Norwegian School of Economics
Bergen, Fall 2020

The Hidden Costs of Passive Investing

An Empirical Study on the Impact of Passive Investing on the Liquidity and Price Efficiency of Norwegian Stocks

Oskar Petter Førde and Henrik Borge Sundby

Supervisor: Thore Johnsen

Master thesis, Economics and Business Administration

Major: Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis marks the culmination of our 5 years of studies at the Norwegian School of Economics. In this regard, we would like to thank our families for their support, and fellow students and friends for making our time in Bergen memorable in so many ways. We also want to thank NHH, and especially the Department of Finance, for a truly inspiring programme of study.

Throughout the process of working on the thesis, we have received invaluable guidance and support from people worthy of acknowledgement. First and foremost we want to thank our supervisor, Thore Johnsen. Without your encouragement and constructive feedback, we would not have been able to write this thesis. We also extend our gratitude to the people at Morningstar for providing essential data, and to Joakim A. Embu (KLP) and Peter Hermanrud (SpareBank 1 Markets) for valuable discussions in the initial phases of our project.

Norwegian School of Economics

Bergen, December 2020

Oskar P. Førde

Oskar Petter Førde

Henrik B Sundby

Henrik Borge Sundby

Abstract

In this thesis, we examine the relationship between passive ownership in the Norwegian stock market and the liquidity and price efficiency of the underlying stocks. We suggest that a shift from active to passive investing should be associated with a decrease in the liquidity of individual stocks, as more shares are held in long-term deposits and more trading is executed on a non-fundamental basis. The same shift, in addition to a reduction in liquidity and an increase in the trading of stocks in large baskets, should all lead to a decrease in the price efficiency of the underlying stocks.

In our analysis, we utilise a panel of 214 publicly listed Norwegian stocks in the period from 2000 to 2020, with monthly ownership and daily financial market data on the stock-level. We find a negative relationship between changes in the passive ownership of a stock and contemporaneous changes in its liquidity. The effect is smaller for larger stocks, which could explain why we find a larger effect than the existing literature using our sample of relatively illiquid Norwegian stocks. Further, we document a negative relationship between changes in passive ownership and changes in price efficiency, as measured by return synchronicity. We find that the effect is smaller for larger firms. Taking into account our sample of relatively small firms, the effect is still small relative to the existing literature. This could be explained by the frequent trading of stocks in baskets by ETFs, which is less prevalent in the Norwegian market.

Keywords – Passive Investing, Liquidity, Price Efficiency, Asset Management

Contents

List of Figures

List of Tables

1 Introduction

During the last few months, the debate on active versus passive asset management has again raged in the Norwegian financial press. On the one hand, there is little doubt that the emergence of the passive investment style has improved the welfare of many individuals, by allowing for inexpensive access to what has been advocated as more or less optimal portfolios. However, commentators often fail to take into account the role of the active investor as a facilitator for the passive. Specifically, through their own self-interest, active investors provide liquidity at what they perceive as reasonable price levels, thus affecting both the liquidity and the price efficiency in the market. In our thesis, we examine the impact of the shift from active to passive on these two elements of the market microstructure in Norway.

In 2000, index funds and passive ETFs held about 5% of the U.S. equity market. Today, they hold 19%. In Norway, the numbers are 0.5% and 5%, respectively. During our 21.5 year sample, U.S. stocks have experienced an average net flow from passive funds of about $440 million every single trading day. Norwegian stocks, on the other hand, has received net flows from passive funds of $2.4 million a day[1] during the same period. Since the 2008 financial crisis — when passive investing in Norway really started to escalate — average net flows have been $3.5 million a day.

When examining the possible impacts of this massive increase, the existing literature mainly focuses on three traits of indexers: (1) the passive holdings of shares over long periods of time, (2) the absence of fundamental analysis and (3) the trading of stocks in large baskets.

Prior literature shows that these traits could influence several aspects of the stock market. More specifically, Israeli et al. (2017) show that there is a positive relationship between ETF ownership and trading costs. Due to the migration of uninformed investors to ETFs, they also find a negative relationship between ETF ownership and price efficiency. These findings correspond well with those of Hamm (2014), Sammon (2020) and Zou (2019). Further, Ben-David et al. (2018) document a positive relationship between the underlying

[1]None of the figures for the Norwegian equity market include The Government Pension Fund Norway (Folketrygdfondet), which we will discuss in detail later.

stock's ETF ownership and volatility, due to a strong arbitrage channel between ETFs and their underlying assets. Da and Shive (2018) claim that increases in ETF ownership lead to an increase in return comovement in stock markets. This relationship is also emphasised by Sullivan and Xiong (2012), who also document an increase in non-diversifiable risk from increases in passive investing.

We examine the impact of passive strategies on the liquidity and price efficiency of Norwegian stocks. We posit that the long-term passive holdings of stocks lead to lower liquidity, as more shares are kept away from trading. This could again negatively affect price informativeness, as higher trading costs induce less trading by informed investors. In addition to this, an increase in the trading of stocks in large baskets could also negatively affect the firm-specific information embedded in prices.

We approach the issue empirically. We focus on the increase in passive investing during the last two decades and use ownership data on the stock-level to examine the impacts of passive inflows on various liquidity and price efficiency metrics. We use the effects discovered in the U.S. market as guidelines for our analyses. The result is a holistic review of the mechanisms through which passive investing might have affected price efficiency in the Norwegian market, partly through its influence on liquidity. This market is different from the American, in that passive investing is both less prevalent and more comprised of traditional buy-and-hold index funds than exchange-traded funds. In addition to this, the average free float share in Norwegian stocks is among the lowest in the western world, and about half of that in the U.S., which could make the liquidity in the market relatively sensitive to increases in long-term passive holdings.

Specifically, our hypotheses are stated as follows:

H1: ***"Does passive investing lead to reduced liquidity in the Norwegian stock market?"***

and

H2: ***"Does passive investing lead to less informative prices in the Norwegian stock market?"***

In examining the first hypothesis, we use a similar approach as Israeli et al. (2017) in applying two proxies for liquidity: the price impact of trading and the relative spread. First, on the stock-level, we analyse the relationship between changes in the passive ownership of a stock and its liquidity through correlational analyses. Second, we use OSEBX inclusions as quasi-exogenous shocks in passive ownership, in order to establish a causal relationship using a difference-in-differences approach. In addressing the hypothesis, we apply the same methodology to four proxies of price efficiency.

We find that annual increases in passive ownership are significantly connected to decreases in liquidity. The interpretation is the same using both proxies and is consistent with our expectations and the findings of Israeli et al. (2017). The estimated relationship is, however, larger in the Norwegian market, which could be a result of the differences in the average level of free float and the differences in the types of passive vehicles analysed. Using a limited set of OSEBX inclusions, we are not able to establish a causal relationship for any of our liquidity proxies.

As for price informativeness, the results are more ambiguous. Using the event-based proxies of Sammon (2020), the relationship between passive ownership and price informativeness is only statistically significant for one out of the three proxies. Nevertheless, using the return synchronicity proxy of Roll (1988) and Durnev et al. (2003), we find a significant and positive relationship, indicating that the firm-specific information component in prices is negatively associated with changes in passive ownership.

We structure the thesis in the following way. In the next chapter, we will provide background information on passive investing, the ways in which it could influence markets and the existing literature on the topic. The third chapter contains a description of the data, variables and methodology used in the analyses. In the fourth chapter, we present and discuss the findings of our analyses. The fifth chapter contains our conclusions, as well as potential limitations with our analyses and suggestions for future research.

2 Theory

In the following chapter, we provide the relevant background for this thesis. We first review a few key traits of passive investing and the case for choosing to invest in such a strategy. Further, we examine the trend in passive investing during the last two decades, before outlining the theoretical background for the link between passive investing and our dependent variables. Lastly, we motivate our hypotheses on the basis of the existing literature on this topic.

2.1 Passive Investing

We classify as passive any strategy that aims to track a predefined market-weighted index of some sort. This includes the strategies of traditional index funds and index ETFs, but not those that are normally considered in the middle of the passive-active continuum, such as smart-beta and factor strategies. We acknowledge that we fail to account for a great deal of undisclosed or privately managed passive investing. In addition to the funds deemed appropriate by the abovementioned criteria, we include The Government Pension Fund Norway ("Folketrygdfondet" or "FTF") as a passive fund in our analyses. We motivate this choice in Section 2.1.3. In the three following sections, we elaborate on the three different types of passive vehicles in our analyses. Figure 2.1 illustrates their distribution in size over the last 20 years.

2.1.1 Index Mutual Funds

The index fund was first introduced by Renshaw and Feldstein (1960). Most mutual funds at the time failed to outperform broad indices (Jensen, 1964), and the authors proposed that instead of attempting to identify high-performing funds, investors should aim for the average return of the stocks in the index. They also pointed out that such a strategy would require little analysis, thus adding economic value for investors through the reduction of fees. The idea was debated throughout the 1960s. In the following decade, Burton Malkiel's work on the random walk and efficient markets hypotheses further fueled the debate, as Malkiel in his book "A Random Walk Down Wall Street" argued that the returns of active funds are in nature mean-reverting, implying that no single investor is able to consistently outperform the market. Consequently, any investor would be better off investing in the passive, index-replicating mutual fund (Malkiel, 1973). Following these arguments, the late John Bogle introduced the first index fund available to retail investors in 1976. Today, this fund is known as the Vanguard 500 Index Fund and is the world's largest mutual fund with a net asset value (NAV) of about $620 billion Vanguard (2020b).

The theoretical idea of the index fund was, and still is, very simple. The fund aims to track the performance of an index, with minimal tracking error. This is achieved by holding a representative basket of the index constituents. In other words, an index fund operating in a frictionless world should always own shares in all the assets that make up the index it tracks. The individual holdings are weighted in the same manner as the index, which is usually by market capitalisation (MCap) with an adjustment factor for free float. Index fund managers only trade when the fund experiences inflows or outflows, when stocks are included or excluded from the index, or in response to SEOs, M&As, buybacks and dividends. The manager pays no attention to the fundamentals of the index constituents and does not intervene in any way to optimize risk or return beyond what is embedded in the index. In theory — as the efficient market hypothesis states that all information about a financial asset is reflected in its price (Fama, 1970) — the fund always transacts at the correct price. Along these lines, one could argue that the index fund "free rides" on the analysis conducted by active investors. This absence of competitive stock-level analysis is what allows index funds to charge low fees, relative to those of active managers (PwC Asset Management, 2017).

In reality, there are a few additional aspects of index investing worth considering. First, there is a trade-off between transaction costs and the number of individual stocks held when minimising tracking error. Index fund managers incur transaction costs when trading, and these costs are larger for less liquid stocks, which often are the same stocks in which the index is underweight. Consequently, the expected tracking error depends on aspects such as the size of the fund, the number of index constituents, and the liquidity of these constituents. This is evident through the strategy of global index funds in the Norwegian market. Through ten constituents, Norway represents about 0.18% of the global equity index[2] (MSCI, 2020). Therefore, most global index funds only hold between two and ten Norwegian stocks, as any deeper diversification is deemed sub-optimal for

[2]MSCI World Index as of 01.09.2020.

such a small part of the overall portfolio. However, as both index technology and scale has improved, transaction costs have decreased, and index investing is today viewed as the most cost-efficient way to gain broad exposure to equity markets. This is one of the reasons why the prevalence of passive investing have increased dramatically over the course of the last 20 years. Today, John Bogle's Vanguard manages about $6.2 trillion of passive funds in total, or about 7% of the global stock market (Vanguard, 2020a).

2.1.2 Exchange-Traded Index Funds

Most of the existing literature on the effects of passive investing is conducted on index ETFs. We treat index ETFs and index mutual funds the same way in our analyses. About 98% of ETFs are in fact passive in the sense that they follow an index with market capitalisation weighting (Zou, 2019). There is not much that separates index ETFs from traditional index funds in terms of investment strategy. Despite this, there is one key difference, which is important in order to understand how ETFs might affect the market for individual stocks and why such a large part of the research on passive investing is focused on ETFs.

Similar to open-ended mutual funds, the number of shares of an ETF varies as the fund experiences inflows or redemptions. A key difference is that holders of an ETF share can trade this share in a secondary market throughout the day. In fact, ETFs are tradeable in the secondary market in the same way as a stock, providing investors with the options to go long or short and to use limit orders and stop-loss orders. Additionally, authorised participants (APs) have the ability to both create and redeem shares of the ETF with the ETF provider at the NAV of the underlying portfolio. This opens up an arbitrage channel, as APs can profit if the ETF price deviates from the NAV. As APs also can trade in the secondary market, the same mechanism will apply here. The result of this arbitrage mechanism is very high liquidity at prices close to the NAV, at any given time of the day. On the contrary, with traditional open-ended mutual funds, investors can only purchase or redeem shares with the fund provider at the end-of-day NAV.

This opens up a new way in which the fund market and the stock market interact. With mutual funds, the belief has been that only the transactions generated to facilitate inflows and outflows could influence the stock market. With ETFs, on the other hand, there is a strong arbitrage channel requiring that movement in ETF prices must be equalled with a similar movement in the underlying stocks, within seconds. The relationship works in both directions. As ETFs have become increasingly popular for trading on broad macroeconomic events, it is evident that volatility in ETFs will, to some extent, propagate to the underlying securities. More specific, to the extent that ETFs attract non-fundamental demand — that is, demand that would disappear had ETFs not existed — ETF ownership will result in non-fundamental volatility for individual stocks due to this arbitrage channel (Ben-David et al., 2018).

In the U.S., ETFs account for 17% of all fund holdings, while the same number in Norway is 8%[3]. The ETF arbitrage channel is the basis for many of the academic contributions we will present in the following, and the prevalence of ETFs could therefore be important when comparing the effects of passive investing in Norway to those in the U.S.

2.1.3 Folketrygdfondet

The Government Pension Fund Norway, or Folketrygdfondet (FTF), is one of two Norwegian state pension funds, and must not be mistaken for the larger Government Pension Fund Global ("The Oil Fund"). FTF has a total assets under management (AUM) of \$13.4[4] billion invested in Norwegian equities but is not passive in the sense that it seeks to minimise tracking error. Instead, FTF's mandate, which is issued by the Norwegian Ministry of Finance, states that the fund shall seek to achieve the highest possible return in the long term. The mandate also states that the annualised volatility of the fund relative to the OSEBX should not exceed three percentage points and that its ownership in any single stock should not exceed 15 per cent (Mandat for Statens pensjonsfond Norge – SPN, 2010). The fund claims to be active in the sense that it seeks to achieve excess returns relative to the index. However, as the fund has grown larger, there seems to be limited room for manoeuvring an active strategy in the Norwegian market, due to transaction costs and the 15 per cent ownership restriction. On the other hand, a strict index strategy would also incur substantial transaction costs every time the index is rebalanced (Johnsen, 2011). Therefore, the fund executes its active strategy through the long-term holding of most index constituents, with a few strategic deviations. This way,

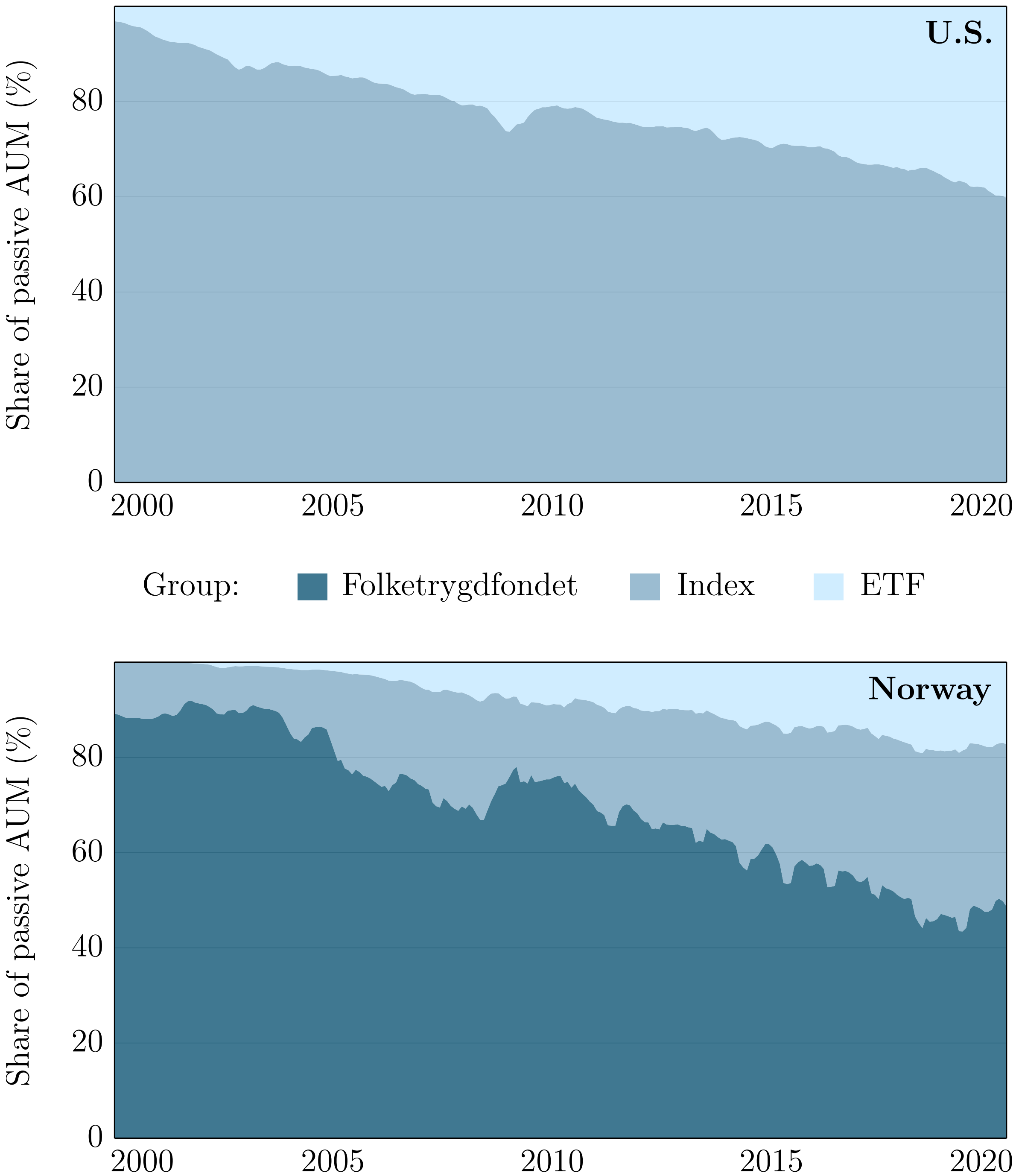
[3]By including Folketrygdfondet, ETFs account for 6.5% of all fund holdings as of June 2020.
[4]Assets under management as of 30.06.2020.

FTF operates similarly to index funds in that they "remove" shares from regular trading when investing in a stock. This is equivalent to a more or less temporary reduction in the free float. The fund's deviations from the index are not crucial for our analysis, but the long-term perspective is. That being said, it is important to emphasise that there are no flows neither in nor out of the FTF. This means that the fund transacts only when reallocating capital between stocks or between asset classes. As we analyse changes in passive ownership, FTF is not necessarily as important for our analyses as Figure 2.1 indicates. Nevertheless, the fund contributes to our data as they have fewer but larger transactions, while index funds and ETFs tend to have more but smaller transactions.

We expect the impact of the FTF to be similar to that of index funds and index ETFs. This the first reason why we include FTF in our analyses. The second reason is the size of the fund. FTF holds about 6% of all listed stocks in Norway, which is about the same as index funds and ETFs combined. This means that by ignoring the fund, we would to some extent overstate the impact of index funds and ETFs.

Figure 2.1: AUM Distribution for U.S. and Norwegian Passive Funds

This chart illustrates the distribution between the main types of passive investment vehicles, based on their share of the total passive ownership in U.S. and Norwegian stocks. The ownership data is aggregated by each individual fund on a monthly basis, based on the Morningstar Direct database and FTF reports.



2.2 The Case for Passive Investing

With the introduction of the passive investment style came also the debate on the right to life for active and passive managers. In this section, we briefly present what we believe to be the most important contributions to this debate.

William Sharpe stated, in his article "The Arithmetic of Active Management", that:

"If active and passive management styles are defined in sensible ways, it must be the case that (1) before costs, the return on the average actively managed dollar will equal the return on the average passively managed dollar and (2) after costs, the return on the average actively managed dollar will be less than the return on the average passively managed dollar." (Sharpe, 1991)

To arrive at this conclusion, Sharpe assumes that all passive investors hold some fraction of the market, with each asset weighted relative to its weight in the market. Active investors are defined as "one who is not passive". Consequently, active investors must in aggregate hold a portfolio identical to that of passive investors. Using simple arithmetics, Sharpe proves the first of his two assertions. Passive investors will, individually and in aggregate, achieve the same return as the market. Active investors must therefore also in aggregate, though not individually, achieve the return of the market. Second, assuming that active managers must charge higher fees, passive managers must achieve a higher return net of fees. Since it was published in 1991, Sharpe's arithmetic has often been utilised as a type of "proof" of why passive management is the sensible choice for most investors.

Some 27 years later, Lasse Pedersen of AQR[5] challenged Sharpe's "Arithmetic", or at least the public interpretation of it. Pedersen outlines a number of scenarios in which he believes that Sharpe (1991) deviates from reality. Sharpe examines a single period where passive managers "start out" with a fraction of the market, hence do not trade at all. Pedersen argues that the absence of trading is unrealistic, as any passive vehicle tracking a market capitalisation-weighted index would have to rebalance its holdings as securities come in and out of the index, for instance in the case of index additions and exclusions, IPOs, M&As, SEOs, buybacks and dividend payouts. In such instances, passive managers would have to initiate trades with an active counterpart, inevitably incurring some form

[5]AQR Capital is an active asset management firm.

of transaction costs. The perhaps most important flaw of Sharpe's theory is the idea of investing in "the market". In reality, the market is an index, for instance the S&P 500 consisting of some 500 stocks. Sharpe's arithmetic would not hold if active managers were allowed to trade stocks outside the index, which in most cases, they are. Pedersen argues that based on these mechanisms, active managers could very well outperform passive managers over time. (Pedersen, 2018)

A different argument for indexed or «basket» investing relates to the positive skew of individual stock returns. While this skew has been recognised in several academic contributions (Albuquerque, 2012; Bessembinder, 2018; Fama and French, 2018), Bessembinder et al. (2019) examine the practical implications for equity investors. They find that only 3.8% of U.S. stocks outperformed the S&P 500 during the course of their lifetime. Strikingly, the top-performing 1.3% of stocks accounted for the entire global stock market's wealth creation in the period from 1990 to 2018. In order for an investor to capture the positive return of the aggregate market, it then becomes obvious that he/she must own a few of these top performers. However, assuming no comparative advantage over other investors, the investor can not identify these in advance. This ultimately means that the only way to achieve this is through an index strategy. This is because a randomly (assuming no competitive advantage) selected portfolio from the index will, in most cases, underperform the index itself, when the median stock's return is lower than that of the index. In their master's thesis, Norang and Agustsson (2018) use a 1985-2017 sample and concludes that the same highly positive skew is present in the Norwegian stock market.

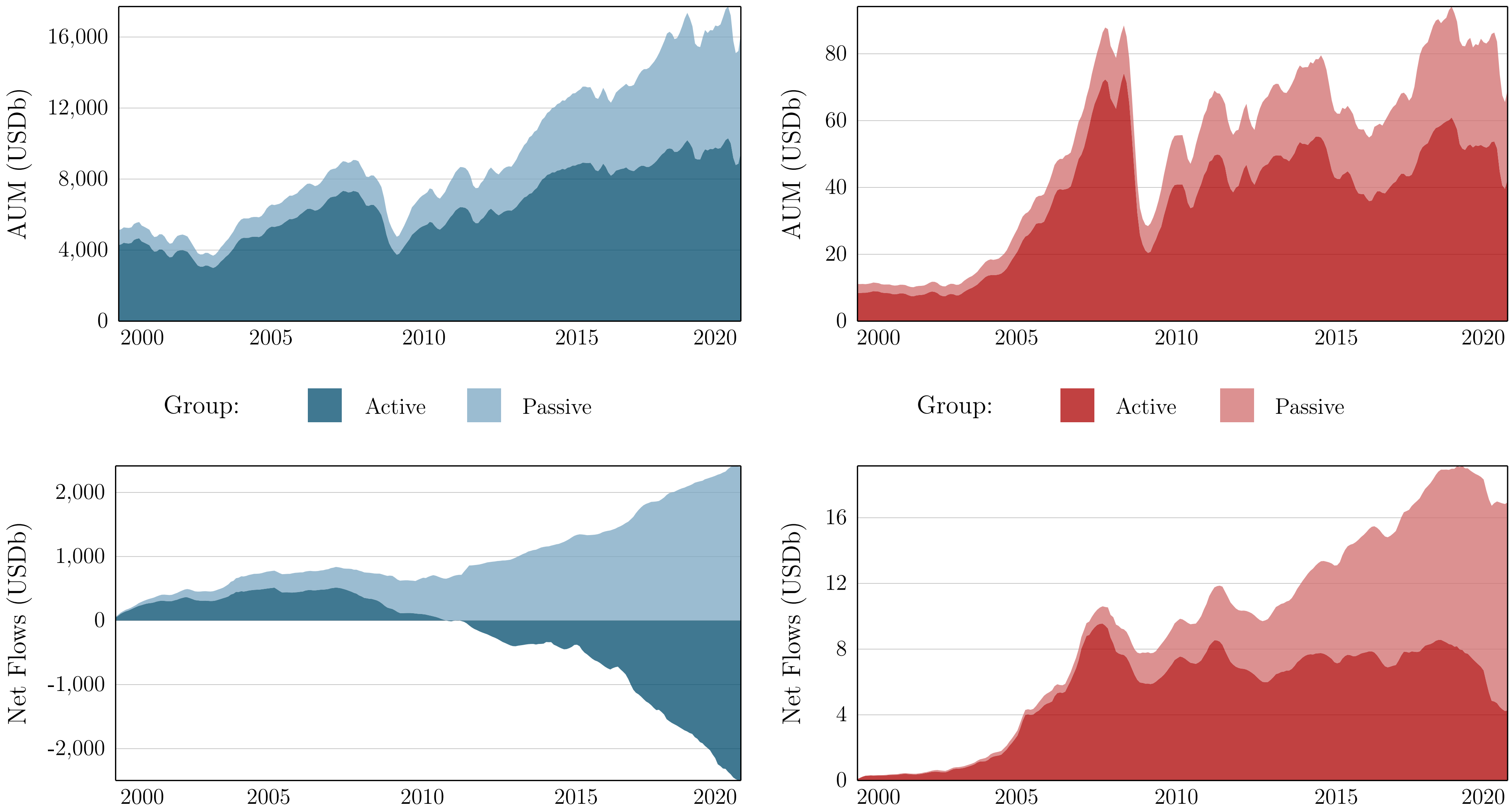
2.3 The Magnitude of Passive Investing

The total magnitude of passive strategies in equity markets is difficult to estimate. In our analyses, we include as passive all index mutual funds and index ETFs, as well as Folketrygdfondet.

Evident from Figure 2.2, the increase in popularity for passive strategies during the last 20 years has been remarkable, especially in the period after the GFC. The U.S. market is the world-leader in indexed equity products, with about half of all fund assets in passive vehicles. In Norway, the level is considerably lower, yet trending in the same direction.

Figure 2.2: AUM and Flows for Passive Funds in U.S. and Norwegian Stocks

These charts illustrate the total assets under management and the total cumulative net flow for passive and active funds in the U.S. and Norwegian markets during our sample period. The leftmost charts (blue) illustrate the U.S. market, while the charts to the right (red) illustrate the Norwegian market. The flows are cumulative aggregates of each individual fund on a monthly basis, based on data from the Morningstar Direct database. Folketrygdfondet is included in the Norwegian charts.



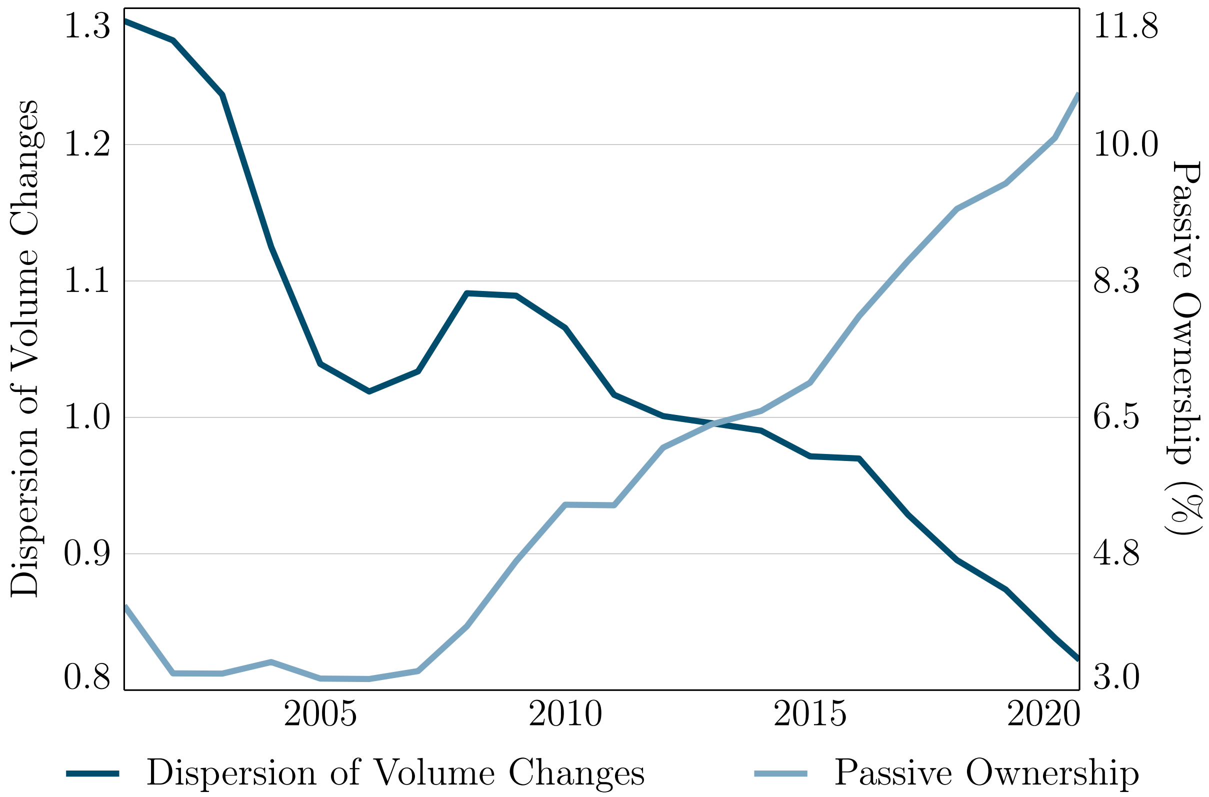
2.3.1 Index Funds, Index ETFs and Folketrygdfondet

In the period from January 2000 to June 2020, the U.S. equity holdings of index funds and index ETFs increased from 17% to 41% of total fund U.S. equity holdings, while increasing from 3% to 24% in Norwegian equities during the same period, excluding Folketrygdfondet. Relative to the total value of the stock market, index funds and ETFs now hold about 20% in the U.S. and 5% in Norway. By including Folketrygdfondet as a passive fund, they hold a total of 9% of the Norwegian market. The net flows into U.S. stocks from passive investing is equivalent to daily average inflows of about $440 million every trading day in the abovementioned period. During the same period, active managers have experienced net outflows of about the same magnitude. In Norway, while active managers have not experienced net flows in any direction during the last 10 years, passive managers have seen inflows of about $10 billion in the same period (or $4.1 million per trading day).

One key difference between passive investments in the U.S. and in Norway, is the distribution between the different types of passive vehicles. As illustrated in Figure 2.1, ETFs now constitute about 40% of passive holdings in the U.S., while less than 20% in Norway. As ETFs are used by investors for shorter horizons than index mutual funds, this means that passive vehicles in the U.S. market trade more frequently than in Norway. On the contrary, passive strategies in Norway are more tilted towards the long-term buy-and-hold strategy. This is relevant, as we believe that both the passive holdings of shares over long periods of time and the frequent trading of large baskets of stocks could influence the price efficiency of the underlying stocks. Sullivan and Xiong (2012) argue that the trading of index mutual funds and index ETFs is visible through the dispersion of volume changes between stocks. Stocks that are subject to passive funds' basket trading experience similar daily volume changes, as passive trading is uniformly a buy or sell order across stocks. In Figure 2.3, we replicate Sullivan and Xiong on our Norwegian sample and illustrate a potential link between the rise in passive investing, and a simultaneous decrease in volume change dispersion possibly caused by an increase in the trading of stocks in large baskets.

Figure 2.3: Dispersion of Volume Changes

This chart illustrates the equally-weighted dispersion of volume changes and the passive ownership share from 2000 to 2020 in the Norwegian equity market. In line with Sullivan and Xiong (2012), we calculate the dispersion of volume changes by taking the natural logarithm of the weekly trading volume divided by that in the previous week. The dispersion of volume changes in a given week is the standard deviation of these observations across all firms. Each annual observation is the average of the respective year, the prior year and the next year. Passive ownership is defined as the total market capitalisation held by passive funds, divided by the total market capitalisation of the market. Stocks with a market capitalisation below 250 NOKm are excluded.



2.3.2 Implicitly Passive

In addition to index mutual funds, ETFs and the FTF, there are other significant passive actors in the market, which we briefly account for in this section.

In 2017, index mutual funds and ETFs' AUM were estimated at about 20% and 23% of global passive AUM, respectively (BlackRock, 2017). The remaining 57% of passive capital were made up of so-called "institutional indexers" and "internal indexers", such as sovereign wealth funds, family offices and insurance providers. Both the degree of passiveness of such vehicles and the size of their assets, are ambiguous and to great extents undisclosed. Consequently, we refrain from analysing this portion of the market.

Next, the passive share of active vehicles is not accounted for, not even in Blackrock's 2017 report. This part of the market is hard to quantify, mainly because the degree of passiveness is ambiguous. On the one hand, the "active share"[6] of active Norwegian funds was stable at around 45% from 2007 to 2016, resulting in a 12-month rolling tracking error of about 5% in the period. Based on these figures, almost half of Norwegian active funds are classified as potential "closet indexers" (Thoresen and Øren, 2017). On the other hand, the passive share of an active manager is not the same as traditional index funds' holdings. While an active manager also holds a representative portion of the index — and probably for the same reasons as an index fund does — he or she is not bound by a rigid mandate and has the choice to adjust or abandon this strategy if the market, for instance, is deemed overpriced or volatile.

2.4 Passive Investing and Liquidity

To understand why an increase in a stock's passive ownership could lead to a decrease in its liquidity, we must examine how and on what basis active and passive investors trade in the market.

As predicted by Glosten (1994), electronic limit order books have neglected the role of the traditional dealer in stock markets and it is evident that the new liquidity providers are traders themselves. After the introduction of electronic markets, traders can position themselves in the order book and provide liquidity at different price levels using limit orders. In theory, active investors are therefore liquidity providers, as they have a notion of when an asset is overpriced or underpriced and are willing to provide liquidity at these levels. On the contrary, a passive strategy does not involve such notions. Passive investors only trade in relation to index rebalancing or to facilitate inflows and outflows. While they do attempt to minimise transaction costs, the most important objective when trading is to gain exposure rapidly after an inflow, and conversely after an outflow. Consequently, passive investors utilise market orders or similar types of trading algorithms, as the price at which they transact does not matter to the same extent. This way, passive investors can be viewed as liquidity demanders, as opposed to active investors (Hachmeister, 2007).

[6] Active share as calculated by Cremers and Petajisto (2009).

An additional trait of passive investors is that their holdings often have a more long-term perspective than those of active investors, meaning that they trade more seldom. While holding about 20% of the U.S. stock market, index mandates only account for about 5% of the trading (Rowley et al., 2018). This might very well be a superior strategy, but it implies that an increase in the passive ownership is analogous to a decrease in the free float of a stock, which is documented to have a negative impact on liquidity (Chan et al., 2004; Ding et al., 2016).

Another way in which funds could affect the liquidity of the underlying securities is through lending out shares, for instance to facilitate demand for short-selling or hedging. Evans et al. (2017) find that active funds, as opposed to passive funds, lose from lending out shares. Further, Massa et al. (2015) document extensive lending activity by passive funds, sometimes even outweighing their operational costs[7]. Sørmo (2016) finds that the lending out of shares is positively associated with liquidity, which means that a shift from active to passive investing could also have positive impacts on liquidity.

2.5 How and Why Prices Become Efficient

In order to establish a connection between passive investing and price efficiency, we rely on the Grossman and Stiglitz (1980) model.

The model builds on a few basic assumptions. First, all prices are made up of two components, where both are random variables, but one is observable by acquiring information at some cost. The other is unobservable for all traders, implying that no amount of information can lead to complete control. Second, traders in the model can either choose to expend resources to become informed or stay uninformed. By acquiring information at some cost, informed traders will have an edge over uninformed ones and will therefore be able to earn a profit by trading with uninformed traders. Through the trading on information, prices become more informationally efficient. The price efficiency of a stock therefore depends on the number of traders choosing to become informed.

Based on these assumptions, the authors famously claim that prices can not be fully

[7] As of 31.12.2019, KLP had NOK 4.8 billion worth of stocks lent out, equivalent to about 2.5% of their total stock holdings. At the same time, FTF had lent out about 7.8% of their total stock holdings. These figures are based on their annual reports.

efficient, because if this was the case there would be no profits from trading, and therefore no incentive for anyone to become informed nor to trade. Instead, they claim that prices are efficient to the extent that traders are motivated to gather information, and to trade on the basis of this. For traders to be motivated — that is, to choose to incur costs to become informed — the returns that they can generate in the market must at least be equal to the costs incurred to become informed.

If too many traders expend resources to become informed, then the compensation would be small relative to the costs, as there would be few uninformed traders left. This would lead to a migration back from informed to uninformed traders over time, and price efficiency would decrease back to its equilibrium level. On the contrary, if too few traders choose to become informed, prices would become so inefficient that the compensation would exceed the costs, and more traders would choose to become informed. Thus, the equilibrium is one where prices are "efficiently inefficient" to not drift too far away from efficiency.

Using this model, there are several ways in which prices could become more (less) efficient. The first and most obvious one is that more (less) informed traders will lead to more (less) efficient prices. However, other factors, like the quality of information or transaction costs in the market, would affect the choice of investors to become informed or not, and thus indirectly also affect price efficiency.

There are two ways in which passive investing could influence this equilibrium. First, uninformed investors are likely the ones switching from trading the underlying securities to an index strategy, as their losses to informed investors are limited when investing in such vehicles (Gorton and Pennacchi, 1993). This leads to a decrease in the returns and an increase in the costs of trading on information, which both lead to a decrease in price informativeness. Second, and more indirectly, the possible increase in trading costs outlined in the previous section is equivalent to an increase in the cost of information, and is also expected to lead to a decrease in price efficiency.

2.6 Literature Review

The prevalence of passive strategies has increased substantially since the global financial crisis, both in the U.S. and Norwegian equity markets. In line with these developments, academics and professionals ask themselves what implications this might have for the microstructure of markets. In the following, we present a few of the ways in which passive strategies are believed to affect the market, and the research that has been published touching on this topic. We emphasise that a few of the aspects are intertwined, but we will nevertheless review the contributions individually.

2.6.1 Index Inclusion Effects

In academic research, the most popular subject in terms of index strategies' effects on markets has been index inclusion effects. Consider for instance the inclusion of the Tesla stock in the S&P 500, which will be effective at the 21^{st} of December 2020. Roughly 4.6 trillion passive dollars track the index (Dans, 2020), which Tesla will constitute around 1.7%[8] of. Evidently, the inclusion has initiated a $77 billion buying pressure in the stock, as passive funds are obliged to hold 16% of the stock's free float immediately after inclusion. This will inevitably affect both the liquidity and the price of the Tesla stock.

An index inclusion effect in the S&P 500 was first discovered by Andrei Shleifer (1986). On average, stocks included in the S&P have risen about 8.8% around the inclusion date. The effect on excluded stocks is stronger, averaging at -15.1% (Petajisto, 2011). Naturally, as passive managers have become aware of this, they have attempted to anticipate possible index inclusions ahead of the announcement, and trade on this information if their mandate allows for it. For instance, KLP — Norway's largest index manager — utilises such a strategy and claims for it to be a source of outperformance over time (Embu, 2020). Despite this, using fund performance data from 2010-2014, Nesse and Aasen (2015) found no consistent outperformance by Norwegian nor U.S. index funds. Such strategies are referred to as enhanced indexing strategies, and could be the reason why Scari (2016) found no link between the increase in the passive share of the market and an increase in index inclusion effects. Instead, Scari found that the index inclusion effect for the S&P 500 index peaked in the late 1990s. An interesting hypothesis is that the reason for this peak was all the new technology firms with a high founder ownership, i.e., low float-to-market capitalisation ratio. In the late 1990s, indices were not float-adjusted, meaning that the amount of passive investments were solely based on market capitalisation. Considering two otherwise identical stocks, one would expect prices of the low free float stock to rise

[8]Calculated based on TSLA market capitalisation and free float as of close 15.12.2020, retrieved from Yahoo Finance.

more from index inclusion than that of the high free float stock. Today, most indices, including the OSEBX and S&P 500, are market capitalisation weighted with adjustments through a float factor, ultimately reducing the cost of replication for passive vehicles.

We emphasise that while our thesis is not a study of index inclusion effects, our analyses examine a few of the same mechanisms. However, we measure the impact of passive ownership, and index membership is only relevant in deciding which stocks that receive the inflows of passive investors. We control for index inclusion effects in our causality analyses in Chapter 4.

2.6.2 The Impact of Passive Investing

The idea of an impact of passive strategies on stock market dynamics apart from the effects at index inclusions, is a relatively recent one. The pioneers of the index fund largely relied on the notion that inefficiencies would be countered by active investors and arbitrageurs, ultimately limiting the impact of passive vehicles on the market. Nevertheless, as the prevalence of passive investing has increased, some researchers argue that this notion is no longer valid.

Hamm (2014) finds that ETF ownership is positively associated with the adverse selection costs of trading. This is due to the migration of uninformed traders from the stock market to the ETF market, which results in a decrease in liquidity for individual stocks. Israeli et al. (2017) posit that through this decrease in liquidity, ETF ownership could also negatively affect the price efficiency of individual stocks. They find significant relationships supporting this proposition using data on U.S. stocks. Sammon (2020) investigates the same effect using novel measures of price informativeness based on the market's reactions to announcements of firm-specific information. His findings largely correspond to those of Israeli et al. (2017). Glosten et al. (2016) also investigate the effect of ETFs ownership on the price informativeness of underlying securities but find that ETF ownership is positively related to price informativeness. The reason for the opposing views lays in the research design. While Israeli et al. (2017) use lagged changes in ETF ownership, Glosten et al. (2016) consider contemporaneous changes. The interpretation of these conclusions could be that while an ETF trade is initiated on the basis of information, the long-term effect of increased passive holdings is a decrease in informational efficiency.

Zou (2019) considers the directionality of the pricing effects of ETF ownership. Instead of merely concluding that prices could become inefficient, she finds a positive link between ETF flows and valuations in the long-term. This view is supported by Wurgler (2010). The hedge fund Logica Capital Partners has expressed similar concerns regarding valuations. While critics of this view claim that mispricing would be countered by active managers, Logica points to the fact that the average cash holdings of active mutual funds is about 5% (Yan, 2006), while passive funds' level is closer to 0.5% (Embu, 2020). Considering a universe of one active and one passive manager, and cash and equity as the only option for the two funds, it becomes evident that a shift from active to passive must drive equity prices upwards (Logica Capital Partners, 2020). The specific effect of this becomes a question of the elasticities and price impact of inflows.

Another line of research conducted on the effect of passive ownership focuses on the increasing comovement of index constituents. Not to be mistaken for research related to index inclusion effects, Sullivan and Xiong (2012) examine the comovement of S&P stocks and link its increase to the increase in passive strategies. Specifically, they find a strong increase in the pairwise correlation between index constituents, and also seemingly clear signs of indexed trading by considering the dispersion of volume changes among stocks. This measure is almost a perfect inverse of the trend in passive ownership, similar to Figure 2.3. Da and Shive (2018) use stock-level data and arrive at the same conclusion that index investing is associated with increased comovement. They claim that ETFs is the most important reason for such observations, as it has become easier for investors to trade on macroeconomic events using ETFs. This trading must propagate to individual stocks. To the extent that the ETFs attract non-fundamental flows, this comovement is a sign of less informative prices. On the contrary, Madhavan and Morillo (2018) claim that the rise in cross-stock correlations is not connected to ETF flows, but rather an increased investor emphasis on the macroeconomic environment.

The last effect of passive investing we consider in this review is the one on volatility. As opposed to those on liquidity, price informativeness and volume, this effect is almost solely attributed to ETFs, and not so much regular mutual funds. Ben-David et al. (2018) argue that demand shocks in the ETF market propagate to the underlying securities, due to arbitrageurs trading whenever the ETF price deviates from the NAV. The idea is that noise traders choose ETFs over trading the underlying securities, due to reduced costs and the fact that there are less informed traders present. Given that the trading in ETFs does not reflect firm-specific information, the arbitrage channel leads to non-fundamental volatility in the underlying securities. This is similar to the arguments of Da and Shive (2018) in relation to comovement. Ben-David et al. (2018) find a significant and causal positive relationship between ETF ownership and volatility in their 2000-2012 sample of U.S. stocks.

2.7 Motivation

We add to the literature by examining the effects of passive investing on both liquidity and price efficiency in the peculiar Norwegian market. Our analyses on liquidity are similar to those of Israeli et al. (2017) and Hamm (2014), however, we expand the existing literature through our focus on all observable forms of passive investing, and not just ETFs. Further, as opposed to Israeli et al. (2017), we use contemporaneous changes and apply a quasi-experiment setting when analysing the possible link between liquidity and price efficiency.

The U.S. market differs from the Norwegian in three important ways. First, passive investing in total is larger in the U.S., with Norway lagging about ten years behind. Second, a significantly larger share of passive holdings is invested with ETFs in the U.S. Lastly, Norwegian stocks have among the lowest free float ratios in the western world.

The difference in the prevalence of passive investing does not necessarily produce different results, however, American researchers have a larger sample of data to draw their conclusions from. As for the differences in the types of passive vehicles utilised and analysed, we expect the passive ownership in Norwegian stocks to be more inelastic and long-term than in the U.S., which could imply a stronger negative impact on liquidity and a weaker effect on price efficiency. Lastly, low floats in Norwegian stocks are expected to further amplify the impact on liquidity, as we define passive ownership relative to non-adjusted market capitalisation.

3 Data and Methodology

In this section, we present both the data and methodology used in our analyses throughout the paper. First, we elaborate the sources of our different types of data and the structure and size of the samples derived from these sources. Second, we present the variables utilised in order to carry out the analyses and provide descriptive statistics for these. Finally, we elaborate on the specific methodology that we apply in our analyses.

3.1 Data Sources and Sample Selection

In total, we utilise data from five different data sets and four different providers. First, we gather financial data on all publicly traded Norwegian stocks from NHH Børsprosjektet (Norges Handelshøyskole, 2020) for the 21.5-year period from 01.01.2000 to 30.06.2020. This includes daily prices, returns, bids and asks, turnovers, shares outstanding and sector classifications (GICS) for all stocks throughout the period. We also collect the daily closing value of the OSEBX. In total, we retrieve this data for 591 stocks over an average of 7.5 years, resulting in a total of 1,118,128 stock-days. We exclude observations of stocks with a market capitalisation below $27.8 million[9] due to low trading activity, reducing the size of the sample to 540 stocks and 853,779 stock-days. 214 of the stocks are still listed with an average life of 9.7 years. This means that we have a total of 521,452 daily observations of stocks that are still listed. This becomes relevant as the fund ownership data only includes stocks that are currently listed, which limits our analyses to these stocks. Figure 3.1 illustrates the relationship between stocks that are currently listed and those that were delisted from the Oslo Stock Exchange at some point in the period.

Second, we collect the stock holdings of all mutual funds and ETFs from the Morningstar Fund Ownership database through the Morningstar Direct suite (Morningstar, 2020). The ownership data is provided on a monthly basis for every individual fund, and includes all individual stock positions on a number of shares basis. The Morningstar ownership data only includes stocks that are currently listed. This could introduce a bias, as stocks from Børsprosjektet that are not currently listed are excluded. However, we do not believe that this sort of "survivorship bias" will affect the analyses conducted, as passive ownership is

[9] $27.8 million is equal to NOK 250 million at a USDNOK of 9.00.

a fairly recent phenomenon, and there is no indication whether delisted (mainly due to bankruptcy or M&A activities) stocks would have had a high or low passive ownership relative to others. Out of more than 500,000 funds, 6,814 have at some point held shares in a Norwegian firm that is currently listed. We emphasise that we are not considering merely the holdings of Norwegian funds, but rather the Norwegian equity holdings of all funds. As passive, we classify all funds which name includes certain strings, such as “INDEX”, “IDX”, “OSEBX” and “MSCI” in both uppercase, lowercase and various combinations of the two[10]. Using our classification rules, 871 out of the 6,814 funds are classified as passive. We merge the ownership data with the financial data by matching the tickers manually. This means that for every stock in every month, we have the exact number of shares outstanding and the number of shares held by passive funds. Currently listed stocks that are not present in the Morningstar database are assigned a passive ownership share of zero. The 6,814 funds have over time held shares in 185 out of the 214 stocks, whereas the passive ones have held 123 stocks.

Third, we collect the annual end-of-year holdings of the Norwegian equity portfolio of Folketrygdfondet. From the fund’s annual reports, we manually extract all individual stock holdings on a number of share basis and merge these with the ticker symbols from the other data sets. The Morningstar data is on a monthly basis, while the data from Folketrygdfondet is on an annual basis. Some of our analyses rely on year-over-year changes between dates that are not end-of-year. Consequently, we assume that Folketrygdfondet invests gradually over the full course of the year when converting the data to a monthly basis. This conversion is in no way optimal but we deem it better than the alternative, which would be to assume that all end-of-year values are representative for the entire year, implicitly assuming that the fund only invests on the 1^{st} of January each year. In our sample of stocks that are currently listed, Folketrygdfondet’s AUM has increased from NOK 10.6 billion in December 1999 to NOK 121.8 billion in June 2020, with 39 portfolio stocks on average[11]. The average stock-level fund ownership for our sample, segmented in passive funds, active funds and the FTF, is provided in Figure 3.2.

[10] Complete list: "INDEX", "IDX", "INDEKS", "PASSIVE", "OSEBX", "OBX", "S&P", "SANDP", "BLOOMBERG", "RUSSELL", "100", "500", "1000", "2000", "3000", "MORNINGSTAR", "FTSE", "MSCI", "STOXX", "BLACKROCK", "BLKROCK", "STATESTREET", "TARGET".

[11] As illustrated in Figure 3.1, our sample size increases with time. In 2000, when the deviations are the largest, FTF held 59 stocks, while only 25 are included in our sample. Similarly, the actual AUM was NOK 19.3 billion in December 1999.

Figure 3.1: Sample Size

This chart illustrates the number of individual stocks in our sample and their status for each month. As only currently listed stocks can be linked to fund ownership data, our sample size is represented by the dark blue area.

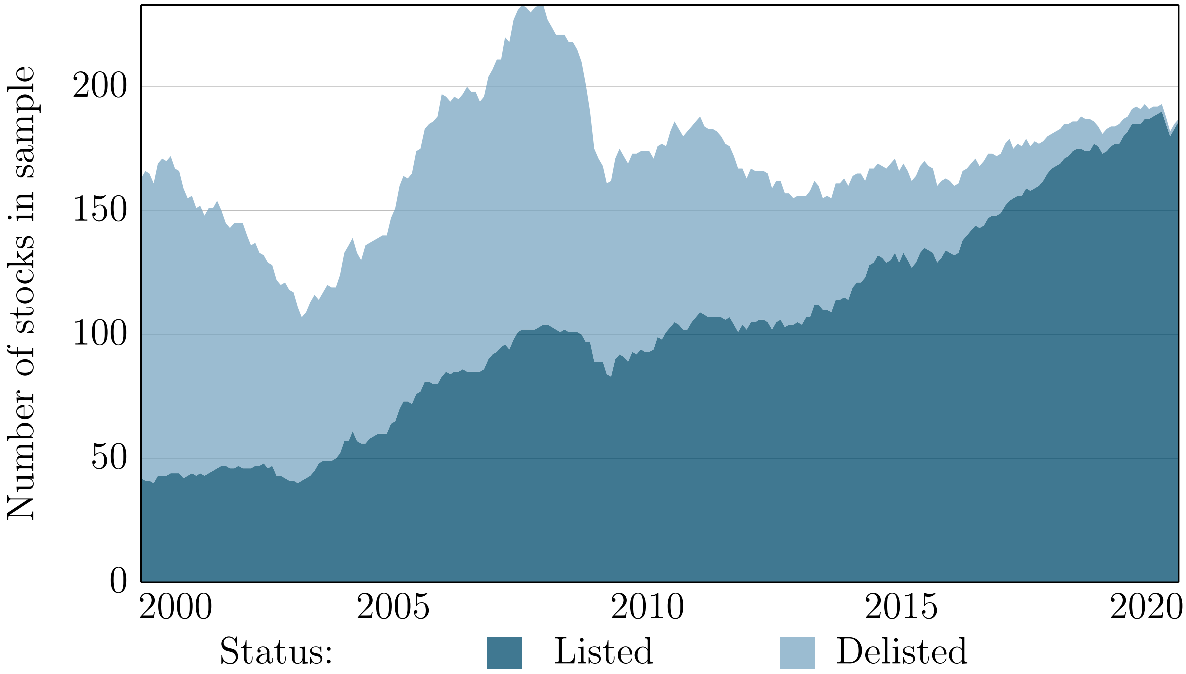
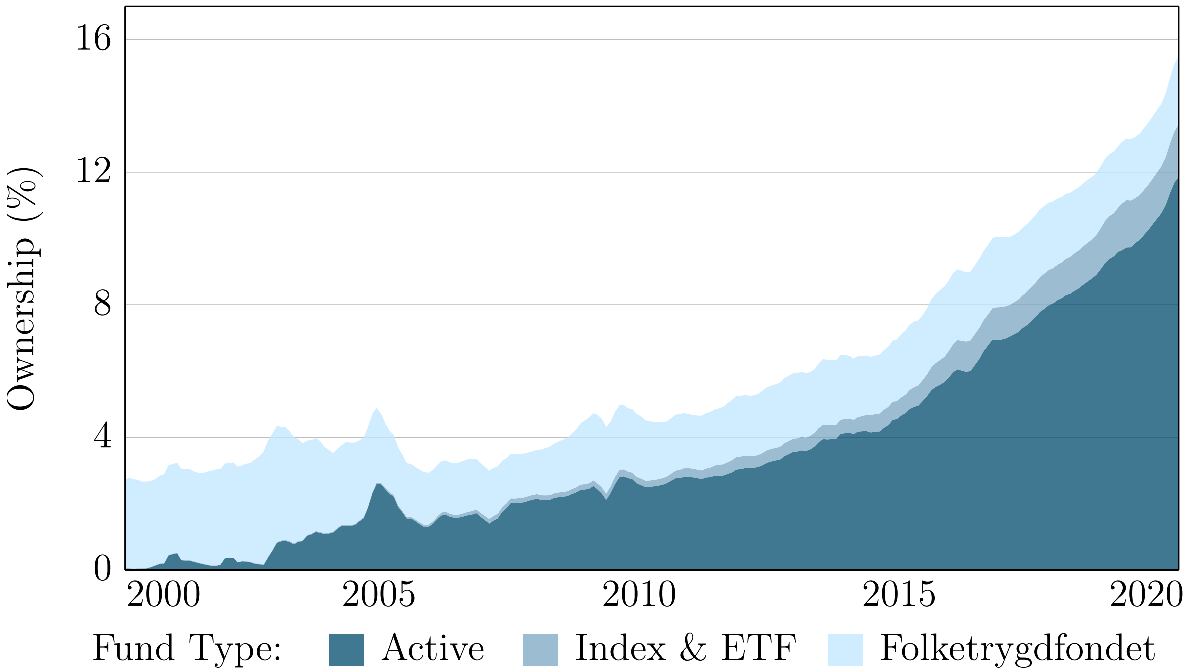


Figure 3.2: Fund Ownership in Sample

This chart illustrates the average fund ownership for each stock in our sample, segmented by fund type. The averages are based on equal weights for all stocks.



Fourth, we gather information on the date and the time of earnings announcements from Thomson Reuters Eikon. If the announcement time is before or during the date's trading hours, the effective date is the same as the announcement date. If the announcement time is later than 16:20[12], the next trading day is classified as the effective date. For announcements where only the date but not the time is specified, we classify as the effective date the current or the following date, based on which of them has the higher trading volume. The data on earnings announcements is matched with the financial data and the ownership data manually by ticker. On average, each stock is linked with 20 announcements.

Lastly, also from Thomson Reuters Eikon, we collect a list of all inclusions and exclusions from the OSEBX. By combining this data with the index constituents in 2001[13] and today, we derive a complete list of all the index constituents at any given time. We link this data with the financial data by ticker symbol and construct dummies indicating whether a stock at any point in time is included in the OSEBX.

3.2 Variables Selection

We aim to examine how increases in passive ownership have affected the liquidity and price efficiency of Norwegian stocks. Consequently, we use passive ownership as an independent variable, and different measures of liquidity and price informativeness as the dependent variable in our regression models. As stock markets are complex systems, we also include carefully selected control variables to account for simultaneous effects. All variables will be presented in the following sections.

3.2.1 Independent Variable: Passive Ownership

First, we need a consistent measure of passive ownership as our independent variable of interest. As discussed in Section 3.1, our data set contains the number of shares held by passive funds and Folketrygdfondet on a monthly basis. We calculate the passive ownership share as the number of shares held by passive vehicles divided by the total number of shares outstanding. We multiply this share by 100 in order to express passive

[12]The OSE trading session ends at 16:20 each day.
[13]OSEBX has served as the benchmark index, replacing the TOTX, since 23.05.2001.

ownership in percentage points. As some of our analyses are based on specific earnings announcement dates within a certain month, we assume the ownership to be constant throughout each month[14].

3.2.2 Dependent Variables: Liquidity

Active traders can be considered liquidity providers as they are always, at least implicitly, present in the order book of a stock. At some high (low) price, an active owner would consider selling (buying) shares, thus providing liquidity in that stock. This is not the case for passive managers, who do not conduct fundamental analyses and use market orders or similar types of algorithms when trading. From February 19^{th} to May 31^{st} this year — during the sharpest economic downfall in modern history — only 5% of Vanguard DC[15] investors and 17% of Vanguard retail investors traded in any meaningful way (Vanguard, 2020). Less than 0.5% liquidated their assets and moved to cash. As such price-inelastic investors hold more of the outstanding shares of a stock, we expect the liquidity of that stock to decrease. In our analyses, we use two proxies of liquidity.

3.2.2.1 Amihud's Illiquidity (*Illiq*)

As our first measure for liquidity, we use the famous illiquidity proxy of Amihud (2002). This is the most renowned liquidity measure in the financial industry and measures the price impact of trading in a low-frequency manner (Goyenko et al., 2009). It is derived using the ratio of the absolute daily stock return to the daily NOK trading volume. Specifically, *Illiq* is defined as:

$$Illiq_{i,t} = \frac{\mid return_{i,t} \mid}{turnover_{i,t}} \tag{3.1}$$

where $\mid return_{i,t} \mid$ is the return of stock i on day t expressed in absolute terms, while $turnover_{i,t}$ is the trading volume of stock i on day t expressed in NOK millions. We use

[14] This assumption does not apply to months where corporate actions (SEOs, stock splits, etc.) have caused changes in the number of shares outstanding within the month. To avoid spikes in *Passive Ownership* in such months, we have manually adjusted the passive ownership share within the month in question. We believe such adjustments are reasonable, as index funds and ETFs minimise tracking error, and would therefore need to participate in SEOs. Such adjustments are not performed on the FTF holdings, as this fund does not necessarily need to adjust their holdings accordingly in such events.

[15] DC = Defined Contribution. A pension plan with defined contributions by the employer and the employee, as opposed to a defined benefit plan.

annual averages in our analyses and expect changes in *Passive Ownership* to be positively associated with changes in *Illiq*.

3.2.2.2 Abdi and Ranaldo's Spread (*Spread*)

Our second proxy for liquidity is the bid-ask spread. More specifically, we use an estimate of the effective spread, introduced by Abdi and Ranaldo (2017). This estimate is similar to, and highly correlated with, the famous Roll (1984) spread measure but based on more easily available information. Further, the Abdi and Ranaldo measure delivers the most accurate estimates for less liquid stocks (Abdi and Ranaldo, 2017)[16]. The effective spread, *Spread*, is defined as:

$$Spread_{i,m} = \frac{1}{N_i} \sum_{t=1}^{N_i} 2\sqrt{max\left(E[(c_{i,t} - \eta_{i,t})(c_{i,t} - \eta_{i,t+1})], 0\right)} \tag{3.2}$$

where $c_{i,t}$ is the daily closing log-price, $\eta_{i,t}$ is the mean of the daily high and low log-prices and N_i is the number of daily observations for stock i in month m. To mitigate the effect of negative roots and to achieve the most accurate spread, we use annual averages of monthly estimates (m). We expect for changes in *Passive Ownership* to be positively associated with changes in *Spread*.

3.2.3 Dependent Variables: Price Informativeness

In the second part of our analysis, we examine the link between passive ownership and price informativeness. As trading costs increase, active investors should, ceteris paribus, choose to trade less on the basis of idiosyncratic information (Grossman and Stiglitz, 1980). This is because an increase in trading costs is equivalent to an increase in the cost of information. An additional effect, proposed by Israeli et al. (2017), is that there will be less uninformed investors left for informed investors to profit from trading with, as these migrate to index investing. Both effects lead to less information being embedded in prices through trading. Further, as more people invest in passive than in active vehicles, we would expect prices to reflect less stock-specific information. In order to analyse these effects, we need to establish proxies of price informativeness. As neither the true value nor all the relevant information of a stock can be quantified, measuring the true price

[16]Relative to the TAQ (NYSE Trade and Quote Data) reported spread.

informativeness is not possible. However, the literature suggests a range of proxies. We focus on two types of proxies. First, we use three proxies attainable through event studies. Specifically, we can focus on events where we are certain that new firm-specific information has become available. That is, earnings announcements. Information on trading and pricing prior to earnings announcements has been used in multiple academic papers as a measure of price informativeness (Foster et al., 1984; Pettit, 1976; Sammon, 2020). The idea behind this is that trading prior to earnings announcements reflects investors' beliefs regarding the information which will be presented by the firm. We use three such event-based proxies introduced by Sammon (2020). The fourth proxy measures the extent to which the price of a stock moves in tandem with the market as a whole. Roll (1988) suggests that return comovement is an inverse function of the idiosyncratic information — relative to more market-wide information — embedded in stock prices. All four proxies will be presented in the following sections.

3.2.3.1 Sammon's Pre-Earnings Abnormal Volume (CAV)

The first measure of price informativeness is the abnormal trading volume in the days leading up to the earnings announcement. When the passive ownership in a stock increases, one would expect the price of the stock to contain less information, as a result of the aggregate market conducting less firm-specific research, and the incentive mechanisms of Grossman and Stiglitz (1980). Sammon compares the trading volume of each day in the month (22 days) leading up to the announcement with the average volume for the previous three months (63 days). More specifically, the first measure, cumulative abnormal volume (CAV), is defined as:

$$CAV_{i,T} = \sum_{t=-22}^{21} AV_{i,t} \tag{3.3}$$

where:

$$AV_{i,t} = \frac{\sum_{t=-22}^{21} V_{i,t}}{\frac{1}{63}\sum_{k=1}^{63} V_{i,t-22-k}} \tag{3.4}$$

where $V_{i,t}$ is the NOK trading volume for stock i at day t. The abnormal volume (AV) for each day in a 22-day period leading up to the announcement is calculated by dividing the respective day's volume by the average daily volume in a fixed window spanning from four to one month prior to the announcement. The cumulative abnormal return in Equation 3.3, which will serve as the dependent variable in our regression, is simply the sum of the abnormal volume (Equation 3.4) in the 22-day period leading up to the announcement date. This means that if there had not been any excess volume, CAV would be equal to 22, indicating that the trading volume leading up to an earnings announcement is equal to the trading volume in the absence of an earnings announcement. We expect changes in CAV to be negatively correlated with changes in *Passive Ownership*.

3.2.3.2 Sammon's Pre-Earnings Drift (*DM*)

A decline in trading volume is not enough to deem pre-earnings prices less informative. Prices could reflect all relevant information, even though the trading volume has decreased. Therefore, we also investigate the relationship between passive investing and pre-earnings announcement drift. The tendency for stock prices to increase (decline) prior to a positive (negative) earnings announcement is documented in academia (Easton et al., 1984). The reason that we experience such patterns is the trading of informed investors on signals of performance or explicit guidance from the firm. Consequently, we would assume that an increase in passive ownership would reduce the drift prior to a stock's earnings announcement, as a smaller share of the market is conducting firm-specific analyses. Sammon's measure of drift captures the extent to which pre-earnings returns reflect the return on the announcement day. It does so by comparing the two returns relative to one another, so that an overall change in volatility would not lead to biased results. Specifically, the drift measure (DM) is defined as:

$$DM_{i,T} = \begin{cases} \frac{1+\sum_{t=-22}^{21} r_{i,t}}{1+\sum_{t=-22}^{22} r_{i,t}}, & \text{if } r_{i,T} > 0 \\ \frac{1+\sum_{t=-22}^{22} r_{i,t}}{1+\sum_{t=-22}^{21} r_{i,t}}, & \text{if } r_{i,T} < 0 \end{cases} \tag{3.5}$$

where $r_{i,t}$ is the return of stock i on day t in the period leading up the announcement date, while $r_{i,T}$ is the return on the announcement date. DM is close to 1 if the drift is strong, i.e., if the return on the announcement day is small but in the same direction as the cumulative return of the pre-announcement period. On the other hand, DM is low if the return on the announcement day is large relative to that of the pre-announcement period, or if the price moves substantially in the opposite direction. Note that if the return on the announcement day is negative, the measure in Equation 3.5 is inverted. This is done so that the interpretation remains the same for negative and positive earnings announcements. We expect changes in *Passive Ownership* to be negatively associated with changes in *DM*.

3.2.3.3 Sammon's Earnings-Day Volatility (*QVS*)

If we believe that prices contain less of the information about earnings in the days prior to the announcement, we would assume that this would be made up for on the announcement day. At the announcement day, much of the uncertainty is removed, and informed traders are now willing to trade on the stock-specific information presented. Therefore, if less information about the announcement is acquired by the market in advance, the announcement itself would have to contain more new information, hence provoke greater volatility. The quadratic variation share (QVS) is a measure of the quarterly volatility occurring on earnings announcement days and is defined as:

$$QVS_{i,T} = \frac{r_{i,T}^2}{\sum_{t=1}^{63} r_{i,t}^2} \tag{3.6}$$

where $r_{i,T}^2$ is the squared return of stock i at earnings announcement date T and $r_{i,t}^2$ is the returns of stock i at day t in quarter T. The ratio is defined as the daily squared return for the earnings announcement day divided by the sum of daily squared returns for the announcement quarter. We expect changes in *Passive Ownership* to be positively associated with changes in (QVS).

3.2.3.4 Stock Return Synchronicity (*Synch*)

With the fourth proxy, we want to examine the relationship between a stock's passive ownership and its return comovement with the market. To do this, we use the adjusted R^2 from a market model regression. Roll famously suggests that low R^2 statistics in his asset pricing regressions could be a sign of high price informativeness. In his view, stock prices can reflect three types of information: market-wide information, sector-wide information and stock-specific information. This means that the residual in a market model with stock-specific returns explained by sector and market returns will describe the amount of firm-specific information embedded in the price. The relationship was formalised by Durnev et al. (2003) and has been recognised as a proxy for price efficiency in several articles (Bramante et al., 2015; Piotroski and Roulstone, 2004). Specifically, we regress the daily returns of each stock against the returns of the benchmark index and the respective stock's sector over the course of a year[17]. The adjusted R^2 of this regression is our proxy for price informativeness. In line with Israeli et al. (2017), we truncate negative values of R^2. High values of R^2 implies that prices move more in tandem with that of the market, and thus contain less firm-specific information (Roll, 1988). The regression from which the adjusted R^2 is derived is:

$$ret_{i,t} = \beta_1 * ret_{OSEBX,t} + \beta_2 * ret_{OSEBX,t-1} + \beta_3 * ret_{SEC,t} + \beta_4 * ret_{SEC,t-1} + \epsilon_{i,t} \quad (3.7)$$

where $ret_{i,t}$, $ret_{OSEBX,t}$ and $ret_{SEC,t}$ are the daily returns at time t of stock i, the OSEBX and the stock's sector, respectively. The model in Equation 3.7 is run once for each stock-year. Stocks with less than 75 daily observations or less than three other stocks in their sector are excluded to avoid a biased measure. In line with Israeli et al. (2017) and Piotroski and Roulstone (2004), we transform the variable to a continuous one using the formula below:

$$Synch_{i,y} = \ln\left(\frac{R^2_{i,y}}{1 - R^2_{i,y}}\right) \quad (3.8)$$

where $R^2_{i,y}$ is the adjusted R^2 from Equation 3.7 for stock i in year y. We expect changes in *Passive Ownership* to be positively associated with changes in *Synch*.

3.2.4 Control Variables

It is natural to assume that there are other factors than passive ownership influencing our dependent variables. In order to deal with issues of endogeneity, we have included various control variables which we will present in the following section.

In all our models, we include market capitalisation as a control variable. We find it intuitive for market capitalisation to be the most important single factor in explaining both liquidity and price informativeness. In relation to liquidity, we expect large stocks to have tighter spreads and a lower price impact per NOK traded. Moreover, in relation to price informativeness, we believe that larger stocks in general are being followed by a larger number of analysts and professional traders. This would imply higher price informativeness. However, larger stocks are also a larger part of indices and sectors, which

[17] We use OSEBX as the benchmark index and the two first digits of the GICS as the sector.

could imply the contrary, especially when using the synchronicity measure. Nevertheless, market capitalisation seems like an integral control variable in this analysis and we express this variable by the natural logarithm.

Across all models, we include the active ownership of a stock. Glosten and Harris (1988) document a significant relationship between fund ownership and liquidity. As we already account for passive ownership, we should also include active, which is defined as *Fund Ownership - Passive Ownership*. We include *Active Ownership* as a control variable in price efficiency analyses as well, due to our reliance on the link between liquidity and price informativeness (Grossman and Stiglitz, 1980). This control variable also addresses the potential role of active investors as liquidity providers. We calculate *Active Ownership* as the total number of shares of a stock held by active funds divided by the stock's total number of shares outstanding.

For the liquidity analyses, we include both volatility and market-adjusted volatility as control variables. In general, we believe that both spreads and price impact might increase in periods of high volatility. However, there might be differences in how liquidity is affected when the overall market volatility changes, and when firm-specific information leads to changes in volatility. For instance, firm-specific volatility could lead to increases in the stock's trading volume to a greater extent than a market-wide rise in volatility could. *Volatility* is defined as the annualised standard deviation of returns, calculated as the standard deviation of a stock's returns multiplied by the square root of the number of trading days. *Idiosyncratic Volatility*, on the other hand, is the return of a stock, less the return of OSEBX, squared. We also include trading volume, *Volume*, as a control variable in the liquidity analyses. This variable is calculated as the total shares traded divided by the total shares outstanding per day. Especially the price impact measure is sensitive to changes in trading volume in small-cap and mid-cap stocks. As illustrated in the correlation matrix, Table 3.1, we emphasise that the link between volume, volatility and liquidity could be problematic when interpreting the coefficients.

Specific to the analyses on price informativeness, we include *Beta* as a control variable. We would not want to interpret high-beta stocks as having less informative prices (Li et al., 2014). That is, for a stock with a relatively high beta, we allow for idiosyncratic information to be deemed less important by the market, relative to market-wide information. We have calculated *Beta* as the covariance between the returns of a stock and OSEBX, divided by the variance of the returns of OSEBX. Lastly, we also include *Idiosyncratic Volatility* in the analyses on the event-based measures, as we believe that high idiosyncratic volatility throughout the year could affect the perceived significance of earnings announcements. We do not include this variable in the fourth price informativeness measure, *Synch*, as the absence of idiosyncratic volatility is equivalent to a high synchronicity value.

3.3 Methodology

In this section, we present the econometric approaches utilised in order to address our research questions. The analysis of each hypothesis is split into two parts. First, we investigate potential links using a panel data regression model both including and excluding control variables and fixed effects. Second, we analyse the same relationships using a difference-in-differences (DiD) approach, in order to potentially establish causal links. In the following, both approaches will be presented.

In line with Hamm (2014), Israeli et al. (2017) and Sammon (2020), we use a first differences approach, considering the changes in the variables instead of their levels. While the common reason for using a first differences approach is limiting issues with endogeneity, we already apply an equivalent method with the inclusion of fixed effects. The reason for our choice is rather that such a design limits the impact of the levels of *Passive Ownership* being non-stationary over time, as well as autocorrelation affecting our coefficients. The use of differences variables throughout our analyses leads to low R^2 in our models. In financial regressions, levels variables can often inflate the R^2, as most variables are sticky and move in one direction over time. Additionally, the adjusted R^2 penalises the use of many independent variables[18]. As we utilise first differences models with both time and stock fixed effects, we therefore expect low R^2s across models. However, we do not deem the adjusted R^2 to be crucial for our analyses, as we measure relatively small effects in very complex structures. Further, we include neither levels nor interaction variables, in order to facilitate intuitive interpretations of our regression models. We do, however, wish to address the large dispersion in market capitalisation in our sample. The inclusion of the level of *Log MCAP* could introduce multicollinearity in our regressions, as large firms

[18] 21 years and 214 unique stocks result in 235 additional independent variables when we include time and stock fixed effects in regressions on the annual sample.

are the ones experiencing the largest inflows of passive capital. Instead, we include for all regression models an additional table where the same models are run on different terciles of the sample, sorted by market capitalisation. These tables are provided in Appendix A2.

We derive the change variables by taking the one-period difference in all variables defined in this chapter and use these new variables in our regressions. In the analyses on liquidity and synchronicity, the Δ denotes the differences between $year_t$ and $year_{t-1}$. In the analyses on earnings announcements, the Δ denotes the differences between $quarter_t$ and $quarter_{t-4}$[19]. We consider the year-over-year change in order to address issues regarding seasonality, as it could for instance be argued that market participants consider 1^{st} quarter announcements more important than 2^{nd} quarter ones.

3.3.1 Correlational Study

The first part of the analysis of both hypotheses consists of panel data regressions, in which we study the correlation between changes in *Passive Ownership* and changes in the abovementioned dependent variables. We use annual or quarterly observations, and include time fixed effects, allowing us to control for structural changes in the information environment on the Oslo Stock Exchange, such as the prevalence of stock trading in general or the prevalence of performance guiding by firms. For instance, imagine if firms in 2001 in aggregate were more inclined than firms in 2019 to reveal information to the market ahead of earnings announcements. This could lead to biased estimates on the relationship between *Passive Ownership* and price informativeness, as *Passive Ownership* has increased consistently in the same period. Using year fixed effects, each year is assigned a unique intercept, which can absorb such differences or control for unique market environments impacting our measures, such as the GFC and the COVID-19 crisis. We also control for differences between stocks, by applying stock fixed effects. The earnings of some firms might for instance be more predictable, and the stock could therefore have high pre-earnings volumes and drift. Some stocks might even have different response coefficients on information. Such time-invariant differences are absorbed by each stock's intercept. In order to control for heteroskedasticity[20], we apply robust standard errors

[19]For observations of $quarter_t$ where the $quarter_{t-4}$ is missing, we take the difference between $quarter_t$ and $quarter_{t-1}$.

[20]The Durbin-Watson test is not significant for any of our models, indicating no autocorrelation. This is primarily through the use of the first differences design.

using the White method. Following the recommendations of Petersen (2009), we cluster standard errors by stock, as our data involves a larger number of stocks than years or quarters.

For all of the dependent variables presented in this chapter, we apply variations of the following model:

$$\Delta Y_{i,t,d} = \alpha + \beta * \Delta PO_{i,t} + \sum_{j=1}^{J} \beta_j * \Delta C_{i,t,j} + \sum_{i=1}^{I} \beta_i * FE_i + \sum_{t=1}^{T} \beta_t * TE_t + \epsilon_{i,t,d} \quad (3.9)$$

where $\Delta Y_{i,t,d}$ is the change in the dependent variable d of stock i from time $t-1$ to t. Similarly, $\Delta PO_{i,t}$ is the change in *Passive Ownership* of stock i from time $t-1$ to time t. $\Delta C_{i,t,j}$ is the change in control variable j for stock i from time $t-1$ to t, while FE_i and TE_t is the stock and year fixed effects, respectively.

3.3.2 Quasi-Experiment

In the second part of the analyses of each hypothesis, we use quasi-exogenous shocks in our passive ownership measure as the treatment in a difference-in-differences (DiD) regression. The shocks are index inclusions[21]. When a stock is included in the OSEBX, funds that track the index must purchase shares, and the passive ownership will increase substantially, as illustrated in Figure 3.3. According to our hypotheses, this should lead to a decrease in liquidity and price efficiency.

Such an approach requires for inclusions to be randomly assigned. The inclusions to the OSEBX are decided by a series of rules and criteria on a semi-annual basis and not by a committee directly, like the S&P 500. These criteria are related to liquidity, sector, market capitalisation and free float, with the objective of the index to best represent the Norwegian stock market. Consequently, the selection of index constituents is not randomly assigned. One could also argue that due to the size of the Norwegian market, the number of potential constituents is so small that it should be possible for market participants to predict which stocks that are going to be included.

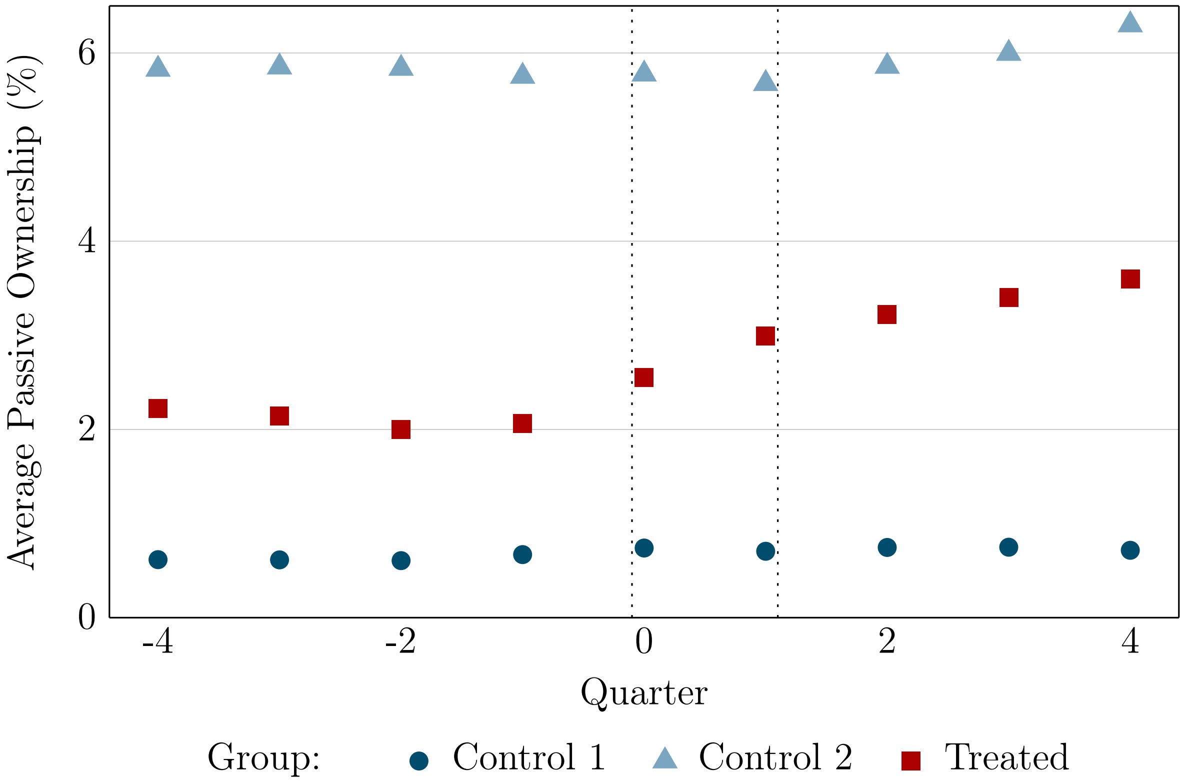
In order to deal with this, we construct two groups of control stocks in addition to the

[21] We only consider inclusions, as exclusions are likely to be based on fundamental aspects of the stock, such as poor performance or the lack of liquidity. Further, it is more difficult to construct reasonable control groups for such stocks.

treated group[22]. The first group (Control 1) contains stocks in the same sector as the treated stock and have a similar market capitalisation[23]. Further, stocks in this group can not be a part of the OSEBX at any time during the four-year period. These stocks are essentially stocks that we believe could have been included instead of the one that was actually included. This way, we can posit that within this group of stocks, inclusions are "as if" randomly assigned. The second group of control stocks (Control 2) are stocks that adhere to the same requirements but are members of the OSEBX during the entire four-year period. This period consists of a two-year period before, and a two-year period after the stock in the treated group was included to the index. The quarter of inclusion and the subsequent quarter are excluded from the sample to limit index inclusion effects, as discussed in Section 2.6.1.

Figure 3.3: Passive Ownership and Index Inclusions

This chart illustrates the average passive ownership of the quasi-experiment sample, segmented by group affiliation. The area between the two dotted lines are excluded from the sample.



[22] With "the treatment" being index inclusion.

[23] A treated stock can be assigned control stocks that are in the same sector while having a market capitalisation of no less than 50% and no more than 200% of the market capitalisation of the treated stock. This means that a treated stock with a NOK 10 billion MCap can be assigned control stocks ranging from NOK 5 billion to NOK 20 billion in MCap.

In total, we have 91 stocks in the treated group. The first control group contains 54 stocks, while the second contains 53 stocks. Figure 3.3 illustrates a 1.65 percentage points increase in passive ownership in the treated group between the two periods. This is expected, and serves as the basis for our natural experiment. At the same time, the two control groups have no specific trend in their passive ownership level during the periods. In line with Sammon (2020), we use the differences between the average of the first and the second period as our sample[24]. This way, the model compares the differences between the two periods for the two groups[25]. The DiD-estimator is a dummy variable indicating whether or not a stock is in the treated group. Consequently, the coefficient of the DiD-estimator will be the difference in the difference between the treated stocks and the control stocks over the two periods. Specifically, the regressions model is defined as:

$$\Delta Y_{i,d} = \alpha + \beta * DiD_i + \sum_{j=1}^{J} \beta_j * \Delta C_{i,j} + \sum_{s=1}^{S} \beta_s * SE_s + \sum_{m=1}^{M} \beta_m * MIE_m + \epsilon_{i,d} \quad (3.10)$$

where $\Delta Y_{i,d}$ is the change in the dependent variable d of stock i. DiD_i is a dummy variable indicating whether or not stock i has been included, and $\Delta C_{i,j}$ is the change in the control variable j for stock i. Finally, SE_s and MIE_m is the sector and month of index inclusion fixed effect, respectively.

3.4 Descriptive Statistics

In this last section of the chapter, we present the descriptive statistics of the data utilised throughout the analyses. First, we present simple statistics for all of the independent variables, including *Passive Ownership*. Second, we describe the dependent variables and illustrate their trends. Third, we describe the data used in the quasi-experiments.

Tables 3.1 and 3.2 provide information on the data used in the first part of the analyses. The statistics are split into two separate tables due to differences in the number of observations, as there are more earnings announcements than stock-years. All continuous change variables, such as Δ *Log MCAP* and Δ *Passive Ownership* are winsorised at the 1 and 99 per cent level to mitigate the influence of statistical outliers in the analyses.

[24]The averages mitigate the downward bias in the standard errors (Bertrand et al., 2004).
[25]The treated group includes treated stocks, and the non-treated group includes control groups 1 and 2.

Table 3.1: Descriptive Statistics for Annual Sample

This table provides the number of observations, means, standard deviations and distributions for the annual sample. The sample consists of pooled cross-sections in the 2001-2020 sample period. *Illiq* is the annual average of daily absolute returns divided by turnover in mNOK. *Spread* is the annual average of the monthly average spread based on log close, high and low prices. *Synch* is the log-transformed R^2 from a market model where stock returns are explained by market and sector returns. The Δ of these dependent variables is the difference between $year_t$ and $year_{t-1}$. *Passive Ownership* is defined as the sum of shares held by passive funds, divided by shares outstanding. *Active Ownership* is defined as the sum of shares held by active funds, divided by shares outstanding. *Log MCAP* is the natural logarithm of market capitalisation. Changes in these independent variables are calculated as end-of-year changes. *Volatility* is the annualised standard deviation of returns, while *Volume* is the average daily shares traded scaled by shares outstanding. *Beta* is the covariance of returns between a stock and OSEBX, divided by the variance of OSEBX. *Idiosyncratic Volatility* is the annual sum of the squared differences between daily returns of a stock and OSEBX. All change variables are winsorised at the 1% level, while Δ*Synch* is winsorised at the 5% level. The (%) indicates that the variable is expressed in percentage points.

	N	Mean	Std. Dev.	Min.	p25	Median	p75	Max.
Dependent Variables								
Illiq (%)	2432	6.41	16.66	0.00	0.15	0.93	4.55	108.19
Δ Illiq (%)	2218	0.22	8.28	-38.70	-0.34	0.00	0.51	38.87
Spread (%)	2432	1.23	1.16	0.00	0.56	0.97	1.60	22.72
Δ Spread (%)	2218	-0.08	1.07	-3.38	-0.60	-0.09	0.43	3.37
Synch	2104	-2.06	2.10	-6.91	-3.19	-1.85	-0.74	7.28
Δ Synch	1882	0.17	1.31	-2.32	-0.58	0.10	0.86	2.99
Passive Ownership Variables								
Passive Ownership (%)	2432	2.47	4.24	0.00	0.00	0.00	3.66	21.32
Δ Passive Ownership (%)	2218	0.22	1.11	-3.55	0.00	0.00	0.14	5.28
Control Variables								
Active Ownership (%)	2432	5.20	7.53	0.00	0.00	1.93	7.54	58.29
Δ Active Ownership (%)	2218	0.80	2.18	-4.49	0.00	0.00	1.18	11.06
Log MCAP	2432	21.54	1.66	19.34	20.17	21.34	22.64	27.01
Δ Log MCAP	2218	0.03	0.56	-1.73	-0.22	0.03	0.30	1.74
Volatility (%)	2418	38.13	15.80	0.00	26.86	35.78	48.95	117.33
Δ Volatility (%)	2196	1.01	13.56	-33.60	-6.69	-0.24	7.47	41.98
Volume (%)	2432	0.32	0.73	0.00	0.04	0.11	0.34	19.36
Δ Volume (%)	2218	0.01	0.37	-1.50	-0.04	-0.00	0.04	1.96
Beta	2235	0.67	0.53	-6.25	0.30	0.61	0.97	7.29
Δ Beta	2021	0.04	0.34	-0.85	-0.16	0.01	0.22	1.06
Idiosyncratic Volatility	2432	0.47	7.46	0.00	0.08	0.14	0.32	362.01
Δ Idiosyncratic Volatility	2218	0.01	0.34	-1.44	-0.05	-0.00	0.06	1.46

Table 3.2: Descriptive Statistics for Quarterly Sample

This table provides the number of observations, means, standard deviations and distributions for the quarterly sample. The sample consists of pooled cross-sections in the 2000-2020 sample period. *CAV* is the sum of daily volume over a period of one month before an earnings announcement, excluding the announcement day, divided by normal volume. *DM* is the cumulative return of the month before the announcement, excluding the announcement day, divided by the cumulative return of the same month including the announcement day. *QVS* is defined as the squared return on the announcement day, divided by the sum of squared returns for the full quarter. *Passive Ownership* is defined as the sum of shares held by passive funds, divided by shares outstanding. *Active Ownership* is defined as the sum of shares held by active funds, divided by shares outstanding. *Log MCAP* is the natural logarithm of market capitalisation. *Idiosyncratic Volatility* is the annual sum of the squared differences between daily returns of a stock and OSEBX. *Beta* is the covariance of returns between a stock and OSEBX, divided by the variance of OSEBX. All change variables are calculated as $quarter_t$ less $quarter_{t-4}$ as long as applicable, and winsorised at the 1% level. The (%) indicates that the variable is expressed in percentage points.

	N	Mean	Std. Dev.	Min.	p25	Median	p75	Max.
Dependent Variables								
CAV	4124	22.33	14.60	0.00	14.02	19.64	26.59	181.95
Δ CAV	4116	0.04	17.67	-60.90	-7.67	-0.02	7.69	60.96
DM	4125	96.21	4.02	56.50	94.78	97.28	98.92	105.76
Δ DM	4117	-0.12	4.75	-15.38	-2.41	-0.01	2.20	14.99
QVS (%)	4124	6.12	8.97	0.00	0.45	2.44	7.97	69.33
Δ QVS (%)	3354	-5.01	8.46	-40.50	-6.80	-1.47	0.50	1.00
Passive Ownership Variables								
Passive Ownership (%)	4125	3.72	4.65	0.00	0.00	1.03	7.19	21.00
Δ Passive Ownership (%)	4125	0.30	1.03	-2.61	0.00	0.00	0.52	4.55
Control Variables								
Active Ownership (%)	4125	5.50	7.13	0.00	0.03	2.94	8.39	49.46
Δ Active Ownership (%)	4125	0.75	1.90	-3.83	0.00	0.08	1.15	9.36
Log MCAP	4125	22.32	1.67	19.34	21.05	22.28	23.32	27.14
Δ Log MCAP	4125	0.07	0.55	-1.68	-0.19	0.06	0.32	1.87
Idiosyncratic Volatility	4125	0.58	11.28	0.01	0.07	0.13	0.26	362.37
Δ Idiosyncratic Volatility	4125	-0.00	0.25	-1.39	-0.04	0.00	0.05	0.95
Beta	4125	0.81	0.47	-0.65	0.46	0.76	1.10	3.00
Δ Beta	4125	0.03	0.29	-0.77	-0.14	0.01	0.18	0.94

The average *Passive Ownership* in the annual sample (Table 3.1) is 2.47%. This is slightly lower than in the samples of Hamm (2014) and Israeli et al. (2017), however, the variance in our data is larger. This is likely due to passive investing being more prevalent in U.S. stocks, especially in the first years of our sample, as illustrated in Figure 2.2. Another reason is that our analyses include relatively smaller firms. The mean of *Log MCAP* is roughly $250 million[26] in the annual data, which is about half of that in Hamm (2014) and Israeli et al. (2017). In the sample related to earnings announcements (Table 3.2) both market capitalisation and *Passive Ownership* are more similar to the existing literature. This confirms our notion that the deviation in the annual data is caused by market capitalisation. This is because large firms are over-represented in the event-based sample, relative to the annual sample. Both changes in *Log MCAP* and changes in *Passive Ownership* for both samples correspond roughly with the data used in the other two papers. The other independent variables, including *Active Ownership*, show little variation between the two samples.

Considering the dependent variables, the most notable deviations from prior literature are those of the liquidity proxies. The average *Spread* in our sample is 123 basis points, however, with a large standard deviation of 116 bps. This high standard deviation is also likely caused by the small-cap firms in our data set. The standard deviation for changes in *Spread* is about twice as large as the one in Israeli et al. (2017). As illustrated in Figure 3.4, *Spread* is larger for smaller firms. From the same figure, we see that *Spread* shows no sign of any long-term trend but is negatively correlated to economic cycles. The same interpretation is assumed for the illiquidity measure. Evident through Table 3.1 and Figure 3.4, small firms lead to an increase in the mean of the illiquidity measure in our sample. Further, the arithmetic mean of the illiquidity measure — as illustrated by the dark blue line in Figure 3.4 — reacts heavily to economic cycles, which could imply that the liquidity of small-cap firms is almost completely drained of liquidity in uncertain markets, such as during the GFC. We emphasise that the analyses will be run on different groups of firms sorted by market capitalisation, in order to control for the outliers in our liquidity measures.

[26] $(e^{21.54}) * USDNOK$, where $USDNOK$ is 9.00.

The price informativeness measures of Sammon (2020), *CAV*, *DM* and *QVS*, are provided in Table 3.2. The volume measure, *CAV*, has a mean of 22.3. This implies that the average volume over the month leading up to an earnings announcement date is not higher than the volume in the rest of the quarter[27]. Figure 3.5 illustrates no significant changes in the pre-earnings volume between the first and the second half of our observations, sorted by date. We would expect for the volume prior to the announcement to be lower in the second period, due to the overall increase in passive ownership. Further, the *DM* level in our sample is slightly lower than in Sammon's sample, which possibly could indicate that Norwegian stocks are priced less efficiently in the pre-earnings period. Evident in Figure 3.5, we show that for earnings announcements with positive news, the drift is significantly weaker in the second period, and the movement on the announcement date is stronger. For observations of negative announcements in the same figure, the trend is ambiguous.

The synchronicity measure has an arithmetic mean of -2.06, which is equivalent to an R^2 of about 0.115 in the market model described in Equation 3.7. Figure 3.4 illustrates how the synchronicity measure varies a great deal over the sample period, and that the value-weighted average of *Synch* is above 0 as opposed to -2.06. A *Synch* of 0 is equivalent to an R^2 of 0.50. Once again, we illustrate the large differences between different levels of market capitalisation. This is not surprising, and merely underlines the limited number of stocks in our sample and its large dispersion in size. Consequently, we believe for changes in *Log MCAP* to be a crucial control variable in our analyses, and possibly also for the effects of passive investing to differ between different size samples.

Lastly, we review the data for the second part of our analyses using the difference-in-differences approach which is provided in Table A3.1 in Appendix A3. This is a limited data set, as only a total of 91 stocks have been added to the index during our sample period. Further, our first control group of stocks that are similar but were not concluded consists of 54 stocks, and the second control group of stocks already included amount to 53 stocks. Unsurprisingly, changes in *Active Ownership* are larger in this sample, as, for instance, active funds with strict mandates could be restricted from holding stocks not included in the index. Changes in both *Volatility* and *Idiosyncratic Volatility* have negative means, while changes in *Beta* are larger than in the full annual sample.

[27] The trading volume in the 22-day period is scaled by the average volume in the rest of the quarter, so that a *CAV* of 22 implies zero abnormal volume.

The average changes in *Spread* and *Illiq* are more negative for the DiD sample than our for the annual sample. As we would expect, this indicates that OSEBX inclusion, ceteris paribus, has a positive effect on liquidity. Changes in *Synch* are more positive for this sample, at 0.26 versus 0.17 in the full sample. These trends are the reason why we apply control variables in our DiD-models. Lastly, the changes in the event-based measures generally correspond to those in the full sample. However, across all dependent variables, the variation is lower than in the quasi-experiment sample than in the full sample. This is probably due to the DiD sample being more homogeneous, i.e., less dispersion in market capitalisation, for instance.

Figure 3.4: Spread, Illiquidity and Synchronicity Measures

These charts illustrate the levels of *Spread*, *Illiq* and *Synch* over time. The arithmetic mean is indicated by dark blue lines, while light blue lines indicate market capitalisation weighted means.

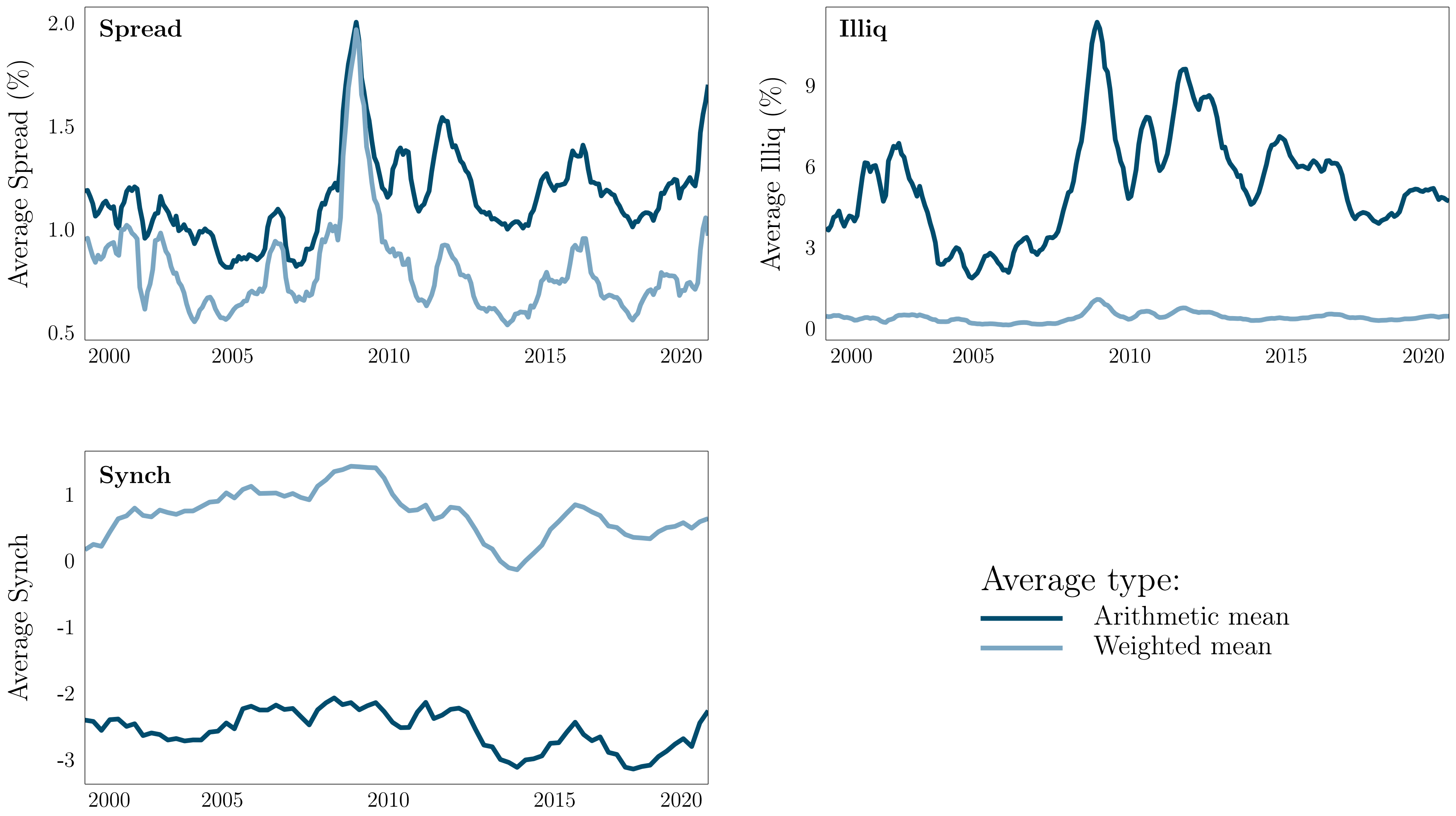
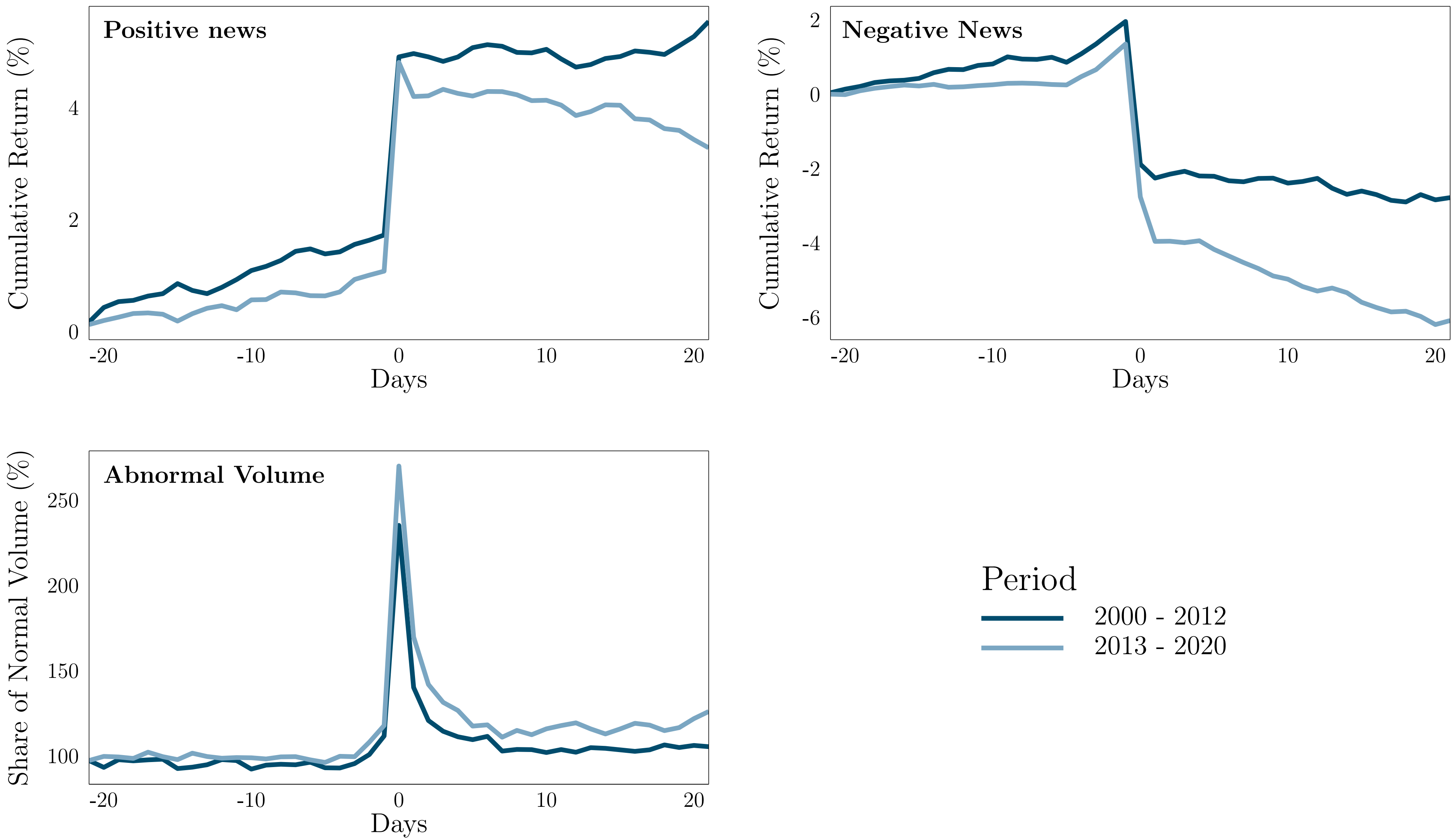


Figure 3.5: Return and Volume at Earnings Announcements

These charts illustrate the equally-weighted cumulative returns for positive and negative news on the earnings announcement day, as well as trading volume scaled by normal volume. The dark blue lines are based on observations from 2000 to 2012, while the light blue lines are based on observations from 2013 to 2020. With this segmentation, each period is based on the same amount of observations. The x-axes indicates the days relative to the announcement day.



4 Analysis

In this chapter, we present the results of our analyses. First, we review the Pearson's correlations of the variables used in the analyses. Second, we present the results on both hypotheses and discuss possible causes and implications.

4.1 Pearson's Correlations

Before analysing the hypotheses, we briefly review the Pearson's correlations between key variables. The correlation coefficients for the two samples (annual and quarterly) are provided in Tables 4.1 and 4.2. The correlations for the DiD-sample is provided in Table A3.2 in Appendix A3.

In general, we note that the Pearson's correlations are low. This is expected, as we use the first-differences design. Using such a design, we alleviate strong correlations due to sticky levels of certain variables. This also means that multicollinearity is likely not an issue in our analyses. While there is no general consensus regarding the level at which multicollinearity becomes an issue, coefficients below 0.7 are generally considered unproblematic (Booth et al., 1994). Additionally, the variance inflation factor (VIF) is never above 2 for any variables, with 5 commonly used as a threshold (Sheather, 2009). The VIFs are provided in Table A4.1 in Appendix A4.

Considering the relationships between changes in *Passive Ownership* and changes in our dependent variables in the annual sample, the correlations do not necessarily correspond with our hypotheses. This is the case for both liquidity proxies. However, we stress that the Pearson correlation between two variables should not be interpreted rigidly, especially when the expected relationship between the independent variable of interest and the dependent variable is small. An analysis of correlation coefficients does not allow for the use of control variables and is therefore likely to involve simultaneous biases. The Pearson's correlations between changes in *Passive Ownership* and changes in *Illiq* and *Spread* are -0.01 and 0.00, respectively. This is not necessarily surprising as the correlation between changes in *Log MCAP* and *Passive Ownership* intuitively is positive. Further, both changes in *Illiq* and *Spread* are negatively correlated with changes in *Log MCAP*, which is also intuitive. The reason for the ambiguous coefficients between changes in *Passive*

Ownership and changes in our liquidity measures could therefore be that stocks receive more inflows from passive vehicles as they grow larger, while at the same time experiencing improved liquidity. Changes in *Log MCAP* could also be the reason why changes in *Passive Ownership* is positively correlated to changes in *Synch*, a relationship consistent with our second hypothesis. This merely points out the importance of controlling for changes in *Log MCAP* in our analyses, which we do across all models.

The quarterly sample includes the three event-based proxies. From Table 4.2, we see that changes in *Passive Ownership* are negatively correlated to changes in pre-earnings trading (CAV). This is consistent with our second hypothesis. However, we see no signs of a similar relationship between changes in *Passive Ownership* and changes in the pre-earning drift (DM) or in the earnings-day volatility (QVS). Once again, we stress that the suggested relationship could be small and that there could be other factors interfering with the correlations. That being said, we note that changes in QVS and changes in DM display a high negative correlation. This is expected, as a weaker drift should indicate more earnings-day volatility. At the same time, changes in CAV is neither correlated to changes in DM nor changes in QVS. This is unexpected, and could indicate that the three measures do not measure price informativeness in the same fashion.

Lastly, we consider the Pearson's correlations for the DiD sample, which is provided in Table A3.2 in Appendix A3. Changes in *Spread* are negatively correlated with *OSEBX Inclusion*. From the literature on index inclusion effects (Petajisto, 2011), we know that there might be positive returns associated with index inclusions. Further, it is intuitive that spreads decrease as firms grow. This way, the negative relationship between changes in *Spread* and *OSEBX Inclusion* might be caused by both variables' correlation with changes in *Log MCAP*. The same arguments could be applied to the correlation between *OSEBX Inclusion* and changes in *Synch*, as a growing firm constitutes a larger part of both the market and the sector on which the R^2 is based. Therefore, we include changes in *Log MCAP* as a control variable in the regression models. The correlation between *OSEBX Inclusion* and changes in *Illiq* is unexpectedly negative, which we believe could be caused by the variable's large sensitivity to volume changes.

Table 4.1: Pearson's Correlation for Annual Sample

This table provides Pearson's correlations for the annual sample. The sample consists of pooled cross-sections in the 2001-2020 sample period. *Illiq* is the annual average of daily absolute returns divided by turnover in mNOK. *Spread* is the annual average of the monthly average spread based on log close, high and low prices. *Synch* is the log-transformed R^2 from a market model where stock returns are explained by market and sector returns. The Δ of these dependent variables is the difference between $year_t$ and $year_{t-1}$. *Passive Ownership* (PO) is defined as the sum of shares held by passive funds, divided by shares outstanding. *Active Ownership* (AO) is defined as the sum of shares held by active funds, divided by shares outstanding. *Log MCAP* is the natural logarithm of market capitalisation. Changes in these independent variables are calculated as end-of-year changes. *Volatility* is the annualised standard deviation of returns, while *Volume* is the average daily shares traded scaled by shares outstanding. *Beta* is the covariance of returns between a stock and OSEBX, divided by the variance of OSEBX. *Idiosyncratic Volatility* (IdVol) is the annual sum of the squared differences between daily returns of a stock and OSEBX. All change variables are winsorised at the 1% level. Δ*Synch* is winsorised at the 5% level.

	Illiq	Δ Illiq	Spread	Δ Spread	Synch	Δ Synch	PO	Δ PO	AO	Δ AO	Log MCAP	Δ Log MCAP	Volatility	Δ Volatility	Volume	Δ Volume	Beta	Δ Beta	IdVol	Δ IdVol
Illiq	1.00																			
Δ Illiq	0.34	1.00																		
Spread	0.16	0.08	1.00																	
Δ Spread	0.04	0.11	0.58	1.00																
Synch	-0.37	0.04	-0.17	0.00	1.00															
Δ Synch	-0.03	-0.11	-0.12	-0.04	0.33	1.00														
PO	-0.21	-0.01	-0.19	-0.02	0.47	-0.02	1.00													
Δ PO	-0.08	-0.01	-0.06	0.00	0.20	0.06	0.37	1.00												
AO	-0.19	-0.03	-0.20	-0.08	0.23	0.07	0.44	0.21	1.00											
Δ AO	-0.10	-0.06	-0.08	-0.04	0.07	0.00	0.18	0.12	0.56	1.00										
Log MCAP	-0.32	-0.07	-0.30	-0.05	0.68	-0.01	0.65	0.23	0.33	0.14	1.00									
Δ Log MCAP	-0.04	-0.17	-0.13	-0.21	-0.07	0.02	0.02	0.08	0.02	-0.01	0.20	1.00								
Volatility	0.05	0.01	0.34	0.03	0.09	0.02	-0.12	-0.03	-0.17	-0.09	-0.15	-0.13	1.00							
Δ Volatility	-0.01	0.02	0.12	0.19	0.04	0.16	-0.01	0.02	0.04	-0.00	-0.03	-0.20	0.42	1.00						
Volume	-0.14	-0.07	0.09	-0.01	0.15	0.04	0.01	0.02	-0.07	-0.03	-0.02	0.04	0.28	0.02	1.00					
Δ Volume	-0.02	-0.13	0.00	-0.03	0.00	0.07	-0.00	0.01	0.03	0.00	-0.00	0.06	0.09	0.10	0.43	1.00				
Beta	-0.30	-0.04	0.01	-0.01	0.65	0.19	0.25	0.11	0.08	0.03	0.34	-0.06	0.41	0.09	0.32	0.09	1.00			
Δ Beta	-0.04	-0.12	-0.01	-0.00	0.19	0.58	-0.02	0.04	0.06	0.03	0.02	0.08	0.11	0.29	0.05	0.16	0.34	1.00		
IdVol	0.00	-0.00	0.05	0.02	0.00	0.02	-0.03	-0.01	-0.03	-0.01	-0.04	-0.03	0.07	0.02	0.02	-0.02	0.03	0.05	1.00	
Δ IdVol	0.06	0.08	0.22	0.21	0.00	-0.07	-0.02	-0.04	-0.00	-0.03	-0.07	-0.19	0.18	0.34	0.11	0.32	0.03	0.04	0.13	1.00

Table 4.2: Pearson's Correlation for Quarterly Sample

This table provides Pearson's correlations for the quarterly sample. The sample consists of pooled cross-sections in the 2000-2020 sample period. *CAV* is the sum of daily volume over a period of one month before an earnings announcement, excluding the announcement day, divided by normal volume. *DM* is the cumulative return of the month before the announcement, excluding the announcement day, divided by the cumulative return of the same month including the announcement day. *QVS* is defined as the squared return on the announcement day, divided by the sum of squared returns for the full quarter. *Passive Ownership* (PO) is defined as the sum of shares held by passive funds, divided by shares outstanding. *Active Ownership* (AO) is defined as the sum of shares held by active funds, divided by shares outstanding. *Log MCAP* is the natural logarithm of market capitalisation. *Idiosyncratic Volatility* (IdVol) is the annual sum of the squared differences between daily returns of a stock and OSEBX. *Beta* is the covariance of returns between a stock and OSEBX, divided by the variance of OSEBX. All change variables are calculated as $quarter_t$ less $quarter_{t-4}$ as long as applicable, and winsorised at the 1% level.

	CAV	Δ CAV	DM	Δ DM	QVS	Δ QVS	PO	Δ PO	AO	Δ AO	Log MCAP	Δ Log MCAP	IdVol	Δ IdVol	Beta	Δ Beta
CAV	1.00															
Δ CAV	0.66	1.00														
DM	0.01	0.01	1.00													
Δ DM	-0.01	-0.01	0.64	1.00												
QVS	-0.02	-0.01	-0.59	-0.42	1.00											
Δ QVS	0.02	-0.01	0.10	-0.42	-0.21	1.00										
PO	-0.02	-0.01	0.03	0.01	0.19	-0.18	1.00									
Δ PO	-0.02	-0.04	0.02	0.00	0.06	-0.05	0.35	1.00								
AO	0.02	0.01	-0.01	-0.00	0.15	-0.18	0.44	0.24	1.00							
Δ AO	-0.01	-0.00	-0.04	-0.01	0.08	-0.12	0.20	0.08	0.55	1.00						
Log MCAP	0.02	-0.03	0.12	0.01	0.17	-0.16	0.62	0.24	0.25	0.10	1.00					
Δ Log MCAP	0.11	0.03	0.13	0.12	0.00	0.05	-0.02	0.07	-0.01	-0.07	0.13	1.00				
IdVol	-0.01	-0.05	-0.01	-0.00	-0.02	0.02	-0.03	-0.01	-0.03	-0.02	-0.05	-0.00	1.00			
Δ IdVol	0.06	0.06	-0.05	-0.00	-0.03	-0.02	0.01	-0.04	0.03	0.01	-0.06	-0.28	0.07	1.00		
Beta	0.02	-0.05	-0.17	-0.00	0.07	-0.09	0.17	0.11	0.02	0.03	0.27	-0.06	0.02	-0.02	1.00	
Δ Beta	-0.01	-0.07	-0.03	0.00	-0.01	-0.01	-0.03	0.02	0.05	0.02	-0.00	0.07	0.04	0.01	0.33	1.00

4.2 Main Results

In this section, we present the results of our analyses on both hypotheses. For each hypothesis, we give a brief summary of our most important findings, before providing a more in-depth examination of the specific models with interpretations of the different variables and their coefficients.

4.2.1 H1: Passive Ownership and Liquidity

Main Findings (Table 4.3 - 4.4). We first review the findings from our analysis of Hypothesis 1. We posit that an increase in passive ownership should negatively affect the liquidity of a stock. As discussed in Section 2.4, a shift from active to passive investing could imply a decrease in liquidity caused by the differing ways in which the two types of investors trade in the market, but also by their different investment horizon. On the contrary, an increase in passive ownership is also associated with an increase in the shares available for lending, which could imply an increase in liquidity. As proxies for liquidity, we use the price impact measure of Amihud (2002) and the spread estimator of Abdi and Ranaldo (2017).

As illustrated in Table 4.3, we find that increases in passive ownership are significantly associated with decreases in liquidity. The interpretation is the same using both proxies, and indicate that a 1% increase in passive ownership is associated with a 9.2 bps increase in the price impact of trading and a 3.1 bps increase in the spread. Applying the models to samples with stocks of different sizes in Tables A2.1 and A2.2 in the Appendices, we find that the effect on the illiquidity measure is considerably smaller for large firms. In general, our findings are consistent with our expectations and indicate that the negative effects of passive ownership on liquidity outweighs the benefits from increases in securities lending shown by Sørmo (2016). Our findings are larger than those of Hamm (2014) and Israeli et al. (2017), which again is consistent with the notion that the effects are larger for smaller and less liquid firms. Using the quasi-experiment setting in Table 4.4, the coefficient is similar to what we those identified using the full sample. However, they are not significantly different from zero and we can not infer a causal relationship, likely due to our limited number of observations.

Table 4.3: Liquidity and Passive Ownership

The regressions in this table depict the relationship between changes in two liquidity proxies and changes in passive ownership. The sample consists of 211 stocks and 2,196 or 2,218 stock-year observations in the period from 2001 to 2020. Models (2) and (4) include control variables and fixed effects for stock and year, while models (1) and (3) include no such elements. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses. Corresponding tables with the same models run on different subsamples are provided in Tables A2.1 and A2.2 in the Appendices.

	Dependent variable:			
	Δ Illiq (%)		Δ Spread (%)	
	(1)	(2)	(3)	(4)
Δ Passive Ownership (%)	0.010	0.092^{*}	0.010	0.031^{**}
	(0.019)	(0.052)	(0.018)	(0.012)
Δ Active Ownership (%)		-0.217^{***}		-0.013
		(0.078)		(0.009)
Δ Log MCAP		-1.797^{***}		-0.269^{***}
		(0.455)		(0.050)
Δ Volatility (%)		-0.037		0.007^{**}
		(0.025)		(0.003)
Δ Volume (%)		-3.105^{***}		-0.132
		(0.732)		(0.081)
Δ Idiosyncratic Volatility		1.855^{**}		0.542^{***}
		(0.819)		(0.138)
Stock FE	No	Yes	No	Yes
Year FE	No	Yes	No	Yes
Observations	2,218	2,196	2,218	2,196
Adjusted R^2	-0.105	-0.079	-0.105	-0.045
F Statistic	0.003	12.505^{***}	0.179	23.325^{***}
	(df = 1; 2006)	(df = 6; 1960)	(df = 1; 2006)	(df = 6; 1960)

Note: *p<0.1; **p<0.05; ***p<0.01

Findings in Correlational Study (Table 4.3). Column (1) and (2) in Table 4.3 use Amihud's illiquidity as the dependent variable, while column (3) and (4) use Abdi and Ranaldo's spread. In models (1) and (3), we exclude all control variables and fixed effects. In models (2) and (4), we control for stock and year fixed effects and include control variables.

In column (1), changes in *Passive Ownership* is positively correlated to changes in *Illiq*, however, the coefficient is not significant at any meaningful cut-off level. Explaining changes in the price impact of trading simply with changes in our measure of passive ownership seems unrealistic. As discussed, such a model likely involves issues with endogeneity. For instance, increased passive ownership is associated with increased market capitalisation, which in turn is negatively correlated to illiquidity. In column (2), we control for changes in market capitalisation and other variables nominated by the literature on liquidity. In line with Hypothesis 1, we now find a positive association between changes in *Passive Ownership* and *Illiq*, significant at the 10% level. Specifically, we find that a one percentage point increase in passive ownership is associated with a 9.2 bps increase in the average impact of every NOK 1 million traded. This represents a 1.4% increase from the sample's mean and is slightly lower than the findings of Israeli et al. (2017), who find a one percentage point increase in ETF ownership to be associated with a 2.1% increase in returns. As illustrated in Figure 3.4, smaller firms are relatively less liquid, especially when using the Amihud proxy. In Table A2.1 in Appendix A2, we provide the model in column (2) for six additional samples. By using only observations of the largest tercile of firms (column (6) in Table A2.1), the coefficient is about 1/3 of that with the full sample, meaning that a one percentage point increase is associated with a 2.7 bps increase in the price impact of every NOK 1 million traded[28]. The coefficient is significant at the 5% level.

Through the control variables we document that, as firms grow, they also tend to experience a decrease in the price impact of trading. This is likely due to several simultaneous effects such as the attraction of new investors due to recent returns, but also more long-term effects such as a more dispersed base of active investors leading to tighter intervals in

[28]The illiquidity measure is very sensitive to deviations in trading volume. Therefore, the absolute value of the coefficient — even the one for the largest 3^{rd} tercile of firms — is in no way applicable to a firm like Equinor, with an average daily trading volume of NOK 900 million during the sample period.

the order book. Increases in *Idiosyncratic Volatility* are associated with increases in *Illiq*, as uncertainty leads to less trading. This is as expected, however, the insignificant and negative link between changes in *Volatility* and changes in *Illiq* is surprising. The interpretation is that the price impact is not affected by changes in the overall volatility, when controlling for idiosyncratic volatility.

In columns (3) and (4) in Table 4.3, we use *Spread* as the proxy for liquidity. Overall, the results are similar to those discovered using *Illiq*. In column (3), we find no significant link between passive inflows and increases in *Spread*. While the coefficient is positive, it is not significantly different from zero. As with *Illiq*, the coefficient is statistically significant after the inclusion of fixed effects and control variables, and indicates that a one percentage point increase in passive ownership is associated with an increase in a stock's spread of 3.1 basis points. This is equivalent to a 2.5% increase from the mean[29]. This is higher than the coefficient of Israeli et al. (2017), who find a 1.6% increase from the mean, with only a slightly lower sample mean. As illustrated in Table A2.2 in Appendix A2, the effect is smaller for large firms (3^{rd} tercile, column (6)), however, the differences between the samples are not nearly as large as for the illiquidity measure. This makes sense when considering Figure 3.4, which shows a relatively small difference between the weighted and the arithmetic mean for the sample, meaning that the dispersion in spreads between stocks is relatively small. For the largest tercile of firms, a one percentage point increase in passive ownership is associated with a 2.8 bps increase in *Spread*.

Both changes in market capitalisation and annual volatility are as expected significantly associated with changes in a stock's spread. Specifically, growth in market capitalisation of 20%[30] is associated with a 4.8 basis points decrease in *Spread*. Further, an increase in *Volatility* of 5 percentage points is associated with an increase of 3.5 basis points in *Spread*. Changes in *Idiosyncratic Volatility* are also positively connected to changes in *Spread*. The interpretations of the control variables follow the ones laid forth in relation to the illiquidity measure.

[29]The average *Spread* for the full sample is 125 bps.
[30]Equivalent to 18% increase in the natural logarithm.

Table 4.4: Liquidity and OSEBX Inclusion

The regressions in this table depict the effects of OSEBX inclusion on two liquidity measures, using a Difference-in-Differences methodology. Models (2) and (4) include control variables and fixed effects for sector and month of inclusion. DiD is a dummy variable indicating treatment. The DiD coefficient is the effect of being included in the OSEBX in the period from 2002 to 2019, as opposed to two control groups of (1) similar stocks that are not included and (2) similar stocks that are included throughout the period. The sample consists of 185/198 differences in averages between a pre-period of two years excluding the inclusion quarter and a post-period of two years excluding the quarter following the inclusion. Standard errors clustered by sector and robust to heteroscedasticity are displayed in the parentheses.

	Dependent variable:			
	Δ Illiq (%)		Δ Spread (%)	
	(1)	(2)	(3)	(4)
DiD	0.060	0.137	−0.060	0.066
	(0.115)	(0.161)	(0.059)	(0.057)
Δ Active Ownership (%)		−0.076**		0.024***
		(0.034)		(0.006)
Δ Log MCAP		−0.434***		−0.083
		(0.125)		(0.071)
Δ Volatility (%)		0.011		0.007***
		(0.008)		(0.003)
Δ Volume (%)		−1.045		0.154
		(0.854)		(0.185)
Δ Idiosyncratic Volatility		−0.004		0.031***
		(0.015)		(0.010)
Sector FE	No	Yes	No	Yes
Index Month FE	No	Yes	No	Yes
Treated	91	91	84	84
Control 1	54	54	50	50
Control 2	53	53	51	51
Adjusted R^2	−0.005	0.215	−0.002	0.485
F Statistic	0.097	3.702***	0.580	10.118***
	(df = 1; 196)	(df = 20; 177)	(df = 1; 183)	(df = 19; 165)

Note: *p<0.1; **p<0.05; ***p<0.01

Findings in Quasi-Experiment (Table 4.4). Considering the results put forth in the first part of our analysis on the link between passive ownership and liquidity, it is important to note that correlation does not imply causation. Therefore, in this second part, we consider OSEBX inclusions as exogenous shocks in the passive ownership of a stock, and use these events as an independent variable in a quasi-experiment setting. The index in question, the OSEBX, includes firms mainly based on sector and market capitalisation (Oslo Stock Exchange, 2020). Consequently, we construct control groups based on these two criteria, and Figure 3.3 illustrates the differences in *Passive Ownership* between the groups in relation to index inclusion. We regress the differences in the average *Spread* and *Illiq* between the pre-inclusion and post-inclusion periods for the treatment group and the control groups. The results of the regressions are provided in Table 4.4, where columns (2) and (4) include control variables and fixed effects for sector and month of inclusion.

In Table 4.4, *DiD* is a dummy variable indicating whether or not an observation is in the treatment group. In columns (1) and (3) we show that the treated stocks experience increases in the illiquidity measure and decreases in the spread measure, relative to the control groups. However, as we believe there might be other variables affecting these proxies, we refrain from analysing these models in further detail. In column (2) and (4), we include fixed effects for sector and month of inclusion, as well as control variables. Now, the *DiD* coefficient is positive for both proxies, and corresponds roughly in size to the *Passive Ownership* coefficients of the first part of the analysis when considering the average increase in passive ownership[31] and the size of the firms in the sample[32]. That being said, none of the two liquidity proxies are significantly different from zero, meaning that we can not conclude on a causal relationship. Unfortunately, the data on index inclusions is not extensive, and a lack of observations is likely the reason why the coefficients are not significant[33]. Thus, we rely on the results of our correlational study and conclude that while we can not infer causality, there is strong evidence in our findings of a positive relationship between a stock's inflows of passive capital, and decreases in its liquidity.

[31]The difference in *Passive Ownership* between the pre-inclusion and post-inclusion periods is about 1.65 percentage points on average for the treated group.

[32]The firms in the quasi-experiment sample are of similar size to those in the 2^{nd} and 3^{rd} tercile of the full sample sorted by market capitalisation.

[33]91/84 treated; 54/50 in control group 1; 53/51 in control group 2.

4.2.2 H2: Passive Ownership and Price Informativeness

Main Findings (Table 4.5 - 4.8). We expect an increase in *Passive Ownership* to be associated with a decline in price informativeness due to three reasons. First, a migration of traders from the market for individual stocks to the market for indexed products would overall imply less trading on idiosyncratic information by the aggregate market. Second, an increase in the costs of trading is equivalent to an increase in the cost of information, which should also lead to less informed trading. Third, we expect the firm-specific component in prices to be negatively affected by an increase in the trading of stocks in large baskets. As proxies for price efficiency, we use the three event-based proxies of Sammon (2020), which relates to the reflection of earnings announcement information in stock prices. Further, we use the R^2 proxy suggested by Roll (1988) and formalised by Durnev et al. (2003), which relates to the synchronicity of stock returns.

With the event-based proxies in Table 4.5, we find ambiguous results. On the one hand, we document a significant relationship between increases in passive ownership and less pre-earnings trading (CAV), which is consistent with Sammon (2020) and indicates a decrease in the overall efforts of the market to trade on information in the pre-earnings period. On the other hand, we find neither a significant decrease in the pre-earning drift (DM) nor an increase in the volatility on earnings announcement days (QVS), which would be expected had the pre-earnings price efficiency dropped. These interpretations are robust to using different samples, as reported in Tables A2.3, A2.4 and A2.5 in the Appendices. In the quasi-experiment in Table 4.7, we document a significant decrease in the pre-earnings trading but find no such evidence for the other two proxies.

When using the R^2 (*Synch*) proxy in Table 4.6, we find a significant relationship between a stock's increases in *Passive Ownership* and increases in its returns' synchronicity with those of the market and its sector. In Tables A2.6 and A7.1 in the Appendices, we show that the effect is smaller for larger stocks. The directionality of the relationship is consistent with the existing literature but the effect seems smaller in our sample. One possibility is that the difference could be caused by the different types of passive vehicles analysed. In the second part of the analysis, we document a significant causal relationship. However, we believe that we might fail to control for increases in *Synch* at index inclusions that are not caused by passive investing.

Table 4.5: Price Informativeness and Passive Ownership

The regressions in this table depict the relationship between changes in three price efficiency proxies and changes in passive ownership. The sample consists of 153 or 172 stocks and 3,354 or 4,117 stock-year-quarter observations in the period from 2000 to 2020. Models (2), (4) and (6) include control variables and fixed effects for stock and year-quarter, while models (1), (3) and (5) include no such elements. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses. Corresponding tables with the same models run on different subsamples are provided in Tables A2.3, A2.4 and A2.5 in the Appendices.

	Dependent variable:					
	Δ CAV		Δ DM		Δ QVS (%)	
	(1)	(2)	(3)	(4)	(5)	(6)
Δ Passive Ownership (%)	−0.780***	−0.853***	−0.0001	−0.0002	−0.010	0.060
	(0.238)	(0.223)	(0.001)	(0.001)	(0.166)	(0.149)
Δ Active Ownership (%)		−0.002		0.0002		−0.239***
		(0.201)		(0.001)		(0.087)
Δ Log MCAP		1.887***		0.010***		0.324
		(0.732)		(0.002)		(0.264)
Δ Idiosyncratic Volatility		4.132**		0.010*		−0.378
		(1.886)		(0.005)		(0.457)
Δ Beta		−4.375***		−0.003		−0.021
		(1.067)		(0.003)		(0.468)
Stock FE	No	Yes	No	Yes	No	Yes
YearQuarter FE	No	Yes	No	Yes	No	Yes
Observations	4,116	4,116	4,117	4,117	3,354	3,354
Adjusted R^2	−0.042	−0.055	−0.044	−0.056	−0.048	−0.071
F Statistic	7.088***	8.466***	0.035	7.664***	0.005	2.150*
	(df = 1; 3943)	(df = 5; 3859)	(df = 1; 3944)	(df = 5; 3860)	(df = 1; 3200)	(df = 5; 3119)
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 4.6: Return Synchronicity and Passive Ownership

The regressions in this table depict the relationship between changes a return synchronicity measure and changes in passive ownership. The sample consists of 197 stocks and 1,882 stock-year observations in the period from 2000 to 2020. Models (2) and (4) include control variables and fixed effects for stock and year, while models (1) and (3) include no such elements. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses. Corresponding tables with the same models run on different subsamples are provided in Table A2.6 in the Appendices.

	Dependent variable:	
	Δ Synch	
	(1)	(2)
Δ Passive Ownership (%)	0.080***	0.039***
	(0.021)	(0.014)
Δ Active Ownership (%)		0.005
		(0.010)
Δ Log MCAP		0.251***
		(0.062)
Δ Beta		2.074***
		(0.110)
Stock FE	No	Yes
Year FE	No	Yes
Observations	1,882	1,882
Adjusted R^2	−0.112	0.267
F Statistic	8.036*** (df = 1; 1684)	225.705*** (df = 4; 1662)
Note:		*p<0.1; **p<0.05; ***p<0.01

Findings in Correlational Study (Table 4.5 - 4.6). The results of our regressions using the event-based proxies and the synchronicity proxy are provided in Tables 4.5 and 4.6, respectively. For all proxies, the first column provides the relationship without control variables nor fixed effects, while the second column includes these elements. First, we interpret the results of the regressions on the event-based proxies.

The first proxy is the pre-earnings volume (*CAV*). Column (1) in Table 4.5 shows that changes in a stock's passive ownership are negatively associated with changes in abnormal trading volume in the month leading up to its earnings announcements. In column (2), the coefficient is slightly larger, suggesting that a one percentage point increase in *Passive Ownership* is associated with a 3.7% decrease in *CAV*. This is equivalent to a loss of 0.85 trading days of volume during the 22-day estimation window. The directionality of our coefficient is consistent with the findings of (Sammon, 2020) on U.S. stocks, however, the coefficient itself is about twice as large. Table A2.3 in the Appendices shows that the effect is smaller for the largest tercile of stocks, indicating a loss of 0.54 trading days. Further, growth in market capitalisation is associated with more pre-earnings trading. This could be explained by the tendency of growing firms to experience an increased interest from analysts and the market as a whole. This means that the efforts of the market to analyse information ahead of the announcement intensify as firms grow. Positive changes in *Beta* are associated with a decrease in *CAV*, which could be caused by the overall market focusing less on firm-specific information and more on macroeconomic factors. The opposite interpretation is assumed for the positive coefficient of changes in *Idiosyncratic Volatility*.

The second proxy is the drift measure (*DM*) used in column (3) and (4) of Table 4.5. This is the ratio of the return in the month of the announcement excluding the announcement day, to the total return of the period including the announcement day[34]. Consequently, it measures the extent to which the information presented was embedded in the price during the month leading up to the announcement. Column (3) and (4) show a negative, but small and insignificant relationship between changes in *Passive Ownership* and changes in *DM*. Therefore, we have no indication of a relationship between the two variables. Despite this, we do find a significant relationship between increases in *Log MCAP* and increases

[34] For announcements with negative returns, the ratio is inverted, in order for the interpretation of the variable to remain the same (lower values of *DM* indicate less informative prices and conversely).

in *DM*. The reason for this relationship is likely similar to the interpretation for *CAV*, i.e., that when firms grow their performance is increasingly analysed by various market participants.

The third measure is the earnings announcement day's share of the quarterly volatility of a stock (*QVS*). This measure builds on the two previous ones. That is, with less informative prices in the period leading up to the announcement, we would expect more volatility on the announcement day. The results are provided in column (5) and (6) of Table 4.5. The coefficient of 0.06 indicates that an increase of one percentage point in *Passive Ownership* is associated with an increase of 0.06 percentage points in *QVS*. This is an increase of about 1% relative to the sample mean of 6.12%. However, the effect is not significant at any of the cut-off levels. For this particular measure, changes in *Active Ownership* is significantly associated to changes price informativeness, and exhibits a negative coefficient. The explanation for such a relationship could be that sophisticated active investors rather trade on privately gathered information prior to the announcement than on the specific announcement day.

The fourth proxy of price informativeness relates to the comovement of a stock's returns to that of the market and its sector. As prices contain less firm-specific information, the correlation with its sector and the market increases (Roll, 1988). Durnev et al. (2003) find that stocks with low return correlations with the market and its sector display higher correlation with future earnings. We follow Durnev et al. (2003) and use as our fourth measure the transformed R^2 of a market model regression, as described in Equation 3.7, with sector and market returns as explanatory variables for individual stock returns.

The results from the regressions on the synchronicity measure are illustrated in Table 4.6. In column (1), we run the regression without fixed effects and control variables. We document a positive relationship between increases in *Synch* and inflows into passive vehicles. The relationship is significant at the 1% level. When including control variables and fixed effects in column (2), the coefficient decreases in size but remains significant at the same level. Changes in *Log MCAP* is an important control variable, as growing firms constitute a growing share of the total market and their sectors, intuitively resulting in a larger R^2 in the regression model upon which the synchronicity measure is based. Further, we find no significant relationship between increases in *Synch* and increases in the *Active*

Ownership. This suggests that the relationship of fund ownership with changes in *Synch* is mainly related to passive investing. Lastly, we control for changes in *Beta*, essentially allowing for stocks with high betas to move more in tandem with the market (Li et al., 2014). The positive association between changes in *Beta* and changes in *Synch* is not surprising and highly significant. Controlling for these effects, we find that a 1 percentage point increase in *Passive Ownership* is associated with a 0.039 increase in *Synch*. This is equivalent to an increase of about 0.4 percentage points in the R^2 of the market model regression. In Table A2.6, the coefficient for the sample of large stocks is about 1/3 of that in the full sample. Despite this, the effects are only slightly smaller for the largest tercile of firms. This is due to a non-linear relationship between the synchronicity measure and the R^2 from the regression[35].

In line with our hypothesis, these results indicate that an increase in passive ownership is associated with a decrease in the firm-specific information component in a stock's price. Such a relationship is economically significant in several ways. First, it implies that there could be a link between increases in passive investing and a less functioning market, i.e., a market in which capital is not efficiently allocated between firms. Second, a market with less firm-specific information is equivalent to a more risky market for investors, as the effects of diversification would be expected to decrease as stocks move more in tandem with the market. In Table A5.1 in the Appendices, we document a significant relationship between increases in a stock's passive ownership and increases in its equally-weighted correlation with other stocks. That being said, less informationally efficient prices should also create opportunities for active investors, as the benefits of trading on firm-specific information should increase as fewer investors choose to do so. If index investors end up holding more and more misvalued stocks, active managers should focus on identifying these stocks with misvalued fundamentals, instead of attempting to mimic the index strategy (also referred to as "closet indexing"). By profiting from such inefficiencies, the active style should increase in popularity, and the aggregate market should at some point in the future reach an equilibrium where prices are just sufficiently inefficient for the two styles to coexist (Grossman and Stiglitz, 1980; Gârleanu and Pedersen, 2018).

[35] The coefficient of 0.014 for the largest tercile indicates an increase in the R^2 of about 0.3 percentage points. The reason for the different interpretation is that the largest stocks have a higher mean of *Synch*. An increase in *Synch* from a higher level is equivalent to a larger increase in the R^2. The relationship between changes in R^2 and changes in *Synch* is illustrated in Table A7.1 in the Appendices.

A potentially omitted variable in the relationship could be the aggregate market's emphasis on systematic, rather than idiosyncratic risk. Specifically, the rise in passive investing could be caused by the market deeming systematic factors more important than idiosyncratic factors. In this case, our findings would be expected, but our second hypothesis would still be wrong. This would essentially imply that despite the indications of the existing literature (Bramante et al., 2015; Durnev et al., 2003; Roll, 1988), increases in R^2 is not equivalent to a decrease in price efficiency. In Figure A6.1 in the Appendices, we segment overall volatility into three component parts: (1) market-specific, (2) sector-specific and (3) firm-specific volatility, using the methodology of Campbell et al. (2001). We do not identify trends confirming either a decrease in idiosyncratic volatility or an increase in systematic volatility. Consequently, we believe that a shift in the market's focus on systematic information seems unlikely. Furthermore, we believe that such an effect caused by passive investing is unlikely in the Norwegian market, considering the relatively limited prevalence of ETFs focused on Norwegian macro factors.

Table 4.7: Price Informativeness and OSEBX Inclusion

The regressions in this table depict the effects of OSEBX inclusion on three event-based measures, using a Difference-in-Differences methodology. Models (2), (4) and (6) include control variables and fixed effects for sector and month of inclusion. DiD is a dummy variable indicating treatment. The DiD coefficient is the effect of being included in the OSEBX in the period from 2003 to 2019, as opposed to two control groups of (1) similar stocks that are not included and (2) similar stocks that are included throughout the period. The sample consists of 120 differences in averages between a pre-period of two years excluding the inclusion quarter and a post-period of two years excluding the quarter following the inclusion. Standard errors clustered by sector and robust to heteroscedasticity are displayed in the parentheses.

	Dependent variable:					
	Δ CAV		Δ DM		Δ QVS (%)	
	(1)	(2)	(3)	(4)	(5)	(6)
DiD	−2.113	−2.248*	0.011	0.011	0.409	0.769
	(1.869)	(1.321)	(0.008)	(0.010)	(0.943)	(0.651)
Δ Active Ownership (%)		−0.101		−0.001		0.384***
		(0.370)		(0.001)		(0.080)
Δ Log MCAP		1.546		0.004		0.620***
		(1.587)		(0.012)		(0.191)
Δ Idiosyncratic Volatility		1.781***		−0.002		0.171
		(0.485)		(0.003)		(0.252)
Δ Beta		1.977		−0.028**		1.968
		(2.202)		(0.011)		(2.054)
Sector FE	No	Yes	No	Yes	No	Yes
Index Month FE	No	Yes	No	Yes	No	Yes
Treated	51	51	51	51	51	51
Control 1	28	28	28	28	28	28
Control 2	41	41	41	41	41	41
Adjusted R^2	0.013	0.090	0.013	0.216	−0.007	0.126
F Statistic	2.577	1.732*	2.614	3.054***	0.196	2.071**
	(df = 1; 118)	(df = 16; 103)	(df = 1; 118)	(df = 16; 103)	(df = 1; 118)	(df = 16; 103)

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.8: Return Synchronicity and OSEBX Inclusion

The regressions in this table depict the effects of OSEBX inclusion on a return synchronicity measure, using a Difference-in-Differences methodology. Models (2) and (4) include control variables and fixed effects for sector and month of inclusion. DiD is a dummy variable indicating treatment. The DiD coefficient is the effect of being included in the OSEBX in the period from 2002 to 2019, as opposed to two control groups of (1) similar stocks that are not included and (2) similar stocks that are included throughout the period. The sample consists of 197 differences in averages between a pre-period of two years excluding the inclusion quarter and a post-period of two years excluding the quarter following the inclusion. Standard errors clustered by sector and robust to heteroscedasticity are displayed in the parentheses.

	Dependent variable:	
	Δ Synch	
	(1)	(2)
DiD	0.670***	0.319**
	(0.125)	(0.154)
Δ Active Ownership (%)		0.036**
		(0.015)
Δ Log MCAP		0.725***
		(0.094)
Δ Beta		1.637***
		(0.633)
Sector FE	No	Yes
Index Month FE	No	Yes
Treated	90	90
Control 1	54	54
Control 2	53	53
Adjusted R^2	0.068	0.504
F Statistic	15.261*** (df = 1; 195)	12.077*** (df = 18; 178)
Note:		*p<0.1; **p<0.05; ***p<0.01

Findings in Quasi-Experiment (Table 4.7 - 4.8). Using the same specifications as previously, we regress changes in the event-based proxies and the synchronicity measure on the treatment dummy indicating whether or not the stock has been added to the OSEBX. The results of the regressions are provided in Tables 4.7 and 4.8, respectively. Both the models using CAV and QVS as proxies display larger coefficients, however, in the same direction as in the correlational analyses. The reason could be that we underestimate the true increase in passive ownership, and that index inclusions also lead to increases in the investments of implicit indexers, as discussed in Section 2.3.2. Column (2) indicates a significant relationship between quasi-exogenous increases in passive ownership and decreases in CAV. More specifically, the increase in passive ownership induced by OSEBX inclusion leads to a decrease in the pre-earnings trading volume equivalent to 2.25 normal trading days. The coefficient for DM, on the other hand, now has a positive sign. This is not expected, as it indicates that the effect of index inclusion is a more precise incorporation of information prior to the announcement. This relationship could for instance be caused by increases in the number of analyst covering a firm after inclusion, which we do not control for. We emphasise that neither the drift measure nor the quadratic variation share provides statistically significant results, making the interpretation uncertain.

One explanation for the differences between our results and the ones of Sammon (2020) could be that index investing in Norway is still too limited for us to document similar effects to those in the U.S. market. Despite this, our regressions on liquidity strongly suggest a correlation between increases in *Passive Ownership* and reductions in liquidity. Consequently, the question becomes whether or not there is a link between liquidity and price efficiency. Either no such link exists, or there are structural differences in earnings announcements between the two countries making these measures less credible for Norwegian stocks. For instance, we note that both the average drift and pre-earnings trading in U.S. stocks are higher than for Norwegian stocks, indicating that U.S. stock prices are more informative. At the same time, the QVS of U.S. stocks is higher than that of Norwegian stocks, indicating the contrary. Adding to this, as discussed in Section 4.1, the correlations between the proxies themselves also indicate inconsistencies.

Lastly, we apply our quasi-experiment setting to the synchronicity measure. As illustrated in Table 4.8, the results are larger than reported in the correlational analysis. In column

(1), we find a significant and positive relationship between *DiD* and changes in *Synch*. We exclude the quarter of inclusion and the subsequent quarter, in order to limit index inclusion effects. However, as we know from the literature, there is likely a positive return effect from being included in an index (Petajisto, 2011). If the higher level of *Log MCAP* is sustained beyond the excluded quarters, changes in *Log MCAP* becomes a crucial control variable. This is because growing firms constitute a growing part of both the market and the sector. We also control for active ownership, as many active investors could have mandates restricting investments outside the OSEBX. Lastly, we control for changes in *Beta*, following the arguments laid forth in the correlational analysis. By controlling for these effects, the *DiD* coefficient is significant at the 5% level and indicates that the increase in passive ownership from OSEBX inclusion leads to an increase in *Synch* of 0.319. This is equivalent to an increase in the market model R^2 of 3.5 percentage points from the mean, which is substantially higher than in the correlational analysis when accounting for the 1.65 percentage points average increase in passive ownership at inclusion.

There are two possible explanations for this. First, our proxy for passive ownership could underestimate the true increase in passive ownership at index inclusions. This is likely not the only reason, as the deviations in our other analyses are considerably smaller. A second reason is that there could be other factors than passive investing affecting the synchronicity at inclusions. On the one hand, the traditional view is that comovement within indices is a result of comovement in fundamental factors. In this view, synchronicity in indices could stem from factors such as more synchronised discount rates being applied to index constituents. On the other hand, a more novel view proposed by Barberis et al. (2005)[36] is that comovement occurs as a result of trading commonality, as specific indices serve as the preferred habitat for certain investors, and that some investment products are linked to specific indices. This relates to passive investing, but also to other investors with the OSEBX as their preferred habitat. Consequently, we can not infer with certainty that passive investing is the only reason for the increase in return synchronicity at inclusion.

We believe that the trading of passive investors causes a great deal of the synchronicity induced by index inclusion. However, we acknowledge that other unobserved factors might interfere with our interpretation in the quasi-experiment setting. On the basis of this, we can not unreservedly infer a causal relationship.

[36]This is referred to as the "habitat" view of comovement.

5 Conclusions

In this final chapter, we first provide our conclusions. Second, we emphasise the most important limitations of our analyses and findings. Lastly, we provide a few suggestions for future research on the impact of passive investing.

5.1 Conclusions

The objective of this thesis is to investigate the impact of passive investing on the liquidity and price efficiency of Norwegian stocks. First, we posit that passive ownership has a negative influence on liquidity, as more shares are held in long-term deposits and more trading is based on price-inelastic demand. Thus, our first hypothesis is:

H1: ***"Does passive investing lead to reduced liquidity in the Norwegian stock market?"***

The results of our analyses on the first hypothesis indicate a significant relationship between increases in the passive ownership of a stock, and decreases in its liquidity. The relationship is robust to using various subsamples, however, the effect seems stronger for the stocks of smaller firms. In general, our findings are consistent with the existing literature, which is exclusively conducted on ETFs in the U.S. (Hamm, 2014; Israeli et al., 2017). That being said, the negative effect of passive ownership on liquidity seems larger in Norway. This could be due to the low free float in Norwegian stocks, or that the passive vehicles present in Norwegian stocks trade less frequently[37].

Further, we claim that the negative impact on liquidity might lead to less trading by informed investors, and thus, less informative prices. Basket trading by index funds and ETFs could amplify this effect. Our second hypothesis is stated as follows:

H2: ***"Does passive investing lead to less informative prices in the Norwegian stock market?"***

In our analyses on the second hypothesis, we use two different types of proxies for price informativeness. With proxies related to earnings announcements, we find somewhat mixed results. Utilising more data through the fourth proxy, we find a statistically significant

[37]In the Norwegian market, ETFs constitute a relatively small part of the total passive ownership.

relationship between increases in passive ownership and decreases in the firm-specific information component of stock prices. This finding is consistent with less trading on idiosyncratic information by informed market participants, with the result being less informative prices. While the directionality of our findings is consistent with the existing literature (Israeli et al., 2017; Sammon, 2020; Qin and Singal, 2015), it seems the negative effect of passive ownership on price efficiency is smaller in Norwegian stocks. This could indicate that the passive investor's trading of stocks in baskets is a more important channel in terms of impact on price efficiency than the long-term holding of shares[38].

In the analysis on liquidity, we fail to establish causality using index inclusions as quasi-exogenous shocks in passive ownership. While the coefficients and their signs were similar to those in the correlational analyses, they were consistently insignificant, likely due to a limited number of observations. Consequently, we are not able to conclude on a cause-effect relationship. In the analyses on price efficiency, we document a significant relationship between index inclusions and increases in return synchronicity. Still, we can not unreservedly infer causality between increases in passive investing and decreases in synchronicity, as we suspect there might be unobservable factors interfering with the relationship.

That being said, we believe our findings are of relevance for both passive and active investors, regulators and commentators. Passive investing is largely based on the notion that prices are efficient due to the efforts of active investors, and that passive vehicles do not affect the markets in which they act. Active investors provide liquidity and contribute to price efficiency directly through their trading. Passive investors, on the other hand, remove liquidity and do not contribute to making prices efficient. This thesis provides empirical evidence of both these externalities of passive investing. As passive investing continues to gain popularity, these negative externalities will continue to affect the structure of the stock market. Inevitably, we believe this will lead to opportunities emerging for active investors, and that there is some form of equilibrium level of efficiency where both styles can coexist and profit.

[38] In general, index ETFs are associated with more trading and less long-term holding of shares, relative to index mutual funds and the FTF.

5.2 Limitations

In the following sections, we discuss the limitations of our thesis, and how these might affect our findings and conclusions.

First of all, there are a few limitations to our samples. The Oslo Stock Exchange is relatively small with an average of 221 stocks listed during our sample period. Further, our fund ownership data only allows for the analysis of currently listed stocks. While we do not believe that this leads to a biased sample, it severely reduces the number of observations. After excluding observations of stocks with a market capitalisation below $27.8 million[39], the final sample is small relative to comparable studies conducted on U.S. stocks[40]. This becomes especially problematic in the quasi-experiment, and we believe that the size of our sample is the primary reason for our inability to establish causality. We deem the full sample sufficiently large, however, it contains stocks that deviate greatly from one another in factors such as size, liquidity, volatility and passive ownership. Ideally, a larger sample could have allowed for analyses conducted on more specific subsamples, which in turn could have led to a more in-depth understanding of the many mechanisms in play.

Further, we are unable to analyse the directionality of the proposed mispricing introduced by passive ownership and trading. This is mainly due to a lack of easily accessible and complete accounting data for Norwegian stocks. Durnev et al. (2003) show that stocks with a low market model R^2 have a higher R^2 in a model where stock returns are explained by future earnings. We are unable to test this due to the absence of accounting data and must therefore rely on the findings of the existing literature.

There are also a few limitations to our passive ownership proxy. Several studies have indicated that many active managers follow a strategy similar to that of index funds, but with a small number of active bets (Cremers and Petajisto, 2009; Thoresen and Øren, 2017). Moreover, there is no reason to believe that other institutional investors have not also tilted towards index investing in recent years. We fail to account for such implicit indexing in our analysis, which means that our specific findings may not be generalisable to

[39] $27.8 million is equal to NOK 250 million at a USDNOK of 9.00.

[40] After excluding both delisted stocks and stocks with a market capitalisation below $27.8 million, our final sample consists of 121 stocks on average.

the true increase in passive ownership. In addition to this, we measure passive ownership as the number of shares held by passive funds over the total number of shares outstanding. In our analysis on liquidity, we would ideally have used the free float as the denominator. The implication is that we are likely to underestimate the effects of passive investing on liquidity in stocks with a below-average free float share, and conversely.

Lastly, in the quasi-experiment, there are always concerns regarding violations of underlying assumptions. For our analyses, this mainly relates to index inclusion effects. Specifically, if other index inclusion effects than increases in passive ownership are affecting our dependent variables, we could find biased estimates. In order to deal with this, we include control variables nominated by both the literature on our dependent variables and on index inclusion effects in general. Further, we exclude both the quarter of inclusion and the subsequent quarter. However, we acknowledge that there might still be other effects arising from index inclusions that we fail to control for.

5.3 Proposals for Future Research

We encourage more research on the impacts of passive investing in global equity markets. That being said, passive investing is also becoming more prevalent in fixed income markets[41]. There is limited research conducted on the role of the passive investor in such markets. Bond indices are also capitalisation-weighted, meaning that passive investors are more exposed to the most indebted firms. Further, we believe the incentive mechanisms of passive inflows in bonds are interesting, and could serve as the topic for future research.

We also encourage more research on the impacts of passive investing in the Norwegian equity market. The inclusion of accounting data could allow for analyses on valuations. For instance, one could ask whether or not the flows from passive investors, given that they are based on non-fundamental factors, are associated with inflated prices over some term. This could provide further insights into the efficiency of the market in light of the ongoing shift from active to passive investing. That being said, we emphasise that analyses on the Norwegian market will likely benefit from longer time series and larger passive ownership shares, meaning that researchers should consider waiting a few years

[41] To our knowledge, this trend is not present in the Norwegian bond market, mainly due to liquidity and transparency reasons.

before conducting their analyses.

Lastly, researchers could examine a possible equilibrium between passive and active investment styles, with the potential impacts of passive investing in mind. We believe that passive investing could become overcrowded at some point in the future, and that this will lead to opportunities of higher returns for active managers. The when and the how of such a turning point could be an interesting topic for financial researchers.

References

Abdi, F. and Ranaldo, A. (2017). A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices. *The Review of Financial Studies*, 30(12):4437–4480.

Albuquerque, R. (2012). Skewness in Stock Returns: Reconciling the Evidence on Firm Versus Aggregate Returns. *The Review of Financial Studies*, 25(5):1630–1673.

Amihud, Y. (2002). Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. *Journal of Financial Markets*, 5(1):31–56.

Barberis, N., Shleifer, A., and Wurgler, J. (2005). Comovement. *Journal of Financial Economics*, 75(2):283–317.

Ben-David, I., Franzoni, F., and Moussawi, R. (2018). Do ETFs Increase Volatility? *The Journal of Finance*, 73(6):2471–2535.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1):249–275.

Bessembinder, H. (2018). Do Stocks Outperform Treasury Bills? *Journal of Financial Economics*, 129(3):440–457.

Bessembinder, H., Chen, T.-F., Choi, G., and Wei, K. C. J. (2019). Do Global Stocks Outperform US Treasury Bills? *Unpublished*.

BlackRock (2017). Index Investing Supports Vibrant Capital Markets. *Unpublished*.

Booth, G. D., Niccolucci, M. J., and Schuster, E. G. (1994). Identifying Proxy Sets in Multiple Linear Regression: An Aid to Better Coefficient Interpretation. *Unpublished*.

Bramante, R., Petrella, G., and Zappa, D. (2015). On the Use of the Market Model R-Square as a Measure of Stock Price Efficiency. *Review of Quantitative Finance and Accounting*, 44:379–391.

Campbell, J. Y., Lettau, M., Malkiel, B. G., and Xu, Y. (2001). Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk. *Journal of Finance*, 56(1):1–43.

Chan, K., Chan, Y.-C., and Fong, W.-M. (2004). Free Float and Market Liquidity: A Study of Hong Kong Government Intervention. *The Journal of Financial Research*, 27(2):179–197.

Cremers, M. and Petajisto, A. (2009). How Active Is Your Fund Manager? *The Review of Financial Studies*, 22(9):3329–3365.

Da, Z. and Shive, S. (2018). Exchange Traded Funds and Asset Return Correlations. *European Financial Management*, 24(1):136–168.

Dans, E. (2020). What Tesla's Entry Onto The S&P 500 Tells Us About The Global Economy. Retrieved 16.12.2020 from: https://www.forbes.com/sites/enriquedans/2020/11/19/what-teslas-entry-onto-the-sp-500-tells-us-about-the-globaleconomy/?sh=1e719e6c5131.

Ding, X., Ni, Y., and Zhong, L. (2016). Free Float and Market Liquidity Around the World. *The Journal of Empirical Finance*, 38(A):236–257.

Durnev, A., Morck, R., Yeung, B., and Zarowin, P. (2003). Does Greater Firm-Specific Return Variation Mean More or Less Informed Stock Pricing? *Journal of Accounting Research*, 41(5):797–836.

Easton, P., Gao, G., and Gao, P. (1984). Pre-Earning Announcement Drift. *Unpublished*.

Embu, J. (2020). Interview On The Strategy of KLP. *Personal communication*.

Evans, R., Ferreira, M. A., and Prado, M. P. (2017). Fund Performance and Equity Lending: Why Lend What You Can Sell? *European Finance Association*, 21(3):1093–1121.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of the American Statistical Association*, 25(2):383–417.

Fama, E. F. and French, K. R. (2018). Long-Horizon Returns. *The Review of Asset Pricing Studies*, 8(2):232–252.

Finansdepartementet (2010). Mandat for forvaltningen av Statens pensjonsfond Norge. Retrieved from: https://lovdata.no/dokument/INS/forskrift/2010-12-21-1790?q=mandat%20spn/.

Foster, G., Olsen, C., and Shevlin, T. (1984). Earnings Releases, Anomalies, and the Behaviour of Security Returns. *The Accounting Review*, 59(4):574–603.

Fox, J. and Monette, G. (1992). Generalized Collinearity Diagnostics. *Journal of the American Statistical Association*, 87(417):178–183.

Glosten, L., Nallareddy, S., and Zou, Y. (2016). ETF Activity and Informational Efficiency of Underlying Securities. *Management Science, forthcoming*.

Glosten, L. R. (1994). Is the Electronic Open Limit Order Book Inevitable? *The Journal of Finance*, 49(4):1127–1161.

Glosten, L. R. and Harris, L. E. (1988). Estimating the Components of the Bid/Ask Spread. *Journal of Financial Economics*, 21(1):123–142.

Gorton, G. and Pennacchi, G. (1993). Security Baskets and Index-Linked Securities. *The Journal of Business*, 66(1):1–27.

Goyenko, R., Holden, C., and Trzinka, C. (2009). Do Liquidity Measures Measure Liquidity. *Journal of Financial Economics*, 92(2):153–181.

Grossman, S. J. and Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, 70(3):393–408.

Gârleanu, N. and Pedersen, L. H. (2018). Efficiently Inefficient Markets for Assets and Asset Management. *The Journal of Finance*, 73(4):1663–1712.

Hachmeister, A. (2007). *Informed Traders as Liquidity Providers: Evidence from the German Equity Market*. Springer, Wiesbaden.

Hamm, S. (2014). The Effects of ETFs on Stock Liquidity. *Unpublished*.

Israeli, D., Lee, C. M. C., and Sridharan, S. A. (2017). Is There a Dark Side to Exchange Traded Funds? An Information Perspective. *Review of Accounting Studies*, 22:1048–1083.

Jensen, M. C. (1964). The Performance of Mutual Funds in the Period 1945-1964. *The Journal of Finance*, 23(2):389–416.

Johnsen, T. (2011). Evaluering av aktiv forvaltning for Statens pensjonsfond Norge. Retrieved from: https://www.regjeringen.no/globalassets/upload/fin/statens-pensjonsfond/eksterne-rapporter-og-brev/2011/tore_johnsen_spn_21.3.2011.pdf.

Li, B., Rajgopal, S., and Venkatachalam, M. (2014). R^2 and Idiosyncratic Risk Are Not Interchangeable. *The Accounting Review*, 89(6):2261–2295.

Logica Capital Partners (2020). Long Volatility Presentation. *Unpublished*.

Madhavan, A. and Morillo, D. (2018). The Impact of Flows into Exchange-Traded Funds: Volumes and Correlations. *The Journal of Portfolio Management*, 44(7):96–107.

Malkiel, B. G. (1973). *A Random Walk Down Wall Street*. Norton, New York.

Massa, M., Zhang, B., and Zhang, H. (2015). The Invisible Hand of Short Selling: Does Short Selling Discipline Earnings Management? *Society for Financial Studies*, 28(6):1701–1736.

Morningstar (2020). Morningstar Direct. Retrieved from: https://www.morningstar.com.

MSCI (2020). MSCI World Constituents as of 01.10.2020. Retrieved 15.12.2020 from: https://www.msci.com/constituents.

Nesse, S. G. and Aasen, I. A. (2015). Indeksfond i USA, Norge og Sverige: en studie på indeksfonds replikeringsevne i USA, Norge og Sverige fra 2010 til 2014. *Unpublished*.

Norang, J. and Agustsson, F. R. (2018). Do Norwegian Individual Stocks Outperform Government Bills? *Unpublished*.

Norges Handelshøyskole (2020). Børsprosjektet. Retrieved from: https://www.nhh.no/forskning/borsprosjektet/.

Oslo Stock Exchange (2020). Oslo Børs: Equity Indices - Index Methodology. *Unpublished*.

Pedersen, L. H. (2018). Sharpening the Arithmetic of Active Management. *Financial Analysts Journal*, 74(1):21–36.

Petajisto, A. (2011). The Index Premium and its Hidden Cost for Index Funds. *Journal of Empirical Finance*, 18(2):271–288.

Petersen, M. A. (2009). Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. *The Review of Financial Studies*, 22(1):435–480.

Pettit, R. R. (1976). The Impact of Dividend and Earnings Announcements: A Reconciliation. *The Journal of Business*, 49(1):86–96.

Piotroski, J. D. and Roulstone, Darren, T. (2004). The Influence of Analysts, Institutional Investors, and Insiders on the Incorporation of Market, Industry, and Firm-Specific Information into Stock Prices. *The Accounting Review*, 79(4):1119–1151.

PwC Asset Management (2017). Asset & Wealth Management Revolution: Embracing Exponential Change. *Unpublished.*

Qin, N. and Singal, V. (2015). Indexing and Stock Price Efficiency. *Financial Management*, 44(4):875–904.

Renshaw, E. F. and Feldstein, P. J. (1960). The Case for an Unmanaged Investment Company. *Financial Analysts Journal*, 16(1):43–46.

Roll, R. (1984). A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *Journal of Finance*, 39(4):1127–1139.

Roll, R. (1988). R^2. *Journal of Finance*, 43(3):541–566.

Rowley, J. R., Hirt, J. M., and Wang, H. (2018). Setting the Record Straight: Truths About Indexing. *Unpublished.*

Sammon, M. (2020). Earnings Announcements and the Rise of Passive Ownership. *Unpublished.*

Scari, C. (2016). On the Changes to the Index Inclusion Effect with Increasing Passive Investment Management. *Unpublished.*

Schleifer, A. (1986). Do Demand Curves for Stocks Slope Down? *The Journal of Finance*, 41(3):579–590.

Sharpe, W. F. (1991). The Arithmetic of Active Management. *Financial Analysts Journal*, 47(1):7–9.

Sheather, S. J. (2009). *A Modern Approach to Regressions with R.* Springer, New York.

Sullivan, R. N. and Xiong, J. X. (2012). How Index Trading Increases Market Vulnerability. *Financial Analysts Journal*, 68(2):70–84.

Sørmo, J. A. (2016). Aksjeutlån og likviditeten i aksjemarkedet. *Unpublished.*

Thoresen, I. and Øren, M. (2017). Aktiv forvaltning i små markeder: En analyse av aktiv forvaltning i nordiske aksjefond over tidsperioden 2007-2016. *Unpublished.*

Vanguard (2020). Cash Panickers: Coronavirus Market Volatility. *Unpublished.*

Vanguard (2020a). Fast facts about Vanguard. Retrieved 16.12.2020 from: https://about.vanguard.com/who-we-are/fast-facts/.

Vanguard (2020b). Vanguard 500 Index Fund Investor Shares. Retrieved 15.12.2020 from: https://investor.vanguard.com/mutual-funds/profile/portfolio/vfinx.

Wurgler, J. (2010). On the Economic Consequences of Index-Linked Investing. *NBER Working Paper.*

Yan, X. (2006). The Determinants and Implications of Mutual Fund Cash Holdings: Theory and Evidence. *Financial Management*, 35(2):67–91.

Zou, Y. (2019). Lost in The Rising Tide: Exchange-Traded Fund Flows and Valuation. *Unpublished.*

Appendices

A1 Variable Definitions

Table A1.1: Variable Definitions

This table provides the definitions of all variables utilised in this paper.

Variable		Definition
$Passive\ Ownership_{i,t}$	=	The number of shares of stock i held by passive funds at time t divided by the total number of shares outstanding of stock i at time t. Expressed in percentage points.
$Spread_{i,m}$	=	The bid-ask spread for stock i in month m, defined as $\frac{1}{N_i}\sum_{t=1}^{N_i} 2\sqrt{max\,(E[(c_{i,t}-\eta_{i,t})(c_{i,t}-\eta_{i,t+1})],0)}$, where $c_{i,t}$ is the daily closing log-price, $\eta_{i,t}$ is the mean of the daily high and low log-prices and N_i is the number of daily observations for stock i in month m. Monthly averages are used to mitigate the effect of negative roots and to achieve the most accurate estimate. Expressed in percentage points.
$Illiq_{i,t}$	=	The Amihud illiquidity measure for stock i at time t, defined as $\frac{\lvert return_{i,t}\rvert}{turnover_{i,t}}$, where $\mid return_{i,t} \mid$ is the absolute return of stock i on day t, while $turnover_{i,t}$ is the trading volume of stock i on day t in NOK millions. Expressed in percentage points.
$CAV_{i,t}$	=	The cumulative abnormal volume in relation to an earnings announcement for stock i at time $t = 0$, defined as $\sum_{t=-22}^{21} \frac{\sum_{t=-22}^{21} V_{i,t}}{\frac{1}{63}\sum_{k=1}^{63} V_{i,t-22-k}}$, where $V_{i,t}$ is the NOK trading volume for stock i at day t.
$DM_{i,t}$	=	The drift measure in relation to an earnings announcement for stock i at time $t = 0$, defined as $\begin{cases} \frac{1+\sum_{t=-22}^{21} r_{i,t}}{1+\sum_{t=-22}^{22} r_{i,t}}, & \text{if } r_{i,T} > 0 \\ \frac{1+\sum_{t=-22}^{22} r_{i,t}}{1+\sum_{t=-22}^{21} r_{i,t}}, & \text{if } r_{i,T} < 0 \end{cases}$, where $r_{i,t}$ is the return on stock i on day t in the period leading up the announcement day, while $r_{i,T}$ is the return on the announcement day.

Variable		Definition
$QVS_{i,t}$	=	The quadratic variation share in relation to an earnings announcement for stock i at time T, defined as $\frac{r^2_{i,T}}{\sum_{t=1}^{63} r^2_{i,t}}$, where $r^2_{i,T}$ are the squared return of stock i at earnings announcement day T and $r^2_{i,t}$ is the return of stock i at day t in quarter T. Expressed in percentage points.
$Synch_{i,t}$	=	The return synchronicity for stock i at time t, defined as the adjusted R^2 in: $ret_{i,t} = \beta_1 * ret_{OSEBX,t} + \beta_2 * ret_{OSEBX,t-1} + \beta_3 * ret_{SEC,t} + \beta_4 * ret_{SEC,t-1} + \epsilon$, where $ret_{i,d}$, ret_{OSEBX} and $ret_{SEC,d}$ is the return at time t of stock, OSEBX and the stock's sector. The adjusted R^2 is transformed to a unbounded variable using: $\ln\left(\frac{r^2}{1-r^2}\right)$.
$Active\ Ownership_{i,t}$	=	The number of shares of stock i held by active funds at time t divided by the total number of shares outstanding of stock i at time t. Expressed in percentage points.
$Log\ MCAP_{i,t}$	=	The natural logarithm of the market capitalisation of stock i at time t.
$Volatility_{i,t}$	=	The annualised standard deviation of returns for stock i at time t. Expressed in percentage points.
$Volume_{i,t}$	=	The daily trading volume, defined as shares traded divided by shares outstanding for stock i at time t. Expressed in percentage points.
$Idiosyncratic\ Volatility_{i,t}$	=	The idiosyncratic volatility of stock i at time t, defined as: $(ret_{i,t} - ret_{OSEBX,t})^2$, where $ret_{i,t}$ and $ret_{OSEBX,t}$ is the daily return of stock i and OSEBX at time t.
$Beta_{i,t}$	=	The beta of stock i at time t, defined as the covariance between stock i and OSEBX returns, divided by the variance of the OSEBX returns.
$DiD_{i,t}$	=	A dummy variable equal to 1 if stock i is included in OSEBX at time t and 0 otherwise.
Δ	=	The change operator, Y_t less Y_{t-1} for annual samples and Q_t less Q_{t-4} for quarterly samples.

A2 Robustness Analyses

Table A2.1: Robustness Analysis for Illiquidity Measure

The regressions in this table depict the relationship between changes in illiquidity (Illiq) and changes in passive ownership for different subsamples. Model (1) is the full sample; Model (2) includes observations of stocks in the OSEBX; Model (3) includes observations of stock not in the OSEBX; Models (4), (5) and (6) include observations of stocks in the 1^{st}, 2^{nd} and 3^{rd} tercile based on market capitalisation, respectively; Model (7) includes observations of stock with a non-zero passive ownership. Across all models, fixed effects for stock and year are included. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses.

	Dependent variable:						
	Δ Illiq (%)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ Passive Ownership (%)	0.092*	0.014	1.021	1.859	−0.076	0.027**	0.007
	(0.052)	(0.011)	(0.723)	(5.640)	(0.109)	(0.011)	(0.011)
Δ Active Ownership (%)	−0.217***	0.002	−0.707***	−1.348**	−0.179**	−0.011	−0.026*
	(0.078)	(0.006)	(0.237)	(0.599)	(0.086)	(0.009)	(0.014)
Δ Log MCAP	−1.797***	−0.102***	−3.553***	−4.250***	−1.608***	−0.251***	−0.339***
	(0.455)	(0.037)	(0.795)	(1.316)	(0.606)	(0.075)	(0.100)
Δ Volatility (%)	−0.037	−0.004	−0.054	−0.083	0.015	0.006	−0.006
	(0.025)	(0.003)	(0.042)	(0.063)	(0.027)	(0.005)	(0.007)
Δ Volume (%)	−3.105***	−0.296***	−4.935***	−4.155***	−1.296***	−0.317***	−0.648***
	(0.732)	(0.085)	(1.123)	(1.274)	(0.417)	(0.108)	(0.133)
Δ Idiosyncratic Volatility	1.855**	0.475**	2.084*	1.622	0.177	0.414	0.646**
	(0.819)	(0.223)	(1.242)	(1.314)	(0.764)	(0.312)	(0.255)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	OSEBX	Not OSEBX	MCAP 1	MCAP 2	MCAP 3	Passive
Observations	2,196	763	1,255	699	735	762	1,089
Adjusted R^2	−0.079	−0.104	−0.124	−0.216	−0.277	−0.120	−0.080
F Statistic	12.505***	6.989***	10.169***	4.820***	2.764**	6.540***	12.417***
	(df = 6; 1960)	(df = 6; 648)	(df = 6; 1055)	(df = 6; 545)	(df = 6; 558)	(df = 6; 640)	(df = 6; 933)

Note: *p<0.1; **p<0.05; ***p<0.01

Table A2.2: Robustness Analysis for Spread Measure

The regressions in this table depict the relationship between changes in spread and changes in passive ownership for different subsamples. Model (1) is the full sample; Model (2) includes observations of stocks in the OSEBX; Model (3) includes observations of stock not in the OSEBX; Models (4), (5) and (6) include observations of stocks in the 1^{st}, 2^{nd} and 3^{rd} tercile based on market capitalisation, respectively; Model (7) includes observations of stock with a non-zero passive ownership. Across all models, fixed effects for stock and year are included. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses.

	Dependent variable:						
	Δ Spread (%)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ Passive Ownership (%)	0.031**	0.030**	0.034	0.122	0.039	0.028**	0.029**
	(0.012)	(0.013)	(0.090)	(0.791)	(0.044)	(0.012)	(0.012)
Δ Active Ownership (%)	−0.013	−0.006	−0.013	−0.071*	−0.002	−0.004	−0.005
	(0.009)	(0.010)	(0.020)	(0.042)	(0.021)	(0.010)	(0.009)
Δ Log MCAP	−0.269***	−0.144**	−0.354***	0.157	−0.493***	−0.185***	−0.245***
	(0.050)	(0.071)	(0.077)	(0.148)	(0.096)	(0.057)	(0.061)
Δ Volatility (%)	0.007**	0.013***	0.006	0.002	0.023***	0.015***	0.013***
	(0.003)	(0.004)	(0.004)	(0.004)	(0.006)	(0.004)	(0.004)
Δ Volume (%)	−0.132	−0.114	−0.058	−0.255*	0.057	−0.131	−0.152
	(0.081)	(0.116)	(0.131)	(0.148)	(0.178)	(0.130)	(0.133)
Δ Idiosyncratic Volatility	0.542***	0.190	0.665***	0.631***	0.098	−0.283	0.049
	(0.138)	(0.216)	(0.195)	(0.167)	(0.247)	(0.415)	(0.264)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	OSEBX	Not OSEBX	MCAP 1	MCAP 2	MCAP 3	Passive
Observations	2,196	763	1,255	699	735	762	1,089
Adjusted R^2	−0.045	−0.109	−0.103	−0.202	−0.152	−0.116	−0.090
F Statistic	23.325***	6.475***	13.724***	5.964***	13.152***	6.969***	10.932***
	(df = 6; 1960)	(df = 6; 648)	(df = 6; 1055)	(df = 6; 545)	(df = 6; 558)	(df = 6; 640)	(df = 6; 933)

Note: *p<0.1; **p<0.05; ***p<0.01

Table A2.3: Robustness Analysis for CAV Measure

The regressions in this table depict the relationship between changes in cumulative abnormal volume (CAV) and changes in passive ownership for different subsamples. Model (1) is the full sample; Model (2) includes observations of stocks in the OSEBX; Model (3) includes observations of stock not in the OSEBX; Models (4), (5) and (6) include observations of stocks in the 1^{st}, 2^{nd} and 3^{rd} tercile based on market capitalisation, respectively; Model (7) includes observations of stock with a non-zero passive ownership. Across all models, fixed effects for stock and year-quarter are included. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses.

	Dependent variable:						
	Δ CAV						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ Passive Ownership (%)	−0.853***	−0.770***	−0.682	−2.996	−0.714***	−0.540**	−0.763***
	(0.223)	(0.204)	(1.108)	(2.325)	(0.268)	(0.245)	(0.218)
Δ Active Ownership (%)	−0.002	−0.027	0.022	−0.423	0.300	0.127	0.120
	(0.201)	(0.188)	(0.471)	(0.577)	(0.290)	(0.170)	(0.204)
Δ Log MCAP	1.887***	2.122***	0.570	2.308	2.019	2.239**	3.151***
	(0.732)	(0.719)	(1.389)	(1.552)	(1.355)	(0.971)	(0.932)
Δ Idiosyncratic Volatility	4.132**	5.466*	3.820	1.880	6.087*	5.613	11.404***
	(1.886)	(2.897)	(2.737)	(2.153)	(3.435)	(6.603)	(3.354)
Δ Beta	−4.375***	−1.982	−7.753***	−8.918***	−3.438**	−1.984	−3.357***
	(1.067)	(1.413)	(1.773)	(1.950)	(1.633)	(1.404)	(1.123)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YearQuarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	OSEBX	Not OSEBX	MCAP 1	MCAP 2	MCAP 3	Passive
Observations	4,116	2,328	1,788	1,372	1,370	1,374	2,606
Adjusted R^2	−0.055	−0.063	−0.126	−0.145	−0.134	−0.104	−0.054
F Statistic	8.466***	5.165***	3.917***	3.762***	2.504**	2.834**	11.055***
	(df = 5; 3859)	(df = 5; 2163)	(df = 5; 1567)	(df = 5; 1179)	(df = 5; 1195)	(df = 5; 1230)	(df = 5; 2417)

Note: *p<0.1; **p<0.05; ***p<0.01

Table A2.4: Robustness Analysis for DM Measure

The regressions in this table depict the relationship between changes in drift measure (DM) and changes in passive ownership for different subsamples. Model (1) is the full sample; Model (2) includes observations of stocks in the OSEBX; Model (3) includes observations of stock not in the OSEBX; Models (4), (5) and (6) include observations of stocks in the 1^{st}, 2^{nd} and 3^{rd} tercile based on market capitalisation, respectively; Model (7) includes observations of stock with a non-zero passive ownership. Across all models, fixed effects for stock and year-quarter are included. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses.

	Dependent variable:						
	Δ DM						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ Passive Ownership (%)	−0.0002	−0.0002	0.002	−0.011**	−0.00000	0.001	0.0003
	(0.001)	(0.001)	(0.004)	(0.005)	(0.001)	(0.001)	(0.001)
Δ Active Ownership (%)	0.0002	0.001	−0.0002	−0.0004	0.0003	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Δ Log MCAP	0.010***	0.010***	0.009**	0.008**	0.010**	0.008*	0.013***
	(0.002)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
Δ Idiosyncratic Volatility	0.010*	0.017**	0.010	0.008*	0.007	0.032	0.031***
	(0.005)	(0.008)	(0.006)	(0.005)	(0.013)	(0.027)	(0.012)
Δ Beta	−0.003	0.006	−0.013***	−0.006	−0.004	0.006	0.002
	(0.003)	(0.004)	(0.005)	(0.006)	(0.007)	(0.004)	(0.004)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YearQuarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	OSEBX	Not OSEBX	MCAP 1	MCAP 2	MCAP 3	Passive
Observations	4,117	2,328	1,789	1,373	1,370	1,374	2,606
Adjusted R^2	−0.056	−0.062	−0.123	−0.150	−0.137	−0.103	−0.058
F Statistic	7.664***	5.468***	4.736***	2.526**	1.747	2.848**	9.019***
	(df = 5; 3860)	(df = 5; 2163)	(df = 5; 1568)	(df = 5; 1180)	(df = 5; 1195)	(df = 5; 1230)	(df = 5; 2417)

Note: *p<0.1; **p<0.05; ***p<0.01

Table A2.5: Robustness Analysis for QVS Measure

The regressions in this table depict the relationship between changes in quadratic variation share (QVS) and changes in passive ownership for different subsamples. Model (1) is the full sample; Model (2) includes observations of stocks in the OSEBX; Model (3) includes observations of stock not in the OSEBX; Models (4), (5) and (6) include observations of stocks in the 1^{st}, 2^{nd} and 3^{rd} tercile based on market capitalisation, respectively; Model (7) includes observations of stock with a non-zero passive ownership. Across all models, fixed effects for stock and year-quarter are included. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses.

	Dependent variable:						
	Δ QVS (%)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ Passive Ownership (%)	0.060	0.105	−0.355	0.472	0.041	−0.103	0.037
	(0.149)	(0.162)	(0.429)	(0.535)	(0.175)	(0.212)	(0.144)
Δ Active Ownership (%)	−0.239***	−0.166	−0.398**	−0.201	−0.117	−0.292**	−0.273***
	(0.087)	(0.116)	(0.169)	(0.146)	(0.183)	(0.136)	(0.100)
Δ Log MCAP	0.324	0.584	0.326	0.528	0.596	0.594	0.316
	(0.264)	(0.419)	(0.336)	(0.379)	(0.549)	(0.907)	(0.402)
Δ Idiosyncratic Volatility	−0.378	−0.757	−0.408	−0.120	−1.065	−7.224**	−2.232**
	(0.457)	(0.746)	(0.602)	(0.304)	(1.396)	(3.113)	(0.975)
Δ Beta	−0.021	−0.560	0.908	−0.597	0.971	0.119	−0.245
	(0.468)	(0.557)	(0.698)	(0.705)	(0.960)	(0.770)	(0.646)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
YearQuarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	OSEBX	Not OSEBX	MCAP 1	MCAP 2	MCAP 3	Passive
Observations	3,354	2,034	1,320	977	1,132	1,245	2,233
Adjusted R^2	−0.071	−0.081	−0.162	−0.209	−0.161	−0.115	−0.078
F Statistic	2.150*	1.179	2.272**	1.048	0.859	1.403	2.219**
	(df = 5; 3119)	(df = 5; 1874)	(df = 5; 1124)	(df = 5; 802)	(df = 5; 970)	(df = 5; 1109)	(df = 5; 2060)

Note: *p<0.1; **p<0.05; ***p<0.01

Table A2.6: Robustness Analysis for Synchronicity Measure

The regressions in this table depict the relationship between changes in return synchronicity (Synch) and changes in return synchronicity for different subsamples. Model (1) is the full sample; Model (2) includes observations of stocks in the OSEBX; Model (3) includes observations of stock not in the OSEBX; Models (4), (5) and (6) include observations of stocks in the 1^{st}, 2^{nd} and 3^{rd} tercile based on market capitalisation, respectively; Model (7) includes observations of stock with a non-zero passive ownership. Across all models, fixed effects for stocks and year are included. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses.

	Dependent variable:						
	Δ Synch						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Δ Passive Ownership (%)	0.039***	0.023*	0.001	0.170	0.071*	0.014	0.028**
	(0.014)	(0.013)	(0.088)	(0.191)	(0.038)	(0.012)	(0.012)
Δ Active Ownership (%)	0.005	−0.006	−0.004	0.013	0.023	−0.009	0.009
	(0.010)	(0.007)	(0.022)	(0.040)	(0.020)	(0.008)	(0.007)
Δ Log MCAP	0.251***	0.256***	0.197*	0.243*	0.295**	0.117*	0.251***
	(0.062)	(0.062)	(0.102)	(0.141)	(0.118)	(0.065)	(0.058)
Δ Beta	2.074***	1.360***	2.692***	2.389***	2.270***	1.287***	1.625***
	(0.110)	(0.097)	(0.189)	(0.256)	(0.169)	(0.084)	(0.098)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	Full	OSEBX	Not OSEBX	MCAP 1	MCAP 2	MCAP 3	Passive
Observations	1,882	735	979	576	633	673	1,038
Adjusted R^2	0.267	0.291	0.198	0.044	0.149	0.266	0.301
F Statistic	225.705***	103.090***	106.549***	42.400***	68.708***	87.624***	148.499***
	(df = 4; 1662)	(df = 4; 623)	(df = 4; 794)	(df = 4; 432)	(df = 4; 468)	(df = 4; 565)	(df = 4; 890)

Note: *p<0.1; **p<0.05; ***p<0.01

A3 DiD: Descriptive Statistics and Correlation Matrix

Table A3.1: Descriptive Statistics for DiD Sample

This table provides the number of observations, means, standard deviations and distributions for the sample used in the causal analysis. The sample consists of pooled cross-sections in the 2002-2019 sample period. *Illiq* is the annual average of daily absolute returns divided by turnover in mNOK. *Spread* is the annual average of the monthly average spread based on log close, high and low prices. *CAV* is the sum of daily volume over a period of one month before an earnings announcement, excluding the announcement day, divided by normal volume. *DM* is the cumulative return of the month before the announcement, excluding the announcement day, divided by the cumulative return of the same month including the announcement day. *QVS* is defined as the squared return on the announcement day, divided by the sum of squared returns for the full quarter. *Synch* is the log-transformed R^2 from a market model where stock returns are explained by market and sector returns. *OSEBX Inclusion* is a dummy variable indicating whether or not a stock is in the treatment group. *Active Ownership* is defined as the sum of shares held by active funds, divided by shares outstanding. *Log MCAP* is the natural logarithm of market capitalisation. *Volatility* is the annualised standard deviation of returns, while *Volume* is the average daily shares traded scaled by shares outstanding. *Beta* is the covariance of returns between a stock and OSEBX, divided by the variance of OSEBX. *Idiosyncratic Volatility* is the annual sum of the squared differences between daily returns of a stock and OSEBX. All change variables are calculated as the two-year average after the OSEBX inclusion date, less the two-year average before, and are winsorised at the 1% level. Δ*Synch* is winsorised at the 5% level. The quarter of index inclusion and the subsequent quarter are excluded from the sample. The (%) indicates that the variable is expressed in percentage points.

	N	Mean	Std. Dev.	Min.	p25	Median	p75	Max.
Dependent Variables								
Δ Illiq (%)	198	-0.07	1.35	-4.27	-0.24	-0.02	0.17	5.91
Δ Spread (%)	185	-1.23	0.53	-3.14	-1.45	-1.08	-0.86	-0.55
Δ CAV	120	-0.32	7.18	-20.17	-4.35	-0.71	4.48	15.83
Δ DM	120	-0.00	0.04	-0.08	-0.02	-0.00	0.01	0.19
Δ QVS (%)	120	1.69	4.98	-9.74	-1.69	1.52	4.74	14.03
Δ Synch	197	0.26	1.24	-2.34	-0.52	0.19	1.15	2.49
DiD Variable								
OSEBX Inclusion	198	0.46	0.50	0.00	0.00	0.00	1.00	1.00
Control Variables								
Δ Active Ownership (%)	198	2.82	4.05	-4.93	0.15	1.67	3.86	22.82
Δ Log MCAP	198	0.28	0.77	-1.83	-0.11	0.28	0.69	2.65
Δ Volatility (%)	198	-0.65	24.54	-60.15	-15.21	-2.41	13.09	63.07
Δ Volume (%)	198	-0.01	0.37	-1.18	-0.08	0.00	0.08	1.68
Δ Idiosyncratic Volatility	198	-0.43	1.77	-5.09	-1.41	-0.22	0.65	3.67
Δ Beta	198	0.13	0.36	-0.70	-0.08	0.11	0.32	1.41

Table A3.2: Pearson's Correlation for DiD Sample

This table provides Pearson's correlations for the sample used in the causal analysis. The sample consists of pooled cross-sections in the 2002-2019 sample period. *Illiq* is the annual average of daily absolute returns divided by turnover in mNOK. *Spread* is the annual average of the monthly average spread based on log close, high and low prices. *CAV* is the sum of daily volume over a period of one month before an earnings announcement, excluding the announcement day, divided by normal volume. *DM* is the cumulative return of the month before the announcement, excluding the announcement day, divided by the cumulative return of the same month including the announcement day. *QVS* is defined as the squared return on the announcement day, divided by the sum of squared returns for the full quarter. *Synch* is the log-transformed R^2 from a market model where stock returns are explained by market and sector returns. *OSEBX Inclusion* (OSEBX inc.) is a dummy variable indicating whether or not a stock is in the treatment group. *Active Ownership* (AO) is defined as the sum of shares held by active funds, divided by shares outstanding. *Log MCAP* is the natural logarithm of market capitalisation. *Volatility* is the annualised standard deviation of returns, while *Volume* is the average daily shares traded scaled by shares outstanding. *Beta* is the covariance of returns between a stock and OSEBX, divided by the variance of OSEBX. *Idiosyncratic Volatility* (IdVol) is the annual sum of the squared differences between daily returns of a stock and OSEBX. All change variables are calculated as the two-year average after the OSEBX inclusion date, less the two-year average before, and are winsorised at the 1% level. Δ*Synch* is winsorised at the 5% level. The quarter of index inclusion and the subsequent quarter are excluded from the sample.

	Δ Illiq	Δ Spread	Δ CAV	Δ DM	Δ QVS	Δ Synch	OSEBX inc.	Δ AO	Δ Log MCAP	Δ Volume	Δ Volatility	Δ IdVol	Δ Beta
Δ Illiq	1.00												
Δ Spread	0.06	1.00											
Δ CAV	-0.10	0.02	1.00										
Δ DM	0.08	-0.28	-0.13	1.00									
Δ QVS	-0.12	0.07	-0.14	-0.47	1.00								
Δ Synch	-0.09	-0.09	-0.05	-0.01	-0.05	1.00							
OSEBX inc.	0.02	-0.06	-0.15	0.15	0.04	0.27	1.00						
Δ AO	-0.26	0.09	0.06	-0.16	0.28	0.02	-0.04	1.00					
Δ Log MCAP	-0.35	-0.38	-0.04	0.16	0.02	0.49	0.21	0.02	1.00				
Δ Volume	-0.22	0.29	0.08	-0.17	-0.08	0.12	-0.04	0.08	-0.04	1.00			
Δ Volatility	0.24	0.50	0.09	-0.41	-0.05	-0.11	0.03	0.08	-0.53	0.24	1.00		
Δ IdVol	0.06	0.26	0.35	-0.14	0.00	-0.18	-0.20	0.05	-0.26	0.08	0.20	1.00	
Δ Beta	0.01	-0.12	0.00	-0.25	0.11	0.52	0.18	-0.01	0.13	0.22	0.20	-0.05	1.00

A4 Variance Inflation Factors

Table A4.1: Variance Inflation Factors

This table provides the variance inflation factor (VIF) for the relationships documented in the paper. All dependent variables are present in the table, with the VIFs from the models in which they are included.

	Correlational Models						DiD Models					
	Δ Illiq	Δ Spread	Δ CAV	Δ DM	Δ QVS	Δ Synch	Δ Illiq	Δ Spread	Δ CAV	Δ DM	Δ QVS	Δ Synch
Δ Passive Ownership (%)	1.20	1.20	1.25	1.25	1.29	1.20						
Δ Active Ownership (%)	1.32	1.32	1.35	1.35	1.47	1.33	1.43	1.38	1.43	1.43	1.43	1.36
Δ Log MCAP	1.56	1.56	1.59	1.59	1.76	1.56	1.87	1.91	1.62	1.62	1.62	1.25
Δ Volume (%)	1.24	1.24					1.19	1.25				
Δ Volatility (%)	1.42	1.42					1.76	1.86				
Δ Idiosyncratic Volatility	1.37	1.37	1.34	1.34	1.44		1.26	1.32	1.29	1.29	1.29	
Δ Beta			1.14	1.14	1.17	1.15			1.37	1.37	1.37	1.29
DiD							1.25	1.25	1.28	1.28	1.28	1.21
Stock Dummies	1.01	1.01	1.01	1.01	1.02	1.01						
Year Dummies	1.08	1.08				1.07						
Year-Quarter Dummies			1.03	1.03	1.03							
Sector Dummies							1.12	1.15	1.15	1.15	1.15	1.11
Month of Index Inclusion Dummies							1.12	1.12	1.18	1.18	1.18	1.10

Note: The VIF's are calculated as the square of $GVIF^{1/(2*Df)}$, where Df is the number of coefficients in the model. This form of the Generalised Variance Inflation Factor (GVIF) reduces the GVIF to a linear measure which is proportional to the inflation caused by multicollinearity in the coefficient's confidence interval (Fox and Monette, 1992). Thus, the general rule of thumb of critical VIF levels (>5) can be applied.

A5 Pairwise Correlation and Passive Ownership

Table A5.1: Pairwise Correlation and Passive Ownership

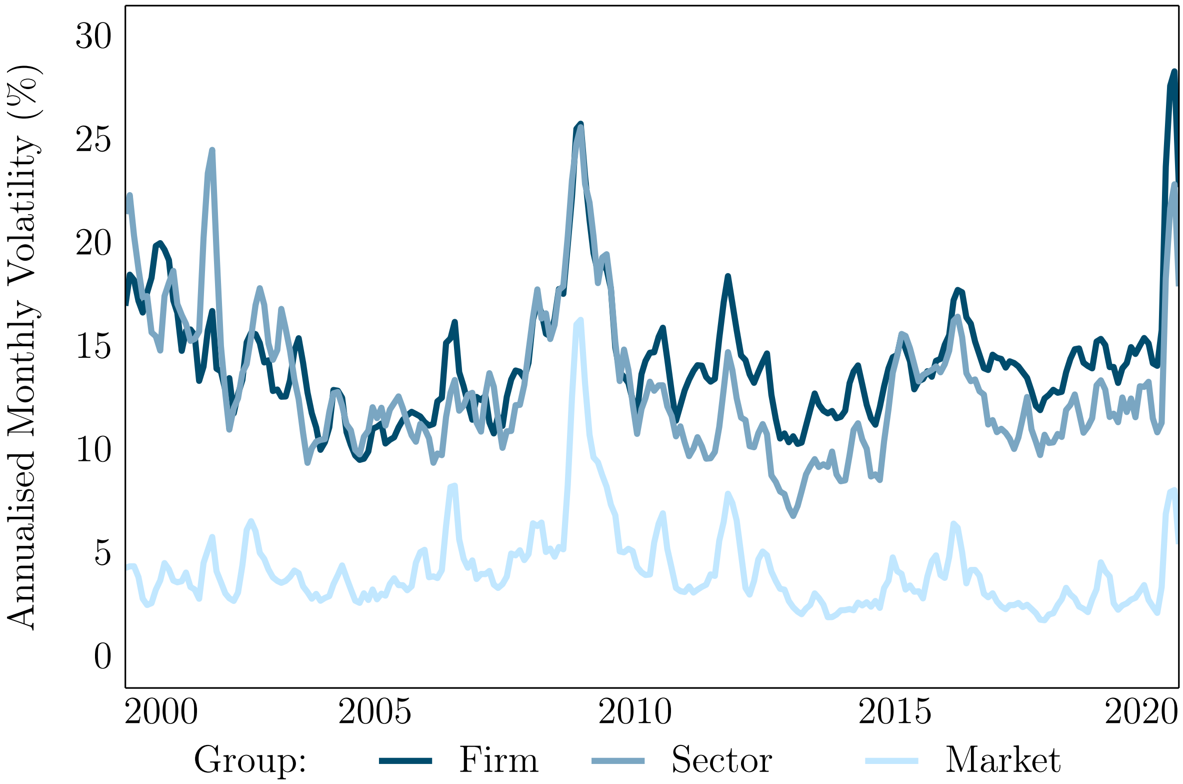
The regressions in this table depict the relationship between changes in equally-weighted pairwise correlation and changes in passive ownership. The sample consists of 205 stocks and 2,070 stock-year observations in the period from 2001 to 2020. Models (2) and (4) include control variables and fixed effects for stock and year, while models (1) and (3) include no such elements. Standard errors clustered by stock and robust to heteroscedasticity are displayed in parentheses. The dependent variable is multiplied by 100 and calculated as the change in the annual average pairwise correlation with all other stocks from one year to another.

	Dependent variable:	
	Δ Pairwise Correlation	
	(1)	(2)
Δ Passive Ownership (%)	0.536***	0.220***
	(0.186)	(0.066)
Δ Active Ownership (%)		−0.008
		(0.034)
Δ Log MCAP		0.176
		(0.201)
Δ Idiosyncratic Volatility		−2.130***
		(0.389)
Δ Beta		8.866***
		(0.381)
Stock FE	No	Yes
Year FE	No	Yes
Observations	2,070	2,070
Adjusted R^2	−0.104	0.321
F Statistic	9.489*** (df = 1; 1864)	240.992*** (df = 5; 1841)
Note:		*p<0.1; **p<0.05; ***p<0.01

A6 Volatility Distribution

Figure A6.1: Annualised Monthly Volatility Distribution

This chart illustrates the 3-month rolling annualised monthly volatility. The volatility is segmented into mutually exclusive components: firm-specific, sector-specific and market-specific volatility, in line with Campbell et al. (2001). Market-specific volatility is calculated using the standard deviation of the MCap-weighted returns of the total market. Sector-specific volatility is defined as the annualised root of the sum of residuals in a market model regressing MCap-weighted sector returns on market returns, which is then weighted by the MCap of each sector. Finally, firm-specific volatility utilises a similar approach, regressing firm returns on the weighted sector returns. The annualised firm volatility is based on the MCap-weighted root of the sum of residuals in this regression.



A7 Return Synchronicity

Figure A7.1: Synchronicity and R^2

This chart illustrates the relationship between a one percentage point increase in passive ownership and the change in the adjusted R^2 derived from the market model, for different levels of *Synch*. The x-axis corresponds to different *Synch* levels and the solid lines represent the change in R^2 from the coefficiencts of Δ *Passive Ownership* from two different regression models. The dotted lines mark the sample mean of the level of *Synch* for the full sample and for the top tercile based on market capitalisation, and the corresponding change in R^2. The change in R^2 is multiplied by 100, thus, the value 0.4 equals 0.4 percentage points.