Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Economics

Bendik Vandvik Aakhus and Ole Martin Grøndahl

Do Norwegian Stocks Outperform the Risk-Free Rate?

An empirical study of the Norwegian stock market in the period 1980-2024

Master's thesis in Master of Science in Economics
and Master of Science in Financial Economics
Supervisor: Snorre Lindset
June 2024

NTNU
Norwegian University of
Science and Technology

Bendik Vandvik Aakhus and Ole Martin Grøndahl

Do Norwegian Stocks Outperform the Risk-Free Rate?

An empirical study of the Norwegian stock market in the period 1980-2024

Master's thesis in Master of Science in Economics
and Master of Science in Financial Economics
Supervisor: Snorre Lindset
June 2024

Norwegian University of Science and Technology
Faculty of Economics and Management
Department of Economics

NTNU
Norwegian University of
Science and Technology

Preface

This study marks the final chapter of our master's degree in financial and social economics at the Norwegian University of Science and Technology (NTNU). The thesis is motivated by Professor Hendrik Bessembinder's study "Do stocks outperform Treasury bills?" from 2018. We found this study interesting and quite eye-opening when it came to the performance of American stocks and high positive skewness of returns distribution. We therefore decided to study whether these results were also valid for the Norwegian stock market.

We would like to thank our supervisor Professor Snorre Lindset for the inspiring assignment, as well as significantly improving the quality of our work through his guidance. We also wish to thank Professor Yabin Wang for her assistance in STATA tools and Professor Espen Sirnes providing guidance and data through the chosen database, TITLON. Finally, we would like to thank friends and family for all their support throughout this writing period.

Abstract

In this thesis, we examine the performance of individual stocks from the Oslo stock exchange (OSEBX) in comparison to the risk-free interest rate (NIBOR). The timeframe is from January 1980 to March 2024, covering 976 unique stocks.

We find that the majority of stocks listed on the Oslo stock exchange generate lifetime buy-and-hold returns lower than the lifetime return of risk-free investments. Furthermore, less than half of stocks listed generate lifetime buy-and-hold returns greater than zero. The smaller companies generate the highest mean return and have the highest positive skew of distribution, supporting the concept of a small-firm effect. We also find that the net wealth creation is mainly consolidated to a few companies. The top 2% wealth creating companies deliver the net gain equal to the entire Oslo stock market since 1980, meaning the bottom 98% collectively generate wealth equal to the risk-free interest rates.

As we increase the number of stocks and time horizon, in a value-weighted portfolio, we find that their performance improves compared to the risk-free interest rate. Portfolios consisting of forty randomly picked stocks (each month), with a lifetime (44 years) buy-and-hold strategy, outperform the risk-free interest rate three out of four times. Comparatively, single stock portfolios with the same time horizon, outperform the risk-free rate only one out of five times.

The high performance of the stock market as a whole, compared to the relatively poor performance of individual stocks, underline the importance of skewness and compounding returns.

Sammendrag

I denne avhandlingen undersøker vi utviklingen til enkeltaksjer fra Oslo Børs (OSEBX), sammenlignet med den risikofrie renten (NIBOR). Tidsrammen er fra januar 1980 til mars 2024, og dekker 976 unike aksjer.

Vi finner at majoriteten av aksjene som er notert på Oslo Børs, genererer en levetidsavkastning som er lavere enn avkastningen på risikofrie investeringer. Videre er det mindre enn halvparten av de børsnoterte aksjene som genererer en levetidsavkastning større enn null. De mindre selskapene genererer den høyeste gjennomsnittlige avkastningen og har den største positive skjevheten i fordelingen, noe som støtter konseptet om en "small-firm effect". Vi finner også at netto verdiskaping i hovedsak er konsolidert til et fåtall selskaper. Topp 2% av selskapene som skaper mest formue, leverer en nettogevinst som tilsvarer hele gevinsten til Oslo Børs siden 1980, noe som betyr at de resterende 98% av selskapene, samlet sett genererer en formue tilsvarende den risikofrie renten.

Når vi øker antall aksjer og tidshorisonten i en verdivektet portefølje, finner vi at antall avkastninger som slår den risikofrie renten øker. Porteføljer bestående av førti tilfeldig valgte aksjer (hver måned), med en livslang (44 år) kjøp-og-hold-strategi, gjør det bedre enn den risikofrie renten tre av fire ganger. Til sammenligning oppnår porteføljer, bestående av enkeltaksjer med samme tidshorisont, høyere avkastning enn den risikofrie renten kun én av fem ganger.

Den sterke utviklingen i aksjemarkedet som helhet, sammenlignet med den relativt svake utviklingen i enkeltaksjer, understreker betydningen av skjevhet og sammensatt avkastning.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background and Research Question

Over the last decade, the proportion of private households investing in stocks and funds has had an upwards trend (Brynestad et al. 2021). While an increasing number of private individuals participate in the stock market, those investing into individual common stocks often find themselves underperforming relative to the market (Wold 2022). The financial newspaper Dagens Næringsliv also reports that in 2023, only four out of 62 actively managed mutual funds managed to perform above the Oslo stock market index (Winther and Tallaksen 2024). Given these insights, it is important to uncover which parameters actually affect the Norwegian stock market. Moreover, could diversification possibly affect the opportunities for excess returns compared to the risk-free interest rate and indices?

Comparisons to the risk-free interest rate may seem unnecessary, since data from 1980 to today shows that the Norwegian stock market as a whole has outperformed the risk-free interest rate. Furthermore, it has been documented that well-diversified portfolios also outperform the risk-free rate on an annual basis. (Ødegaard 2021) However, by examining investments in individual stocks, the results may be completely different. Bessembinder (2018) conducted a study in the US stock market, and found that the majority of individual stocks actually underperform compared to the risk-free interest rate (US 1-month Treasury bill). In his conclusion, Bessembinder states that the reason why the market can beat the risk-free rate, while the majority of common stocks underperform, is due to a large positive skewness in the distribution of stock returns. In other words, a relatively small group of stocks contribute to the majority of the excess return in the US stock market. We have chosen to base our thesis mainly on Bessembinder's study, examining whether these results are applicable for the Norwegian stock market as well. We have therefore chosen the following research question for this master's thesis:

"Do Norwegian stocks outperform the risk-free rate?"

2 Theory and Literature Review

In this chapter we introduce the theoretical framework for our thesis, with formulas for stock returns, volatility, skewness and wealth creation. It also presents theory regarding time interval, diversification, small-firm effect, changes in the stock market and bootstrapping.

2.1 Stock Returns

Stock returns can be described as the gain or loss achieved by investing in a company over a period of time. This return is often expressed as a percentage change, but can also be shown as a change in monetary value or in logarithmic form. According to Siddikee (2018), two methods are usually used to calculate equity returns; arithmetic and logarithmic. As we do not need the additive effect of logarithmic returns, we will mainly use arithmetic returns in this thesis. Disregarding any transaction costs, the formula for arithmetic return is defined by

$$s_t^i = \frac{P_t^i - P_{t-1}^i}{P_t^i} = \frac{P_t^i}{P_{t-1}^i} - 1, \tag{2.1}$$

where s_t^i is the return of stock i in the time interval $(t-1, t)$, P_t^i is the price of stock i at time t, and P_{t-1}^i is the price of stock i at the previous time $t-1$.

In this study, we use the last price of the month, adjusted for dividends, splits, and other corporate actions, in order to calculate the stock returns. This adjusted price is calculated by multiplying the multiplication factors for dividends and corporate actions with the closing price, expressed as

$$A_{t-1}^i = P_{t-1}^i \cdot \prod_{k=1}^{K} C_k^i \cdot \prod_{u=1}^{U} L_u^i, \tag{2.2}$$

where A_{t-1}^i is the adjusted price for stock i at previous time $t-1$, C_k^i is the adjustment factor for corporate actions for stock i in the period between adjustments k, L_u^i is the adjustment factor for dividends actions for stock i in the period between adjustments.[1] Dividends are adjusted for as this is a payment to the investor and is therefore included in the actual return an investor receives. Corporate actions such as splits are adjusted for as they reduce the stock price by the split ratio and can therefore give a false impression of the price development. For example, consider a case where the price of one share was NOK 20 one month, but they increase the number of shares the next month by a ratio of 2:1. The price would then be NOK 10 the next month, assuming no growth. Using formula (2.1), the stock split would result in a return of -50%, when in fact there has been no return at all. These adjustments are done for all previous stages, which means that adjustments for dividends and corporate actions have a retroactive effect. The adjusted price is therefore very different from the "actual" price at the time of purchase and functions as a theoretical price better suited for calculating historical returns. We see this suitability clearly in the case of Seabird, which had an extraordinary high adjusted price of NOK 57.1 million at Initial Price Offering (IPO), while the closing price at the same time was NOK 28.3. If we use the price and not the adjusted price, Seabird's lifetime return would have been -85.2%, while in reality - adjusted for dividends, splits and other corporate actions - it has actually had a lifetime return of -99.9%. As we are primarily basing our results on arithmetic returns, formula (2.1) is therefore extended to include the adjustment factors for dividends and corporate actions in formula (2.2), and is expressed as

$$r_t^i = \frac{A_t^i}{A_{t-1}^i} - 1, \tag{2.3}$$

where r_t^i is the arithmetic adjusted return for stock i in the time interval $(t-1, t)$.

2.2 Volatility

In order to calculate the skewness in the distribution of returns we first find the standard deviation of the returns, also known as volatility. Volatility is defined by Daly (2008) as

[1] See "End of Day US Stock Prices" at https://data.nasdaq.com/databases/EOD#anchor-adjustment-overview and "Corporate Action Methodology" at https://www.lseg.com/content/dam/ftse-russell/en_us/documents/methodology/corporate-actions-methodology.pdf for the calculation of dividend and corporate adjusted multiplication factors.

"the changeableness of the variable under consideration; the more the variable fluctuates over a period of time, the more volatile the variable is said to be". The volatility is important to consider when analyzing returns in relation to risk. Daly further presents the common measure of stock return volatility - abbreviated as the standard deviation of returns -, shown as

$$\sigma = \sqrt{\sum_{1}^{T}(R_t - \bar{R})^2/(T-1)}, \tag{2.4}$$

where σ is the standard deviation of the sample returns R_t, $\bar{R}$ is the sample mean return, and T is the number of observations.

2.3 Skewness

In statistical and empirical analysis, the assumption of normal distribution is often necessary in order to perform traditional parametric tests. Through large numbers of observations, it is assumed that the median equals both the mode and the average of given observations. When this is not the case, the distribution of the observations is skewed, i.e. non-normal. Through their research on the behaviour of stock returns, Yan and Han (2019) conclude that "...normal distribution is not suitable for modelling distribution of stock returns". When observations differ from normal distribution, we can estimate the level of difference through a *skewness* test.

Skewness is described by D'Agostino and Belanger (1990) as *nonnormality*, and "non-normality of a population can be described by values of its central moments differing from the normal values". Skewness can graphically be presented as in Figure 2.3.1:

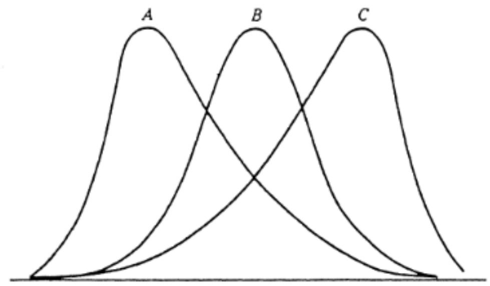

Figure 2.3.1: Sketch showing general position of mean, median and mode in a population, from D'Agostino and Belanger (1990), pp. 317. Illustration of distribution with skewness $\sqrt{\beta_1}$: A, $\sqrt{\beta_1} > 0; B, \sqrt{\beta_1} = 0; C, \sqrt{\beta_1} < 0$.

Furthermore, D'Agostino and Belanger (1990) present the formula for skewness coefficient, expressed as

$$Skew = \sqrt{\beta_1} = \frac{E(X-\mu)^3}{\sigma^3}, \tag{2.5}$$

where σ is the standard deviation - solved in formula (2.3) - μ is the mean, E is the expected value operator. For normal distribution: $Skew = \sqrt{\beta_1} = 0$. When the total observations hold a relatively small group of observations with high values, we find a skewness similar to curve A in Figure (2.3.1) with a distribution tail towards the high value observations. In these cases, the mean will be higher than the median. When choosing single stock portfolios with high positive skewness in the stock return observations, the mean returns may not be a good representation of expected returns.

2.4 Time Interval

A common phrase in financial investment is *"Time in the market beats timing the market - almost always."* (Fisher 2018). By simply increasing the time horizon of any given portfolio, the expected volatility of stock returns in said portfolio are assumed to be reduced and thereby a risk reduction is established. With a reduced level of risk, the expected returns are assumed to increase. This concept of risk reduction is called *time diversification*. However, according to Olsen and Khaki (1998): "Supporters of the concept argue that risk decreases as investment horizon increases; detractors suggest that the concept is false on the face of it." They state that the debate "stems from the profession's failure to accept a common definition of risk", but argue that "time diversification is consistent both theoretically and empirically with current conceptions of risk and rationality." Comparing investment horizons on U.S. stocks and U.S. T-bills, Hansson and Persson (2000) argue that their results show gains from time diversification: "The weights for stocks in an efficient portfolio were significantly larger for long investment horizons than a one-year horizon".

In the present thesis, we will therefore analyze the possible effects from time diversification through the same time intervals performed by Bessembinder (2018), which is *yearly*, *decade* and *lifetime*.

2.5 Diversification

Lee, Lee, and Lee (2010) state that companies are risk aligned to their own sector, but less so to other sectors. By increasing the number of stocks from different sectors in a given portfolio, as long as cashflows are not perfectly correlated, a "smoothing or diversification will take place". Through the belief that companies generate expected stock returns based on systematic (market) and unsystematic (firm-specific) factors, the total security risk can be presented as

$$Total\ security\ risk = Systematic\ risk + Unsystematic\ risk. \tag{2.6}$$

By that reasoning, as the number of stocks in a given portfolio increases, "the unsystematic components tend to cancel each other as they are all residuals from the relationship of security returns with the overall market return" (Lee, Lee, and Lee 2010). Graphically, we can then view the assumed process as shown in Figure 2.5.1:

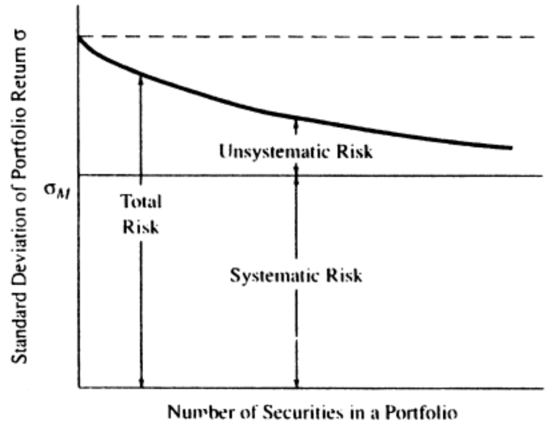

Figure 2.5.1: Connection between number of securities/stocks in an equal-weighted portfolio and the risk/standard deviation of portfolio return, Figure provided by Lee, Lee, and Lee (2010)

Statman (1987) presents, through the works of Elton and Gruber (1977), the reduction in expected standard deviation (SD) of annual portfolio returns, when increasing the number of stocks in an equal-weighted portfolio. In portfolios with only one stock, the expected *SD* is at 49.236, whilst portfolios with one hundred stocks have much lower levels of expected *SD*, at 19.686. Statman concludes that in order to consider the portfolio well diversified with insignificant unsystematic risk, the portfolio should include at least 30 stocks. This conclusion, Statman acknowledges, differs greatly from earlier results arguing that the "benefits of diversification for stock portfolios are exhausted when the number of stocks reaches 10 or 15" (Statman 1987). Even though expected *SD* is reduced upon an increase in the number of stocks in a portfolio, it does not necessarily imply a reduction in skewness as well. However, according to the findings of Simkowitz and Beedles (2010), "raw portfolio skew decreases as the number of assets in the portfolio increases", which shows a tendency of reduced skewness following a stock diversification.

In our thesis, we therefore include portfolios of Norwegian stocks in order to analyze more than single stock performance. We will estimate portfolios comprising of *one*, *five*, *ten*, *twenty* and *forty* stocks.

2.6 Wealth Creation

Wealth creation is defined by Lepak, Smith, and Taylor (2007), expanding upon the work of Bowman and Ambrosini (2000), as "the monetary amount realized at a certain point in time, when then exchange of the new task, good, service or product takes place". Wealth creation will in our findings be assessed as the creation of wealth as the market capitalization of a given company changes from its listing date on the Norwegian stock exchange. In reality, if all stocks were to be sold simultaneously, a significant portion of the perceived value would disappear, assuming a lack of potential buyers. Such a scenario is considered unlikely and our thesis will therefore defer from any judgments as to whether or not the stated market capitalization of a company represents its true worth.

Finally, given the opportunity that any investment can be put into risk-free interest rate with a guaranteed given return, the real wealth creation is considered as the change in market capitalization beyond what would be achieved with an investment into risk-free rate, shown as

$$W_{t,T}^{j} = M_{T}^{j} - (M_{t}^{j} \cdot Rf_{t,T}), \tag{2.8}$$

where $W_{t,T}^{j}$ is the net wealth creation for company j in the period (t, T), M_{T}^{j} is the market capitalization for company j at final time T, M_{t}^{j} is market capitalization for company j at first time t, and $Rf_{t,T}$ is the risk-free interest rate obtained in the period (t, T). It should be noted that in the present thesis, formula (2.8) will not include any accumulated market wealth prior to our first time period, i.e. January 1980.

2.7 Small-Firm Effect

A commonly used estimation tool of a company's value is the capital asset pricing model (CAPM). Developing upon the model, Fama and French (2013) introduced the importance of firm size through the Five-Factor Asset Pricing Model (FFAPM). However, this model carries one significant flaw. Banz (1981) found through his analysis of the

American stock market, from 1936 to 1980, significantly higher risk-adjusted returns in smaller firms compared to larger firms. These findings challenge the standard FFAPM estimations, and has later been coined as the small-firm-effect (SFE). Banz concludes the possible reason for high risk-adjusted returns in small capitalization firms could be related to the lower level of detailed company information published. To a risk-averse investor, the lack of information about a company might increase the perceived risk.

Booth and Smith (1987) tested the relation between the small-firm effect and skewness preference in investors. Their findings indicate that the “small-firm effect cannot be fully attributed to tax effects, benchmark error, or incorrect assumptions of the CAPM about investor risk aversion”, somewhat supporting Banz’ theory.

2.8 Changes in the Stock Market

Fama and French (2004) observed the development of characteristics in new listings at the American stock market from 1980 to 2001. They observed an upwards trend in the number newly listed stocks each year towards 2001. Furthermore, they also observed that among the newly listed stocks in the period 1980-2001, the “profitability becomes progressively more left skewed, and growth becomes more right skewed.” (Fama and French 2004). According to Fama and French, the changes of priority from profitability to growth for companies at the beginning of their listing significantly reduces the survival rates of new listings. As growth increasingly becomes the focus, they also found that a larger portion of new listings were smaller companies.

Building upon Fama and French, Fink et al. (2010) studied the volatility in the average public firm during the Internet boom, and note that “the spike in firm-specific risk in the late 1990s can be explained by the interaction of 2 reinforcing factors: a dramatic increase in the number of new listings, and a simultaneous decline in the age of the firm at initial public offering (IPO)”.

2.9 Bootstrapping

Bessembinder (2018) utilized bootstrapping in his analysis of expected stock performance in the American stock market. Mooney and Duval (1993) defined bootstrapping as a "computationally intensive, nonparametric technique for making probability-based inferences". The inferences are made about a population characteristic, θ, based on an estimator, $\hat{\theta}$, using a sample drawn from that population. The difference between bootstrapping and traditional parametric approaches is that bootstrapping "... employs large numbers of repetitive computations to estimate the shape of a statistics' sampling distribution ... This allows the researcher to make inferences in cases where such analytic solutions are unavailable, and where such assumptions are untenable." (Mooney and Duval 1993).

In the present thesis, given Bessembinder (2018)'s findings, we expect to not be able to utilize the normal distribution assumption which is necessary to proceed with traditional parametric approaches. However, by using bootstrapping, we can make certain inferences given the sampling distribution estimated through repetitive computation, in this case 20,000 simulations. The practical process of bootstrapping will be further discussed in section 3.3.

3 Methodology

In this chapter we introduce the quantitative methodology applied to investigate if Norwegian individual common stocks outperform the risk-free interest rate. In addition to extending the formulas from Chapter 2 to annual, decade and lifetime horizons, we also present a stochastic process to create portfolios of various sizes.

3.1 Buy-and-Hold Strategy

The returns on which this thesis is based are derived from monthly adjusted stock prices. In this respect, monthly returns will be the buy-and-hold return if one examines the dataset with a monthly time horizon. For periods longer than one month, these returns must be linked together so that they are representative of the time period. The linking is done by extending formula (2.3) to a geometric formula that multiplies 1 + the returns, and is expressed as

$$R^i_{t,T} = (1 + r^i_t) \cdot (1 + r^i_{t+1}) \cdot ... \cdot (1 + r^i_T) - 1, \tag{3.1}$$

where $R^i_{t,T}$ is the buy-and-hold return for stock i in the time interval (t, T), and r^i_t is the monthly return for stock i at time t. This form of composition includes the compounding effect that would be achieved by passively investing in a share over a longer period. In essence, the buy-and-hold return simulates an investor who buys a share at the start of a period and - given that the stockbroker reinvests the dividends received by the investor, - remains completely passive. As the investor is forced to hold the stock for the agreed upon time period, any attempts to time the market within the period is therefore not possible.

In addition to calculating total buy-and-hold returns for the stocks, we have also - in accordance with Bessembinder (2018) - calculate the simple summed returns for its indicative qualities. The simple sum of returns works as an indicator of whether or not the mean buy-and-hold return is positive, and is expressed as

$$S^i_{t,T} = \sum_{t=1}^{T} r^i_t, \tag{3.2}$$

where $S^i_{t,T}$ is the simple summed return for stock i in the time interval (t, T).

Next, we calculate the geometric mean of monthly returns, which works as an indicator for the centralization of our buy-and-hold returns. Skewness is a major part of the estimations in the present thesis, and it is therefore crucial to have an indicator which more accurately displays the centralization when skewness is high. The mean geometric return works better for centralization compared to the mean buy-and-hold return since it is not affected by extreme values to the same extent. This sturdiness comes from the fact that the T-th root of the buy-and-hold return is used, which compacts the exponential effect that compounding returns possess. The formula for geometric mean is expressed as

$$G^i_{t,T} = \sqrt[T]{(1 + r^i_t) \cdot (1 + r^i_{t+1}) \cdot ... \cdot (1 + r^i_T)} - 1 = \sqrt[T]{1 + R^i_{t,T}} - 1, \tag{3.3}$$

where $G^i_{t,T}$ is the geometric mean monthly return for stock i in the time interval (t, T). As well as providing us with a measurement for centralization, the formula serves a second purpose. As we will later discuss in section 4.1, obtaining the geometric mean monthly return for stocks that delist for a shorter period of time - instead of the total return for the delisting period -, is crucial for the construction of market indices, as well as the bootstrap simulations we perform.

Finally, we calculate the annualized return of each stock. The annualized return represents what the annual average return of a stock is over a specific period, which in terms of our scope of interest will be applied to the lifetime buy-and-hold returns. This calculation is performed because the stocks in the dataset have varying lifetimes, and when comparing stock performances we need a return that makes the basis of comparison fair. Since our returns are listed in monthly intervals, the formula for annualized return can - for our purposes - be expressed as

$$Y^i_{t,T} = (1 + R^i_{t,T})^{\frac{12}{T}} - 1, \tag{3.4}$$

where $Y^i_{t,T}$ is the annualized return for stock i in the time interval (t, T).

3.2 Equal-Weighted and Value-Weighted Market Indices

An index measures the return of a certain number of stocks using standardized metrics and methodology. They often involve all the stocks of certain categories, such as the S&P 500, which is a value-weighted index measuring the performance of the 500 highest capitalized American companies. In our thesis, we measure the Oslo stock market through both equal-weighted (EW) and value-weighted (VW) indices, calculated through the formulas

$$EW_t = \sum_{i=1}^{I} (\frac{r^i_t}{N_{t-1}}), \tag{3.6}$$

and

$$VW_t = \sum_{i=1}^{I} (r^i_t \cdot \frac{M^i_{t-1}}{TM_{t-1}}), \tag{3.7}$$

where EW_t is the equal-weighted index at time t, r^i_t is the monthly return of stock i at time t, N_{t-1} is the total number of stocks listed at time $t-1$, VW_t is the value-weighted index at time t, M^i_t is the market capitalization of stock i at time $t-1$, and TM_{t-1} is the total market capitalization of all the stocks listed at time $t-1$. In essence, these formulas construct market indices where the weighting is rebalanced as soon as current months number of listed stocks and their respective market capitalization is known. In order to link these market indices over the same period as the stocks' buy-and-hold returns, we use formula (3.1).

3.3 Bootstrapping

Bessembinder (2018) describes the method by which his results have been obtained, which we present through a simplified example in Table 3.3.1. As stated in the theory section, a bootstrap simulation infers results from a given sample by repeating stochastic simulations/tests, with replacement. In more practical terms, each month we draw x number of random stock returns, weigh them based on their corresponding market capitalization, and store the value-weighted portfolio return. In our example, we have five simulations. We also have the risk-free interest rate, R_f, and the value-weighted index, VW, which both hold returns equal to the ones calculated outside of the simulations. Furthermore, in the example we have four time periods, where to each time period in the simulation there is a corresponding portfolio return.

Table 3.3.1: Example of bootstrap results where $Sim.1$ is the first simulation, R_f is the risk-free interest rate and VW is the value-weighted index.

	Sim.1	Sim.2	Sim.3	Sim.4	Sim.5	R_f	VW
Period 1	0.2	0.3	0.5	0.7	0.9	0.2	0.3
Period 2	-0.4	0.5	0.4	0.6	0.4	0.4	0.5
Period 3	0.6	-0.7	0.1	0.9	-0.4	0.2	0.9
Period 4	0.8	0.9	0.2	0	-0.5	0.4	0.8
Total return	1.2	1.0	1.2	2.2	0.4	1.2	2.5
Mean return	0.3	0.25	0.3	0.55	0.1	0.3	0.625

3.3.1 Returns

At the crosspoint of simulation 1 and period 1, we find in this example, a return of 0.2, i.e. 20%. This represents the portfolio return for period 1 in simulation 1. The portfolio is weighted to the stocks' market capitalization from the previous period. All stocks have their market capitalization divided by the total market capitalization of all the stocks in the portfolio. The return is calculated as presented in formula (2.3). Each simulation then sums up its respective returns, such as simulation 1 which has a return of 1.2, i.e. 120%. That total return is then divided by the number of periods, in our example four, which gives an average period return to simulation 1, as 0.3, i.e. 30%.

We then report the average return for all 20,000 simulations by calculating the mean of all the simulated mean returns, in our case: $(0.3 + 0.25 + 0.3 + 0.55 + 0.1)/5 = 0.3 \rightarrow$ 30%.

3.3.2 Median and skewness

Similarly to the final stage of calculating returns, we obtain median and skewness through the mean return row. In our example, this gives us a median of $0.1, 0.25, \mathbf{0.3}, 0.3, 0.55 \rightarrow$ 30%, and a skewness - through formula (2.5) - of: 0.4929. This means we would see a positively skewed distribution as its larger than zero, similar to curve A in Figure 2.3.1.

3.3.3 Positive returns, risk-free interest and value weighted index

In the analysis section, we will present the percentage of portfolios that have returns above zero, the risk-free rate, and the value-weighted index. Let us assume each period in Table 3.3.1 is yearly and the total number of years in the dataset are included. The total number of observations are therefore 20 (four periods times five simulations). The calculation of the percentage of returns that outperform the benchmarks is therefore performed as follows:

1. For each period (rows), the five observations, one for each simulation, are individually compared to zero, risk-free rate and value-weighted index. We compare for each period, as the stocks returns, risk-free rate, and the value-weighted index have their own periodically unique return. In our example, period 1 therefore gives us a % $>$ 0 equal to 100%, % $>$ rf equal to 80%, and % $>$ VW equal to 60% (equal to the benchmark does not count). We then proceed to compare the returns for the next periods.

2. We average the ratios found in step 1, across all the simulations. In our example this will result in a % $>$ 0 equal to $(1 + 0.8 + 0.6 + 0.8)/4 = 80\%$, % $>$ rf equal to $(0.8 + 0.4 + 0.4 + 0.4)/4 = 50\%$, and % $>$ VW equal to $(0.6 + 0.2 + 0 + 0.2)/4 = 25\%$, across all the simulations.

3.3.4 Yearly, decade, lifetime

As argued by Olsen and Khaki (1998), time diversification is an important aspect to reduce risk and acquiring higher average returns. We therefore collapse the monthly stock returns, risk-free rate, and value-weighted index into yearly, decade and lifetime buy-and-hold returns using formula (3.1). Looking at Table 3.3.1, assuming each period is one decade, we find the average portfolio decade return in the same procedure as in section 3.3.1. Likewise, the calculations comparing $\% > 0$, $\% > R_f$, and $\% >$ VW are calculated as in section 3.3.3.

3.3.5 1, 5, 10, 20, 40 stock portfolios

As argued by Lee, Lee, and Lee (2010), the number of stocks in a given portfolio reduce the level of unsystematic risk through diversification. In order too see if the same results apply to our dataset, we performa bootstrap of 20,000 simulations for each level of portfolio size, 1, 5, 10, 20, and 40 stocks. As described in section 3.3.1, we weigh the returns based on the corresponding market capitalization in the period prior to the return has been received. We then proceed to sum the weighted returns within the period, in order to obtain the value-weighted return for the portfolio, in that period. Once the value-weighted portfolio return is obtain, we follow the steps in sections 3.3.1 and 3.3.3, in order to observe the effects of diversification.

3.4 Composition of the Dataset

In the following tables, we illustrate the composition of data regarding a single stock; RGI Antilles. Table 3.4.1 shows the monthly matching of returns with the risk-free interest rate, value-weighted and equal-weighted indices. Using formula (3.1), we link the monthly returns and the corresponding benchmarks to annual, decade and lifetime horizons, as shown in Table 3.4.2. As the time horizon of RGI Antilles extends over two calendar years, the table shows two annual buy-and-hold returns; one for 1996 and one for 1997. Furthermore, since the lifespan of RGI Antilles does not extend over two whole decades, the decade buy-and-hold return is equal to the lifetime one. As the benchmarks are linked using the same formula, the annual, decade and lifetime benchmarks only have observations in the same period as the buy-and-hold return.

Table 3.4.1: The table shows the composition of the monthly data for RGI (Antilles), matched with the risk-free rate, value-weighted and equal-weighted index at the same period.

RGI (ANTILLES)

Date	Adjusted price	Monthly return	Market capitalization	Monthly interest	Monthly VW index	Monthly EW index
1996-07-31	67		4,520*			
1996-08-30	63	-0.0597	4,250*	0.0042	0.0328	0.0315
1996-09-30	66	0.0476	4,452*	0.0043	0.0199	0.0075
1996-10-31	61.5	-0.0682	4,149*	0.0042	0.0358	0.0604
1996-11-29	66	0.0732	4,452*	0.0037	0.0717	0.0856
1996-12-30	65	-0.0152	4,456*	0.0035	0.0446	0.0326
1997-01-27	75.5	0.1615	5,175*	0.0029	0.0967	0.1347

* represents number in millions

Table 3.4.2: The table shows the composition of monthly data for RGI (Antilles), linked to annual, decade and lifetime horizons.

RGI (ANTILLES)

Variable	1996-08-30	1996-09-30	1996-10-31	1996-11-29	1996-12-30	1997-01-27
Cumulative return	-0.0597	-0.0149	-0.0821	-0.0149	-0.0299	0.1269
Annual return					-0.0299	0.1615
Decade return						0.1269
Lifetime return						0.1269
Cumulative interest	0.0042	0.0085	0.0128	0.0165	0.0201	0.0231
Annual interest					0.0201	0.0029
Decade interest						0.0231
Lifetime interest						0.0231
Cumulative VW index	0.0328	0.0528	0.0905	0.1688	0.2209	0.3391
Annual VW index					0.2209	0.0967
Decade VW index						0.3391
Lifetime VW index						0.3391
Cumulative EW index	0.0315	0.0393	0.1020	0.1964	0.2355	0.4018
Annual EW index					0.2355	0.1347
Decade EW index						0.4018
Lifetime EW index						0.4018

4 Data

The data for this thesis is obtained from the TITLON database. The database is a charity project run primarily by Professor Espen Sirnes, through the University of Tromsø, and contains many different types of assets listed on the Norwegian stock market. When we first started the project, the database was limited to the time period 1980-2020, as OSEBX was acquired by Euronext in 2020, and access to data was limited. Fortunately, TITLON entered into a partnership with Euronext at the turn of 2023-2024, making data available right up to the time of writing. TITLON was chosen as our source of data as it is the database - we have access to - with the longest observational period, as well as a broad range of variables.

4.1 Filtering Process

With the implementation of four years of data, collected from the very beginning, it is natural that it contains some errors in certain variables. For this reason, we have filtered the dataset to the best of our knowledge in order to ensure that it is representative of the Norwegian stock market.

Initially, we started with a dataset consisting of daily observations of 1,053 stocks. In order to match the reporting to the risk-free interest rate as closely as possible, the dataset was converted to show the last recorded price of the month. Furthermore, we removed the stocks that either had no market capitalization, or market capitalization that was set to 0. This was done since market capitalization is a requirement for some of the analyses that will be performed later in the thesis. The lack of market value was due to the fact that the number of outstanding shares was not stated, and this applied to 56 stocks. Furthermore, we excluded eight stocks that only had one monthly observation. This was done because, according to formula (2.1), at least two observations of adjusted price is required to calculate a monthly return. Furthermore, we excluded seven assets that were classified as ETFs. ETFs are Exchange Traded Funds and are portfolios consisting of more than one single stock, which is not what we wish to analyze in this thesis.

When sorting the dataset by monthly return, we found some abnormally high returns, i.e. Gambit with an apparent monthly return of 66.64 or 6,664%. We then discovered that some shares were delisted from the stock exchange and relisted a period later with the same International Securities Identification Number (ISIN). To adjust for this, we used formula (3.3) for stocks that were delisted for less than or equal to six months in order to obtain the geometric mean monthly return for the period they were delisted. We then added observations in the delisting period containing the geometric mean monthly return, as well as the average development in market capitalization. Adding these observations ensured that the weighting of market indices were rebalanced correctly in accordance with our monthly intervals. In addition to correcting the weighting of market indices, this approach also ensured that the bootstrap simulations exclusively picked monthly returns. Picking a buy-and-hold return for the whole delisting period would simulate an investor who held the same stock for more than one month, when the purpose of bootstrapping is to pick a random stock each month. For stocks that were delisted for longer than six months before relisting, we set the monthly return equal to missing for the first observation after relisting.[2] This approach for dealing with periodic delisting corrected the majority of the abnormally high returns we found, but there were still some we considered very high. We set a limit of 600% monthly return, where all stocks with higher returns than this were thoroughly checked against other databases, such as Eikon Datastream, Yahoo Finance, Google Finance and Nordnet. When crossreferencing with these other databases, we found nine stocks that had significantly different returns from our dataset. After a conversation with Professor Espen Sirnes, we found that five of these returns were the result of a comma error in the adjusted price towards the end of the stocks' life, and were thus corrected. The remaining four; "Crew Gold Corporation", "Crew Gold Corporation New Shares", "Nydalens Compagnie B" and "Wentworth Resources" were excluded from the dataset, as their high returns could not be verified in other databases.

Finally, we found that most of the stocks imported from Euronext (2020-2024) had a

[2] We set the limit at six months as it may be realistic that missing observations could be due to an error on the part of the database. For stocks with a delisting period longer than six months, it is considered that the stock was completely delisted, the shareholders received the delisting price and lost their shares in the company. In other words, no return is achieved during the period the share is delisted for longer than six months.

missing company-id. Fortunately, TITLON provided an internal code for these companies that referred to the correct company at an earlier date, where a company-id was present. By using the code in the Appendix, we were able to link company-id together for all companies that had observations before 2020. The remaining companies that were listed for the first time in the period 2020-2024 and lacked company-id, we set the company-id equal to the internal code so that they were identified by a unique id.

Filtering the data as explained above, provided us with a final dataset consisting of 976 stocks, which is the data basis for this present thesis. Figure 4.1.1 shows the historical development of the number of stocks in our dataset from January 1980 to March 2024. The lowest amount was in August 1980 with 36 unique stocks. Interestingly, during the Covid-19 pandemic, the number of stocks listed was actually growing, and in fact skyrocketed at the beginning of 2021. The growth reached an all-time peak in April 2022 with 350 unique stocks listed in the same month. Throughout the lifespan of our dataset, we have an average of 213 stocks listed every month.

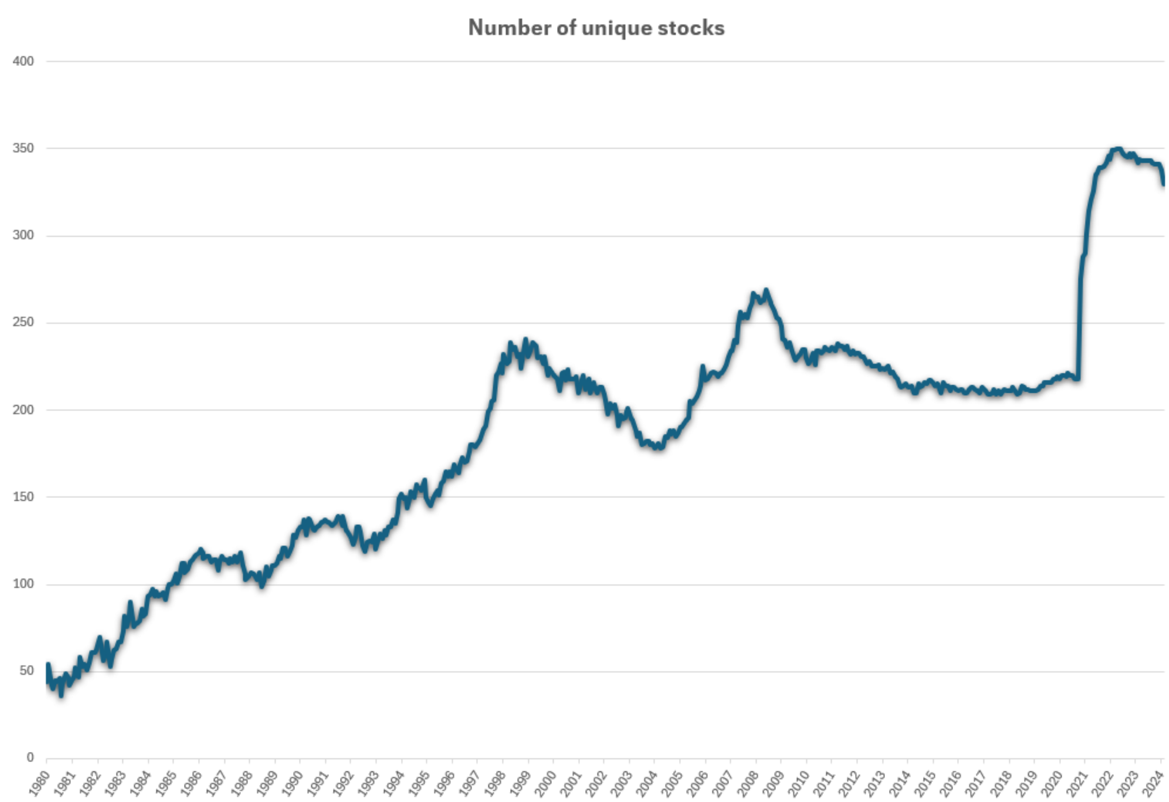

Figure 4.1.1: The graph shows the historical development of total number of unique stocks in our filtered dataset from January 1980 to March 2024.

4.2 Descriptive Statistics

Table 4.2.1 shows the monthly descriptive statistics on which our thesis is based. All numbers are so-called "nominal", as they are not adjusted for inflation or transaction costs. Adjusted price, which is our basis for monthly returns, include some extreme values, as mentioned in section 2.1, as the price is adjusted for dividends and other corporate actions, most noteworthy; stock splits. The largest value of NOK 57.9 million comes from Seabird Exploration in June 2007. This is because the company carried out reverse splits many times through its lifetime, which has kept the closing price somewhat more stable, while the market value declined, due to a lower number of outstanding shares. Seabird is also the biggest contributor to why the mean adjusted price is so high, as by removing the stock, the average adjusted price goes from NOK 18,341 to 6,110, while the median remains approximately the same.

Furthermore, as shown in Table 4.2.1, the number of observations of risk-free rate, value-weighted, and equal-weighted indices are reduced to the number of observations of the monthly return. In other words, we have removed the observations of the risk-free interest rate, value-weighted, and equal-weighted market indices if it is the first observation of a stock, as it has not yet had the opportunity to earn a monthly return. As mentioned in section 4.1, this also applies to stocks that are delisted, and then relisted with the same ISIN more than six months later. Matching the number of observations in this way prevents the comparison of a non-existent return with the benchmarks, as even though it only applies to 0.96% of our dataset, it may corrupt our results.

Table 4.2.1: The table shows the descriptive statistics for the stocks used in this thesis. All numbers are reported as monthly observations. rf is the risk-free rate, EW is the equal-weighted market index and VW is the value-weighted market index. Certain variables are rounded to four decimals for visual clarity. Price and adjusted price are in NOK.

Variable	N	Mean	SD	Min	Median	Max
Price	99,152	110.14	409.91	0.0005	45	25,000
Adjusted price	101,453	18,341.82	682,070.2	0.0005	26.44	57.90*
Market capitalization	101,453	5,820*	29,200*	23.43	747*	1,220,000*
Monthly return	100,477	0.0106	0.1785	-0.9937	0	8.2000
Monthly rf	100,477	0.0041	0.0033	0.0002	0.0032	0.0143
Monthly VW index	100,477	0.0082	0.0577	-0.2387	0.0135	0.2383
Monthly EW index	100,477	0.0103	0.0584	-0.2178	0.0109	0.1937

* represents number in millions

4.3 Risk-Free Interest Rate

Risk-free interest generally refers to the interest rate that can be achieved by tying up capital in an asset without exposure to risk factors. We know from the financial crisis in the US in 2008 that even the banks' interest rates are not always risk-free in practice. However, the general notion in economic theory is that interest rates obtained from certain government bonds, treasury bills, or bank products can - for theoretical purposes - be regarded as risk-free. In line with Ødegaard (2021), we choose to use NIBOR 3-month as our risk-free interest rate. This is because we have a dataset with monthly observations, and NIBOR is the closest we come to matching our observation horizon. The alternative - government bonds - are only issued with a maturity of at least one year. NIBOR is an acronym for Norwegian Interbank Offered Rate and is essentially an indication of the money market rate at which banks are willing to lend Norwegian kroner to other reliable banks without providing any form of collateral. NIBOR was first introduced at the beginning of 1986, and finding a risk-free rate before this time can be somewhat tricky. We choose to use the Interbank Overnight (O/N) rate from 1980 to 1986, obtained from Norges Bank (Norges Bank 2019). Data from Norges Bank on Interbank O/N and NIBOR 3-month extends to 2013, and the remaining data on NIBOR 3-month is obtained from Statistics Norway (SSB 2024).

Figure 4.3.1 shows the historical development of our risk-free interest rate from January 1980 to March 2024. The graph shows a declining trend until the Covid-19 pandemic, with some large spikes in the 80s, and one in 1992 that was due to a fixed exchange rate policy from 1986 that “cracked” in 1992 and led to a currency crisis (Gjedrem 1999). In the wake of the Covid-19 pandemic, we see that interest rates are rising at a drastic rate, as a result of the level of inflation brought about by the low “pandemic interest rate”.

Figure 4.3.1: The graph shows the historical development of the monthly risk-free rate from January 1980 to March 2024.

4.4 Market Indices

In addition to comparing individual stock returns with the risk-free interest rate, we also compare with market indices also known as market portfolios, consisting of all the stocks in the dataset. As mentioned in section 2.5, well-diversified portfolios, i.e. portfolios consisting of many stocks, will reduce the unsystematic risk. As these are indices that consists of the entire dataset, the total risk will be approximately equal to the systematic risk. Even though - unlike risk-free interest rates - they carry some risk, the lack of unsystematic risk makes them natural assets to compare with.

When calculating the value-weighted market index, the previous months market value of a stock is used, and compared with the total market value of all the stocks in the same period, in order to obtain the weighting. The weighting determines how much of the portfolio should be invested in a specific stock. For example, throughout its lifetime, Equinor has had a market value that corresponds to approximately 20% of the entire market at any given time. This means that approximately 20% of the return of the value-weighted market index consists of Equinor's return each month. For the equal-weighted market index, all stocks - regardless of size - receive an equal weighting based on previous months total number of stocks.

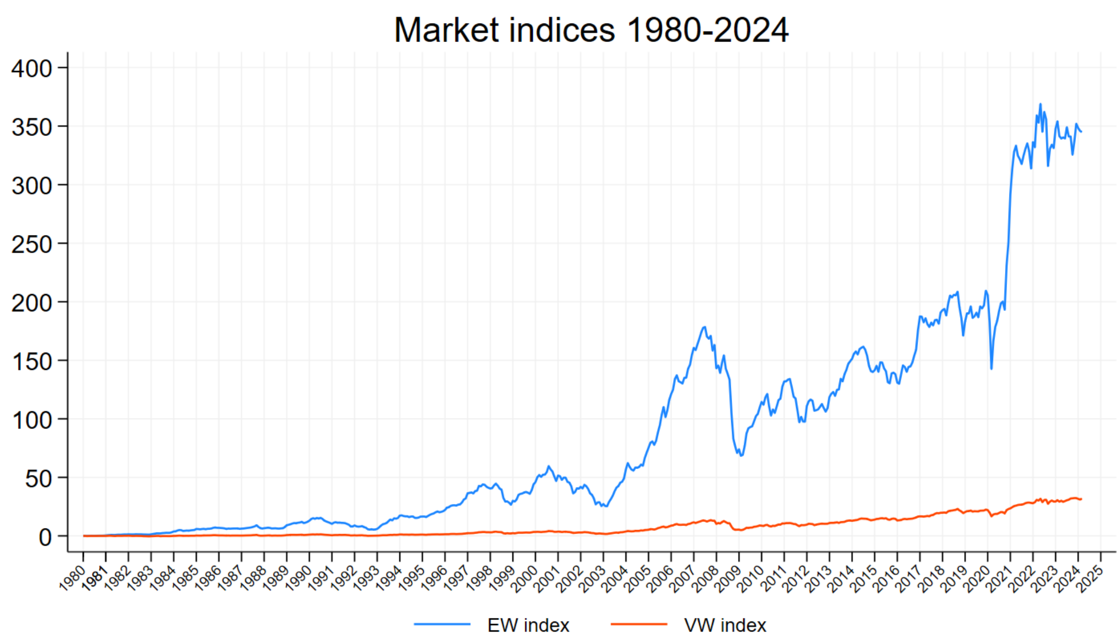

Figure 4.4.1: The graph shows the historical development of buy-and-hold returns of equal-weighted and value-weighted indices from January 1980 to March 2024.

Figure 4.4.1 shows the development of the equal-weighted and value-weighted market indices from January 1980 to March 2024. The figure shows that the equal-weighted index rapidly overtook the value-weighted one, and from that point on, the spread in performance has holistically grown larger. A possible explanation for these widely different developments could be the small-firm effect as explained in section 2.7. When comparing our market indices with the ones Swade et al. (2023) utilized using CRSP for the American stock market, we find that our results, both in descriptive statistics as well as the graphed developments, correspond well. However, it should be noted that when comparing with Ødegaard (2021) and Kristiansen (2019), we see that our market indices are quite different. These differences may be due to a different approach in data filtering, frequency of rebalancing, as well as a different data basis, as we know Ødegaard retrieved his data directly from Oslo Børs Information when this was still possible.

5 Analysis

5.1 Distribution of Buy-and-Hold Returns

In this section, we analyze the actual buy-and-hold returns from our dataset with different time horizons. If a stock is delisted prematurely, the returns are calculated to the last appearance. Moreover, the risk-free interest rate, value-weighted and equal-weighted indices are also only calculated to the stock's last appearance. This ensures that the basis for comparison is the same; for example, it would be wrong to compare the return of a stock that only lives for seven months, with a risk-free rate over an entire calendar year. For each time horizon greater than one month, the buy-and-hold returns are linked using formula (3.1). This simulates an investor who remains completely passive after the initial purchase (assumes dividends is reinvested by the broker). We thus do not take into account speculative positions where the investor makes changes to the investment based on news or similar factors. Furthermore, as explained in section 3.1, summed and geometric returns are also calculated, using formulas (3.2) and (3.3) respectively.

Table 5.1.1: The entire dataset, all stocks from January 1980 to March 2024 are included. Annual refers to calendar year. Decade refers to non-overlapping whole decades. Lifetime refers to the stocks first to last appearance. If the stock is delisted prematurely to the horizon, the return is calculated to the stocks delisting date. The risk-free rate, equal weighted and value weighted index is matched to each stocks return, for all time periods. rf is the risk-free rate. SD is the standard deviation. EW is the equal weighted index. VW is the value weighted index. Certain variables are rounded to four decimals for visual clarity.

Panel A: Individual stocks, monthly horizon (N=100,477):

Variable	Mean	Median	SD	Skewness	%>0
Buy-and-hold return, rf:	0.0041	0.0032	0,0033	0.9941	100.0%
Buy-and-hold return, stocks:	0.0106	0.0000	0.1785	5.3451	47.3%
	%>rf	%>VW index	%>EW index		
Buy-and-hold return, stocks:	46.6%	46.3%	45.7%		

Panel B: Individual stocks, annual horizon (N=9,526):

Variable	Mean	Median	SD	Skewness	%>0
Sum return:	0.1072	0.0837	0.6457	0.9586	59.4%
Buy-and-hold return, rf:	0.0453	0.0294	0.0398	1.1864	100.0%
Buy-and-hold return, stocks:	0.1554	0.0292	0.8265	7.0165	53.6%
Geometric return, stocks:	-0.0021	0.0032	0.0677	-0.5895	53.7%
	%>rf	%>VW index	%>EW index		
Buy-and-hold return, stocks:	48.9%	44.8%	42.4%		

Panel C: Individual stocks, decade horizon (N=1,793):

Variable	Mean	Median	SD	Skewness	%>0
Sum return:	0.5695	0.4779	1.4833	0.2424	70.6%
Buy-and-hold return, rf:	0.3148	0.1357	0.4626	2.8252	100.0%
Buy-and-hold return, stocks:	1.1813	0.0717	5.5153	11.7434	53.7%
Geometric return, stocks:	-0.0062	0.0019	0.0471	-1.8142	53.7%
	%>rf	%>VW index	%>EW index		
Buy-and-hold return, stocks:	46.8%	40.5%	35.9%		

Panel D: Individual stocks, lifetime horizon (N=976):

Variable	Mean	Median	SD	Skewness	%>0
Sum return:	1.0441	0.6942	2.2582	0.6617	70.8%
Buy-and-hold return, rf:	0.8803	0.2155	1.8675	4.0977	100.0%
Buy-and-hold return, stocks:	7.8011	-0.0009	65.01261	15.7175	49.8%
Geometric return, stocks:	-0.0091	-0.0001	0.0435	-0.5645	49.8%
	%>rf	%>VW index	%>EW index		
Buy-and-hold return, stocks:	43.0%	34.7%	29.4%		

5.1.1 Monthly Returns

Panel A of Table 5.1.1 shows the summary of monthly observations across our dataset. The results show the overall distribution of returns, thus taking both time series and cross-section into account simultaneously. We see that the average return is 1.06%, while the average risk-free rate is 0.41% on a monthly basis. This means that, overall, an average excess monthly return of 0.65% is achieved by investing in the stock market compared to the risk-free rate. As mentioned in section 2.2, volatility is a term for risk, and is often expressed as the standard deviation, i.e. the variation from the average. In the panel, it we see that individual stocks carry a far greater variation with 17.85% in standard deviation, compared to 0.32% for the risk-free interest rate. From Table 4.2.1 we also see that the monthly equal-weighted and value-weighted market indices have 5.84% and 5.77% standard deviation, respectively.

Furthermore, panel A shows that with a monthly time horizon, individual stocks outperform the risk-free interest rate only 46.6% of the time, the value-weighted index 46.3% of the time, and finally the equal-weighted market index 45.7% of the time. This is firstly supported by the median, which for the returns is equal to 0, while they are all positive for the risk-free rate, value-weighted and equal-weighted market indices. However, it should be noted that looking at the median of the equal-weighted and value-weighted market indices is not as important, as these indices - unlike risk-free interest rates - can, and have produced negative returns. Furthermore, the explanation for why individual stocks underperform compared to risk-free interest rates is partially found when looking at the number of returns that produce positive results. Monthly individual stocks have a greater tendency to produce negative returns, as only 47.3% of all returns are positive. This essentially means that of the returns that are actually positive, only 0.7% are lower than the risk-free rate.

So, with individual monthly returns that are predominantly negative and underperforming compared to the risk-free rate, how can the average return be outperforming? As shown in Figure 2.3.1, skewness denotes the symmetry of distribution in a dataset. With a normal distribution, the mean will be equal or close to the median. What is considered a high skewness varies depending on the type of dataset, but also among different researchers. Nevertheless, according to Byrne (2013) and Hair et al. (2010), a very liberal interval for skewness coefficients in a normally distributed univariate analyses is between -2 and +2. Since the average monthly return is far above the median, we assume a high positive skewness, which is verified as we calculate a skewness coefficient of 5.35 for the monthly returns. This high positive skewness suggests that the majority of monthly returns are below and around the median, while it has a long tail, which indicates that there is a minority of stocks producing “extreme returns”. This supports Bessembinder (2018)’s claim that a minority of stocks is what ensures that the market as a whole outperforms the risk-free interest rate.

5.1.2 Annual Returns

Panel B in Table 5.1.1 displays the summary of annual buy-and-hold returns across our dataset. For each stock, we have calculated the sum return, geometric return and buy-and-hold return. All returns, including risk-free interest rate, value-weighted and equal-weighted indices, are all calculated for whole calendar years, i.e. January 31st to December 31st, or to the last observation date if a stock is prematurely delisted. By looking at the panel we see that annual buy-and-hold returns for stocks have an average of 15.54%. In comparison, the risk-free rate has an average buy-and-hold return of 4.53%, which corresponds to an average excess return of 11.01% when investing in stocks. Although a high excess return is achieved on average, only 48.9% of individual annual returns actually outperform the risk-free interest rate. This is reflected in the high standard deviation of 82.65%, as well as the skewness with a coefficient of 7.02. Compared to the value-weighted and equal-weighted indices, the results are even worse, as only 44.8% outperform the value-weighted index and 42.4% outperform the equal-weighted index.

In contrast to the monthly horizon, the majority of annual buy-and-hold returns are positive, specifically 53.6% are. As explained in section 3.1, geometric mean return act as a measure of centrality, and in this case with a coefficient of -0.0021, indicates that the majority of returns are centered slightly below zero. This centrality is also verified in the graphical representation of the distribution in Figure 5.1.1. In the figure, we also observe the long tail that a high positive skewness coefficient refers to. The graph is set up in such a way that it has a pillar for returns within the interval -100% to 500%, where each pillar represents a return interval of 2%. Annual returns above 500% are excluded in this graph due to issues with visual clarity, and this affects 34 observed returns. The highest achieved annual buy-and-hold return in our dataset is 2,328% by Opticom ASA in the period 1999-2000. If we include Opticom ASA, the x-axis of the graph would extend to 24, thus reducing the visualization of what we wish to portray, namely the distribution and centralization of the majority of the returns. The data basis for this graph is therefore 9,492 annual buy-and-hold returns.

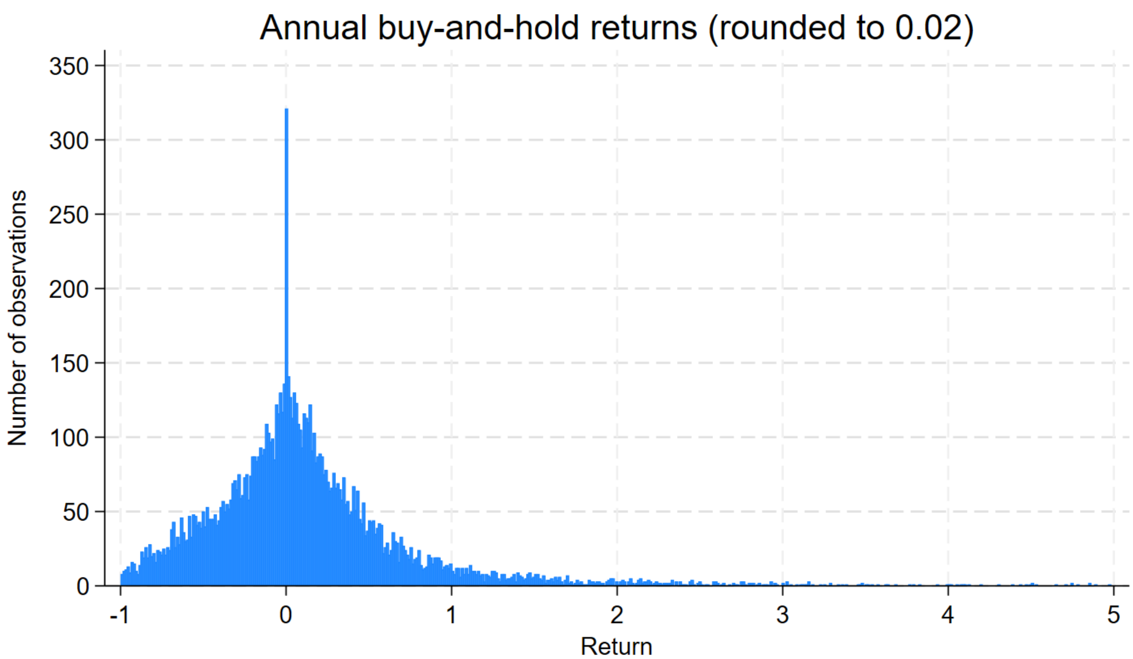

Figure 5.1.1: The histogram shows the frequency distribution of annual buy-and-hold returns rounded to 0.02 (or 2%). The graph represents all observations (9,492) with returns in the interval -1.0 to 5.0 (or -100% to 500%). Observations with higher return than 5.0 (34) are not included in the histogram. If a stock lists or delists within the calendar year, the returns are calculated for the period in which it is available.

5.1.3 Decade Returns

Panel C in Table 5.1.1 summarizes the statistical properties of our decade buy-and-hold returns. As with annual returns, sum return, geometric and buy-and-hold return is calculated for each stock. All returns refer to whole decades i.e. January 31st 1980 to

December 31st 1989, or the last observation within the decade, if the stock is delisted prematurely. For the latest decade, i.e. the one commencing in 2020, our data falls short of covering an entire decade, as we are at the time of writing in March 2024. As with stocks delisted prematurely, the period 2020-2024 is treated similarly. We find in panel C, that the average decade buy-and-hold return is 118.13%, while in comparison the average risk-free rate is 31.48%. These returns equates to an average excess return of 86.65% when investing in the market compared to the risk-free rate. While these figures may seem tempting to an investor, it turns out that only 46.8% of individual stocks outperform the risk-free rate over a ten-year horizon. Compared to the value-weighted and equal-weighted indices, the number of buy-and-hold returns that outperform also decrease, to 40.5% and 35.9% respectively.

With a median of 7.17% and the number of positive returns increased from the annual buy-and-hold returns, it is easy to think that the centering is in the positive range. The geometric return, however, shows otherwise. With a coefficient of -0.0062, somewhat lower than that of the annual returns, the centralization has actually moved downwards. These results are also verified graphically in Figure 5.1.2, where we see that far more returns are towards -1 (-100%) compared to the annual distribution. Extending the time horizon from annual to decade, both the standard deviation and the skewness has increased. This is somewhat more difficult to decipher from the graph, as the interval is from -100% to 500%, and returns beyond this interval are thus not included. Omitting returns above 500%, affects a total of 97 returns, and the data basis for the graph is thus 1,696. The data basis is naturally lower than for the annual histogram, as the decade histogram consists of a maximum of five buy-and-hold returns per stock, compared to 45 for annual one. In addition to less buy-and-hold returns, the interval remains the same, while returns have been given ten years to compound. An example of this compounding effect is the holding company Ganger Rolf, which achieved the highest decade buy-and-hold return of 8,614% in the period 1980-1990, over 6,000% higher than the biggest achiever in annual buy-and-hold returns. Worth noting, as with many stocks in our dataset, the stock was delisted in 2016 which may be due to a series of years with poor performance towards the end. In fact, in the period 2010-2016 the stock lost 58.4% of its value.

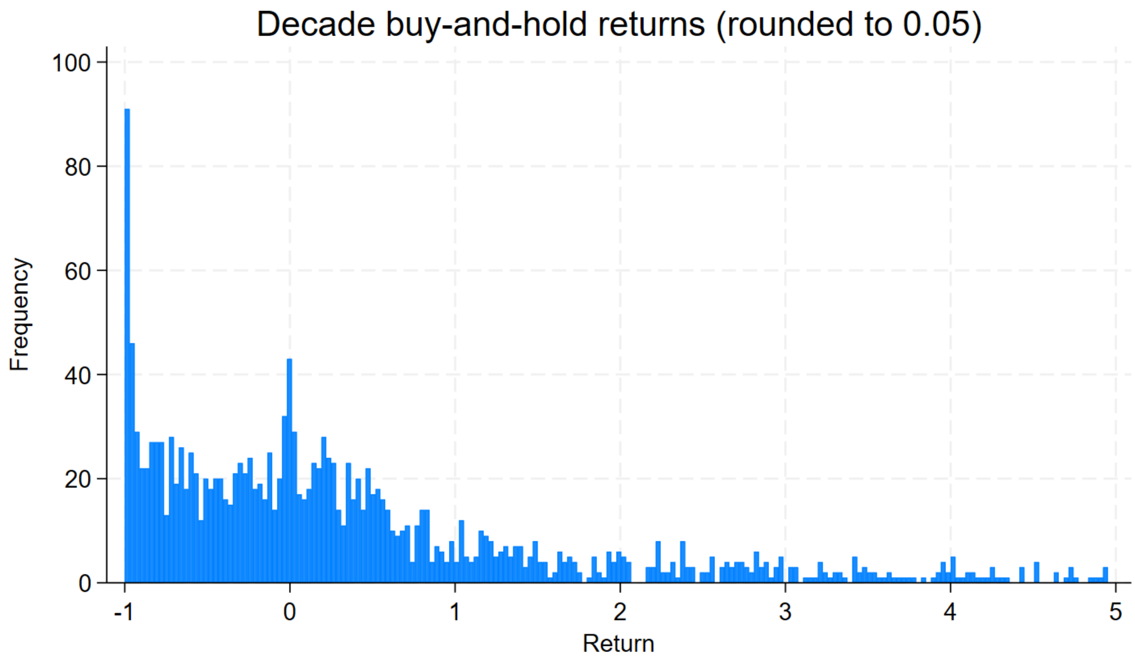

Figure 5.1.2: The histogram shows the frequency distribution of decade buy-and-hold returns rounded to 0.05 (or 5%). The graph represents all observations (1,696) with returns in the interval -1.0 to 5.0 (or -100% to 500%). Observations with higher return than 5.0 (97) are not included in the histogram. If a stock is lists or delists within the whole decade, the returns are calculated for the period in which it is available.

5.1.4 Lifetime Returns

Panel D in Table 5.1.1 shows the summary of lifetime buy-and-hold returns across our dataset. All returns are computed for the whole lifetime of our dataset, i.e. January 31st 1980 to March 31st 2024. As the majority of our stocks delist during this period, the stock returns and their corresponding risk-free rate, value-weighted and equal-weighted market indices will in most cases refer to a shorter period. The average lifetime of the stocks in our dataset is eight years and five months as of March 2024. If the stocks that are still listed, and likely to achieve a longer lifespan are removed, the average lifespan of the stocks is seven years and one month. This decrease in average lifetime is due to the fact that many of the larger companies - based on market capitalization - have generally had a long lifetime, and some are still active today. It is also worth mentioning that the Norwegian government is a major shareholder in some of these large companies, which may be one reason for the long lifespan of some of the companies. As of December

2023, the government had direct ownership in 69 companies in Norway (Regjeringen 2023). These are usually companies that are critical to the Norwegian infrastructure, such as Telenor, Equinor, Bane NOR etc. Not all companies owned by the state are listed in our dataset, as many of them are private limited companies, in which the state has 100 percent ownership. An example of a company that owes its longevity to the state is DNB. In the 1990s, the Norwegian government took over the shares in DNB, as in other financial institutions, after a finance crisis that led to major losses at most banks. In this takeover, the Norwegian government invested large sums of money to ensure the company's continued performance, which in turn has led to the Norwegian government still having a significant stake of 34 percent of the shares in DNB today.

From the panel, we see a confirmation of the trend that time horizon has a clear effect on buy-and-hold returns. Increasing the time horizon reduces the number of individual stocks that outperform the risk-free rate, value-weighted and equal-weighted indices. We note however, that there is one exception to this trend. The number of returns that outperform the risk-free rate increases in the transition from monthly to annual time horizon. From panel D, we also find that the difference between the average buy-and-hold return of stocks and the risk-free rate is increasing. For lifetime horizons, the average buy-and-hold return for stocks is 780.11%, while for the risk-free interest rate it is 88.03%. In addition to previous results, the results from panel D, indicate that Bessembinder (2018)'s assertion that a few stocks drive the entire market, also applies to the Norwegian stock market. An interesting finding from Table 5.1.1 is the development in the number of stocks that generate positive returns. There is a clear trend from monthly to annual to decade, which may indicate that Hansson and Persson (2000)'s claim that time diversification leads to a lower risk of losses is true. However, in the transition from decade to lifetime, we see a contradictory result. The proportion of returns that are positive goes from 53.7% to 49.8%. In addition, the median buy-and-hold return reaches an all-time high with decade horizon, and an all-time low with lifetime horizon. Therefore, given time horizons of monthly, annual, decade or lifetime, it is possible that the greatest effect of time diversification in the Norwegian stock market occurs by using a decade time horizon.

With a mean geometric return of -0.0091 we demonstrate the drastic exaggeration the arithmetic average buy-and-hold return can cause if used as an indicator of centralization. This exaggeration is also shown visually in Figure 5.1.3, which shows the frequency distribution of lifetime buy-and-hold returns. As with decade horizon, the numbers are rounded to the nearest 0.05, or 5%, but as the compounding effect is so large on stocks over 44 years, the interval for this graph has been increased to -1 to 10.0 or -100% to 1,000%. Thus, the graph shows data after an exclusion of 82 stocks, and the data basis is therefore 894 stocks. The stock with the highest lifetime return in our dataset is Orkla ASA, which achieved a lifetime buy-and-hold return of 101,365% (yes, you read that right), in the period January 31st 1980 to March 31st 2024. With a high average buy-and-hold return and a low centering point, there are signs that the skewness will be high. This is confirmed in the panel with a skewness coefficient of 15.72, and is due to the large compounding effect of longer time horizons. The centering point is somewhat lower than previous time horizons, while stocks with lifetime returns such as Orkla causes the distribution tail to be very long.

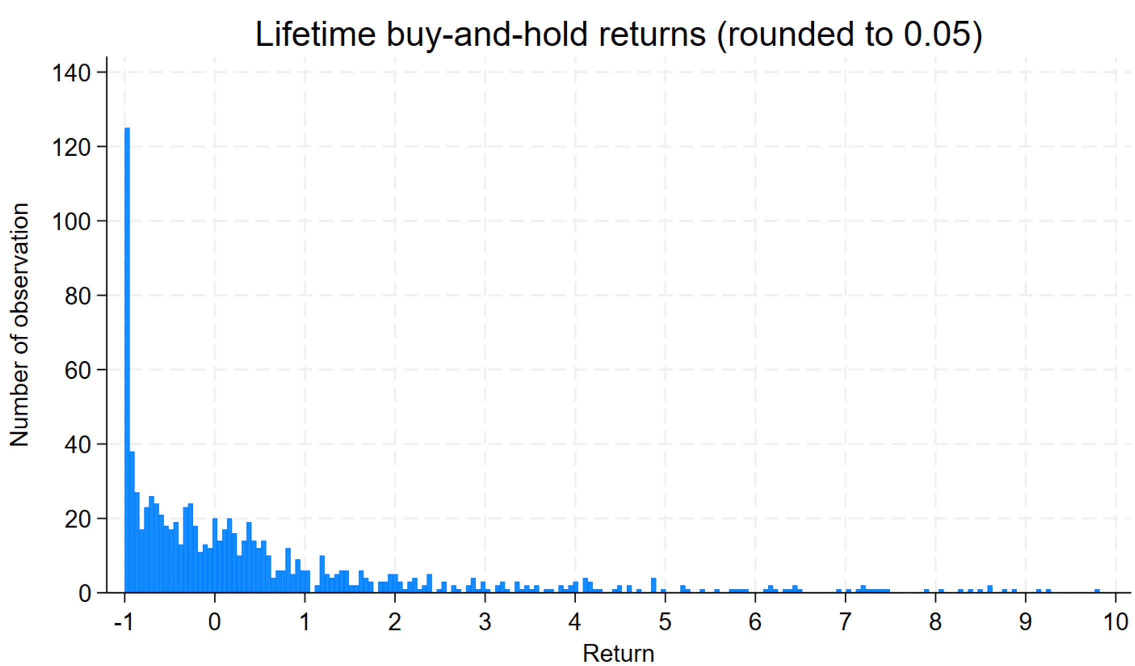

Figure 5.1.3: The histogram shows the frequency distribution of lifetime buy-and-hold returns rounded to 0.05 (or 5%). The graph represents all observations (894) with returns in the interval -1.0 to 10.0 (or -100% to 1,000%). Observations with higher return than 10.0 (82) are not included in the histogram. If a stock lists or delists within the dataset's lifetime, the returns are calculated for the period in which it is available.

5.2 Distribution of Lifetime Returns by Listing Status

As two thirds (66.26%) of the stocks in our dataset delists during the period January 1980 through March 2024, it is important to analyze which effect listing status has on our results. The database does not categorize the stocks by listing status, nor the reason for delisting, which the CRSP database does for the US market. We therefore categorize the dataset into two panels based only on listing status, i.e. whether the stock is still active as of March 2024 or not, as shown in Table 5.2.1.

Table 5.2.1: The table shows lifetime buy-and-hold returns split into two panels; A listed and B delisted stocks. The risk-free rate, equal weighted and value weighted index is matched to each stocks' return, for all time periods. rf is the risk-free rate. SD is the standard deviation. EW is the equal weighted index. VW is the value weighted index. Certain variables are rounded to four decimals for visual clarity.

Panel A: Individual stocks, listed lifetime horizon (N=330):

Variable	Mean	Median	SD	Skewness	%>0
Sum return:	1.4698	0.8982	2.5137	0.7753	72.4%
Buy-and-hold return, rf:	0.8576	0.0948	2.2626	4.1746	100.0%
Buy-and-hold return, stocks:	15.9115	0.0663	107.0862	10.0329	51.2%
Geometric return, stocks:	-0.0097	0.0012	0.0324	-1.0468	51.2%
	%>rf	%>VW index	%>EW index		
Buy-and-hold return, stocks:	47.3%	31.5%	30.3%		

Panel B: Individual stocks, delisted lifetime horizon (N=646):

Variable	Mean	Median	SD	Skewness	%>0
Sum return:	0.8273	0.5782	2.0852	0.4353	69.9%
Buy-and-hold return, rf:	0.8918	0.2893	1.6317	3.6422	100.0%
Buy-and-hold return, stocks:	3.6708	-0.0137	22.3735	11.3538	49.1%
Geometric return, stocks:	-0.0088	-0.0004	0.0482	-0.4746	49.1%
	%>rf	%>VW index	%>EW index		
Buy-and-hold return, stocks:	40.9%	36.3%	29.0%		

Panel A shows the summary of lifetime buy-and-hold returns for stocks that are still active on the stock exchange today. In total, there are 330 stocks that are still active, and these have together achieved an average lifetime buy-and-hold return of 1,592%. Compared to the lifetime returns in Table 5.1.1, these 330 stocks provide an average excess return of 812% - almost double - compared to the market as a whole. The skewness in panel A is somewhat lower than for the overall dataset, and this may be due to a significantly higher median. This in turn may indicate that the center point of the distribution is generally slightly higher for the stocks in panel A, while the spread in buy-and-hold returns remain approximately the same.

An interesting finding here is our observation that the proportion of returns that outperform the risk-free interest rate and the equal-weighted index has increased compared to the overall dataset, while the proportion that outperform value-weighted index has decreased. One reason for this could be that the overall dataset includes companies who do not survive, and thus has more returns that cannot be compared with the remarkably low interest rates we have experienced in recent years. We also present evidence of this in panel B, where the proportion of returns outperforming the risk-free rate goes down, compared to panel A. The fact that fewer stocks outperform the value-weighted index compared to the overall dataset may - like with the risk-free rate - be due to the time horizons. As shown in Figure 4.3.1, the compounding effect on the market indices has led to a extraordinary development in recent years, especially for the equal-weighted index. However, even though the equal-weighted index has performed exceptionally well overall, its performance since 2021 has stagnated. The indices are matched to the stocks' buy-and-hold returns, which leads to stocks that were listed after the exponential growth until 2021, being compared to a reasonably stagnant equal-weighted market index. In addition, the value-weighted one has performed slightly better in recent years, which might explain the decrease in proportion of returns outperforming the value-weighted index. When comparing the proportion of returns greater than the risk-free rate in panel B, the argument stays the same as with panel A, only in reverse. Since many of the stocks are delisted before the low interest rates we experienced during the Covid-19 pandemic, between 2020 and 2022, there are fewer returns compared to this low interest rate. As for the proportion of returns greater than the value-weighted and equal-weighted index, the explanation follows the same logic as with the risk-free rate. More lifetime buy-and-hold returns are compared to the exponential growth of the equal-weighted index, while the value-weighted index - in the same period - did not perform as well.

Panel B in Table 5.2.1 summarizes the statistics for lifetime buy-and-hold returns for stocks that delists in the period January 1980 through March 2024. At the time of writing, a total of 646 stocks have been delisted in this period, and together they have an average lifetime buy-and-hold return of 325%. At first glance, one would think that this is natural, as delisted stocks include more stocks with almost -100%, or bankruptcy. However, this is wrong, as 16.66% of delisted stocks deliver a lifetime return of less than 90%, while in comparison, 14.8% of the listed stocks deliver a lifetime return of less than 90%. Furthermore, when looking at the proportion of returns greater than zero, there is only a 2.1 percentage point difference between listed and delisted stocks. These results shows that there are reasons other than bankruptcy that cause delisted stocks to deliver a lifetime return that is approximately 1,200% lower than for listed stocks. There are a number of reasons why a company is no longer present on the stock exchange (Oslo Børs 2021). If one believes in a relatively efficient market and that the majority of companies on the Oslo Stock Exchange are not delisted due to legal issues with the Financial Supervisory Authority, it is reasonable to assume that a large proportion of companies that are delisted, delists due to mergers, acquisitions, privatizations or simply a choice to change the stock exchange they are listed. One example is Sbanken, who was delisted in March 2022 due to a merger with DNB, and achieved a lifetime return of 174.69% from November 2015 to delisting. Since the merger with Sbanken, DNB has experienced a return of 20.8%. Considering the price development of both stocks before (positive for Sbanken and negative for DNB) and after the Sbanken merger, it can be assumed that Sbanken accounted for a substantial amount of the return achieved by DNB. In other words, if Sbanken had not merged with DNB and was still listed, the lifetime return could have been greater than what it achieved up to the delisting date, given that we believe the price development for Sbanken would have remained the same prior to the merger. Such delisting reasons - other than financial weakness and legal problems - could thus “cut” off the lifetime return, potentially transferring the return to a company that is still listed. This argument would be consistent with both ours and Bessembinder (2018)’s table of lifetime returns based on listing status.

5.3 Distribution of Returns by Market Capitalization

The relationship between market capitalization and return has been extensively researched since it was first introduced by Rolf W. Banz (1981). As explained in section 2.7, the “small-firm effect” is significant and well documented over the years. In essence, it means that smaller companies tend to have a higher risk-adjusted return compared to larger companies. In our analysis, smaller companies will deliver an even higher average return, as our returns are not risk-adjusted according to CAPM or similar models. Although smaller companies deliver an average return that is higher than that of larger companies, it is not necessarily the case that they deliver better results than larger companies when compared with risk-free interest rates and market indices. In addition, it is interesting to see whether this small-firm effect, compared to the benchmarks, has the same development depending on the time horizon used. The following section is therefore dedicated to analyzing the performance of companies based on market capitalization compared to various benchmarks, over various horizons.

Table 5.3.1 summarizes our findings of buy-and-hold returns based on market capitalization at monthly, annual and decade horizon. The returns are grouped into deciles according to their respective market value in the period before the return was achieved. The deciles are ranked from smallest (1) to largest (10) based on market value. The ranking is therefore performed on the basis of the market capitalization at the same time as the starting price is drawn, i.e. in $t-1$. Additionally, this grouping process also means that if a stock is to qualify for inclusion in the table, it must have survived from one period to the next. However, this does not mean that the stock must have survived an entire decade, but that it has survived the transition from one decade to the next. The benchmarks are matched to the returns so that the basis for comparison is correct, as shown in Table 5.1.1. Lifetime returns, however, are not included, since over a period of just over 44 years the original market value will have little impact on the lifetime return. Omitting lifetime returns also excludes any problems with grouping that may arise with companies that are listed towards the very end of the dataset.

Table 5.3.1: The table shows monthly, annual and decade buy-and-hold returns grouped by firm size. Stocks are divided into deciles based on market capitalization from previous period. 1 is the 10% smallest firms, and 10 is the 10% largest. The risk-free rate, equal weighted and value weighted index is matched to each stocks return, for all time periods. rf is the risk-free rate. EW is the equal weighted index. VW is the value weighted index. Certain variables are rounded to four decimals for visual clarity.

Panel A: Individual stocks by market capitalization, monthly horizon:

Group (mkt. cap)	Mean	Median	Skewness	%>0	%>rf	%>VW index	%>EW index
1 (N=10,281)	0.0273	0.0000	6.6827	43.7%	43.2%	44.9%	44.5%
2 (N=10,029)	0.0098	0.0000	3.3468	43.5%	42.9%	43.8%	43.4%
3 (N=10,058)	0.0075	0.0000	5.5156	44.7%	44.4%	44.7%	44.4%
4 (N=10,050)	0.0075	0.0000	1.3359	45.9%	45.1%	45.2%	44.1%
5 (N=9,951)	0.0082	0.0000	2.1297	47.2%	46.4%	46.1%	45.6%
6 (N=10,119)	0.0072	0.0000	1.0776	47.5%	46.9%	46.6%	45.4%
7 (N=10,089)	0.0086	0.0000	0.7791	49.5%	48.7%	47.3%	46.3%
8 (N=10,015)	0.0080	0.0024	0.5770	50.5%	49.7%	47.7%	47.5%
9 (N=10,070)	0.0088	0.0063	0.4481	52.0%	51.1%	48.6%	47.9%
10 (N=9,813)	0.0083	0.0082	1.1892	53.5%	52.2%	49.8%	48.6%

Panel B: Individual stocks by market capitalization, annual horizon:

Group (mkt. cap)	Mean	Median	Skewness	%>0	%>rf	%>VW index	%>EW index
1 (N=871)	0.2849	0.0298	6.5802	52.7%	49.0%	45.5%	43.5%
2 (N=852)	0.2197	0.0349	4.7644	54.1%	48.5%	44.2%	42.3%
3 (N=851)	0.1999	0.0209	4.2033	52.4%	48.4%	43.0%	42.4%
4 (N=854)	0.1441	0.0417	2.3797	53.9%	49.2%	44.0%	41.7%
5 (N=844)	0.1695	0.0278	2.8449	53.4%	48.7%	44.9%	42.4%
6 (N=858)	0.1495	0.0647	1.9431	54.5%	50.6%	44.9%	41.8%
7 (N=858)	0.1646	0.0437	13.7203	53.5%	49.7%	45.1%	42.8%
8 (N=847)	0.1343	0.0429	1.4008	57.9%	51.4%	47.9%	42.7%
9 (N=856)	0.1462	0.0673	1.3398	57.2%	52.5%	47.5%	44.6%
10 (N=834)	0.0921	0.0679	0.5085	59.6%	54.1%	46.9%	44.7%

Panel C: Individual stocks by market capitalization, decade horizon:

Group (mkt. cap)	Mean	Median	Skewness	%>0	%>rf	%>VW index	%>EW index
1 (N=83)	2.3873	0.2688	7.6187	55.4%	49.4%	38.6%	37.3%
2 (N=82)	1.1460	0.2793	2.8743	57.3%	47.6%	46.3%	45.1%
3 (N=82)	0.9722	0.2855	6.6711	58.5%	51.2%	41.5%	37.8%
4 (N=81)	0.6431	0.1618	2.7893	58.0%	49.4%	41.9%	35.8%
5 (N=80)	1.0447	0.3301	5.3143	63.8%	55.0%	50.0%	45.0%
6 (N=83)	0.6939	0.1552	4.1084	53.0%	50.6%	43.4%	34.9%
7 (N=81)	0.6715	0.1853	1.6122	59.3%	48.1%	50.6%	45.7%
8 (N=82)	1.0606	0.1854	2.4654	63.4%	46.3%	48.8%	41.5%
9 (N=82)	1.1083	0.3805	3.0235	65.9%	56.1%	43.9%	42.7%
10 (N=79)	0.6138	0.2403	2.8462	63.3%	51.9%	41.8%	32.9%

Panel A verifies Banz (1981)'s theory and NBIM (2012)'s survey that smaller firms deliver higher mean buy-and-hold returns. However, the trend is not linear since all groups outperform group 6. Moreover, there are fluctuations in the highest returns among them. On the other hand, there is a clear trend in the proportion of returns that are greater than zero, and those that are greater than the risk-free interest rate. We observe that companies with a lower market capitalization - apart from group 1 outperforming group 2 - deliver fewer returns greater than zero, and the same applies to the proportion greater than the risk-free interest rate and value-weighted market index. A comparison with the equal-weighted index reveals some interesting findings. The proportion of returns that outperforms the equal-weighted index seems to fluctuate a lot more when we iterate through the groups. Although the proportion fluctuates, we see the same trend as when comparing the returns to the other benchmarks. Thus, there is an interesting trend in which the companies with the 10% lowest market capitalization outperform the neighboring groups of companies - in terms of market capitalization - when compared with all of the benchmarks. We also see that the skewness is generally higher among the smaller companies. The difference in skewness may be due to the difference between the median and the mean is greater among the smaller companies, while the standard deviation experiences a proportionally weaker development than the difference in median and the mean.

Panel B expands the horizon and summarizes the statistics of annual buy-and-hold returns grouped by market capitalization. Apart from a few small fluctuations in average buy-and-hold returns among the middle groups, we see that the small-firm effect also applies to our dataset on an annual basis. The distribution of skewness is somewhat surprising, as group 7 has some comparatively extreme values, while beyond this group we see that market capitalization affects skewness very clearly. In contrast to panel A, all groups deliver a majority of returns greater than zero. It is difficult to draw a firm conclusion on the role of market capitalization as a result of these observations, since the proportion greater than zero seems to jump up and down for each group we iterate downwards. In fact, when examining the proportion of returns greater than the risk-free interest rate and the equal-weighted market index, we see the same development and fluctuations as with the proportion greater than zero. However, these fluctuations are not present when observing the proportion of returns greater than the value-weighted index. Comparing with the value-weighted benchmark we see that - apart from group 1 and 2 - the trend is clear and linear, where higher market capitalization increases the proportion of returns greater than value-weighted index. Interestingly, the 10% smallest companies outperform the neighboring six groups in returns larger than the value-weighted index. This development is surprising, as we would expect smaller companies' returns to be more volatile, given that the proportion of returns greater than zero is lower compared to almost every other group, while the mean return being highest. Overall, expanding the time horizon from monthly to annual increases the proportion of returns greater than zero and the risk-free rate, while the proportion of returns greater than the market indices is approximately the same for the smaller companies, and lower for the medium and bigger companies. With all the fluctuations in the benchmarks, it is difficult to draw a linear conclusion based on the relationship between market capitalization and our benchmark on an annual horizon, although the general trend remains clear.

If we extend the time horizon all the way to decade, we obtain some interesting results. Panel C summarizes our market capitalization-grouped data at a decade level, and we see here that - with a decade horizon - we can still to some extent, conclude that Banz (1981)'s theory of the small-firm effect is valid. Our conclusion is based on the phenomenon that Banz described, that the small-firm effect is non-linear, and that this effect was only observable at absolute smallest companies. The rest - average sized and large companies - displayed the equal return. In our panel C we observe that the 10% smallest clearly deliver the highest returns, and the 10% largest deliver the lowest, but in the groups between them there is no clear trend or development direction. We see - as with monthly and annual horizon - similar fluctuations throughout all the statistical properties as well as the benchmark for risk-free rate and greater than zero. When examining the proportion of returns greater than the market indices, the results for decade horizon differs quite a lot from previous horizons. Returns greater than value-weighted index is lowest for the smallest companies, and the biggest companies are not far behind. When we iterate through the groups, the fluctuations are more apparent, as we see the highest proportion of returns larger than the value-weighted index in group 7 at 50.6%, while the lowest is in group 1 at 38.6%. The proportion of returns greater than the equal-weighted index share similar traits - as with the comparison to value-weighted index -, however, the largest 10% of companies are the ones performing worst compared to the equal-weighted index. Most interesting of results is that group 7 - which had the highest skewness on annual horizon - now for some reason, has the lowest. This remarkable development may be due to the fact that some companies in this group have been particularly sensitive to macro-economic conditions that have led to some extreme buy-and-hold returns with an annual time horizon. When the time horizon increases to decades, these same stocks may not have survived the transition from one decade to the next, and are therefore not included in panel C. Not surviving the transition, could explain the sharp reduction in skewness coefficients, going from 13.7203 at annual horizon, to 1.6122 at the decade horizon.

Based on the results presented in Table 5.3.1, we can see a clear small-firm effect in both the average return, but also in the comparison with the various benchmarks at the monthly horizon. In the extension to the annual horizon, we will still see a small-firm effect in average returns, but it is difficult to draw a firm conclusion from these data on the market capitalization's effect when comparing the buy-and-hold returns with our benchmarks. Finally, when the horizon is extended to ten years, it is clear that market capitalization has no apparent effect on our benchmarks. Although the approximate linear trend shown in monthly and annual horizons do not appear in the decade horizon, we still - in accordance with Banz' argument of non-linearity - conclude that there is a small-firm effect on the mean buy-and-hold returns based on the visual representation of the table.

5.4 Distribution of Lifetime Returns by Decade of Initial Appearance

As with all capitalist economies, the Norwegian stock market is sensitive to macroeconomic factors, and is therefore always changing. Because the market is always changing, certain factors or characteristics that defined a company at its IPO might persist to this day. It is therefore important to analyze companies by IPO decade to see whether certain macroeconomic factors which defined the stock market at that time has lead to better or worse results in the long run. Table 5.4.1 summarizes lifetime buy-and-hold returns for our dataset, based on which decade the stocks are first introduced to the market. Stocks that were listed before 1980 are treated as if they were introduced in the period January 1980 to December 1989, since the dataset does not extend further back in time.

Table 5.4.1: The table shows all lifetime buy-and-hold returns based on the decade the stocks were first introduced. Entire dataset is used, and stretches from January 1980 to March 2024. The risk-free rate, equal weighted and value weighted index is matched to each stocks return, for all time periods. The annualized return is the mean annualized return for all the stocks within the grouping. rf is the risk-free rate. EW is the equal weighted index. VW is the value weighted index. Certain variables are rounded to four decimals for visual clarity.

Lifetime buy-and-hold returns by decade of initial appearance

Initial decade	N	Mean	Median	Skewness	%>0	%>rf	%>VW	%>EW	Months lived	Annualized return
1980-1989	160	36.4599	1.5273	6.5039	65.0%	46.9%	50.0%	25.6%	200	0.0407
1990-1999	259	5.4361	0.2422	6.5799	58.5%	49.6%	39.2%	32.7%	119	0.0375
2000-2009	235	1.2382	-0.2798	5.1847	43.2%	40.3%	33.1%	33.1%	99	0.0157
2010-2019	124	0.9376	0.0364	4.0851	51.6%	48.4%	29.8%	30.6%	83	-0.0386
2020-2024	198	-0.1313	-0.3787	3.3781	32.8%	31.3%	21.2%	23.2%	30	-0.1511

From Table 5.4.1, we see that the mean lifetime buy-and-hold return is highest among stocks that were introduced in the 1980s. In fact, we see a clear trend that the longer it's been since the stocks were first introduced to the stock market, the higher the average lifetime buy-and-hold return they have. The fact that older stocks have a higher mean lifetime return is a natural result because of the exponential effect that compounding has. Thus, even though 87.5% of the stocks introduced in the 80s are delisted today, they have on average the longest lifespan of 200 months lived on the stock exchange. Given the difference in average lifespan, we examine the mean annualized return in order to make the basis of comparison fair. From the table we see similarities between the mean lifetime buy-and-hold and the mean annualized return, where the stocks who were introduced prior and during the 1980s, on average provides the highest annual return. Furthermore, we see a declining trend in the mean annualized return when iterating from oldest to newest decade of initial appearance. A particularly interesting finding is that stocks that listed during and soon after the Covid-19 pandemic, hold a remarkably low average annualized return of -15.11%. The declining trend in annualized return confirms Fama and French (2004)'s findings, which they argue is due to the fact that newer listed stocks' cross-sectional profitability is less compared to older listings', and that the asset growth rate is stronger. These characteristics result in a sharp decline in average lifespan, which Fink et al. (2010) also found in their studies on the idiosyncratic volatility during the internet boom. Fink et al. (2010) argue that most of the increase in unsystematic risk during the boom can be explained by newer listings having a tendency to be younger, and therefore less economically established - compared to older listings - at the date of listing.

When we compare the lifetime buy-and-hold returns with the benchmarks, we observe varying results. The proportion of returns greater than zero tend to be higher with older listings, with a spike among stocks that listed in the decade 2010-2019. In three of the five decades reported, over half of the lifetime buy-and-hold returns exceed zero, while the stocks listed in the decades 2000-2009 and 2020-2024, generated lifetime returns where only 43.2% and 32.8% exceeded zero respectively. In comparison with the risk-free rate, all of the groups - by decade of initial appearance - report proportions of returns where less than half exceed the risk-free rate. Stocks listed in 1990s perform the best compared to the risk-free rate, while the ones listed in 2020-2024 perform the poorest. The fact that the stocks listed in 1990s perform best - compared to the risk-free rate - is mainly due to the fact that they generate a relatively large proportion of high returns, and the especially high risk-free rate we experienced in the 80s, declined during this period. Even though the stocks listed in 2020-2024 are generally compared with a low risk-free rate, they generate a large proportion of fairly low lifetime buy-and-hold returns, which results in only 31.3% outperforming of the risk-free rate. An interesting finding is that the stocks listed in 2010-2019 perform second best compared to the risk-free rate, while the mean annualized return is negative. This result could suggest that either - compared to the other groups - the majority of returns that are greater than zero (51.6%) are closer to zero, or that the returns that are below zero are more negative, or more plausibly, a combination of the two conditions.

The proportion of lifetime buy-and-hold returns greater than the value-weighted index show a rapid decline for the stocks listed after the 1980s, whereas comparing the returns with the equal-weighted index shows an interesting fluctuation between the groups. In fact, at the turn of the millennium we see a shift in where the proportion of returns that outperform the equal-weighted index is higher than the proportion outperforming the value-weighted one. A possible explanation for this shift could be the combination of equal-weighted index being more volatile - as shown in 4.4 -, and the fact that newer stocks on average have a lower lifespan. This will in turn yield more lifetime returns compared to an equal-weighted index in decline.

5.5 Bootstrapped Portfolios

As previously mentioned, the average lifespan of the stocks in our dataset is eight years and five months. Of the 976 stocks included in this thesis, only two stocks; "Orkla" and "Norsk Hydro" have a lifespan equivalent to the length of our dataset. In addition, only 87 of the stocks live longer than half of the period 1980-2024. These short lifespans suggests that the long-term buy-and-hold returns calculated, are limited to the lifespan of the stocks, and therefore does not necessarily reflect whole periods. In order to observe the buy-and-hold returns for whole long-term periods we make use of bootstrap simulations. Bootstrapping is a stochastic process we utilize in order to draw a random stock each month for the entirety of of out dataset. In essence, this process simulates and investor who at all times holds a random stock, and its buy-and-hold return is therefore not limited by the lifespan of a company.

In accordance with Bessembinder (2018), we repeat the process of selecting a random stock for each month in our dataset 20,000 times to ensure that potential random spikes in the distribution of returns are equalized. The monthly returns obtained from bootstrapping are then linked to annual, decade and lifetime buy-and-hold returns. Lastly, the linked buy-and-hold returns are then compared to zero, as well as the risk-free rate and the value-weighted index obtained by holding the assets in the same period. In addition to examining the effects of bootstrapping a single-stock portfolio, we also create value-weighted portfolios of 5, 10, 20, and 40 stocks in order to see the effects of diversification.

Table 5.5.1: The table shows annual, decade and lifetime returns for bootstrapped portfolios of various sizes. Each portfolio is a result of 20,000 simulations, in which monthly returns, risk-free rate and VW index are calculated to the respective horizon. The risk-free rate and value weighted index is matched to each portfolios return, for all time periods. rf is the risk-free rate. VW is the value weighted index. Certain variables are rounded to four decimals for visual clarity.

Panel A: Bootstrapped 1-stock portfolio

	Annual			Decade			Lifetime		
	Mean	Median	Skewness	Mean	Median	Skewness	Mean	Median	Skewness
Holding return	0.1943	0.1833	1.3150	3.2639	1.6255	12.2430	251.1803	-0.1241	104.4948
%>0	52.41%			49.37%			48.54%		
%>rf	47.65%			36.77%			20.45%		
%>VW index	44.61%			34.25%			14.48%		

Panel B: Bootstrapped 5-stock value weighted portfolio

	Mean	Median	Skewness	Mean	Median	Skewness	Mean	Median	Skewness
Holding return	0.1470	0.1452	0.2052	1.8133	1.4283	2.9078	81.7689	16.5418	15.2209
%>0	60.01%			73.22%			94.21%		
%>rf	53.96%			51.89%			53.54%		
%>VW index	48.42%			41.50%			36.35%		

Panel C: Bootstrapped 10-stock value weighted portfolio

	Mean	Median	Skewness	Mean	Median	Skewness	Mean	Median	Skewness
Holding return	0.1379	0.1368	0.1609	1.5107	1.2883	2.4155	65.4504	22.7241	10.9188
%>0	62.06%			81.92%			98.46%		
%>rf	55.79%			56.98%			62.64%		
%>VW index	49.18%			43.34%			41.25%		

Panel D: Bootstrapped 20-stock value weighted portfolio

	Mean	Median	Skewness	Mean	Median	Skewness	Mean	Median	Skewness
Holding return	0.1294	0.1290	0.0770	1.2961	1.1844	1.3699	50.1210	26.1346	5.7192
%>0	63.59%			88.29%			99.72%		
%>rf	57.14%			60.49%			69.25%		
%>VW index	49.43%			44.58%			43.41%		

Panel E: Bootstrapped 40-stock value weighted portfolio

	Mean	Median	Skewness	Mean	Median	Skewness	Mean	Median	Skewness
Holding return	0.1232	0.1227	0.0959	1.1749	1.1168	0.9721	41.6849	27.9286	3.4365
%>0	65.13%			93.45%			99.99%		
%>rf	58.56%			63.79%			76.77%		
%>VW index	49.57%			45.44%			44.23%		

Panel A in Table 5.5.1 shows the results of bootstrapping a single-stock portfolio. The mean buy-and-hold returns are significantly higher than to those reported in 5.1.1, which is caused by the fact that buy-and-hold returns are extended to whole periods.

By randomly selecting a stock each month, the mean buy-and-hold return in the period 1980-2024 is 251.2, or 25,180%, compared to 780% by holding one random stock for its lifespan. Interestingly, the proportion of returns that outperform the benchmarks is lower. The fact that fewer returns outperform the benchmark supports the previous argument that the reduction in the buy-and-hold periods lead to a higher chance that more returns are compared to a period where the risk-free rate and the value-weighted index is in decline, or stagnant. While we see differences between the two tables, they also share similarities in that the proportion of returns outperforming the benchmarks is reduced when extending the buy-and-hold period. By expanding the period in which we are able to compound returns, we expect in accordance with Bessembinder (2018)'s findings, that the skewness increases. In our findings, however, these results are only true for the decade and lifetime buy-and-hold returns, while the skewness in annual returns is lower than the one reported earlier, in 5.1.1.

Panel B to E in Table 5.5.1 presents the results from bootstrapping varying sized value-weighted portfolios. Looking at the development in skewness as we increase the portfolio size, we see that the findings from Simkowitz and Beedles (2010) about diversification reducing skewness, also applies here. From a single-stock portfolio to a 40-stock portfolio, the skewness goes from 1.3150 to 0.0959 with an annual horizon. In fact, the same pattern applies to all time horizons, and the skewness of lifetime buy-and-hold returns is reduced mostly due to diversification, going from 104.5 with a single-stock portfolio, to 3.4 with a 40-stock portfolio. The returns and median are also greatly affected by diversification. For annual and decade horizons, the mean return and median share the same declining trend as portfolio size increases, while in contrast, the lifetime median buy-and-hold return increases with portfolio size. This is quite a puzzling result since following Bessembinder (2018), we expect to see the median increase for all time horizons as we increase the portfolio size.

When comparing the buy-and-hold returns with our benchmarks, the value of diversification becomes even more apparent. Going from one to five stocks in a portfolio, increases the proportion of returns outperforming the benchmarks significantly, with the highest effect seen in lifetime buy-and-hold returns. In fact, at any given time horizon, all portfolios consisting of more than one asset, provides a proportion of returns outperforming the risk-free rate greater than 50%. Furthermore, when comparing the returns with zero, we see that a significant majority is greater for all time horizons, with as many as 99.99% of lifetime buy-and-hold returns in a 40-stock portfolio being greater than zero. This can, however, not be said in comparison with the value-weighted index. Although there is a significant increase in the proportion of returns that outperform the value-weighted index going from one to five stocks, less than half of the portfolio returns at any given time horizon outperform the value-weighted index. As number of stocks in the portfolio increases, so does the proportion of returns outperforming the value-weighted index. However, as with the other benchmarks, the growth in percentage of portfolios outperforming the value-weighted index becomes smaller and smaller, supporting Lee, Lee, and Lee (2010)'s claim of nullification of unsystematic risk through diversification.

5.6 Wealth Creation

So far in this thesis, we have seen repeated evidence that the majority of single-stock returns underperform compared to the risk-free interest rate. As a final segment, we shift our focus from the distribution of stock returns to the distribution of wealth creation on the Oslo Stock Exchange in the period 1980-2024. As explained in section 2.6, we calculate the wealth creation as the lifetime monetary change in market capitalization in excess of the risk-free interest rate achieved over the same period. For companies with several share classes, such as “Hafslund Ser. A” and “Hafslund Ser. B”, we add the market values together to create a total market capitalization for the company each month. By combining the share classes, we obtain a total data basis for wealth creation of 845 companies. Totalling the wealth creation of all the companies in the database, we find that the total wealth creation in the market is NOK 1,996 billion in the period 1980-2024. The average lifespan of the companies in our dataset is 9.8 years, where seven of these companies have a lifetime equal to our time period 1980-2024.[3]

[3]Company lifespan must not be confused with stock lifespan. A few stocks in our dataset - not necessarily due to having multiple share classes - change their ISIN throughout the period 1980-2024 while retaining the same company-id.

Table 5.6.1: The table shows total lifetime wealth creation for the top 50 firms on the OSEBX. Wealth created is the excess (compared to the risk-free rate) difference in market capitalization from first to last appearance. Wealth created is displayed in NOK millions. Name displayed is the most recent name linked to the company-id. For companies with multiple share classes, the market capitalization is summed up into one total market capitalization per company-id. Annualized return is calculated using the ISIN (or stock) with the longest life span in months, per company-id.

Wealth creation

Company-id	Name	Wealth created (NOK millions)	% of total	Cumulative % of total	ISIN	Annual return	Start	End	Months lived
1309	EQUINOR	557000	27.9%	27.9%	NO0010096985	11.4%	2001-06	2024-03	274
1772	DNB BANK	313000	15.7%	43.6%	NO0010031479	15.9%	1980-01	2024-03	531
8376	AKER BP	166000	8.3%	51.9%	NO0010345853	13.5%	2007-12	2024-03	196
2202	ADEVINTA	131000	6.6%	58.5%	NO0003031809	7.2%	1992-12	2024-03	376
2331	KONGSBERG GRUPPEN	113000	5.6%	64.2%	NO0003043309	19.6%	1993-12	2024-03	364
5063	MOWI	101000	5.1%	69.2%	NO0003054108	-5.9%	1997-07	2024-03	321
6046	ROYAL CARIBBEAN CRUISES	91900	4.6%	73.8%	LR0008862868	8.0%	1997-08	2016-03	224
8288	SALMAR	86100	4.3%	78.2%	NO0010310956	23.2%	2007-05	2024-03	203
1107	ORKLA F	72200	3.6%	81.8%	NO0003733800	16.9%	1980-01	2024-03	531
2354	SCHIBSTED SER. B	67400	3.4%	85.2%	NO0003028904	13.8%	1992-07	2024-03	381
7760	YARA INTERNATIONAL	59800	3.0%	88.2%	NO0010208051	14.6%	2004-03	2024-03	241
6039	FRONTLINE	47100	2.4%	90.5%	BMG3682E1921	6.0%	1997-07	2024-03	321
7187	SUBSEA 7	47000	2.4%	92.9%	LU0075646355	5.2%	1997-06	2024-03	322
11217	GJENSIDIGE FORSIKR	47000	2.4%	95.2%	NO0010582521	15.4%	2010-12	2024-03	160
1955	UNI STOREBRAND PRF	42100	2.1%	97.3%	NO0003053605	7.7%	1980-01	2024-03	531
6047	AKER	38500	1.9%	99.3%	NO0010234552	19.3%	2004-09	2024-03	235
8620	BAKKAFROST	36200	1.8%	101.1%	FO0000000179	25.8%	2010-03	2024-03	169
11186	WALLENIUS WILHELMS	33500	1.7%	102.8%	NO0010571680	19.9%	2010-06	2024-03	166
1797	SPAREBANK 1 SR-BK	33500	1.7%	104.4%	NO0006000009	15.7%	1994-05	2024-03	359
12741	HAFNIA LIMITED	29100	1.5%	105.9%	BMG4233B1090	38.4%	2019-11	2024-03	53
6233	TELENOR	27700	1.4%	107.3%	NO0010063308	9.0%	2000-12	2024-03	280
7148	LERØY SEAFOOD GP	26800	1.3%	108.6%	NO0003096208	18.4%	2002-06	2024-03	262
2392	TOMRA SYSTEMS	26000	1.3%	109.9%	NO0005668905	8.7%	1985-01	2024-03	471
12424	GOLDEN OCEAN GROUP	25800	1.3%	111.2%	BMG4032A1045	4.1%	2004-12	2024-03	232
6068	AMERSHAM	24000	1.2%	112.4%	GB0002747532	11.8%	1997-10	2004-03	78
7968	SEADRILL	23500	1.2%	113.6%	BMG7945E1057	-21.0%	2005-11	2024-03	221
1103	CHR. BANK OG KREDITKASSE	23300	1.2%	114.8%	NO0003003501	-9.4%	1980-01	2001-01	253
1695	SPAREBANK 1 SMN	19500	1.0%	115.8%	NO0006390301	15.8%	1994-05	2024-03	359
2345	PROTECTOR FORSIKRG	17700	0.9%	116.6%	NO0010209331	22.6%	1993-03	2024-03	373
5122	NETCOM	17500	0.9%	117.5%	NO0003057507	48.2%	1996-05	2000-10	54
1937	HAFSLUND NYCOMED F	17200	0.9%	118.4%	NO0004306408	17.0%	1980-01	2017-08	452
5117	NORDIC SEMICONDUC	16200	0.8%	119.2%	NO0003055501	19.1%	1996-04	2024-03	336
11306	BORREGAARD	16000	0.8%	120.0%	NO0010657505	24.4%	2012-10	2024-03	138
5097	STOLT-NIELSEN B	15900	0.8%	120.8%	BMG850801025	6.0%	1996-02	2024-03	338
7628	NORWEGIAN AIR SHUT	15400	0.8%	121.6%	NO0010196140	-16.9%	2003-12	2024-03	244
8252	ALGETA	14900	0.7%	122.3%	NO0010239437	34.5%	2007-03	2014-03	85
2455465	HÖEGH AUTOLINERS	14600	0.7%	123.0%	NO0011082075	122.6%	2021-11	2024-03	29
5103	BLUENORD	14200	0.7%	123.7%	NO0010379266	-28.6%	1987-05	2024-03	443
8594	FLEX LNG	13900	0.7%	124.4%	BMG359472021	14.2%	2009-10	2024-03	174
6051	AF GRUPPEN B	13700	0.7%	125.1%	NO0003078107	18.7%	1997-09	2024-03	319
2400	VEIDEKKE	13700	0.7%	125.8%	NO0005806802	15.8%	1986-06	2024-03	454
2079575	CADELER	13400	0.7%	126.5%	DK0061412772	22.1%	2020-11	2024-03	41
1642	SPAREBANKEN VEST	12700	0.6%	127.1%	NO0006000900	11.1%	1995-01	2024-03	351
6802	GOLAR LNG	12700	0.6%	127.8%	BMG9456A1009	13.8%	2001-07	2012-08	134
8004	PETROBANK ENERGY AND RESOURCES	12700	0.6%	128.4%	CA71645P1062	56.5%	2006-06	2008-07	26
6070	TGS	11600	0.6%	129.0%	NO0003078800	8.3%	1997-10	2024-03	318
2305	ATEA	11500	0.6%	129.6%	NO0004822503	4.0%	1985-03	2024-03	469
6053259	DOF GROUP	10900	0.5%	130.1%	NO0012851874	136.8%	2023-06	2024-03	10
1827	SPAREBANKEN NOR N	10700	0.5%	130.6%	NO0006390103	24.1%	1991-06	2002-09	136
12241	BW LPG	9800	0.5%	131.1%	BMG173841013	17.1%	2013-11	2024-03	125

Table 5.6.1 shows the top 50 companies sorted by wealth creation in descending order. The reported start and end dates, as well as number of months lived pertains to the company, not any particular stock linked to the company. Furthermore, we assign the annualized return to the longest-lived stock for each company. Certain companies are reported with a negative annualized return, while the wealth creation is positive. This may occur as a result of using the stock that lives the longest, while in reality

there may be another stock - which lives shorter - that accounts for a positive return. Furthermore, some companies in Table 5.6.1 - such as Norwegian Air Shuttle - have undergone complex financial crises where they have been required to raise capital.[4] When raising capital, it is possible that the market capitalization increases even if the adjusted share price decreases. This is due to the fact that new shares issued can lead to a price loss that is disproportionate to the number of shares issued, e.g. a price loss of -20%, while the number of shares is increased by 50%. The highest ranked company in our dataset is Equinor with a wealth creation of NOK 557 billion, which corresponds to 27.5% of the total wealth creation in the market. By adding up the top three performing companies; Equinor, DNB Bank and Aker BP, we find that they collectively account for over half of the wealth creation in the Norwegian stock market. In fact, if we add up the top 17 largest contributors to wealth creation, we pass 100% of the net wealth creation. Essentially, this means that the top 2% of companies - ranked by wealth creation - collectively account for all of the net wealth creation in the Norwegian stock market. The remaining 98% of companies in our dataset, thereby collectively generate lifetime monetary gains that equals the risk-free interest rate.

As we have seen earlier in the analysis, time horizon is a major contributor to increased skewness in the distribution of buy-and-hold returns. Combining the time horizon with large initial market capitalization appears to increase the spread in the distribution of wealth creation. The top and bottom 2% have an average lifespan of 26.3 and 17.2 years respectively, while the remaining 96% have an average lifespan of 9.3 years. Furthermore, the average market capitalization at IPO for the top and bottom 2% are NOK 14.2 billion and NOK 24.0 billion respectively, while the middle 96% have an average market capitalization at IPO of NOK 1.6 billion. Our findings are thus that a company with a high market capitalization at IPO and a long lifetime will on average either generate the largest gains or the largest losses in monetary value, depending on positive or negative excess return respectively.

[4] See "Litt om Norwegian emisjonen" at `https://aksjenorge.no/aktuelt/2020/05/11/nas_mai2020/` for the story of how new share issues affected shareholders in Norwegian Air Shuttle.

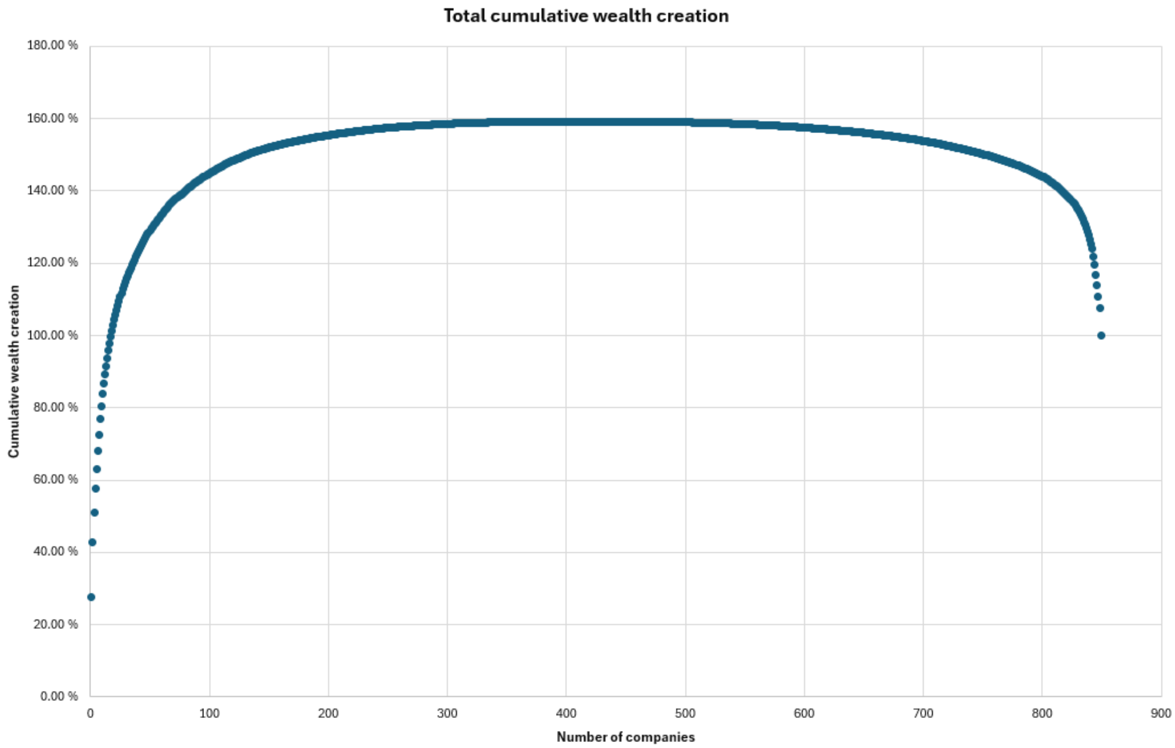

Figure 5.6.1: The graph shows the total cumulative net wealth creation of companies listed on the OSEBX from January 1980 to March 2024. Companies are sorted from largest to smallest wealth creation.

Figure 5.6.1 shows the cumulative net wealth creation for all the companies in our dataset. As the companies are sorted from largest to smallest wealth creation, the graph is ascending until the companies with zero or negative wealth creation are included. By design, the curve naturally asymptotes at 100%, as this is the net wealth created in the market. Of the 845 companies in our dataset, only 48.9% of them generate a positive wealth creation, which is why we see a declining development in the graph, past the top 414 companies. The curve reaches a maximum of 161.1% by including the 414 companies that generate positive wealth, which in turn means that the gross wealth creation in the Norwegian stock market is 61.1% larger than the net wealth creation. In monetary value, this gross wealth creation equates to NOK 3,213 billion, compared to NOK 1,996 billion in net wealth creation. While the top 2% of companies collectively account for all the net wealth creation in the stock market, the bottom 2% collectively account for 34.4% of the net wealth destruction in the Norwegian stock market. The bottom 2% of companies thereby offset all of the collective wealth created by the middle 96% of companies.

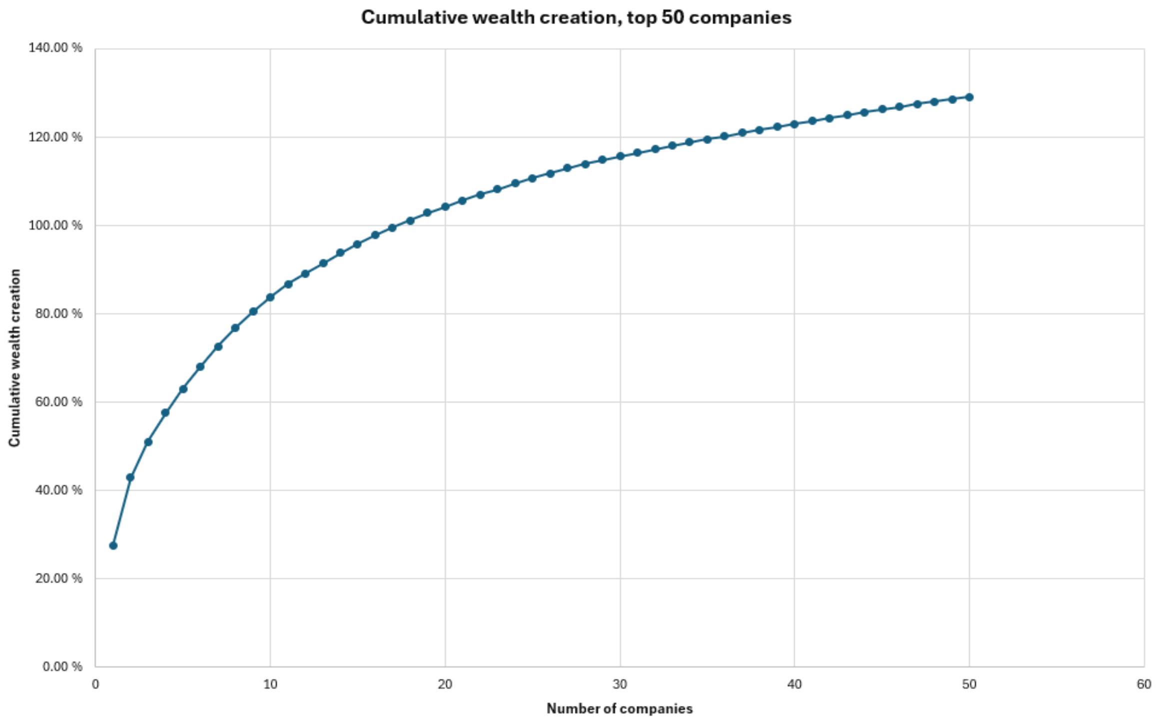

Figure 5.6.2: The graph shows the cumulative net wealth creation of the top 50 (by wealth creation) companies listed on the OSEBX from January 1980 to March 2024. Companies are sorted from largest to smallest wealth creation.

Figure 5.6.2 shows the cumulative percentage of wealth creation for the top 50 companies in our dataset. Similar to Figure 5.6.1, the companies are sorted from largest to smallest in wealth creation. All the companies included generate positive wealth in the period 1980-2024, and the curve therefore reaches its peak at 50 companies. These companies collectively provide a gross wealth creation of 131.1%. In other words, the top 5.9% of companies in our dataset generate a monetary wealth equivalent to NOK 2,620 billion.

6 Conclusion

In this thesis, we have studied the performance of Norwegian stocks in comparison to the risk-free interest rate. In the period 1980-2024, the Norwegian stock market as a whole, both in value-weighted and equal-weighted indices, outperform the risk-free interest rate. However, only 43% of individual Norwegian stocks generate lifetime buy-and-hold return which outperform the risk-free interest rate. Furthermore, less than half of Norwegian stocks generate a lifetime buy-and-hold return greater than zero. The excess return generated by the market as a whole, is due to high returns generated by a small group of stocks, which emphasises the importance of positive skewness in the distribution of individual returns.

We find that the performance of the smallest stocks (in terms of market capitalization), generate the highest mean returns in monthly, annual and decade horizon. These returns support the signs of a small-firm effect. Furthermore, as we progress through the timeline, we find that newly listed stocks tend to generate lower annualized returns.

The wealth creation in the Norwegian stock market is, to a considerable extent, carried by a few companies. 2% of the top value creating companies, represent the net gain for the entire Norwegian stock market. The remaining 98% generate wealth equal to the risk-free rate. The top three Norwegian stocks; Equinor, DNB Bank and Aker BP represents half of net cumulative wealth creation throughout our timeline, i.e. January 1980 to March 2024.

Through bootstrapping, we demonstrate the effect of diversification by increasing the number of stocks in a portfolio, and the time horizon. The time horizon positively affects mean buy-and-hold returns and skewness, while the number of stocks positively affects the portfolio performance, compared to the risk-free interest rate. Single-stock portfolios, with a lifetime horizon, only outperforms the risk-free interest rate 20.45% of the time, whilst portfolios with forty randomly chosen stocks each month, with the same time horizon, outperforms the risk-free interest rate 76.77% of the time.

References

Banz, R.W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics* 9 (1), pp. 3–18. DOI: https://doi.org/10.1016/0304-405X(81)90018-0.

Bessembinder, H. (2018). Do stocks outperform Treasury bills? *Journal of Financial Economics* 129 (3), pp. 440–457. DOI: https://doi.org/10.1016/j.jfineco.2018.06.004.

Booth, J.R. and Smith, R.L. (1987). An examination of the small-firm effect on the basis of Skewness preference. *Journal of Financial Researchs* 10 (1), pp. 77–86. DOI: https://doi.org/10.1111/j.1475-6803.1987.tb00477.x.

Bowman, C. and Ambrosini, V. (2000). Value creation versus value capture: Towards a coherent definition of value in strategy. *British Journal of Management* 11 (1), pp. 1–15. DOI: https://doi.org/10.1111/1467-8551.00147.

Brynestad, E.I., Got, S.C.W., Oppedal, M.E., Pålsson, V., and Vegsund, H.C. (2021). *Endret spareadferd under pandemien*. Available at: https://www.ssb.no/nasjonalregnskap-og-konjunkturer/finansregnskap/artikler/endret-spareadferd-under-pandemien (Accessed 17 01 2024).

Byrne, B.M. (2013). *Structural Equation Modeling With AMOS*. 2nd ed. New York: Routledge.

D'Agostino, R.B. and Belanger, A. (1990). A Suggestion for Using Powerful and Informative Tests of Normality. *The American Statistician* 44 (4), pp. 316–321. DOI: https://doi.org/10.2307/2684359.

Daly, K. (2008). Financial volatility: Issues and measuring techniques. *Physica A* 387 (11), pp. 2377–2393. DOI: https://doi.org/10.1016/j.physa.2008.01.009.

Elton, E.J. and Gruber, M.J. (1977). Risk Reduction and Portfolio Size: An Analytical Solution. *Journal of Business* 50 (4), pp. 415–437. DOI: http://dx.doi.org/10.1086/295964.

Fama, E.F. and French, K.R. (2004). New Lists: Fundamentals and Survival Rates. *Journal of Financial Economics* 73 (2), pp. 229–269. DOI: https://dx.doi.org/10.2139/ssrn.279302.

Fama, E.F. and French, K.R. (2013). A Five-Factor Asset Pricing Model. *Fama-Miller Working Paper Series*. DOI: https://dx.doi.org/10.2139/ssrn.2287202.

Fink, J., Fink, K.E., Grullon, G., and Weston, J.P. (2010). What Drove the Increase in Idiosyncratic Volatility During the Internet Boom? *Journal of Financial and Quantitative Analysis* 45 (5), pp. 1253–1278. DOI: https://doi.org/10.1017/S0022109010000487.

Fisher, K. (2018). *You don't need perfect timing, just time, to earn big returns in the stock market*. URL: https://eu.usatoday.com/story/money/columnist/2018/09/16/stock-market-2018-avoid-timing-returns/1303307002/ (visited on 05/29/2024).

Gjedrem, S. (1999). *Utfordringer i den økonomiske politikken*. Available at: https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Foredrag-og-taler/1999/1999-01-28/ (Accessed 25 04 2024).

Hair, J.F., Black, W.C., Babin, B.J., and Anderson, R.E. (2010). *Multivariate Data Analysis*. 7th ed. New York: Pearson.

Hansson, B. and Persson, M. (2000). Time Diversification and Estimation Risk. *Financial Analysts Journal* 56 (5), pp. 55–62. DOI: http://dx.doi.org/10.2469/faj.v56.n5.2390.

Kristiansen, H. (2019). *Gir enkeltaksjer avkastning utover risikofri rente?* Available at: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2612719 (Accessed 25 04 2024).

Lee, C.-F., Lee, A.C., and Lee, J. (2010). *Handbook of Quantitative Finance and Risk Management*. 1st ed. New York: Springer.

Lepak, D.P., Smith, K.G., and Taylor, M.S. (2007). Introduction to Special Topic Forum: Value Creation and Value Capture: A Multilevel Perspective. *The Academy of*

Management Review 32 (1), pp. 180–194. DOI: https://doi.org/10.5465/amr.2007.23464011.

Mooney, C.Z. and Duval, R.D. (1993). *Statistical inference using the Bootstrap*. Sage Publications, Inc., pp. 1–40. DOI: https://doi.org/10.4135/9781412983532.

NBIM (2012). *A Survey of the Small-firm Effect*. Available at: https://www.nbim.no/globalassets/documents/dicussion-paper/2012/discussionnote_12-12.pdf (Accessed 06 04 2024).

Norges Bank (2019). *Historical monetary statistics for Norway*. Available at: https://www.norges-bank.no/en/topics/Statistics/Historical-monetary-statistics/ (Accessed 25 04 2024).

Ødegaard, B.A. (2021). *Empirics of the Oslo Stock Exchange. Basic, descriptive, results 1980-2020*. Available at: https://ba-odegaard.no/wps/empirics_ose_basics/empirics_ose_basics_2021_03.pdf (Accessed 25 04 2024).

Olsen, R.A. and Khaki, M. (1998). Risk, Rationality, and Time Diversification. *Financial Analysts Journal* 54 (5), pp. 58–63. DOI: https://doi.org/10.2469/faj.v54.n5.2212.

Oslo Børs (2021). *OSLO RULE BOOK II - ISSUER RULES*. Available at: https://www.euronext.com/sites/default/files/2021-03/Oslo%20Rule%20Book%20II%20-%20Issuer%20Rules%20-%20Effective%20date%201%20March%202021.pdf (Accessed 04 03 2024).

Regjeringen (2023). *Hva staten eier*. Available at: https://www.regjeringen.no/no/tema/naringsliv/statlig-eierskap/selskaper---ny/id2604524/?expand=factbox2607470 (Accessed 04 05 2024).

Siddikee, N. (2018). Effect of daily dividend on arithmetic and logarithmic return. *The Journal of Finance and Data Science* 4 (4), pp. 247–272. DOI: https://doi.org/10.1016/j.jfds.2018.06.001.

Simkowitz, M.A. and Beedles, W.L. (2010). Diversification in a three-moment world. *Journal of Financial and Quantitative Analysis* 13 (5), pp. 927–941. DOI: `https://doi.org/10.2307/2330635`.

SSB (2024). *Renter i banker og kredittforetak*. Available at: `https://www.ssb.no/statbank/table/10701` (Accessed 25 04 2024).

Statman, M. (1987). How many stocks make a diversified portfolio? *The Journal of Financial and Quantitative Analysis* 22 (3), pp. 353–363. DOI: `http://dx.doi.org/10.2307/2330969`.

Swade, A., Nolte, S., Shackleton, M., and Lohre, H. (2023). Why do equally weighted portfolios beat value-weighted ones? *Journal of Portfolio Management* 49 (5), pp. 167–187. DOI: `https://doi.org/10.3905/jpm.2023.1.482`.

Winther, P. and Tallaksen, S.S. (2024). *Kun fire av 62 aktivt forvaltede fond slo indeksfond i 2023*. Available at: `https://www.dn.no/marked/dn-investor/fond/aktive-fond/kun-fire-av-62-aktivt-forvaltede-fond-slo-indeksfond-i-2023/2-1-1580592` (Accessed 13 02 2024).

Wold, G.C. (2022). *Aksjeinvestering for nybegynnere: – Med bare en hundrelapp i måneden, kan man få ballen til rulle*. URL: `https://www.dagsavisen.no/tema/%C3%B8konomi/2022/01/24/aksjeinvestering-for-nybegynnere/` (visited on 05/20/2024).

Yan, H. and Han, L. (2019). Empirical distributions of stock returns: Mixed normal of kernel density? *Physica A* 514 (C), pp. 473–486. DOI: `https://doi.org/10.1016/j.physa.2018.09.080`.

Appendix

Appendix - Step-by-step programming

This appendix shows the step-by-step coding we have used in order to fill the tables and figures in our thesis. We are by no means efficient programmers, - as seen by the code provided - but following these steps will replicate the necessary statistics in order to repeat this study with a different lifespan.

Tools: TITLON, STATA, Excel and Python

1. Download all equity data from TITLON.uit.no from 1980 to current date. Download data on NIBOR 3-month from Norges Bank (1986-2013) and from SSB (2013-2024). Download the data on Interbank Overnight rate (1980-1986) from Norges Bank.

2. (STATA) Transform data from daily to monthly:

```
replace name = upper(name)
gen date2 = date(date, ’YMD’)
format date2 %td
sort isin date2
gen yearmonth = ym(year(date2), month(date2))
bys isin yearmonth (date2) : keep if _n==_N
drop date2
```

3. (STATA) Remove observations with a market capitalization equal to zero or missing:

```
drop if missing(mktcap)
drop if mktcap==0
```

4. (STATA) Remove stocks that only appear once:

```
by isin, sort: gen no\_mnth=\_N
sort no\_mnth
drop if no\_mnth==1
```

5. (STATA) Remove all observations marked as ’ETF’:

```
drop if sector==’ETF’
```

6. Find stocks with abnormal returns above 600% in one month. Compare returns to

other databases. If not verified, remove the stocks.

7. (Python) Calcualte the return, and add observations with geometric mean return for stocks that delist for less than six months. Set return to zero for stocks that delist longer than six months:

```
import pandas as pd
import numpy as np
df = pd.read_csv('merged.csv', sep=';')
df = df.sort_values(['isin', 'yearmonth'])
df['date'] = pd.to_datetime(df['date'], format='%Y-%m-%d')
df['date'] = df['date'].dt.strftime('%d.%m.%Y')
last_yearmonth = {}
def difference(row):
    isin = row['isin']
    yearmonth = row['yearmonth']
    if isin not in last_yearmonth:
        result = 0
    else:
        result = yearmonth - last_yearmonth[isin]
    last_yearmonth[isin] = yearmonth
    return result
df['difference'] = df.apply(difference, axis=1)
df['return']=df.groupby('isin')['adjustedprice'].pct_change()
def fill_missing_returns(df):
    df = df.sort_values(by=['isin', 'yearmonth'])
    def fill_isin_missing(group):
        rows_to_add = []
        for i in range(1, len(group)):
            current_difference = group.iloc[i]['difference']
            if 1 < current_difference <= 6:
                previous_yearmonth = group.iloc[i - 1]['yearmonth']
                previous_date = pd.to_datetime(group.iloc[i - 1]['date'], format='%d ↩
                    .%m.%Y')
                next_return = group.iloc[i]['return']
                interpolated_return = (1 + next_return) ** (1 / current_difference)  ↩
                    - 1
                last_mktcap = group.iloc[i - 1]['mktcap']
                first_mktcap = group.iloc[i]['mktcap']
                for j in range(1, current_difference):
                    interpolated_yearmonth = previous_yearmonth + j
                    interpolated_mktcap = last_mktcap + (first_mktcap - last_mktcap) ↩
                         / current_difference * j
                    interpolated_date = previous_date + pd.DateOffset(months=j)
                    rows_to_add.append({
                        'isin': group.iloc[i]['isin'],
                        'yearmonth': interpolated_yearmonth,
```

```
                         'return': interpolated_return,
                         'mktcap': interpolated_mktcap,
                         'difference': 1,
                         'companyid': group.iloc[i]['companyid'],
                         'internalcode': group.iloc[i]['internalcode'],
                         'symbol': group.iloc[i]['symbol'],
                         'name': group.iloc[i]['name'],
                         'date': interpolated_date.strftime('%d.%m.%Y')
                     })
                 group.at[group.index[i], 'return'] = interpolated_return
                 group.at[group.index[i], 'difference'] = 1
        if rows_to_add:
            new_rows = pd.DataFrame(rows_to_add)
            group = pd.concat([group, new_rows]).sort_values(by='yearmonth'). ↩
                reset_index(drop=True)
        return group
    df = df.groupby('isin').apply(fill_isin_missing).reset_index(drop=True)
    return df
df = fill_missing_returns(df)
df.loc[(df['difference'] > 6) | (df['difference'] == 0), 'return'] = 0
df.to_csv('titlondata.csv', index=False, sep=';')
```

8. (Excel and Python) Manually add the 'yearmonth' variable to the files containing risk-free interest rate. Merge the files with risk-free interest rate with the dataset containing stock information:

```
import pandas as pd
import numpy as np
from functools import reduce
stockreturn_df = pd.read_csv('titlondata.csv', sep=';')
nibor86_df = pd.read_csv('nibor86.csv', sep=';')
nibor13_df = pd.read_csv('nibor13.csv', sep=';')
ON80_df = pd.read_csv('ON80.csv', sep=';')
dataframes = [stockreturn_df, nibor86_df, nibor13_df, ON80_df]
merged_df = reduce(lambda left, right: pd.merge(left, right, on='yearmonth'), ↩
    dataframes)
merged_df.to_csv('merged.csv', index=False, sep=';')
```

9. (Python) Calculate necessary statistics in order to be able to group the data in deciles based on market capitalization:

```
import pandas as pd
import numpy as np
import datetime as dt
df = pd.read_csv('merged.csv', sep=';')
df['date'] = pd.to_datetime(df['date'], format='%Y-%m-%d')
```

```
df['date'] = df['date'].dt.strftime('%d.%m.%Y')
df['date'] = pd.to_datetime(df['date'], format='%d.%m.%Y')
df = df.sort_values(['isin', 'yearmonth'])
df['year'] = df['date'].dt.year
df['decade'] = (df['year'] // 10) * 10
df['last_mktcap_year'] = df.groupby(['isin', 'year'])['mktcap'].transform('last')
df['last_mktcap_decade'] = df.groupby(['isin', 'decade'])['mktcap'].transform('last' ↩
    )
last_indices_year = df.groupby(['isin', 'year'])['date'].idxmax()
last_indices_decade = df.groupby(['isin', 'decade'])['date'].idxmax()
df['prevyr'] = pd.NA
df['prevdec'] = pd.NA
for idx in last_indices_year:
    next_year_row = df[(df['isin'] == df.loc[idx, 'isin']) & (df['year'] == df.loc[ ↩
        idx, 'year'] + 1) & (df['date'] == df[(df['isin'] == df.loc[idx, 'isin']) & ↩
        (df['year'] == df.loc[idx, 'year'] + 1)]['date'].max())]
    if not next_year_row.empty:
        df.loc[next_year_row.index, 'prevyr'] = df.loc[idx, 'last_mktcap_year']
for idx in last_indices_decade:
    next_decade_row = df[(df['isin'] == df.loc[idx, 'isin']) & (df['decade'] == df. ↩
        loc[idx, 'decade'] + 10) & (df['date'] == df[(df['isin'] == df.loc[idx, ' ↩
        isin']) & (df['decade'] == df.loc[idx, 'decade'] + 10)]['date'].max())]
    if not next_decade_row.empty:
        df.loc[next_decade_row.index, 'prevdec'] = df.loc[idx, 'last_mktcap_decade']
df.drop(['last_mktcap_year', 'last_mktcap_decade'], axis=1, inplace=True)
df.to_csv('predecile.csv', index=False, sep=';')
```

10. (STATA) Create a column which groups the dataset in deciles based on market capitalization (Also possible in Python, however we were not able to get the correct grouping):

```
ssc install egenmore
gen prevmnth=.
replace prevmnth=mktcap[_n-1] if yearmonth>yearmonth[_n-1] & isin==isin[_n-1]
egen monthlydecile=xtile=(prevmnth),by(yearmonth) nq(10)
egen yearlydecile=xtile=(prevyr),by(year) nq(10)
egen decadedecile=xtile=(prevdec),by(decade) nq(10)
```

11. (Python) Calculate necessary statistics in order to fill Tables 5.1.1, 5.2.1, 5.3.1, and 5.4.1.

```
import pandas as pd
import numpy as np
import datetime as dt
df = pd.read_csv('merged.csv', sep=';')
df['date'] = pd.to_datetime(df['date'], format='%Y-%m-%d')
```

```
df['date'] = df['date'].dt.strftime('%d.%m.%Y')
df['date'] = pd.to_datetime(df['date'], format='%d.%m.%Y')
df = df.sort_values(['isin', 'yearmonth'])
first_yearmonth_per_isin = df.groupby('isin')['yearmonth'].first()
last_yearmonth_per_isin = df.groupby('isin')['yearmonth'].last()
months_alive_per_isin = (last_yearmonth_per_isin - first_yearmonth_per_isin) + 1
df['no_mnth'] = df['isin'].map(months_alive_per_isin)
df['totmktcap'] = df.groupby('yearmonth')['mktcap'].transform('sum')
df['valueweight'] = df['mktcap'] / df['totmktcap']
df['vw_return'] = df['return'] * df['valueweight']
df['unik'] = df.groupby('yearmonth')['yearmonth'].transform('count')
df['null'] = (df['difference'] != 0) & (df['difference'] <= 6)
df['null'] = df.groupby(df['yearmonth'])['null'].transform('sum')
df['lv_return'] = df['return'] / (df['null'])
df['ew'] = df.groupby('yearmonth')['lv_return'].transform('sum')
df['vw'] = df.groupby('yearmonth')['vw_return'].transform('sum')
df['positive']=(df['return']>0).astype(int)
df['biggervw']=(df['return']>df['vw']).astype(int)
df['biggerint']=(df['return']>df['interest']).astype(int)
def calculate_cumulative_returns(df, column_name, grouping_cols):
    df[f'cumulative_{column_name}'] = (1 + df[column_name]).groupby(df[grouping_cols ↩
        [0]]).cumprod() - 1
    df[f'yearly_cumulative_{column_name}'] = (1 + df[column_name]).groupby([df[ ↩
        grouping_cols[0]], df['year']]).cumprod() - 1
    last_row_year = df.groupby(grouping_cols + ['year']).tail(1).index
    df.loc[last_row_year, f'yearly_last_cumulative_{column_name}'] = df.groupby( ↩
        grouping_cols + ['year'])[f'yearly_cumulative_{column_name}'].transform(' ↩
        last')
    df[f'decade_cumulative_{column_name}'] = (1 + df[column_name]).groupby([df[ ↩
        grouping_cols[0]], df['decade']]).cumprod() - 1
    last_row_decade = df.groupby(grouping_cols + ['decade']).tail(1).index
    df.loc[last_row_decade, f'decade_last_cumulative_{column_name}'] = df.groupby( ↩
        grouping_cols + ['decade'])[f'decade_cumulative_{column_name}'].transform(' ↩
        last')
    last_row = df.groupby(grouping_cols[0]).tail(1).index
    df.loc[last_row, f'lifetime_last_cumulative_{column_name}'] = df.groupby( ↩
        grouping_cols[0])[f'cumulative_{column_name}'].transform('last')
calculate_cumulative_returns(df, 'return', ['isin'])
calculate_cumulative_returns(df, 'vw', ['isin'])
calculate_cumulative_returns(df, 'ew', ['isin'])
calculate_cumulative_returns(df, 'interest', ['isin'])
df['last_obs'] = df.groupby('isin')['yearmonth'].transform('last')
df['listing_status'] = 'delisted'
df.loc[df['last_obs'] == df['yearmonth'].max(), 'listing_status'] = 'listed'
current_compounded_interest = 0
for yearmonth, group in df.groupby('yearmonth'):
    first_monthly_interest = group['interest'].iloc[0]+1
```

```
    if current_compounded_interest == 0:
        current_compounded_interest = first_monthly_interest
    else:
        current_compounded_interest *= first_monthly_interest
    df.loc[group.index, 'total_comp_monthly_interest'] = current_compounded_interest ↩
        - 1
current_compounded_ew = 0
for yearmonth, group in df.groupby('yearmonth'):
    first_monthly_ew = group['ew'].iloc[0]+1
    if current_compounded_ew == 0:
        current_compounded_ew = first_monthly_ew
    else:
        current_compounded_ew *= first_monthly_ew
    df.loc[group.index, 'total_comp_monthly_ewmarket'] = current_compounded_ew - 1
current_compounded_vw = 0
for yearmonth, group in df.groupby('yearmonth'):
    first_monthly_vw = group['vw'].iloc[0]+1
    if current_compounded_vw == 0:
        current_compounded_vw = first_monthly_vw
    else:
        current_compounded_vw *= first_monthly_vw
    df.loc[group.index, 'total_comp_monthly_vwmarket'] = current_compounded_vw - 1
def create_comparison_column(df, col1, col2, new_column_name):
    mask = df[col1].notna() & df[col2].notna()
    df[new_column_name] = (df.loc[mask, col1] > df.loc[mask, col2]).mean()
create_comparison_column(df, 'return', 'ew', 'monthly_vs_ew')
create_comparison_column(df, 'return', 'vw', 'monthly_vs_vw')
create_comparison_column(df, 'return', 'interest', 'monthly_vs_interest')
create_comparison_column(df, 'yearly_last_cumulative_return', ' ↩
    yearly_last_cumulative_ew', 'yearly_vs_ew')
create_comparison_column(df, 'yearly_last_cumulative_return', ' ↩
    yearly_last_cumulative_vw', 'yearly_vs_vw')
create_comparison_column(df, 'yearly_last_cumulative_return', ' ↩
    yearly_last_cumulative_interest', 'yearly_vs_interest')
create_comparison_column(df, 'decade_last_cumulative_return', ' ↩
    decade_last_cumulative_ew', 'decade_vs_ew')
create_comparison_column(df, 'decade_last_cumulative_return', ' ↩
    decade_last_cumulative_vw', 'decade_vs_vw')
create_comparison_column(df, 'decade_last_cumulative_return', ' ↩
    decade_last_cumulative_interest', 'decade_vs_interest')
create_comparison_column(df, 'lifetime_last_cumulative_return', ' ↩
    lifetime_last_cumulative_ew', 'life_vs_ew')
create_comparison_column(df, 'lifetime_last_cumulative_return', ' ↩
    lifetime_last_cumulative_vw', 'life_vs_vw')
create_comparison_column(df, 'lifetime_last_cumulative_return', ' ↩
    lifetime_last_cumulative_interest', 'life_vs_interest')
def calculate_cumulative_sum(df, column_name, grouping_cols):
```

```
    df[f'sum_{column_name}'] = df.groupby(grouping_cols[0])[column_name].cumsum()
    df[f'yearly_sum_{column_name}'] = df.groupby([grouping_cols[0], df['year']])[column_name].cumsum()
    last_row_year = df.groupby(grouping_cols + ['year']).tail(1).index
    df.loc[last_row_year, f'yearly_last_sum_{column_name}'] = df.groupby(grouping_cols + ['year'])[f'yearly_sum_{column_name}'].transform('last')
    df[f'decade_sum_{column_name}'] = df.groupby([grouping_cols[0], df['decade']])[column_name].cumsum()
    last_row_decade = df.groupby(grouping_cols + ['decade']).tail(1).index
    df.loc[last_row_decade, f'decade_last_sum_{column_name}'] = df.groupby(grouping_cols + ['decade'])[f'decade_sum_{column_name}'].transform('last')
    last_row = df.groupby(grouping_cols[0]).tail(1).index
    df.loc[last_row, f'lifetime_last_sum_{column_name}'] = df.groupby(grouping_cols[0])[f'sum_{column_name}'].transform('last')
months_in_year = df.groupby(['isin', 'year'])['yearmonth'].transform(lambda x: x.max() - x.min() + 1)
months_in_decade = df.groupby(['isin', 'decade'])['yearmonth'].transform(lambda x: x.max() - x.min() + 1)
months_in_life=df.groupby(['isin'])['yearmonth'].transform(lambda x: x.max() - x.min() + 1)
last_row_year = df.groupby(['isin','year']).tail(1).index
last_row_decade = df.groupby(['isin','decade']).tail(1).index
last_row_life = df.groupby(['isin']).tail(1).index
df.loc[last_row_year, 'yearly_geometric'] = ((1 + df.groupby(['isin', 'year'])['yearly_last_cumulative_return'].transform('last')) ** (1 / months_in_year)) - 1
df.loc[last_row_decade, 'decade_geometric'] = ((1 + df.groupby(['isin', 'decade'])['decade_last_cumulative_return'].transform('last')) ** (1 / months_in_decade)) - 1
df.loc[last_row_life, 'life_geometric'] = ((1 + df.groupby(['isin'])['lifetime_last_cumulative_return'].transform('last')) ** (1 / months_in_life)) - 1
calculate_cumulative_sum(df, 'return', ['isin'])
def create_positivity_mean_columns(df, col, new_column_name):
    mask = df[col].notna()
    df[new_column_name] = (df.loc[mask, col] > 0).mean()
create_positivity_mean_columns(df, 'return', 'no_pos_mnth_return')
create_positivity_mean_columns(df, 'yearly_last_cumulative_return', 'no_pos_year_return')
create_positivity_mean_columns(df, 'decade_last_cumulative_return', 'no_pos_decade_return')
create_positivity_mean_columns(df, 'lifetime_last_cumulative_return', 'no_pos_life_return')
create_positivity_mean_columns(df, 'interest', 'no_pos_mnth_interest')
create_positivity_mean_columns(df, 'yearly_last_cumulative_interest', 'no_pos_year_interest')
create_positivity_mean_columns(df, 'decade_last_cumulative_interest', 'no_pos_decade_interest')
```

```
create_positivity_mean_columns(df, 'lifetime_last_cumulative_interest', '↩
    no_pos_life_interest')
create_positivity_mean_columns(df, 'yearly_last_sum_return', 'no_pos_year_sum')
create_positivity_mean_columns(df, 'decade_last_sum_return', 'no_pos_decade_sum')
create_positivity_mean_columns(df, 'lifetime_last_sum_return', 'no_pos_life_sum')
create_positivity_mean_columns(df, 'yearly_geometric', 'no_pos_year_geo')
create_positivity_mean_columns(df, 'decade_geometric', 'no_pos_decade_geo')
create_positivity_mean_columns(df, 'life_geometric', 'no_pos_life_geo')
def fromstatus(df, col, new_column_name, status):
    status_mask = df['listing_status'] == status
    mask = df[col].notna() & status_mask
    df[new_column_name] = (df.loc[mask, col] > 0).mean()
fromstatus(df, 'lifetime_last_cumulative_return', 'no_pos_life_return_delisted', '↩
    delisted')
fromstatus(df, 'lifetime_last_cumulative_return', 'no_pos_life_return_listed', '↩
    listed')
fromstatus(df, 'lifetime_last_cumulative_interest', 'no_pos_life_interest_delisted', ↩
     'delisted')
fromstatus(df, 'lifetime_last_cumulative_interest', 'no_pos_life_interest_listed', '↩
    listed')
fromstatus(df, 'lifetime_last_sum_return', 'no_pos_life_sum_delisted', 'delisted')
fromstatus(df, 'lifetime_last_sum_return', 'no_pos_life_sum_listed', 'listed')
fromstatus(df, 'life_geometric', 'no_pos_life_geo_delisted', 'delisted')
fromstatus(df, 'life_geometric', 'no_pos_life_geo_listed', 'listed')
def comparestatus(df, col1, col2, new_column_name, status):
    status_mask = df['listing_status'] == status
    mask = df[col1].notna() & df[col2].notna() & status_mask
    df[new_column_name] = (df.loc[mask, col1] > df.loc[mask, col2]).mean()
comparestatus(df, 'lifetime_last_cumulative_return', 'lifetime_last_cumulative_ew',  ↩
    'life_vs_ew_listed','listed')
comparestatus(df, 'lifetime_last_cumulative_return', 'lifetime_last_cumulative_vw',  ↩
    'life_vs_vw_listed', 'listed')
comparestatus(df, 'lifetime_last_cumulative_return', '↩
    lifetime_last_cumulative_interest', 'life_vs_interest_listed', 'listed')
comparestatus(df, 'lifetime_last_cumulative_return', 'lifetime_last_cumulative_ew',  ↩
    'life_vs_ew_delisted', 'delisted')
comparestatus(df, 'lifetime_last_cumulative_return', 'lifetime_last_cumulative_vw',  ↩
    'life_vs_vw_delisted', 'delisted')
comparestatus(df, 'lifetime_last_cumulative_return', '↩
    lifetime_last_cumulative_interest', 'life_vs_interest_delisted', 'delisted')
df['monthlydecile_pos'] = df.groupby('monthlydecile')['return'].transform(lambda x:  ↩
    (x > 0).mean())
df['yearlydecile_pos'] = df.groupby('yearlydecile')['yearly_last_cumulative_return' ↩
    ].transform(lambda x: (x > 0).mean())
df['decadedecile_pos'] = df.groupby('decadedecile')['decade_last_cumulative_return' ↩
    ].transform(lambda x: (x > 0).mean())
df['monthlydecile_vw'] = df['return'] > df['vw']
```

```
df['monthlydecile_vw'] = df.groupby('monthlydecile')['monthlydecile_vw'].transform('
    mean')
df['monthlydecile_int'] = df['return'] > df['interest']
df['monthlydecile_int'] = df.groupby('monthlydecile')['monthlydecile_int'].transform
    ('mean')
df['monthlydecile_ew'] = df['return'] > df['ew']
df['monthlydecile_ew'] = df.groupby('monthlydecile')['monthlydecile_ew'].transform('
    mean')
df['yearlydecile_vw'] = df['yearly_last_cumulative_return'] > df['
    yearly_last_cumulative_vw']
df['yearlydecile_vw'] = df.groupby('yearlydecile')['yearlydecile_vw'].transform('
    mean')
df['yearlydecile_int'] = df['yearly_last_cumulative_return'] > df['
    yearly_last_cumulative_interest']
df['yearlydecile_int'] = df.groupby('yearlydecile')['yearlydecile_int'].transform('
    mean')
df['yearlydecile_ew'] = df['yearly_last_cumulative_return'] > df['
    yearly_last_cumulative_ew']
df['yearlydecile_ew'] = df.groupby('yearlydecile')['yearlydecile_ew'].transform('
    mean')
df['decadedecile_vw'] = df['decade_last_cumulative_return'] > df['
    decade_last_cumulative_vw']
df['decadedecile_vw'] = df.groupby('decadedecile')['decadedecile_vw'].transform('
    mean')
df['decadedecile_int'] = df['decade_last_cumulative_return'] > df['
    decade_last_cumulative_interest']
df['decadedecile_int'] = df.groupby('decadedecile')['decadedecile_int'].transform('
    mean')
df['decadedecile_ew'] = df['decade_last_cumulative_return'] > df['
    decade_last_cumulative_ew']
df['decadedecile_ew'] = df.groupby('decadedecile')['decadedecile_ew'].transform('
    mean')
df.to_csv('withoutwealth.csv', index=False, sep=';')
```

12. (Python) Bootstrap 20,000 different value-weighted portfolio returns per 'year-month' in the dataset (set 'n=1' to the desired portfolio size):

```
import pandas as pd
import numpy as np
from datetime import datetime
df = pd.read_csv('withoutwealth.csv', sep=';')
df = df.sort_values(by='yearmonth')
num_simulations = 20000
simulated_returns = []
unique_months = df['yearmonth'].unique()
for sim_num in range(1, num_simulations + 1):
    selected_stocks = df.groupby('yearmonth').sample(n=1, replace=True)
```

```
    total_market_value = selected_stocks.groupby('yearmonth')['mktcap'].transform('sum')
    selected_stocks['weight'] = selected_stocks['mktcap'] / total_market_value
    selected_stocks['weighted_return'] = selected_stocks['return'] * selected_stocks['weight']
    monthly_returns = selected_stocks.groupby('yearmonth')['weighted_return'].sum().tolist()
    simulated_returns.append(monthly_returns)
    if sim_num % 10 == 0:
        print(f'Simulation {sim_num} done at {datetime.now()}.')
simulated_returns = np.array(simulated_returns).reshape(-1, len(unique_months)).T
simulated_returns_df = pd.DataFrame(simulated_returns)
simulated_returns_df.columns = [f'sim{sim_num}' for sim_num in range(1, num_simulations + 1)]
simulated_returns_df.to_csv('1_stock_bootstrap.csv', index=False, sep=';')
```

13. (Python) Link the bootstrap returns to annual, decade and lifetime buy-and-hold returns and compare them to the benchmarks in order to fill Table 5.5.1:

Annual:

```
import pandas as pd
import numpy as np
df = pd.read_csv('1_stock_bootstrap.csv', sep=';')
return_columns = [col for col in df.columns if col.startswith('vwmean')]
for col in return_columns:
    df[col] = df.groupby('year')[col].transform(lambda x: (1 + x).cumprod() - 1)
df['interest'] = df.groupby('year')['interest'].transform(lambda x: (1 + x).cumprod() - 1)
df['vw'] = df.groupby('year')['vw'].transform(lambda x: (1 + x).cumprod() - 1)
df = df.groupby('year').tail(1)
def compare_interest(row):
    vwmean_cols = [row[col] for col in return_columns]
    interest_val = row['interest']
    proportion_greater = sum([1 for val in vwmean_cols if val > interest_val]) / len(vwmean_cols)
    return proportion_greater
def compare_vw(row):
    vwmean_cols = [row[col] for col in return_columns]
    vw_val = row['vw']
    proportion_greater = sum([1 for val in vwmean_cols if val > vw_val]) / len(vwmean_cols)
    return proportion_greater
def compare_zero(row):
    vwmean_cols = [row[col] for col in return_columns]
    zero_val = 0
    proportion_greater = sum([1 for val in vwmean_cols if val > zero_val]) / len(
```

```
        vwmean_cols)
    return proportion_greater
df['retvs_int'] = df.apply(compare_interest, axis=1)
df['retvs_vw'] = df.apply(compare_vw, axis=1)
df['retvs_0'] = df.apply(compare_zero, axis=1)
df['retvs_int']=df['retvs_int'].mean()
df['retvs_vw']=df['retvs_vw'].mean()
df['retvs_0']=df['retvs_0'].mean()
for col in return_columns:
    df[col]=df[col].mean()
df=df.tail(1)
df['avgret']=df[return_columns].mean(axis=1)
df=df.transpose()
df.to_csv('1_stock_bootstrapyr.csv', index=False, sep=';')
```

Decade:

```
import pandas as pd
import numpy as np
df = pd.read_csv('1aksjeboot.csv', sep=';')
return_columns = [col for col in df.columns if col.startswith('vwmean')]
for col in return_columns:
    df[col] = df.groupby('decade')[col].transform(lambda x: (1 + x).cumprod() - 1)
df['interest'] = df.groupby('decade')['interest'].transform(lambda x: (1 + x). ↩
    cumprod() - 1)
df['vw'] = df.groupby('decade')['vw'].transform(lambda x: (1 + x).cumprod() - 1)
df = df.groupby('decade').tail(1)
def compare_interest(row):
    vwmean_cols = [row[col] for col in return_columns]
    interest_val = row['interest']
    proportion_greater = sum([1 for val in vwmean_cols if val > interest_val]) / len ↩
        (vwmean_cols)
    return proportion_greater
def compare_vw(row):
    vwmean_cols = [row[col] for col in return_columns]
    vw_val = row['vw']
    proportion_greater = sum([1 for val in vwmean_cols if val > vw_val]) / len( ↩
        vwmean_cols)
    return proportion_greater
def compare_zero(row):
    vwmean_cols = [row[col] for col in return_columns]
    zero_val = 0
    proportion_greater = sum([1 for val in vwmean_cols if val > zero_val]) / len( ↩
        vwmean_cols)
    return proportion_greater
df['retvs_int'] = df.apply(compare_interest, axis=1)
df['retvs_vw'] = df.apply(compare_vw, axis=1)
```

```
df['retvs_0'] = df.apply(compare_zero, axis=1)
df['retvs_int']=df['retvs_int'].mean()
df['retvs_vw']=df['retvs_vw'].mean()
df['retvs_0']=df['retvs_0'].mean()
for col in return_columns:
    df[col]=df[col].mean()
df=df.tail(1)
df['avgret']=df[return_columns].mean(axis=1)
df=df.transpose()
df.to_csv('1_stock_bootstrapdec.csv', index=False, sep=';')
```

Lifetime:

```
import pandas as pd
import numpy as np
df = pd.read_csv('1aksjeboot.csv', sep=';')
return_columns = [col for col in df.columns if col.startswith('vwmean')]
for col in return_columns:
    df[col] = df[col].transform(lambda x: (1 + x).cumprod() - 1)
df['interest'] = df['interest'].transform(lambda x: (1 + x).cumprod() - 1)
df['vw'] = df['vw'].transform(lambda x: (1 + x).cumprod() - 1)
df=df.tail(1)
def compare_interest(row):
    vwmean_cols = [row[col] for col in return_columns]
    interest_val = row['interest']
    proportion_greater = sum([1 for val in vwmean_cols if val > interest_val]) / len ↩
        (vwmean_cols)
    return proportion_greater
def compare_vw(row):
    vwmean_cols = [row[col] for col in return_columns]
    vw_val = row['vw']
    proportion_greater = sum([1 for val in vwmean_cols if val > vw_val]) / len( ↩
        vwmean_cols)
    return proportion_greater
def compare_zero(row):
    vwmean_cols = [row[col] for col in return_columns]
    zero_val = 0
    proportion_greater = sum([1 for val in vwmean_cols if val > zero_val]) / len( ↩
        vwmean_cols)
    return proportion_greater
df['retvs_int'] = df.apply(compare_interest, axis=1)
df['retvs_vw'] = df.apply(compare_vw, axis=1)
df['retvs_0'] = df.apply(compare_zero, axis=1)
df['retvs_int']=df['retvs_int'].mean()
df['retvs_vw']=df['retvs_vw'].mean()
df['retvs_0']=df['retvs_0'].mean()
df['avgret']=df[return_columns].mean(axis=1)
```

```
df=df.transpose()
df.to_csv('1_stock_bootstraplife.csv', index=False, sep=';')
```

14. (Python) Link the company-id so that each company has its own unique id (not the same as ISIN). Calculate the wealth creation based on the company-id, and the annualized return based on the company-id's ISIN with the longest lifespan:

```
import pandas as pd
import numpy as np
import datetime as dt
df = pd.read_csv('withoutwealth.csv', sep=';')
df['date'] = pd.to_datetime(df['date'], format='%Y-%m-%d')
df['date'] = df['date'].dt.strftime('%d.%m.%Y')
df['date'] = pd.to_datetime(df['date'], format='%d.%m.%Y')
df = df.sort_values(['isin', 'yearmonth'])
df['companyid']=df['companyid'].fillna(0)
df.loc[df['internalcode'] != 0, 'companyid'] = df[df['internalcode'] != 0].groupby('internalcode')['companyid'].transform(lambda x: next((val for val in x if val != 0), 0))
df['companyid'] = df.groupby('symbol')['companyid'].transform(lambda x: next((val for val in x if pd.notnull(val)), 0))
df['companyid'] = df.groupby('isin')['companyid'].transform(lambda x: next((val for val in x if pd.notnull(val)), 0))
df['companyid'] = df.groupby('name')['companyid'].transform(lambda x: next((val for val in x if pd.notnull(val)), 0))
df.loc[df['companyid']==0, 'companyid']=df['internalcode']
df['new_mktcap'] = df.groupby(['companyid', 'yearmonth'])['mktcap'].transform('sum')
df['decadenr']=df.groupby('isin')['decade'].transform('first')
df['max_isin_by_companyid'] = df.groupby('companyid')['isin'].transform(lambda x: x[df.loc[x.index, 'no_mnth'].idxmax()])
mask = df['isin'] == df['max_isin_by_companyid']
df['chosen_lifetime_return'] = df['lifetime_last_cumulative_return'].where(mask)
df['chosen_lifetime_return'] = df.groupby('companyid')['chosen_lifetime_return'].transform('max')
df['chosen_no_mnth']=df['no_mnth'].where(mask)
df['chosen_no_mnth'] = df.groupby('companyid')['chosen_no_mnth'].transform('max')
df['annualized_chosen_return'] = (1+df['chosen_lifetime_return']) ** (12 / (df['chosen_no_mnth'])) - 1
is_last = df.groupby(['companyid', 'yearmonth'])['mktcap'].cumcount() == df.groupby(['companyid', 'yearmonth'])['mktcap'].transform('count') - 1
df = df[is_last]
last_row_comp=df.groupby(['companyid']).tail(1).index
df=df.sort_values(['companyid','yearmonth'])
df['cumcompint'] = (1 + df['interest']).groupby(df['companyid']).cumprod() - 1
df.loc[last_row_comp, 'wealth']=(df.groupby(['companyid'])['new_mktcap'].transform('last'))-((df.groupby(['companyid'])['new_mktcap'].transform('first'))*(1+df.
```

```
    groupby(['companyid'])['cumcompint'].transform('last')))
df=df.sort_values('yearmonth')
df['totalwealth']=df['wealth'].sum()
df['perctotal']=df['wealth']/df['totalwealth']
df=df.sort_values('wealth', ascending=False)
df['cumwealth']=df['perctotal'].cumsum()
df['first_month'] = df.groupby('max_isin_by_companyid')['date'].transform('min').dt.
    to_period('M').dt.strftime('%Y-%m')
df['last_month'] = df.groupby('max_isin_by_companyid')['date'].transform('max').dt.
    to_period('M').dt.strftime('%Y-%m')
df.to_csv('withwealth.csv', index=False, sep=';')
```

15. (STATA) Lastly, summarize all the different returns based on desired conditions for example 'decadenr' which provides the statistics based on initial decade of appearance.

NTNU

Norwegian University of
Science and Technology