Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
NTNU Business School

Benjamin Hashimi

Fundamental Indexation in the U.S. and Norwegian Equity Markets

New Evidence and Extension of the Methodology

Master's thesis in Economics and Business Administration
Supervisor: Khine Kyaw
June 2020

NTNU
Kunnskap for en bedre verden

Preface

This thesis is a part of the two-year program for the Master of Science degree at the Norwegian University of Science and Technology. My objective is to evaluate alternative passive investment strategies by backtesting the fundamental indexation model and modifying it. This important research area came to my attention after reading a Bloomberg article written by Reed Stevenson in 4. September 2019, "The Big Short's Michael Burry Explains Why Index Funds Are Like Subprime CDOs." The article sparked the curiosity to dive deep into the mechanics of passive investment, especially the weighting system. My keen interest in asset management has been a great motivation throughout this semester.

I am also deeply grateful for the valuable guidance of Associate Professor Khine Kyaw during this project. Without her wise supervision and support, this thesis would not have succeeded.

NORWEGIAN SCHOOL OF SCIENCE AND TECHNOLOGY

TRONDHEIM, JUNE 2020

BENJAMIN HASHIMI

Abstract

This thesis aims to investigate whether alternative index weighting based on financial metrics is a better option than the standard cap-weighting practice. In this thesis, I contribute to the existing literature by mainly (1) incorporating non-financial metrics (Environmental, Social and Governance combined score) ESG (2) developing the model by screening and weighting the index portfolio with financial efficiency measures, to give the index portfolio a growth tilt (3) backtesting the original study of Arnott et al. (2005) with a newer dataset (4) implementing the model on a new market (i.e.,Oslo Stock Exchange).

Employing various risk-adjusted performance measures, I document the superior risk and return profile of the fundamentally weighted index over the cap-weighted. For the U.S. equity market, I observed a significant five-factor alpha above 2.00 percentage points (pps). The fundamentally weighted indexes outperformed conventionall index in both absolute and relative sizes. The results from contribution (2) yielded an annualized five-factor alpha of 6.561 pps significant on a 5% confidence level. I also backtested the model on the Norwegian stock market, where one of the indexes surpassed the OSEBX by 187% (2002-2019) with less volatility. On average, every single alternative weighted index outperformed the OSEBX in terms of absolute returns. No significant empirical evidence is found in support of ESG as an additional benefiting metric to the risk and return profile. However, the results indicate a tendency that ESG portfolios provide an excess return.

Keywords – Fundamental indexation, passive investment, ESG investing, index tracking, index fund, growth companies, portfolio management

Contents

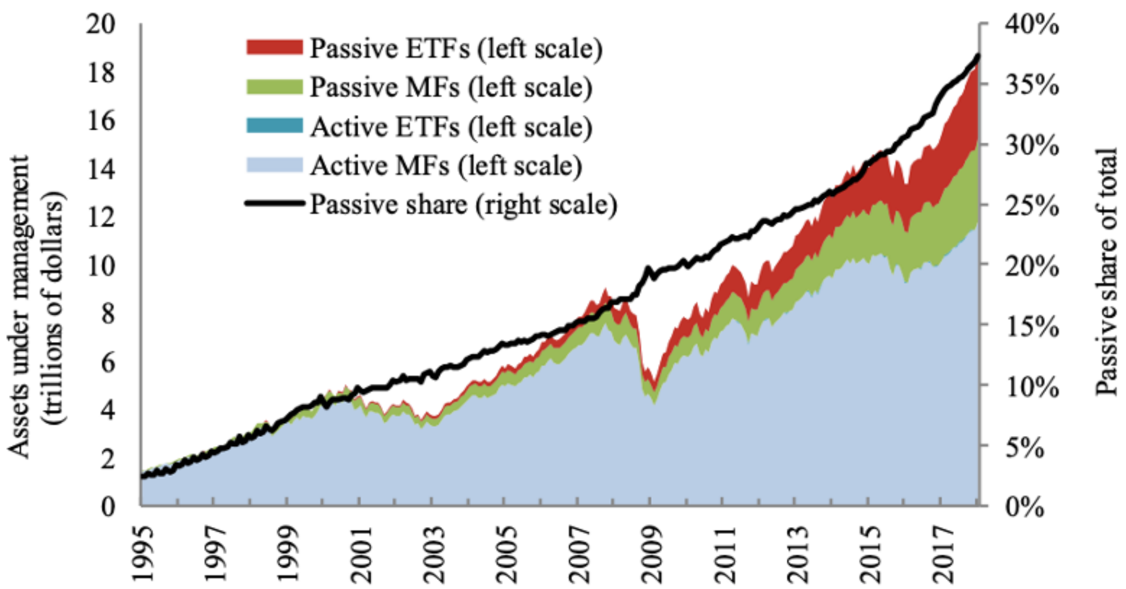
List of Figures

List of Tables

1 Introduction

Capital flows into passive investment strategy has increased substantially in the past decades; at the end of September 2019 assets under management for funds tracking U.S equity indexes surpassed actively managed funds[1]. The $ 4.27 trillion invested in passive funds have steadily grown since the launch of the first index-tracking fund in 1976 (Fichtner et al., 2017). Previous researchers have highlighted two reasons for this significant growth; relatively lower costs of management, and evidence of underperformance of active management strategies on average (Anadu et al., 2019; Fichtner et al., 2017). The regulatory focus on investment fees is also a contributing factor, that encourages and further empower the financial industry to provide low-cost passive products to both retail and institutional investors.[2]

Figure 1.1: Total assets in active and passive strategies



Source: Federal Reserve Board

The majority of these passive mutual funds track capitalization-weighted indexes that buy and sell stocks depending on the relative capitalization weight of each stock in an index. Therefore, weighting an index by its market-cap should perhaps reflect all investors' current information and view, enabling investors holding the index fund to free-ride on the cumulative knowledge of all active investors (Liu and Wang, 2018). This simple capital

[1]https://www.wsj.com/articles/index-funds-are-the-new-kings-of-wall-street-11568799004

[2]Sushko and Turner (2018) point out transparency-driven (MiFID 2) regulations passed by the European Union in 2018 as a legislative that promote low-cost investment vehicles.

allocation strategy bears a possible risk of overweight overvalued stocks and underweight undervalued Arnott et al. (2005), and clashes with the most elementary principle of investment; buy low and sell high. The weighting-risk can be substantial in times of sector bubbles as we witnessed during the dot-com era (Arnott et al., 2011, p. 8). They exemplify this point by referring to the so-called "Axis of Wealth Destruction," which consisted of Cisco, AOL, and Lucent Technologies in the early 2000s. These companies lost a significant portion of their market value when the dot-com bubble burst in late 2000. Especially the Cisco systems stock was badly mispriced; the company's weight in the Russell 1000 index increased from 1.7% (1999) to 4.1% (2000), while its percentage of the economy 0.1% to 0.2% and the P/E multiple expanded from 81.8 to 181.9. The opposite occurs for undervalued stocks, Barclays percent of the FTSE 100 index decreased from 3.1% (2007) to 0.8% (2009); meanwhile, its economic footprint[3] increased from 2.8% to 3.1% (Kalesnik, 2014).

However, it is not evident for an investor to determine when and why the market price is inefficient. Consequently, the "buy low and sell high" method can be a very costly approach given the information cost. To overcome these challenges, Arnott et al. (2005) released a new index design aiming to separate the connection between companies' index weights and stock prices. The concept uses the company's fundamental metrics to size each position in a passive portfolio. The theoretical justification for this method is the belief in the mean-reversion of stock prices and that the market price is a noisy approximation of the underlying value (Arnott et al., 2011, p. 21).

Motivated by the Arnott et al. (2005), In this thesis, I shed new light on the fundamentally weighted index (FWI). Here I investigate how Arnott et al. pioneering index design can be further developed with additional metrics and applied to other equity markets. This thesis consists of four analyses attempting to investigate if the FWI model is a better alternative to the cap-weighting. To begin, I replicate the original study of Arnott et al. (2005) using a newer dataset from 1978 to 2019 and conduct a comprehensive analysis of these results. I find that the FWI portfolios produce on average superior returns compared to conventional indexes with similar and often less volatility. A yearly significant Fama & French five-factor (FF5F) alpha over 2.0% is also reported. The FF5F regression analysis reveals results that confront common critiques of the FWI approach. By combining various

[3] The "footprint" is measured as percent of economy

accounting metrics in a composite index, the portfolio diminishes the factor loadings to value premium—the composite index [4] has no significant value-factor (HML) loading, contradicting the common value-tilt critique of (Perold, 2007; Jun and Malkiel, 2008; Blitz and Swinkels, 2008). Once successfully replicated the FWI model and cross-examined it with the 2005 results of Arnott et al. and (Walkshäusl and Lobe, 2010). The near identical results made me confident to improve the model further.

The second contribution to the FWI literature is the development of an alternative index portfolio with growth style. By following the methodology of Clausen and Hirth (2016), I use the return on tangible assets (ROTA) ranking as a screening and weighting scheme for the FWI portfolio. I construct the fundamental index with the same sector composition as the NASDAQ 100. A geometric annual mean of 22.05% with a 23.77% volatility is observed, surpassing the NASDAQ 100 with its 10.75% geometric return and 25.49% standard deviation from 1979 to 2019. Once adjusting for the bearing systematic risks, predominantly market-risk and SMB, the FF5F alpha is 6.561 percentage points on a 5% confidence level. The portfolio has a defensive downside during market downturns and, at the same time, outperforms with the following rebounds.

My third supplement to the FWI methodology is the incorporation of non-financial metrics (ESG combined score) in the screening and weighting process. The results from this model indicate that the ESG factor has a positive impact on portfolio returns, but the findings are not statistically significant. On average, the ESG-friendly indexes (e.g., S&P 500 ESG and FWI composite ESG) have higher Sharp-ratios and surpass the non-ESG portfolios in absolute and relative measures. However, once I adjusted for the additional factor risks, all indexes showed negative alpha values. The FWI-ESG index had the smallest negative alpha value, thus coming as the best performing index. Lastly, I test the FWI model for the Norwegian equity market, which only has been examined once by Walkshäusl and Lobe (2010). I find that fundamental indexes provide an average of over 2.00 pps three-factor alpha compared to the OSBEX (0.076 pps). The FWI methodology has a distinctive tilt towards big-cap businesses; the SMB factor loading is, on average, above negative 0.5, where some indexes being significant on a 10% level.

[4]A composite index is an equal weight of many fundamental metrics; see chapter 5.1

2 Literature review

In the following chapter, first, I present the literature that this master thesis relies upon and additional empirical evidence from international markets that supports the primary research of (Arnott et al., 2005). As well as critiques of the methodology. Next, essential financial theories are discussed, and divers alternative theories and hypotheses are represented.

2.1 Empirical evidence

The methodology of fundamental indexation is quite new, first proposed by Robert D. Arnott, Jason Hsu, and Philip Moore in 2005. This new method has been replicated to other markets with great success. In this subsection, I will present a handful of studies conducted on different markets in various time periods.

2.1.1 The pioneers of the fundamental weighted index

Arnott et al. (2005) proposed an alternative approach to the standard cap-weighted and called it "fundamental indexation." They weighted the index by its accounting fundamentals, such as; trailing five-year gross revenue, equity book value, trailing five-year gross sales, trailing five-year gross dividends, trailing five-year cash flow, and total employment (Arnott et al., 2005). Those factors are defined as market-valuation-indifferent (MVI), where the primary purpose is to avoid the problem associated with mispriced stocks. These fundamental indexes were constructed with American companies in a period of 42 years (1962-2004) which covered different economic and market environments. With that, they demonstrated fundamental indexes superior performance over the traditional cap-weighted market index. These returns were, on average 1.97 percentage points greater than the S&P 500 yearly throughout the whole period and 2.15 pps higher than the reference index, which was a self-constructed cap-weighted index with the exact same constituents as the FWI index. Sales as a weighting factor showed to be the highest performing index, that beat the reference portfolio by 2.56 pps a year. However, there was some difference in various market environments. Fundamental index outperformed the cap-weighted indexes during the bear market but not bull markets. The question they asked was what if the fundamental index had a value stock bias relative to the cap-weighted indexes and the opposite. The conclusion pointed out four sources that might explain the excess return: (1) superior construction method of the portfolio, (2) inefficiency of market price, (3) additional exposure to distress risk, or (4) a mixture. They also assume that these results are likely to endure in the future.

This new index design suggests that the market price is a noisy approximation to its fundamental value (Chen et al., 2007). According to Arnott et al. (2005), weighting by firm-specific fundamentals is a better reflection of the economic state because of the inefficiency of the market price as an accurate indicator for the incremental performance of a company, which determines the long-term stock price return. Hsu (2004) suggest that when stock prices do not reflect the underlying firm value, cap-weighted indexes are sub-optimal because undervalued stocks will have a relatively smaller market capitalization than their fair value, hence smaller portion in the index. Vice versa, for overvalued stocks, which cause the return drag. Treynor (2005) also agrees on the price as a noisy factor that does not efficiently reflect the underlying value, which implies the sub-optimal characteristics of a standard cap-weighted index.

2.1.2 International evidence

The concept of FWI has been examined for international and regional stock markets; Filipozzi and Tomingas (2017) for the Baltic states and Estrada (2008); Walkshäusl and Lobe (2010) for international markets. Estrada showed that a dividend-weighted index outperformed the cap-weighted index by 1.9% a year over 32 years for 16 countries. Filipozzi and Tomingas (2017) backtested the fundamental indexation model versus the cap-weighted OMX Baltic Benchmark Gross Index (OMX BB GI) for 2006-2016. They used the same constituents as the OMX BB GI and showed that alternative weighting outperformed the benchmark with 2.1 pps annually. Usually, past studies have focused on large stock markets, where the "value stocks" and small-cap excess returns are well documented Basu (1977); Barr Rosenberg and Lanstein (1984). The methodology in the Filipozzi and Tomingas (2017) study differs from Arnot et al. (2005) where they used the same constituents as the OMX BB GI for liquidity and sector exposure reasons. Because the Baltic market is significantly smaller and less liquid than the American stock market.

Walkshäusl and Lobe (2010) have undergone the most comprehensive global study of the fundamental indexation by applying the method across 50 countries in developed and emerging markets. They found that all global fundamentally indexes outperformed their cap-weighted peers. 46 out of 50 countries specific fundamentally-weighted indexes yield higher returns versus cap-weighted for the same amount of risk. They performed robustness and factor tested the findings and decomposed them with the single-factor and Fama & French's three-factor model, as well as Carhart's four-factor model.

The risk-adjusted performance of the composite index measured in the Sharpe-ratio was positive for almost every country except; Argentina, Philipines, Taiwan, and Sri-Lanka. Many of the positive Sharpe-ratios were significant on a 5% and 1% level. For the Norwegian stock market, a fundamental composite index returned 16.03% on average with a volatility of 23.45%, but the results were not significant at 5% nor 10%. Interestingly, the three-factor models' results showed that (HML factor) was decisive for all global portfolios and highly significant. The exposure to the value premium was from 0.19 for sales to 0.33 for employee weighted index, which implied that, for the most part, low B/M stocks generated the excess returns. Further on, size factor loading were positive for a few portfolios, but the degree of the size factor was below the value factor 0.02 (sales) to 0.09(employees). They concluded that fundamental indexation is a unique method and should not be mistaken as a value strategy, as the critics suggest.

Arnott et al. (2005) were not the first ones that emphasized the inefficiency of the cap-weighted index nor the first to explore the idea of fundamental weighting. Gibbons et al. (1989) and Zhou (1991) used likelihood-based tests to demonstrate the weakness of cap-weighted indexes in the American stock market. Moreover, Haugen and Baker (1991) studied the efficiency of the Wilshire 5000 index by constructing low-volatility portfolios. They found that for 1972-1989 there were alternative equity indexes based on the constituents of Wilshire 5000, with superior return and lower risk relative to the cap-weighted version. Major asset managers have also explored the reweighting method of existing indexes. For example, Barclays, Goldman Sachs, and Global Wealth Allocation managed reweighted portfolios of the S&P500 index (Arnott et al., 2005). Arnott argues that this strategy is not sufficient since it requires that companies be large in both capitalization and the other selected financial metric.

2.1.3 Critiques of the FWI model

Previous researchers on the subject have expressed various reasons for the superior performance of FWI. However, the increased exposure to Fama and Frenchs risk factors is pointed out to be the main reason and not market mispricing, as stated by the early researchers (Jun and Malkiel, 2008). The associated risk factors are mainly value and size factors (Filipozzi and Tomingas, 2017; Estrada, 2008; Perold, 2007). The FWI has a bias towards the “value effect,” which can be considered as an umbrella term for companies with conditional, price-dividend ratio, price-book ratio, and price-earnings ratios (Arnott et al., 2007). Perold (2007) reasons that FWI is a strategy with a value tilt. Hence the chance for overweighting “value stocks” with low market capitalization is higher than of cap-weighted indexes. Later studies by Chow et al. (2011) and Jun and Malkiel (2008) approves that of Perold and point out the increased exposure towards value stocks to be a significant contributing factor. Jun and Malkiel (2008) also see that the positive alpha is explained by the Fama-French three-factor model, the value, and size premium. Interestingly, they discovered a mean-reversion in the performance of the FWI strategy.Walkshäusl and Lobe (2010), on the other hand, found evidence that contradicts Jun and Malkiel. They adjusted the returns of the FWI for value and size tilts and still observed that five out of eight fundamental global indexes exhibited a significant positive alpha at a 5% level, which indicates that the FWI is more than only a sophisticated value strategy.

2.2 Theoretical framework

In this subsection, I will exhibit key financial theories to give a comprehensive understanding of different theories and their criqtiques.

2.2.1 The rationale of capitalization weighting

In order to comprehend the predominant role of cap-weighted indexes, we need to look at Sharpe's (1964) (CAPM), which was built on Markowitz's 1959 modern portfolio theory (MPT) and can be considered as the intellectual basis of cap-weighting. The theory linked the market equilibrium and pricing of assets and introduced the concept of the "market-clearing portfolio" where supply equals demand at any given time. In the world of CAPM, all investors are facing the same opportunity set and can allocate along the efficient frontier. Therefore, investors are able to hold the same optimal portfolio. Since the entire investor base is holding the same portfolio of risky assets, the market portfolio must then be a value-weighted market portfolio of risky assets. Hence the weights of each stock are the total market value of all the outstanding units of that asset divided by the overall market value of all risky assets (Fama and French, 2004). Additionally, Phillips and Ambrosio (2008) argues that market indexes aim is to estimate the overall market condition and should, therefore, be cap-weighted. Arnott et al. (2005) also expresses the benefits of a cap-weighted index referring to the simplicity of having little active management of index-funds since market capitalization is highly correlated with trading liquidity, which reduces transaction costs. The other significant benefits are automatic rebalancing and better liquidity of the fund. All of these factors are decreasing asset management costs, which deliver a superior net fee return (Rowley Jr et al., 2018).

2.2.2 Critiques of the CAPM

The CAPM model has undergone various studies since its first appearance, and many asset-pricing anomalies are discovered which reject the positive linear relationship between beta and return. Banz (1981) revealed a contrary in the CAPM and found a relation between market capitalization size of a firm and its average return. Companies with low market-cap had higher average returns compared to firms with larger market cap. The results remain even when considering that small-cap companies' inherent higher risk and beta. Basu (1977) discussed the relationship between price-earnings ratio and excess return. He proved that over 14 years (1957-1971), the low P/E portfolios had, on average higher absolute and risk-adjusted rates of return than the high P/E stocks. Later in 1981, he examined the relationship between earnings yield and firm size. NYSE companies of high E/P had earned, on average higher risk-adjusted returns than of low E/P companies Basu et al. (1981). Further on, Barr Rosenberg and Lanstein (1984) documented the "value effect." They divided companies between high book to market (B/M) and growth companies with low (B/M). The results showed that high (B/M) companies the "value stocks" outperformed growth companies when adjusted for market risk.

2.2.3 The efficient market hypothesis and its critiques

The CAPM withholds many assumptions that are questionable in practice, one of them the efficient market hypothesis (EMH). The EMH implies that stock prices reflect all relevant information, where the market price of a security is an aggregated sum of all available information among all investors. When new information appears, the news is discounted into the market price without delay, which conjectures that the investor base is wiser than any single investor (Fama, 1965). Thus, actively managing a portfolio by stock picking buying and selling is not a profitable investment strategy without taking additional risk. The market price forms as a "random walk"; hence price formation is unpredictable. Therefore, neither fundamental nor technical analysis of stocks would yield a higher return than holding a selection of random securities, at least not without similar risk. The difference between an informed expert and an uninformed speculator vanish with this hypothesis in regard. As a result, the market portfolio should be a passive cap-weighted index because of its mean-variance feature. Recent studies conducted on the performance of actively managed funds, where they measure net return after fees, reveal that on average active managers underperform the market portfolio over time (Carhart, 1997; Busse et al., 2014).

Even though the EMH has strong support in modern finance theory, many anomalies have been reported in the past. For example: "Weekend effects, January effect, size and value effect, A day-end transaction price anomaly, monthly effect, etc" (Dimson et al., 1988). Some of these anomalies diminished since discovered, which supporters of the EHM use to argue for the validity of the hypothesis. However, few still remain in the market

Marquering et al. (2006) and continues to challenge the theory.

Considering fully efficient markets in a strong form where all information is available, even the private ones have been revised. Efficiency in a market exists in various forms (Malkiel and Fama, 1970). They extended the EMH and categorized it into three basic forms; strong form includes private information, semi-strong contains all public information, and the weak form only reflects past prices. The question of which state of the EMH the US market exists in is a long debate. A comprehensive survey of the literature supporting the weak-form by Lim and Brooks (2011) exposes the vast studies supporting the weak-form. Yen and Lee (2008), which addressed the same topic with a survey and proved that the "golden era" of the EMH is over. The school of behavioral finance gets traction with its more dynamic and lose assumptions on human behavior.

2.2.4 The noisy market hypothesis

Siegel (2006) introduced an alternative hypothesis to the EMH. This new hypothesis claims that it exists short-term shocks “noise” in the financial markets which prevent the market price of securities from reflecting intrinsic value. He argues that market participants such as speculators, momentum traders, and often insiders or institutional investors are the ones generating this noise. They are not speculating (creating noise) but trading for functional purposes, such as diversification, liquidity, or taxes. These trades are characterized as noise, which causes temporary shocks on the prices of securities. Such shocks can last for days or years, and their irregularity is challenging for investors to generate excess returns.

Further on, Siegel (2006) argues that the noisy market hypothesis may explain the size and value anomalies. In his paper, Siegel shows that a total market dividend-weighted index remained unchanged during the dot-com bubble, while Russell’s 3000 index value decreased by almost 50% between the height of March 2000 to October 2002.

2.2.5 The adaptive market hypothesis

Andrew W. Lo worked on a new market behavior hypothesis in the early 2000s. The Adaptive Market Hypothesis (AMH), which is from the school of behavioral finance, does not reject the EMH in its theoretical aspect (Lo, 2004). Nevertheless, treat it as an ideal state of a market without frictions that often come with regulation, transaction costs, and irrational behavior by market participants. The AMH suggests that participants in the market execute deals based on previous experience and heuristic decisions of what could be an optimal solution. They learn as the market evolves, and the knowledge is acquired by positive and negative feedback. Under the AMH, market prices reflect the amount of information delivered by the quantity and quality of the participants in the market. As stated in the AMH, arbitrage and profitable trading events may occur but vanish when investors exploit them. Changes in the market environment cause the occurrence of these opportunities. As the economy develops, participants are adapting to the changes and learn from their mistakes.

2.2.6 Mean-reversion in the stock price

The early observers of financial markets supposed that security prices could differ from their fundamental values (De Bondt and Thaler, 1989). In a mean-reverting economy, stock prices are tending towards the mean over a long horizon, and short-term fluctuation exists in the market. The difference between market and fundamental value exists but is temporary. Therefore, in a mean-reverting economy, the stock prices revert to its fundamental values over time. The short-term variations of security prices are caused by systematic "irrationality," which is generated by irrational "noise" traders De Bondt and Thaler (1989).

Poterba and Summers (1988) investigated mean-reversion theory in the American stock market and found that stock returns are positively serially correlated over the short horizon, and negatively autocorrelated over a long horizon. The results revealed the random walk hypothesis's poor validity when they examined the whole dataset for 1926-1985. They designated noise trading as a likely reason and suggested that evaluating such theory requires another firm-specific information than the only stock return.

3 Extension

This thesis mainly draws on the pioneering study of Arnott et al. (2005), "Fundamental indexation." Here, I contribute to the existing literature in four ways. First, I replicate the original study of Arnott et al. within four decades for the American stock market. This period overlaps with the initial study until 2004, which is done for two reasons; proof of methodology and gaining new empirical evidence from 2004 to 2019. The interesting part is to analyze the performance for the period after the financial crisis in 2008, especially during falling interest-rates, expansion of the IT-sector and quantitative easing, which has tilted index investors towards growth companies with high market capitalization.

Secondly, I backtest the FWI model for the Norwegian stock market (2003-2019), that to my knowledge, only has been tested once by Walkshäusl and Lobe (2010) (1988-2007). They only represented results for the composite index and did not compare it with conventional indexes such as Oslo Børs Benchmark Index (OSEBX). Therefore I use the OSEBX as a benchmark index since it is the primary index in Norway, and many passive fund managers track this index. Thus, showing an alternative to the predominant position of cap-weighted indexes in Norway. According to the Norwegian Fund and Asset Management Association, the capital flows to mutual funds tracking the OSEBX had significant growth in the last 20 years[5]. Hence, researching for alternative strategies is profoundly relevant.

Furthermore, I employ the methodology of Clausen and Hirth (2016) to capture growth companies by their fundamental metrics. I use the same definition of a growth company as Ardishvili et al. (1998), which defines the growth rate of over 5% as a growth business. The general definition is that if a company grows more than the relative economy, it is, therefore, designated as a growth. The companies in the FWI index, which I later present in this thesis, have a yearly sales growth over 5%. In this part, I further develop the original method of Arnott et al. by using financial ratios instead of absolute accounting metrics. Considering one of the common critiques of the FWI is that the model tilts the portfolio towards mature companies with poor growth forecasts (Arnott et al., 2011, p. 151). To encounter this critique, I follow the same approach as Clausen and Hirth,

[5]From the monthly statistics of VFF: https://www.vff.no/siste-mÃĕned

where I use the intangible driven earnings to weight an index with the same sector exposure as the NASDAQ 100. This study intends to show that it exists an alternative metric to capture growth companies than solely inflated market capitalization.

Finally, I extend the model by including non-financial metrics—ESG combined score as a weighting and screening factor. In this thesis, I consider ESG as a fundamental metric that gives us valuable information about a company. As previous research on passive investment strategy with ESG screening has given us opposing results. Kurtz et al. (2011) found a positive relationship between ESG scores and stock returns. Whereas Brammer et al. (2006) did not saw the similarity. Kurtz et al. (2011) examined the longterm returns between the KLD 400 Social Index and the S&P 500. They observed that ESG screened index outperformed the benchmark in between 1992-1999 but underperformed in the first decade of the 2000s. They concluded that the KLD index had a significant systematic factor bias during the 90s which was the sole driver for the excess returns. Before-mentioned that ESG scores are positively correlated to the size factor since big corporations withhold higher ESG disclosure standards, and are likely profiting from the economic of scale regarding ESG implementation. Besides, the ESG scoring method also favors particular sectors (Giese et al., 2016).

This topic is highly relevant for the current asset management practice due to the increasing role of passive investment strategies in the Norwegian and U.S. equity markets Anadu et al. (2019) . Hence, the importance of evaluating diverse passive investment strategies are crucial for all kind of investors. For instance, the equity share of the Government Pension Fund Global has a similar investing style as an index fund where the exposure within each position is determined by the relative market capitalization [6]. Private pension providers also utilize this strategy a survey cited in the Financial Times [7] showed that passively managed funds account for 34% of the AUM. Thus, heavily concentrating on only one strategy can make them vulnerable, given the outstanding obligations of these funds. My intention is not to build a case against market capitalization as a weighting method, but to explore alternative weighing techniques.

[6] From the Norges Bank Investment Management: https://www.nbim.no/en/the-fund/how-we-invest/equity-management/

[7] https://www.ft.com/content/f75459e3-3a6d-383e-843b-6c7141e8442e

3.1 Research question

1. *Is the fundamentally weighted index superior to the market-value-weighted indexes?* This is the main research question that motivates this master thesis.

2. *Can it be further improved?*

To answer these questions, I construct alternative hypothetical indexes based on fundamental metrics and compare their performance characteristics with popular conventional indexes from the Norwegian and American markets (i.e., S&P 500, NASDAQ 100 and OSEBX).

3.2 Hypotheses

I form my hypotheses based on results from previous studies, where I expect that past results have endured. In the original study of Arnott et al. (2005), they argue for the persistence of the significant superiority of the FWI model. This argument is more or less verified by later research on other markets and different periods Walkshäusl and Lobe (2010); Estrada (2008); Filipozzi and Tomingas (2017). Thus, I assume that the self-constructed FWI indexes will outperform standard market indexes in absolute returns.

- **Hypothesis I:** *Fundamentally weighted indexes have a better relative and absolute risk-and-return profile than conventional capitalization-weighted market indexes.*
  This hypothesis is derived from assumptions that the equity markets are not fully efficient and answers to my main research question. Here I assume that past results have endured, and replicating the Arnott et al. (2005) with a newer dataset should give somehow equivalent results.

- **Hypothesis II:** *By incorporating ESG in the screening and weighing process, the risk and return profile of the index portfolio improves.*
  To test H_{II}, I compare results from ESG-screened indexes with non-screened. The idea is to examine the effect of ESG-screening and as an additional weighting element. For this part, I am only able to examine for the period after 2002 since it was the first year Thomson Reuters Eikon published the first ESG rankings (Reuters, 2019). The rationale of this hypothesis is that supplementary non-financial information

in ESG combined score will reduce the risk and improve returns. As the study of Ashwin Kumar et al. (2016) showed that ESG-companies inherited less risk and excess returns compared to their peers in the same industry in the U.S.

- **Hypothesis III:** *Fundamentally weighted indexes do not withhold a significant value-stock bias compared to cap-weighted.* Here I address the most common critique of the FWI methodology. When using multiple financial metrics to weight an index, factor loadings to specific risk-factors decreases, since a single metric weighting would have size biases. The value-premium factor loadings for a composite index will eventually reduce when using multidimensional sizes. As previous critiques of the FWI methodology have pointed out HML as the primary explanation for the superior return of the methodology.

- **Hypothesis IV:** *There exists a better method to get growth exposure in an index than only relying on the market capitalization.* The famously NASDAQ 100, which is by far the most growth tilted equity index, has a P/E ratio of 29.65[8] with the highest sector exposure (over 50%) within technology [9]. The $9.8 Trillion market value of the index at the end of 2019 is backed by only $342.8 Billions earnings (Nasdaq Factset, 2019) and the total market value is almost half of the total GDP of the U.S economy[10]. In order to be this true, these 100 companies have to generate almost half of the American GDP in risk adjusted terms in future. Which means that investors are betting that only a handful of the listed companies would generate future earnings that are a significant portion of the entire U.S. economy—thus prepaying for this future success that is uncertain. Subsequently, this growth tilt is only expectations and nothing is certain. So to separate the price from growth prospects, in this thesis, I build a fundamentally growth tilted index based on past growth from small and medium-sized businesses. The FWI growth index is built such that it has the same sector pool and 100 companies.

[8]From the WSJ as of 17. June 2020 https://www.wsj.com/market-data/stocks/peyields

[9]https://indexes.nasdaqomx.com/docs/NDX-vs-SPX_2\%20pager.pdf

[10]The current dollar GDP of the U.S. in 2019 was $ 21.06T according to Bureau of Economic analysis https://www.bea.gov/news/2019/gross-domestic-product-1st-quarter-2019-advance-estimate

The rest of this thesis is organized in the subsequent order: in chapter 4, I explain the data gathering and cleaning process for the U.S. and Norwegian stock market. Section 5 describes the methodology used to construct different indexes, and the performance metrics I use to examine each index. Later on, in chapter 6, I present the results for each market and indexes. Chapter 7 will discuss the results from an economic point of view and explain the reasons for the performance and underperformance of each index. Finally, in chapter 8, I summarize the discussion and give a practical interpretation of the results. An evaluation of limitations and ideas for further research is also described in chapter 7 and 8.

4 Data

To answer the previous chapter's hypotheses, I construct fundamentally weighted indexes for the U.S. and Norwegian stock markets. For the U.S. market, I construct three types of indexes. One group with only accounting metrics, where I reconstruct the original method of but with a newer dataset. The other group consists of companies with an ESG score, where I form indexes ranging from 2003-2019. The third group consists of one index, intending to replicate the same company and sector exposure as the NASDAQ 100. Lastly, for the Norwegian stock market, I build one group of indexes with accounting metrics for 16 years (2003-2019).

4.1 Data gathering for the U.S.

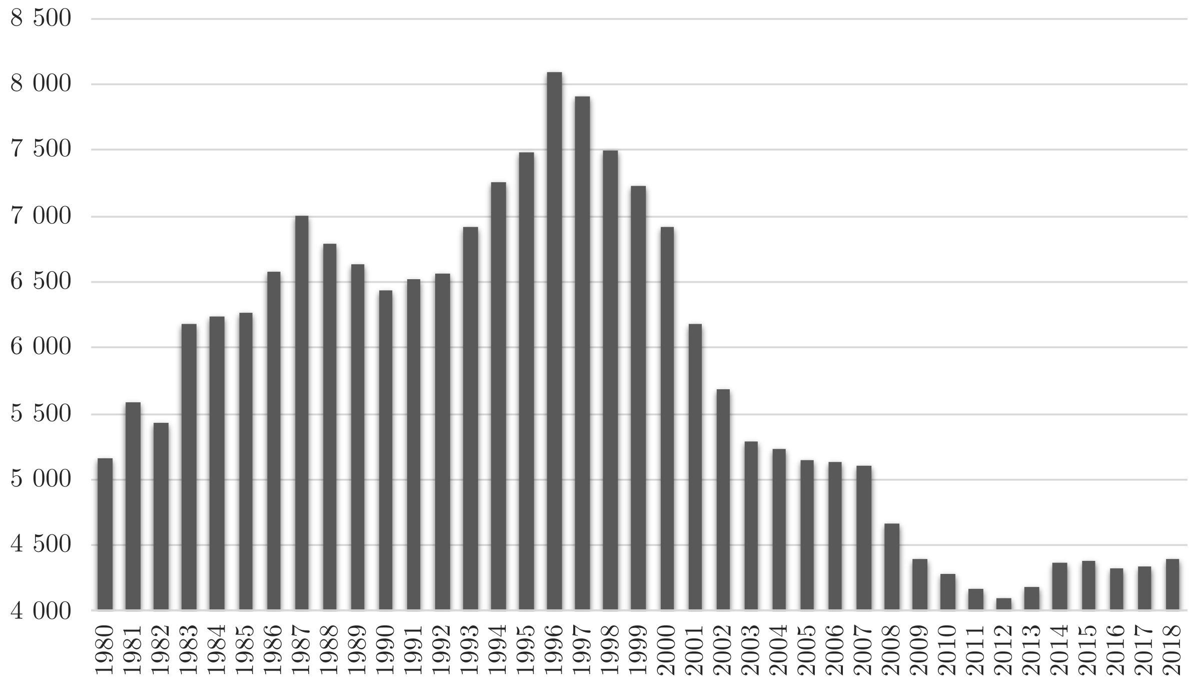
I obtain annual data for all publicly traded U.S. companies by using the CRSP/Compustat merged (CCM) via the Wharton Research Data Service (WRDS). The CRSP database covers all major stock exchanges in the U.S. such that constituents are from different sectors and industries. Hence, the investment universe is nationwide and is not limited by a particular stock exchange market. CRSP data is given by the calendar year whereas the Compustat is in the fiscal year so that index returns are calculated at the end of each calendar year. Henceforth, I avoid a likely look-ahead bias (Walkshäusl and Lobe, 2010). The initial data sampling process has no criteria or restrictions, I have used the option for "search the entire database," where I have included both active and inactive companies to prevent a possible survivorship bias. The data range is from 1975 to 2019, which cover over four decades of stock market data with different economic and market environments.

I have downloaded the following metrics for all U.S. domestic companies from WRDS database:

- GIC sector codes
- Company status active/inactive
- Closing price annually by the calendar year
- Adjusting factor for the closing price
- Company shares outstanding
- Book values per share
- Dividend
- Revenue
- Cash flow
- Net income

All companies are identified by its respective GVKEY code. Total observations are 268 140 firm years, with 24 478 unique companies, of which 20 081 are inactive. It represent the entire feasible investment universe when constructing indexes, but once adding additional screening factors, the number of companies decreases. For example, there are some companies with missing accounting data prior to the 1980s, but they are usually small and new companies. The number of listed companies has also decreased since its peak in 1997, thus limiting our investment universe.

Figure 4.1: Number of listed public companies in the U.S.



4.2 The U.S. ESG data

I have used Thomson Reuters Eikon to retrieve ESG data and accounting metrics for U.S. listed companies annually. The ESG data and accounting metrics are given in fiscal years, whereas stock prices are by calendar year. Such following the same data structure as the Compustat database. In the screening processes, I used two criteria; at least one year with an ESG score between 2002 and 2020, and set the U.S. as the country of incorporation. The first combined ESG ranking was released in 2002 (Thomson Reuters 2019). ESG scores are continuous data, which means that companies tend to receive ESG grade when first received one. It prevents rebalancing due to missing ESG values. Total companies that satisfy this condition accounts for 2 410. The fact that accounting data in the ESG indexes and non-ESG indexes are from two different data providers (e.g.,Compustat and Worldscope) should not affect the results (Ulbricht and Weiner, 2005; Walkshäusl and Lobe, 2010). They do not find statistical nor procedural limitations in Worldscope versus Compustat.

Refintiv (former Thomson Reuters) calculates over 400 company-level ESG measures, of which they choose a subclass of 178 most equal and related fields to make the overall company scoring. These measures are then grouped into ten categories and weighted proportionally. That makes the total ESG score a comprehensive evaluation of a company and is calculated in an automatic, data-driven, and objective method (Reuters, 2019). Hence, assembly the overall ESG score unbiased of a certain sector and is, therefore, a better metric to use than dividing each of the pillars into individual scores.

Table 4.1: Thomson Reuters ESG score calculation

Pillar	Category	Indicators in Scoring	Weights
Environmental	Resource Use	20	11 %
	Emission	22	12 %
	Innovation	19	11 %
Social	Workforce	29	16 %
	Human Rights	8	4.5%
	Community	14	8 %
	Product Responsibility	12	7 %
Governance	Management	34	19 %
	Shareholders	12	7 %
	CSR Strategy	8	4.5%
Total		**178**	**100 %**

4.3 The Norwegian stock market data

I followed the same procedure as the ESG data gathering when obtaining market data for Norway. I filtered all public traded companies (active and inactive) with "Country of Incorporation" in Norway as a criterion, which provided 2 236 firm years' observations with 260 unique companies.

I then downloaded the following company data:

- GIC sector name
- Annual closing price by the calendar year
- Company shares outstanding
- Book value
- Dividend
- Revenue
- Cash flow
- EBITDA

The closing price from Thomson Reuters was adjusted for stock split, reverse stock split and dividend. I used EBITDA instead of net income, since net income values fluctuated and had negative values. Thus keeping the index free of short positions. I also replaced all negative values for trailing three-years fundamental metrics with zero for having only long positions.

5 Methodology

According to Lo (2016), conventional indexes are created as hypothetical portfolios to present a particular market or a segment of a market. The purpose of a stock index is to determine the prices of the market or segment and is usually market-cap-weighted. The stock index has at least two distinct function in modern finance. First informative; indexes deliver a cumulative measure of the constituents' performance to feature economy-wide drivers of the market. Secondly, indexes work as a benchmarking measure to evaluate asset managers; they have the option to track the stock index or actively pick stocks in order to beat the index. To achieve the second function, the index has to fulfill two criteria:

1. Transparency, meaning that every characteristic of the index must be available for the public. Such that investors are able to replicate the index and achieve the same reported return as the index.

2. Systematic and rule-based, so that the construction of the index must be independent of any discretion or subjective judgment.

By the classification of Lo (2016), our fundamental indexation method is within the theoretical definition of an index.

5.1 The Fundamental indexation model

When replicating the fundamental indexation method of Arnott et al. (2005), it is not sufficient to barely rearrange the constituents of a cap-weighted index by fundamental weights. They argue that *"if we simply reweight the stocks in the S&P 500 or the Russell 1000 by book value, we miss a large number of companies with substantial book value that are trading at a low price-to-book ratio"*. Thus, the portfolio is concentrated towards stocks that are large in both market-cap and book value. Their solution to this problem is to rank all companies by individually fundamental metrics and then pick the top 1000 companies for each metric. The constituents of the index are in this way included by their relative metric. Thus gives us the following equation:

$$\omega_{AS,i,t} = \frac{\max\{0, AS_{i,t-1}\}}{\sum_{i=1}^{N} \max\{0, AS_{j,t-1}\}} \tag{5.1}$$

Where ω_{AS_i} is the weight of each company corresponding to its accounting size (AS), and AS_i is the trailing five-year average of the accounting size. The denominator is aggregated sum of the top 1000 companies for each accounting size in the particular year.

The accounting sizes I use to construct indexes are five-year average trailing: cash flow, dividend, revenue and net income except for the book value, which is in single-year. Arnott et al. (2005) did not specify the reason for using one-year book-value, but the logical reason is that book value is less volatile than the other metrics. The advantage of using five-year trailing metrics is to reduce portfolio turnover since single year data tends to fluctuate more. When they used single-year data, the difference in annual return with a five-year trailing average was within ± 10 bps, while turnover increased by more than 2 pps.

Further on, the four metrics of each company is combined in a composite index with equal weights, and then the top 1000 companies of the overall composite are selected. This composite index represents a more robust construction than using a single metric (Arnott et al., 2011, p.76). Robust in the sense of eliminating biases that come with a single metric weight and, at the same time, creating a multidimensional measure of a company. The composite weights of non-dividend-paying companies are averaged by three, thus not discriminating against those companies. The percentage of the U.S. companies paying dividends have decreased in the last decades Kahle and Stulz (2017) and (Arnott et al., 2011, p.78). On the other hand, stock buyback as a mechanism to repay investors has increased due to changes in taxation. Therefore, the decision for not paying dividends is made for other purposes than solely a company's ability to repay its investors. Excluding those companies may affect the index's ability to capture companies with growth ambitions.

The composite index construction can mathematically be expressed by the following equation:

$$\overline{\omega}_{\mathbf{comp,i,t}} = \begin{cases} \frac{1}{4}\sum_{j=1}^{4}\big(\omega_{BV,i,t}, \omega_{CF,i,t}, \omega_{REV,i,t}, \omega_{DIV,i,t}\big)_{DIV>0} \\ \frac{1}{3}\sum_{j=1}^{3}\big(\omega_{BV,i,t}, \omega_{CF,i,t}, \omega_{REV,i,t}\big)_{DIV=0} \end{cases} \tag{5.2}$$

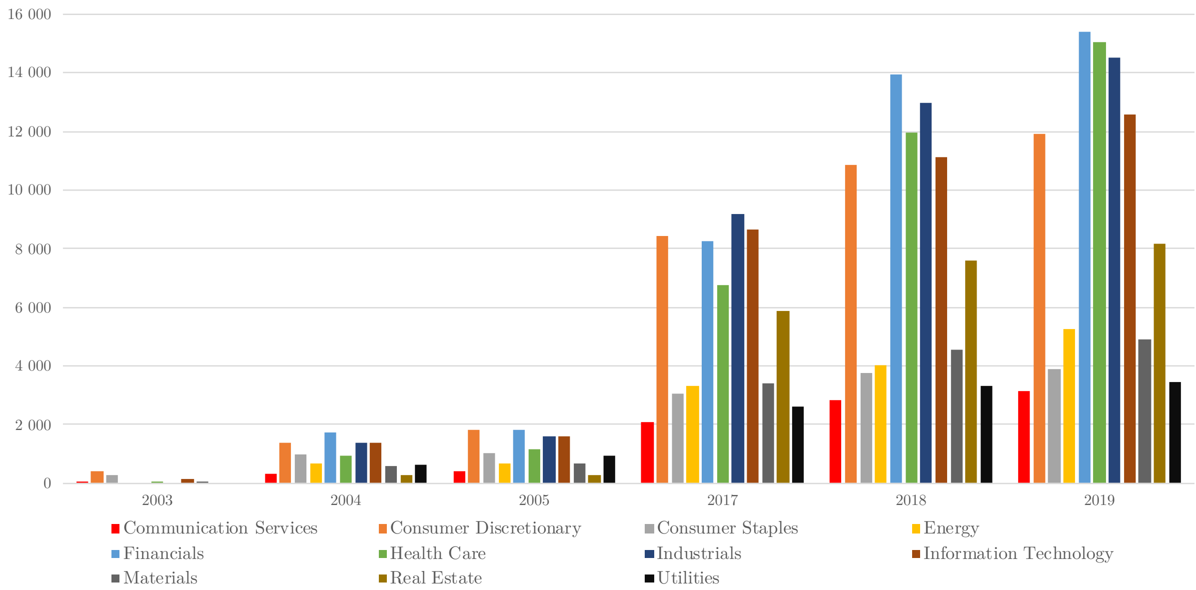
All indexes are rebalanced annually and are held constant throughout the year. The returns are calculated by using the end-day closing price of the last trading day.

I use the same model for the Norwegian and American markets with minor changes. The constituents of the Norwegian FW index varies from 35 to 50 stocks, due to negative accounting metrics where I have replaced all negative values with zero. I have also used three instead of five when calculating trailing averages, thus following the same approach as (Filipozzi and Tomingas, 2017). The three-year trailing average is a better choice since major companies listed on the Oslo Stock Exchange are affected by business cycles (Hillestad, 2007). Therefore, by using three years, the drags that particular "good" years create would diminish. Besides that, the three-years trailing average is a more representative state of the Norwegian economy, due to cycles stocks.

5.2 ESG screened index

The ESG data I have retrieved from Thomson Reuters Eikon is densely sector biased as we see in the figure (5.1). Consumer discretionary, financials, health care, and information technology are sectors with the highest combined ESG-scores and with most companies, especially in the early 2000s. However, it has changed; firms from various sectors are reporting their ESG-status. The investment universe of ESG companies is therefor expanding, which enables diversification to other sectors.

Figure 5.1: Total combined ESG score from each sector



To isolate the ESG-factor as a weighting scheme, I follow a four-step procedure. First, I sort all companies yearly by its GIC sector codes. Then I subtract each company's ESG score from the sector median. Next, I rank all companies yearly by exceeding ESG-scores. Lastly, I choose the top 250 companies each year and use their relative ESG-score to determine each company's weight in the index.

Moreover, I use the relative ESG score to incorporate it as a weighing factor in a composite index, which consists of book value, total revenue, cash flow, and relative ESG score. All companies in the composite index are above their sector medians as well as having strong financial fundamentals.

5.3 Capturing growth companies with the FWI

In this section, I will present an alternative model to capture the growth factor by following the same approach as (Clausen and Hirth, 2016). This alternative model is based on the efficiency of a company to use its already existing intangible, which can be used as an additional factor to explain the market capitalization.

The idea of this index is to capture growth companies by core-fundamental performance. The benchmark of this index is the NASDAQ 100, which consists of 100 companies from across six sectors: information technology, consumer discretionary, healthcare, consumer staples, industrials, and telecommunication services. The NASDAQ 100 index has a distinctive sector tilt towards information technology, with over 50% weight [11].

Clausen and Hirth (2016) found that R&D expenses and the earnings-based intangibility measure are positively correlated with the market capitalization of a firm. Their new measure gauged the relative productivity of already existing intangibles, in contrast to R&D expenses, which measure the investment in new intangibles—this new measure is determined in a three-step process:

- The return on tangible assets is calculated for each firm-year. Which uses property, plant, and equipment (PP&E) in the denominator rather than total assets. The fraction is defined as:

$$\textbf{ROTA} = \frac{EBITDA}{\text{Net PP\&E}} \tag{5.3}$$

  This equation reveals how efficient a firm performs in terms of EBITDA per tangible asset. A company with high ROTA is probably using the most efficient internal processes, such as; skillful workers, efficient computer systems, well-known brand names, etc. Thus they assume that high ROTA companies are more intangible intensive.

- Next, they adjust for variations in the market cycles and cross-industry variations with substracting the by-industry-and-year median ROTA from each ROTA, which is normalized by the by-industry-and-year standard deviation to control for changes in the variation.

- In the third step, each company is ranked according to its ROTA. Thus, avoid the absolute size of returns, because the ROTA measure is a noisy size of the intangible-driven earnings.

To replicate the same sector exposure as the NASDAQ 100, I delete all companies with sic codes (6000-6999) financial sector, which remove 31.224 firm years observations, (4900-4991) 7.837 firm years observations from the energy sector, (1520-1731) from real estate and utilities (4000-4900) deleting 11.096 observations—leaving a total of 146.994 firm years observations from the remaining sectors.

This method of indexing is considerably different from the original index design of Arnott

[11] From the NASDAQ global information services: https://www.nasdaq.com/docs/Nasdaq-100Index.pdf

et al. (2005) in such a way that it uses fractional measures instead of absolute accounting metrics. Further on, I modify the (5.3) equation by using three-year trailing averages of EBITDA and Net PP&E. Due to substantial fluctuation when using single year figures, by doing so, I decrease the volatility of the index with 7 % annually standard deviation and turnover, at the same time. The portfolio returns were not affected much by this procedure.

5.4 Risk adjusted performance measures

In order to measure and compare the performance of each index, I report the return series in risk-adjusted amounts. In this subsection, I will present the risk-adjusted metrics which I use to examine cap-weighting versus FWI.

5.4.1 Sharp ratio

Sharpe-Ratio is the most well-known performance measures and was first introduced by William Sharpe in 1966, in the article "Mutual Fund Performance." The "reward-to-variability-ratio" shows the relationship between the excess return of an asset and its risk. These measures are calculated as ex-post by using the historical standard deviation and the average return of the market price. In this thesis, I have used the one-year treasury-bill in Norway and the USA as the risk-free proxy.

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p} \tag{5.4}$$

One of the advantages of the Sharpe-Ratio is that it ables me to compare all the indexes against each other. Where other measures use relative risk sizes such as beta, the Sharpe-Ratio uses only absolute volatility and risk free rate as a reference portfolio. Thus, cross-examining is possible on a macro-level.

5.4.2 Treynor Ratio

The numerator in Treynor-Ratio is the same as Sharpe, but the difference is in the risk-size. The Treynor-Ratio's risk-size is the beta, which often is calculated with a well-diversified market index. Treynor does not explain all variations in the return series, but only the part that correlates with the market index. I have used the Fama and French mkt-rf factor to calculate the beta for each alternative indexes.

$$\text{Treynor Ratio} = \frac{R_p - R_f}{\beta_p} \tag{5.5}$$

5.4.3 Information ratio

The Information-Ratio (IR) exposes the excess return of a portfolio beyond its benchmark relative to the standard deviation of the excess return, also called tracking error (TE). The benchmark is often a value-weighted market index, and the TE shows the consistency of the excess return. A low TE means that the portfolio beats its reference consistently. If the IR < 0, it indicates that the portfolio has underperformed, and vice-versa if IR > 0.

$$\text{Information Ratio} = \frac{R_p - R_b}{TE_p} \tag{5.6}$$

The tracking error is calculated with the following equation:

$$TE = \sqrt{\frac{\sum_{i=1}^{n}(R_p - R_b)^2}{N-1}} \tag{5.7}$$

5.4.4 Jensen's alpha

The model is based on the CAPM to predict an expected rate of return by using the β coefficient to measure the risk of a holding and $R_m - R_f$ as the equity risk premium. Positive alpha values mean that the portfolio yields abnormal return that can not be explained by the systematic market-risk projected by β_p.

$$\alpha_j = E(R_p) - \{E(R_f) + \beta_p * [E(R_m) - E(R_f)]\} \tag{5.8}$$

5.4.5 Fama & French five-factor model

I use the five-factor model to capture various risk exposure to the FWI indexes. The five-factor model exhibits the RMW (robust minus weak) and CMA (conservative minus aggressive) in addition to (1) CAPM beta, (2) SMB, and (3) HML (Fama and French, 1992, 1993, 2015). The five-factor is an extension of the three-factor model from 1992 Fama and French (2015); they added these two factors to improve the explanatory power of the regression analysis. Because Titman et al. (2004); Novy-Marx (2013) stressed the shortcomings of the three-factor model regarding profitability and investment grade to explain the variation in average security returns.

RMW and CMA factors follow the same methodology as the HML factor. The RMW is calculated by averaging the returns of two portfolios consisted of small and big companies with robust operating profitability (OP), and subtracting it from average returns of a portfolio with big and small companies with a weak OP.

$$RMW = \left(\frac{SR + BR}{2}\right) - \left(\frac{SW + BW}{2}\right) \tag{5.9}$$

CMA factor is constructed by subtracting the average return of a portfolio with two groups of companies with a conservative and aggressive investment policy.

$$CMA = \left(\frac{SC + BC}{2}\right) - \left(\frac{SA + BA}{2}\right) \tag{5.10}$$

HML the value factor is calculated by:

$$HML = \left(\frac{\text{Small Value + Big Value}}{2}\right) - \left(\frac{\text{Small Growth - Big Growth}}{2}\right) \tag{5.11}$$

The size factor SMB is constructed by averaging returns of three portfolios with small companies minus the average of portfolios with big companies. The sizes are based on the market capitalization of a firm.

$$SMB = \left(\frac{SV + SN + SG}{3}\right) - \left(\frac{BV + BN - BG}{3}\right) \tag{5.12}$$

The five-factor model, with all its components, can then be expressed as:

$$R_{pt} - R_{ft} = \alpha + \beta(r_m - r_f) + \beta_i^S(SMB) + \beta_i^V(HML) + \beta_i^p(RMW) + \beta_i^{Inv}(CMA) + \epsilon_{it} \tag{5.13}$$

Where the Rp-Rf is the return of the index portfolio minus the yield of the U.S. treasury bill, Rm is the return of the cap-weighted market portfolio. Fama & French use the CRSP database to construct the market portfolio. Alpha is the excess return that can not be explained by the risk factors. Hence, revealing the true performance of an investment strategy. The FF5F annual data is withdrawn from Kenneth French's web site[12] at Dartmouth from 1978 to 2019.

For the Norwegian equity market, I employ the three-factor model instead of the five-factor model. Because of the CMA and RMW factors are not available from the database I used[13]. On the other hand, including additional explanatory variables could weaken the efficiency of our model. Previous researchers on the Norwegian market have expressed the limited size of the overall equity market; hence the SMB and market factors are pointed out as the most important risk factors in Norway (Skjeltorp et al., 2008).

[12] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[13] Calculated by Bernt Arne Ødegaard: http://finance.bi.no/~bernt/financial_data/ose_asset_pricing_data/index.html

6 Empirical Results and Analysis

In this chapter, I exhibit the results and analysis of the replication and extensions of the methodology. First, I present the results for the recalculation of the original study by Arnott et al., where I have backtested the FWI model from 1978 until 2019. Later on, I present the results from my contribution to the literature by incorporating the ESG combined score as a factor into the FWI method and backtested it between 2003-2019. Further on, I show how the FWI method can be modified to capture growth companies by a fundamental metric. Lastly, I show the outcome for the implementation of the FWI in the Norwegian stock market.

6.1 Results from the replication

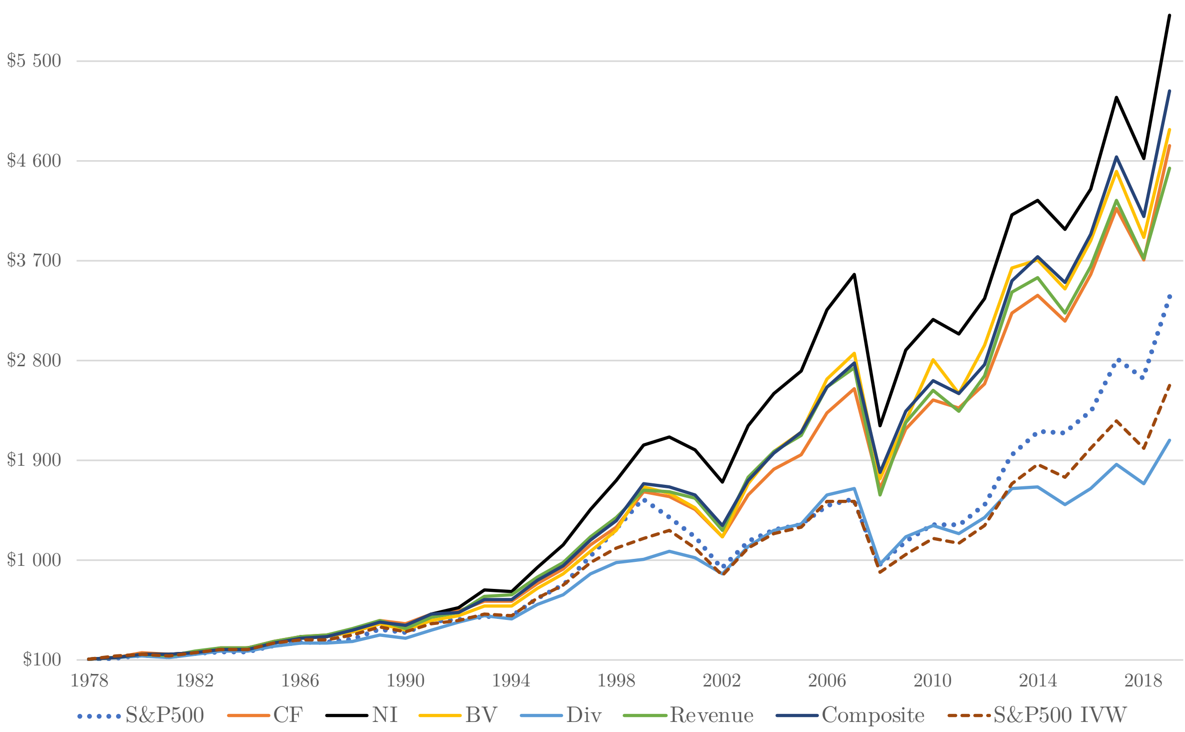
For benchmarking purposes, I use the S&P 500 market capitalization-weighted and the S&P 500 intrinsic value-weighted, which are two different variants of the index with the same company composition but weighted by other metrics [14]. The cap-weighted version is the most well-known and is often used to construct passive mutual funds or exchange-traded funds. Whereas the intrinsic value-weighted as the name suggests weights each company by its intrinsic value. I use these benchmarks to show that the FWI approach has two distinct features in the screening and weighting process. Because it is not sufficient to only rearrange constituents of a traditional market capitalization index by other metrics. The much-researched alternative to the cap-weighting is equal-weighting, and random portfolios (Arnott et al., 2011, p. 16). In this thesis, I do not construct such indexes as the phenomenon is well-documented (Gibbons et al., 1989; Zhou, 1991; Haugen and Baker, 1991). I instead use the intrinsic value-weighted version of the famously S&P 500. The objective is to investigate whatever a reshuffling of the S&P index is a better solution than the FWI approach.

Figure 6.1 illustrates the dollar growth of \$ 100 (USD) invested in fundamentally and capitalization-weighted strategies from 1978 until 2019. As the graph shows, all FWI indexes outperform their cap-weighted counterparts and follow the same trend line. The best performing index is the net income weighted index, with an ending value of \$ 5

[14] From the Standard & Poor's methodology library: https://us.spindices.com/documents/methodologies/methodology-sp-us-indices.pdf

909 and is $ 2 512 more than the S&P 500 cap-weighted. It's necessary to notice that the indexes are not adjusted for transaction costs and should only be considered as hypothetical return series. The same applies to the S&P 500 and other conventional indexes, Arnott et al. suggest to drop transaction costs from the equation since it's not a common practice to include such costs when constructing indexes.

Figure 6.1: Accumulated Growth of $ 100 for the whole period



From table 6.1, we can notice that the majority of the FWI portfolios exhibit better returns and lower standard deviation, which is in accordance with (Arnott et al., 2005). Except for the dividend weighted index, where the geometric return deviates by 5.41% annually from the findings in (2005). Likely the volatility is almost the same. It's also worth mentioning that the dividend payout is not reinvested. A dividend reinvested strategy would probably give higher returns, but in these analyses, I only focus on returns from capital gain. Moreover, the dividend index has also limited downside risk and upside return. The next best downside risk has cash flow and composite indexes. All of the minimum returns are from during the financial crisis, where the equity market had its worst year since 1973-4. The rebound that occurred in 2009 is the max return point for every FWI portfolios except for the dividend, which had the best performing year in 2003.

All indexes are slightly negatively skewed with excess kurtosis, indicating fat tails. Additionally, it shows that the standard deviation is not an appropriate risk measure (Tsiang, 1972); hence I use other risk measures to provide a complete description of risk. The S&P 500 IVW, revenue, and dividend indexes are highly skewed, with the highest excess kurtosis values. That indicates extended tails to the left, which means that the probability of extreme negative returns is relatively higher than that of normal distribution.

Table 6.1: Descriptive statistics for the replication

	S&P 500	S&P 500 IVW	Cash Flow	Net Income	Book Value	Revenue	Composite	Dividend
$\overline{X}_A$	9,97 %	9,19	10,70	11,30	10,95	10,81	10,95	8,53 %
$\overline{X}_G$	8,76 %	8,04	9,62	10,20	9,70	9,51	9,88	7,50 %
Volatility	15,57 %	14,97	14,66	14,75	15,65	15,91	14,56	13,99 %
Skewness	-0,80	-1,10	-0,82	-0,94	-0,92	-0,89	-0,91	-1,06
Kurtosis	3,71	4,61	3,60	4,07	3,81	4,28	3,74	4,73
Min $_{\text{Retrun}}$	-38,49 %	-41,24	-35,66	-38,11	-39,47	-41,98	-35,95	-31,09 %
Max $_{\text{Return}}$	34,11 %	32,53	32,80	31,00	38,76	41,79	31,38	29,54 %

The return series are calculated by using the annually closing price of the FWI and the S&P indexes. The period spans from 1978 to 2019, and all numbers are annualized. Volatility is measured in the standard deviation of returns. $\overline{X}_A$ and $\overline{X}_G$ stand for arithmetic and geometric annual returns. The Kurtosis value is in absolute number, and is not substracted from 3.

Table 6.2 is a comprehensive overview of the relative performances of the FWI portfolios versus their benchmarks. Here I have used the one year T-bill as a risk-free proxy to calculate the Ratios. The overall performance of the fundamental indexes is above the S&P 500, excluding the dividend weighted. These indexes also have low relative betas to the cap-weighted indexes. Moreover, as shown in Jensen's alpha row, the CAPM-alpha are all prominently higher for the FWI portfolios. However, the alpha decreases once we include additional systematic risk factors exhibited in subtable B. The tracking error (TE) expresses the consistency of the outperformance, which indicates the volatility of the excess returns generated by FWI portfolios. As we see from the information ratio (IR), the model has the ability to produce consistent risk-adjusted excess returns. The best performing single metric index is net income weighted that surpasses benchmarks by 1.3, and 2.1 percentage points annually—besides, the composite index with the next best performance, which is somehow as expected given the past results of (Arnott et al., 2005).

The panel B display results from the Fama and French five-factor regression using the equation (5.13) from the methodology chapter. It unveils the systematic factor risks associated with these two strategies. As we see, the FWI strategy produces significant five-factor adjusted alphas. Whereas the net income weighted index inherent the highest alpha through the whole period surpasses the composite with 0.003 pps. The alpha value for the composite is also different from zero on a 5% confidence level. The market exposure is significant on a 1% level for all indexes, with a factor loading of less than 1.00 for the majority of FWI indexes indicating a defensive market exposure compared to the S&P 500. The SMB factor shows that the S&P 500 is highly concentrated towards the big cap companies, reasonably accurate given the cap-weighting methodology. FWI indexes, on the other hand, do not have such bias but are likely to include small-cap companies. Since the FWI technique use market valuation indifferent metrics and includes up to 1000 companies regardless of their market capitalization. Moreover, the additional factors are impacting the FWI methodology differently. The revenue index has the most distinctive SMB bias, following by book value. The HML loadings are, on average, high for the FWI portfolios, where dividend index has the highest tilt towards value stocks. As anticipated, given the design of that index. Implying the factors' high explanatory power for the FWI dividend returns, which means that the index is likely to overweight companies with high book values relative to the market-cap. The profitability factor (RMW), on the other hand, impacts the FWI indexes differently. Three indexes bear negative factor loading, but none of them are statistically significantly different from zero. Although, the dividend index is significant on a 5% confidence level, with the highest factor loading of the whole population, which is theoretically correct given past results of (Nissim and Ziv, 2001). They found that the dividend level provides information about the profitability of a company.

Moreover, the investment ratio of companies is the single factor that influences most of the indexes. However, we cannot draw any statistical conclusion given the relatively high standard error and no difference from zero. Apart from that, we can see from the S&P 500 IVW that the CMA factor explains much of the returns for companies with high book value and residual income. It is showing the same characteristics as the constituents of

FWI indexes. This regression pattern reveals that single metric weighting is biased by one or more of the four systematic risk factors. However, once multiple financial measures are combined, the factor exposure decreases (i.e., Composite index). With these results, I cannot reject the **Hypothesis** I and III from chapter 3. As I have documented in this analysis, the FWI approach withholds better relative and absolute returns over the long term. Besides, the fact that the composite index does not possess a significant HML loading is also proved here, hence confronting the model's most common critiques.

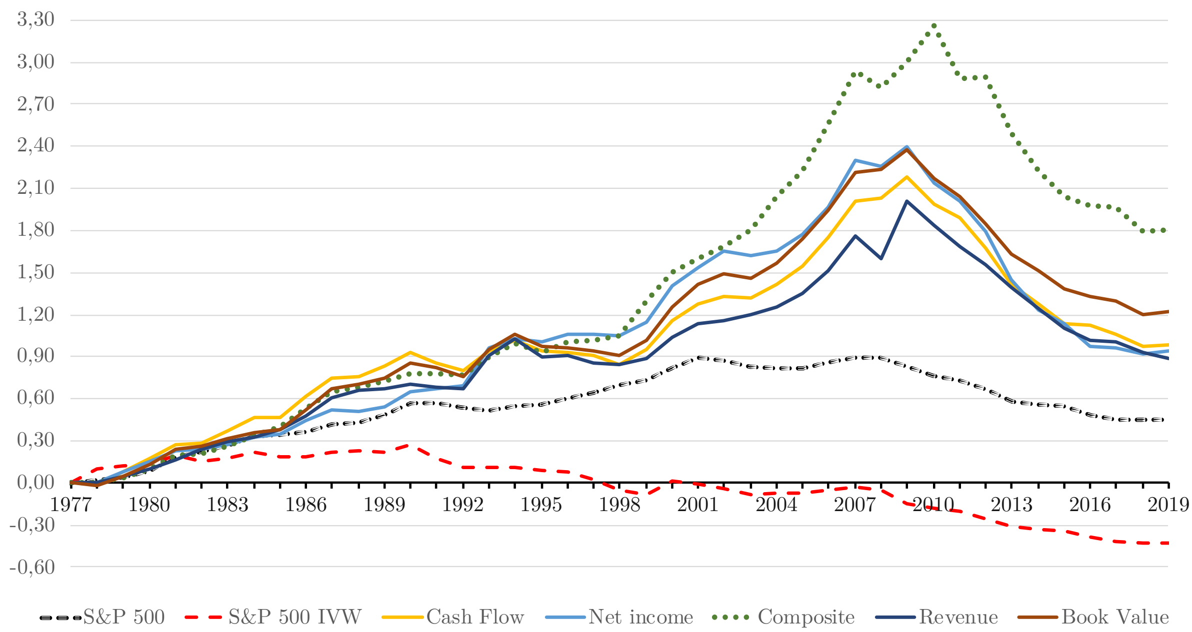
Table 6.2: Performance analysis for the replicated indexes

	S&P 500	S&P 500 IVW	Cash Flow	Net Income	Book Value	Revenue	Composite	Dividend
A. *Relative performance measures*								
Sharpe Ratio	0,351	0,317	0,422	0,461	0,412	0,397	0,443	0,288
Treynor Ratio	0,055	0,052	0,069	0,077	0,069	0,068	0,073	0,049
Jensen's Alpha (pps)	1,287	0,277	2,381	2,604	2,321	2,233	2,810	-0,004
Beta to S&P 500	1,000	0,896	0,893	0,889	0,930	0,921	0,886	0,816
Benchmarking vs S&P 500								
Excess return (pps)	-	-0,770	0,725	1,300	0,980	0,840	0,970	-1,441*
Tracking Error	-	0,054	0,053	0,056	0,063	0,072	0,053	0,067
Information Ratio	-	-0,143	0,138	0,236	0,156	0,117	0,185	-0,215
Benchmarking vs S&P500 IVW								
Excess return (pps)	0,770	-	0,015	2,1***	1,741**	0,016*	1,76**	-0,660
Tracking Error	0,054	-	0,056	0,051	0,061	0,062	0,054	0,048
Information Ratio	0,143	-	0,268	0,412	0,289	0,260	0,321	-0,138
B. *Fama & French five-factor analysis*								
alpha (pps)	0,861*	-0,007	2,016*	2,309*	1,539**	1,623	2,306**	-1,454
	(0.006)	(0,009)	(0,012)	(0,013)	(0,011)	(0,013)	(0,012)	(0,011)
Mkt-rf	1,017	0,957	0,967	0,995	0,928	0,957	0,967	0,935
	(0,031)	(0,044)	(0,059)	(0,063)	(0,061)	(0,063)	(0,058)	(0,052)
SMB	-0,204***	-0,083	0,070	0,000	0,098	0,152*	0,031	0,008
	(0,044)	(0,061)	(0,082)	(0,088)	(0,085)	(0,088)	(0,081)	(0,072)
HML	0,020	0,213***	-0,015	0,147	0,268*	0,077	0,059	0,231*
	(0,051)	(0,072)	(0,097)	(0,103)	(0,100)	(0,103)	(0,095)	(0,085)
RMW	0,030	0,058	0,021	0,080	-0,127	-0,063	-0,013	0,187**
	(0,054)	(0,075)	(0,102)	(0,109)	(0,106)	(0,109)	(0,101)	(0,090)
CMA	0,062	0,234**	0,170	0,084	0,171	0,173	0,101	0,139
	(0,079)	(0,110)	(0,149)	(0,158)	(0,154)	(0,159)	(0,146)	(0,131)
$R^2_{adjusted}$	0,996	0,928	0,846	0,852	0,845	0,870	0,898	0,896

The stars symbolized with * show the statistical significance level —such that *, **, and *** should be interpreted as 10%, 5%, and 1% levels. The excess return statistics are calculated with paired t-test. The value-weighted CRSP portfolio, as calculated by Fama and French, is used to determine betas for the Treynor-Ratios' denominator as well as Jensen's alpha. All indexes have significant exposure to the mkt-rf factor on a 1% confidence level. The comma symbol is used as a decimal separator. The values in parentheses under the coefficients in panel B is the standard error from FF5F regression. All rows with pps are given in percentage points

Figure 6.2 illustrates the historically accumulated alphas adjusted for FF5F risks. The graph is calculated by subtracting achieved returns from expected returns. The composite index represented by the dot green line is the best alpha generating index. The index follows the same trend line as the rest of FWIs but surpassing the S&Ps with a good margin. Moreover, the alpha has diminished for the FWI indexes in the last decade since its peak in 2010.

Figure 6.2: Cumulative five-factor alpha



Given the superior performance of the FWI model, especially the composite index, in the following table 6.3, I show the index's time-dependent performance. I have used the rolling regression technique to estimate the variations of the FF5F equation (5.13) with a ten years rolling window. The purpose of this analysis is to show time variations of each risk factor and the alpha-generating capabilities of the composite index over time, also how it changes during different economic environments. Here, I report the results with a heatmap; low values are colored red and high in green. This spectrum is used for all metrics without the market factor. For the mkt-rf, the color scale is the opposite in order to illustrate the defensive nature of the FWI versus the overall market.

As we see in Table 6.3, the index behaves differently during bull and bear markets and fluctuations in interest rates. An ex-post definition of a bull market is a 20% rally from the previous low and vice versa for bear (Arnott et al., 2005). The 1990s technology/media/telecommunications (TMT) driven bull market shows to had a neutral exposure on the index. We can see that the portfolio had a tilt towards low beta companies, thus kept the overall market risk relatively small. Interestingly, the end of the TMT rally, with the following recession, seems to have slightly impacted the index. Whereas the S&P 500 and the more TMT concentrated NASDAQ100 fell by over 20% respectively each year between 2000-3, the FWI composite decreased by an average of 7%. The following bull market did change the composition of the factor exposure; the HML seems to be the single factor that contributed most, which is a common behavior among investors, with the "flight to quality" mentality shifting investors towards bonds and value stocks after periods with high volatility (Vayanos, 2004). The positive trend of the HML factor is kept almost the same since 2002, with a slight break during the financial crisis. Moreover, the SMB factor is impacting the index negatively after the TMT bubble until 2013, but has regained in the last bull rally. The excess alpha is, to some degree, connected with interest rates, for example, the yield of 90 days T-bill. During low-interest rates, particularly in the last decade and after the financial bubble, the alpha decrease or being negative. The same occurred in the early 2000s and the start of the 90s. Under other conditions, when the interest rates increase, the alpha also improves, for instance, right after the dot-com bubble and rising interests from 2003 to the height in mid-2007. With that being said, I do not find any significant evidence to draw a statistical conclusion. But the pattern is indicating a connection.

Table 6.3: Rolling regression for the composite index with 10 years window

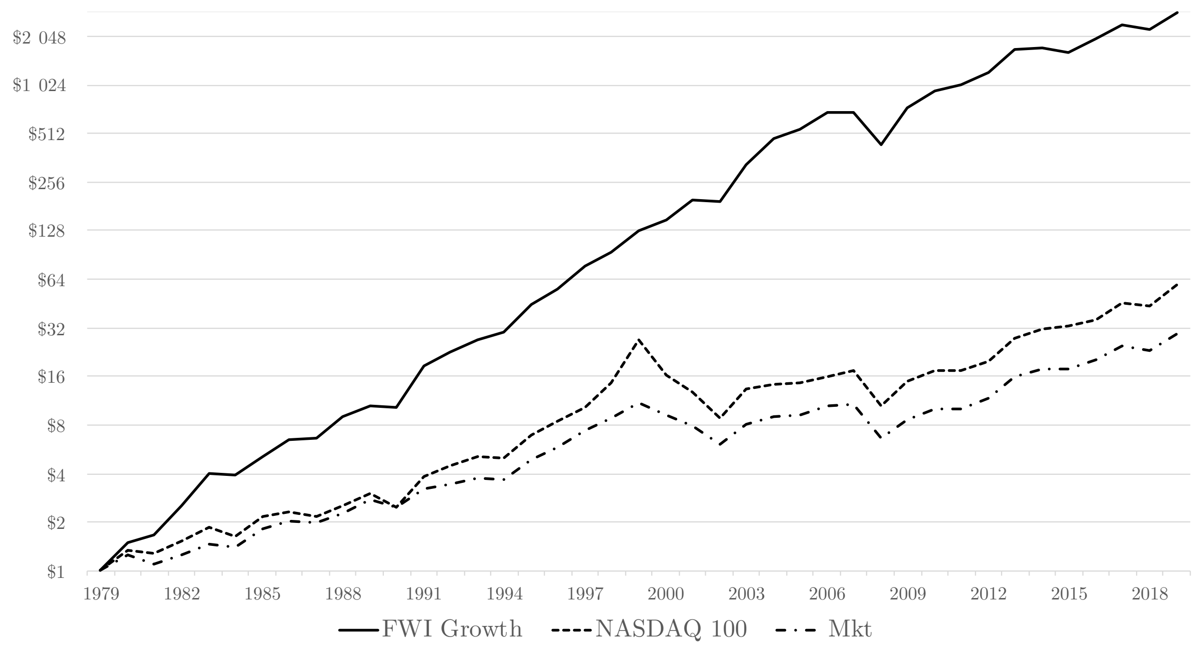
Year	R^2	R^2_{adj}	Mkt-rf	SMB	HML	RMW	CMA	α
1988	0,952	0,893	0,803	-0,131	-0,118	-0,305	0,008	0,106
1989	0,946	0,879	0,726	0,055	-0,183	0,090	0,262	0,073
1990	0,953	0,894	0,868	-0,214	0,140	-0,196	0,115	0,048
1991	0,983	0,962	0,898	0,490	-0,927	2,356	2,305	-0,168
1992	0,984	0,965	0,898	0,509	-0,960	2,406	2,369	-0,172
1993	0,933	0,850	0,943	-0,365	0,150	-0,654	-0,247	0,084
1994	0,926	0,834	0,923	-0,317	0,107	-0,609	-0,167	0,084
1995	0,907	0,790	0,795	0,215	-0,368	-0,267	0,774	0,056
1996	0,949	0,886	0,768	0,237	-0,332	-0,324	0,577	0,064
1997	0,949	0,885	0,810	0,116	-0,151	-0,338	0,325	0,059
1998	0,952	0,893	0,821	0,190	-0,211	-0,271	0,501	0,055
1999	0,945	0,877	0,864	0,165	-0,281	-0,042	0,788	0,033
2000	0,929	0,839	0,841	0,145	-0,204	-0,092	0,615	0,037
2001	0,954	0,896	0,789	0,060	-0,148	-0,194	0,635	0,038
2002	0,986	0,968	0,707	0,061	0,254	-0,393	0,092	0,062
2003	0,981	0,958	0,519	-0,035	0,535	-0,561	-0,339	0,096
2004	0,980	0,954	0,564	-0,031	0,422	-0,478	-0,204	0,084
2005	0,976	0,946	0,665	0,014	0,270	-0,311	-0,068	0,065
2006	0,976	0,946	0,590	-0,031	0,376	-0,462	-0,106	0,074
2007	0,956	0,901	0,596	-0,026	0,269	-0,371	-0,094	0,081
2008	0,981	0,956	1,119	-0,156	-0,021	0,215	0,169	0,058
2009	0,990	0,977	1,084	-0,178	0,147	0,198	0,081	0,044
2010	0,977	0,948	1,026	-0,176	0,266	0,113	-0,073	0,043
2011	0,969	0,931	0,954	-0,242	0,290	-0,161	-0,124	0,050
2012	0,956	0,900	0,863	0,038	0,165	-0,390	-0,587	0,046
2013	0,936	0,856	0,849	-0,064	0,178	-0,205	-0,529	0,033
2014	0,926	0,833	0,875	0,216	0,292	0,025	-0,700	0,017
2015	0,915	0,809	1,003	0,161	0,305	0,439	-0,405	-0,017
2016	0,948	0,883	0,956	0,239	0,029	0,286	-0,377	-0,027
2017	0,983	0,962	0,972	0,333	0,102	0,334	-0,240	-0,045
2018	0,966	0,923	0,965	0,284	0,070	0,249	-0,200	-0,041
2019	0,983	0,961	1,042	-0,099	0,047	0,298	0,125	-0,058

6.2 Results for the growth index

Figure 6.3 exhibits the accumulated dollar growth of the cap-weighted NASDAQ 100, the market factor, and the alternative growth index. The Y-axis is given in the log-scale in order to distinguish the difference between each portfolio. As mentioned in the methodology chapter, FWI growth is constructed such that it reflects the exact same sector composition and number of companies of the NASDAQ 100. The portfolio consists of 100 companies from across seven sectors, weighted by price indifference metric. The Mkt index is the CRSP portfolio as calculated by Fama & French and is included to illustrate the differences.

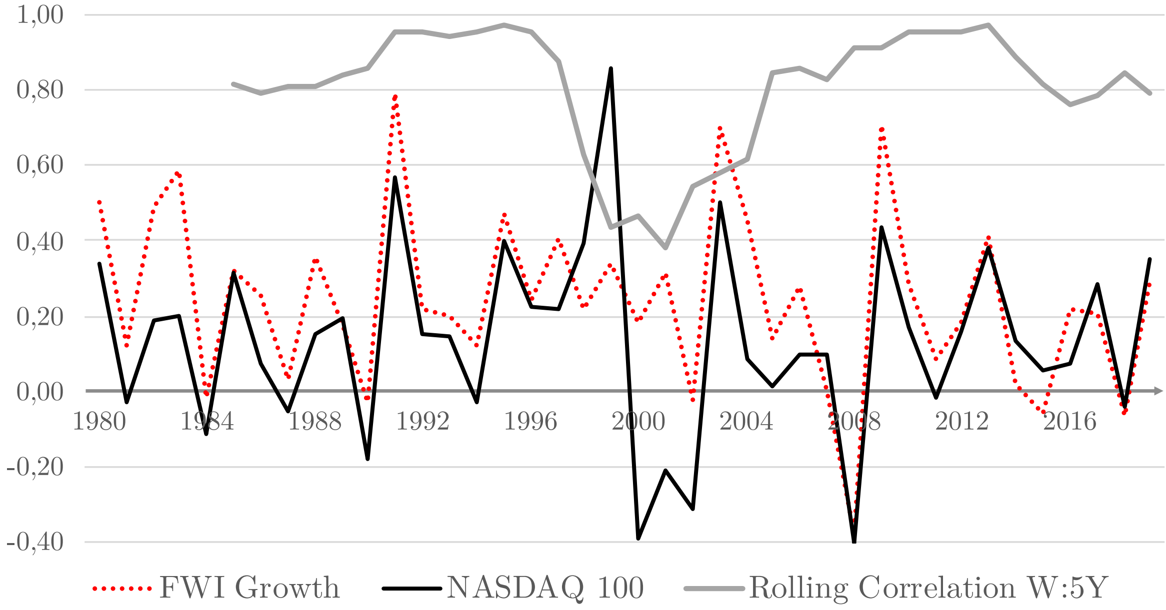
The FWI growth index, with its exponential increase, outperforms the NASDAQ 100 with a factor of 48.6 and has an ending value of $ 2,889. Versus the NASDAQ 100, which accumulates a total dollar value of 59.36.

Figure 6.3: Semi-logarithmic graph illustrating the growth of 1$



The comovement of these two indexes is better illustrated in figure 6.4, which exhibits yearly returns. From this graph, we can see that the FWI growth correlates with the NASDAQ with a high coefficient until the mid-90s. But the correlation decreases as the tech-sector inflates at the end of the 90s. We can also see how the downside risk is minimized when we use price indifferent weighting metrics. As well as the 2003 and 2009 rebounds that are more than the cap-weighted index.

Figure 6.4: Yearly return series of NASDAQ 100 and the FWI growth



The descriptive statistic in Table 6.4 shows the statistical properties of each portfolio. The FWI Growth outperforms the benchmark and have relatively lower volatility than the NASDAQ 100. With an average geometric return of 22.05% the portfolio will double the invested capital approximately every 3.5 years, when we leave the transaction costs out of the equation. Both indexes are slightly positively skewed, indicating a higher mean over the median, and we can expect a larger right-handed tail. The excess kurtosis of both indexes is greater than the zero, and is leptokurtic distributed with heavy tails. Moreover, the maximum and minimum returns show that extreme variations can appear with these indexes. But downside risks are relatively smaller considering the upside potential than the other indexes discussed previously.

Table 6.4: Descriptive statistics for the FWI Growth and NASDAQ 100

	$\overline{X}_A$	$\overline{X}_G$	Volatility	Skewness	Kurtosis	Min $_{\text{Return}}$	Max $_{\text{Return}}$
NASDAQ 100	13,70 %	10,75 %	25,49 %	0,1376	3,6515	-40,54 %	85,59 %
FWI Growth	24,35 %	22,05 %	23,77 %	0,1407	3,3493	-37,81 %	79,50 %

The return series are calculated by using the annually closing price of the NASDAQ 100 and FWI growth index. The period spans from 1979 to 2019, and all numbers are annualized. Volatility is measured in the standard deviation of returns. $\overline{X}_A$ and $\overline{X}_G$ stand for arithmetic and geometric annual returns. The Kurtosis value is in absolute number and is not substracted from 3.

Table 6.5. exhibits the relative performance of the FWI growth versus its benchmark. The corresponding performance ratios for the FWI portfolio are the double of the cap-weighted NASDAQ 100; with an average CAPM alpha of 9.81% pps the FWI index has generated abnormal returns that can not be explained solely by the market movement. Whereas both indexes have generated high alphas consistently over time, with a volatility of the excess returns of 0.13 and 0.157. The IR is considerably higher than the NASDAQ 100, which is as expected, given the profoundly significant excess performance of the FWI portfolio. An IR of 0.933 is considered as an outstanding performance, counting past empirical findings from the U.S. market Goodwin (1998) where the average IR of funds with growth and small-cap strategies have been at 0.25 and 0.41.

Furthermore, the panel B reveals the factor risks associated with these indexes. Both indexes generate, on average, substantially significant alpha-values when adjusted for the risk factors. The FF5F regression can only explain 81.5% of the variations in FWI growth returns, which is the weakest explanatory power among all regression analyses executed for the U.S. I also performed a regression with three factors to investigate the effects of two additional factors, leaving the CMA and RMW factors. The adjusted R^2 decreased typically to 0.80, and the alpha increased to 9.34; also, the residuals increased. Thus, I choose to report with five factors.

The market factor is significant at 1% confidence level for both portfolios. While the FWI index bears a higher market risk beta than the NASDAQ 100, predictably bearing in mind the constituents of the index. The index is heavily concentrated in small-medium-enterprises, where the median company's market cap is around 287 Million USD. The 10th and 90th percentiles are \$ 98.68 million and \$7.303 billion; small enterprises are, in general, more volatile than the big-caps. In addition, both indexes have negative HML factor loading; the NASDAQ 100 has a notable loading of -0.242. As foreseen, considering the index's composition of growth companies from relatively young sectors. Interestingly the RMW and CMA systematic risk-factor exposure are differing for these two indexes. The profitability premium is significant and positive for the FWI portfolio, indicating that firm profitability is a consistent determinant of index returns.

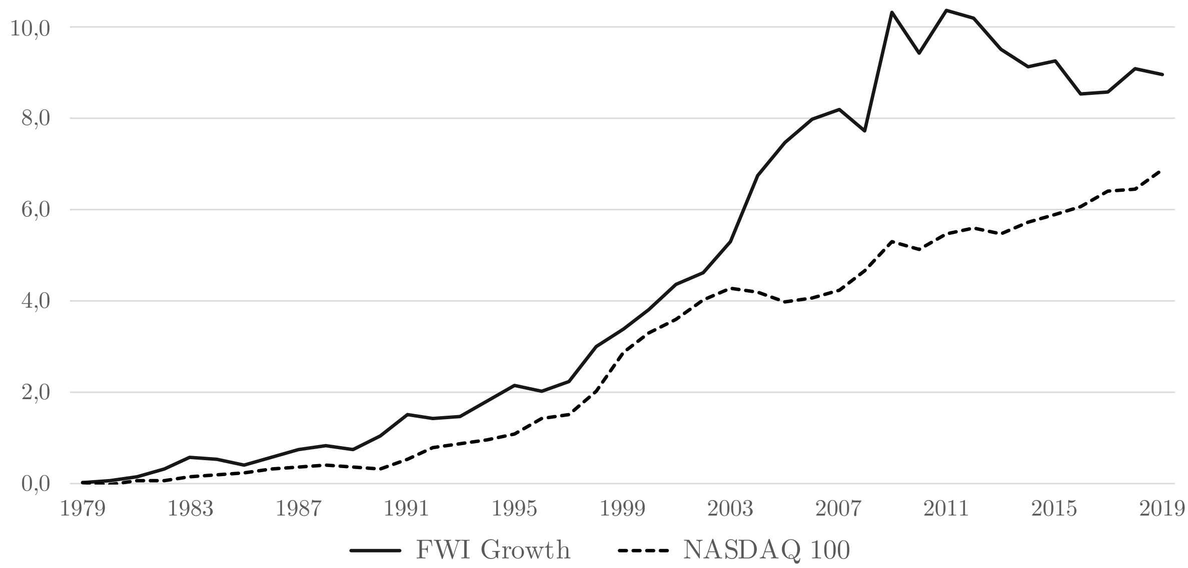
Table 6.5: Performance measures of FWI Growth vs. NASDAQ 100

	NASDAQ 100	FWI Growth	Mkt
A. *Relative performance measures*			
Sharpe Ratio	0,369	0,844	0,270
Treynor Ratio	0,092	0,172	-
Jensen's Alpha (pps)	4,431	9,810	-
Market Beta	1,019	1,165	1,000
Benchmarking vs Mkt			
Excess return (pps)	8,89**	15,56***	-
Tracking Error	0,130	0,157	-
Information Ratio	0,377	0,993	-
Benchmarking vs NASDAQ 100			
Excess return (pps)	-	10,651***	-
Tracking Error	-	0,191	-
Information Ratio	-	0,557	
B. *Fama & French five-factor analysis*			
alpha (pps)	3,579** (0,019)	6,561** (0,027)	
Mkt-rf	1,023 (0,077)	1,169 (0,112)	
SMB	0,345** (0,114)	1,068*** (0,166)	
HML	-0,242* (0,125)	-0,048 (0,183)	
RMW	-0,716*** (0,132)	0,365* (0,192)	
CMA	-0,436** (0,193)	0,226 (0,281)	
$R^2_{adjusted}$	0,927	0,815	

The stars symbolized with * show the statistical significance level *, **, and *** should be interpreted as 10%, 5%, and 1% levels. The excess return statistics are calculated with paired t-test. Fama & French's value-weighted CRSP portfolio (Mkt) is used to calculate betas for the denominator of the Treynor equation as well as the Jensen's alpha. All indexes have significant exposure to the mkt-rf factor on a 1% confidence level. The comma symbol is used as a decimal separator. The values in parentheses under the coefficients in table B is the standard error from FF5F regression. PPS stands for percentage points.

The graph in figure 6.5 displays the accumulated FF5F adjusted alphas for these two strategies. The FWI growths' somehow "extreme" returns diminish when we fix for the bearing systematic risk-factors. The accumulated alpha is nevertheless positive and relatively large compared to the indexes from the replication. But the difference between NASDAQ 100 is narrowed, and the gap created in the early 2000s is closing. With these results, I have robust statistical significant evidence that supports the **Hypothesis IV** synthesized in chapter 3. It exists a better method to gain growth exposure in an index than only relying on the market price.

Figure 6.5: Cumulative five-factor alpha NASDAQ 100 vs FWI Growth

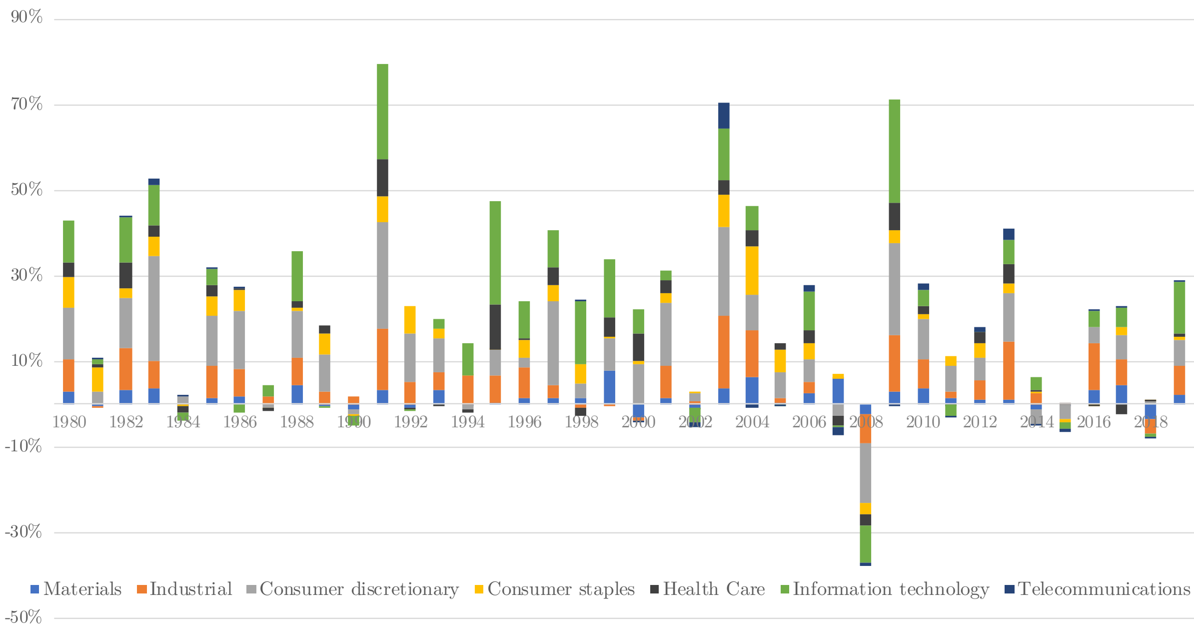


The reason behind this impressive performance is many folded. A subsequent comprehensive inquiry of the company and sector composition of the FWI growth index can reveal one of the contributing factors; the historical sector bets. The model's ability to capture new profitable growth cases is distinctive and remarkable; in the following figure 6.6, we can see the historical weighted returns from each sector. Here we have four leading sectors that account for the majority of the historical returns. Especially the information technology and consumer discretionary are the two best performing areas, which also have been the best sectors in the economy in past decades[15]. Technology companies are also relatively more intangible intensive Ciftci et al. (2014) since our model

[15] https://eresearch.fidelity.com/eresearch/markets_sectors/sectors/si_performance.jhtml?tab

overweight companies that have high ROTA ranking, which is an indirect measure of the productivity of the intangible assets (Clausen and Hirth, 2016). It is, therefore, a reasonable explanation of the extraordinary performance of our index. We have also to consider the structural shift in the U.S. economy, from the tangible to intangible intensive economy (Ciftci et al., 2014). Corrado et al. (2009) found that just 8% of economic growth can be assigned to the regular 'bricks and mortar' assets investment. Since these sectors are relatively young, the most productive companies from these sectors with high ROTA ranking are classified as intangible-intensive firms. The higher ranking also pays of as higher returns in the stock prices.

Figure 6.6: The FWI growth index's weighted sector returns



6.3 FWI ESG results

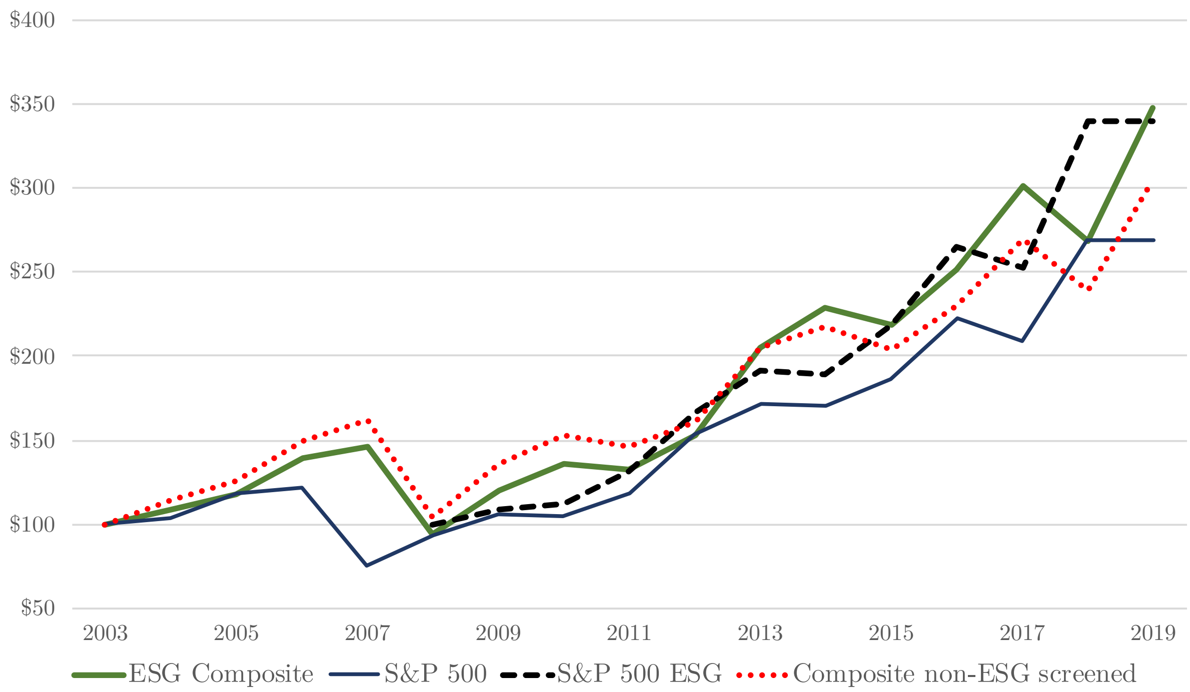
This subsection is dedicated to the results from my third contribution to the literature; ESG screening and weighting as a non-financial fundamental metric. The methodology is backtested from 2003 to 2019 and is compared against commercially available cap-weighted indexes. Here I have used two variants of the S&P500 indexes again; the S&P 500 ESG is a newly launched index from Standard Poor's[16] and is excluding companies that do not satisfy a set of ESG related criteria. The index is cap-weighted like the ordinary S&P 500;

[16]From the S&P Dow Jones Index Methodology: https://www.spglobal.com/_media/documents/the-sp-500-esg-index-integrating-esg-values-into-the-core.pdf

as of December 2018, almost 183 constituents of the S&P 500 were excluded, totaling 24.36% of the index's overall market cap. I used the Standard Poor's website to retrieve annually historical backtested data. The start date is from 2009, which does not overlap with the rest of the indexes precisely, thus resulting in the relatively smaller risk measures in table 6.6. Since the financial crisis of 2008-9 is not accounted for in the return series. I use this index to examine the role of ESG and see what if the weighting mechanics or ESG is a decisive factor for index returns.

Figure 6.7 illustrates the hypothetical growth of $ 100 invested in each of these four passive strategies. The ESG screened indexes accumulate most returns and have an ending value of respectively $ 347.50 and 339.41. Surprisingly quite similar, but when I adjust for the lagging period. The ESG Composite surpasses the S&P 500 ESG by $30.58 in ending value, but the non-ESG indexes are still underperformers. Indicating a positive tendency of ESG on returns.

Figure 6.7: Growth of 100$ invested in ESG screened indexes and alternatives



The ESG composite index is a composition of cash flow, book value, revenue, and combined ESG score. The index constituents vary through the period, whereas at the beginning, the index consisted of only 54 companies. But the number of companies was increased to 250 as more companies satisfied the criteria described in chapter 5.2. The composite non-ESG screened is the same index as presented earlier in this chapter.

Table 6.6 A summarizes the statistical features for the whole period, starting from 2003 to the end of 2019. The ESG-screened composite index poses an excess return over one pps annually, with almost the same volatility. Thus implies a definite tendency for ESG as a screening and weighting factor. This tendency is somehow vague if we look at the performance in the last decade, as shown in subtable B. The ESG friendly portfolios have performed relatively better than the non-ESG. Still, the standard deviation of returns remains almost the same and even increase for the FWI ESG. Further on, the FWI indexes are slightly negatively skewed, which indicates the higher median over mean. These indexes have longer tails to the left and had relatively higher downside risk to upside potential in the last decade. The last decades' bull market had three years of modest negative intra-year return. As an index investor with a long-term perspective, this relatively limited period is biased by the overall equity-market rally; therefore, the effect of ESG-factor is challenging to distinguish.

Table 6.6: Descriptive statistics for ESG screened indexes

	ESG - Composite	Composite non-ESG	S&P 500	S&P 500 ESG
A. 2004-2019				
$\overline{X_A}$	9,58 %	8,63	8,27	
$\overline{X_G}$	8,10 %	7,20	6,89	
Volatility	16,68 %	16,31	15,70	
Skewness	-1,15	-1,21	-1,49	
Kurtosis	2,12	3,14	3,93	
Min Retrun	-35,88 %	-35,95	-38,49	
Max Return	33,86 %	31,30	29,60	
B. 2010-2019				
$\overline{X_A}$	12,10 %	9,11	11,82	13,59
$\overline{X_G}$	11,23 %	8,36	11,22	13,00
Volatility	13,78 %	12,68	11,53	11,57
Skewness	-0,188	-0,130	0,077	0,081
Kurtosis	2,26	2,04	2,13	2,12
Min Retrun	-11,02 %	-11,56	-6,24	-4,63
Max Return	33,86 %	27,62	29,60	34,40

The return series are calculated by using the annually closing price of the indexes. The period is divided into two subperiods, and all numbers are annualized. Volatility is measured in the standard deviation of returns. $\overline{X}_A$, and $\overline{X}_G$ stand for arithmetic and geometric annual returns. The Kurtosis value is in absolute number and is not subtracted from 3.

The following table 6.7 provides further clarity on the effect of ESG as an extra factor for both screening and weighting purposes. The relative performance ratios are higher for the ESG portfolios, despite after adjusting period for the S&P 500 ESG. Thus, reflecting the ESG screened portfolios' consistent excess returns on a 10% confidence level over the S&P 500 non-ESG. The tracking-error also confirms the persistence of the excess returns, with a relatively low standard deviation. Moreover, the information ratio is also considerably high with values over 0.40 and 1.5. Looking only at the subtable A, the S&P 500 ESG looks like the best case. However once we adjust for the additional systematic factor-risks from, the dependable performance reveals other features. The cap-weighted ESG portfolio has the highest negative alpha with a value of negative 1.665 pps. The best index in terms of alpha is the fundamentally weighted ESG index following by the regular S&P 500. The FF5F regression do explain much of the return variations with adjusted squared R values over 0.95 except for the FIW Composite non-ESG. All indexes had almost perfectly movement with the market-factor.

The considerably high (negative) SMB exposure for ESG portfolios is of no surprise, especially the S&P 500 ESG, with a significant loading of 0.743 at a 5% level. The FWI ESG portfolio reflects the same, with a median market-capitalization of 44 billion USD and the 90th percentile of $ 300 billion. As mentioned previously, large companies are more ESG focused than small-cap companies Giese et al. (2016) since they are likely profiting from the economies of scale in the ESG implementation process.

Moreover, the value premium is relatively higher for the fundamentally weighted indexes, consistent with previous findings in this chapter. Interestingly, the ESG version of the cap-weighted index inherent a bigger value factor loading, but the result is not significant. The same for the profitability and investment factors, which are relatively high. Intuitively, it can indicate that profitable companies are more likely to adapt an ESG-friendly business process. Furthermore, the CMA factor, which explains the difference between firms with low and high investment policies, has a positive relationship with the returns of ESG-portfolios but not significant. The factor loading is highest for the S&P 500 ESG, but the high value of standard error (SE) indicates some misfitting of the data by the model. The relatively high SE may come from limited observations, as we have only 10 data points, although this issue can be fixed by using monthly data.

These results appear that including ESG in the FWI framework has a positive impact on the index's returns. But, once we adjust for bearing systematic risk, the analysis indicates that the excess returns are derived from additional risk. Thus I do not find statistically significant evidence to keep the **Hypothesis** II.

Table 6.7: ESG Index Performance analysis

	Composite	ESG - Composite	S&P 500	S&P 500 ESG
A. *Relative performance measures*				
Sharpe Ratio	0,347	0,396	0,337	0,975
Treynor Ratio	0,056	0,063	0,049	0,118
Jensen's Alpha (pps)	0,310	1,102	0,934	5,624
Benchmarking vs S&P 500				
Beta	0,998	1,047	1,000	0,959
Excess return (pps)	0,366	1,30*	-	1,76*
Tracking Error	0,045	0,053	-	0,056
Information Ratio	0,081	0,449	-	1,534
Benchmarking vs S&P 500 ESG				
Beta	1,051	1,139	0,951	1,000
Excess return (pps)	-4,470	-1,487	-1,760	-
Tracking Error	0,035	0,043	0,035	-
Information Ratio	-1,377	-0,928	-1,534	-
B. *Fama & French five-factor analysis*				
alpha (pps)	-1,201 (0,034)	-0,350 (0,023)	-0,602 (0,004)	-1,665 (0,008)
Mkt-rf	0,998 (0,112)	1,018 (0,070)	1,010 (0,038)	1,003 (0,078)
SMB	0,069 (0,278)	-0,174 (0,173)	-0,2441** (0,096)	-0,743** (0,191)
HML	0,206 (0,174)	0,186 (0,108)	0,034 (0,060)	0,127 (0,114)
RMW	0,236 (0,351)	0,144 (0,219)	-0,016 (0,121)	0,350 (0,183)
CMA	-0,280 (0,301)	0,033 (0,015)	0,030 (0,104)	0,359 (0,228)
R^2 adjusted	0,912	0,956	0,984	0,952

The stars symbolized with * show the statistical significance level—such that *, **, and *** should be interpreted as 10%, 5%, and 1% levels. The excess return statistics are calculated with paired t-test. Fama French's cap-weighted CRSP portfolio is used to calculate betas for the denominator of the Treynor equation as well as the Jensen's alpha. All indexes have significant exposure to the mkt-rf factor on a 1% confidence level. The comma symbol is used as a decimal separator. The values in parentheses under the coefficients in table B is the standard error from FF5F regression. The time-series for S&P 500 ESG regression is from 2010-2019. PPS stands for percentage points.

6.4 The Norwegian FWI

The end of this chapter is devoted to the study of fundamental indexation technique for the Norwegian market. As I have mentioned earlier in the thesis, the FWI model has been only examined once for the Norwegian equity market by Walkshäusl and Lobe (2010) in a limited scope. In this subsection, I present a comprehensive analysis of the FWI in Norway within almost two decades, covering various market conditions.

The following figure 6.8 shows the theoretical growth of 100 kr invested in these strategies. Every index constructed with the FWI technique outperforms the benchmark index OSEBX. They all have the same trend-line and are affected by the same macro-shocks, the Financial crisis (2008-9), European debt crisis (2011), and the oil price plunge (2014-16). The best single weighting factor is the total revenue, which exceeds the OSEBX by 187% with a lower yearly standard deviation.

Each index is constructed by using a three-year trailing average of each accounting size and is rebalanced annually. The maximum weight of a company is 10% since a handful of companies account for over 50% of the accounting values. Therefore, by having no limit, the index performance would be heavily skewed towards specific sectors and companies. The $0 \leq W_i \leq 0.1$ weight constraint is also in accordance with the legislative compliance in Norway cf. verdipapirfondloven § 6-2 [17]. By doing so, I am able to give the findings from this thesis a practical implication. The composite index is equal-weights of cash flow, revenue, EBITDA, and book value.

[17] Lov 25. november 2011 nr. 44 om verdipapirfond (verdipapirfondloven – vpfl.)

Figure 6.8: Growth of 100 kr invested in alternative and traditional indexes

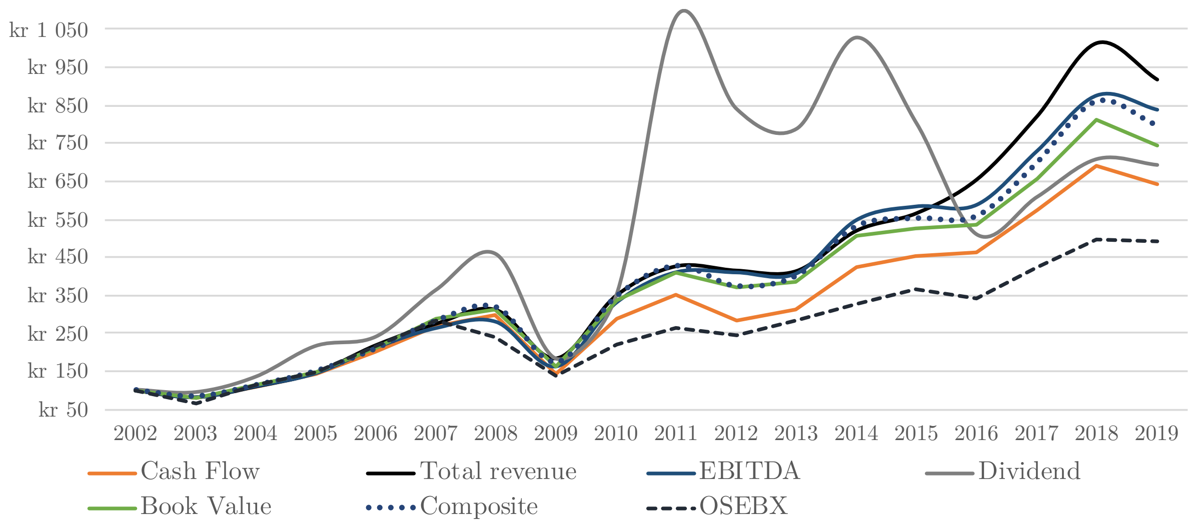


Table 6.8 exhibits detailed statistics for each index. The average arithmetic means of FWI portfolios are above 2.23 pps when excluding the dividend index. The trade off between additional risk and excess return is favorable for revenue, EBITDA, and the composite index. Every FWI indexes are positively skewed, where the dividend index has the highest value, indicating heavily fat tails, which reflects the extreme variations in returns. The frequency of returns exceeding the mean is greater for all FWI indexes. The excess kurtosis is also positive for every index except for the dividend and OSEBX. When omitting the dividend index, the best maximum and minimum values are in revenue index by having the least downside risk and second-best upside potential.

The composite risk and return results are almost the same as the findings of (Walkshäusl and Lobe, 2010). They observed an average arithmetic return of 16.03% and volatility of 23.45% producing a Sharpe Ratio of 0.570 [18]. By treating the returns from the financial crisis in 2008 and the subsequent rebound as outliers, the volatility drops to 20.73%.

[18]The results are for the composite index, consisted of book value, cash flow, dividends and sales. The backtested time period was (1988-2007)

Table 6.8: Descriptive statistics for the Norwegian market

	Cash Flow	Revenue	EBITDA	Dividend	BV	Composite	OSEBX
$\overline{X}_A$	15,61 %	16,28	16,28	23,88	16,39	16,54	13,98 %
$\overline{X}_G$	11,55 %	13,94	13,31	21,26	12,52	12,99	9,81 %
Volatility	30,86 %	27,34	29,64	59,41	31,13	29,72	28,05 %
Skewness	0,34	0,39	0,83	1,23	0,67	0,57	-0,052
Kurtosis	4,28	4,13	4,87	1,69	4,69	4,52	2,59
Min Retrun	-51,66 %	-40,77	-41,88	-60,28	-47,84	-45,56	-42,66 %
Max Return	59,27 %	68,41	66,58	91,75	80,45	62,54	67,65 %

The return series are calculated by using the annually closing price of the indexes and are reweighted once a year. Volatility is measured in the standard deviation of returns. $\overline{X}_A$, and $\overline{X}_G$ stand for arithmetic and geometric annual returns. The Kurtosis value is in absolute number and is not subtracted from 3.

In panel A of Table 6.9, we can comprehend each strategy's real performance in relative terms. Five out of six indexes can show a better Sharpe Ratio than the benchmark. The best Sharpe value is held by the revenue weighted portfolio, following by the multi-metric, composite index. The Treynor ratio and CAPM alpha for the FWI portfolios is also above the benchmark, here the EBITDA index has the highest alpha value of them all. Further on, as expected, the majority of the FWIs have a defensive movement with the OSBEX, and at the same time, produced excess returns. The additional returns from the FWI strategy appear to be consistent over time, addressing the positive information ratio.

The comparatively excellent performance of the revenue weighted index still persists after adjusting for the systematic risk factors; size and value premium. As we see in panel B, the alpha value of 4.22 pps for the revenue index, which is significant at a 10% level, is the best performing single metric. All indexes have an excess alpha over two percentage points, which is considerably higher than the OSEBX. The three-factor model does explain over 80% of the return variations for the fundamentally-weighted indexes and over 95% cap-weighted benchmark. Moreover, the beta loading to market risk is neutral and, in many cases, defensive for the FWI constructed indexes. Two out of six indexes have a factor loading over 1.00, and (revenue, EBITDA, book value, and composite) indexes are less riskier than the market portfolio. Moreover, all six indexes are tilted towards big-cap companies, with (cash flow, dividend, and book value) being the ones with the highest negative exposure to small-cap firms. The value premium factor loading is relatively higher for the FWI portfolios, whereas the revenue and EBITDA indexes are less exposed to the HML factor than the OSEBX.

These results are more or less as expected; the value tilts that the FWI model creates are recorded previously in this thesis for the American equity market. It is also consistent with the past results Walkshäusl and Lobe, which observed a highly significant HML exposure for global FWI indexes. However, negative SMB exposure is surprising, especially when the loadings are above 0.5 for five out of six indexes. This is reasonable given the 50th percentile market-cap of the FWI index being at 2.458 billion USD and 90th percentile at 32.07 billion, which is by the definition of Oslo Stock Exhange[19] big-cap companies.

Table 6.9: Performance measures for the Norwegian market

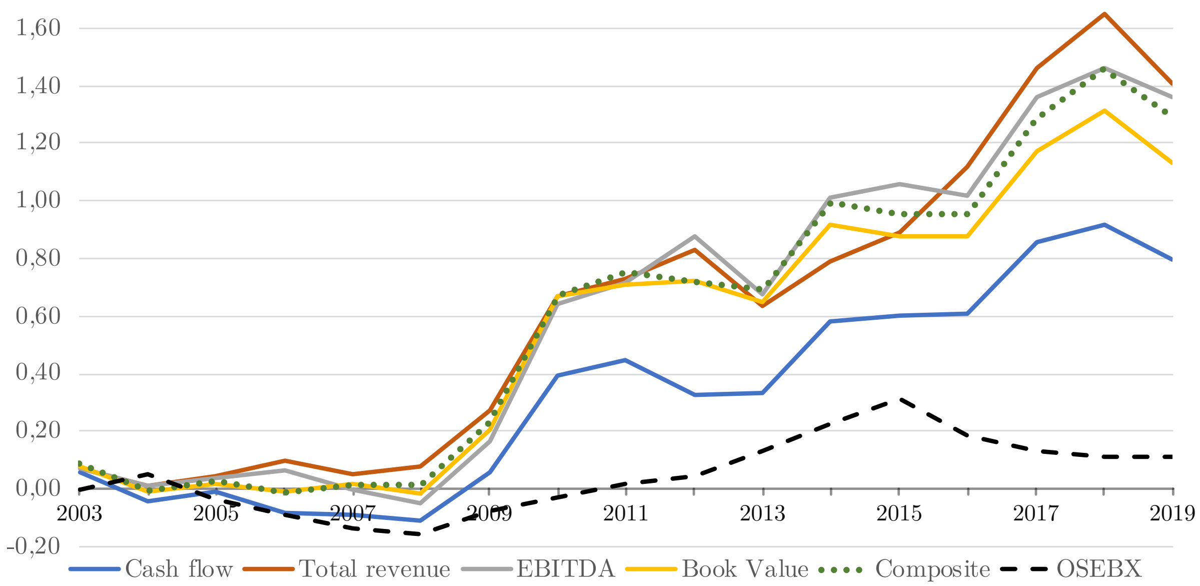
	Cash Flow	Revenue	EBITDA	Dividend	BV	Composite	OSEBX
A. *Relative performance measures*							
Sharpe Ratio	0,400	0,477	0,440	0,347	0,422	0,447	0,382
Treynor Ratio	0,129	0,159	0,145	0,144	0,137	0,146	0,107
Jensen's Alpha (pps)	5,850	7,596	7,700	4,341	6,967	7,373	1,450
Benchmarking vs OSEBX							
Beta	0,956	0,821	0,899	1,138	0,959	0,909	1,000
Tracking Error	0,153	0,156	0,158	0,516	0,157	0,155	-
Information Ratio	0,106	0,148	0,145	0,375	0,153	0,165	-
B. *Fama & french three-factor analysis*							
alpha (pps)	2,100	4,220*	3,450	2,345	2,820	3,520*	0,076
	(0,065)	(0,074)	(0,052)	(0,112)	(0,098)	(0,079)	(0,021)
Mkt-rf	1,044	0,925	0,964	1,178	0,994	0,989	1,005
	(0,116)	(0,109)	(0,135)	(0,156)	(0,112)	(0,122)	(0,108)
SMB	-0,584*	-0,388	-0,555	-1,819	-0,634*	-0,582	0,005
	(0,324)	(0,304)	(0,378)	(0,491)	(0,356)	(0,341)	(0,303)
HML	0,197	0,043	0,122	0,319	0,230	0,216	0,127
	(0,274)	(0,257)	(0,319)	(0,180)	(0,301)	(0,289)	(0,256)
R^2 adjusted	0,857	0,837	0,789	0,879	0,832	0,830	0,950

The stars symbolized with * show the statistical significance level—such that *, **, and *** should be interpreted as 10%, 5%, and 1% levels. The mkt factor is used to calculate market betas for the denominator of the Treynor equation as well as the Jensen's alpha. All indexes have a significant exposure to the mkt-rf factor on a 1% confidence level. The comma symbol is used as a decimal separator. The values in parentheses under the coefficients in table B is the standard error from FF3F regression. The systematic risk data for the regression is retrieved from:
http://finance.bi.no/~bernt/financial_data/ose_asset_pricing_data/index.html

[19]From Equity Indices - Index Methodology retrived from: https://www.oslobors.no/obnewsletter/download/5dd28398ba24cdbe4e3aa64b4a390dae/file/file/2020-01-01\%20Oslo\%20B\%C3\%B8rs\%20-\%20Equity\%20Indices\%20-\%20Index\%20Methodology.pdf.

The graph in figure 6.9 illustrates the total three-factor alpha produced by each of these strategies. As we see revenue weighted index is the best performing, with an accumulated alpha of 140.86% versus the OSEBX that yielded a total alpha of 11.16% between 2003 to 2019. The revenue index's total return for the whole period has been a whopping 819.98% were over one-quarter of the returns have been a pure performance that can not be explained by the additional systematic-risk factors. These results are in line with past observations in the U.S. market (Arnott et al., 2005). They recorded that revenue as a single metric weighting scheme was the second-best index with an ending value of $ 182.05 right behind sales with (184.95).

Figure 6.9: Accumulated three-factor alpha in the Norwegian market



7 Discussion

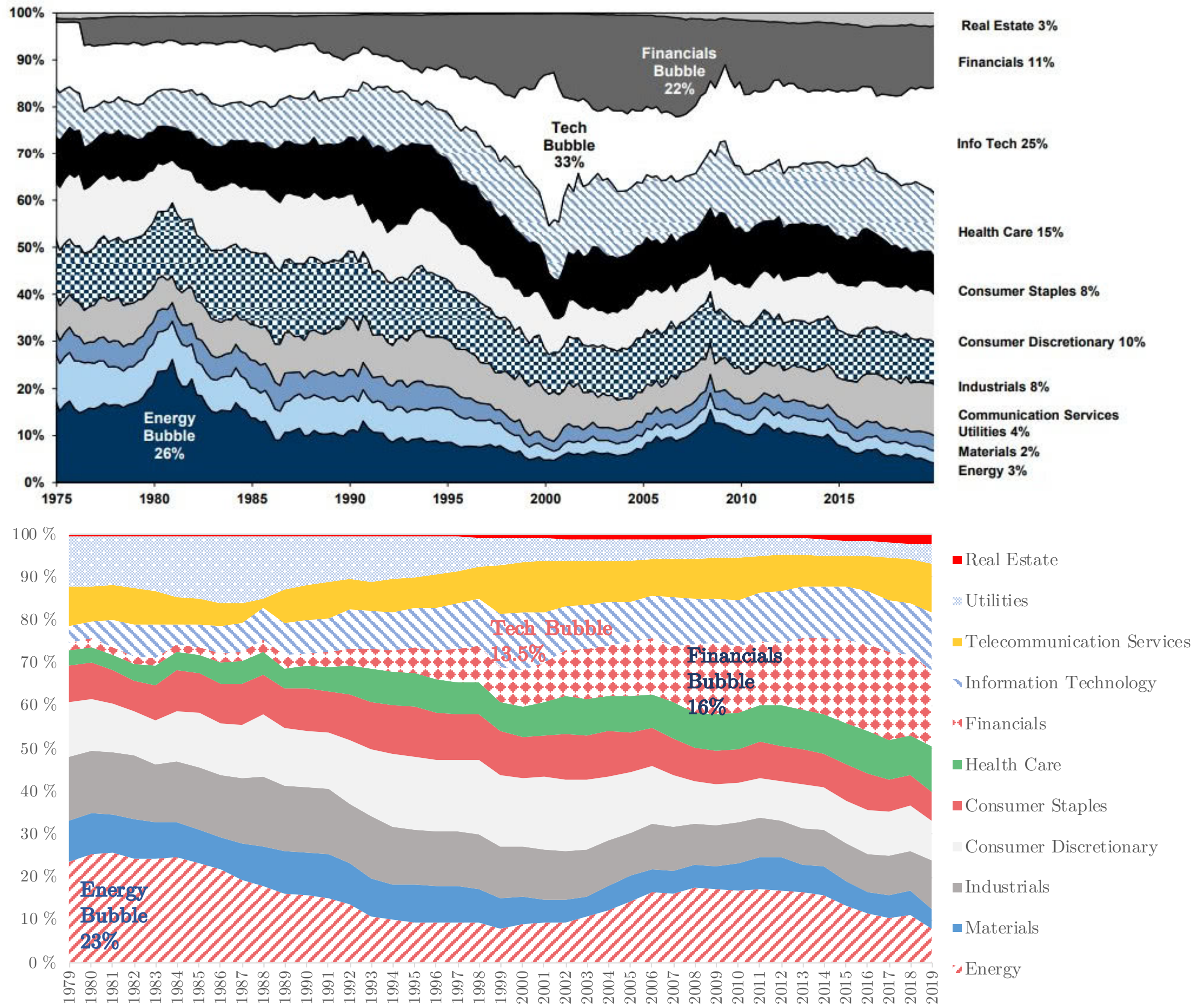
The emerging position of the FWI methodology and supporting evidence of the superior qualities are posing new challenges for both the practitioner and academia. Hence, the new empirical evidence presented in the previous chapter can be interpreted in different directions. In this chapter, I evaluate the findings from various perspectives. First, I start to address the historical sector "bets" that each indexing strategy makes. So to explain how structural changes in the economy affect the sector allocation. Then, I mention some of the FWI critiques that I have not encountered in this thesis.

7.1 Historical sector composition and bubbles

If we decompose past sector allocation that each of the models makes, we are able to evaluate the underlying features of these strategies. One of the central critiques of the cap-weighting approach is that the model overweight overvalued stocks and underweight undervalued. This critique is, without a doubt, a severe imperfection of the model. The idea of always trusting the price as an accurate indicator of the fair value is the rationale that drives these sector bubbles. As well as the mechanical rebalancing of the passive strategy that only relies on one factor to judge an investment. To illustrate this claim, I have calculated the actual sector weights from each industry and highlighted specific years of importance. Then we can compare each of the strategies and discuss the flaws.

The graphs in Figures 7.1 illustrate the sector allocation of each strategy. On average, the sector weights are quite similar. However, the FWI weights are smoothed out and do not has these spikes. S&P, on the other hand, tends to oscillate and is profoundly affected by the sector bubbles. The one sector bubble that deviates most is from the dot-com period. S&P index had a total exposure of 33% towards the information technology sector, whereas the FWI model assigned only 13.5%. The 20% deviation is considerable and causes a return drag for the index investor. The sector tilts from the energy, and financial bubbles are also notable. We can see from the FWI graph that these sectors had strong financials, but the market sentiment affected the prices to inflate. The largest sector in the FWI index is financials (17.7%) followed by IT (13.75%) as of the beginning of 2020, which deviates a lot with the S&P 500. Especially the IT sector that accounts for 25% of the index, with a Shiller P/E of 34.10 as of May 2020. The historical sector growth is also a good indication of the overall direction of the economy. That reflects the structural shift of the American economy towards service sectors. Health care, financials, and IT sectors have the highest growth rate among all sectors in the FWI index, with an average compounded annual growth rate (CAGR) of 4.15%. Hence, guiding the future expectation for equity returns.

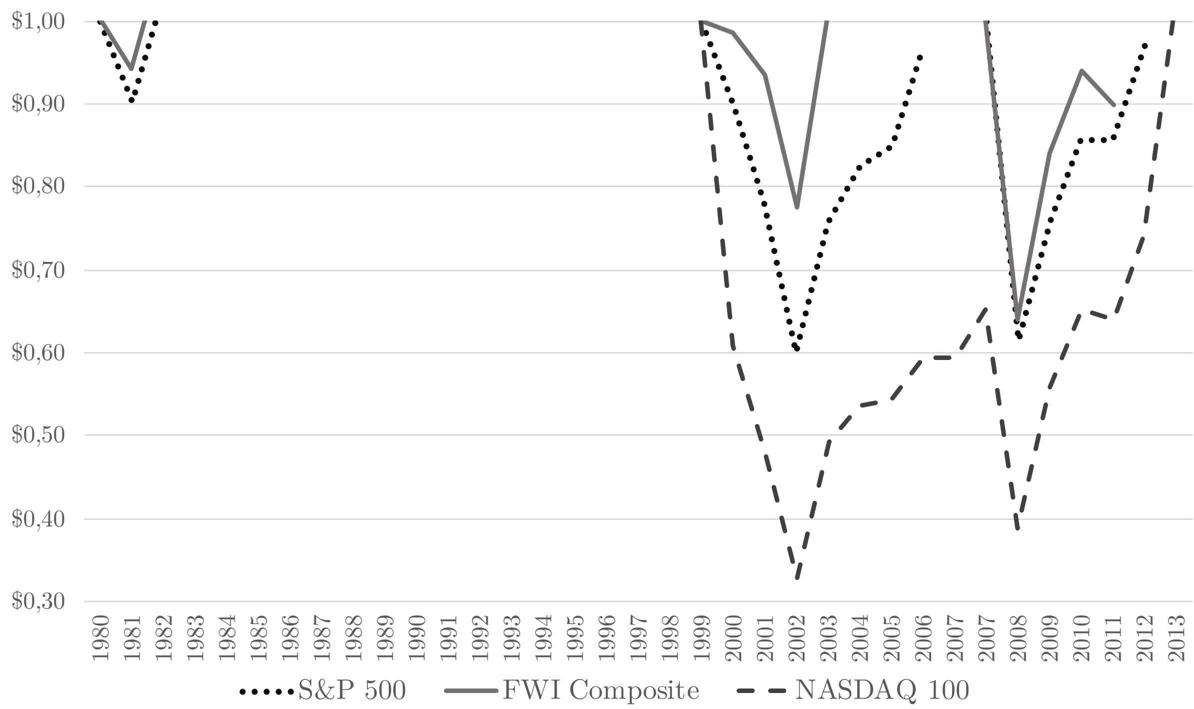
Figure 7.1: Sector composition of the S&P 500 (1974-2020E) and The FWI Composite



The upper graph is from the Goldman Sachs Global Investment Research as of March 31, 2020. The next graph is my calculations for the FWI composite index by assigning each company weight to its corresponding GIC sector code. The sector values are annually and aggregated company weights of each GIC sector. The FWI index has no company weight constraints, whereas the highest value is recorded for AT&T Corporation from 1979 with an index weight of 6.1%.

If we assume that the index investor has a long-term horizon over five years, then he/she should be indifferent to these bubbles? Since we have over a five-year horizon and the market is capable of evaluating securities fairly—this investment approach and assumptions can be very costly when an investor turns a blind eye to the fundamentals. In the following Figure 7.2, I have illustrated the recovery time of market crashes. As we see, the crashes hardly affect a cap-weighted index, and the recovery is also slower. If we further assume that an investor had foolishly trusted the market price and invested a dollar at the height of these bubbles. The regain of the one-dollar value would be different during each crash. The energy bubble crash had little impact on both indexes. However, the dot-com bubble burst, which was primarily a growth crash, did strike the cap-weighted portfolios heavily. The NASDAQ 100 used over 14 years to gain up to the historical valuation heights of 1999. Whereas, the S&P did regain by late 2006 but was hit by another crash that it used over five years to recover. The financial crisis, which was to some extent, a value crash did affect the FWI index, but the recovery was faster. I have left the FWI growth out of this graph since it was not affected by the 2000s crash and the rebound of the 2009 did recover all the loses of 2008. The limited downside of the FWI model is also pointed out as one distinctive advantageous of the methodology by other studies (Pysarenko et al., 2019; Arnott et al., 2005; Balatti et al., 2017).

Figure 7.2: Market crashes and rebounds



7.2 Limitations of the methodology

We can not discuss an investment strategy without pointing out the practical limitations that a practitioner may face. Blitz and Swinkels (2008) express major barriers that a fund manager can face when replicating the index. They point out three notable restrictions; disequilibrium, absence of buy-and-hold strategy, and subjective choice of weighting metrics. They argue that the FWI portfolio cannot be held in equilibrium by every investor, and states that *"For every stock that is overweighted by fundamental investors, there must, by definition, be some other investor who actively underweights the same stock and vice versa."* The idea of "market-clearing-portfolio" from the CAPM theory heavily influences this critique. Blitz and Swinkels do not account for the flaws of this MCP utopia and that irregularity exists in the market, as expressed previously in this thesis. In a world where investors have different views on the market, we should expect that the cap-weighting and FWI strategies coexist—thus enabling investors to benefit from each other.

The lack of a passive buy-and-hold strategy is a drawback that may oppose some challenges for the practitioner. Since an FWI portfolio requires some rebalancing strategy, considering the continuous changes in the market price that push weights away from their pre-calculated target levels. Therefore, in the presence of transaction costs, these constant rebalancing would drain the returns and probably cause a negative momentum-factor exposure Blitz and Swinkels (2008). On the other hand, the FWI approach is contra-trading against investor sentiment; therefore, the costs may be surprisingly low (Arnott et al., 2011, p. 178). Nonetheless, negative momentum-factor exposure will persist. In this thesis, I have not considered the effect of transaction costs in my calculations, since the whole concept of an index is hypothetical. Each investor's unique circumstances also complicate the implementation details mentioned in this paragraph—such as; account size, vehicle preferences, brokerage channel, and custodial arrangements. This issue can be a proposal for further research. In this thesis, I have documented the superior hypothetical performance of the FWI methodology. However, the most exciting aspect now is to examine the performance of a tradable FWI fund. Lastly, the critique of the selection process of which fundamentals figures are used to construct an index. The opponents of the FWI claim that this selection process is subjective, which is correct. In this thesis, I have used different accounting metrics for the American and Norwegian markets because of each market's different characteristics and the availability of the data. All these choices have been subjective, for no doubt. I have also shown the shortcomings of single metric weighting and how factor biased these indexes get. However, when using multiple metrics to measure each position's weights, the index portfolios get more robust. Then, I can argue for the objectivity of the composite index, since a multidimensional measure of company size is more reliable than relying upon one single-metric.

7.3 Concluding discussion

From an economic-centric standpoint, cap-weighting is growth tilted to the same extent as the fundamental weighting is value tilted relative to the market. Consequently, giving us two completely different equity indexes; a market-weighted index and an economic-weighted index (Arnott et al., 2019). The distinctive differences are in the weighting and rebalancing mechanics of these two strategies. The FWI methodology has a built-in buy low and sells high discipline during rebalancing; Whenever the price sours more than the underlying fundamentals, it sells and buys more when the price decreases, vice versa for the cap-weighting. These indexes also have entirely different investment styles, and when we weigh stocks according to the size of the business, growth stocks will trade by premium multiples, so they are reweighted down. Whereas value stocks are trading "cheap" relative to their economic footprint, these are weighted up. Therefore the FWI model always has a start value tilt, but its a constantly changing value exposure since FWI concentrating against every stock's price movements. With cap-weighting, on the other hand, we are systematically overweight every overvalued stock and underweight every undervalued stock. The problem is to determine which stock is over-or undervalued. However, by using another weighting scheme, we can randomize this error. On the other hand, with the FWI model, each stock's weight is not correlated with the over or undervalued factor. Nevertheless, once we tie the weight to the price, we create a start correlation between what the error is in the price and the weight in the portfolio. With cap-weighting, the model overloads the overvalued stocks even though we do not know which one. We lack the same knowledge when using fundamental indexation, but the error is randomized when utilizing price indifferent weighting schemes, so they cancel. Thus, instead of a systematic drag on performance, its a systematic boost.

8 Conclusion

In this thesis, I investigated the FWI methodology in various forms. First, I successfully replicated the Arnott et al. (2005) study and showed new empirical evidence, that proves the endurance of the model. The results indicate the weak risk and return profile of conventional cap-weighted indexes. I also repeated the process on the Oslo Stock Exchange and reported an alternatively better passive portfolio with higher returns and less risk. The relatively higher values for SMB (negative exposure) and HML (positive factor loading) somehow indicate the index's shortcomings. Further research can modify the weighting constraints of the FWI methodology for small stock markets. I used a maximum 10% limit of each position, but this constraint still tilted the portfolio towards big-cap.

The results from my contribution exhibited how additional non-financial factors and efficiency ratios can significantly impact the model. I documented a significant 6.5 alpha on a 5% confidence level. Section 6.2 demonstrates how the market price is an inefficient method to get growth exposure in an index. However, my other contribution, which incorporated combined ESG-factors to the index, did not reveal significantly better results and only slightly higher returns than the S&P 500 and FWI Composite non-ESG. Although the increased annual volatility does not give better risk-adjusted returns[20] when compared to the cap-weighted ESG index. These results somehow indicate that the ESG as an additional parameter can improve a portfolio. This underperformance is presumably caused by (1) a weak portfolio construction method as described in (5.2) or (2) poor data quality (section 4.2). Although, when we adjusted the absolute returns for additional distress risks[21], the FWI ESG index came as the best index with the least negative alpha. So given these results, I have not sufficient robust evidence to draw any statistical conclusions.

[20] Refering to section A of Table 6.7
[21] Fama and French factors

References

Anadu, K., Kruttli, M. S., McCabe, P. E., Osambela, E., and Shin, C. (2019). The shift from active to passive investing: Potential risks to financial stability? *Available at SSRN 3244467*.

Ardishvili, A., Cardozo, S., Harmon, S., and Vadakath, S. (1998). Towards a theory of new venture growth. In *Babson entrepreneurship research conference, Ghent, Belgium*, pages 21–23.

Arnott, R., Amie Ko, C., and Treussard, J. (2019). Standing alone against the crowd: Abandon value? now?!?

Arnott, R., Hsu, J., Liu, J., and Markowitz, H. (2007). Does noise create the size and value effects? *Manuscript, University of California, San Diego*.

Arnott, R. D., Hsu, J., and Moore, P. (2005). Fundamental indexation. *Financial Analysts Journal*, 61(2):83–99.

Arnott, R. D., Hsu, J. C., and West, J. M. (2011). *The fundamental index: A better way to invest*. John Wiley & Sons.

Ashwin Kumar, N., Smith, C., Badis, L., Wang, N., Ambrosy, P., and Tavares, R. (2016). Esg factors and risk-adjusted performance: a new quantitative model. *Journal of Sustainable Finance & Investment*, 6(4):292–300.

Balatti, M., Brooks, C., and Kappou, K. (2017). Fundamental indexation revisited: New evidence on alpha. *International Review of Financial Analysis*, 51:1–15.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, 9(1):3–18.

Barr Rosenberg, K. R. and Lanstein, R. (1984). Persuasive evidence of market inefficiency. *Journal of portfolio management*, 11:9–17.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The journal of Finance*, 32(3):663–682.

Basu, S. et al. (1981). The relationship between earnings' yield, market value and return for nyse common stocks: Further evidence.

Blitz, D. and Swinkels, L. (2008). Fundamental indexation: An active value strategy in disguise. *Journal of Asset Management*, 9(4):264–269.

Brammer, S., Brooks, C., and Pavelin, S. (2006). Corporate social performance and stock returns: Uk evidence from disaggregate measures. *Financial management*, 35(3):97–116.

Busse, J. A., Goyal, A., and Wahal, S. (2014). Investing in a global world. *Review of Finance*, 18(2):561–590.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.

Chen, C., Chen, R., and Bassett, G. W. (2007). Fundamental indexation via smoothed cap weights. *Journal of Banking & Finance*, 31(11):3486–3502.

Chow, T.-m., Hsu, J., Kalesnik, V., and Little, B. (2011). A survey of alternative equity index strategies. *Financial Analysts Journal*, 67(5):37–57.

Ciftci, M., Darrough, M., and Mashruwala, R. (2014). Value relevance of accounting information for intangible-intensive industries and the impact of scale: The us evidence. *European Accounting Review*, 23(2):199–226.

Clausen, S. and Hirth, S. (2016). Measuring the value of intangibles. *Journal of Corporate Finance*, 40:110–127.

Corrado, C., Hulten, C., and Sichel, D. (2009). Intangible capital and us economic growth. *Review of income and wealth*, 55(3):661–685.

De Bondt, W. F. and Thaler, R. H. (1989). Anomalies: A mean-reverting walk down wall street. *Journal of Economic Perspectives*, 3(1):189–202.

Dimson, E. et al. (1988). *Stock market anomalies*. CUP Archive.

Estrada, J. (2008). Fundamental indexation and international diversification. *The Journal of Portfolio Management*, 34(3):93–109.

Fama, E. F. (1965). The behavior of stock-market prices. *The journal of Business*, 38(1):34–105.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of*.

Fama, E. F. and French, K. R. (2004). The capital asset pricing model: Theory and evidence. *Journal of economic perspectives*, 18(3):25–46.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

Fichtner, J., Heemskerk, E. M., and Garcia-Bernardo, J. (2017). Hidden power of the big three? passive index funds, re-concentration of corporate ownership, and new financial risk. *Business and Politics*, 19(2):298–326.

Filipozzi, F. and Tomingas, R. (2017). Performance evaluation of fundamental indexation strategies on the nasdaq omx baltic stock exchange. *Research in Economics and Business: Central and Eastern Europe*, 9(2).

Gibbons, M. R., Ross, S. A., and Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, pages 1121–1152.

Giese, G., Ossen, A., and Bacon, S. (2016). Esg as a performance factor for smart beta indexes. *The Journal of Index Investing*, 7(3):7–20.

Goodwin, T. H. (1998). The information ratio. *Financial Analysts Journal*, 54(4):34–43.

Haugen, R. A. and Baker, N. L. (1991). The efficient market inefficiency of capitalization–weighted stock portfolios. *The Journal of Portfolio Management*, 17(3):35–40.

Hillestad, O.-C. (2007). Nøkkeltallsanalyse av oslo børs.

Hsu, J. C. (2004). Cap-weighted portfolios are sub-optimal portfolios. *Journal of investment Management*, 4(3).

Jun, D. and Malkiel, B. G. (2008). New paradigms in stock market indexing. *European Financial Management*, 14(1):118–126.

Kahle, K. M. and Stulz, R. M. (2017). Is the us public corporation in trouble? *Journal of Economic Perspectives*, 31(3):67–88.

Kalesnik, V. (2014). The second generation of index investing. *Smart Beta*, pages 25–29.

Kurtz, L. et al. (2011). The long-term performance of a social investment universe. *The Journal of Investing*, 20(3):95–102.

Lim, K.-P. and Brooks, R. (2011). The evolution of stock market efficiency over time: a survey of the empirical literature. *Journal of Economic Surveys*, 25(1):69–108.

Liu, H. and Wang, Y. (2018). Index investing and price discovery. *Available at SSRN 3166685*.

Lo, A. W. (2004). The adaptive markets hypothesis. *The Journal of Portfolio Management*, 30(5):15–29.

Lo, A. W. (2016). What is an index? *The Journal of Portfolio Management*, 42(2):21–36.

Malkiel, B. G. and Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417.

Markowitz, H. (1959). *Portfolio selection: Efficient diversification of investments*, volume 16. John Wiley New York.

Marquering, W., Nisser, J., and Valla, T. (2006). Disappearing anomalies: a dynamic analysis of the persistence of anomalies. *Applied Financial Economics*, 16(4):291–302.

Nissim, D. and Ziv, A. (2001). Dividend changes and future profitability. *The Journal of Finance*, 56(6):2111–2133.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28.

Perold, A. F. (2007). Fundamentally flawed indexing. *Financial Analysts Journal*, 63(6):31–37.

Phillips, C. and Ambrosio, F. (2008). The case for indexing. *Investment Counseling and Research, The Vanguard Group*.

Poterba, J. M. and Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of financial economics*, 22(1):27–59.

Pysarenko, S., Alexeev, V., and Tapon, F. (2019). Predictive blends: Fundamental indexing meets markowitz. *Journal of Banking & Finance*, 100:28–42.

Reuters, T. (2019). Thomson reuters esg scores. *URL: https://www.refinitiv.com/en/financial-data/company-data/esg-research-data*.

Rowley Jr, J. J., Walker, D. J., and Ning, S. Y. (2018). The case for low-cost index-fund investing. *Vanguard white paper (April). https://institutional. vanguard. com/iam/pdf/ISGIDX. pdf.*

Sharpe, W. F. (1964). A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, 39(1):119–138.

Siegel, J. (2006). The 'noisy market'hypothesis. *Wall Street Journal*, 14:A14.

Skjeltorp, J., Næs, R., and Ødegaard, B. (2008). Hvilke faktorer driver kursutviklingen på oslo børs. *Norsk Økonomisk Tidsskrift*, 2.

Sushko, V. and Turner, G. (2018). The implications of passive investing for securities markets. *BIS Quarterly Review, March.*

Titman, S., Wei, K. J., and Xie, F. (2004). Capital investments and stock returns. *Journal of financial and Quantitative Analysis*, 39(4):677–700.

Treynor, J. (2005). Why market-valuation-indifferent indexing works. *Financial Analysts Journal*, 61(5):65–69.

Tsiang, S.-C. (1972). The rationale of the mean-standard deviation analysis, skewness preference, and the demand for money. *The American Economic Review*, 62(3):354–371.

Ulbricht, N. and Weiner, C. (2005). Worldscope meets compustat: A comparison of financial databases. *Available at SSRN 871169.*

Vayanos, D. (2004). Flight to quality, flight to liquidity, and the pricing of risk. Technical report, National bureau of economic research.

Walkshäusl, C. and Lobe, S. (2010). Fundamental indexing around the world. *Review of Financial Economics*, 19(3):117–127.

Yen, G. and Lee, C.-f. (2008). Efficient market hypothesis (emh): past, present and future. *Review of Pacific Basin Financial Markets and Policies*, 11(02):305–329.

Zhou, G. (1991). Small sample tests of portfolio efficiency. *Journal of Financial Economics*, 30(1):165–191.

Appendix

A1 Appendix A

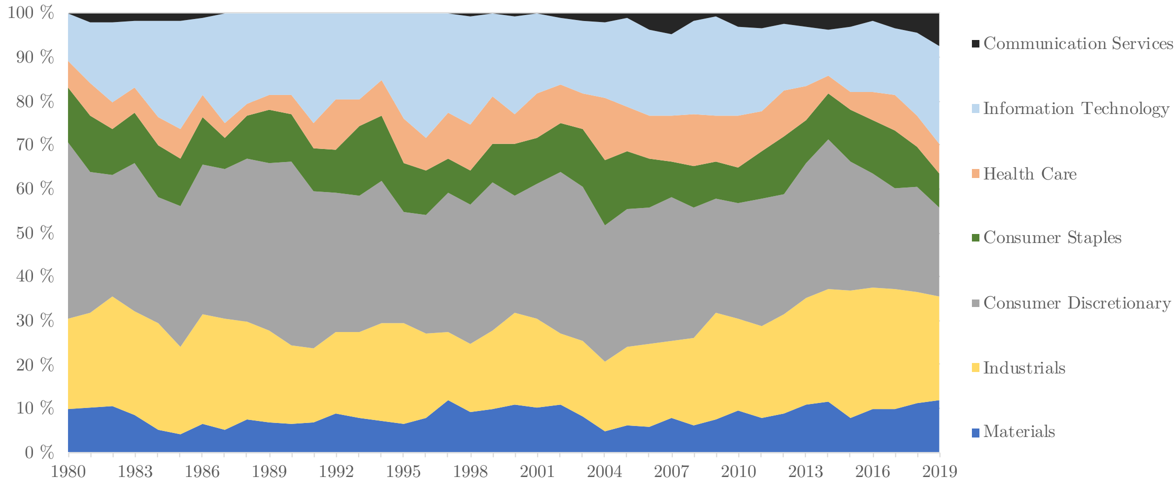


Figure A1.1: Growth Index - Sector Exposure

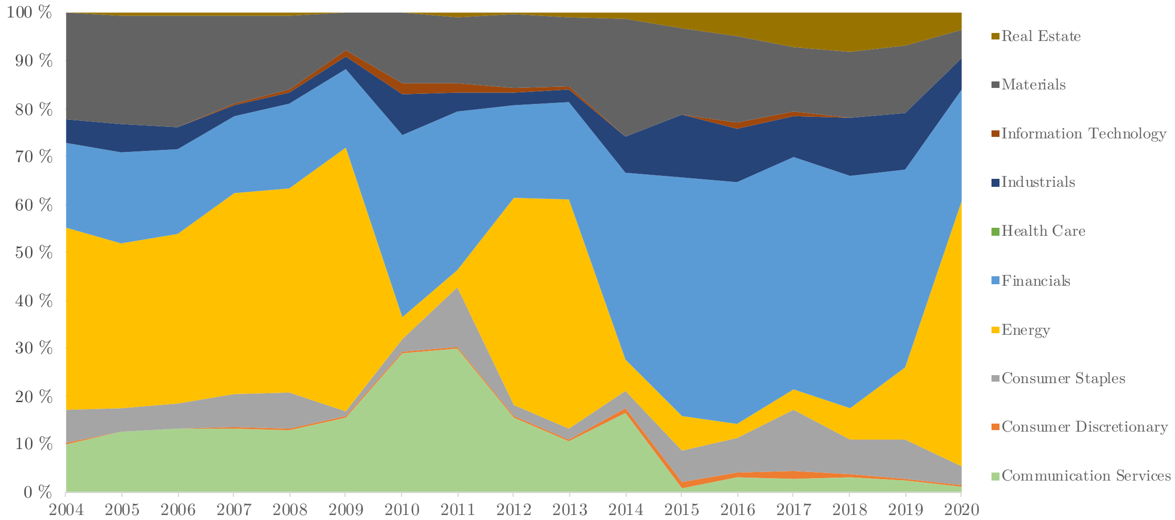


Figure A1.2: The Sector Exposure of the Norwegian Composite Index

Growth Index - Rolling FF5F analysis

Year	R^2	$R^2{}_{adj}$	Mkt-rf	SMB	HML	RMW	CMA	α
1990	0,96	0,91	0,60	1,76	-0,73	1,19	2,00	0,09
1991	0,96	0,90	0,59	2,58	-1,48	2,45	2,63	0,06
1992	0,95	0,88	0,62	3,03	-2,36	4,61	4,44	-0,14
1993	0,93	0,85	0,71	1,76	-0,91	0,92	0,96	0,16
1994	0,92	0,82	0,80	1,72	-0,93	0,78	0,95	0,15
1995	0,95	0,89	1,08	0,68	0,04	0,37	-0,87	0,19
1996	0,96	0,91	0,95	0,92	-0,19	-0,43	-1,05	0,25
1997	0,95	0,88	1,01	0,71	0,08	-0,40	-1,39	0,24
1998	0,96	0,92	1,00	0,76	0,03	-0,25	-1,31	0,22
1999	0,93	0,85	1,11	0,66	-0,11	0,46	-0,56	0,15
2000	0,89	0,76	1,36	0,85	-0,74	0,88	0,94	0,11
2001	0,77	0,48	1,06	0,63	-0,52	0,55	0,86	0,14
2002	0,81	0,58	1,09	0,64	-0,14	0,34	0,43	0,14
2003	0,97	0,93	0,50	0,42	1,27	-1,03	-0,74	0,28
2004	0,97	0,93	0,43	0,39	1,35	-1,12	-0,84	0,30
2005	0,89	0,75	0,95	0,71	0,38	-0,12	0,00	0,18
2006	0,92	0,82	1,31	0,90	-0,07	0,56	0,16	0,15
2007	0,95	0,89	1,82	1,00	-0,10	1,04	0,27	0,10
2008	0,98	0,97	1,63	1,23	0,03	0,84	0,11	0,09
2009	1,00	0,99	1,79	1,24	0,05	1,17	0,09	0,06
2010	0,96	0,91	1,60	1,32	0,35	0,89	-0,42	0,06
2011	0,96	0,91	1,61	1,41	0,37	0,96	-0,45	0,06
2012	0,96	0,91	1,52	1,59	0,27	0,66	-0,88	0,06
2013	0,93	0,84	1,67	1,00	0,40	1,87	-0,63	-0,02
2014	0,92	0,81	1,71	1,36	0,54	2,16	-0,85	-0,04
2015	0,92	0,83	1,69	1,32	0,48	2,11	-0,78	-0,04
2016	0,93	0,85	1,57	1,22	-0,32	1,60	-0,32	-0,05
2017	0,90	0,78	1,41	1,26	-0,36	1,14	0,07	-0,05
2018	0,84	0,64	1,16	1,39	-0,43	0,86	0,01	0,00
2019	0,97	0,94	1,01	0,65	-0,19	0,65	0,46	0,00

Table A1.1: Growth Index - 10 years rolling window

	S&P 500	S&P 500 IVW	Cash Flow	Net Income	Book Value	Revenue	Composite	Dividend	NASDAQ 100	FWI Growth
A. *Fama & French three-factor*										
alpha (pps)	1,313**	0,002	2,345**	2,681**	2,308**	1,825	2,567**	0,002	4,538**	8,563***
	(0,005)	(0,007)	(0,009)	(0,010)	(0,010)	(0,010)	(0,009)	(0,009)	(0,019)	(0,022)
Mkt-rf	0,984	0,916	0,851	0,874	0,911	0,933	0,935	0,924	1,192	1,067
	(0,028)	(0,041)	(0,053)	(0,056)	(0,057)	(0,057)	(0,078)	(0,049)	(0,100)	(0,117)
SMB	-0,204***	-0,076	0,080	-0,007	0,141	0,178**	0,057	-0,010	0,374**	0,972***
	(0,042)	(0,061)	(0,079)	(0,084)	(0,086)	(0,085)	(0,078)	(0,073)	(0,159)	(0,187)
HML	0,076**	0,332***	0,066	0,198**	0,124*	0,228**	0,103*	0,342***	-0,493***	0,161
	(0,033)	(0,047)	(0,062)	(0,065)	(0,066)	(0,066)	(0,060)	(0,057)	(0,117)	(0,137)
R^2 adjusted	0,945	0,918	0,828	0,843	0,824	0,843	0,849	0,867	0,838	0,756
B. *Carhart four-factor*										
alpha (pps)	1,031**	-0,002	2,102**	2,481**	2,214*	1,664	2,435**	-0,001	3,900*	7,421***
	(0.005)	(0,008)	(0,011)	(0,011)	(0,012)	(0,014)	(0,012)	(0,010)	(0,0211)	(0,024)
Mkt-rf	1,013	0,929	0,861	0,885	0,912	0,919	0,923	0,912	1,107	1,157
	(0,028)	(0,041)	(0,055)	(0,057)	(0,059)	(0,056)	(0,058)	(0,052)	(0,103)	(0,121)
SMB	-0,190***	-0,062	0,081	0,004	0,143	0,162*	0,060	-0,004	0,400**	0,944***
	(0,041)	(0,061)	(0,082)	(0,085)	(0,088)	(0,086)	(0,081)	(0,075)	(0,165)	(0,193)
HML	0,090	0,346***	0,067	0,211	0,126*	0,212**	0,106	0,348***	-0,478***	0,145
	(0,032)	(0,048)	(0,064)	(0,067)	(0,069)	(0,068)	(0,067)	(0,059)	(0,120)	(0,140)
MOM	0,046	0,048	0,002	0,042	0,005	-0,053	0,010	0,021	0,061	-0,064
	(0,023)	(0,034)	(0,045)	(0,047)	(0,049)	(0,048)	(0,044)	(0,041)	(0,086)	(0,101)
R^2 adjusted	0,968	0,925	0,834	0,851	0,848	0,852	0,862	0,876	0,843	0,789

Table A1.2: Three factor and Carhart four factor analysis

NTNU

Kunnskap for en bedre verden