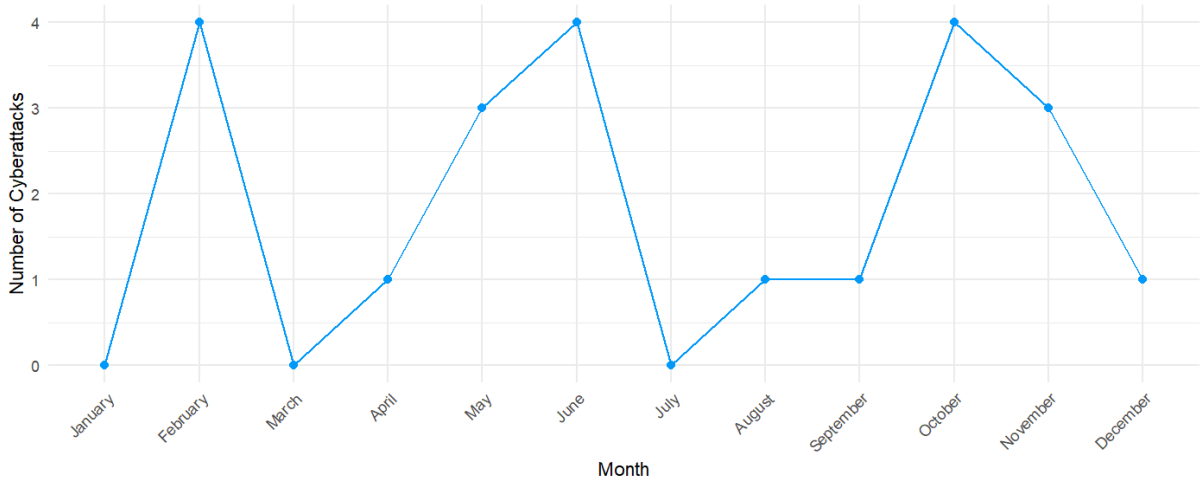
Figure 2: Cybersecurity Risk Across Industries



while the Health Care industry comprises 24 companies. This contrasts with the Energy sector, which includes only a single company, Vestas Wind Systems, which also experienced a cyberattack during our study period. This variation in the number of companies per industry can explain why the Energy sector, which traditionally relies less on information technology systems than the Industrials, has a higher average cybersecurity risk score. Furthermore, excluding the Energy sector, the average cybersecurity risk score variation across industries is less than 0.1, indicating that all industries today face quite similar cybersecurity threats, reflected by the digital age we live in.

Combined, the results correlate quite intuitively with cybersecurity risks: cybersecurity risk has increased over time, and due to the digital era we live in, all industries utilize digital information systems to some extent and therefore face similar cybersecurity threats. A total of 23 cyberattacks were conducted from

Figure 3: Number of Cyberattacks Per Month



This figure displays the seasonality of cyberattacks in Scandinavia by displaying the number of cyberattacks per month between 2018 to 2022

2017 to 2021 across Nordic large-cap firms. Figure 3 displays the number of cyberattacks per month over the period and reveals a somewhat clear seasonal pattern, with peaks occurring in certain months. The number of attacks appears to be higher during the autumn and winter months, as well as in spring and early summer, only to decrease again in July—a month when firm activity is typically lower since many employees are on vacation during this time in Scandinavia. This seasonality could help inform cybersecurity strategies and resource allocation, enabling firms to bolster their defenses during these high-risk periods. It also provides insights into potential attacker behaviors and may indicate times when increased vigilance is necessary.

5.2 Financial Performance of Cybersecurity Risk Sorted Portfolios

In Table 1, we provide a summary of the metrics utilized in our analysis, including cybersecurity risk score, size, price-to-book ratio, geographical location, and the average annual financial returns across different categories of firms over the specified period. Firms are categorized into five groups: all firms, the 23 firms that have been subject to a cyberattack, and firms across three distinct portfolios. Each portfolio is sorted based on their cybersecurity risk score as outlined earlier in the methodology section. As anticipated, the cybersecurity risk scores are highest in Portfolio 1, which contains firms with the highest cybersecurity risk. This is followed by firms that have experienced a cyberattack, which logically

exhibit higher risk scores due to their exposure to actual cyber incidents. Interestingly, firms that have been attacked tend to be larger in size and have a higher price-to-book ratio than others. Additionally, a larger proportion of these companies are based in Stockholm compared to companies in Portfolio 3, which has the lowest cybersecurity risk score. When examining the three portfolios, it is observable

Table 1: Summary Metrics by Category

Category	***Cyber Risk Score***[1]	***Size***[2]	***PB Ratio***[3]	***In Stockholm***[4]	***Return***[5]
All_firms	0.696	7588	5.94	39.8%	0.143
Firms_attacked	0.739	18110	12.0	45.0%	0.094
Firms_P1	0.748	11380	4.08	35.7%	0.083
Firms_P2	0.713	6591	7.12	51.8%	0.132
Firms_P3	0.623	4635	6.67	31.5%	0.216

1. For all firms in each category the cybersecurity risk score was calculated by taking the median value over the time period 2018-2022. 2. Defined as the average market cap for each firm during the time period 2018-2022 in U.S.Dm. 3. Defined as the average price-to-book ratio over the time period. 4. % of firms who have their HQ located in Stockholm 5. The average annual return for each category during the time period

that firms in the high-risk cybersecurity portfolio are typically larger compared to those in the other portfolios. This suggests that size may be a significant factor in cyberattack susceptibility. Furthermore, firms in the low cybersecurity risk portfolio display a higher average annual return (21.6%) compared to those in the high-risk category (8.3%). In addition, firms which have been subject to a cyberattack have the second to lowest average annual return (9.4%), furthering indicating that high risk cybersecurity firms have lower financial returns than low risk firms during the study's time period. This disparity in financial performance could indicate that higher cybersecurity risk may detract from a firm's overall profitability or that investors reward lower-risk firms with higher valuations. The inverse relationship between cybersecurity risk and financial performance highlights that cybersecurity risk can be viewed as a negative factor when predicting a firm's financial return.

The comparative risk-return profiles of four portfolios, sorted by cybersecurity risk and including a long-short strategy, highlight distinct investment profiles. Portfolio 1, which is categorized as the high-risk cybersecurity portfolio, exhibits the lowest financial returns and also the lowest volatility among the three portfolios. In contrast, Portfolio 3, the low cybersecurity risk portfolio, displays the highest average annual return but also the highest volatility. So far, the observed average annual return appears to be dependent on the portfolios' risk profiles; Portfolio 3 offers higher returns but also entails greater risk. Consequently, we constructed a long-short portfolio strategy, which involves going long in Portfolio 3 and short in Portfolio 1. This long-short portfolio demonstrates higher returns compared to the high and mid-risk cybersecurity portfolios, Portfolios 1 and 2, while also exhibiting lower volatility. This configuration effectively contrasts the typical risk-return relationship, suggesting that the different cybersecurity risk profiles of the different portfolios can mitigate risk while enhancing returns.

To analyze the financial returns of the different cybersecurity risk-sorted portfolios, we conducted a CAPM regression, as outlined earlier in the methodology section. In this context, β is defined as the correlation of the portfolio's returns with the market return, which in this case is the average return of all Nordic large-cap firms. As displayed in Table 2, high-risk cybersecurity firms have a lower β compared to low-risk cybersecurity firms, which exhibit the highest β. This indicates that the latter have a higher risk profile, aligning well with the increased return volatilities of these firms.

Figure 4: Portfolio Volatility Versus Avg. Annual Return

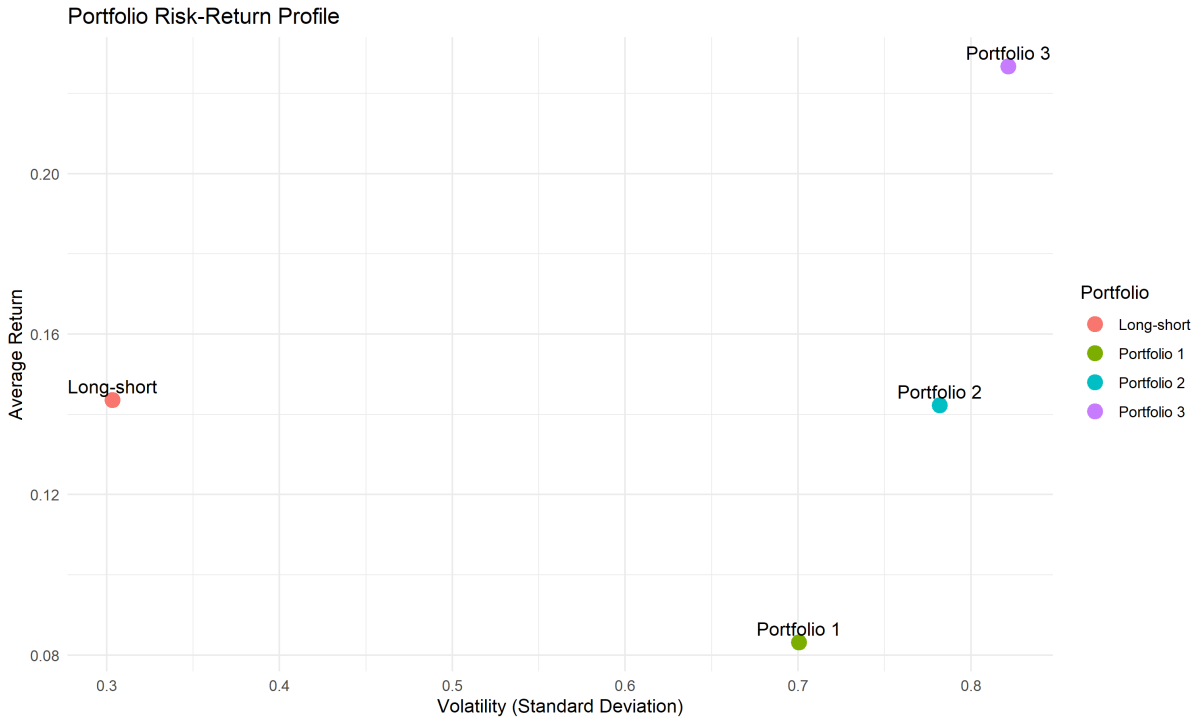


Table 2: Alpha, Beta, and Their Significance for Portfolios 1-4 under CAPM

Portfolio	***Alpha***	***Beta***
Portfolio 1	-0.005415***	0.9206***
Portfolio 2	-0.002029	1.035***
Portfolio 3	0.004399*	1.081***
Long-short	0.009801***	0.161***

Significance levels: * p<0.1, ** p<0.05, *** p<0.01

Portfolio 1 shows a significantly negative alpha, implying that these firms underperform the market. Conversely, Portfolio 3 displays a positive alpha, though not significant, suggesting marginal overperformance relative to the market. More interestingly, the long-short portfolio not only has a significantly positive alpha, the highest across all four portfolios, but also the lowest β, indicating that this portfolio has the lowest risk while also outperforming the market as a whole. Since these findings are statistically significant with a p-value below 0.01, they underscore the robustness of the results under CAPM. This suggests that a portfolio strategy of going long in low-risk cybersecurity stocks and short in high-risk stocks can generate positive abnormal returns not explained by increased market risk. Furthermore, this observation contradicts the Efficient Market Hypothesis, which posits that higher returns should be accompanied by higher risks. This suggests that cybersecurity risk is not efficiently priced in the market.

Table 3 extends the financial analysis to the Fama-French three-factor model, which accounts for size and value effects in addition to market risk, as outlined by the methodology section. The variation in alpha across portfolios indicates differing levels of abnormal returns that are not explained by the model; however, none of the alphas are statistically significant, thus these observations are not robustly supported by the data. Previously under CAPM, it was observed that Portfolio 1 had a statistically

Table 3: Alpha, RM-RF, SMB, HML, and Their Significance for Portfolios 1-4 under Fama-French Model

Portfolio	*Alpha*	*Market Premium*	*SMB*	*HML*
Portfolio 1	-0.001073	1.06***	-0.4473***	-0.1374***
Portfolio 2	-0.003174	1.107***	-0.2315***	0.09475*
Portfolio 3	0.0005132	1.133***	-0.1662**	0.2178***
Long-short	0.001586	0.07366	0.2811***	0.3551***

Significance levels: * p<0.1, ** p<0.05, *** p<0.01

significant negative alpha, indicating underperformance relative to the market. While this alpha is no longer significant in the Fama-French model, Portfolio 1 exhibits significantly negative size (SMB) and value (HML) factors, suggesting that these factors explain the observed underperformance rather than a high cybersecurity risk score.

Moving to the long-short portfolio, the significant positive alpha observed under CAPM is also no longer significant in the Fama-French framework. Similarly, this portfolio shows positive and significant size and value factors, which then explain the previously observed overperformance of the long-short portfolio. Thus, significant factor loadings on SMB and HML indicate that size and Book-To-Market equity are meaningful in explaining the returns of these portfolios during the observed period.

These findings suggest a limitation in the robustness of the impact of cybersecurity risk scores on the financial performance of Nordic large-cap firms. Moreover, it implies that the Efficient Market Hypothesis still holds true, as any excess returns observed across the portfolios can be explained by the three factors present: market risk, size, and value. In addition, our analysis implies that cybersecurity risk does not have a significant impact on Nordic large cap firms' stock returns and the absence of a cybersecurity risk premium. This analysis underlines the importance of considering multiple factors when assessing the financial implications of any custom risk score and the potential limitations of using cybersecurity risk scores alone to predict financial performance.

Table 4: Portfolio Performance under High Cybersecurity Risk Periods

Portfolio	*Avg Return High Risk*	*Avg Return Low Risk*	*Significance*
Portfolio 1	-0.10	0.15	
Portfolio 2	-0.12	0.25	*
Portfolio 3	0.02	0.31	
Long-short	0.12	0.15	

Significance levels: * p<0.1, ** p<0.05, *** p<0.01

Finally, Table 4 contrasts the performance of portfolios during periods of high and normal cybersecurity risk, as determined by Google search trends on cyberattacks in Sweden. The data reveals that firms in Portfolio 1 experience a negative average return of -10% during high-risk periods, yet they achieve positive returns of 15% during lower-risk periods. Interestingly, Portfolio 3 is the only group that generates a positive average return during high-risk periods, indicating that its lower cybersecurity risk profile enables it to maintain positive performance even under adverse conditions. Furthermore, all portfolios generate higher returns during low-risk periods, suggesting that periods of elevated cybersecurity risk generally impact all stocks negatively, likely because investors may be more cautious about investing during these times.

We also conducted statistical t-tests to compare the returns of the two periods to validate the robustness of the results. Unfortunately, none of the results were statistically significant, which limits the conclusions that can be drawn from this analysis. This lack of significance highlights the need for further research, possibly with larger data sets or alternative methodologies, to better understand the financial impact of cybersecurity risk periods on stock performance.

6 Summary and interpretation

This thesis set out to study how cybersecurity risk affects stock returns within Nordic large-cap firms, specifically quantifying the risk for each firm and analyzing how these risks are reflected in stock returns. Our empirical findings provide both descriptive and statistical results to illuminate the potential relationship between organizational cybersecurity risk and financial performance. Our results contrast previous findings but are also lacking in statistical robustness and face central limitations.

6.1 Interpretation of Descriptive Evidence

Our results indicate a steadily increasing trend in the average cybersecurity risk scores among Nordic large-cap index firms from 2018 to 2022, with a significant increase in risk scores from 2021 to 2022. This aligns with Florackis et al. (2023), who also observed an increase in cybersecurity risk over time, suggesting that we are living in an era of heightened digital threats. Potentially this highlight organizations growing reliance on digital systems and the escalating nature of cyber threats. Our study also identified a seasonal pattern in cyberattacks, with higher frequencies in certain months. The seasonal pattern can be compared with the seasonal pattern of national Swedish cyberattacks during 2023 which show a peak during the early winter months corresponding somewhat to our results (MSB. 2023) which should be interpreted with caution due to the relatively small sample size (23 cyberattacks).

We found that larger firms with higher price-to-book ratios are more susceptible to cyberattacks, which is consistent with Florackis et al. (2023) and Kamiya et al. (2021), who noted that larger firms tend to be targeted more frequently. This trend could be due to the hypothesis that larger firms are more lucrative targets for hackers, potentially offering larger ransoms. However, our study diverges from Florackis et al. (2023) in terms of industry-specific risks. While the Florackis et al. (2023) study showed that the energy industry had the lowest cybersecurity risk score, our analysis ranked it as the highest. This discrepancy could be attributed to differences in the datasets and industry definitions used. In addition, it could be explained by differing number of companies per industry. In Appendix *Table 5* we display the % of cyberattacks per industry in our results. It then becomes apparent that the energy sector which included only one company, Vestas, had experienced a cyberattack, thus skewing the results. Furthermore, the Industrials and Health Care industries which had the highest number of cyberattacks in our dataset, scored lower on our cybersecurity risk scale. This could be attributed to the fact that the % of cyberattacks in these industries are similar or lower compared to the other industries in our dataset possibly explaining why although these industries have the largest number of cyberattacks they also consist of a larger number of companies compared to the other industries, explaining why they do not have the highest cybersecurity risk score.

6.2 Interpretation of Cybersecurity Risk Impact on Stock Returns

Our analysis revealed that firms with higher exposure to cybersecurity risk tend to underperform financially compared to their lower-risk counterparts under the CAPM, suggesting a challenge to the Efficient

Market Hypothesis (EMH). However, when applying the Fama-French three-factor model, this underperformance was attributed not to cybersecurity risk but to size and the Book-To-Market (BTM) factors, as the alpha for the long-short portfolio lost its statistical significance while remaining positive. Thus, our findings indicate that the observed over/underperformance of low/high cybersecurity risk stocks in the Nordic large cap index is attributed to size and BTM factors. Therefore, our study indicates that the Efficient Market Hypothesis holds and that cybersecurity risk does not entitle a risk premium for Nordic large cap firms. This contrasts with previous studies which found that there exists a cybersecurity risk premium which is accurately priced into stock returns with investors demanding larger returns for holding high cybersecurity risk stocks (Florackis et al. 2023, Jiang et al. 2020, Helene et al. 2023). The discrepancy in our results could potentially be explained by limitations such as a smaller dataset, leading to less reliable outcomes. Limitations of this study are explained in greater detail in section 6.4. Our findings also highlight the importance of utilizing a three factor model instead of a one factor model as that could have generated faulty conclusions in our study.

Our results also show that firms in the low cybersecurity risk portfolio displayed higher average annual returns compared to high-risk firms. This inverse relationship between cybersecurity risk and financial performance suggests that higher cybersecurity risk may detract from a firm's profitability, aligning with Kamiya et al. (2021), who found that cyberattacks involving financial information loss led to significant shareholder value losses. Similarly, Jiang et al. (2020) found that high cyber risk stocks tend to underperform the market during heightened cybersecurity risk periods, indicating the market's sensitivity to perceived cyber threats. This also aligns with our results which show that the underperformance of high risk cybersecurity stocks is particularly pronounced during periods of heightened cybersecurity threats based on google search data, suggesting a market sensitivity to real-time risk escalations. Such findings are consistent with the Efficient Market Hypothesis, which predicts that stock prices reflect all available information, including the evolving risk landscape.

6.3 Study Limitations

There are notable limitations in our study compared to previous research with the main being: geographical restriction and small sample size, lack of cybersecurity risk disclosures in Nordic firms, and a possibly inaccurate cybersecurity risk score.

6.3.1 Geographical Restriction and Small Sample Size

Unlike previous studies broader focus on U.S. markets, our study is confined to the Nordic region, which might not experience the same level of cybersecurity threats or have the same regulatory and technological frameworks.

A notable limitation of our study was the relatively small sample size, primarily due to the manual methods required for data collection. In contrast to the approach taken by Florackis et al. (2023), who could leverage the SEC's website and a web crawling algorithm to automatically download all the firms' 10-K forms, our process involved manually downloading each company's annual report from their respective websites taking a substantial amount of time. Additionally, identifying whether any firms had been subjected to a cyberattack required a manual desktop search, which is less efficient and more prone to oversight compared to a centralized database. This labor-intensive process restricted both the number of firms included in our dataset and the time period over which we could conduct our analysis.

Implementing a centralized database in the EU, similar to what previous studies utilized for the U.S. market, could substantially enhance cybersecurity research capabilities. Such a database would not only

facilitate more comprehensive data collection by automating the retrieval of financial documents and records of cybersecurity incidents but would also provide a valuable resource for shareholders and other market participants seeking transparent and up-to-date information on firms' cybersecurity exposures and incidents.

6.3.2 Lack of Cybersecurity Risk Disclosures in Nordic firms

Furthermore, another significant limitation is that, unlike U.S. firms, Nordic firms are not mandated to include a specific segment on cybersecurity risk disclosure in their annual reports. In 2023, the SEC (2023) adopted new regulations requiring all firms to incorporate cybersecurity risk disclosures within their annual reports if a company has been subject to a cyberattack. In contrast, Nordic firms may choose to voluntarily include such disclosures, but far from all do. Our review of the Nordic large-cap list revealed that many annual reports lacked mentions of "cyber", "cyberattack", "cybersecurity risk", or "cyber incident", suggesting that these firms opt not to address cybersecurity risks at all. Given the pervasive nature of cybersecurity threats, it is highly improbable that these firms are completely free from such risks. This lack of disclosure more likely indicates a deficiency in transparency, which poses significant challenges for our study.

The absence of explicit cybersecurity information in their reports complicates the construction of accurate and meaningful cybersecurity risk scores based on textual analysis. This omission not only skews our cybersecurity risk scores but also potentially misleads investors and other stakeholders about the true risk profile of these firms which could lead to mispriced investments, as the market may not be fully aware of the latent risks these companies face. To address this issue, a regulatory shift in the Nordic countries mirroring the SEC's approach could be instrumental. Mandating comprehensive cybersecurity disclosures would not only improve transparency but also enhance the robustness of cybersecurity risk assessments like ours.

6.3.3 Possible Inaccuracies of our Cybersecurity Risk Score

Our cybersecurity risk score, while informative, is not without its limitations. The initial data conversion process from PDF to text, necessary for our analysis, could potentially lead to loss of information. Unlike the uniform structure of U.S. 10-K reports, Nordic firms' annual reports vary significantly in format, making textual analysis more challenging. We employed an R package to convert these PDFs into text files, a process that might introduce errors or omit crucial data due to the varying quality and layout of the source documents. Additionally, U.S. 10-K reports include specific sections dedicated to risk disclosure, which simplifies the process of extracting relevant data. In contrast, Nordic firms do not follow a standardized format for disclosing risks, which necessitated the analysis of entire annual reports, often exceeding 200 pages. This extensive analysis required significant computational resources and prolonged processing times—over 10 hours in our case—further limiting the reproducibility of our results. Such extensive processing demands might require a more advanced computer system to replicate our findings effectively. These constraints not only impacted the efficiency and scalability of our study but also raised concerns about the generalizability and accessibility of our approach. To mitigate these issues in future studies, developing more robust text extraction tools and possibly advocating for standardized reporting formats in the Nordic region could improve the accuracy and reproducibility of cybersecurity risk assessments. Due to these restrictions, textual analysis might not have been the best method to construct our cybersecurity risk score and we could have taken a different approach as outlined by Jiang et al. (2020) to generate more reliable results.

Our cybersecurity risk score model, developed alongside Florackis et al. (2023), is built on the critical assumption that firms experiencing a cyberattack will reflect this heightened cybersecurity risk in their annual reports, and that other firms facing similar levels of risk will use analogous language. However, Hilary et al. (2016) presents a contrasting view, suggesting that firms subjected to cyberattacks do not necessarily increase their cybersecurity risk disclosures. A significant limitation of Hilary's et al. (2016) study, though, is that their dataset ends in 2014. Since then, our results and those of previous studies have demonstrated that cybersecurity risk has significantly increased over recent years, prompting firms to take cybersecurity risk disclosures more seriously (Florackis et al. 2023, IMF 2024, MSB 2023, Jiang et al. 2020).

6.4 Further Research and Possible Improvements

6.4.1 Further Research

Multiple previous studies have utilized and validated their cybersecurity risk scores by predicting the likelihood of a company experiencing a future cyberattack (Jiang et al. 2020, Florackis et al. 2023, Rustam et al. 2023). Unfortunately, due to time constraints, we were unable to undertake this predictive validation in our study. However, incorporating such an analysis would be a valuable extension of our work. By predicting future cyberattacks, our cybersecurity risk score could serve as a proactive tool for companies, enabling them to potentially assess their vulnerability to cyber threats. Moreover, predicting future cyberattacks would significantly bolster the validity of our cybersecurity risk score as it would demonstrate that it correlates to real risk of being subject to a cyberattack.

Moreover, Lending et al. (2018) emphasized the importance of corporate governance and social responsibility in mitigating cyber risks, finding that firms with higher standards in these areas are less likely to suffer data breaches. It would be interesting to investigate these differences within Nordic large-cap firms to see if firms with lower cybersecurity risk scores and those not subject to cyberattacks generally exhibit better corporate governance and social responsibility practices compared to high cybersecurity risk firms. Additionally, it would be interesting to examine whether breached firms in the Nordic region alter their cybersecurity disclosures in annual reports post-breach, and how these changes impact their financial returns, paralleling the findings of Jing et al. (2023). Their study showed that increased cybersecurity risk disclosures following a breach can lead to significant negative market reactions if the disclosures are not sufficiently comprehensive. Considering the decreased transparency and disclosure requirements in Scandinavia compared to the U.S., such an investigation could reveal whether the lack of mandatory cybersecurity risk disclosures following a cyberattack impacts investor perception and firm valuation.

Due to the proximity and economic similarities of the Scandinavian countries, we chose not to weight the risk-free rates of each country based on the number of represented companies in our dataset. While we acknowledge that this could have affected our results, we believe the impact would be minimal due to the relatively small differences in interest rates across Scandinavian countries and due to the fact that most companies in our dataset were Swedish.

6.4.2 Possible Improvements of Our Cybersecurity Risk Score

To enhance the accuracy of our cybersecurity risk score, several improvements could be considered:

1. **Weighted Keywords:** Placing greater emphasis on specific keywords might refine the risk assessments. For instance, the term "cyberattack" might be weighted more heavily than "information security" due to its direct implication of risk. Adjusting the weight of keywords according to their relevance and impact could lead to more precise risk profiles.

2. **Expanded Keyword List:** Increasing the number of cybersecurity-related keywords could improve the comprehensiveness of our analysis. While our current list includes over 450 terms, expanding this to include a broader array of terms, similar to Florackis et al. (2023) list of over 6000 words, could capture a wider spectrum of cybersecurity-related discussions within annual reports.

3. **Extracting Cybersecurity Sentences and Comparing Lengths:** Developing a method to systematically extract sentences or segments specifically related to cybersecurity could provide deeper insights into how extensively firms discuss and disclose cybersecurity issues. Comparing the length and detail of these disclosures across firms could serve as an indicator of transparency and the seriousness with which firms approach cybersecurity. Longer, more detailed discussions could imply a higher level of engagement with cybersecurity issues, potentially correlating with better preparedness and risk management practices. This metric could be used to enhance the cybersecurity risk score by incorporating a qualitative assessment of the depth of disclosure.

Implementing these changes could potentially increase the sensitivity and specificity of the cybersecurity risk score, making it a more reliable tool for investors and analysts. These enhancements would also align our methodology more closely with previous studies in textual analysis, leading to more nuanced insights into how cybersecurity risks are perceived and managed across different geographies as almost all previous research on cybersecurity risk has focused on the U.S. market.

7 Concluding remarks

This thesis has examined the interplay between cybersecurity risks and stock market responses within the Nordic region, providing insights into how digital threats impact investor behavior and market dynamics. Our investigation reveals that cybersecurity risk does not significantly affect Nordic Large Cap stocks' returns. Instead, size, value, and market risk factors are more significant in explaining the outperformance of our cybersecurity risk sorted portfolios. Our findings contrast with previous cybersecurity studies (Florackis et al. 2023, Jiang et al. 2020) but align with Hilary et al. (2016) in that cybersecurity risk does not seem to translate into significant changes in stock returns.

However, the study is not without its limitations. One major constraint was the geographical restriction to the Nordic region, which might not experience the same level of cybersecurity threats or have the same regulatory and technological frameworks as other regions like the U.S. Additionally, the relatively small sample size, primarily due to manual data collection methods, limited our analysis. Unlike Florackis et al. (2023), who leveraged the SEC's web crawling algorithm to download all 10-K forms, we manually retrieved each company's annual report, which was time-consuming and less efficient. Identifying cyberattacks also required a manual search, further constraining our dataset. Another significant limitation is the lack of mandated cybersecurity risk disclosures in Nordic firms. Unlike the SEC's 2023 regulations requiring cybersecurity risk disclosures in the U.S., Nordic firms often choose not to include such information voluntarily. This lack of transparency complicates the construction of accurate cybersecurity risk scores and potentially misleads investors about the true risk profile of these firms.

To address these limitations, future research could benefit from automated data collection methods and a centralized EU database for cybersecurity incidents. This would enhance the comprehensiveness and efficiency of data collection. Additionally, regulatory changes in the Nordic region mandating detailed cybersecurity risk disclosures could improve transparency and the robustness of cybersecurity risk assessments. Further research could also explore the predictive validity of the cybersecurity risk score by forecasting future cyberattacks. This would help in strengthening the measure's validity and providing

firms with a proactive tool for assessing their vulnerability to cyber threats. Investigating the role of corporate governance and social responsibility in mitigating cyber risks within Nordic large-cap firms could also yield valuable insights.

In conclusion, while our cybersecurity risk score provides informative insights into the relationship between digital threats and financial performance, the study highlights the need for improved data collection methods, regulatory enhancements, and further research to fully understand and quantify the impact of cybersecurity risks on stock returns. Implementing these changes could significantly enhance the robustness, reliability, and practical relevance of cybersecurity risk assessments, providing valuable guidance for investors, firms, and policymakers in managing and mitigating cybersecurity threats in an increasingly digitalized business environment.

References

Banz, W. Rolf (1981). "The relationship between return and market value of common stocks". In: *Journal of financial economics* 9, pp. 3–18.

Bauer, D. Michael, Huber, Daniel, Rudebusch, D. Glenn & Wilms, Ole (2022). "Where is the carbon premium? Global performance of green and brown stocks". In: *The Journal of Climate Finance* 1.

Cohen, Lauren, Malloy, Christopher & Nguyen, Quoc (2020). "Lazy Prices". In: *Journal of Finance* 3, pp. 1185–1772.

Fama, Eugene (1970). "Efficient Capital Markets: A Review of Theory and Empirical Work". In: *The Journal of Finance* 25, pp. 383–417.

Fama, Eugene & French, Kenneth (1993). "Common risk factors in the returns on stocks and bonds". In: *Journal of Financial Economics* 33, pp. 3–56.

Fama, Eugene & MacBeth, D. James (1973). "Risk, return, and equilibrium: empirical tests". In: *Journal of Political Economy* 81, pp. 607–636.

Finansdepartementet (2024). *Utredning föreslår operativ krishantering vid cyberattacker och andra allvarliga driftsstörningar i betalningssystemet.* URL: https://www.regeringen.se/pressmeddelanden/2024/01/utredning-foreslar-operativ-krishantering-vid-cyberattacker-och-andra-allvarliga-driftsstorningar-i-betalningssystemet/.

Florackis, Chris, Louca, Christodoulos, Michaely, Roni & Weber, Michael (2023). "Cybersecurity Risk". In: *The Review of financial studies* 36, pp. 351–407.

Gentzkow, Matthew, Kelly, Bryan & Taddy, Matt (2019). "Text as Data". In: *Journal of economic literature* 57, pp. 535–574.

Gerding, Erik (2023). "Cybersecurity Disclosure". In: *SEC*. URL: https://www.sec.gov/news/statement/gerding-cybersecurity-disclosure-20231214.

Hanley, W. Kathleen & Hoberg, Gerard (2010). "The Information Content of IPO Prospectuses". In: *The Review of Financial Studies* 23, pp. 2821–2864.

Hilary, Gilles, Segal, Benjamin & Zhang, H. May (2016). "Cyber-Risk Disclosure: Who Cares?" In: *Georgetown McDonough School of Business Research* 2852519.

IMF (2024). *Global Financial Stability Report*. URL: https://www.imf.org/en/Publications/GFSR.

Jakobsson, Josefin (2024). "Regeringens CSRD-förslag kan bryta mot EU-rätt". In: *Dagens Industri*. URL: https://www.di.se/hallbart-naringsliv/regeringens-csrd-forslag-kan-bryta-mot-eu-ratt/.

Jamilov, Rustam, Rey, Helene & Tahoun, Ahmed (2023). "The Anatomy of Cyber Risk". In: *Institute for New Economic Thinking Working Paper Series* 206.

Jiang, Hao, Khanna, Naveen, Yang, Qian & Zhou, Jiayu (2020). "The Cyber Risk Premium". In: *Management studies* Forthcoming.

Jing, Chen, Elaine, Henry & Jiang, Xi (2023). "Is Cybersecurity Risk Factor Disclosure Informative? Evidence from Disclosures Following a Data Breach". In: *Journal of business ethics* 187, pp. 199–224.

Kamiya, Shinichi, Jun-Koo, Kang, Jungmin, Kim, Milidonis, Andreas & Stulz, M. René (2021). "Risk management, firm reputation, and the impact of successful cyberattacks on target firms". In: *Journal of Financial Economics* 139, pp. 719–749.

Lending, Claire, Minnick, Kristina & Schorno, J. Patrick (2018). "Corporate Governance, Social Responsibility, and Data Breaches". In: *The Financial Review* 53, pp. 413–455.

Leovy, Jill (2017). "Cyberattack cost Maersk as much as $300 million and disrupted operations for 2 weeks". In: *Los Angeles Times*. URL: https://www.latimes.com/business/la-fi-maersk-cyberattack-20170817-story.html.

Lindh, Adam & Sundström, Ellen (2024). "Hackerattacken kan ta veckor att lösa – hundratals myndigheter och företag drabbade". In: *Aftonbladet*. URL: https://www.aftonbladet.se/nyheter/a/ab6Ek7/hackerattacken-100-tal-myndigheter-och-foretag-drabbade.

Lintner, John (1965). "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets". In: *The review of economics and statistics* 47, pp. 13–37.

Loughran, Tim & McDonald, Bill (2011). "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks". In: *The Journal of finance (New York)* 66, pp. 35–65.

Mossin, Jan (1966). "Equilibrium in a Capital Asset Market". In: *Econometrica* 34, pp. 768–783.

MSB (2024). *EU förändrar cybersäkerhetsområdet, Årsrapport It-incidentrapportering 2023*.

NICCS (2024). *Vocabulary*. URL: https://niccs.cisa.gov/cybersecurity-career-resources/vocabulary.

Riksbanken (2024). *Styrränta, in- och utlåningsränta*. URL: https://www.riksbank.se/sv/statistik/rantor-och-valutakurser/styrranta-in--och-utlaningsranta/.

Rosenberg, Barr, Reid, Kenneth & Lanstein, Ronald (1985). "Persuasive evidence of market inefficiency". In: *Journal of Portfolio Management* 11, pp. 9–17.

Sharpe, F. William (1964). "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk". In: *The Journal of finance (New York)* 19, p. 425.

Stattman, Dennis (1980). "Book values and stock returns". In: *The Chicago MBA: A Journal of Selected Papers* 4, pp. 25–45.

Treynor, Jack (1962). "Toward a Theory of Market Value of Risky Assets".

8 AI Transparency Statement

We have utilized CHAT-GPT4 in order to improve the grammar of our writing and to find synonyms where necessary. In order to do this we have uploaded parts of our text we were unsure of gramatically and asked CHAT-GPT4 to fix any grammar, style, and syntax errors. In addition to helping with grammatical errors CHAT-GPT4 also helped us when we ran into an error with our code - helping us pinpoint the bug and fix it.

9 Appendix

Table 5: N cyberattacks / total companies per industry

Industry	**Percentage of cyberattacks**
Consumer Goods	14.29%
Industrials	17.02%
Technology	15.38%
Health Care	16.67%
Financials	8.00%
Real Estate	0.00%
Consumer Services	15.38%
Basic Materials	10.00%
Energy	100.00%
Telecommunications	28.57%
Utilities	0.00%

Table 6: List of Nordic Large Cap Companies Subject to a Cyberattack (2017-2021)

Year	**Company**
2017	Ericsson, Sandvik, APMM, Swedbank
2019	Addtech, Demant, SKF, Vitrolife
2020	AstraZeneca, Beijer, NCC, Truecaller
2021	Axfood, Elekta, H&M, Nokia, Nolato, Swedbank, Tietoevry, Volvo, Vestas