Copenhagen Business School

M.Sc. Finance and Strategic Management

Master's Thesis

The Effect of Passive Investment on Market Efficiency

An Empirical Study of the U.S. Equity Market (1989-2018)

Authors:
Ingrid Alexandra Louise Petersson
91826

Kasper S.R. Thomsen
93074

Supervisors:
Theis Ingerslev Jensen
Ph.D. Student

Lasse Heje Pedersen
Finance Professor

May 15, 2019

108 Pages

212,590 Characters

Abstract

The past decades have seen the rise of an investment revolution. Just 40 years ago, the first index mutual fund was founded, today passive management occupies 30% of the market, and by 2024 at the latest, Moody's expect passive investment to surpass active management. The impacts of this natural experiment are yet to be identified.

In this paper, we attempt to identify the effect of increased passive investments on market efficiency. First, we conduct an extensive literature review, which provides the foundation for the development of three hypotheses. The literature provides support for either more, less, or equally efficient markets. In order to reach conclusions regarding our hypotheses, we collect 52,153,354 observations, a total of 437,101,186 data points, and conduct 658,118 unique event studies over the 30-year period ranging from 1989-2018. We use the results from the event study and their significance levels as a proxy for the level of market efficiency. Subsequently, we run two regression models in an additional attempt to isolate the effect of increased index investing.

We find no evidence to suggest that the increase in indexing should have made highly indexed securities more or less efficiently priced compared to a basket of otherwise similar firms. Thus, we disprove our first and second hypotheses. However, we find weak evidence to support our final hypothesis, which suggests that the net effect of increased indexing is negligible. Thus, we find no conclusive empirical evidence to answer our research question. Furthermore, our results seem to suggest that securities have, on average, become more efficiently priced. The results seem to suggest that the detrimental implications of increased indexing on market efficiency proposed by some scholars and practitioners are yet to materialise. We go on to consider how these results fit in with the literature on market efficiency in general. We conclude that this thesis provides evidence in favour of the efficiently-inefficient view on market efficiency.

Acknowledgements

This thesis was made possible by the guidance and support, we received from multiple sources. First, this thesis would not have been possible without the thorough supervision we received from our advisor Theis Ingerslev Jensen. He always made himself available and contributed greatly to the direction of this thesis. Second, we would like to thank Jesper Rye Rasmussen for feedback on the code and assisting in the creation of the lambda expressions that made the generation of the more complex variables efficient even in a dataset of this magnitude.

Table of Contents

1 Introduction

The composition of investors in asset markets are changing. The past 40 years have seen the birth and rise of an investment revolution in which investors have been moving funds aggressively from active to passive management. Moody's (2017) report that the proportion of passively managed funds have increased from 4% in 1996 to around 30% today. The implications of this are yet to be identified.

This revolution was initiated in 1975, when John Bogle was inspired by Paul Samuelson (1974) work. Samuelson, a Nobel Laureate, challenged the performance of active management, and suggested that someone should set up an index tracking the S&P 500. That someone turned out to be John Bogle. He created the first value-weighted index mutual fund – and it was a complete flop (Zweig, 2016). Initially, the fund raised just $11m – less than 10% of the goal. Active managers campaigned against it referring to it as "unamerican" and "a sure way to mediocrity" (ibid.). But Bogle persisted. Consequently, the 'First Index Investment Fund' was listed on the 31st of August 1976. The fund initially struggled to attract attention but continued to grow steadily. By 1985, the fund had $511m under management, before soaring in popularity during the bull run of the late 80's and early 90's. In 1995, assets under management had grown to $55bn. Consequently, Vanguard was able to reduce fees and attract additional funds. Today, Vanguard is one of the largest asset managers and have more than $5.3tn under management.

Researchers and practitioners continue to conjecture about the effects of the index revolution on asset markets. Some active managers claim that index investing is "worse for society than Marxism" (Kawa, 2016), because of its implications for capital allocation. Researchers show that index investing has increased the correlation of equities and thus systematic risk (Sullivan & Xiong, 2012). Furthermore, the value-weighted approach leaves investors overexposed to overvalued securities and underexposed to undervalued companies (Brown, 2018). Before passing, Bogle suggested that increased indexing may have detrimental effects on corporate governance (Bogle, 2018). Recently, researchers have gained an interest in the effects of increased passive ownership on market efficiency. However, no conclusive evidence has emerged.

In this paper, we address this research gap by conducting a long-term empirical study of asset prices from 1989-2018. In this paper, we attempt to tease out the effect of an increased proportion of passive investments on asset pricing efficiency. In particular, we test whether the S&P 500 have become more or less efficiently priced compared to a basket of similar firms.

1.1 Motivation

Market efficiency is one of the fundamental ideas of modern finance. As such, we have encountered the theory multiple times throughout our time at Copenhagen Business School (CBS). However, what has been absent is a definitive answer to the obvious question, “are markets efficient?”. This question has interested us continuously during our time at CBS, and we have taken every opportunity granted to us to explore it further. However, we are yet to arrive at a conclusion as new avenues of thought continues to emerge. In particular the opposing view of behavioural finance and the unifying view of ‘efficiently inefficient’ have been great inspirations and motivators for this paper.

More recently, the rise of index investing and its effect on market efficiency have caught our interest. Index investing has been touted as the ‘democratisation of investing’ (Novick, 2017). Furthermore, the rise can be regarded as a signal from investors, who realise that they are unable to beat the market. Index investing offers many advantages to investors, such as a low-cost, mediocre market return, many people, especially practitioners, have pointed out the potential downsides of increased index investing. For instance, a research analyst at Sanford C. Bernstein claims that passive investing “is worse than Marxism” (Kawa, 2016) due to its effect on capital allocation.

The rise of indexing may also have adverse effects on market efficiency. Bleiberg, Priest, and Pearl (2017) describe multiple avenues by which increased indexing could reduce the level of market efficiency. Much research has been conducted regarding the effects of indexing on asset markets in general. However, so far, no definitive empirical evidence has been collected on the effect of increased passive investments on market efficiency. In this paper, we attempt to identify this effect.

1.2 Research Question

The past decades have seen an exponential growth in the proportion of index investing (Moody's Investors Service, 2017). Practitioners and academics have discussed the implications for market efficiency, however, so far, no decisive evidence have been collected to prove or disprove the effect. Proponents of active management suggest that index investing reduces the number of market participants, the liquidity of markets, and thus the asset pricing efficiency of those markets. However, conversely, the proponents of index investing argue that any inefficiencies arising will be arbitraged away by arbitrageurs. Hence, markets will revert to their efficient equilibrium. In this paper, we explore this question in detail and attempt to tease out the effect of index investing on market efficiency.

Does the increase in passive investment have an effect on the level of stock market efficiency?

1.3 Overview of Paper

The paper will proceed as follows. In Section 2, we provide an initial introduction of the topic by distinguishing passive and active investing as well as presenting the historical background of index investing. In Section 3, we conduct an extensive literature review on market efficiency, behavioural finance, and our methodology. In Section 4, we develop three hypotheses based on the literature review. In Section 5, we go through our methodological approach and group definitions. In Section 6, we describe the data preparation process extensively. In Section 7, the data collected is presented. In Section 8, the results of our tests are presented, while we discuss them in Section 9. In Section 10, we consider limitations of our paper and conduct robustness tests. In Section 11, we discuss and suggest further research, which could be conducted in this field. Finally, in Section 12, we provide a conclusion.

2 Setting the Stage

This section of the paper will provide an introduction to index investing and its historical background. The purpose of the section is to provide a clear distinction between active and passive investing as well as a definition of index investing. Furthermore, we set out to establish how we will contribute to the literature on market efficiency.

2.1 Distinguishing Active and Passive Investments

This paper will discuss index investing, passive investing, and active investing in detail. To ensure this discussion takes place on a common understanding, we will start out by defining these terms.

An investor is defined as active if a fraction of his portfolio deviates from the benchmarking index (Cremers & Petajisto, 2009). Investors invest actively in a pursuit to exploit market inefficiencies and with the goal to beat the market. However, this pursuit is costly. Transaction costs associated with trading, information acquisition costs, etc. are expensive undertakings and will reduce the profit of the strategy. According to French (2008), 0.67% of the aggregate value of the market is spent yearly trying to achieve excess return. In the long-run, these costs accumulate and have large implications for long-term returns. Furthermore, it must hold that, before costs, the average return of active investors must be zero (Sharpe, 1991). That is, when I win, you lose, and vice versa.

Passive funds, however, refrain from making active decisions regarding the allocation of funds. Instead it tracks an index. A strategy which John Bogle (Arnold, 2018) refers to as the most boring investment strategy. Hence, investors in an index fund acknowledges the challenge of beating the market and instead buys the market (Vanguard, 2019). To quote Bogle (2007), "Don't look for the needle in the haystack. Just buy the haystack". The most well-known index fund is the Vanguard S&P 500, which was the first index mutual fund, and tracks the 500 largest firms on the U.S. equity market as determined by Standard and Poors. Compared to active investing, an index fund has much lower fees. Managing an index fund requires much less resources since an index fund is rule-based and only requires adjustments of the portfolio to match the index, when the composition of the index changes. Consequently, costs related to trading and management are much lower. These savings are passed on to investors with fees as low as 0 bps (Rosenbaum, 2019). This has a huge effect on long-run aggregate returns. In this paper, we will use the terms passive investing and index investing interchangeably. Though all passive investments are not index investments, all index investments are passive investments.

The distinction between active and passive investing is not clear cut. Inigo-Fraser Jenkins, the head of the quantitative research team at Sanford C. Bernstein & Co., claims that "there is no such thing as passive investing" (Burger, 2018). Algorithms, which are an important tool in passive investing, are built by people and correspondingly subject to the same biases as people. Furthermore, investors cannot let go of the thought of beating the market. Rather than investing in the traditional value-weighted indexes, smart-beta strategies with different weights (volatility, value, momentum, liquidity, etc.) have been developed (Financial Times, 2018). Though these are passive, rule-based indexes, they resemble something closer to active management compared to the traditional index fund. Furthermore, because there are so many passive funds/ETFs (upwards of 3 million according to some estimates (ibid.)), choosing the fund becomes an active decision (Chancellor, 2018). Furthermore, Pedersen (2018) challenges the definition of passive investing. He argues that, when new shares are issued, firms enter and exit, and indexes are reconstituted even the passive investor has to trade regularly. Hence, the decision regarding when to rebalance the portfolio to mimic the index becomes an active decision.

2.2 Historical Background

Traditional investing has been disrupted. In 1975, John Bogle created the first value-weighted index mutual fund available to the public (Zweig, 2016). He had recently lost his job as CEO of Wellington Management and started his own investment company, Vanguard. Initially, Vanguard was a traditional active-management firm, but when Bogle read Paul Samuelson (1974) paper contesting active management, he was inspired. Samuelson claimed that even the best money managers were not able to consistently outperform the market. He went on to suggest that someone should set up a mutual fund tracking the S&P 500 (ibid.). That someone turned out to be John Bogle. His first fund, however, was a complete flop. Originally, Bogle set out to raise \$150m, but at the time of the underwriting just \$11m had been contributed to the fund (Zweig, 2016). The fund was critiqued and characterised as 'unamerican', 'a sure way to mediocrity', but Bogle persisted. So, on the 31st of August 1976 the first index fund was made available to the public.

It was not an immediate success. Ten years after inception, the fund had just \$511m under management (Vanguard, 2019). However, as assets under management grew, Vanguard was able to lower their fees significantly from 65bps in 1976 to less than half by 1990, thus significantly undercutting traditional active mutual funds (ibid.). Low fees turned out to be essential in attracting investments. Something Bogle attributed to "the relentless rules of humble arithmetic", i.e. fees have compounding interest. In the long, run excessive fees can wipe out even large gains of active management (Bogle, 2005). A period of exponential growth in the fund followed. By 1995, assets under management reached $55 billion – today, the comparable number is more than $5 trillion (Vanguard, 2019). Resultingly, Vanguard have been able to reduce their fees even more. In 2018, the expense ratio of the 'Vanguard Total Stock Market Index' was just 4 bps (Rosenbaum, 2019). However, Vanguard were not the only beneficiaries of this revolution.

Several competitors have emerged. Schwab, Fidelity, and iShares have set up similar products and have encroached on Vanguards turf. In addition to the index focused companies, the big investment banks such as Goldman Sachs and J.P. Morgan have also introduced ETFs with competitive expense ratios. In 2018, Fidelity one-upped all of their competitors by launching the first no-fee index fund (ibid.). Fintech companies and robo-advisors, such as Betterment, have been created as a result of this development. They provide a platform, which encourages retail investors to invest passively rather than actively.

The media have increasingly focused on this topic. Index investing is often referred to as the 'democratisation of investing' (Novick, 2017). Several stories exhibiting the superiority of index investing has emerged. For example, in 2007, Warren Buffett suggested that an S&P 500 index fund would be able to beat a basket of hedge funds net of fees, costs, and all expenses (Perry, 2018). He offered a wager worth $500,000 to anyone who was willing to take the other side of the bet. Ted Seides, a hedge fund manager, accepted the bet. The wager ran from 1st January 2008 to 31st December 2017. The return of the index fund eclipsed that of the hedge funds. By the end of the bet, $100,000 invested in the index would have grown to $225,586, while the hedge funds would have grown to just $147,889 (ibid.).

More entertainingly, the economist Burton Malkiel (1973) claimed that a portfolio managed by "a blindfolded monkey throwing darts at a paper would do just as well as one carefully selected by experts". In a 14-year period from 1988 to 2001, the Wall Street Journal decided to test this hypothesis. Their results came out inconclusive, due to the small sample size and the endogenous design of the test (Jasen, 2002). However, in 2013 Arnott, et al. (2013) published a more formalised study simulating 100 monkeys throwing darts. In their study, the average monkey beat the average investor by 1.7% per year since 1964. This result has been heavily cited in the popular media (e.g. Ferri, 2012; Economist, 2014).

The stories published in popular media has made the retail investor increasingly aware of the power of index investing. Thus, media have contributed to the rise of index investing and its popularity looks unstoppable. The investment revolution, which had its humble beginnings in the mid-70s are about to dominate the investment landscape. In 2017, Moody's (2017) reported that passive investments occupied upwards of 30% of the market and forecast that it will be the dominant investment type by 2024.

The effects of this are yet to be seen. However, researchers and practitioners are already greatly debating these effects. Sanford C. Bernstein & Co. claims that passive investing is "worse than Marxism" (Kawa, 2016), because it weakens the market mechanisms connected to allocation of capital. Proponents of agency theory argues that more passive investing will reduce the monitoring ability and incentive of investors, thus reducing the oversight of management (Appel, Gormley, & Keim, 2016). Appel, Gormley, and Keim (ibid.) suggest that this could lead to more fraud, worse performance, etc. Index investing includes some detrimental mechanism, e.g. the buying and selling of many assets simultaneously. Sullivan and Xiong (2012) shows that this has increased correlation of equities. Furthermore, the value-weighted approach leaves investors overexposed to overvalued companies following significant rallies, such as the FAANG (Facebook, Amazon, Apple, Netflix, and Google) stocks in the most recent years (Brown, 2018). Finally, it is argued that the rise of indexing may have an adverse effect on market efficiency (Bleiberg, Priest, & Pearl, 2017).

3 Theoretical Background

This section of the paper will provide a theoretical background of the preceding research conducted in this field. Based on this literature review, we will go on to develop our hypotheses in the subsequent section. This section will proceed as follows: First, we will discuss the extensive literature on market efficiency, including a sub-section on price discovery and liquidity. Second, we will discuss the opposing view of behavioural finance namely that markets can be inefficient, and anomalies can arise. Third, we go on to consider what evidence has arisen on the ability to beat the market and which view of market efficiency this support. Fourth, we discuss the middle ground, i.e. that markets can be considered 'efficiently inefficient'. Fifth, we will briefly consider a different contributor to efficiency and a different component of active management, namely corporate governance. Sixth, we will look at the literature on indexing and its effects. Finally, we will provide examples of our methodology being applied in other research papers.

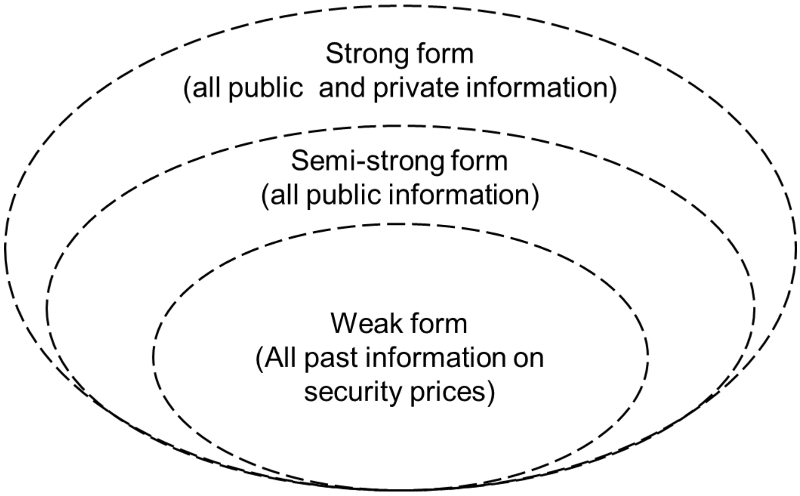
3.1 Market Efficiency

In theory, index investing is the only way to invest for a proponent of market efficiency. If markets are efficient, they cannot be beaten, i.e. the optimal investment is in the index. In this section, we will discuss the extensive literature on market efficiency as well as some of the mechanisms influencing it.

3.1.1 The Efficient Market Hypothesis

At the core, the efficient market hypothesis relates to the degree to which information is being reflected in the price (Grossman & Stiglitz, 1980). The idea of efficient capital markets was formalised in 1970's, when Fama (1970) published his seminal paper on the subject. He states that efficient markets work "[…] under the assumption that security prices at all times fully reflect all available information […]" (ibid.). The idea, which he dubbed 'the efficient market hypothesis' (EMH), has been widely appraised and critiqued in the literature. Proponents of the theory argue that investors are rational value-maximising agents, who make optimal decisions based on the information at hand. Yet, they cannot anticipate forthcoming information. Conversely, opponents disagree with the underlying assumption of rationality. Instead, they propose an alternative behavioural framework as explanation for market anomalies. In this subsection, we will briefly review the three forms of market efficiency discussed by Fama (1970): weak form, semi-strong form, and strong form. The three forms of market efficiency are illustrated in Figure 1.

Figure 1 Illustrative Depiction of the Efficient Market Hypothesis



Source: Own creation

The weak form EMH argues that security prices fully reflect all historical price information. Consequently, investors cannot obtain abnormal profits based on past information alone, i.e. prices can be said to have no memory and thus follow a random walk (Titan, 2015). Traditionally, weak form EMH has been tested by measuring autocorrelation among returns and by comparing the return of different trading strategies (Naseer & Tariq, 2015).

Fama's (1970) initial study measured autocorrelation on the Dow Jones Index in the period 1957-1962 and found that these correlations were zero. This indicated linear independence and thus consistency with weak form EMH. Contrastingly, Campbell, Lo, and MacKinlay (1997) found short-term serial correlation to be non-zero, which suggests that there is short-term momentum in stock prices. This is inconsistent with weak form EMH. DeBondt and Thaler (1985) finds that markets tend to overreact and underreact. They find that there is evidence of a return reversal. The researchers attribute this phenomenon to pessimism and optimism among investors. Later, Jegadeesh and Titman (1993) find that investors can earn a superior return by going long past losers and shorting past winners. This is a strict violation to weak form market efficiency.

Fama and French (1988) on the other hand found that 25% of the 3-5 year return of small firms could be predicted, while 40% of the return of large firms long-term returns can be explained by past returns. The correlation constituted a reversal, since the long-term returns were negatively correlated to the past returns. Fama and French (1988) interpret their results as a mean reversion to the fundamental value, i.e. there are two components of stock prices, a permanent and a temporary. Hence, in contrast to DeBondt and Thaler (1985) and Jegadeesh and Titman (1993), they consider it consistent with weak form EMH. In summary, the evidence on weak form EMH is mixed and no definitive consensus has been reached among researchers. Consequently, weak form market efficiency remains the most contested form of market efficiency.

Semi-strong market efficiency builds on the assumptions of weak form EMH. Furthermore, it adds that all publicly available information must be reflected in the price. That is, the semi-strong form implies that investors can only obtain an abnormal return with non-public information on the underlying assets. To test the semi-strong form, researchers employ event studies (Naseer & Tariq, 2015). Event studies test how stock prices react to new information (Campbell, Lo, & MacKinlay, 1997). One event which is often tested in the literature, as well as in this paper, is earnings announcements.

In 1968, Ball and Brown (1968) conducted the first test of post-earnings announcements drift. In perfectly efficient markets, the new information should be seamlessly and immediately incorporated into the price. That is, prices should revert to a random walk immediately. However, Ball and Brown (1968) found that prices 'drifted' into place. When this drift is significant, it is a sign of market inefficiencies, i.e. their finding suggests that semi-strong form does not hold. More recent studies like DeBondt and Thaler (1985) as well as Brandt, Kishore, Santa-Clara, and Venkatachalam (2008) have confirmed their findings. DeBondt and Thaler (1985) find that markets tend to overreact to extreme news. Resultingly, rather than reverting to a random walk post-announcement, as predicted by EMH, the price drifts into place. Brandt et al. (2008) devise a trading strategy, which buys and sells shares based on a proxy for earnings surprise. That strategy produces an average abnormal return of 7.55% per year in the period 1987-2004. This is a strict violation of semi-strong form market efficiency.

Finally, strong form EMH elaborates on weak and semi-strong form by assuming that all information, including private information, is reflected in the price. This would imply that "no individual has higher expected trading profits than others because he has monopolistic access to information" (Fama, 1970). That is, trading on insider information will not enable superior returns. Most empirical evidence suggests that markets are not strongly efficient. Chowdhury, Howe, and Lin (1993) find that insiders can use their superior information prior to public announcements to make a superior consistent return. Similarly, Pettit and Venkatesh (1995) find that insiders are able to time their transactions such that they get an above market return. That is, this evidence suggests that the market does not reflect all information and these findings contradict strong-form market efficiency.

There is largely “no consensus among economists regarding any of the three forms of EMH” (Titan, 2015). However, in general most evidence invalidates strong and semi-strong form efficiency. Yet, weak form efficiency continues to be the most highly debated form of market efficiency.

3.1.2 Price Discovery

Price discovery is the function by which information is reflected in the price (Bunzel, et al., 2017). For markets to be efficient, a well-functioning price discovery function is required. The price discovery function relies on the wisdom of the crowds (Bleiberg, Priest, & Pearl, 2017). To be wise, crowds must meet two criteria: (1) people must have diverse opinions, and (2) those opinions must be formed independently of one another. When market participants become too similar, mispricing is more likely. This phenomenon is referred to as ‘the madness of the crowds’ (Mackay, 1841).

In an increasingly indexed world, it is not hard to imagine that the diversity of market participants could suffer as ‘noise traders’, non-sophisticated investors, rely less on active investing and more on passive investing. Liu and Wang (2018) develops an equilibrium model to test these effects. They predict that the outcome will depend on the cause of the rise in index investing. If indexing increases as a result of increased participation costs in the non-index market, then price discovery and efficiency is reduced. However, if it is due to increased price transparency in the non-indexed market and thus lower profitability, then price discovery is improved. If the rise in indexing is exogenous, however, then the model predicts that price discovery is reduced in both the indexed and non-indexed market.

One part of price discovery, which is often neglected is the ability to express your pessimistic opinion. Short-selling enables traders to do so. However, it is costlier, riskier, more difficult, and thus less common. Nonetheless, shorting does have an impact on the price. Consider an example from Pedersen (2015, pp. 120-21):

> “There exist two types of investors, type 1 and type 2, who differ in their views on how cyclical the stock really is. Type 1 investors believe the stock will be worth 80 in a recession and 120 in a boom. Given that these are equally likely, they value the stock at 100. Type 2 investors believe that the stock is more cyclical [recession=60, boom=140]. […] type 2 investors also value the stock at 100”.

Now, suppose that prices are set by the most optimistic trader, because short-selling is impossible, what is the price? It seems intuitive that the price should be 100 – both traders value the stock as such. However, since prices are set by the most optimistic trader, the price will be 80 in a recession and 140 in a boom, i.e. the price today is (140+80)/2=110. That is, because both investors expect to sell to a 'greater fool' and shorting is impossible, a 10% bubble emerges in this example. Boehmer and Wu (2012) finds that stock prices are more accurate, when short sellers are more prominent. Furthermore, information is reflected in the price faster and negative earnings surprises exhibit less post-earnings announcement drift when shorting flow is larger (ibid.).

An additional obstacle to an efficient price discovery function is the inherent frictions of trading. As described, price discovery is the mechanism by which information is reflected in the price. However, information is costly to acquire (Garleanu & Pedersen, 2018). Furthermore, trading is associated with transaction costs, which occur as a compensation for liquidity (Aiyagari & Gertler, 1991). These frictions have a detrimental effect on the efficiency of the price discovery mechanism.

3.1.3 Liquidity

For markets to be efficient and prices to adapt to new information, they must be liquid (O'Hara, 2003), i.e. there must be a mechanism that matches buyers and sellers. Imagine an extreme example (ibid.), where buyers and sellers arrive to the market on different days. If buyers arrive on Monday, but there are no sellers, while sellers arrive on Tuesday, and there are no buyers, no trades will occur. However, this deadlock can be solved by a middle hand, who sells on Monday and buys on Tuesday, thus enabling all market participants to trade (Grossman & Miller, 1988). However, for taking this risk, she requires compensation – the bid-ask spread.

The effect of this transaction cost on market efficiency has been widely debated in the literature. Two opposing arguments have emerged. On the one hand, the cost is too small a part of the risk premium to matter (Aiyagari & Gertler, 1991). On the other hand, researchers have empirically shown that asset prices do in fact reflect liquidity costs (Amihud & Mendelson, 1986). The cost of liquidity can be thought of as a tax born by investors. It seems apparent, that if these costs are large enough, they will impact the ease and volume of investing and thus the asset pricing efficiency (O'Hara, 2003).

When stocks are included and excluded from an index, it has profound implications for their liquidity. Hegde and McDermott (2003) finds that after the inclusion to the index, a sustained increase in liquidity follows. They attribute this to a decrease in transaction costs and more information. Conversely, when a stock is excluded from the index, it becomes less liquid. Furthermore, they conduct an event study to determine the return effects of inclusion. They find that firms included in an index exhibit a positive cumulative abnormal return. They attribute this mostly to the decrease in the effective spread. Hence, the evidence put forth by Hegde and McDermott (2003) suggests that inclusion in an index reduces transaction costs, increases liquidity, and increases asset pricing efficiency.

A prerequisite for a liquid market is a high number of active buyers and sellers (Economides, 1995). It is often 'noise traders', who provide this liquidity (O'Hara, 2003). When index investing increases as a proportion of total funds under management, the proportion of active traders is reduced. Rationally, 'noise traders' should be overrepresented among the investors, who migrates from active to passive investing (Garleanu & Pedersen, 2018). Hence, more indexing may reduce liquidity and thus the efficiency of markets.

3.2 Market Inefficiencies

While passive investors rely on market efficiency, active investors attempt to exploit inherent mispricing arising from market inefficiencies to earn an above normal return. This section will provide the opposing view to the literature discussed previously. While market efficiency assumes that all market actors are rational, the literature on market inefficiencies abandons the assumption of homo economicus and borrows findings from psychology to explain the actions of market actors. In this section, we will discuss the increasingly accepted literature on behavioural economics and the market anomalies which have been identified.

3.2.1 Behavioural Finance

Becker (1976) describes the pillars of rational choice as: (1) people have rational, consistent preferences, (2) they maximise utility, and (3) make independent decisions based on all available information. These assumptions provide the foundation for decision-making in neoclassical economics. Until the 1970's, this belief was recognised as the truth in economics. However, at this point, a new school of thought began to develop. Behavioural economics challenged the neoclassical views on decision-making and instead suggested that human beings make sub-optimal decisions due to incomplete information and inability to process such information. People are influenced by biases, heuristics, context, and social influence. Consequentially, choice is inconsistent and irrational. Based on this line of thinking, an opposition to market efficiency led by Robert Shiller, Amos Tversky, Daniel Kahneman, and Richard Thaler has emerged.

Shiller (1981) founded the school, which would come to be known as behavioural finance, when he challenged the efficient market hypothesis. When he tested the reaction to dividend changes, he found that markets reacted inconsistently with the model predicted by the efficient market hypothesis (ibid.). Bounded rationality is the foundation of behavioural finance. The concept of bounded rationality was introduced in 1982 by Herbert Simon (1982). He suggests that people are not perfectly rational, rather they are bounded in their rationality by time, available information, and thinking capacity.

Kahneman (2003; 2011) develops Simon's idea by introducting the 'dual system theory'. The theory suggests that human beings makes decisions in two ways: (1) intuitive, which requires little effort and is subject to heuristics and biases, (2) reasoning, which is effortful, thoughtful, and converging towards rational. In practice, most decisions are intuitive (2003; 2011). Consequently, people often make fast and flawed decisions. Frederick (2005) exemplifies this by asking a relatively simple question: "A ball and a bat costs $1.10 [together]. The bat costs $1 more than the ball. How much does the ball cost?". Though seemingly simple, Mastrogiorgio and Petraca (2014) find that 43.3% of students in an experiment fail to answer it correctly (the ball costs $0.05, while the bat costs $1.05). This is attributed to system one, intuition. In a modified and more complex problem, which invites more computation ('the banana and bagel' problem), students employed system two, reasoning, and answered correctly 80% of the time (ibid.).

System one enables biases to arise. A meta study by Kumar and Goyal (2015) describes the most common ones. (I) Overconfidence, people are too confident in their own abilities, knowledge, and skills. This leads to excess trading and lower returns. (II) Disposition effect, investors are more likely to sell successful stocks and hold onto unsuccessful ones. A related bias is the sunk cost fallacy. (III) Herding, people are social beings and follows the herd, imitating the judgement and actions of others. (IV) Home bias, people believe they have superior knowledge about things in their immediate adjacency. This leads to excessive investing in domestic companies and employers.

Similarly, Kahneman and Tversky (1979) developed 'prospect theory', which found that people value gains and losses of the same amount differently. Specifically, our willingness to accept risk is influenced by whether it is framed as a loss or a gain. This is known as loss aversion. This is a stark contradiction to rationality. Thaler (1985) found that people label financial resources differently, though money is fungible. Subsequently, money is treated and spend differently. A phenomenon he dubs 'mental accounting'.

3.2.2 Noise Traders

Noise traders are often described as 'liquidity traders' or 'random allocators'. Noise traders provide liquidity in the market and are assumed to allocate their trades uninformed and randomly in most financial models, e.g. Grossman & Stiglitz (1980) and Garleanu & Pedersen (2018). On average, these noise traders always lose. This raises the obvious question: "why are they so stupid?" (O'Hara, 2003). Behavioural finance may provide the answer in that everyone is influenced by biases. However, researchers argue that noise traders are especially prone to biases and thus to making suboptimal decisions. This has profound implications for asset pricing and market efficiency.

De Long, Schleifer, and Summer (1990) develops a framework, which links the proportion of noise traders to the level of market efficiency. Their framework rests on the assumption that noise traders are particularly prone to biases and as a result thereof makes more irrational decisions and investments. Their framework suggests that as the proportion of noise traders increases, the level of market efficiency decreases. Consequently, prices will diverge from fundamental values. Concurrently, as the proportion of noise traders increase, the proportion of sophisticated investors and arbitrageurs decrease. Furthermore, due to the irrational behaviour of the noise traders, sophisticated investors are deterred from investing against them. Thus, market efficiency may decrease as the proportion of noise traders increase. Conversely, in the context of index investing, it is not farfetched to imagine that noise traders are especially prone to migrating away from active investing and towards passive investing (Israeli, Lee, & Sridharan, 2017). That is, the proportion of noise traders will decrease.

3.2.3 Market Anomalies

Due to the inability of human beings to make rational, optimal decisions, market anomalies have arisen. Many of these shortcomings challenge the idea of market efficiency, according to behaviouralists. Harvey, Liu, and Zhu (2015) conducted a meta study, which provides an overview of the most acknowledged anomalies identified so far. In addition to the size and value effects identified by Fama and French (1993), the most known anomalies are discussed below.

Return over the previous period seems to have a predictive effect on future returns. This is in contrast to weak form market efficiency. Heston and Sadka (2008) find that in the short-term, there is a return reversal. In particular, winners over the past month will go on to underperform, while losers will overperform. In the medium-term, Jegadeesh and Titman (1993) find that stocks which have done well in the past 3-12 months tend to overperform in the subsequent month. A strategy designed to exploit this anomaly realises a 12% compounded return on average per year. Carhartt (1997) and Asness (1994) come to similar findings. In the long-term, there appears to be yet another reversal. DeBondt and Thaler (1985) find that stocks which have done well over the past 60 months ('past winners'), excluding the past 12 months, underperform in the future, conversely 'past losers' overperform. They attribute this anomaly to initial overreaction, availability bias, and recency bias.

Strong proponents of market efficiency would, however, argue that the above anomalies are consistent with market efficiency. For example, as discussed earlier, Fama and French (1988) argue that stock prices are composed of a permanent and temporary component, where the drifts discussed above correspond to the temporary component. Furthermore, in a recent series of research papers originating in AQR Capital Management, the existence of the size (Alquist, Israel, & Moskowitz, 2018), value (Asness, Frazzini, Israel, & Moskowitz, 2015), and momentum (Asness, Frazzini, Israel, & Moskowitz, 2014) effects are discussed. While the research concludes that value and momentum are strong factors, providing abnormal return, especially when combined, the size effect is more scrutinised. While they put its existence into doubt, they also conclude that if it existed, it has diminished. Believers in market efficiency would argue that the market has efficiently, though slowly, adapted to these factors.

3.3 Beating the Market

If markets are perfectly efficient and reflect all information, then they cannot consistently be beaten – except by chance. However, if markets participants are influenced by behavioural biases, which influence decision-making, then prices may deviate from their fundamental value. Correcting these deviations will enable arbitrageurs to achieve an abnormal return. The debate concerning the ability to beat markets have rallied among researchers and practitioners. This debate was initiated, when EMH was proposed by Fama (1970), while Samuelson (1974) claimed that no money managers could consistently beat the market.

Beating the market is made particularly difficult by one condition: active investing before costs are a zero-sum game (Sharpe, 1991; French, 2008), i.e. the gain of one investor is the loss of another investor. That is, when costs are accounted for, aggregate returns are negative (Sharpe, 1991; French, 2008). Hence, holding the market and being passive will always, on average, be superior to active investing. However, as Pedersen (2018) proves, this may be true for one period, but in the long-term with market changes, the argument falls apart. In the long-term, even passive managers have to trade. Namely, by trading they have to adjust to the inclusion and exclusion of firms in an index.

Hence, if passive investors buy at a premium and sell at a discount, active managers can beat passive managers (before fees). This is often the case, when a security enters or exits an index (Petajisto, 2010). Furthermore, Pedersen (2018) identifies multiple additional cases in which an active manager can, in theory exploit trading by passive managers. Hence, Pedersen (2018) disproves Sharpe's (1991) and French's (2008) equality, showing that active management is not a zero-sum game. However, the vast majority of evidence seems to suggest that beating the market consistently is very difficult.

Samuelson (1974) wrote the first impactful paper on the ability of investors to beat the market. He showed that even the best money managers are not able to beat the market consistently. Subsequently, multiple esteemed scholars have researched this topic. Carhart (1997) find that the superior performance exhibited by mutual funds can be explained almost exclusively by the momentum effect. Furthermore, he proposes that his results do not suggest that mutual fund managers can, on average, consistently beat the market. More recently, Fama and French (2010) find that only a few funds are able to provide a benchmark-adjusted return large enough to cover their fees. Furthermore, when superior performance is exhibited, it is inconsistent. That being said, some scholars find contrasting evidence. Kosowski, Timmermann, Wemers, and White (2006) find that a significant part of stock pickers in mutual funds can pick stocks to cover their costs. Additionally, their performance was consistent.

Hedge funds, however, have proven to overperform more consistently. Kosowski, Naik, and Teo (2007) find that hedge fund performance cannot be explained by luck and their performance persists across annual horizons. Their results are supported by Jagannathan, Malakhov, and Novikov (2010) and Fung, Hsieh, and Ramadorai (2008). In private equity, the same pattern emerges. Kaplan and Schoar (2005) and Korteweg and Sorensen (2017) find that private equity consistently outperforms their benchmark index and return a positive return after fees. However, it should be noted that private equity and hedge funds represent perhaps the most sophisticated investors.

3.4 Efficiently Inefficient

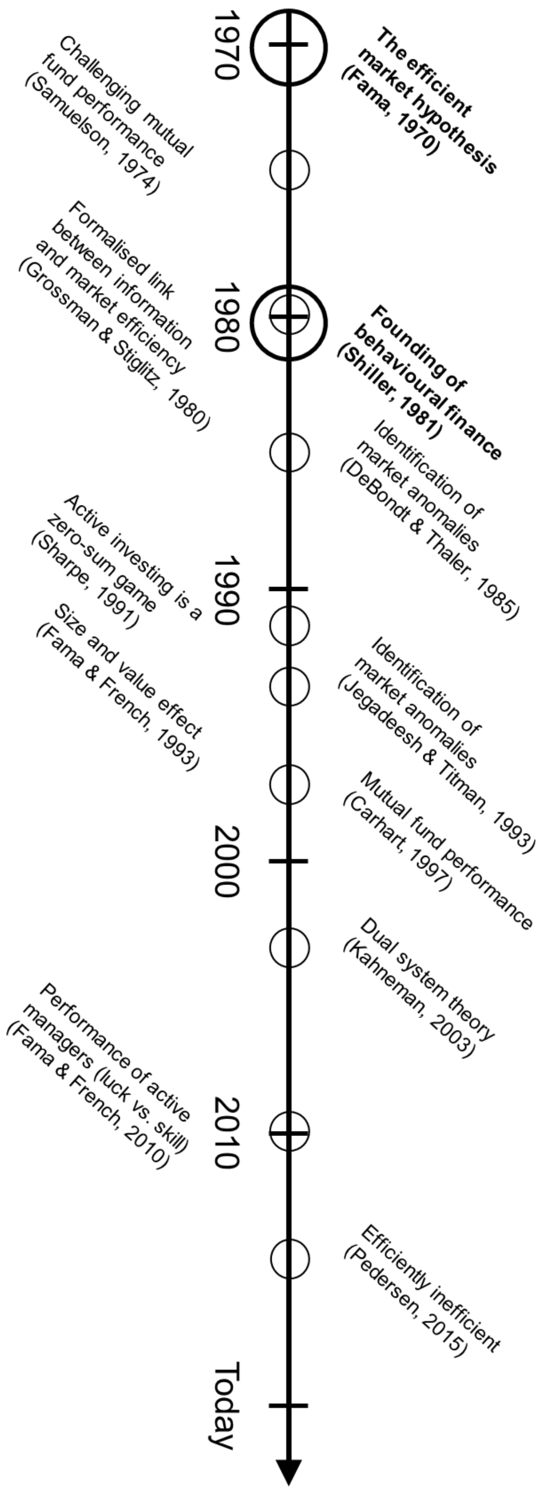
Having presented the vast literature on market efficiency, inefficiency, and the corresponding empirical results, which discuss the ability to beat markets, it is natural to ask: what are markets? For starters, it appears as if the Nobel committee have not picked sides in the debate. In 2013, they awarded the Nobel Memorial Prize in Economic Science to Eugene Fama (strong proponent of market efficiency) and Robert Shiller (strong opponent of market efficiency). However, the reality may be somewhere in between the two extremes. Pedersen (2015) and Garleanu & Pedersen (2018) presents a unifying view, which suggests a trade-off between the cost of information acquisition and return. Since markets can only be efficient, when they incorporate all information (Grossman & Stiglitz, 1980), this trade-off becomes paramount.

In their paper, Garleanu and Pedersen (2018) applies the Grossman-Stiglitz (1980) model on the market for asset managers. The theory assumes that there are overperforming asset managers and underperforming asset managers, just like there are overperforming securities and underperforming securities. However, acquiring the information enabling you to pick the right asset manager includes a costly search process. Skilled asset managers outperform after fees, while unskilled ones underperform (Evans & Fahlenbrach, 2012). Hence, there exists a trade-off between cost and return. Garleanu and Pedersen (2018) assume that noise traders take a random positon and thus invests partly in unskilled asset managers, while informed institutional investors invests solely in skilled asset managers.

The Garleanu and Pedersen (2018) model explains how and why some managers outperform in contrast to the consensus among strong efficient market proponets (Fama, 1970; Sharpe, 2013). In the model, outperformance is compensation for the higher search costs. Hence, they conclude with a compelling argument, consistent with Grossman and Stiglitz (1980): markets are so efficient that not everyone can achieve an abnormal return, however, they are so inefficient that active investors with a comparable advantage in information acquisition and computation can achieve an abnormal return and be compensated for their efforts. This becomes the efficiently-inefficient equilibrium level of market inefficiency – or the "equilibrium degree of disequilibrium" (Grossman and Stiglitz, 1980). This is consistent with the evidence on hedge fund (Kosowski, Naik, & Teo, 2007) and private equity performance (Kaplan & Schoar, 2005; Harris, Jenkinson, & Kaplan, 2014). There are larger search costs in those segments, which would enable them to achieve more consistent abnormal returns.

The preceeding discussion suggests that markets are not perfectly efficient and that they can move along the efficient-inefficient continuum. In particular, the level of market efficiency appears to be influnced by search costs, the number of traders, and the type of traders. The literature review so far is summarised in Figure 2 on the following page.

Figure 2 Overview of Literature Review Relating to Market Efficiency



Source: Own creation

3.5 Corporate Governance

A component of efficiency, which is often neglected is corporate governance. Investors have an innate responsibility of monitoring the companies, to reduce agency costs (Fama & Jensen, 1983), and to unveil new information. Institutional owners are becoming increasingly important as passive investing continues to increase (Aghion, Reenen, & Zingales, 2013). Hence, it is interesting to consider the mechanisms of monitoring and information acquisition in this context.

But first, we must consider how shareholders can exert power (Fichtner, Heemskerk, & Garcia-Bernando, 2017). Firstly, shareholders can influence decision-making directly by voting at the general assembly. Being an institutional investor with large block holdings enables them to influence management. Hence, this has great potential for index funds – in theory. Secondly, you can sell your shares, or 'walk the Wall Street walk'. This will negatively impact the share price and thus send a signal to management. Thirdly, investors can voice their concerns directly to management, e.g. during analyst calls (ibid.).

For index funds, however, the ability to exert power may be more limited. It is evident that they cannot sell individual shares since this would violate their mandate (Appel, Gormley, & Keim, 2016). Furthermore, it is questionable whether they will spend resources on monitoring and exercising their voting power. In theory, the funds could reduce their costs and fees, if they neglected to do so and only focussed on providing a cheap index fund. Passive owners could continue to be passive monitors and free ride on the efforts of active management. However, Appel, Gormley, and Keim's (2016) evidence suggests that this is not the case. They find that passive mutual funds have a positive effect on governance in terms of the number of independent directors, more equal voting rights, and removal of takeover defences.

In fact, passive managers do fulfil their fiduciary roles as monitors on behalf of clients and they do vote in proxy votes (Fichtner, Heemskerk, & Garcia-Bernando, 2017). Furthermore, because they seek long-term returns, they set high standards for managers. McCahery, Sautner, and Starks (2016) finds that 63% of large institutional investors have had direct discussions with management over the past five years, while 45% had discussions with the board of a company without the presence of management. In addition, the three largest institutional investors have publicly stated that they want to become active, monitoring shareholders, e.g. Vanguard publishes an annual stewardship report (Vanguard, 2018). In conclusion, passive managers do exert their power through the means they have.

By monitoring the companies actively and simultaneously having large voting blocks, passive investing may have a positive effect on market efficiency by unveiling more information.

3.6 Indexing and Its Effects

Indexing has been a defining feature of 21st century investing. In the recent decade, it has attracted the attention of many scholars. In particular, researchers have been examining and theorising about the effects of increased indexing. In this section, we consider the effects of indexing already exposed in the pre-existing literature.

Israeli, Lee, and Sridharan (2017) consider the effects of ETFs on market efficiency. They attempt to relate the increase in ETF ownership with lower levels of noise traders and the effect this will have on the price discovery function. They use three proxies for market efficiency: price synchronicity, future earnings response coefficient (the association between current firm-specific return and future earnings), and analyst coverage (ibid.). They find that more ETF ownership of a stock reduces its liquidity and thus widens bid-ask spreads. This disincentivises active managers to acquire information and adjust prices, because doing so is costlier. Furthermore, returns become more synchronised, returns are less predictive of future earnings, and analyst coverage falls. This evidence suggests that markets do in fact become less informationally efficient, when passive investing increases.

Anomalous findings from Belasco, Finke, and Nanigian (2011) suggest that companies in the S&P 500 are overvalued compared to industry peers. Specifically, they find that indexed firms have higher P/E ratios compared to peers outside the index. They go on to conclude that indexes may distort prices and drive them away from fundamental value (ibid.). Consistently, Petajisto (2010) find that firms, which are included in the S&P 500 or Russell 2000 experience an increase of 8.8% and 4.7%, respectively, following the announced inclusion to the index. Conversely, an exclusion is associated with a decrease of similar magnitude. Since inclusion or exclusion from an index does not change the perceived fundamentals of the firm, it is surprising to see such a significant effect. Hence, it would suggest that the required rebalancing of the index distorts prices (ibid.). However, in a more recent paper, Patel and Welch (2017) found that the initial price increase following inclusion has become smaller in the most recent years.

"Index investing is worse than Marxism", so claims the financial research firm Sanford C. Berstein and Co. LLC (Kawa, 2016). Their claim relates to the allocation of capital. Imagine for a moment that all investing is passive. This would have immense implications for the IPO market (Pedersen, 2018). Since passive investors would automatically buy the market, new entries would not be priced efficiently. Resultingly, everyone could take their firms public at the price they wished, and their stocks would be bought by the passive investors. Inevitably, financial markets would crash. This is obviously not an efficient allocation of capital. That being said, it is unlikely that the market will ever become exclusively passively invested. However, more passive investing may still reduce the capital available to venture capital (de Planta, 2017) and hinder creative destruction (Schumpeter, 1942).

The buying and selling of all index components simultaneously have profound implications for diversification, asset correlation, and construction of a portfolio. Sullivan and Xiong (2012) find that correlation of equities have increased as the proportion of passive investments have increased – Israeli, Lee, and Sridharan (2017) later confirms this. Since beta and systematic risk is a function of correlation, it is not surprising that they find that those two risk factors have increased as well. In addition, Wurgler (2011) finds that after a stock has been included in an index, its returns are more correlated with that index.

Furthermore, the value-weighted index approach creates some pitfalls for investors. Firstly, it leaves investors overexposed to certain industries. Consider the tech stocks in the S&P 500. When the sector was at its highest in the beginning of 2018, it represented 24% of the index, with 11% represented by the FAANGs (Leary, 2018). Since these stocks were rapidly growing, they had low book-to-market (B/M) ratios. Fama and French (1993) find that high B/M-ratios outperform firms with low B/M-ratios. Thus, indexing leaves investors overexposed to overvalued companies and underexposed to undervalued companies (Brown, 2018). Indeed, if that is the case, an equal weighted index or a reverse value-weighted index should be able to beat the value-weighted index. Brown (2018) shows that a reverse value-weighted index would have returned 11.4% since 1997, compared to 8.2% of the normal value-weighted index. Moreover, this overexposure may increase idiosyncratic risk, because large cap firms attract more capital regardless of their underlying performance (de Planta, 2017). Since, large cap firms are more similar, they will be exposed to the same risk factors and thus be subject to the same idiosyncratic risks. Hence, the benefits of diversification will be diminished (Fichtner, Heemskerk, & Garcia-Bernando, 2017).

3.7 Applied Methodology in the Literature

This paper conducts a test of market efficiency. We will rely on event studies of post-earnings announcement drift as a proxy for market efficiency. In this section, we will discuss how this methodology has been applied in the literature previously. Furthermore, we discuss the literature on factor models, since they constitute a key component of event studies, namely the models for 'normal return' and thus 'abnormal return'. Our study expands on the event study literature, by performing an OLS regression inspired by the difference-in-difference methodology. Hence, we also consider the literature on this methodology.

3.7.1 Event Studies & Post-Earnings Announcement Drift

Event studies have been widely applied in the financial literature. In particular, it has been used as a test of market efficiency. The most popular method of examining market efficiency is by conducting a test of post-earnings announcement drift (PEAD) using an event study (Kothari & Warner, 2007). The methodology was incepted in 1933 by Dolley according to Kothari and Warner (2007), but did not gain popularity until Fama, Fisher, Jensen, and Roll (1969) conducted a test of stock splits in 1969. Since then, the methodology has been soaring in popularity (Kothari & Warner, 2007). This rise has particularly been driven by more available data and more sophisticated models for measuring normal return. Short-horizon event studies are said to provide "[…] the cleanest evidence we have on market efficiency." (Fama, 1991). However, the results deriving from long-horizon event studies continue to be scrutinised (Kothari & Warner, 2007). Over the long-term, failure to appropriately adjust for risk can make large economical differences. No strong model has been developed to account for this risk. Furthermore, determining significance becomes difficult in the long-run because long-horizon returns depart from the normality assumption on which significance tests are based (ibid.).

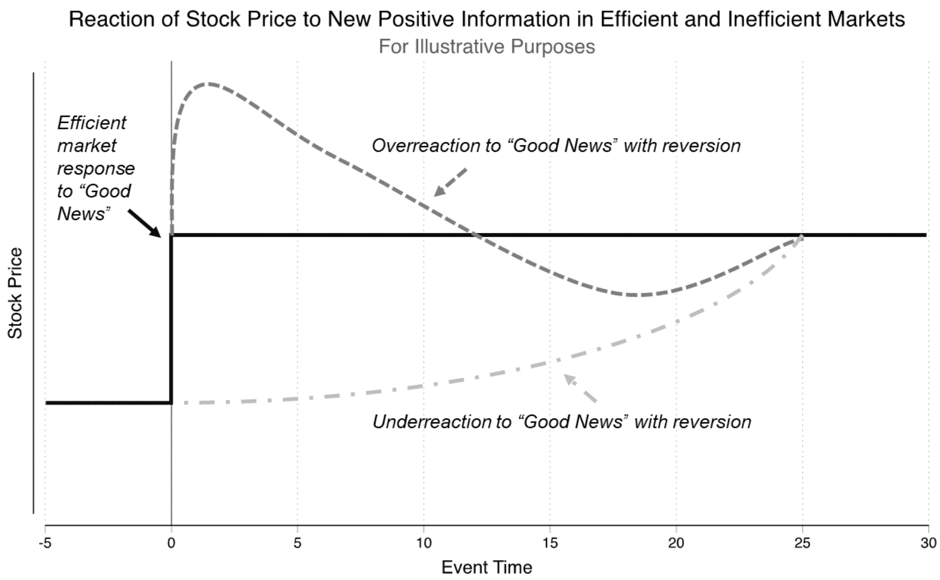
Event studies have been used to identify anomalies, to test for market efficiency, and to examine behavioural biases. PEAD has often been used as a proxy for market efficiency (Ball & Brown, 1968). If the drift following an earnings announcement is significant, it is considered a deviation from market efficiency. Since the anomaly was initially identified, researchers have discussed whether the drift was a result of a delayed price response or failure to account for risk in normal returns models. That is, whether the anomaly in fact contested the theory of efficient markets.

Using an event study, Bernard and Thomas (1989) provided some of the first evidence to suggest that PEAD is in fact a delayed price response and thus evidence of market inefficiencies. Similarly, Brandt et al., (2008) make use of this methodology, when testing whether new information from earnings announcement is efficiently reflected in the price.

DeBondt and Thaler (1985) conducted a long-term event study, when they identified long-term reversal and short-term momentum. In a more recent paper, Chen, Kelly, and Wu (2018) apply the PEAD methodology to test how market efficiency is influenced by exogenous shocks to the information environment. Interestingly, their results support the 'efficiently inefficient' hypothesis developed by Pedersen (2015). Chen, et al. (2018) find that when analyst coverage declines, hedge fund participation mitigates the reduction in information by increasing their information acquisition and trading behaviour. Hence, in their framework, there exists a substitution effect between sophisticated investors and public information providers.

If markets were perfectly efficient, they should react instantaneously and effortlessly to new information. However, that is rarely the case. As illustrated in Figure 3, stock prices often drift into place, either via an overreaction or an underreaction. This often takes place after an earnings announcement. This drift is what is tested in an event study of PEAD.

Figure 3 Illustrative Depiction of Efficient and Inefficient Reaction of Stock Price to New Information



Source: Own creation

Researchers have attempted to explain the drifts resulting from earnings announcements through behavioural finance. One explanation for the initial overreaction is the 'representativeness bias' identified by Kahneman and Tversky (1974). The bias occurs, when people attach too much value to recent events and overvalue their representativeness of the entire population. That is, recently published information is more representative than less recent information. Hence, traders overreact to this new information and overvalue the stock. Conversely, underreaction can be explained by the 'conservatism bias'. Edwards (1968) defines the bias as the inability of people to update their opinion in the light of new information. Hence, in the case of stock prices, new information is not processed accordingly, and as a result thereof prices react inefficiently.

3.7.2 Factor Models: CAPM, Fama-French Three Factor + Momentum

When conducting our event study, a critical component is the calculation of normal return. There are multiple ways to do this. One way is to rely on factor models. The most well-known factor model is the capital asset pricing model (CAPM). This is an equilibrium model, where the return of an asset is the linear function of its covariance with the market, β, as defined by Sharpe (1964) and Linter (1965). More formally, the CAPM can be expressed as follows:

$$CAPM = r_f + \beta * (r_m - r_f) \tag{1}$$

, where r_f is the risk-free rate, β is the covariance with the market, and $(r_m - r_f)$ is the market premium.

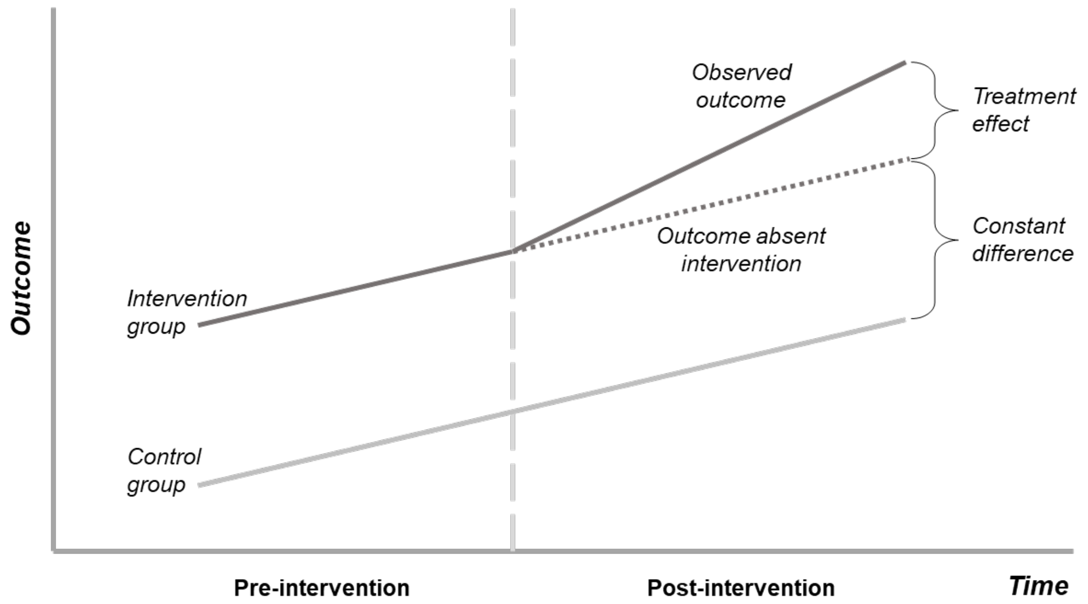
Fama and MacBeth's (1973) research supports the traditional CAPM by finding a positive relationship between risk and return as suggested by the CAPM. However, the relationship is weak for individual cases, and stronger for the overall market. They find that the net effect is linear, and thus conclude that the CAPM holds in the long-run. However, the literature critiquing the simplicity of the CAPM is far greater than the literature in support. For example, Roll (1977) argue that the CAPM cannot be successfully tested until the exact composition of the market is known. In a more recent paper, Fama and French (2004) argue that this is flawed and consequently applications using the standard CAPM is invalidated.

Nevertheless, the traditional CAPM continues to form the basis of modern asset pricing theory. Subsequently, multiple factors have been added to the model. Most famously Fama and French (1993) developed the Fama-French Three Factor model. This model elaborates on the CAPM by adding the size factor and the value factor. In the late 1990's, Carhart (1997) and Asness (1994) add a fourth factor, momentum. More recently, Frazzini and Pedersen (2014) added the betting-against-beta factor and Asness, Frazzini, and Pedersen (2018) added the quality-minus-junk factor.

3.7.3 Difference-in-Difference Estimation

Difference-in-difference test (D-in-D) is a statistical method, which compares two groups in an experiment (Hill, Griffiths, & Lim, 2011, pp. 275-86). The main merit of the method is that it attempts to controls for all the effects, which influence both groups. Historically, D-in-D have predominantly been used in social sciences to test the effect of different policy measures. This was also the case in the first influential study, where Ashenfelter and Card (1985) used D-in-D to estimate the effects of a governmental training program on subsequent participant earnings. They realised that to make any meaningful conclusion, they had to control for any economy-wide moves, which may influence trainee earnings during the observation period. Furthermore, there may have been a self-selection bias between those, who were undergoing the training and those who were not. Hence, to control for this, they devised a comparison group as control, i.e. they had a treatment group and a control group. This allowed them to isolate the effect of the treatment in the form of training as exhibited in Figure 4.

Figure 4 Illustrative Depiction of Difference-in-Difference Test



Source: Own creation

The figure illustrates how the two groups perform in the pre- and post-intervention period. The intervention group exhibits a treatment effect following this intervention. The treatment effect is the difference between the expected outcome, absent the intervention, and the observed outcome. The effect is thereby the difference in the difference.

Though the methodology is widely applied in social sciences, it is rarely used in a purely financial setting. However, it is sometimes used in the research on private equity. Bernstein and Sheen (2016) considers the effect of private equity buyouts on operational performance of restaurants. They use the results of health inspections as a proxy for operational performance as the dependent variable. That is, in their study, the treatment effect is the acquisition, while the control group are adjacent, similar restaurants. Hence, they are able to control for any effects, which may influence their dependent variable other than the treatment effect. That is, they test whether restaurants owned by PE firms perform better in health inspections compared to non-PE owned restaurants, while controlling for exogenous variables.

However, D-in-D is sometimes criticised for being too simplistic, e.g. as it is clear from Figure 4, it assumes that there is a constant difference between the two groups, absent the intervention. That is, that they are similar in all other manners except the intervention and they will be influenced in the same way by any shocks (Bertrand, Duflo, & Mullainathan, 2004). This has profound implications for the standard errors derived from an OLS regression model due to serial correlation. However, it has been widely ignored by researchers. Furthermore, the interventions are rarely random. Hence, there may be endogeneity between the interventions themselves (ibid.).

4 Hypotheses Development

In the following section, we will develop our hypotheses based on the extensive literature review conducted in the previous section. We provide theoretical foundations for three hypotheses: (I) more efficient markets, (II) less efficient markets, and (III) equally efficient markets. Broadly speaking, the hypotheses will be based on the interaction of indexing with noise traders, price discovery, and corporate governance. In particular, how indexing influences the cost of information acquisition and through that price discovery and market efficiency.

In this paper, we attempt to identify the effect of increased index investing on market efficiency. However, for this discussion to be valid, one has to adhere to the belief that markets can be less than perfectly efficient. This has been heavily discussed in the literature by researchers like Fama (1970) and Shiller (1981), who are two of the most prolific researchers on the subject. Yet, they do not see eye-to-eye on the subject. Fama is considered the father of the efficient market hypothesis, while Shiller was one of the first to provide opposing evidence. However, this discussion continues enthusiastically among researchers and practitioners alike.

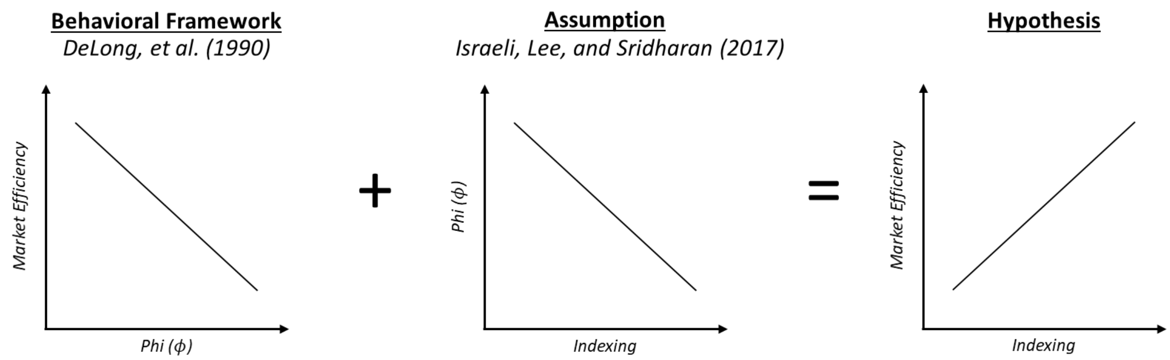
We submit to the understanding that markets are not perfectly efficient. Consequently, we can go on to consider what effect indexing has had on the level of market efficiency. That is, has the increased proportion of passive investments had an effect on market efficiency? Based on the literature review, we develop three hypotheses, which relates index investing to market efficiency. Hence, this paper contributes to the large pre-existing literature on market efficiency. Furthermore, we provide some of the first empirical evidence on the effect of index investing on market efficiency.

4.1 Hypothesis I – Increasingly Efficient

Our first hypothesis considers how noise traders and their migration towards passive investing has influenced market efficiency. We hypothesise that markets have become more efficient as a result of increased indexing. First, we consider how the migration of noise traders from active to passive investing has influenced market efficiency. Second, we discuss how more indexing may increase the ability to short-sell. Finally, the aggregation of ownership and its effect on corporate governance is considered in a market efficiency perspective.

Israeli, Lee, and Sridharan (2017) assume that noise traders are more likely than institutional investors to migrate towards passive investments, such as index investments. As we discussed in the literature review, noise traders are more prone to biases and thus makes less optimal asset allocations. It is our assumption that, as the proportion of passive investment has increased over the past years, the proportion of noise traders have declined. It is this assumption, which our first argument for more efficient markets rests upon. DeLong, et al. (1990) develops a framework, which links the proportion of noise traders (Phi, ϕ) to market efficiency. As the proportion of noise traders increases, the level of market efficiency decreases. This enables us to link indexing, noise traders, and market efficiency. This relationship is illustrated in Figure 5.

Figure 5 Illustration of Link between Indexing, Noise Traders, and Market Efficiency

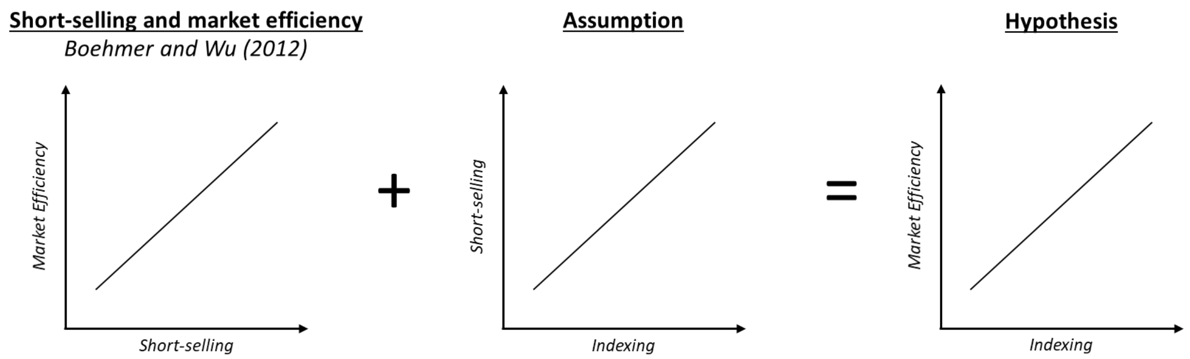


Source: Own creation

As the proportion of passive investing has increased, the proportion of stocks held by large institutional investors have also increased. In theory, this could have profound implications for the ability of investors to borrow and thus short stocks. In 2018, the asset manager Fidelity launched the first no-fee index fund (Rosenbaum, 2019). It is natural to consider how this is possible from an economic point-of-view. One source of income is the lending fee from lending stocks to short-sellers. The launch of Fidelity's no-fee index suggests that short-selling has become more prevalent.

Boehmer and Wu (2012) find that stock prices are more accurate, when short-selling is more prevalent. This is a result of improved price discovery. This enables us to link indexing, short-selling, and market efficiency. This relationship is illustrated in Figure 6.

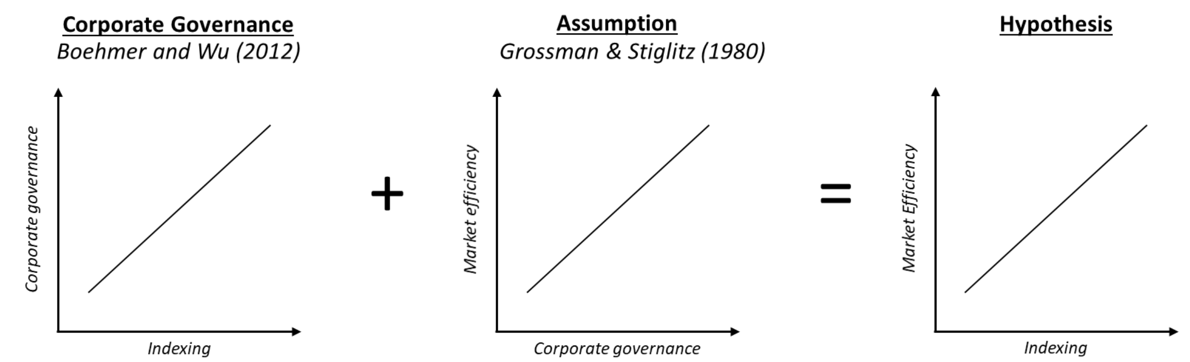
Figure 6 Illustration of Link between Indexing, Short-Selling, and Market Efficiency



Source: Own creation

As the proportion of index investing increases, the ownership of stocks will consolidate with a few large institutional investors. An increased ownership stake will increase their incentive to become active monitors of the company and to unveil new information (Garleanu & Pedersen, 2018). Fichtner, Heemskerk, and Garcia-Bernando (2017) find that passive mutual funds act as effective fiduciaries on behalf of their investors. Hence, they are effective monitors and unveil new information (Appel, Gormley, & Keim, 2016). Consequently, the price will reflect more information and the stock will be more efficiently priced (Grossman & Stiglitz, 1980). This enables us to link indexing, corporate governance, and market efficiency. This relationship is illustrated in Figure 7.

Figure 7 Illustration of Link between Indexing, Corporate Governance, and Market Efficiency



Source: Own creation

The effects of noise traders, short-selling, and corporate governance provide the foundation of our first hypothesis:

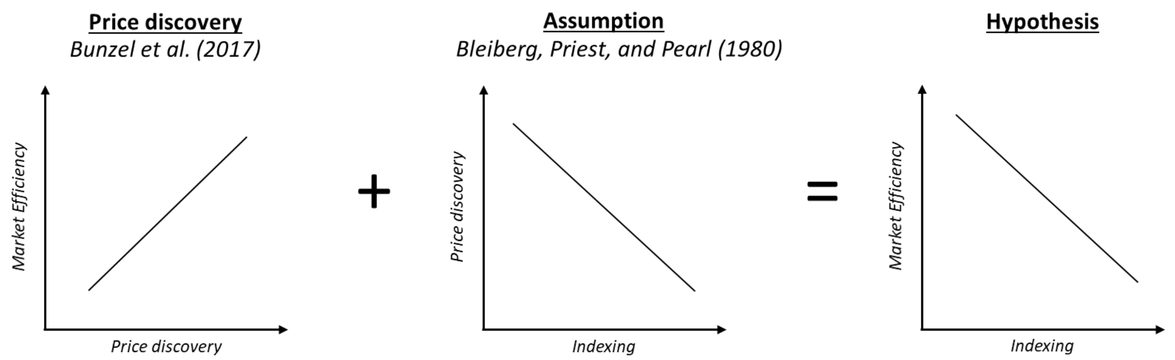
> *Hypothesis I: Markets have become more efficient as a result of increased indexing. This development is driven by the migration of noise traders, the increased prevalence of short-selling, and an increased monitoring power and incentive for information acquisition of institutional investors.*

4.2 Hypothesis II – Decreasingly Efficient

Contrary to Hypothesis I, Hypothesis II considers the negative effects of the migration of noise traders and that impact on liquidity and price discovery. Consequently, we hypothesise that markets have become less efficient as a result of increased indexing. The migration of noise traders from active to passive investing is hypothesised to have profound negative effects on price discovery and liquidity and thus on market efficiency.

The price discovery function relies on the wisdom of the crowds (Bunzel, et al., 2017). When indexing increases, noise traders will likely migrate from active to passive investing (Israeli, Lee, & Sridharan, 2017). When this takes place, the market place will consist of fewer active traders with less diverse opinions. This will reduce the wisdom of the crowds and thus hamper the price discovery function (Bleiberg, Priest, & Pearl, 2017). This enables us to link indexing, price discovery, and market efficiency. This relationship is illustrated in Figure 8.

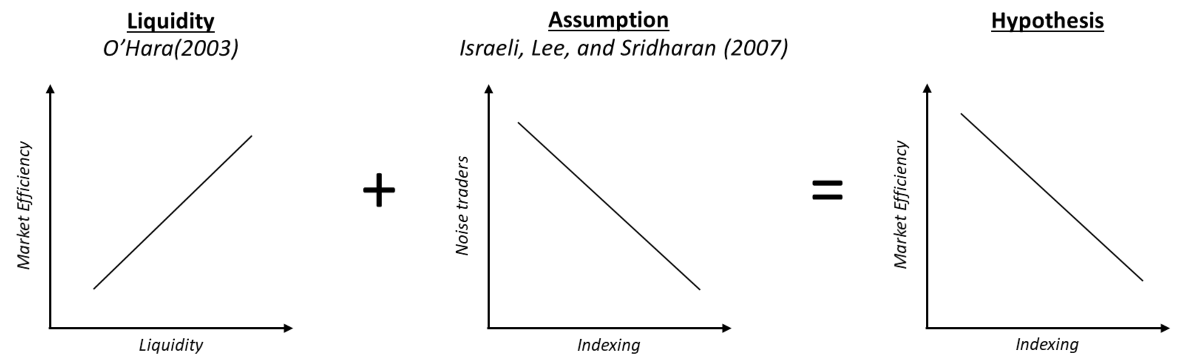
Figure 8 Illustration of Link between Indexing, Price Discovery, and Market Efficiency



Source: Own creation

If we continue to assume that noise traders will become less prevalent as indexing increases, we can also think of effects on the liquidity of markets. For markets to be liquid, there must be a high number of active buyers and sellers. Noise traders often provide this liquidity (O'Hara, 2003). Hence, when there are fewer noise traders, it can be assumed that liquidity will be lower. Furthermore, liquidity is a prerequisite for price discovery, and thus efficient markets (ibid.). This enables us to link indexing, liquidity, and market efficiency. This relationship is illustrated in Figure 9.

Figure 9 Illustration of Link between Indexing, Liquidity, and Market Efficiency



Source: Own creation

These effects of price discovery and liquidity provide the foundation for our second hypothesis:

> *Hypothesis II: Markets have become less efficient as a result of increased indexing. This development is driven by the migration of noise traders and the subsequent negative effect on price discovery and liquidity.*

4.3 Hypothesis III – Efficiently Inefficient

One final option, which we have to consider is that all, or none, of the previously introduced effects influence market efficiency. Consequently, the net effect is negligible. Thus, markets will remain equally efficient and markets are unaffected by increased indexing. If the net effect does not materially influence information acquisition costs or expected returns, we should not see market efficiency move along the efficiently-inefficient frontier (Garleanu & Pedersen, 2018).

> *Hypothesis III: Markets remain equally efficient as a result of increased indexing. The net effect of increased indexing is negligible and does not influence the level of market efficiency.*

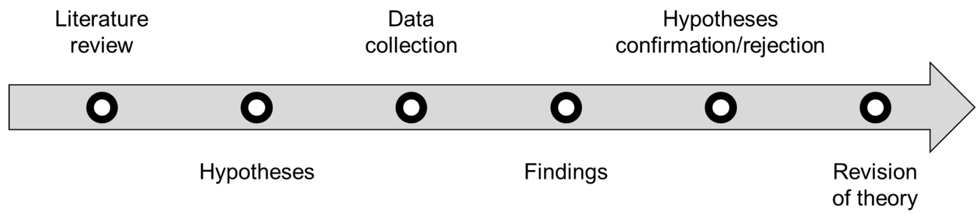
5 Methodology

This section of the paper will focus on the research design and methodology applied in our study. We strive to make the study valid, replicable, and reliable. First, we will discuss our scientific approach. Second, we will present the event study methodology, and discuss how it is used to measure market efficiency. Third, we will present the difference-in-difference methodology, which inspires our regression analysis. Fourth, we will define the groups used to tease out the effects of indexing. Fifth, we will discuss how we are going to apply the event study and regression analysis. Finally, our data management and collection processes are discussed. We will revert to the data preparation in the subsequent section.

5.1 Scientific Approach

Bryman and Bell (2011) present two research methods: 1) deductive and 2) inductive. Our thesis is deductive in nature. That is, our approach deduces hypotheses based on previous research. We then go on to test these hypotheses based on available data, before we accept or reject the hypotheses. The approach is illustrated below in Figure 10.

Figure 10 Scientific Approach



Source: Own creation inspired by Bryman and Bell (2011)

Furthermore, Bryman and Bell (2011) emphasise three important criteria when conducting research: 1) validity, 2) reliability, and 3) replicability. Validity relates to whether the methodology measures what it is intended to. That is, whether the hypotheses can be validly confirmed or rejected based on the results of the analysis. Furthermore, whether the results can be extrapolated beyond the sample and whether they have any impact. Reliability relates to whether the analytical methods are stable, and the data collection is reliable. Replicability is concerned with the ease of reproducing the study by other researchers (ibid.). This paper strives to adhere to these three criterions. Therefore, the data will be gathered from trusted and accessible sources, while the methodology will build on existing research. Furthermore, to ease the process of replication, we will document our approach thoroughly and provide the full code in Appendix V to Appendix IX.

5.2 Event Study

The event study methodology has been widely used in finance, economics, and accounting to determine the effects of an event. In finance, the main purpose of the event study is to determine whether the price of a stock reacts efficiently to news. According to the efficient market hypothesis, the price of a stock should react immediately and effortlessly to reflect new information (1970). An event study tests whether this is in fact the case. It does so by defining normal return, observing actual return, and computing abnormal return. It then tests whether cumulative average abnormal return for a given period is significant. A significant drift following an event is an indication of an inefficient market. Arriving at the drifts and their significance levels are the objective of the event study. In this section of the paper, we will discuss the event study methodology. We will rely on Campbell, Lo, and MacKinlay's (1997, pp. 149-78) seven-step procedure (5.2.1 through 5.2.7) for event studies.

5.2.1 Event Definition

First and foremost, the event to be investigated must be defined. In the finance literature, the most prevalent events tested are earnings announcements, stock splits, M&A notices, and dividend announcements. Having defined the event, one must define the duration of the event window. Typically, the event window is expanded across the event day to include the subsequent day. This is done to capture effects taking place after market closures. It is common to include the period prior to the event in the event window as information can leak prior to the announcement. This is done to capture any informational effects taking place prior to the official event date. In long event studies, it is custom to use an event window, which is symmetrical around the event date. For instance, Campbell, et al. (1997) employ an event window ranging from [-30;30]. The cumulative average abnormal return is aggregated across the entire period. However, if one is only concerned with the post-event effect, then it is insignificant whether the event window starts 5 or 30 days prior to the event.

5.2.2 Selection Criteria

When the event has been defined, one must consider the selection of firms in the study. These firms must be selected based on some carefully selected criteria. These criteria must reflect the scope and purpose of the study. However, these criteria are often constrained by data availability. The criteria could include, among other, listing on an exchange, market capitalisation, legal domicile of the firm, industry code, number of employees, etc. When selecting the sample, it is paramount that you are aware of different biases. Inability to consider e.g. selection bias and survivor bias may hamper the validity of the results. Furthermore, it is paramount that the data collection remains random, such that the sample is random. When the sample has been selected, it is common to provide an elaborate overview of the data.

5.2.3 Normal and Abnormal Return

To determine the effect of an event, it is required to create a measure of abnormal return. Conceptually, abnormal return can be thought of as the actual return less the 'normal' return:

$$\epsilon_{it}^{*} = R_{it} - E[R_{it}|X_t] \tag{2}$$

, where ϵ_{it}^{*} is the abnormal return, R_{it} is the actual return, $E[R_{it}|X_t]$ is the estimated normal return for firm *i* in time period *t*.

Normal return is defined as "[…] the return that would have been expected if the event did not take place […]" (Kothari & Warner, 2007). Hence, it is paramount for the validity of the model that normal returns are defined optimally. There are multiple ways to calculate normal returns. In the following sections, we will consider the most common ones and discuss their advantages and disadvantages. Two distinct groups of models can be defined (Campbell, et al. 1997): economic models and statistical models.

5.2.3.1 Economic Models of Normal Return

Economic models are characterised by being based on the assumptions regarding investors and in contrast to statistical models, they are not based purely on statistical assumptions (Campbell, et al. 1997). The advantage of economic models is the ability to calculate more accurate measures of normal return by applying parametric economic restrictions. The two most common economic models are the capital asset pricing model (CAPM) and the arbitrage pricing model (APT).

The CAPM assumes that the expected return of a security is a function of the return on a market portfolio, while the intercept is the risk-free rate (Sharpe, 1964; Lintner, 1965). However, this assumed intercept adds a restriction to the model. This has negative effects on the variance of the error term, which consequently will be larger compared to the market model (Cable & Holland, 1999). The CAPM-model was commonly used in event studies in the 1970's, but has subsequently decreased in popularity. Because it has become relatively costless to use the market model, the use of the CAPM model has almost ceased. Cable and Holland (1999) found that the CAPM was only preferred to the market model in 14% of their tests, while it was wrong in 43%.

The APT-approach, however, uses multiple factors and often builds on the CAPM. To calculate normal returns, the model assumes that the expected return is determined by its covariance with multiple other factors (Campbell et al., 1997). In contrast to the CAPM, a properly chosen APT does not impose false restrictions on mean returns. However, the additional factors complicates the implementation of the approach in the event study (ibid.). One popular model is the Fama-French three factor model (1993):

$$3\,Factor = \left(R_f + \alpha + \beta_1 * \left(R_m - R_f\right)\right) + \beta_2(SMB) + \beta_3(HML) + \epsilon \tag{3}$$

, where R_f is the risk-free rate, $\left(R_m - R_f\right)$ is the market premium, *SMB* is an abbreviation for small-minus-big (the size factor), *HML* is an abbreviation for high-minus-low (the value factor), and ϵ is the error term.

This model extends on the traditional CAPM by accounting for the size and value factor. The model can be expanded further by adding factors like momentum (MOM) (Asness, 1994; Carhartt, 1997), betting-against-beta (BAB) (Frazzini & Pedersen, 2014), and quality-minus-junk (QMJ) (Asness, Frazzini, & Pedersen, 2018). However, adding additional factors beyond the market factor rarely adds much explanatory power (Cable & Holland, 1999). Nonetheless, the expanded model can be used as a robustness check.

5.2.3.2 Statistical Models of Normal Return

In contrast to the economic models, statistical models rely solely on statistical assumptions about behaviour of asset prices. Statistical models assume that asset returns are jointly multivariate normal and independently and identically distributed through time (Campbell, et al., 1997). We will go on to discuss the constant-mean-return model and the market model.

The constant-mean-return model is perhaps the simplest model. The model assumes that normal return is the mean return in the estimation period, plus an error term. Despite its simplicity, Brown and Warner (1980) finds that it often yields similar results to more sophisticated models. The model can be defined as follows:

$$R_{it} = \mu_i + \epsilon_{it} \tag{4}$$

, where R_{it} is the return in period *t* for security i, μ_i is the mean return for equity i, and ϵ_{it} is the error term for equity i in period *t*. Furthermore, it is assumed that the error term is 0.

The market model in contrast assumes that the return of a security is linearly associated to the return of a market portfolio. In comparison with the mean-market-model, it is more sophisticated. The model can be defined as:

$$R_{it} = \alpha_i + \beta_i * R_{mt} + \epsilon_{it} \tag{5}$$

, where R_{it} is the return in period *t* for security *i*, R_{mt} is the period *t* return of the market index, ϵ_{it} is the error term for equity *i* in period *t*, while α_i and β_i are the parameters of the market model. Again, we assume that the error term is 0. This model is very similar to the CAPM. However, the non-restrictive intercept is a key distinction, which greatly enhances its predictive ability.

A prerequisite for the applicability of the model is that R_{mt} captures a broad-based stock market index. Campbell et al. (1997) suggest that variables computed by CRSP, such as value-weighted return with dividends, are popular choices. The main merit of the model is that it manages to reduce the error term by accounting for variation in equity prices, which are prompted by general market movements. In practice, it does so by calculating a beta-value for each stock in the period prior to the event. This should improve the measure of normal returns and thus increase the ability to detect event effects (ibid.).

5.2.3.3 Abnormal Return

Abnormal return is defined as the return above or below the return expected by the normal return model (Campbell, et al., 1997). In our thesis, we will employ the market model. Hence, abnormal return is defined as:

$$AR_{it} = R_{it} - \hat{\alpha} - \widehat{\beta_i} * R_{mt} \tag{6}$$

, where AR_{it} is abnormal return for security *i* at time *t*, R_{it} is the actual return of security *i* at time *t*, and $\hat{\alpha} - \hat{\beta} * R_{mt}$ is the expected normal return of security *i* at time *t*.

We define average abnormal return (AAR) and its variance as (Kliger & Gurevich, 2014):

$$AAR_t = \frac{1}{N}\sum_{i=1}^{N} AR_{it} \tag{7}$$

$$var(AAR_t) = \frac{1}{N^2}\sum_{i=1}^{N} \sigma^2(AR_{it}) \tag{8}$$

We then go on to aggregate over time. We do so by calculating cumulative average abnormal return (CAAR). We define CAAR and its variance as (Kliger & Gurevich, 2014):

$$CAAR(t_1, t_2) = \sum_{t=t_3}^{t_4} AAR_t \tag{9}$$

$$var(CAAR(t_1, t_2)) = \sum_{t=\mathrm{t}_3}^{t_4} var(AAR_t) \tag{10}$$

Having defined CAAR and its variance, we now have the tools required to test the significance of the event. This enables us to make inferences regarding market efficiency.

To draw inferences from the event, we need to aggregate abnormal return over time and across securities. Table 1 provides an intuitive overview of this process. Abnormal return is aggregated across time downwards in the table, while it is aggregated across securities horizontally in the table. The first step is to aggregate over securities. We do that by defining average abnormal return on a given day.

Table 1 Aggregation of Cumulative Average Abnormal Returns from X to Y

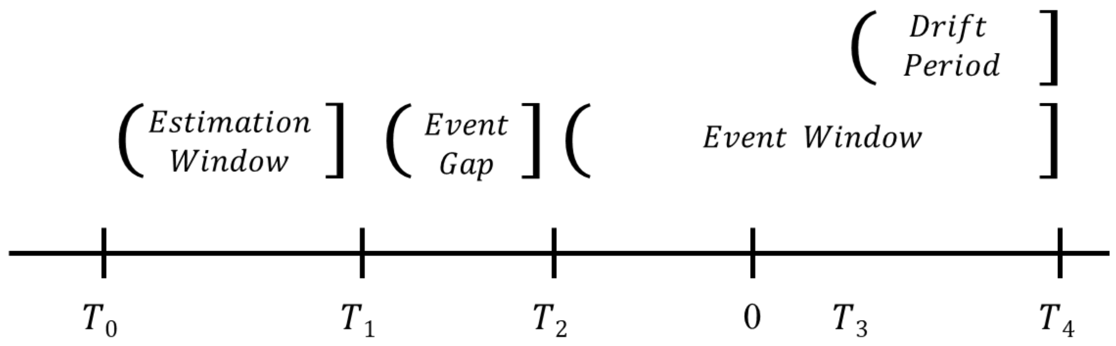
Day / / Event	1	2	3	…	N	Average
T=-X	AR	AR	AR		AR	AAR
…						
-1	AR	AR	AR		AR	AAR
0	AR	AR	AR		AR	AAR
1	AR	AR	AR		AR	AAR
…						
T=Y	AR	AR	AR		AR	AAR
[X; Y]	CAR	CAR	CAR	…	CAR	CAAR

Source: Own creation

5.2.4 Estimation Procedure

The normal performance must be estimated over a time period. This period is generally referred to as the estimation window. The length of the estimation window varies greatly from study to study and there is no general consensus about its duration. However, it is most common to use a period prior to the event window to estimate it. Furthermore, to avoid the event from influencing the normal return predictions, it is common to include a gap between the estimation window and the event window. This is referred to as the event gap. The entire event study timeline is illustrated in Figure 11.

Figure 11 Illustrative Depiction of Event Study Timeline



Source: Own creation inspired by Campbell, et al. (1997, p. 151)

5.2.5 Testing Procedure

The literature distinguishes between two types of tests: parametric and non-parametric. Parametric tests assume that daily returns follow a normal distribution, while non-parametric tests do not adhere to this assumption. Fama (1976) as well as Brown and Warner (1985) show that daily returns often exhibit fat-tails, i.e. they are not normally distributed. Furthermore, parametric tests do not account for event-induced volatility or cross-sectional correlation of abnormal returns. However, they are still valued and applied by multiple event study scholars (e.g. Brown & Warner, 1985; Ahern, 2009; Kliger & Gurevich, 2014). Ahern (2009) cites standardised t-test as one of the leading parametric tests. Furthermore, it is widely applied in literature (ibid.). Therefore, this thesis will apply the standardised t-test, θ_1, as our parametric test. The t-test for testing CAAR can be defined as follows (Kliger & Gurevich, 2014):

$$\theta_1 = \frac{CAAR_{(t_1;t_2)}}{\sigma(CAAR_{t_1;t_2})} \sim N(0,1) \tag{11}$$

, where $CAAR$ and $\sigma(CAAR_{t_1;t_2})$ are defined by Equation (9) and (10).

This measure provides a t-statistic for CAAR in period t_1 to t_2. The drift is deemed significant if the t-statistic surpasses the critical t-value. If that is the case, we can conclude that the drift is significant and thus suggests a violation of market efficiency.

5.2.6 Presentation of Results

When the abnormal returns have been accumulated and tested, they must be presented in an eligible manner. The returns will be presented in a table as well as by a graph, which is customary in research utilising event studies. With these we will attempt to make inferences about the resulting significance levels and effects.

5.2.7 Interpretation and Conclusions

Finally, the results of the event study methodology will be interpreted. Particular attention is given to the significance of the CAAR drifts. Conclusions should relate to the research questions proposed and the hypotheses developed. Furthermore, it is important to consider new insights and implications of the study. Finally, we will discuss the robustness and limitations of results, the event study methodology, and the inputs used in the particular study.

5.3 Regression Analysis

In this section, we will discuss the methodology underpinning our regression analysis. First, we will explain the underlying assumptions and applications of the difference-in-difference estimation. We take inspiration from this. However, our test is not strictly a D-in-D test, because we do not have a clear pre- and post-treatment period. Second, we go on to describe the applied OLS regression and its underlying assumptions.

5.3.1 Difference-in-Difference Estimation

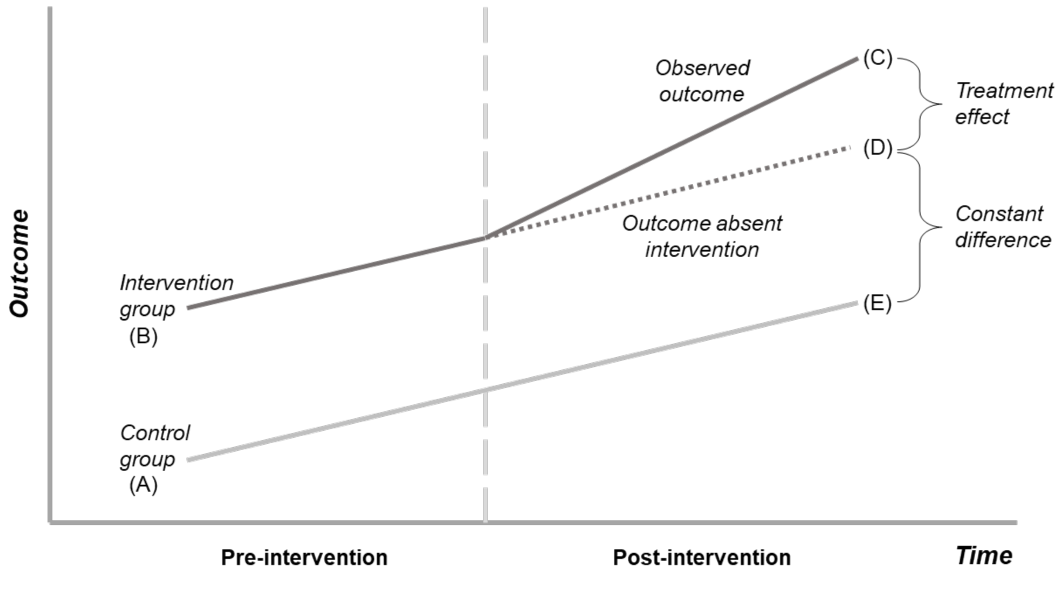
When making decisions, it is natural to think, “What would have happened if I had done B rather than A? Or what would have happened, had I done nothing?”. In terms of policy, politicians and economists often ask themselves similar questions. In social sciences, we can often use laboratory experiments to observe outcomes under different circumstances. However, in real life natural experiments, the counterfactual is impossible to observe (Hill, Griffiths, & Lim, 2011, pp. 275-86). However, in 1985, Ashenfelter and Card developed an econometric approach, which uses a control group to predict the counterfactual result and thus enables us to tease out the effect of a policy change (Ashenfelter & Card, 1985). This method was called difference-in-difference estimation (D-in-D).

That is, rather than considering the changes in the absolute level of one variable, the D-in-D estimator considers the changes in the difference between two or more groups.

The objective of D-in-D is to tease out the effect of an intervention, while controlling for underlying trends. For instance, what effect does the change in minimum wage have on unemployment. If one only observed the unemployment level before and after the intervention, there could be a plethora of confounding variables, which could influence the outcome other than the intervention. Hence, the D-in-D estimation controls for this by using a control group. For instance, one example could be two adjacent American states, which are similar on all material parameters except the intervention. The state where the intervention takes place is the treatment group, while the adjacent state is the control group. This is a classical natural experiment, where the D-in-D approach is applicable (consider e.g. Card & Krueger, 1994).

D-in-D relies on one strong underlying assumption. Absent the intervention, the control group and the treatment group will follow a common trend. The methodology rests on this assumption (Angrist & Pischke, 2015, pp. 175-208). This assumption can be tested by observing data prior to or following the intervention. If the two groups exhibit the same trend, it is a fair assumption. In practice, it can be done by making a visual comparison of the involved groups (ibid.). More formally, the assumption can be tested by including interaction of time indicator variables and the treatment effect in the pre-treatment period.

Figure 12 Illustrative depiction of Difference-in-Difference estimation



Source: Own creation inspired by Hill, Griffiths, & Lim, 2011, p. 282

To fundamentally understand the D-in-D estimation, it is useful to consider an illustrative depiction as in Figure 12. On the y-axis, we have the outcome, which in our earlier example would be unemployment, while the x-axis represents time. Before the treatment, we can observe point $y = B$ and after the treatment we can observe point $y = C$. However, we cannot separate out the effect of the treatment. Thus, we introduce the control illustrated by the line between $y = A$ and $y = E$, $\overline{AE}$. The common trend is represented by the dashed line connecting point $y = B$ and $y = D$, $\overline{BD}$. This line follows the same trend as the control group, $\overline{AE}$. To isolate the treatment effect, $\delta = \overline{CD}$, we need the inputs from A, B, C, and E. The difference-in-difference effect then becomes (Hill, Griffiths, & Lim, 2011, pp. 275-86):

$$\begin{aligned}\delta = \quad & (C - E) - (B - A) \\ & \left(treatment_{after} - control_{after}\right) \\ & \qquad - (treatment_{before} \\ & \qquad - control_{before})\end{aligned} \tag{12}$$

So, the estimator becomes the difference in the difference for the before and after outcome for the treatment and control groups. The D-in-D estimator, δ, can then be found using a regression.

$$\begin{aligned}y_{it} = \beta_1 + \beta_2 TREAT_i + \beta_3 AFTER \\ + \delta(TREAT * AFTER) + \epsilon_{it}\end{aligned} \tag{13}$$

, where β_1 is the outcome for $control_{before} = A$. TREAT is an indicator variable for inclusion in the treatment group, which accounts for fixed differences between the groups. AFTER is an indicator variable for time, which controls for trends which are common across all groups. Equation (12) can then be rewriten as:

$$\delta = \left((\beta_1 + \beta_2 + \beta_3 + \delta) - (\beta_1 + \beta_3)\right) - ((\beta_1 + \beta_2) - \beta_1) \tag{14}$$

5.3.2 OLS Regression

Ordinary Least Squares (OLS) is a method for which one can estimate the coefficients of $\beta_0, \beta_1, \dots, \beta_k$ in the multiple regression model (Stock & Watson, 2015). The estimators that minimize the sum of squared mistakes for the regression are called OLS estimators and are denoted $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$. When minimizing the sum of squared mistakes, the OLS estimators are chosen so that the estimated regression line is as close as possible to the observed data (ibid.).

OLS is a common regression method. There are four important assumptions that needs to hold for the OLS estimators $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ to be jointly normally distributed for large samples.

1. u_i has conditional mean zero given $X_{1i}, X_{2i}, \dots, + X_{ki}$; that is, $E(\, u_i | X_{1i}, X_{2i}, \dots, + X_{ki}) = 0$
2. $(X_{1i}, X_{2i}, \dots, + X_{ki}, Y_i)$, i = 1, …, n, are independently and identically distributed (i.i.d.) draws from their joint distribution.
3. Large outliers are unlikely: $X_{1i}, X_{2i}, \dots, + X_{ki}$ and Y_i have nonzero finite fourth moments.
4. There is no perfect multicollinearity

Source: Stock and Watson (2015, p. 247)

The first assumption implies that on average the population Y_i falls on the population regression line. Therefore, the expected value of u_i, the error term, is zero for any value of the regressors. The second assumption holds if the data for the regression is collected through simple random sampling. Thus, the sample collection determines if the data is independently and identically distributed. If the third assumption holds it is unlikely that the statistical inferences drawn from the OLS regression is distorted by a few observations. This can be controlled through investigating the distribution of the variables in the sample. The fourth assumption of the least squares assumptions is about multicollinearity. Multicollinearity is when one of the regressors is a perfect linear function of the other regressors (Stock & Watson, 2015). This cannot be the case for the assumption to hold. Modern statistical applications like STATA and R will pick up multicollinearity and remove the regressor that is a linear extension of the others. This can also be controlled through examining how the data within the sample is correlated between each other.

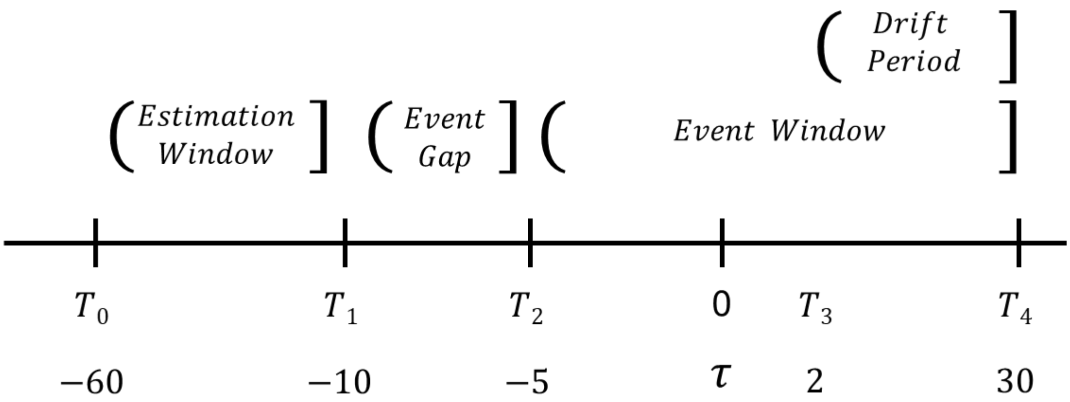
5.4 Application of Models

In this section of the paper, we will provide an overview of how, we are going to apply the methodologies presented above. We will start out by describing the inputs to the event study. First, we will discuss the inputs to a test of the underlying trend for the news groups (Test 1). Second, we will apply the methodology to our research question (Test 2). Third, we will discuss the inputs to the regression model (Test 3).

5.4.1 Test 1 – General Event Study of Underlying Trend

First, we want to test whether there is an underlying trend in stock prices. To determine this, we will conduct an event study of post-earnings announcement drift (PEAD). Hence, we define the *event* as the earnings announcement. This event is the centre of the event study timeline, illustrated in Figure 13.

Figure 13 Illustrative Depiction of Event Study Timeline



Source: Own creation inspired by Campbell, et al. (1997, p. 151)

The inputs to our event study are based on prior research and then modified to fit the requirements of this particular study. We set the *estimation window* to the period [-60; -10]. The objective of the estimation window is to provide adequate data to enable the creation of a credible normal return measure. Hence, the longer an estimation window the better. However, because the event takes place four times annually, we must make sure that the estimation window does not overlap with another event. Therefore, we find that this 50-day estimation window is optimal. The *event window* is defined as the period from 5 days prior to the earnings announcement to 30 days after, [-5; 30]. Our study is only concerned with the aggregation of CAAR following the event, so we reject the convention of choosing a symmetric event window. We then insert an *event gap,* [-10; -5] between the estimation window and the event window to avoid that the event influences the estimation. The *drift period* is then defined as the period from 2 days after the event to 30 days after the event, [2; 30]. We select day two to avoid the initial effect of the announcement, while we allow a long drift to aggregate sufficient data.

To model normal returns, we use the market model. This model appears to be the most widely adopted in literature (e.g. Asquith & Mullins, 1986; Bayless & Chaplinsky, 1996; Campbell, et al., 1997). We will then go on to test whether AAR and CAAR of the drift period is significant using the standardised t-test as defined in Equation (10). Again, we choose this test because it is widely applied in literature (e.g. Brown and Warner, 1985; Ahern, 2009; Kliger & Gurvich, 2014).

A prerequisite for conducting an event study of PEAD is the segregation of event reactions into news groups. Typically, the events are split into three groups: Neutral News, Good News, and Bad News. The segregation is based on a criteria for abnormal returns around the event. Berkman and Truong (2008) find that more than 40% of earnings announcements of Russell 3000 firms in the period 2000-2004 have been made after market closure. Hence, effects on the stock price are not observed until one day later. Furthermore, information or expectations often spill into the market place prior to the earnings annoucement. Thus, there is often an effect prior to the earnings annoucement date (Campbell, Cowan, & Salotti, 2010). This has wide implications for how researchers can measure the earnings announcement effect and thus define different news groups. To mitigate this, we define the news groups based on a three-day window and define CAR3 (Novy-Marx, 2015) as the three-day cumulative logarithmic abnormal return:

$$CAR3 = \Sigma_{j=-1}^{1}\left(\log(R_{i,t+j} - R_{m_{i,t+j}})\right) \tag{15}$$

, where R_{it} is the actual return of security i at time *t+j* (captured by the variable RET) and R_{mt} is the market return at time *t+j* (captured by the variable VWRETD).

That is, CAR3 captures the abnormal return from the day prior to the event, $t-1$, to one day following the event, $t+1$. To define the news groups, Chan, Jegadeesh, and Lakonishok (1996) suggest taking a decile approach. Good News are defined as the upper 10% decile of CAR3 following an earnings announcement, while Bad News are defined as the lower 10% decile. Chordia, Goyal, Sadka, Sadka, and Shivakumar (2009) as well as Bernard and Thomas (1989) employ a similar approach. In this paper, we will use the same methodology.

5.4.2 Test 2 – The Effect of Indexing

The objective of the event study is to determine whether the efficiency of stock prices differ based on the extent to which the stock is owned by index funds. To determine the effect of indexing, we employ the same methodology as above, but conduct an additional group definition. We split the data into groups characterised by their indexation, 'indexation groups'. The definition of these groups are paramount to the validity of this study. Therefore, we have devoted a full section to the matter, please refer to Section 5.5 for elaborate group definitions. Finally, to examine the effect over time, we will also conduct event studies of different time periods. Whereas the general event study tests whether the groups exhibit different significance of drifts over a 30-year period, splitting the data into multiple timeperiods will allow us to see how this metric has changed throughout time.

Due to the size of the event study we will conduct and the processing power this would require, we have decided to use the event study application at WRDS. This application requires us to determine the inputs. In Section 6, we discuss the extensive effort required to create the data files. Subsequently, we will also have to aggregate CAAR and determine the significance of the drift, CAAR [2;30]. Hence, it is only the event study itself, which will be outsourced to the platform.

5.4.3 Test 3 – Regression Analysis

To attempt to draw further inferences from the event study we will conduct a third test. This will enable us to isolate the effect of increased indexing on S&P 500 firms vis-à-vis a basket of similar firms assuming that they follow a common trend.

In Test 3, the following two regression models will be tested utilising STATA.

$$(1) \qquad \widehat{CAR}3_{it} = \hat{\beta}_0 + \hat{\beta}_1 SP_{it} + \hat{\beta}_2 t + \hat{\beta}_3 SP\_x\_t + \hat{\epsilon}$$

$$(2) \qquad \widehat{CAR}3_{it} = \hat{\beta}_0 + \hat{\beta}_1 SP_{it} + \hat{\beta}_2 t + \hat{\beta}_3 (SP\_x\_t) + \hat{\beta}_4\, ln\_MV + \hat{\epsilon}$$

The dependent variable in both regression models are CAR3. CAR3 is the logarithm of the abnormal return summed over the day leading up to the event, the event day, and the day after the event. We choose CAR3, because it is a good proxy for earnings response, and thus the extent to which markets react efficiently to earnings announcements. For regression model (1) the independent variables are the S&P 500 dummy variable, *SP*, a time indicator, *t*, and an interaction variable between the S&P 500 dummy and the time indicator, *SP_x_t*. The same goes for regression model (2) but we additionally want to utilise the natural logarithm of Market Value, *ln_MV*, as a control variable. Besides that, there is of course the intercept and the error term.

To go further into detail on the selected independent variables, the S&P 500 dummy is 0 if the observation i.e. the PERMNO, is not in the S&P 500 index. Hence, the S&P 500 dummy is 1 for S&P 500 firms. The time indicator is 1 for year 1989, 2 for year 1990, and so on until 2018 which is 30. The interaction variable is then the S&P 500 dummy and the time indicator multiplied. We utilise the natural logarithm of Market Value as an attempt to minimise potential skewness from heavily impacting the results.

When running a regression, it is important to attempt to reduce the risk of violating the corresponding regression assumptions. For our two regression models these are the four least squares assumptions for multiple regressions discussed in detail earlier in the methodology. If these four assumptions hold, then the OLS estimators $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ are jointly normally distributed for large samples (Stock & Watson, 2015). In the results section for Test 3, we will go through if our dataset can be assumed to adhere to the least squares assumptions for multiple regressions. This is done through examining the descriptive statistics, the correlation, and the distribution of the variables.

5.5 Group Definition

The objective of our thesis is to attempt to tease out the effect of increased indexing on stock price efficiency. An event study of PEAD provides a good proxy for market efficiency. Furthermore, a regression analysis is applied to consider the effects over time. However, to get reliable, valid results we have to segregate each news group into multiple groups based on their indexation. Hence, this section is related to *sample selection* in the event study terminology. This will allow us to test the effect of increased indexing on stock pricing efficiency over time. This section is only concerned with the definition of the groups. The subsequent section will discuss how we conduct the data management in practice.

5.5.1 Group 1 – S&P 500

Group 1 should represent the most passively owned stocks in the investment universe. As a group, the S&P 500 index lives up to this requirement. The first index mutual fund formed by John Bogle in 1976 (Zweig, 2016) was based on the S&P 500. Since then, it has continued to be the most heavily indexed group of stocks (The Wall Street Journal, 2016). Historical data enables us to identify, which firms were included in the index in the 1989-2018 period (Appendix VIII provides the code for identifying S&P 500 constituents utilising R). Thus, we are able to compose and test this group. We define whether a firm is included in this group as: if firm *i* is included in the S&P 500 at time *t*, then include in Group 1. Furthermore, though the inclusion of firms in the index is to some extent subjective, there are clear requirements to included firms (S&P Dow Jones Indices, 2019). Knowledge about these criteria will enable us to create a group of firms, which are as similar as possible except for the fact that they are not included in the index.

5.5.2 Group 2 – Similar to S&P 500

It is paramount for the validity of this study that Group 2 reflects Group 1 on as many parameters as possible. Optimally, they would be similar along every dimension except for the fact that Group 1 firms are included in the S&P 500, while Group 2 firms are not included. However, devising such a group is a difficult and demanding task. In addition, we need to be wary of datamining. It is possible to continue to adjust the groups to fit our needs, however, it may result in a non-random sample. Furthermore, the group must be sufficiently large as to enable us to draw inferences. Hence, there is a trade-off between on the one hand similarity and validity, and on the other hand the ability to efficiently assign firms to the group and the size of the group. When making decisions about which variables to adhere to and which to disregard, we always keep the informational effect in mind. That is, how will the exclusion of this variable influence the informational efficiency for a given firm following an earnings announcement.

The requirements for inclusion into the S&P 500 are published routinely by Standard and Poor's. Hence, we can observe some quantitative variables on, which we can sort the data. Appendix I provides an overview of these variables for 2019 (S&P Dow Jones Indices, 2019). Note that these criteria are continuously updated. In particular, the market capitalisation threshold is updated regularly to mirror the highs and lows of the general stock market. For instance, the threshold was set at \$5bn on 18th July 2007 prior to the Great Recession and reduced to just \$3bn on 18th December 2008. On 20th February 2019, the threshold was set at its highest level ever, \$8.2bn (ibid.).

We have determined that the two criteria most influential to informational efficiency is market capitalisation and liquidity. A higher market capitalisation has profound informational effects. For instance, larger firms are subject to more media and analyst scrutiny. Furthermore, it can be assumed that they have more owners, who participate in price discovery and actively buy and sell the share to reflect new information. We also use liquidity as a criterion since it is a prerequisite for informational efficiency (O'Hara, 2003). When there is more liquidity, it is easier for buyers and sellers to meet. Consequently, more trades will occur at the price, which traders value the stock at. Thus, prices will reflect more information.

We use market capitalisation as one of the sorting variables for defining Group 2. However, we have not been able to collect data for the threshold value for the full 30-year period, 1989-2018. Instead, we use the lowest market capitalisation value for a stock in Group 1 in a given year as the threshold. We are aware that the value of a stock can potentially vary significantly between rebalancing periods. That is, a stock could potentially meet the threshold value on rebalancing day and then fall 50% in value the next day. This would set the threshold value too low. However, since we cannot observe the true threshold value set by S&P, this is the second-best option.

Standard and Poor's provide two liquidity measures in their inclusion criteria (S&P Dow Jones Indices, 2019). For 2019, stocks must meet the two following liquidity criteria: (1) The ratio of dollar value traded to float-adjusted market capitalisation must be 1.00. (2) The share must be traded a minimum of 250,000 times per month on average in each of the previous six months leading up to the evaluation. However, as with market capitalisation, we cannot observe these criteria historically. Hence, we need to apply a different criterion. We create a measure for dollar volume, i.e. price multiplied by volume. We then aggregate this for the 64 days leading up to the event, the event date, and the following day. We refer to this as Dollar Liquidity:

$$Dollar\ Liquidity = \sum_{j=-64}^{1} (price_{t+j} * volume_{t+j}) \tag{16}$$

We use 66 days since this is approximately equivalent to a quarter of a year in terms of trading days. We then use the same approach as for market capitalisation. We find the lowest dollar liquidity value for Group 1 in a given year and use this as an additional threshold value for Group 2. Hence, we impose an additional requirement for inclusion in Group 2. To be included in Group 2, a security needs to meet both threshold levels. Securities with high market values, but low liquidity cannot be considered for Group 2. Appendix II provides an overview of the threshold values for Market Value and Dollar Liquidity.

In summary, Group 2 are all stocks, which are not in the S&P 500, but still meet the threshold value for both Dollar Liquidity and Market Value. However, when we impose these two criteria, we find that Group 2 is still very large and returns significantly different summary statistics than Group 1. Due to the large sample size, we initially wrote the code for a subset of the sample, 2018. To guide our definition of Group 2, we generated the following summary statistics (Table 2 and Table 3):

Table 2 Summary Statistics for 2018 Data – Market Value

	Group 1	**Group 2**	**Group 2.1**	**Group 2.2**
n	1,127	3,443	1,000	2,443
Min ($m)	2,281	2,282	8,432	2,282
Mean ($m)	37,035	16,107	45,012	4,275
Median ($m)	17,948	5,024	19,184	3,861
σ ($m)	73,369	45,124	76,360	1,587

Table 3 Summary Statistics for 2018 Data – Dollar Liquidity

	Group 1	**Group 2**	**Group 2.1**	**Group 2.2**
n	1,127	3,443	1,000	2,443
Min ($m)	497,377	501,144	520,889	501,144
Mean ($m)	19,802,750	7,322,277	17,762,560	3,048,726
Median ($m)	10,750,430	3,021,774	9,058,052	2,391,299
σ ($m)	42,829,800	19,076,430	32,842,560	2,960,838

Because we have the luxury of plentiful data, we continue to limit the size of Group 2. We set a stricter restriction for inclusion in Group 2 in terms of Market Value. In particular, we rank all of the firms in Group 2 on Market Value and split the group into two. Group 2.1 then becomes the 1,000 firms with the highest market capitalisation for each year, while Group 2.2 is the remaining part of Group 2. The objective of Group 2 is to create a group, which is similar to Group 1. So, we limit Group 2.1 to a similar number of observations as Group 1. With Group 1 having 1,127 observations and Group 2.1 having 1,000 observations in 2018. As is evident from Table 2 and Table 3, the groups are more similar after this additional data split. Since Group 2.2 does not add any additional value to our study, we will not revert to it later in this paper.

Table 4 provides the summary statistics for all groups in the full sample. Whereas Group 1 has 27,076 observations, Group 2.1 has 1,000 per year for 30 years, in total 30,000. We find that the groups have become increasingly similar after the creation of Group 2.1. We will revert to the description of the data in much detail later in the Section 7.

Table 4 Summary Statistics for Market Value and Dollar Liquidity | Full Sample

	Market Value					Dollar Liquidity				
	All	**Group 1**	**Group 2**	**Group 2.1**	**Group 3**	**All**	**Group 1**	**Group 2**	**Group 2.1**	**Group 3**
Mean ($m)	2,691	18,225	7,098	29,396	425	1,272,428	10,559,770	3,070,077	10,634,620	178,425
Median ($m)	221	9,586	1,900	13,435	186	56,626	6,141,937	919,245	4,665,351	41,236
σ ($m)	14,326	34,506	23,753	48,778	1,045	7,063,970	18,924,560	11,051,770	23,629,280	665,349
n	658,118	27,076	159,750	30,000	324,716	658,118	27,076	159,750	30,000	324,716
Proportion	100 %	4 %	24 %	5 %	49 %	100 %	4 %	24 %	5 %	49 %

5.5.3 Group 3 – Other Stocks

Finally, we will have a group of other stocks, which we can utilise as a control group. Hoe, Xue, and Zhang (2018) suggest that one removes all stocks, which have a market capitalisation of less than the 20th percentile of the stocks listed on the NYSE. 'Microcaps' represent 3.2% of the aggregate market capitalisation, but 60.7% of total stocks. These firms have the highest equal-weighted returns and largest dispersion of returns. Hence, they are outliers which may distort our results (ibid.). Our dataset does not allow us to easily remove the 20th percentile of NYSE stocks. Instead, we use 10% of the minimum market capitalisation in the S&P 500 as the threshold value. For 2018, 10% of lowest Market Value S&P 500 security is $220m whereas the 20th percentile on the NYSE is $260m. This discrepancy is not decisive for our results. The goal of both approaches is to remove the smallest most volatile companies and we find that our solution is effective at that. Hence, we rely on our method.

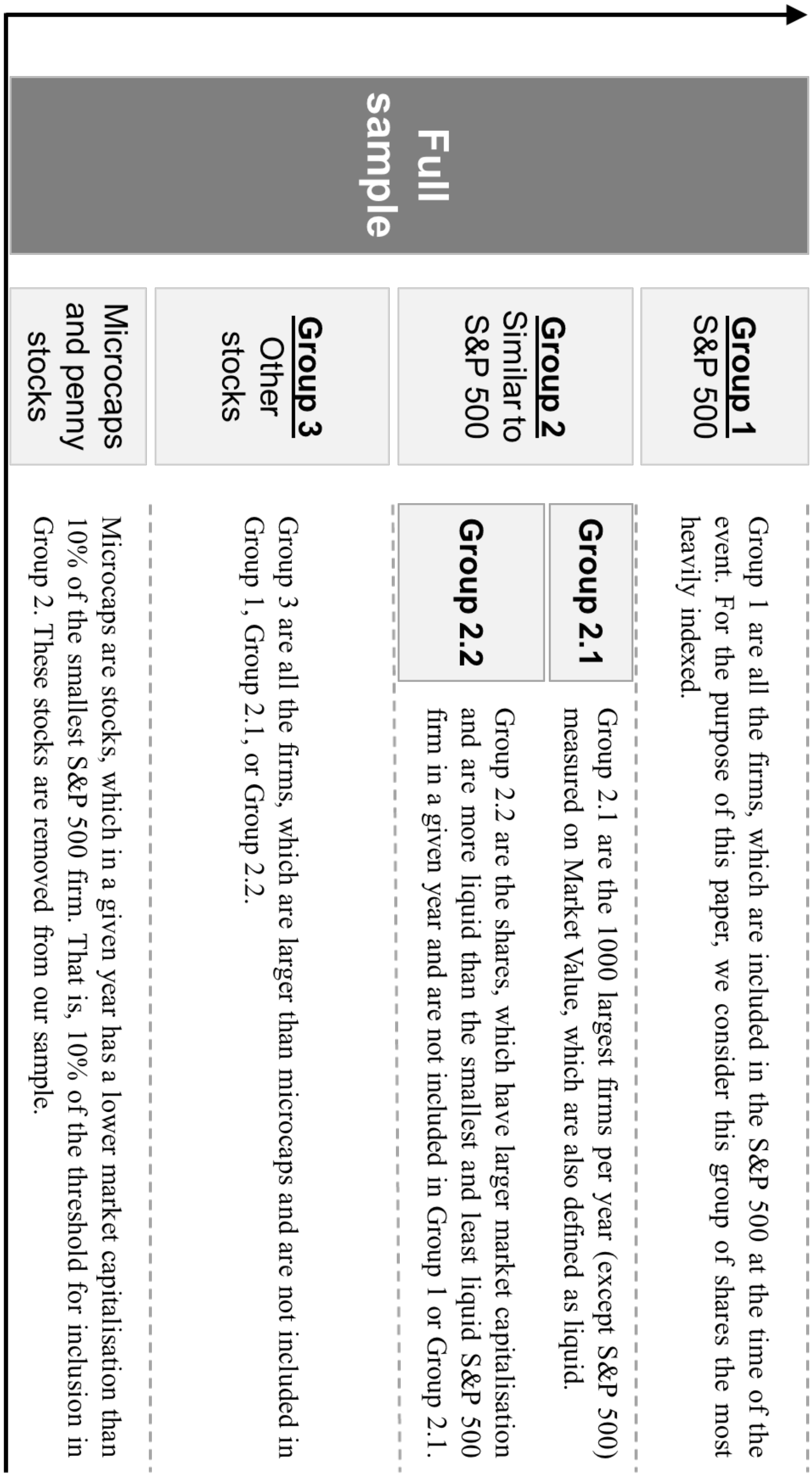
These microcap firms will be removed from our sample. That is, Group 3 will consist of all shares, which are not included in Group 1 or Group 2, but have a Market Value larger than 10% of the S&P 500 threshold for a given year. The summary statistics in Table 5 show that the 1st quartile of Market Value has been increased significantly. As has the mean and median. This suggests that we have been successful in removing microcaps.

Table 5 Summary Statistics for Market Value and Dollar Liquidity | Group 3

	Market Value		**Dollar Liquidity**	
	Group 3 Pre	**Group 3 Final**	**Group 3 Pre**	**Group 3 Final**
Mean ($m)	304	425	129,530	178,425
1st Quartile ($m)	32	78	3,478	10,115
Median ($m)	98	186	18,118	41,236
σ ($m)	886	1,045	589,772	665,349
n	471,292	324,712	471,292	324,712
Proportion	100%	69%	100%	69%

Figure 14 on the following page provides an abbreviated description of the groups.

Figure 14 Illustrative Depiction of Group Definitions



Source: Own creation

5.6 Data Management

We employ multiple programming languages such as Python, R, and STATA to handle the vast amount of data and produce our results. Furthermore, we utilise the event study application of WRDS. In total, we collect 52,153,354 observations, 437,101,186 data points, conduct 658,118 unique event studies, and use 3 programming languages in order to arrive at our results. In this section of the paper, we will describe our sample selection and where we collect data. The data preparation will be discussed subsequently.

5.6.1 Sample Selection

Our sample contains data from the Center for Research in Security Prices (CRSP) database in the period 1st January 1989 to 31st of December 2018. This database contains daily historical data for all stocks listed on the NYSE, NYSE MKT, NASDAQ, AMEX, and Arca exchanges. These exchanges include between 4,000 and 8,000 individual stocks at any point in time (Doidge, Karolyi, & Stulz, 2015). This provides a large amount of data for our study.

In our thesis, we limit the selection to stock exchange codes number 1, 2, and 3. These codes corresponds to the New York Stock Exchange (NYSE), the American Stock Exchange (AMEX), and NASDAQ, respectively. Thus, we limit our sample to the largest, most liquid, and most efficient exchanges. Furthermore, we remove all microcaps (Hou, Xue, & Zhang, 2018). These firms exhibit the highest dispersion of returns and constitute outliers, which may influence our results. Hence, as discussed when defining the groups, we remove these firms.

This is a longitudinal study of the effects of increased index investing. Hence, we have decided to use a long sampling period. Our study is conducted on the 30-year period from 1st January 1989 to 31st of December 2018. This long period enables us to test our hypotheses. Additionally, the vast amount of data provides more confidence to potential findings. In order to assess the effect over time, we will split the dataset into multiple shorter time periods as well as news groups. However, this will reduce the amount of data points in each group. This may particularly pose a problem for Group 2.1 Bad News and Good News, which are the smallest groups (931 and 1,124 observations, respectively). We are confident that this is still enough data points to make inferences. The composition of the groups in the full sample and different time periods will be discussed extensively in Section 7.

When selecting a sample, it is important to mitigate sampling biases. If one fails to do so, it may result in non-random samples. The most common sampling bias in empirical work is selection bias.

In our sample, we select to focus on the NYSE, NASDAQ, and AMEX exchanges and thus disregard the other U.S. exchanges. We do not consider this a problem, since these exchanges are the most liquid and important for the U.S. equity market. Hence, we still consider our sample representative of the U.S. equity market. Another common sampling error is self-selection bias. This bias is most common in laboratory experiments, where subjects self-select to be part of the experiment. In our study, this bias is less pronounced, since firms do not volunteer actively to be part of the sample. However, it could be argued that firms listed on any of the included exchanges may differ from those on the excluded exchanges. Nonetheless, since we include the largest and most respected exchanges, we do not consider this a problem. Furthermore, we consider our sample to be random since, we do not limit our sample to a single exchange or any specific years or dates. Instead, we download all variables across multiple U.S. exchanges for an extensive time period.

5.6.2 Data Collection

Table 6 Data Collection and Variable Specification

Name	***Abbreviation***	***Description***	***Source***
Stock return	RET	Holding period returns with dividends	CRSP
Market return	VWRETD	Value-weighted returns with dividends on market	CRSP
Price	PRC	Closing price on a given day	CRSP
Number of shares outstanding	SHROUT	The number of publicly held shares	CRSP
Volume	VOL	The total number of shares of a stock sold on a day	CRSP
Stock exchange code	EXCHG	Indicates at what exchange the common stock is traded	CRSP
Security identifier	PERMNO	Is a unique security identifier	CRSP
Security identifier	LPERMNO	Is a unique security identifier	CRSP/Compustat
Event date	RDQ	Indicates the date of reported quarterly earnings data	CRSP/Compustat
Location	LOC	Identifies where the company headquarters is located	CRSP/Compustat
Industry code	SIC	Is the standard company identification code	CRSP/Compustat
Stockholder's equity	SEQQ	Stockholder's equity (quarter end)	CRSP/Compustat
Deferred taxes	TXDITQ	Deferred taxes and investment tax credit (quarter end)	CRSP/Compustat
Preferred equity	PSTKQ	Total preferred/preference stock (quarter end)	CRSP/Compustat
Profits	IBQ	Income for the quarter before special items	CRSP/Compustat
Sales	SALEQ	Sales per quarter	CRSP/Compustat

This empirical study of stock prices in the U.S. equity market over a 30-year time-horizon requires millions of observations from a variety of sources. We collect data from Center for Research in Security Prices (CRSP) and CRSP/Compustat merged. The data we collect is required to (1) identify the earnings announcement date, (2) compute a measure for abnormal return enabling us to categorise earnings surprises, and (3) define the indexation groups. In Table 6, we provide an overview of the data variables collected. To ease the effort of replicating our study, we provide the variable abbreviation, a short description, and the source. Subsequently, we will describe why these variables were chosen, and how they contribute to this thesis.

The data has been retrieved from the Wharton Research Data Services (WRDS) platform, which Copenhagen Business School has access to. This database grants access to CRSP, CRSP/Compustat merged as well as a variety of other quantitative data sources. We choose to collect data from CRSP since it is the prime source of daily data on securities. CRSP/Compustat merged is chosen instead of Compustat because it enables us to retrieve the unique linking variable and security identifier LPERMNO and thus merge the data we retrieve from the CRSP and Compustat databases. Had we retrieved the data from the standard Compustat database, we would have been required to merge on the ticker variable. This is problematic because multiple securities can be connected to one ticker across time.

We collect data on stock return, RET, and market return, VWRETD, from CRSP in order to calculate a measure for abnormal return ($AR = R_{firm} - R_{market} = RET - VWRETD$) used in the categorisation of earnings surprises. CRSP includes multiple measures of market return. However, according to Campbell, et al. (1997) VWRETD is a good proxy for market return. Based on their assessment, we decide to use the same variables. For stock return, we choose RET. We find that the variable is comparable to VWRETD because it is including dividends. Furthermore, it provides return for a given stock on a given day. Hence, it is a valid proxy for stock return.

When cleaning the data for microcaps, we require a measure of market capitalisation. This measure is also required, when we create our indexation groups. In order to get that, we download data on price, PRC, and shares outstanding, SHROUT, from CRSP. Furthermore, we collect data on the number of shares of a stock traded daily, VOL, from CRSP. This measure is also required, when we create our indexation groups. Our sample is limited to stocks registered at AMEX, NASDAQ, and NYSE. To make this limitation, we limit the data collection to exchange codes, EXCHG, 1, 2, and 3. This is done when downloading the data from CRSP.

The most important variable from CRSP/Compustat merged is the earnings announcement date, RDQ. This date indicates, when the earnings announcement was made and is thus pivotal for the event study. We then collect a unique security identifier, LPERMNO, which enables us to link the data from CRSP and the merged database. To evaluate how similar Group 1 and Group 2 are, we collect additional descriptive data.

We collect data on the industry code of the company, SIC, and the nationality of the security proxied by the location of the companies' headquarters, LOC. We choose to collect these variables in case they are helpful in describing the groups. Furthermore, we collect SALEQ as a variable for the quarterly sales of the firm. To capture book value, we rely on Hou, Xue, and Zhang (2018) and compute it as stockholder's equity, SEQQ, plus balance sheet deferred taxes and investment tax credit, TXDITCQ, minus book value of preferred stock, PSTKQ. We elaborate on this in the data preparation section. Finally, to capture profits, we download the variable for quarterly income before extraordinary items, IBQ (ibid.).

6 Data Preparation

The main task of this thesis was to prepare the data applied in the event study. Writing the code took over three months and it has been tweaked back and forward multiple times to test what the best approach was to generate the variables needed for the analysis and how to split up the dataset practically. The main code was written in Python utilising the Jupyter Notebook interface. Working through notebooks was the preferred choice due to the vast size of the dataset. That is due to the ease of being able to visualise every script and making sure that each specific transformation of the data was successful one step at the time.

To begin with, the code was written on a subset of the data where year 2018 was downloaded. This was done to be able to give a better overview of how the data behaved and that no data was lost when, for example, merging the datasets. With a smaller dataset one can run and rerun one step of the code quickly and modify it to attain the preferred outcome. Once the code for the subset was done it could be just slightly tweaked to fit the full dataset. It took over 48 hours to run the full dataset, which has over 52 million rows, through all the steps. With limited RAM-memory and space on our laptops it took multiple attempts. We struggled with kernels dying and we had to save down the data frame in the middle of some parts of the code to clear the memory. Subsequently, we had to read in the modified file and continue.

The main code, which prepares the downloaded data from the WRDS platform to the final text files which can be plugged into the event study application, consists of three parts. Part I consists of 6 steps where the data is loaded in and cleaned. Part II consists of 16 steps where the CRSP and Compustat data is merged and some of the variables that are needed for the analysis are created and modified. Part III contains 25 steps. Here additional variables are created before the data is split up into different indexation groups, news groups, and time intervals. We also generate summary statistics and text files which are utilised in the event study. To ease the replication and the transparency of this study all code is displayed for each step with comments in Appendix V to Appendix VII. In this section, the three parts of the code will be described in words together with formulas to explain the data preparation and the corresponding decisions in greater detail.

6.1 Part I

In Part I Step 1, the two libraries pandas and numpy are imported. These libraries hold useful functions within data manipulation and data analysis. In Step 2, the data from CRSP is loaded in using the pd.read_csv function which reads a csv-file. The csv-file that is read in was generated from the CRSP database on the WRDS platform. In Step 2, the data is later displayed to see that the column variables are read in correctly and then the shape of the data frame is printed. For the CRSP data frame there is a total of 52,153,354 rows and 8 columns. Finally, the null values are summed up for each variable of the data frame. The sums can be seen in Table 7. Since the number of nulls in the data frame are a very small fraction of the full dataset all rows that have a null are dropped in Step 3. The final shape of the cleaned CRSP data frame is then 52,108,189 rows and 8 columns. Thus, with just over 40,000 rows dropped it is clear that the null in one variable corresponds to the null in another.

Table 7 Null Values for the CRSP Data Frame

PERMNO	0
date	0
EXCHCD	0
PRC	43,916
VOL	44,672
RET	43,916
SHROUT	16,428
vwretd	0

In Step 4, the Compustat data is loaded in from the csv-file generated by the CRSP/Compustat merged database on the WRDS platform. The merged database was selected because extensive time and effort has been put in to making this database compatible with the CRSP data. This has been done through creating a linking variable, LPERMNO, that matches the PERMNO variable from the CRSP database. This is a unique identifier. Hence, it is a better merging variable compared to the ticker, because the ticker variable can be inherited by other firms. After the data is loaded in, it is displayed to observe that the variables are read in correctly. The size of the Compustat data frame is 863,998 rows and 23 columns. The variables not mentioned in Table 6 on page 55 are default variables when downloading data from the CRSP/Compustat merged database. These are variables that help describe the dataset.

Table 8 exhibits the sum of the nulls for each variable generated in Step 4. It is clear that this data frame contains a lot of nulls in proportion to its size. The variable RDQ, which is the date of the reported quarterly earnings, has 191,795 nulls. This is equal to 22% of the observations. This pattern is consistent throughout the dataset and the majority occurs for SIC codes 6722 and 6726. These are management investment offices and unit investment trusts. This leads us to believe that they adhere to different rules for reporting. The variable RDQ is fundamental to the event study and cannot be substituted. Therefore, all rows with RDQ equal to null is dropped in Step 5. The results of dropping these rows can be seen in Table 8. It is clear that a lot of nulls in other variables are corresponding to the nulls in RDQ as they are removed as well through the drop. However, there is one variable that still has a significant number of nulls and that is TXDITQ. This will have an effect on Book Value, which is derived from this. In Step 6, which is the last step of Part I, two csv-files are saved down using the pd.to_csv function. One for the cleaned CRSP data frame and one for the cleaned Compustat data frame.

Table 8 Null Values for the Compustat Data Frame

	Before Drop	After Drop
GVKEY	0	0
LPERMNO	0	0
DATEDATE	0	0
FYEAR	0	0
FQTR	0	0
INDFMT	0	0
CONSOL	0	0
POPSRC	0	0
DATAFMT	0	0
TIC	25	23
CONM	0	0
CURCDQ	0	0
DATAQTR	920	394
DATAFQTR	0	0
RDQ	191,795	0
IBQ	145,353	1,713
PSTKQ	153,014	7,316
SALEQ	146,497	2,946
SEQQ	149,577	4,135
TXDITQ	263,873	111,981
COSTAT	0	0
LOC	0	0
SIC	0	0

6.2 Part II

In Part II, the CRSP and Compustat data frames are merged. After that variables that are needed for the split into groups and analysis are created. Step 1 is to import the pandas and numpy libraries. In Step 2, the cleaned CRSP and the cleaned Compustat csv-files are read in separately. We then check the size and make sure that the variables are loaded in correctly. In Step 3, pd.isnull().sum() prints out the nulls in each file and we can check that there are no unexpected nulls in the data frames. After this step, one can conclude that the correct files were loaded in and thus can move on to merging the data frames.

The data frames need to be merged on the unique identifier which is PERMNO for the CRSP data frame and LPERMNO for the CRSP/Compustat data frame. Furthermore, they need to be merged on the specific RDQ date. To ease this merge a new variable called PERMNODATE is created in Step 4. PERMNODATE is the PERMNO + DATE or LPERMNO + RDQ for each row. One example of how the final variable PERMNODATE looks for a security is '1000119890103'. Where 10001 is the PERMNO for that security and 1989-01-03 is the reported quarterly earnings date. In Step 5, the CRSP data frame and the Compustat data frame is left joined on the variable PERMNODATE. A left join is a way to merge two sets of data, where all records from the left data frame is returned together with the matching records from the right data frame. In our case, CRSP is the left data frame and thus all those observations are kept. Rows which have a matching PERMNODATE will get additional variables added to the rows. The resulting merged data frame has a shape of 52,109,068 rows and 34 columns. This data frame is then saved down to a csv-file which enables us to clear the memory. The csv-file alone is 5.5 GB, which together with the already loaded in data creates memory issues.

Once the kernel is restarted, the csv-file with the merged data can be read in again in Step 6. This time only the most necessary variables are loaded in and the data frame now holds 23 columns compared to the previous 34. In Step 7, the variables VOL, PRC, SHROUT, RET, and VWRETD are converted into numeric variables. Python has previously regarded them as 'objects' as it is easier to read in a large dataset when it does not expect each observation to be of a certain format. We now convert each observation in these columns to the type 'numeric'. In Step 8, a new variable for absolute price is created, i.e. the absolute value of the variable PRC. This is done in order to not get any negative values of Market Value.

In Step 9, abnormal return called aRET is calculated as seen in Equation (17).

$$aRET = RET - VWRETD \tag{17}$$

In Step 10, a new variable is created, Log_aRET, which is the logarithm of the abnormal return, aRET. This is done using the numpy function np.log1p which corresponds to log(1 + x). This is a way to work around the fact that the abnormal return can be negative. In Step 11, the variable Dollar Volume is created. As exhibited in Equation (18), Dollar Volume is defined as the absolute price times volume, which is the total number of shares sold on a day.

$$Dollar\ Volume = Abs\_Price * \ VOL \tag{18}$$

In Step 12, a list of the columns in the merged data frame is printed. This is needed to be able to perform Step 13, where the CAR3 variable is created. Equation (15) on page 45 displays the definition of the three-day cumulative abnormal return which we call CAR3. It can be seen in the equation that it is the sum of $\log(R_{i,t+j} - R_{m_{i,t+j}})$, which in the practical case here is the Log_aRET variable generated in Step 10 from one day prior to the event to one day after the event. In Step 13, a function was created, which utilises the lambda expression. This is a type of anonymous function that takes in argument x and returns another function, the lambda, that takes another argument y. The code for Step 13 can be seen in Figure 15. This is a snippet of the full code for Part II, which can be seen in full in Appendix VI. In words, the code takes for each x, which is a row where RDQ is not null, it sums the row -1 to row 2 for the Log_aRET variable. This implies one day before the event, the event day which is defined as row 1, and the day after the event. Here we use the iloc indexer which is a way to select index rows and columns based on their position in the data frame. It is important to understand how the iloc indexer works in that is takes [row:column-1].

Figure 15 Code for Creating the CAR3 variable

Step 13

```
# Create the CAR3 variable
# Remember that iloc takes y-1 for columns
def ret_func(row):
    return dfMerged.iloc[row.name-1:row.name+2,24].sum()

dfMerged['CAR3'] = dfMerged.apply(lambda x: ret_func(x) if pd.notnull(x.rdq) else np.nan,
                                  axis=1)
```

In Step 14, a csv-file is generated to be able to make sure that the CAR3 variable was generated correctly with the right column being summed. This csv-file will also work as a back-up of the latest steps in case the kernel dies due to low memory. In Step 15, the Dollar Liquidity variable is created as stated earlier in Equation (16) on page 50. In practice the code runs a function of the same type used to create CAR3 but here the Dollar Volume variable is summed over 66 days for the rows that have an RDQ date. Thus, 64 days prior to the event, the event day, and the day after the event are summed. In Step 16, the last step of Part II, a data frame exclusively containing rows with CAR3 values are created. This is essentially also the rows that have an RDQ date. The shape of the CAR3 only data frame consists of 658,188 rows and 29 columns.

6.3 Part III

In Part III, the last variables needed for the analysis and the splitting into groups are created. Once these variables are created, the groups are split up, before further splits into news groups and time intervals are conducted. Finally, text files which serve as the input for the event study application maintained by WRDS are generated. In Step 1, the pandas and numpy libraries are imported. In Step 2, the Only_CAR3 dataset is read in. In Step 3, we confirm that the shape of the data which is loaded in is correct, i.e. 658,118 rows and 29 columns. Here summary statistics are generated that will be thoroughly described in Section 7. In Step 4, two additional variables are generated that will assist in the sorting into groups and the analysis. These are Market Value and Book Value. Market Value plays a key role in the group split as it is one of the two threshold variables together with Dollar Liquidity. Market Value is defined as seen in Equation (19). Where the variable SHROUT is the number of shares outstanding. Book Value is constructed as stated in Equation (20), where SEQQ is the stockholder's equity, TXDITQ is deferred taxes and investment tax credit, and PSTKQ is the total preferred stock. All steps of Part III of the code can be seen in Appendix VII.

$$Market\ Value = Abs\_Price\ x\ SHROUT \tag{19}$$

$$Book\ Value = seqq + txditq - pstkq \tag{20}$$

In Step 5, the S&P 500 dataset is loaded in as a data frame. The csv-file that is loaded in was generated through a script run in R. The full code of how to obtain the constituents of the S&P 500 can be seen in Appendix VIII. The data frame consists of 755 rows and 3 columns with PERMNO, START, and ENDING. This represents the date when the security first entered the index and the date when it dropped out. This S&P 500 data frame is used in Step 6 to flag if a PERMNO in the full dataset was in the S&P 500 at that specific time point. The script uses the np.piecewise function which is part of the numpy library. Given a set of conditions and corresponding functions the piecewise function evaluates each function on the input data whenever the condition is true. Thus, if the PERMNO is in the S&P 500 data frame then check if the RDQ date is between the START and ENDING and flag if true. The script flags the ones that are not in the S&P 500 index with a 0. We instead want this to be flagged as a null as it allows us to use specific null-sorting functions. Converting the zeros into nulls is done in Step 7.

To be able to move on and start splitting the data into groups, we need to create a variable that is only the year of the RDQ date. For example, if the RDQ date is '19980508' we want the new year variable to state '1998'. The creation of the new year variable is done through utilising the datetime library and the very versatile lambda expression when working through the dataset. In Step 9 of the code, it is time to generate the first group – the S&P 500 group. This is done through creating a new Group 1 data frame. Here, all the observations from the full dataset that are not null for the S&P 500 flag variable are allocated. That is, the ones that are flagged are placed in Group 1.

The full code for Step 10 can be seen in Figure 16 on the following page. The first thing the script does is to generate the threshold levels for Market Value and Dollar Liquidity. The code groups the data frame by year and finds the smallest value per year within Group 1. This is done for the data in Group 1 which is the S&P 500 firms. These threshold levels are saved down into a data frame which can be seen in Appendix II. Based on these levels, the data is divided into Group 2 and Group 3. First all the data from the full dataset that is not flagged for the S&P 500, and thus not in Group 1, is stored in a data frame. Then the code circles through this data frame and divides into Group 2 if the security is above both the threshold level for Market Value and Dollar Liquidity for the specific YEAR of the row that was derived in Step 8 based on the RDQ date. It puts the row into Group 3 if the security fails to be above at least one of the threshold levels. In Step 11, we save down the threshold levels to a new data frame. In Step 12, we make sure that the split of the groups was correct by comparing the number of rows in each group and see if it adds up to the number of rows for the full dataset. Thus, we ensure that we did not lose any data.

Figure 16 Code to Find Thresholds and Generate Group 2 and Group 3

Step 10

```
# Find the thresholds for MarketValue and Dollar_Liquidity for each year
# The threshold level is equal to the smallest of the S&P500 firms for each year
dfThreshold =
dfGroup1[['year','MarketValue','Dollar_Liquidity']].groupby(by=['year']).min().reset_i
ndex()

dfDataSPFlagNull = dfData.loc[pd.isnull(dfData['SP500flag'])]

# Create Group 2 and Group 3 based on the thresholds
# Group 2 has both MarketValue & Dollar_Liquidity larger than the threshold
# Group 2 fails to be above the threshold on at least one of the two

dfGroup2 = pd.DataFrame(columns=dfDataSPFlagNull.columns)
dfGroup3 = pd.DataFrame(columns=dfDataSPFlagNull.columns)

for x in dfDataSPFlagNull.year.unique():
    ThresholdMV = dfThreshold.MarketValue.loc[dfThreshold.year == x].values[0]
    ThresholdLiq = dfThreshold.Dollar_Liquidity.loc[dfThreshold.year == x].values[0]
    # Group 2
    tempdfGroup2 = dfDataSPFlagNull.loc[(dfDataSPFlagNull.year == x) &
                             ((dfDataSPFlagNull.MarketValue >= ThresholdMV)&
                             (dfDataSPFlagNull.Dollar_Liquidity >= ThresholdLiq))]
    dfGroup2 = pd.concat([dfGroup2, tempdfGroup2]).reset_index(drop=True)

    # Group 3
    tempdfGroup3 = dfDataSPFlagNull.loc[(dfDataSPFlagNull.year == x) &
                              ((dfDataSPFlagNull.MarketValue < ThresholdMV)|
                              (dfDataSPFlagNull.Dollar_Liquidity < ThresholdLiq))]
    dfGroup3 = pd.concat([dfGroup3, tempdfGroup3]).reset_index(drop=True)
```

In Step 13, summary statistics are generated that will be thoroughly described under the Data Presentation Section 7. In Step 14, Group 2.1 is created which is the top 1000 rows with the largest Market Value per year from Group 2. This is done by grouping on year and using lambda to loop through the dataset and pick the 1000 largest securities per year. Summary statistics is also generated for Group 2.1 in this step.

In Step 15, we remove microcaps to attempt to avoid them from skewing the results. This is done through utilising the thresholds that we previously found for the smallest level for each year of the S&P 500 constituents. Using this data to find a cut-off point also makes sure that it varies with time. From the threshold level for each year we take 10%. For example, for year 2000 the Market Value threshold level is $268.232m. The cut-off point for year 2000 is 10% of that, which is $26.823m. We generate these cut-off levels for all years and use them to generate a final Group 3 data frame that has 324,716 rows. Thus, dropping 146,576 rows that are below the yearly cut-off levels.

In the following steps, the groups will be further split into news groups. First, in Step 16 the lower and upper boundaries for negative and positive news are found. The upper boundary is the 90^{th}-percentile of CAR3 and the lower boundary is the 10^{th}-percentile of CAR3. This results in the upper level being 0.08840 and the lower level being -0.09601. Second, in Step 17, 18, 19, and 20 the full dataset, Group 1, Group 2, Group 2.1, and Group 3 are split further up into news groups utilising these upper and lower boundaries. Good News is defined as positive earnings surprises and thus above the upper level. Bad News are the ones with CAR3 below the lower boundary and Neutral News is classified as the ones that are above or equal to the lower and below or equal to the upper boundary. After these steps there will be one positive, one neutral, and one negative news group for each of the four main groups. Summary statistics for the 12 different news groups are also generated during these steps and will be analysed in Section 7 where the data is presented.

In Step 21, a new data frame is created for each news group that only contains data on the two columns PERMNO and RDQ. These new data frames are then saved down into text files that can be plugged in to the event study application on the WRDS platform. The shape of the input data frames is also thoroughly controlled in Step 21 as it would be an easy mistake in this very repetitive part of the code to write for example dfESINPUT_Negative_Group2 instead of dfESINPUT_Neutral_Group2 or put a 3 where the 2 should be.

In the four last steps, the groups that are already split into news groups are further split into time-intervals. The split is done through the DATE_FINAL variable that is a conversion of the RDQ date to divide the news groups into three time-intervals. The DATE_FINAL variable is created in Step 22 as a mirror of the RDQ date but in a format that Python can use for the splits. Instead of the RDQ date being '20060708', the DATE_FINAL has the format '2006-07-08'.

In Step 23, the code for the 1989-1998 time-interval is run. In Step 24, the code for 1999-2008 is run. In Step 25, the code for 2009-2018 is run. In these steps, new data frames that contain the two columns PERMNO and RDQ are also generated and saved down into text files that are the basis for the event study.

These final data frames for example looks like dfPositive_Group1_89to98, dfPositive_Group1_99to08, and dfPositive_Group1_09to18 resulting in a total of 36 time-interval groups. Here again the shape of the data frames are thoroughly inspected so that no simple typos create incorrect data frames. This was the last step of Part III, which concludes the walk through of the main code for the data preparation. To ease the replication and transparency of this event study the main code with small comments are provided in Appendix V to Appendix VII. The code for the extraction of the constituents of the S&P 500 can be found in Appendix VIII.

7 Data Presentation

The purpose of this section of the paper is to outline the properties of the full dataset used in this thesis. The objective is to show that Group 1 and Group 2.1 are very similar. As alluded to when we defined the groups, the two most important variables indicating informational efficiency are Market Value and Dollar Liquidity. Hence, this section will focus on those two. However, because this argument is central to the thesis, we will provide comparisons along multiple other variables. First, we compare all of the groups for the full 1989-2018 dataset on Market Value and Dollar Liquidity. Then, we make the same comparison after having split the data into news groups. Finally, we provide a more elaborate comparison of Group 1 and Group 2.1 along multiple other variables.

7.1 Market Value and Dollar Liquidity

Table 9 provide summary statistics for the merged dataset for the full 1989-2018 time period before news definition. It is evident that the groups differ widely on the two parameters of informational efficiency, which we have chosen. For instance, the average Market Value for all groups are \$2,691m, while the same metric is \$425m for Group 3 and \$18,225m for Group 1. The same pattern emerges for Dollar Liquidity, where Group 1 and Group 2.1 are substantially more liquid than Group 2 and 3. In summary, we have successfully created groups, which differ in their Market Value and Dollar Liquidity.

Table 9 Summary Statistics for Merged Data before News Definition

	Market Value					**Dollar Liquidity**				
	All	**Group 1**	**Group 2**	**Group 2.1**	**Group 3**	**All**	**Group 1**	**Group 2**	**Group 2.1**	**Group 3**
Mean (\$m)	2,691	18,225	7,098	29,396	425	1,272,428	10,559,770	3,070,077	10,634,620	178,425
Median (\$m)	221	9,586	1,900	13,435	186	56,626	6,141,937	919,245	4,665,351	41,236
σ (\$m)	14,326	34,506	23,753	48,778	1,045	7,063,970	18,924,560	11,051,770	23,629,280	665,349
n	658,118	27,076	159,750	30,000	324,716	658,118	27,076	159,750	30,000	324,716
Proportion	100%	4%	24%	5%	49%	100%	4%	24%	5%	49%

Table 10, Table 11, and Table 12 provide the same statistics after we have separated the data into different news groups. The data provides several interesting insights. First, it is evident that the mean and median Market Value are lower in the Good and Bad News groups. Since we define the news groups by their CAR3 following an earnings surprise, this would seem to indicate that firms with lower Market Values react more strongly to earnings announcements. That is, the market included less information about the results of the earnings report prior to its announcement.

In contrast, earnings announcements of higher Market Value firms were more likely to have a lower CAR3 and thus be categorised as a Neutral News event. This would seem to indicate that more information has been unveiled about high Market Value firms prior to their earnings announcement.

Table 10 Summary Statistics for Bad News Dataset (1989-2018)

	Market Value					Dollar Liquidity				
	All	Group 1	Group 2	Group 2.1	Group 3	All	Group 1	Group 2	Group 2.1	Group 3
Mean ($m)	1,212	13,644	3,431	22,756	318	975,741	12,875,260	3,337,382	18,356,350	273,864
Median ($m)	139	6,771	1,215	12,258	159	46,111	7,161,952	1,146,197	8,750,757	74,326
σ ($m)	7,730	27,998	11,714	34,278	569	5,663,109	21,954,030	12,988,680	40,398,170	807,530
n	65,812	1,632	10,911	931	31,026	65,812	1,632	10,911	931	31,026
Proportion	100%	2%	17%	1%	47%	100%	2%	17%	1%	47%

Table 11 Summary Statistics for Neutral News Dataset (1989-2018)

	Market Value					Dollar Liquidity				
	All	Group 1	Group 2	Group 2.1	Group 3	All	Group 1	Group 2	Group 2.1	Group 3
Mean ($m)	3,078	18,801	7,738	29,938	448	1,342,107	10,279,410	3,070,702	10,197,000	162,426
Median ($m)	254	9,946	2,055	13,628	191	59,270	6,037,605	887,985	4,453,399	37,116
σ ($m)	15,567	35,135	25,189	49,552	1,128	7,258,343	18,533,050	11,045,290	22,638,400	657,053
n	526,494	23,780	135,994	27,945	261,570	526,494	23,780	135,994	27,945	261,570
Proportion	100%	5%	26%	5%	50%	100%	5%	26%	5%	50%

Table 12 Summary Statistics for Good News Dataset (1989-2018)

	Market Value					Dollar Liquidity				
	All	Group 1	Group 2	Group 2.1	Group 3	All	Group 1	Group 2	Group 2.1	Group 3
Mean ($m)	1,068	14,498	3,432	21,441	341	1,011,684	12,295,390	2,836,398	15,118,820	216,527
Median ($m)	108	7,617	1,332	11,383	170	47,386	6,600,570	1,038,136	7,636,491	56,498
σ ($m)	6,914	30,415	12,401	37,119	583	6,727,856	20,914,240	9,165,626	27,141,620	563,422
n	65,812	1,664	12,845	1,124	32,120	65,812	1,664	12,485	1,124	32,120
Proportion	100%	3%	20%	2%	49%	100%	3%	20%	2%	49%

Since the highest Market Value stocks are in the Neutral group, we would also expect these stocks to have higher Dollar Liquidity. However, that is not the case. The summary statistics reveal that the most liquid stocks are found in the Bad and Good News groups. This may be driven by the way we calculate Dollar Liquidity as expressed in Equation (16). Recall that Dollar Liquidity is the cumulative Dollar Volume from the earnings announcement date +1 and 64 days back. Hence, this includes the earnings announcement date and the following day. It could be that because the earnings announcement is so surprising, this drives a lot of trading, which increases the measure of Dollar Liquidity. However, we do not have any data to support this hypothesis. Lastly, it should be noted that the amount of data in each group differs significantly for each group. However, even when we split the data into three separate time periods, we still have sufficient data to draw inferences.

7.2 Group 1 and Group 2.1

In this section, we provide descriptive statistics for Group 1 and Group 2.1. These groups are central to our thesis. For our results to be reliable, we must show that these groups are similar. In this section, we will point out the similarities and the discrepancies between the two groups. In Table 13, we provide summary statistics for Market Value, Book Value, Dollar Volume, Dollar Liquidity, and CAR3.

Table 13 Summary Statistics for Merged Data before News Definition (1989-2018)

	Market Value		**Book Value**		**Dollar Volume**		**Dollar Liquidity**		**CAR3**	
	Group 1	**Group 2.1**	**Group 1**	**Group 2.1**	**Group 1**	**Group 2.1**	**Group 1**	**Group 2.1**	**Group 1**	**Group 2.1**
Mean ($m)	18,225	29,396	6,940	13,724	286,410	283,858	10,559,770	10,634,620	-0.08 %	0.16 %
1st Quart. ($m)	5,042	7,638	1,644	2,696	56,900	32,653	2,708,502	1,677,533	-3.04 %	-2.25 %
Median ($m)	9,586	13,435	3,490	6,189	143,284	106,097	6,141,937	4,665,331	0.12 %	0.20 %
3rd Quart. ($m)	18,214	28,819	7,256	13,585	312,078	288,811	11,451,680	11,103,950	3.33 %	2.75 %
σ ($m)	34,506	48,788	12,054	24,676	555,589	673,162	18,924,560	23,629,280	6.95 %	5.22 %
n	27,076	30,000	20,442	23,219	27,076	30,000	27,076	30,000	27,076	30,000
Proportion	4%	5%	3%	4%	4%	5%	4%	5%	4%	5%

Firstly, when considering the two liquidity measures, Dollar Volume and Dollar Liquidity, we find that the two groups look very similar. Here, Dollar Volume is the average volume on event days. In particular, the mean value of the two groups are within 2% of each other. In terms of the remaining distribution, Group 2.1 exhibits lower values for all quartiles. However, the values are still not meaningfully different. Furthermore, the standard deviation of Group 2.1 is larger for both measures, but not to an unfavourable extent. In summary, we consider the two groups to be similar to a satisfying extent on this parameter.

In terms of Market Value, the summary statistics reveal that Group 2.1 exhibits a higher mean value and higher values for all quartiles. Furthermore, it also has a higher standard deviation. However, Group 2.1 still looks more similar to Group 1 than Group 2 does. Whereas the mean value of Group 1 is approximately 38% lower than Group 2.1 it is 150% higher than Group 2. That is, the additional refinement provides more similar groups. That being said, we recognise that the groups remain fairly dissimilar on this measure. This could be mitigated by extending Group 2.1 to include more securities. However, as we discuss later in this section, this would reduce the similarities in terms of liquidity and make the groups more dissimilar in terms of number of observations. Furthermore, if we continue to change the sample to fit our needs, it may compromise the randomness of the sample and could potentially be categorised as datamining. Finally, it could be argued that there is some level of Market Value beyond which the marginal informational effect of higher Market Value is negligible. In conclusion, we find that though the two groups are not perfect on this measure, they remain satisfactory to the purpose.

The first thing we should note regarding Book Value is that we have not been able to compute this measure for all observations – approximately 25% remain blank. This data loss may not be evenly distributed across the sample. That is, one should be careful when extrapolating conclusions from the summary statistics. However, the summary statistics for Book Value again show a higher value for Group 2.1 than for Group 1. The discrepancy is approximately on the same level for all quartiles across the two metrics, which would seem to indicate that the two groups have the same distribution of book-to-market value. However, because Group 2.1 has a high standard deviation and is left-skewed compared to Group 1, the mean value is higher. We would have preferred to provide summary statistics for the book-to-market ratio of the two groups. However, because we are not able to compute Book Values for approximately 25% of the sample, we have refrained from doing so. In the data preparation section, we showed that TXDITQ continued to provide null values. Consequently, since the null values may not be evenly distributed across the sample, the summary statistics may be misleading.

CAR3 can to some extent be considered an alternative proxy for market efficiency around the earnings announcement date. In our sample, Group 1 and Group 2.1 are similar on this metric. However, Group 2.1 appears to exhibit less extreme reactions to earnings announcements. That is, whereas the median of the two groups is very similar, the 1st and 3rd quartile of CAR3 for Group 2.1 is less extreme. Furthermore, the standard deviation is lower for Group 2.1. Nonetheless, they remain fairly similar and as we saw in Table 10 and Table 12 we still get an adequate number of observations in each news group. It could appear as if the higher Market Value of Group 2.1 might be what is driving the lower CAR3. However, we tested this hypothesis. In Appendix III, you can see summary statistics for an altered Group 2.1, where we take the top 1,500 firms ranked on market value instead of the top 1,000. It is evident that the groups are more similar on Market Value, but they remain dissimilar to the same extent on CAR3. Furthermore, they become more dissimilar in terms of Dollar Liquidity and number of observations. Thus, we decide to continue with our current definition of Group 2.1.

8 Empirical Results

This section of the paper will provide a quantitative overview of the results derived from our tests. At this point we will merely present the results and not discuss the implications hereof. The discussion will be conducted in the subsequent section. This section proceeds as follows. First, we conduct Test 1, which consider whether there is an underlying trend in any of the news groups. Second, we discuss the results of Test 2, the general event study of post-earnings announcement drift (PEAD). This section will also include results of shorter time-periods. Third, we present the results of Test 3, the regression analysis. The code for preparing the graphs used in this section are provided in Appendix IX.

8.1 Test 1 – PEAD: Underlying Trend

The first test reports results of an event study of post-earnings announcement drift prior to the creation of indexation groups. We conduct this event study to see whether there is an underlying trend in the data for the different news groups. That is, whether either of them exhibits a positive or negative drift absent the additional division into indexation groups. This test will inform the interpretation of the subsequent results. Please note that the tables provide results for the drift period [2;30], while Figure 17 illustrate the same period.

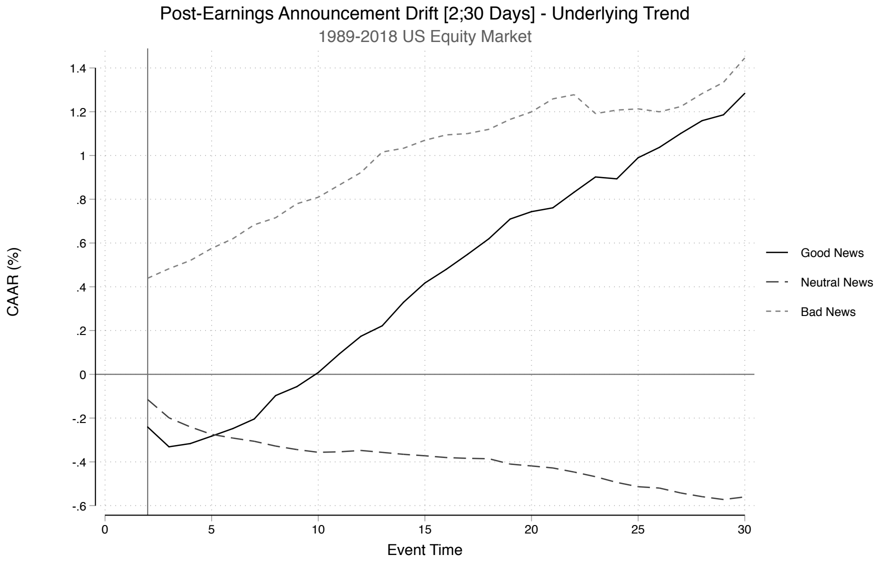
Table 14 – CAAR and t-statistic for 1989-2018 – Period [2;30]

	All Groups
Bad News	1.45 %
	6.88 ***
No News	-0.56 %
	-4.31 ***
Good News	1.29 %
	7.28 ***

* Denotes significance at the 10%-level,
** Denotes significance at the 5%-level,
*** Denotes significance at the 1%-level.

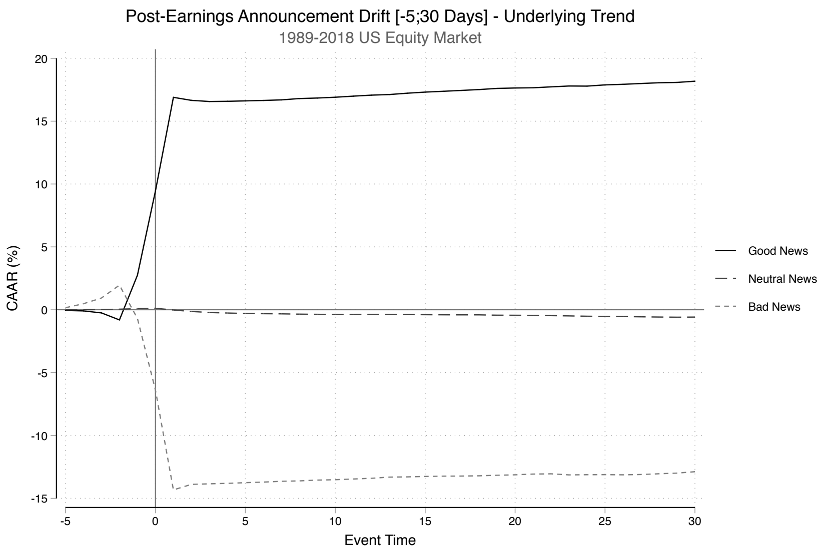
Table 14 reports the empirical results of Test 1. The event study yields highly significant results for all news groups. The drifts of all news groups are significant at the 1%-level. The drifts amount to 1.45%, -0.56%, and 1.29% for Bad, Neutral, and Good News respectively. It is interesting, and surprising, that the drift following the announcement of Bad News is positive – especially at that magnitude. In Section 10, Limitations, we discuss a possible mistake, we made when conducting this test, which we suspect is driving the peculiar results. Overall the results seem to suggest that there is an underlying trend in the different news groups. Figure 17 illustrates the results graphically, while Figure 18 illustrates the [-5;30] period. The solid line depicts the Good News group, the dashed line depicts the Neutral News group, and the dotted line depicts the Bad News group. According to the efficient market hypothesis, securities should react immediately to new information and revert to a random walk (1970). However, as the table and figures suggest, the securities often exhibit a significant drift following the earnings announcement.

Figure 17 PEAD: 1989-2018 | CAAR [2;30]



Source: Own creation

Figure 18 PEAD: 1989-2018 | CAAR [-5;30]



Source: Own creation

8.2 Test 2 – PEAD: The Effect of Indexing

In this section, we will present the results of the event study of PEAD. First, we will present the results for the full time-period, 1989-2018. This will give us an overview of the general effects in the different news and indexation groups. Second, we will go through the results for the three sub-periods, 1989-1998; 1999-2008; 2009-2018. This is intended to give us a general impression of how the drifts of the groups and its significance have developed as indexing has become more prevalent. Please note that the tables throughout this section provide results for the drift period [2;30], while the figures illustrate the same period as well as the [-5;30] day period.

8.2.1 PEAD: 1989-2018

In Table 15, we provide the empirical results of the event study of PEAD for the full time period, 1989-2018. We aggregate drift for the [2;30] day period and test its significance. Figure 19 illustrates the [2;30] day drift period. In particular, we are interested in whether there are any differences between Group 1 and Group 2.1.

Table 15 – CAAR and t-statistic for 1989-2018 | Period [2;30]

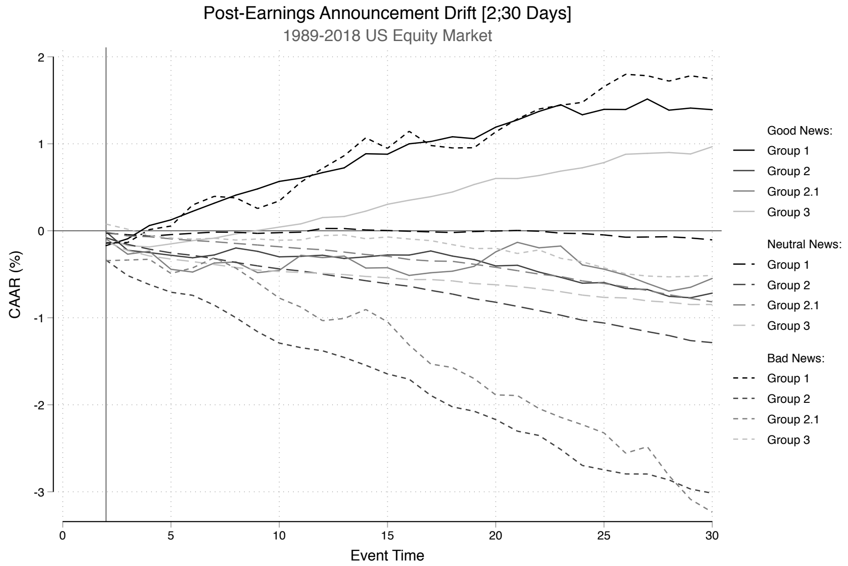
	Group 1	Group 2	Group 2.1	Group 3
Bad News	1.75 %	-3.02 %	-3.23 %	-0.51 %
	3.69 ***	-14.66 ***	-5.49 ***	-2.99 ***
Neutral News	-0.10 %	-1.29 %	-0.82 %	-0.85 %
	-1.69 *	-21.28 ***	-15.85 ***	-8.05 ***
Good News	1.39 %	-0.72 %	-0.55 %	0.97 %
	4.02 ***	-4.37 ***	-1.20	6.90 ***

* Denotes significance at the 10%-level, ** Denotes significance at the 5%-level, *** Denotes significance at the 1%-level.

All drifts exhibited by Group 1 are significant at the 10%-level, while Bad News and Good News are significant at the 1%-level. The drifts for the [2;30] day period are 1.75%, -0.10%, and 1.39% for Bad, Neutral, and Good News, respectively. It is surprising to see that the drift of Group 1 | Bad News is positive with a high magnitude. Group 2 exhibits a similar trend with all drifts being significant at the 1%-level. The drifts amount to -3.02%, -1.29%, and -0.72% for Bad, Neutral, and Good News, respectively. In general, the drifts are more negative than the same drifts for Group 1. It is surprising to find that the drift of Group 2 | Good News is negative. In particular, the drift for Group 2 | Bad News is large. In Group 2.1, we find the first non-significant result. While Bad and Neutral News remain significant at the 1%-level, the drift for Good News is not significant. For Bad and Neutral

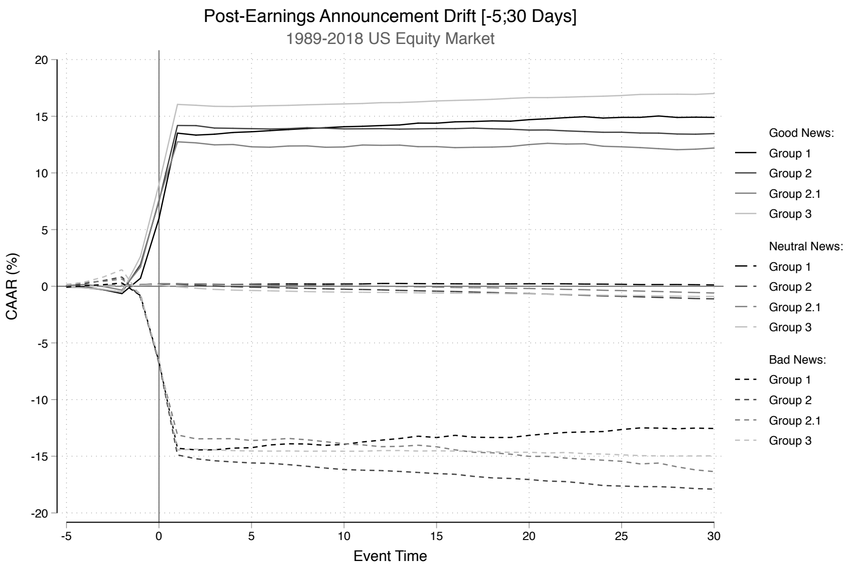
News, the drifts amount to -3.23% and -0.82%, respectively. Since Group 2.1 is a subset of Group 2, it is not surprising that the results are similar. However, it is evident that the additional refinement of the group has resulted in a higher variance of CAAR and thus a lower t-statistic. In Group 3, all drifts are significant at the 1%-level. CAAR for the [2;30] day period amount to -0.51%, -0.85%, and 0.97% for Bad, Neutral, and Good News, respectively.

Figure 19 PEAD: 1989-2018 | CAAR [2;30]



Source: Own creation

Figure 20 PEAD: 1989-2018 | CAAR [-5;30]



Source: Own creation

Figure 19 provide an illustrative depiction of the post-earnings announcement drift and the aggregation of CAAR for the [2;30] day drift period. Figure 20 provides the same illustration for the [-5;30] day period. The solid lines depict Good News, the dashed lines depict Neutral News, and the dotted lines depict Bad News.

8.2.2 PEAD: How the Effect Develops Over Time

In this section, we will discuss the results of the event study of post-earnings announcement drift for three separate time periods. First, we will look at the results from each period separately and consider the effect across news and indexation groups. Then, we provide an overview of the periods and consider whether there has been an overall trend over time. The objective of this overview is to consider whether the emergence of indexing has influenced our indexation groups differently.

8.2.2.1 PEAD: 1989-1998

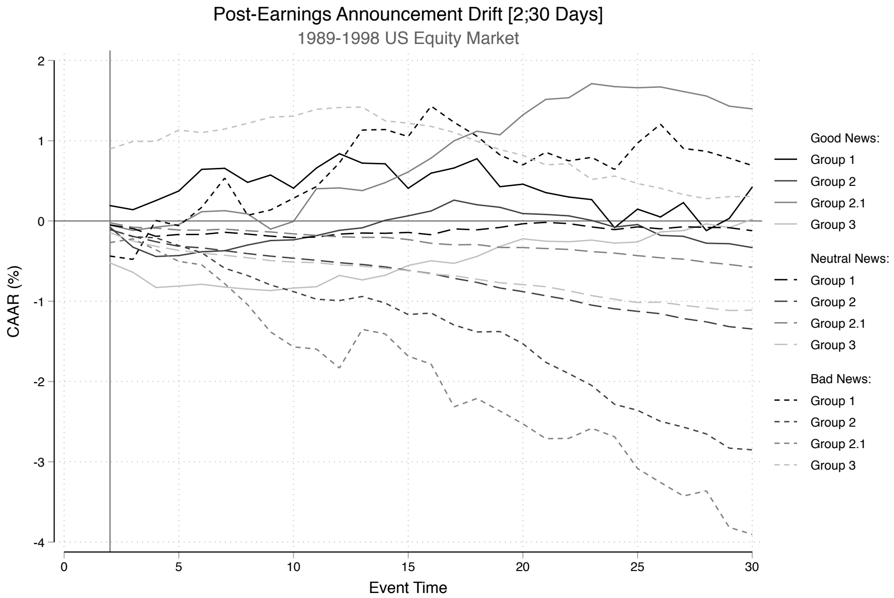
Table 16 – CAAR and t-statistic for 1989-1998 | Period [2;30]

	Group 1	Group 2	Group 2.1	Group 3
Bad News	0.69 %	-2.85 %	-3.90 %	0.31 %
	0.52	-7.93 ***	-4.12 ***	0.95
Neutral News	-0.12 %	-1.35 %	-0.58 %	-1.11 %
	-0.61	-21.90 ***	-2.31 **	-9.96 ***
Good News	0.42 %	-0.33 %	1.40 %	0.02 %
	0.34	-1.18	1.86 *	0.08

* Denotes significance at the 10%-level, ** Denotes significance at the 5%-level, *** Denotes significance at the 1%-level.

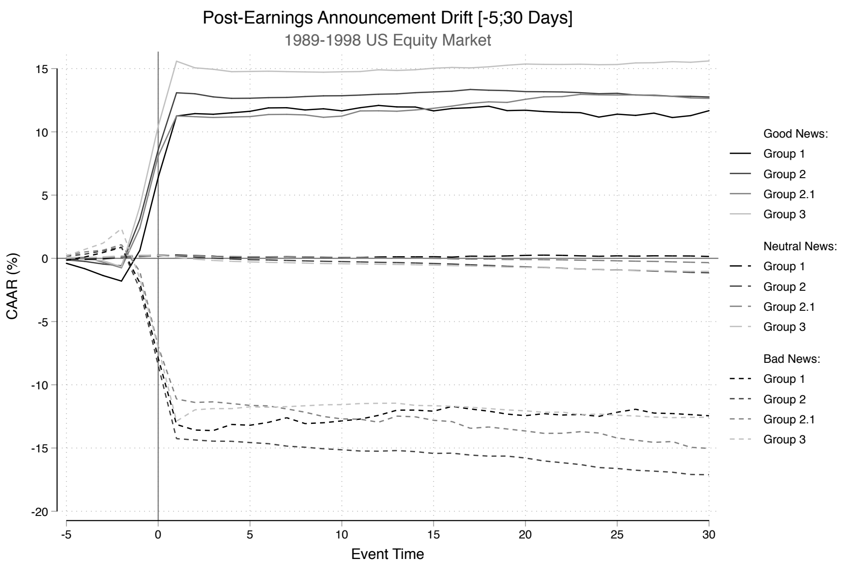
The first thing one notice, when considering the results of the 1989-1998 period is that the drifts are in general smaller and less significant than the full time-period. For instance, none of the CAARs for Group 1 are significant in this time-period. Compared to the full time-period, Group 2 exhibits drifts of similar direction, magnitude, and significance for Bad and Neutral News (-2.85% and -1.35%, respectively), while the drift for Good News is insignificant. Group 2.1 exhibits a large drift significant at the 1%-level for Bad News amounting to -3.90%. The drift of Neutral News is significant at the 5%-level and exhibits a small drift of -0.58%, while the drift of Good News is much larger than for the full-time period but only significant at the 10%-level. Finally, the drift for Group 3 is only significant for Neutral News and amounts to 1.11%. Figure 21 illustrates these results graphically, while Figure 22 shows the [-5;30] day period.

Figure 21 PEAD: 1989-1998 | CAAR [2;30]



Source: Own creation

Figure 22 PEAD: 1989-1998 | CAAR [-5;30]



Source: Own creation

8.2.2.2 PEAD: 1999-2008

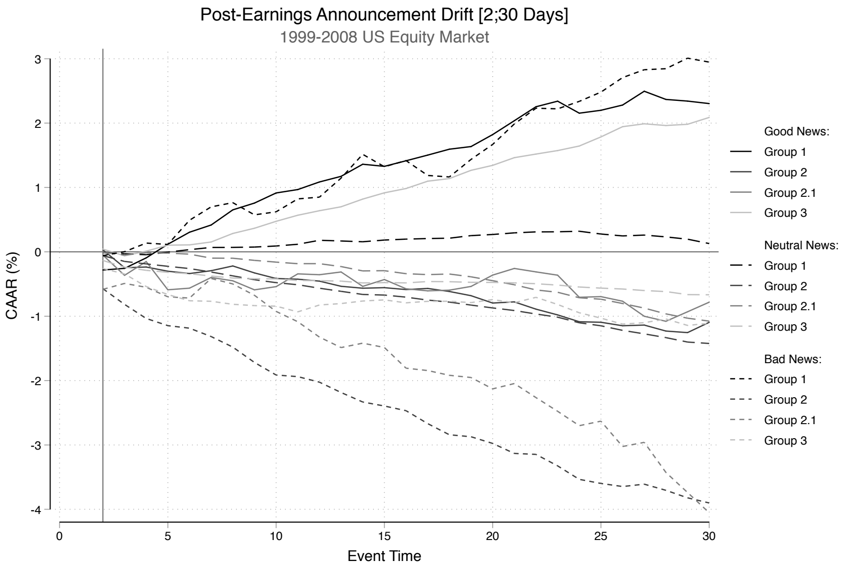
Table 17 – CAAR and t-statistic for 1999-2008 | Period [2;30]

	Group 1	Group 2	Group 2.1	Group 3
Bad News	2.95 %	-3.90 %	-4.05 %	-1.11 %
	3.51 ***	-11.61 ***	-3.98 ***	-4.12 ***
Neutral News	0.13 %	-1.42 %	-1.08 %	-0.67 %
	1.06	-19.62 ***	-6.92 ***	-7.03 ***
Good News	2.30 %	-1.09 %	-0.78 %	2.09 %
	3.95 ***	-4.05 ***	-1.02	9.30 ***

* Denotes significance at the 10%-level, ** Denotes significance at the 5%-level, *** Denotes significance at the 1%-level.

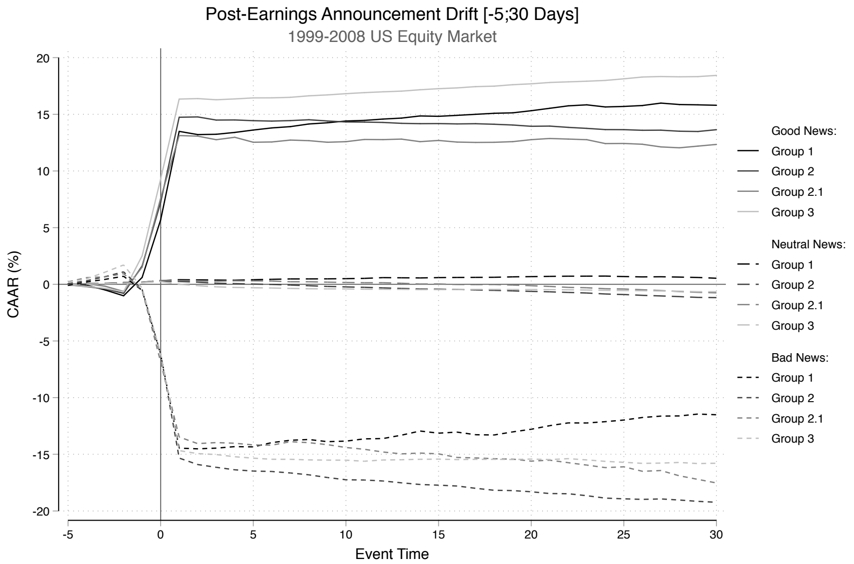
In the 1999-2008 time-period, the significance of the drifts reverts to the level of the full time-period. However, the magnitude of the drifts is in general larger. Group 1 exhibits large drifts significant at the 1%-level for Bad and Good News amounting to 2.95% and 2.30%, respectively. The drift of Neutral News is insignificant. Again, we find that the drift for Group 1 | Bad News is positive. Consistent with the full time-period, Group 2 exhibits highly significant negative drifts for all news groups. The drifts amount to -3.90%, -1.42%, and -1.09% for Bad, Neutral, and Good News, respectively. The results of Group 2.1 are also consistent with the full time-period. Bad and Neutral News exhibit significant drifts at the 1%-level of -4.05% and -1.08%, respectively. Good News remains not significant. All drifts of Group 3 are significant at the 1%-level. The drifts for Bad, Neutral, and Good News amount to -1.11%, -0.67%, and 2.09%, respectively. Figure 23 illustrates these results, while Figure 24 shows the full drift period [-5;30].

Figure 23 PEAD: 1999-2008 | CAAR [2;30]



Source: Own creation

Figure 24 PEAD: 1999-2008 | CAAR [-5;30]



Source: Own creation

8.2.2.3 PEAD: 2009-2018

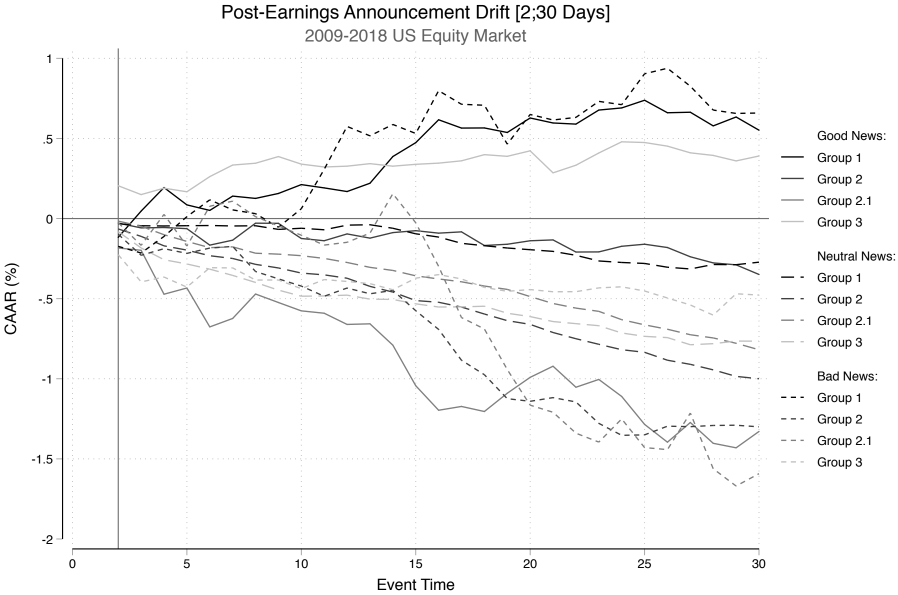
Table 18 – CAAR and t-statistic for 2009-2018 | Period [2;30]

	Group 1	Group 2	Group 2.1	Group 3
Bad News	0.66 %	-1.30 %	-1.59 %	-0.48 %
	1.26	-4.39 ***	-2.06 **	-1.81 *
Neutral News	-0.27 %	-1.00 %	-0.82 %	-0.77 %
	-3.97 ***	-19.41 ***	-10.96 ***	-9.21 ***
Good News	0.55 %	-0.35 %	-1.33 %	0.39 %
	1.40	-1.47	-2.16 **	2.09 **

* Denotes significance at the 10%-level, ** Denotes significance at the 5%-level, *** Denotes significance at the 1%-level.

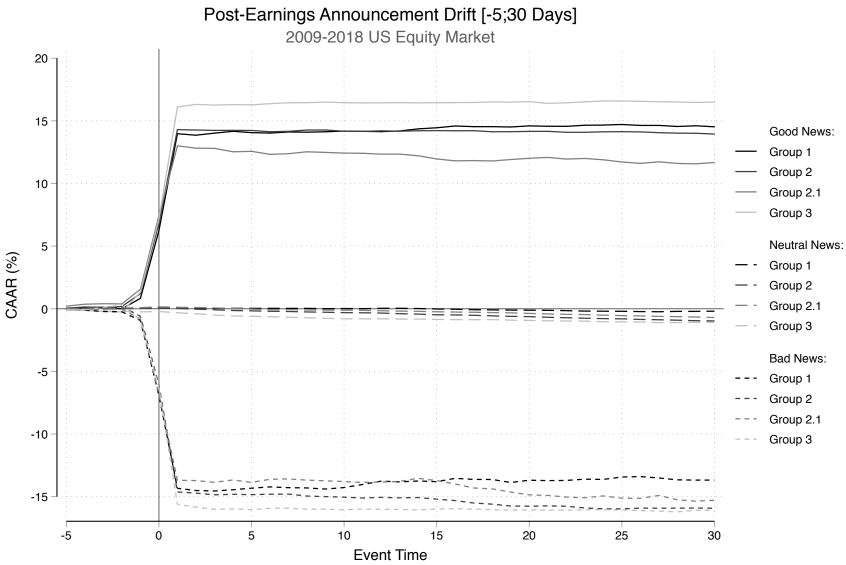
In the most recent time-period, there are fewer significant drifts compared to the previous period. Group 1 only exhibits a significant drift for Neutral News, which amounts to -0.27%. Group 2 exhibits smaller drifts compared to the previous time-periods. The drifts of Bad and Neutral News amount to -1.30% and -1.00%, respectively, while Good News is insignificant. All drifts of Group 2.1 are significant at the 5%-level, while Neutral News are also significant at the 1%-level. The drifts amount to -1.59%, -0.82%, and -1.33% for Bad, Neutral, and Good News, respectively. Group 3 | Bad News exhibit a small negative drift of -0.48% significant at the 10%-level. Neutral News are significant at the 1%-level, while Good News is significant at the 5%-level. They exhibit a drift of -0.77% and 0.39%, respectively. Figure 25 illustrate these results graphically, while Figure 26 shows the full period, [-5;30].

Figure 25 PEAD: 2009-2018 | CAAR [2;30]



Source: Own creation

Figure 26 PEAD: 2009-2018 | CAAR [-5;30]



Source: Own creation

8.2.2.4 PEAD: The Effect Across Time-Periods

Table 19 provides an overview of the three time periods. In general, it appears as if there are more significant drifts in the second period. For Group 1 and Group 3, the drifts are insignificant in the first period, before becoming significant at the 1%-level in the second period, and finally becoming less significant in the final period. Furthermore, the magnitude of the drifts is fairly large in the first period, before they become larger in the second period, while falling to their lowest level, on average, in the final period. In general, Group 2.1 exhibits larger and more significant drifts than Group 1 across all time-periods.

Table 19 – CAAR and t-statistic for Three Time Periods | Period [2;30]

	Group 1	**Group 2**	**Group 2.1**	**Group 3**
Bad News				
1989-1998	0.69 %	-2.85 %	-3.90 %	0.31 %
	0.52	-7.93 ***	-4.12 ***	0.95
1999-2008	2.95 %	-3.90 %	-4.05 %	-1.11 %
	3.51 ***	-11.61 ***	-3.98 ***	-4.12 ***
2009-2018	0.66 %	-1.30 %	-1.59 %	-0.48 %
	1.26	-4.39 ***	-2.06 **	-1.81 *
Neutral News				
1989-1998	-0.12 %	-1.35 %	-0.58 %	-1.11 %
	-0.61	-21.90 ***	-2.31 **	-9.96 ***
1999-2008	0.13 %	-1.42 %	-1.08 %	-0.67 %
	1.06	-19.62 ***	-6.92 ***	-7.03 ***
2009-2018	-0.27 %	-1.00 %	-0.82 %	-0.77 %
	-3.97 ***	-19.41 ***	-10.96 ***	-9.21 ***
Good News				
1989-1998	0.42 %	-0.33 %	1.40 %	0.02 %
	0.34	-1.18	1.86 *	0.08
1999-2008	2.30 %	-1.09 %	-0.78 %	2.09 %
	3.95 ***	-4.05 ***	-1.02	9.30 ***
2009-2018	0.55 %	-0.35 %	-1.33 %	0.39 %
	1.40	-1.47	-2.16 **	2.09 **

* Denotes significance at the 10%-level, ** Denotes significance at the 5%-level, *** Denotes significance at the 1%-level.

8.3 Test 3 – Regression Analysis

In this section, we will go through the results of the regression analysis performed to further draw inference from the event study. First, we need to assess if we can assume coherence to the four least squares assumptions for multiple regression models, which were described in detail in the methodology section. Therefore, we will now examine summary statistics, the correlation matrix, and the distribution of the variables used in our regression models.

In Table 20, the descriptive statistics for both the dependent and the independent variables are displayed. All variables have 57,076 observations. The dependent variable, CAR3, has a minimum of -1.17661 and a maximum of 1.22318 in the sample. There is a risk that Market Value is right skewed or contain outliers and therefore the natural logarithm of Market Value was created which displays descriptive statistics much closer to the normal distribution.

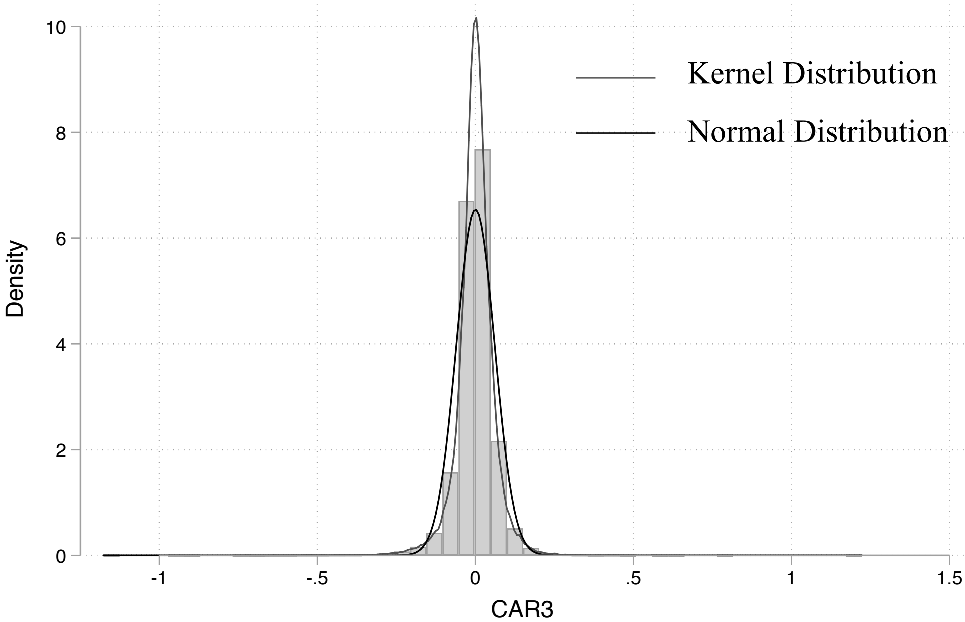
Table 20 Summary Statistics for Regression Variables

	n	Mean	Std. dev.	Min	25^{th}	Median	75^{th}	Max
CAR3	57,076	0.00045	0.06099	-1.17661	-0.02576	0.00166	0.02982	1.22318
SP	57,076	0.47439	0.49935	0.00000	0.00000	0.00000	1.00000	1.00000
t	57,076	17.55277	8.21012	1.00000	11.0000	18.00000	25.00000	30.00000
SP_x_t	57,076	9.40574	11.01737	0.00000	0.00000	0.00000	20.0000	30.00000
MarketValue	57,076	24,000,000	43,000,000	63,000	6,300,000	11,000,000	23,000,000	1,100,000,000
ln_MV	57,076	16.34542	1.06796	11.04836	15.66040	16.25323	16.96212	20.77635

Let us check the distribution of the variables as well before we conclude anything. Figure 27 on the following page displays the distribution of the variable CAR3. The black line in the figure represents the normal distribution. There is also a grey kernel imposed to better reflect the distribution of the histogram. From this graphical representation, we can conclude that the CAR3 variable is approximately normally distributed.

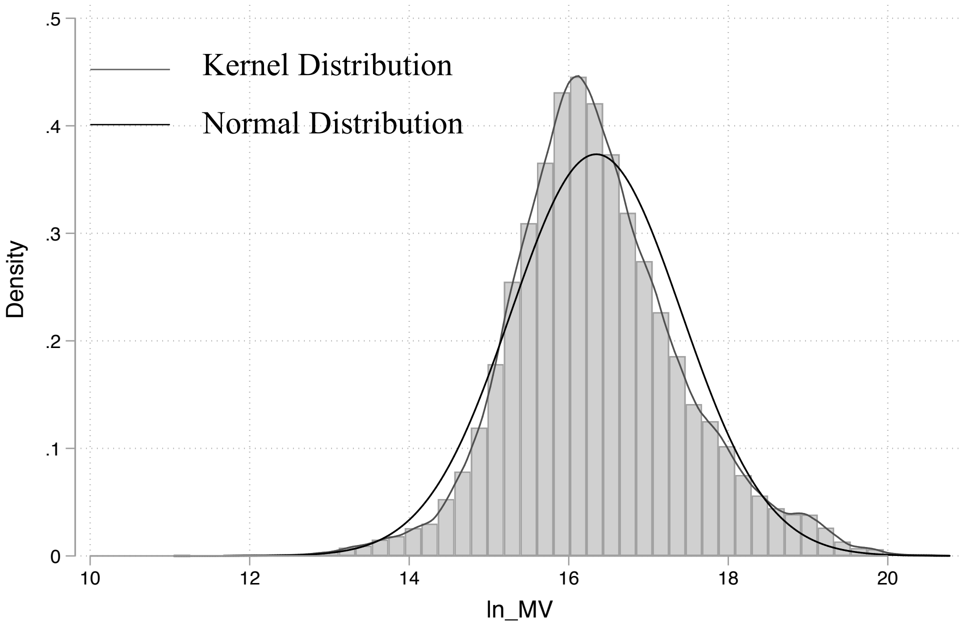
In Figure 28 on the following page, the distribution of the natural logarithm of Market Values is illustrated. The log-transformation of Market Value was done in order to avoid any potential skewness from distorting the results. From the graphical representation of ln_MV, it can be assumed that the variable has an approximately normal distribution.

Figure 27 Distribution of CAR3



Source: Own creation

Figure 28 Distribution of ln_ MV



Source: Own creation

We only examined the graphical representations of the distribution for the two variables ln_MV and CAR3 as the other variables in the two regression models should be normally distributed by construction.

If we move on to examining the correlation matrix in Table 21. All variables are significantly correlated with CAR3 on the 0.1% significance level.

Table 21 Correlation Matrix

	CAR3	SP	t	SP_x_t	ln_MV
CAR3	1				
SP	-0.0194***	1			
t	-0.0167***	0.263***	1		
SP_x_t	-0.0209***	0.899***	0.495***	1	
ln_MV	0.0146***	-0.226***	0.381***	-0.0662***	1

The S&P 500 dummy, time, and the interaction variable between the S&P 500 dummy and time all have a small negative correlation with CAR3. Whereas, the natural logarithm of Market Value has a small positive correlation with CAR3. Since CAR3 and ln_MV have a significant correlation at the 0.1%-level, it makes sense to include ln_MV as a control variable for Regression (2). All the significant correlations are weak except the correlation between the interaction variable and the variables it is the interaction between. However, it is not a perfect linear function of them. With the other correlations being weak, we can assume that there is no perfect multicollinearity in this sample.

Through the preceding analysis of the descriptive statistics, the distribution of the variables, and the correlation matrix we have reduced the risk of violating the four least squares assumptions. Through log-transforming Market Value and examine the graphical representation of the dependent variable, CAR3, and ln_MV we can conclude that they are approximately normally distributed. Thus, we can assume that we do not violate Assumption 3 of no large outliers. Through the correlation matrix, we learnt that there is no perfect multicollinearity in the sample and thus we assume that Assumption 4 holds. Further, we downloaded all data of the CRSP and merged CRSP/Compustat databases for 30 years for the three largest exchanges in the U.S. We selected all data and selected a large timeframe. Furthermore, we did not ourselves select the thresholds for the group splits, they are based on the sample data. We believe that our sample is independently and identically distributed draws from their joint distribution. Thus, we assume that Assumption 2 is not violated. In accordance with Assumption 1, we assume that the conditional distribution of u_i given X_i has a mean of zero. Thus, the model estimates will be unbiased. Further, in the regression models, we will control for heteroskedasticity through using heteroskedasticity-robust standard errors. We assume that all four assumptions hold so that the OLS estimators $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ will be jointly normally distributed for large samples.

The following two regression models were tested in Test 3.

(1) $$\widehat{CAR}3_{it} = \hat{\beta}_0 + \hat{\beta}_1 SP_{it} + \hat{\beta}_2 t + \hat{\beta}_3 SP\_x\_t + \hat{u}$$

(2) $$\widehat{CAR}3_{it} = \hat{\beta}_0 + \hat{\beta}_1 SP_{it} + \hat{\beta}_2 t + \hat{\beta}_3 (SP\_x\_t) + \hat{\beta}_4\, ln\_MV + \hat{u}$$

Heteroskedasticity robust standard errors were applied for both regressions with the regression results being presented in Table 22. In Regression (1) it is only time and the constant that are statistically significant. These are statistically significant at the 1%-level with very small coefficients. This is also clear from the very low R-squared of 0.0005 the independent variables do not contribute much in explaining the dependent variable. The coefficient for time is very small with a negative effect on CAR3.

Table 22 – Regression Results | 1989-2018

	(1)	**(2)**
C	0.0029575 (5.54)***	-0.016124 (-3.12)***
SP	-0.00000864 (-1.36)	-0.0006501 (-0.48)
t	-0.0000892 (-2.87)***	-0.0001588 (-4.23)***
SP_x_t	-0.0018165 (-0.14)	-0.0000227 (-0.36)
ln_MV	- -	0.0012164 (3.71)***
n	57,076	57,076
R^2	0.0005	0.0009

t-statistics in parenthesis

*Denotes significance at the 10%-level, ** Denotes significance at the 5%-level,

*** Denotes significance at the 1%-level.

For Regression (2), time and the constant remain significant at the 1%-level. However, we find that when we control for market capitalisation the R-squared value, and thus the predictive power of the model, is improved. However, it remains very limited. The coefficient for ln_MV is significant at the 1%-level. In this extended model, the dummy, SP, and the interaction variable, SP_x_t, remain insignificant. The magnitude of the coefficient for time becomes larger. Thus, it exerts an increasingly negative effect on CAR3.

9 Discussion and Implications

In this part of the paper, we will discuss the results presented in the preceding section. Firstly, we will consider how the magnitude and direction of our results compare to the existing literature on post-earnings announcement drift. Second, we will discuss the interpretation of our results in the context of our research question. Third and finally, we will discuss the implications.

9.1 A Comparison of Post-Earnings Announcement Drift

In this section, we will consider to what extent our results are similar to that of existing literature on post-earnings announcement drift. However, it should be noted that any comparison is problematic since it is difficult to find comparable studies. For instance, different studies use different estimation periods, different event windows and drift periods, different news definitions, different normal returns models, and different samples. Furthermore, studies of PEAD does not necessarily employ traditional event studies – more often trading strategies are devised. Nonetheless, we will attempt to provide comparable results to validate the results yielded in our study. Lastly, it should be noted that researchers continue to discuss whether PEAD is in fact a delayed price response or a compensation for risk (Bernard & Thomas, 1989). That is, whether any PEAD identified is the result of an inadequate risk-adjustment in the normal returns model.

Bernard and Thomas' (1989) seminal paper on post-earnings announcement drift on the NYSE and AMEX exchanges provides a good initial comparison. Their study employs an event study of PEAD, where they categorise earnings announcements based on a measure for standardised unexpected earnings (SUE). SUE is a measure for the difference in actual earnings compared to the forecast of earnings from a basket of analysts standardised by the variation previous earnings. They categorise Good News as the securities with the 10% highest SUEs and Bad News as the securities with the 10% lowest SUEs. However, where we aggregate abnormal returns over the [2;30] day period, they aggregate abnormal returns in the [0;60] day period. They find that Good News exhibit a drift of 2.19%, while Bad News exhibit a drift of -3.13%. The higher (lower) the SUE of the news is, the higher (lower) is the return. Compared to our results, the magnitude of the drifts is similar. However, the direction of the Bad News drift is different to our Test 1.

Similarly, Brandt et al. (2008) conduct a study of PEAD on the NYSE and AMEX exchanges. However, they do not conduct a traditional event study, instead they create a trading strategy. They do, however, report CAAR for the [1;30] day period following earnings announcements.

Where we use the market model to predict normal returns, they use the return on a benchmark size-adjusted book-to-market Fama-French portfolio. Furthermore, they define earnings surprises using SUE and earnings announcement return (EAR). EAR is the three-day cumulative abnormal return around the earnings announcement date (similar to CAR3, but with a different normal return measure). Where we use a decile approach to categorise earnings announcements, they use a quintile approach. Consistent with Bernard and Thomas (1989), they find that Bad News exhibit a negative drift, while Good News exhibit a positive drift. The drifts using the SUE (EAR) categorisation amount to -0.51% (-0.39%) and 1.08% (1.09%) for Bad and Good News, respectively. While the direction of the drift for Bad News is different from our study, the magnitudes of both drifts are similar.

Chordia, Goyal, Sadka, Sadka, and Shivakumar (2009) likewise conduct a study of PEAD, where they are interested in the effect of liquidity. Like the previous studies, and our study, they collect data from NYSE, NASDAQ, and AMEX on exchange codes 1, 2, and 3. Furthermore, they remove the most illiquid stocks by removing stocks trading for less than $5. In comparison, we remove microcaps. Just like Bernard and Thomas (1989) and Brandt et al. (2008), they categorise earnings based on SUE. However, they go on to perform a subsequent categorisation on liquidity. That is, they form portfolios for a combination of each decile of each category. For instance, the highest decile of SUE and the highest decile of liquidity. However, they do not report abnormal returns, instead they report raw returns accumulated over a month. Consistent with literature, they find that Good News (1.31%) firms provide higher returns than Bad News (0.62%) firms. Interestingly, when they form their portfolios of news and liquidity deciles, they find that for the most liquid decile of securities, the highest return comes from the lowest SUE decile. This is consistent with one of our surprising results. Namely that our Group 1 | Bad News exhibits a positive drift across all time periods. Recall that Group 1 is the S&P 500 firms, which are highly liquid.

In summary, the evidence presented in this section seems to suggest that the result derived from our Test 1 | Bad News is inconsistent and surprising. As alluded to earlier, we will discuss possible sources of this result in our Limitations. However, in general the direction and magnitude of the drifts reported in this paper is consistent with literature. It is particularly interesting that Chordia, et al. (2009) finds supporting evidence of the peculiar result deriving from Group 1 | Bad News. In conclusion, we perceive our results to be consistent with literature.

9.2 The Effect of Indexing

Does the increase in passive investment have an effect on the level of stock market efficiency?

In this paper, we employ an event study of post-earnings announcement drift and a difference-in-difference inspired regression in an attempt to identify the effect of increased indexing on stock market efficiency. Based on an extensive literature review, we developed three hypotheses, which suggest how the efficiency of markets may have been influenced. In this section, we interpret the results, and discuss the findings of our thesis in relation to our research question and hypotheses.

9.2.1 Interpretation of Results

The results of our event study for the full time-period, 1989-2018, are highly significant. 10 out of 12 groups exhibit a drift significant at the 1%-level. This would seem to indicate that the price response to earnings announcements in the full time-period is inefficient. This is not surprising since PEAD has been documented in the literature repeatedly. This paper, however, is interested in whether the rise of index investing over the past 30 years have influenced the efficiency of markets. To answer that question, we devised two groups which are highly comparable in terms of quantitative informational efficiency indicators. However, we conjecture that Group 1, the S&P 500 firms, has become increasingly indexed compared to Group 2.1. Hence, it is interesting to compare the results for these two groups.

The two groups appear to exhibit fairly dissimilar drifts. Group 1 | Good News exhibit a positive drift amounting to 1.39% (4.02***), whereas Group 2.1 | Good News exhibit a non-significant negative drift of -0.55% (-1.20). However, the result for Bad News is more surprising. Consistent with literature, Group 2.1 exhibits a negative drift (-3.23% (-5.49***)) following the announcement, however, Group 1 exhibits a large, significant positive drift (1.75% (3.69***)). Chordia et al. (2009) find that the most liquid Bad News securities exhibit a positive return following an earnings announcement. Nonetheless, this does not seem to explain the discrepancy between Group 1 and Group 2.1 since they are equally liquid according to our definition. Furthermore, it would be surprising to find that the investor is rewarded with a risk premium for holding liquid stocks. In general, it is unclear whether there is any pattern in the efficiency of the two groups in the full time-period. That being said, what we are most interested in is whether the efficiency of the groups develop differently as time passes.

If we shift our attention to the three time-periods tested, there appears to be a general pattern across news and indexation groups. Markets appear to become less efficient in the second period, while being mostly efficient in the most recent period. This is illustrated by lower and less significant drifts in the final period. However, to answer our research question, we are more interested in the development of Group 1 vis-à-vis Group 2.1. Group 1 | Bad News exhibits a similar sized insignificant drift in the first and last period. In comparison, Group 2.1 exhibits a substantially lower drift, which is also less significant. This would seem to indicate that Group 2.1 | Bad News has become more efficiently priced in the course of the 30-year period. For Neutral News, the magnitude of the drift does not change substantially for either of the groups, while the drifts of both groups have become more significant. Finally, the results for Good News does not seem to be decisive either. Group 1 exhibits a non-significant drift at a similar magnitude in both periods, while the drift of Group 2.1 has become increasingly negative. Since these results are unclear, and do not provide any concluding evidence regarding our hypotheses, we elaborate the model and conduct a difference-in-difference inspired regression in an attempt to tease out the effect of increased indexing.

The results of the D-in-D inspired regression analysis supports the interpretation of the event study. Since the control variable ln_MV has a significant effect and contributes to the explanatory power of the model, we will discuss the results of Regression (2). The interaction variable between time and the S&P dummy shows a small insignificant effect (-0.0000227(-0.36)). This would seem to suggest that being included in the S&P 500, and thus being increasingly passively held, has not had an effect on development over time in the level of stock pricing efficiency of the stock – negative nor positive. However, as the results of the event study indicated, time in isolation has had an effect. The variable for time has a small significant impact on the level of stock pricing efficiency. We see that the coefficient is -0.0001588 (-4.23***), which means that CAR3 has been reduced as time has passed. The effect may appear small, but if we aggregate it over 30 years, it amounts to -0.004764. Recall from the summary statistics that for Group 1 and Group 2.1 the mean CAR3 amounted to -0.0008 and 0.0016, respectively. Thus, the effect is fairly large. Since a smaller CAR3 indicates a more efficient response to earnings announcements, we can conclude that our results seem to indicate that security prices have, for Group 1 and Group 2.1, become more efficient.

9.2.2 Concluding Remarks

Our results do not provide any conclusive evidence regarding our research question. We find no evidence to support Hypothesis I (increasingly efficient) or Hypothesis II (decreasingly efficient). Hypothesis III suggests that the (1) net effect of increased indexing is negligible or (2) increased indexing does not change the equilibrium between information acquisition costs and expected return. That is, Hypothesis III proposes that we should not see an effect of increased indexing on the level of stock pricing efficiency. We find weak evidence to support this hypothesis.

Noise traders were central to our hypotheses development. We hypothesised that the development in the proportion of noise traders and that effect would be instrumental in determining the effect on asset pricing efficiency. However, as we discussed, it could have both positive (additional short-selling, less uninformed traders, and more corporate governance) and negative consequences (less liquidity and price discovery) for efficiency. The results collected by this thesis seems to indicate that these effects are either non-existent or have had a net neutral effect. Consequently, our results seem to indicate that the impact of increased indexing is negligible.

The theory behind efficiently-inefficient markets rests on the equilibrium between information acquisition costs and expected returns (Garleanu & Pedersen, 2018). That is, for the level of asset pricing efficiency to change, either one of these factors need to change. The results of our event study and regression analysis seem to indicate that markets have in general become more efficient in the course of the 30-year period. It is likely that this development is driven by reduced costs of information acquisition. In particular, the development in IT and the spread of the internet has made more information easily available to a much wider audience of private and institutional investors. Hence, more information can be collected and reflected in the price at a lower cost. Consequently, prices will reflect more information and in general be more efficiently priced.

Finally, we must consider the option that there is in fact an effect of increased indexing on asset pricing efficiency, but that we have been unable to identify it. The limitations section of this paper will be committed to discussing short-comings in this thesis and the research design. In conclusion, this paper does not find any evidence to support that the level of indexing should have influenced the development in the level of stock pricing efficiency.

9.3 Implications

While researchers are questioning the implications of increased index investing on market efficiency and practitioners are calling indexing “worse than Marxism” (Kawa, 2016), our results seem to indicate that the rise of indexing thus far does not pose a threat to market efficiency. Hence, this would indicate that thus far the consequences suggested by active managers are largely unfounded. That being said, a market composed of 100% passive investing would have detrimental effects on market efficiency and capital allocation (Pedersen, 2018). Hence, it is interesting to consider whether there is an optimal level of indexing beyond the current level.

According to Sharpe’s (1991) arithmetic of asset management, active management is a net negative endeavour. Consequently, investors should continue to move money out of active management and into passive management. Thus, continuing to fuel the index revolution and eventually creating a market of exclusively passive managers. As suggested by practitioners and researchers alike, this would have wide reaching consequences for the markets ability to allocate capital and reflect information.

Contrastingly, Pedersen (2018) sharpens Sharpe’s arithmetic and shows that active managers can profit from active management, because even passive managers have to trade. These profits will be more pronounced, when active managers are superior in information acquisition. As markets become increasingly passively held, active managers should be able to continue to profit in aggregate. Pedersen (2018) goes on to conclude that he expects the fraction of passive management to increase to a proportion less than 100%. Our results seem to support this notion. In the past 30 years, the proportion of passive management has continued to increase. However, so far, the market powers of capital markets have continued to correct prices. Consequently, the gap between the efficiency of S&P 500 firms (Group 1) and similar firms (Group 2.1) has not expanded. Our results seem to imply that there is an equilibrium level of passive investment beyond the current level – and likely below 100%. Thus, we provide empirical evidence to support the conclusions reached by Pedersen (2018).

Though we only provide weak evidence for Hypothesis III, it is interesting to consider how our results fit in with the literature on market efficiency at large. The theory on the strong efficient market hypothesis would implicate that markets at all times reflect all available information (Fama, 1970). That is, markets cannot become more or less efficient – they will always remain perfectly efficient. However, the results of this thesis seem to suggest that markets have in fact become more efficient in the course of the past 30 years. This result is inconsistent with the strong form efficient market hypothesis. Furthermore, according to the semi-strong form of market efficiency, all public information should be reflected in the price immediately (ibid.). However, the majority of the drifts derived from the event study are significant. In theory, prices should react immediately and effortlessly to new information. Nonetheless, since our results seem to indicate that prices appear to be drifting into place, this is inconsistent with the semi-strong form efficient market hypothesis.

In contrast, our results seem to support the efficiently inefficient hypothesis (Pedersen, 2015). It is likely that the proportion of noise traders have fallen as the proportion of index investing has increased (Israeli, Lee, & Sridharan, 2017). In theory, this should impact liquidity, price discovery, and thus the efficiency of markets. However, that is not what we find. Instead, markets seem to adapt and restore equilibrium (as in classic microeconomic supply and demand models). If additional inefficiencies arise, arbitrageurs come in and apply superior information acquisition to correct prices and reap an adequate return. Consequently, the level of market efficiency reverts to its initial level. Conversely, if markets become more efficient, less active managers will be able to reap an adequate return. Consequently, they will go out of business, reducing the level of information acquisition, and thus markets will revert to its initial level. We are unaware, which mechanism is taking place. We can only observe that the difference in the level of efficiency between S&P 500 firms (Group 1) and similar firms (Group 2.1) appears to remain throughout the 30-year period tested.

10 Limitations and Robustness Check

In this section, we will debate the limitations of the research design employed in this paper. Furthermore, we will discuss what we have done to mitigate them and why we have chosen to move forward regardless. Secondly, we will test whether the results of the event study are robust to the test of a different normal returns measure – the Fama-French Three Factor + Momentum model.

10.1 Limitations

Throughout this study, we have done our utmost to make it robust and comparable to other studies. However, our study is still limited in some veins. First, we use CAR3 to categorise news rather than standardised unexpected earnings (SUE). In the literature on earnings surprises, this measure has often been used to capture earnings surprises. SUE is defined as:

$$SUE = \frac{X_{i,q} - E(X_{i,q})}{\sigma_{i,q}} \tag{21}$$

, where $X_{i,q}$ is the actual earnings of firm i at quarter q, $E(X_{i,q})$ is the expected earnings of the same firm at the same time, while $\sigma_{i,q}$ is the standard deviation of earnings over the preceding quarters (Brandt, et al., 2008). Expected earnings can either be derived from seasonal random walk with drift models (ibid., Bernard & Thomas, 1990) or by averaging analyst forecasts (Campbell, et al., 1997). By using a different metric for earnings surprises, our study becomes less comparable.

Second, when we define earnings surprises, we use a non-time variant measure of CAR3. That is, we use the same threshold throughout the time period. In Test 3, we show that markets have on average become more efficient over time. Hence, we fail to take account of this, when we define news. That being said, we do not think it alters our results significantly. However, we do recognise that it is inconsistent to assume a stable CAR3, while simultaneously showing that it is changing. Furthermore, late in the thesis process, we realised that the CAR3 thresholds were based on the full dataset, i.e. including microcaps. Thus, the news definition was more extreme than could have otherwise been expected. However, to correct this error, we would have been required to rewrite the code and rerun the entire event study. A process we did not have time to do.

Third, this thesis is based on more than 430 million data points. This vast amount of data hampers our ability to get an overview of the data and challenges our ability to work with it. Consequently, we cannot meticulously go through it and check for any patterns in null values or whether functions work perfectly. Furthermore, working with such a large dataset turned out to be an almost impossible task with the resources we had at hand. Consequently, we have been unable to meticulously go through the data and verify all functionalities.

Fourth, the data sources used are incomplete. Consequently, we experience some data loss throughout our data preparation process. This data loss is most impactful, where we lose quarterly announcement dates, RDQs. During the initial data cleaning process, the amount of observations is reduced from 863,998 to 672,203 due to missing RDQs. For instance, we know that there should be 500 firms in the S&P 500, which we assume file four quarterly reports per year. This should yield 60,000 observations over the 30-year period. However, in our dataset we only have 27,076 observations for S&P 500 firms. If this data loss was more pronounced in Group 1 compared to Group 2.1 (or conversely) it may skew our results. Furthermore, we use the same databases as other studies (Campbell, et al., 1997; Hoe, Xue, Zhang, 2018; Poulsen, 2018). That is, it must be recognised in the research community that these are the best sources of data – incomplete as they may be.

Fifth, the similarity of our groups can be contested. It is paramount for the validity of our study, that we trust that Group 1 and Group 2.1 are similar. It is one of the underlying assumptions of the D-in-D inspired regression (Angrist & Pischke, 2015, pp. 175-208). In this study, we have focussed on matching the components of the groups based on Market Value and Dollar Liquidity. These criteria were chosen because they are quantitative and good indicators of informational efficiency. Furthermore, we provide extended summary statistics on other variables to allow for further comparisons.

Sixth, when we decide how to limit our Group 3 and remove microcaps, we had the choice of removing firms with market values below the 20th percentile of NYSE (Hou, Xue, & Zhang, 2018) or remove securities with a price of less than \$5 (Cohen & Frazzini, 2008). We choose to omit convention and instead use our own threshold level. We choose to do so because it enables us to use the same methodology for the Group 2 and Group 3 threshold levels. Furthermore, it did not require us to extract new data. However, we recognise that it makes the study less comparable to other literature. That being said, we are convinced that it yields similar results to the two other methodologies. Furthermore, when we conduct Test 1, we have not removed microcaps from this dataset. Test 1 is conducted on the full dataset, while microcaps are not removed until the definition of Group 3. This may potentially skew the results and explain the surprising upward drift for Bad News.

Seventh, it is possible that our results are subject to omitted variable bias. That is, they could be driven by some factor, which we have not accounted for. In the following section, we apply a different measure of normal return. Particularly, we apply the Fama-French Three Factor + Momentum model to predict normal return. Thus, we are able to test whether the results of the event study were driven by either of the four factors.

Eighth, our study is conducted on the U.S. equity market exclusively. That is, the findings of this paper may not apply to other equity markets. The U.S. equity market is different from other equity markets globally – in particular in terms of the degree of indexing. Ninth, our study applies equal-weighted returns rather than value-weighted returns in the event study. This may put greater emphasis on returns of the lowest market capitalisation firms in each indexation group. As the summary statistics show, there is considerable variance in market value of the stocks in each group. Tenth, this study employs a relatively short estimation window for normal return. We do this to avoid the estimation window of one event to overlap with the event window of a previous events.

10.2 Fama-French Three Factor + Momentum Controlled Study

In the literature, multiple factors explaining abnormal return has been identified. For instance, Fama and French (1993) find that the size and value factor can explain a lot of alpha generated by some strategies. Subsequently, Carhart (1997) and Asness (1994) unveil the momentum factor, which also provide some explanatory power. In this section of the paper, we employ the Fama-French Three Factor + Momentum (FF3+M) model to predict normal return. Cable and Holland (1999) find that more extended models rarely add much explanatory power. However, it is interesting to see whether our results can be explained by the choice of normal returns model.

Table 23 – CAAR from Market Model and Fama French 3 Factor + Momentum | 1989-2018 – Period [2;30]

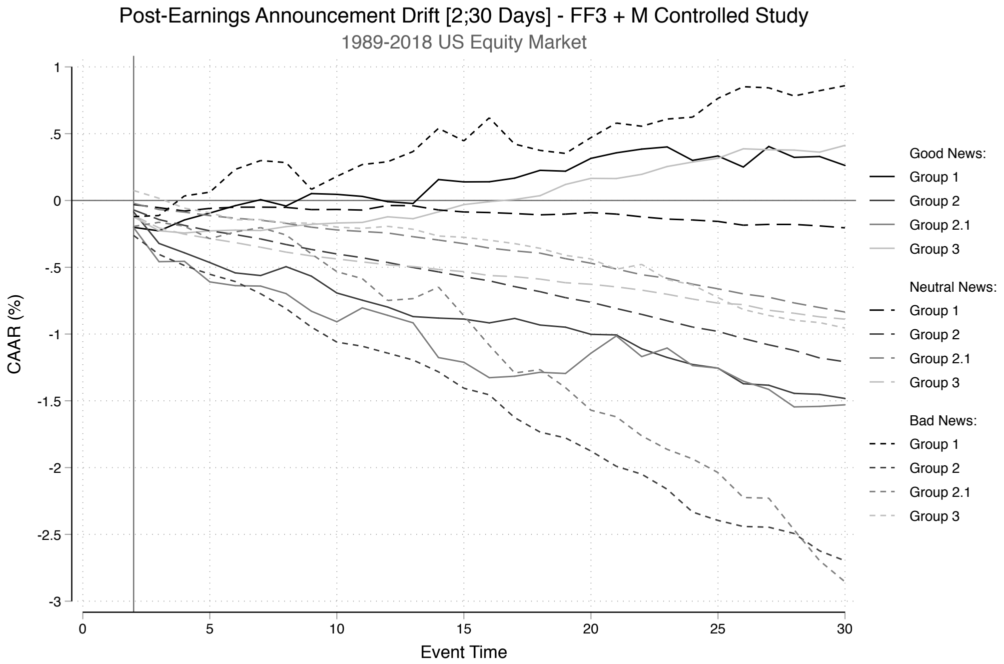
	Group 1	Group 2.0	Group 2.1	Group 3
Bad News				
Market Model	1.75 %	-3.02 %	-3.23 %	-0.51 %
	3.69 ***	-14.66 ***	-5.49 ***	-2.99 ***
FF3 + MOM	0.86 %	-2.70 %	-2.86 %	-0.95 %
	1.82 *	-12.95 ***	-4.95 ***	-5.41 ***
No News				
Market Model	-0.10 %	-1.29 %	-0.82 %	-0.85 %
	-1.69 ***	-21.28 ***	-15.85 ***	-8.05 ***
FF3 + MOM	-0.20 %	-1.21 %	-0.84 %	-0.89 %
	-3.31 ***	-19.44 ***	-16.04 ***	-8.14 ***
Good News				
Market Model	1.39 %	-0.72 %	-0.55 %	0.97 %
	4.02 ***	-4.37 ***	-1.20	6.90 ***
FF3 + MOM	0.26 %	-1.48 %	-1.53 %	0.41 %
	0.22	-8.85 ***	-3.30 ***	2.85 ***

* Denotes significance at the 10%-level, ** Denotes significance at the 5%-level, *** Denotes significance at the 1%-level.

Table 23 provides an overview of the results derived from the market model compared to the FF3+M model. Figure 29 on the following page illustrates the [2;30] day period graphically. Broadly speaking, the results are unaffected by the choice of normal returns model. However, it is interesting to see that the drift and significance diminish considerably for Group 1 Bad News as well as Good News. One of the results, which stood out in the initial test was the positive drift of Group 1 | Bad News.

When we control for the size, value, and momentum factor, we find that this drift is much smaller and less significant. This would seem to indicate that the results were driven by one of these factors.

Figure 29 PEAD: 1989-2018 | CAAR [2;30] | FF3+M Controlled Study



Source: Own creation

The summary statistics unveiled that the mean market value was considerably lower for Bad as well as Good News. This would suggest that there are more small firms measured by market capitalisation in these groups. That is, it could be that the omittance of the size factor from our original normal returns model may have resulted in the drift appearing larger and more significant. Appendix IV shows the summary statistics for the book value of Group 1 subject to the different news categories. They show that the Book Value falls just as the Market Value, when news is more extreme. For instance, mean Book Value of Group 1 | Bad News is 75% of Group 1 total, whereas Market Value of Group 1 | Bad News is also 75% of Group 1 total. However, the Book Value of Group 1 | Good News falls more than the Market Value of Group 1 Good News. This could explain why the FF3+M model changes the results more for Good News than Bad News. That is, that the lower Book Value compared to Market Value indicates a higher average book-to-market value in this group. Hence, the value factor may explain some of the results for Group 1 | Good News, but not for Group 1 | Bad News. It would be optimal to compute the book-to-market value, but for reasons discussed in the Data Presentation section, we have been unable to do so.

11 Further Research

There are multiple options for further research. Firstly, this study is solely concerned with the U.S. equity market. The study could be replicated and conducted on data from other equity markets to see whether similar patterns persist. However, it should be noted that the U.S. equity market is considerably more passively held than most other markets in the world. Secondly, this study is exclusively focussed on equities. However, passive investors also invest in other asset classes such as fixed income. The study could potentially be replicated with a focus on different asset classes. Thirdly, our study uses CAR3 to categorise earnings announcements. As we discussed previously, other researchers employ SUE when conducting this categorisation. It would be interesting to see whether the results are robust to a different earnings announcement metric. Fourthly, this study uses equal-weighted returns, which puts too much emphasis on low market capitalisation firms. Further research could explore how the effects appear, when using value-weighted returns. Fifthly, as pointed out in the limitation section, this paper has multiple short-comings. Further research could be conducted to address these. Lastly, it is evident that this paper does not provide decisive evidence on the effect of indexing. Hence, additional research in the field is required to determine whether the rise of indexing has impacted the level of market efficiency.

12 Conclusion

The past decades have seen an almost exponential increase in assets under management held by passive investors. Today, passive investments occupy more than 30% of the market and is said to grow to beyond half of the market by 2024. The effects of this has become an important point of discussion. The purpose of this paper was to determine whether the rise of passive investments has influenced stock pricing efficiency. In doing so, this paper has contributed to the emerging literature on the effect of indexing and the rise of passive management. Furthermore, we contribute to the vast literature on market efficiency and post-earnings announcement drift (PEAD).

We conduct three tests to identify the effect of increased indexing on asset market efficiency. First, we conduct an event study of PEAD without conducting a subsequent sub-division of the news groups. That is, we test whether there is an underlying trend in the different news groups. Perplexingly, and in contrast to literature, we find that Bad News exhibit a significant upwards drift. Furthermore, consistently with literature, Good News exhibit a significant positive drift. Second, we revert to our research question, and test whether there is an effect on increased indexing. Again, we conduct an event study of PEAD, but continue to divide news groups into indexation groups. Our results do not seem to suggest that there is a significant difference in the development of the different indexation groups. Finally, to quantify this impression, we conduct a difference-in-difference inspired regression, where we compare S&P 500 firms (highly indexed) with similar firms, which are not included in the index (less indexed). We find no evidence to support that stocks which are more indexed exhibit a different development vis-à-vis less indexed, but otherwise similar, securities. Instead, our results seem to indicate that the net effect of increased indexing on the level of stock pricing efficiency is negligible. That is, we find no support for Hypothesis I and Hypothesis II, while providing weak evidence for Hypothesis III. Furthermore, our results seem to indicate that markets have become more efficient during the 30-year period examined.

We go on to consider what implications our results have for index investing and the literature on market efficiency. Our results seem to suggest that indexing does not pose a threat to market efficiency and capital markets. Hence, on this dimension, there is no hinderance to the continued growth of index investing. However, consistent with literature, we believe that there is an equilibrium level of indexing beyond the current level, but below 100%. Furthermore, the drifts exhibited in the event study are inconsistent with the literature on strong and semi-strong market efficiency. The strong form efficient market hypothesis would imply that security prices at all times reflect all available information. However, we find weak evidence to suggest that markets have become more efficient. Likewise, semi-strong form efficient market hypothesis suggest that markets reflect all publicly available information and reacts immediately and seamlessly to new information. However, the evidence from the event study shows that prices tend to drift into place. Hence, our study provides additional evidence, which opposes strong and semi-strong form efficient market hypothesis. Contrastingly, our results seem to suggest that markets adapt and adjust to new contingencies through the market mechanisms of the trade-off between information acquisition costs and expected return. That is, our results seem to support the efficiently inefficient hypothesis of market efficiency.

In conclusion, the results of this thesis suggest that the level of market efficiency has been unaffected by the increase in indexing. However, we do not provide conclusive evidence to answer our research question. Hence, additional research in the field must be conducted to determine, whether there is a causal relationship between the rise of passive investments and the level of market efficiency.

13 Bibliography

Aghion, P., Reenen, J., & Zingales, L. (2013). Innovation and Institutional Ownership. *American Economic Review, 103*(1), 277-304.

Aiyagari, S., & Gertler, M. (1991). Asset return and transactions cost and uninsured individual risk. *Journal of Monetary Economics*, 309-31.

Alquist, R., Israel, R., & Moskowitz, T. (2018). Fact, Fiction, and the Size Effect. *Journal of Portfolio Management*, 34-61.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 223-49.

Angrist, J., & Pischke, J.-S. (2015). *Mastering Metrics: The Path from Cause to Effect.* Princeton: Princeton University Press.

Appel, I., Gormley, T., & Keim, D. (2016). Passive Investors, Not Passive Owners. *Journal of Financial Economics, 121*(1), 111-141.

Arnold, C. (2018, 05 31). *The George Washington Of Investing Wants You For The Revolution*. Retrieved from NPR.com: https://www.npr.org/2015/10/21/443192311/the-george-washington-of-investing-wants-you-for-the-revolution

Arnott, R., Hsu, J., Kalesnik, V., & Tindall, P. (2013). The Surprising "Alpha" from Malkiel's Monkey and Upside-Down Strategies. *Journal of Portfolio Management*, 91-105.

Ashenfelter, O., & Card, D. (1985). Using the Longitudinal Structure of Earnings to Estimate the Effect of Traning Programs. *The Review of Economics and Statistics*, 648-660.

Asness, C. (1994). Variables that Explain Stock Returns. *University of Chicago (Ph.D. Dissertation)*.

Asness, C., Frazzini, A., & Pedersen, L. H. (2018). Quality Minus Junk. *Review of Accounting Studies*, 1-79.

Asness, C., Frazzini, A., Israel, R., & Moskowitz, T. (2014). Fact, Fiction, and Momentum Investing. *Journal of Portfolio Management*, 75-92.

Asness, C., Frazzini, A., Israel, R., & Moskowitz, T. (2015). Fact, Fiction, and Value Investing. *Journal of Portfolio Management*, 34-52.

Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting Numbers. *Journal of Accounting Research*, 159-78.

Becker, G. (1976). *The Economic Approach to Human Behaviour.* Chicago: University of Chicago Press.

Belasco, E., Finke, M., & Nanigian, D. (2011). The Impact of Passive Investing on Corporate Valuations. *Managerial Finance*.

Berkman, H., & Truong, C. (2008). Event Day 0? After-Hours Earnings Announcements. *Journal of Accounting Research*, 71-103.

Bernard, V., & Thomas, J. (1989). Post-Earnings Annoucement Drift: Delayed Price Response or Risk Premium. *Journal of Accounting Research*, 1-36.

Bernstein, S., & Sheen, A. (2016). The Operational Consequences of Private Equity Buyouts: Evidence from the Restaurant Industry. *The Review of Financial Studies*, 2387-2418.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Difference-In-Difference Estimates. *The Quarterly Journal of Finance*, 249-75.

Bleiberg, S., Priest, W., & Pearl, D. (2017). *The Impact of Passive Investing on Market Efficiency.* Epoch Investment Partners.

Boehmer, E., & Wu, J. (2012). Short selling and the price discovery process. *The Review of Financial Studies*, 287-322.

Bogle, J. (2005). The Relentless Rules of Humble Arithmetics. *Financial Analysts Journal*, 22-35.

Bogle, J. (2007). *The Little Book on Common Sense Investing.* Chicago: John Wiley & Sons.

Bogle, J. (2018, 11 29). *Bogle Sounds a Warning on Index Funds*. Retrieved from Wall Street Journal: https://www.wsj.com/articles/bogle-sounds-a-warning-on-index-funds-1543504551

Brandt, M., Kishore, R., Santa-Clara, P., & Venkatachalam, M. (2008). Earnings Announcements are Full of Surprises. *SSRN Electronic Journal*.

Brown, A. (2018, 05 06). *Buying and Holding the Market has Many Pitfalls*. Retrieved from bloomberg.com: https://www.bloomberg.com/view/articles/2018-05-06/trying-to-mimic-the-stock-market-can-be-perilous

Brown, S., & Warner, J. (1980). Measuring Security Price Performance. *Journal of Financial Economics*, 205-38.

Brown, S., & Warner, J. (1985). Using Daily Stock Returns - The Case of Event Studies. *Journal of Financial Economics*, 3-31.

Bryman, A., & Bell, E. (2011). *Business Research Methods.* Oxford University Press.

Bunzel, T., Cohen, S., Madhavan, A., Matthews, S., Novick, B., & Sethi, J. (2017). *Index Investing Supports Vibrant Capital Markets.* BlackRock.

Burger, D. (2018, 02 27). *Investing in Index Funds Is No Longer Passive*. Retrieved from bloomberg.com: https://www.bloomberg.com/news/articles/2018-02-27/passive-becomes-the-new-active-as-indexing-rules-everything

Cable, J., & Holland, K. (1999). Modelling Normal Returns in Event Studies: A Model-Selection Approach and Pilot Study. *The European Journal of Finance*, 331-41.

Campbell, C., Cowan, A., & Salotti, V. (2010). Multi-Country Event Study Methods. *Journal of Banking and Finance*, 3078-90.

Campbell, J., Lo, A., & MacKinlay, C. (1997). *Econometrics of Financial Markets.* Princeton: Princeton University Press.

Campbell, J., Lo, A., & MacKinlay, C. (1997). *Econometrics of Financial Markets.* Princeton: Princeton University Press.

Carhartt, M. (1997). On Persistence in Mutual Fund Performance. *Journal of Finance*, 57-82.

Chan, L., Jegadeesh, N., & Lakonishok, J. (1996). Momentum Strategies. *Journal of Finance*, 1681-1713.

Chancellor, E. (2018, 03 22). *Breakingviews - Chancellor: Index investment riskier than it looks*. Retrieved from Reuters.com: https://www.reuters.com/article/us-funds-indexes-breakingviews/breakingviews-chancellor-index-investment-riskier-than-it-looks-idUSKBN1GY1RZ

Chen, Y., Kelly, B., & Wu, W. (2018). Sophisticated Investors and Market Efficiency: Evidence from a Natural experiment. *The National Bureau of Economic Research (working paper)*.

Chordia, T., Goyal, A., Sadka, G., Sadka, R., & Shivakumar, L. (2009). Liquidity and the Post-Earnings Announcement Drift. *Financial Analysts Journal*, 18-32.

Chowdhury, M., Howe, J., & Lin, J.-C. (1993). The Relation Between Aggregate Insider Trading and Stock Returns. *The Journal of Financial and Quantitative Analysis*, 431-37.

Cohen, L., & Frazzini, A. (2008). Economic Links and Predictable Returns. *Journal of Finance*, 1977-2011.

Cremers, K., & Petajisto, A. (2009). How Active is your Fund Manager? A New Measure that Predicts Performance. *Review of Financial Studies, 22*, 3329-3365.

De Long, B., Schleifer, A., Summer, L., & Waldmann, R. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 703-38.

de Planta, R. (2017, 05 17). *The hidden dangers of passive investing.* Retrieved 06 12, 2018, from Financial Times: https://www.ft.com/content/15dd3552-3fad-11e7-82b6-896b95f30f58

DeBondt, W., & Thaler, R. (1985). Does the Stock Market Overreact? . *The Journal of Finance*, 793-805.

Doidge, C., Karolyi, A., & Stulz, R. (2015). The U.S. Listing Gap. *National Bureau of Economic Research*, 464-87.

Economides, N. (1995). How to enhance market liquidity. *Financial Networks*, 2-12.

Edwards, W. (1968). Conservatism in Human Information Processing. In D. Kahneman, & A. Tversky, *Judgement Under Uncertainty: Heuristics and Biases.* New York: Cambride University Press.

Evans, R., & Fahlenbrach, R. (2012). Institutional investors and mutual fund governance: Evidence from retail-institutional fund twins. *Review of Financial Studies*, 3530-71.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance*, 383-417.

Fama, E. (1976). *Foundations of Finance.* New York: Basic Books, Inc. Publishers.

Fama, E. (1991). Efficient Capital Markets: II. *Journal of Finance*, 1575-1617.

Fama, E., & French, K. (1988). Permanent and Temporary Components of Stock Prices. *Journal of Political Economy*, 246-73.

Fama, E., & French, K. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *The Journal of Financial Economics*, 3-56.

Fama, E., & French, K. (2004). The Capital Asset Pricing Model: Theory and Evidence. *The Journal of Economic Perspectives*, 25-45.

Fama, E., & French, K. (2010). Luck versus Skill in the Cross-Section of Mutual Fund Returns. *Journal of Finance*, 1915-47.

Fama, E., & Jensen, M. (1983). Separation of ownership and control. *Journal of Law and Economics, 26*(2), 301-325.

Fama, E., & MacBeth, J. (1973). Risk, Return, and Equilibrium: Empirical Tests. *The Journal of Political Economy*, 607-36.

Fama, E., Fisher, L., Jensen, M., & Roll, R. (1969). The Adjustment of Stock Prices to New Information. *International Economic Review*, 1-21.

Fichtner, J., Heemskerk, E., & Garcia-Bernando, J. (2017). Hidden power of the Big Three? Passive index funds, re-concentration of corporate ownership, and new financial risk. *Business and Politics, 19*(2), 298-326.

Financial Times. (2018, 05 31). *Definition of Smart Beta*. Retrieved from ft.com/lexicon: http://lexicon.ft.com/Term?term=smart-beta

Frazzini, A., & Pedersen, L. H. (2014). Betting Against Beta. *Journal of Financial Economics*, 1-25.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 25-42.

French, K. (2008, August). Presidential Address: The Cost of Active Investing. *The Journal of Finance, 63*(4), 1537-1573.

Fung, W., Hsieh, D., & Ramadorai, T. (2008). Hedge funds: Performance, risk, and capital formation. *Journal of Finance*, 1777-1803.

Garleanu, N., & Pedersen, L. H. (2018). Efficiently Inefficient Markets for Assets and Asset Management. *Journal of Finance*, 1663-1712.

Grossman, S. J., & Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *American Economic Review*, 393-408.

Grossman, S., & Miller, M. (1988). Liquidity and market structure. *Journal of Finance*, 617-33.

Hegde, S., & McDermott, J. (2003). The liquidity effects of revisions to the S&P 500 index: an empirical analysis. *Journal of Financial Markets*, 413-59.

Heston, S., & Sadka, R. (2008). Seasonality in the Cross-Section of Stock Returns. *Journal of Financial Econmomics*, 418-45.

Hill, C., Griffiths, W., & Lim, G. (2011). *Principles of Econometrics.* Hoboken: John Wiley and Sons, Inc.

Hou, K., Xue, C., & Zhang, L. (2018). Replicating Anomalies. *Review of Financial Studies, Working Paper*.

Israeli, D., Lee, C., & Sridharan, S. (2017). Is there a Dark Side to Exchange Traded Funds (ETFs)? An Information Perspective. *Review of Accounting Studies*, 1048-83.

Jagannathan, R., Malakhov, A., & Novikov, D. (2010). Do hot hands exist among hedge fund managers? An empirical evaluation. *Journal of Finance*, 217-55.

Jasen, G. (2002, 04 18). *Journal's Dartboard Retires After 14 Years of Stock Picks*. Retrieved from wsj.com: https://www.wsj.com/articles/SB10190809017144480

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Market Efficiency. *The Journal of Finance*, 65-91.

Kahneman, D., & Tversky, A. (1974). Judgement Under Uncertainty: Heuristics and Biases. *Science*, 1124-31.

Kahnemann, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decisions Under Risk. *Econometrica*, 263-91.

Kaplan, S., & Schoar, A. (2005). Private equity performance: Returns, persistence, and capital flows. *Journal of Finance*, 1791-1823.

Kawa, L. (2016, 08 23). *Bernstein: Passive Investing Is Worse for Society Than Marxism.* Retrieved 06 15, 2018, from Bloomberg.com: https://www.bloomberg.com/news/articles/2016-08-23/bernstein-passive-investing-is-worse-for-society-than-marxism

Kliger, D., & Gurevich, G. (2014). *Event Studies for Financial Research - A Comprehensive Guide.* New York: Palgrave MacMillian US.

Korteweg, A., & Sorensen, M. (2017). Skill and luck in private equity performance. *Journal of Financial Economics*, 535-62.

Kosowski, R., Naik, N., & Teo, M. (2007). Do Hedge Funds Deliver Alpha? A Bayesian and Bootstrap Analysis. *Journal of Financial Economics*, 229-64.

Kosowski, R., Timmermann, A., Wermers, R., & White, H. (2006). Can Mutual Fund "Stars" Really Pick Stock? New Evidence From a Bootstrap Analysis. *Journal of Finance*, 2551-95.

Kothari, S., & Warner, J. (2007). Econometrics of Event Studies. In *Handbook of Corporate Finance* (pp. 1-33). Elsevier.

Kumar, S., & Goyal, N. (2015). Behavioural Biases in Investment Decisions - A Systematic Literature Review. *Qualitative Research in Financial Markets*, 88-108.

Leary, E. (2018, 02 01). *The Perils of Index Investing*. Retrieved from kiplinger.com: https://www.kiplinger.com/article/investing/T041-C009-S002-the-perils-of-investing-in-index-funds.html

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments. *The Review of the Economics and Statistics*, 13-37.

Liu, H., & Wang, Y. (2018). Index Investing and Price Discovery. *White Paper*, NA.

Mackay, C. (1841). *Extraordinary Popular Delusions and the Madness of Crowds.* Glasgow: Harriman House.

Malkiel, B. (1973). *A Random Walk Down Wall Street.* Princeton: W. W. Norton & Company.

Mastrogiorgio, A., & Petraca, E. (2014). 2014. *Mind and Society*, 135-48.

McCahery, J., Sautner, Z., & Starks, L. (2016). Behind the Scenes: The Corporate Governance Preferences of Institutional Investors. *The Journal of Finance, 71*(6), 2905-2932.

Moody's Investors Service. (2017). *Passive Market Share to Overtake Active in the US No Later than 2024.* New York: Financial Institutions.

Naseer, M., & Tariq, Y. (2015). The Efficient Market Hypothesis: A Critical Review of the Literature. *IUP Journal of Financial Risk Management*, 48-63.

Novick, B. (2017, 01 08). *How Index Funds Democratize Investing*. Retrieved from wsj.com: https://www.wsj.com/articles/how-index-funds-democratize-investing-1483914571

Novy-Marx, R. (2015). Fundamentally, Momentum is Fundamental Momentum. *Working Paper, National Bureau of Economic Research*.

O'Hara, M. (2003). Presidential adress: Liquidity and price discovery. *Journal of Finance*, 1335-54.

Patel, N., & Welch, I. (2017). Extended Stock Returns in Reponse to S&P 500 Index Changes. *The Review of Asset Pricing Studies*, 172-208.

Pedersen, L. (2018). Sharpening the Arithmetic of Active Management. *Financial Analysts Journal*, 21-37.

Pedersen, L. H. (2015). *Efficiently Inefficient: How Smart Money Invests and Market Prices are Determined.* New York: Princeton University Press.

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

Perry, M. (2018, 01 04). *Warren Buffett Wins $1m Bet Made A Decade Ago That The S&P 500 Stock Index Would Outperform Hedge Funds, And It Wasn't Even Close*. Retrieved from seekingalpha.com: https://seekingalpha.com/article/4135244-warren-buffett-wins-1m-bet-made-decade-ago-s-and-p-500-stock-index-outperform-hedge-funds

Petajisto, A. (2010). The Index Premium and the Hidden Costs for Index Funds. *Journal of Empirical Finance*, 271-88.

Pettit, R., & Venkatesh, P. (1995). Insider Trading and Long-Run Return Performance. *Financial Management*.

Pischke, J.-S. (2005, 10). Empirical Methods in Applied Economics - Lecture Notes.

Roll, R. (1977). A Critique of the Asset Pricing Theory's Tests Part I: On Past and Potential Testability of the Theory. *The Journal of Financial Economics*, 129-76.

Rosenbaum, E. (2019, 01 16). *Fidelity one-ups Vanguard, Schwab and iShares, becoming first company to offer a no-fee index fund*. Retrieved from CNBC.com: https://www.cnbc.com/2018/08/01/fidelity-one-ups-vanguard-first-company-to-offer-no-fee-index-fund.html

S&P Dow Jones Indices. (2019). *S&P U.S. Indices: Methodology.*

Samuelson, P. (1974). Challenge to Judgment. *The Journal of Portfolio Management*, 17-19.

Schumpeter, J. (1942). *Capitalism, Socialism, and Democracy.* New York: Harper & Bros.

Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance*, 425-42.

Sharpe, W. (1991). The Arithmetic of Active Management. *Financial Analysts Journal*, 7-9.

Shiller, R. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review*, 421-36.

Simon, H. (1982). Models of Bounded Rationality. *MIT University Press*.

Stock, J., & Watson, M. (2015). *Introduction to Econometrics.* Edinburgh: Pearson Education Limited.

Sullivan, R., & Xiong, J. (2012). How Index Trading Increases Market Vulnerability. *Financial Analysts Journal*, 70-84.

Thaler, R. (1985). Mental Accounting and Consumer Choice. *Marketing Science*, 199-214.

The Wall Street Journal. (2016, 10 17). *The Passivist*. Retrieved from wsj.com: https://graphics.wsj.com/passivists/

Titan, A. (2015). The Efficient Market Hypothesis: review of specialized literature and empirical research. *Procedia Economics and Finance*, 442-449.

Vanguard. (2018). *2018 Investment Stewardship Report.* Valley Forge.

Vanguard. (2019, 01 16). *A remarkable story*. Retrieved from about.vanguard.com: https://about.vanguard.com/who-we-are/a-remarkable-history/

Wurgler, J. (2011). The Economic Consequences of Index Investing. In W. Allen, R. Khurana, J. Lorsch, & G. Rosenfeld, *Challenges to Business in the Twenty - First Century: The Way Forward* (pp. 1-15).

Zweig, J. (2016, 08 31). *Birth of the Index Mutual Fund: 'Bogle's Folly' Turns 40*. Retrieved from Wall Street Journal: https://blogs.wsj.com/moneybeat/2016/08/31/birth-of-the-index-mutual-fund-bogles-folly-turns-40/

14 Appendix

14.1 Appendix I – S&P 500 Inclusion Criteria Overview

NB. These guidelines were enacted as of 20th February 2019, i.e. after end point of our data collection period. Notice that these guidelines are updated regularly. (S&P Dow Jones Indices, 2019)

Criteria	Description	Adherence	Comments
Domicile	Only common stocks of US companies are eligible. Defined as: • Files 10-K reports • Has a plurality of assets or revenue in the US • Primary listing must be on a US exchange • However, the committee can include any firm if they deem it American	Not required	There could be some informational discrepancies between firms considered American and firms which are not.
Primary listing	Must have its primary listing on: • NYSE • NYSE Arca • NYSE American • NASDAQ Global Select Market • NASDAQ Select Market • NASDAQ Cappital Market • Investors Exchange (IEX) • CBOE BZX • CBOE BYX • CBOE EDGA • CBOE EDGX	Required	This is also the sources from which we gather data. Furthermore, there must be more information and liquidity on these exchanges.
Multiple share classes	The company cannot have multiple share classes of common stock. However, firms are grandfathered in. I.e. if they are included and issue a new share class, they may remain in the index.	Not required	To the best of our knowledge, there is no informational disadvantage to having multiple share classes.
Market cap	Companies must have a market capitalisation of USD 8.2bn or more to be included in the S&P 500.	Required	Large informational effects of being a large company.
Liquidity	• The ratio of annual dollar value traded to float-adjusted market capitalisation must be at least 1.00. • The share must be traded a minimum of 250,000 times per month in each of the six months leading up to the evaluation.	Required	Liquidity is a prerequisite for efficiency.
IWF	Investable weight factor. The number of shares floated relative to the total number of outstanding shares. Must be above 0.5. IWF is defined as: $availabe\ floated\ shares\ /\ total\ shares\ outstanding$	Not required	To the best of our knowledge, there is no informational disadvantage to having a low IWF.
Financial viability	The sum of the earnings of the four most recent quarters must be positive – as should the most recent quarter.	Not required	To the best of our knowledge, there is no informational disadvantage to having worse financial results.
IPOs	Should be traded at least 12 months prior to being considered for any index. This condition is waived for spin-offs of currently included companies.	Not required	There are many IPO effects.

14.2 Appendix II – Threshold Levels

Table 24 – Threshold Levels for Group Inclusion

$m	Group 2 min. MV	Group 3 min. MV	Group 2 min DL
1989	226	23	43,364
1990	63	6	14,213
1991	157	16	27,313
1992	128	13	22,237
1993	173	17	21,155
1994	147	15	14,496
1995	198	20	55,608
1996	352	35	48,734
1997	427	43	91,188
1998	490	49	159,988
1999	153	15	94,372
2000	268	27	163,831
2001	644	64	185,229
2002	215	22	169,731
2003	398	40	168,555
2004	550	55	333,953
2005	481	48	410,949
2006	775	78	596,690
2007	801	80	832,337
2008	547	55	938,311
2009	244	24	502,539
2010	1,212	121	940,037
2011	824	82	822,978
2012	1,193	119	533,436
2013	2,046	205	651,296
2014	3,353	335	888,394
2015	1,598	160	1,398,549
2016	1,789	179	717,645
2017	2,316	232	380,372
2018	2,281	228	497,377

14.3 Appendix III – Additional Group Comparison

Table 25 Summary Statistics for Merged Data before News Definition (1989-2018) | Additional

	Market Value			Dollar Liquidity			CAR3		
	Group 1	Group 2.1	Top-1500	Group 1	Group 2.1	Top-1500	Group 1	Group 2.1	Top-1500
Mean ($m)	18,225	29,396	21,193	10,559,770	10,634,620	8,047,882	-0.08 %	0.16 %	0.19 %
1st Quart. ($m)	5,042	7,638	4,819	2,708,502	1,677,533	1,207,916	-3.04 %	-2.25 %	-2.28 %
Median ($m)	9,586	13,435	8,272	6,141,937	4,665,331	3,267,244	0.12 %	0.20 %	0.22 %
3rd Quart. ($m)	18,214	28,819	18,734	11,451,680	11,103,950	7,928,654	3.33 %	2.75 %	2.84 %
σ ($m)	34,506	48,788	41,499	18,924,560	23,629,280	19,754,150	6.95 %	5.22 %	5.57 %
n	27,076	30,000	30,000	27,076	30,000	30,000	27,076	30,000	30,000
Proportion	4%	5%	5%	4%	5%	5%	4%	5%	5%

14.4 Appendix IV – Additional Group Comparison

Table 26 Summary Statistics for Merged Data Before News Definition

	Market Value (Group 1)				Book Value (Group 1)			
	Total	Bad	No	Good	Total	Bad	No	Good
Mean ($m)	18,225	13,644	18,801	14,498	6,940	5,266	7,208	4,821
Median ($m)	9,586	6,771	9,946	7,617	3,490	2,435	3,722	2,252
σ ($m)	34,506	27,998	35,135	30,415	12,054	9,928	12,321	9,584
n	27,076	1,632	23,780	1,664	20,442	1,240	17,914	1,288
Mean-to-mean	1.00	0.75	1.03	0.80	1.00	0.75	1.04	0.69

14.5 Appendix V – Code for Data Preparation: Part I

```
# IN PART I THE DATA IS LOADED IN FROM CRSP AND COMPUSTAT AND THEN CLEANED
```

Step 1

```
# Import pandas and numpy libraries
import pandas as pd
import numpy as np
```

Step 2

```
#Read in CRSP, check the data, shape and null values
df_crsp = pd.read_csv('CRSP_Full.csv', dtype={'PERMNO':object, 'date':object,
'EXCHCD': object, 'PRC':object, 'VOL':object, 'RET':object, 'SHROUT':object,
'vwretd':object})

df_crsp

df_crsp.shape
#(52153354, 8)

df_crsp.isnull().sum()
```

Step 3

```
# Drop the null values
dfCrsp = df_crsp.dropna(how='any')

dfCrsp.shape
#(52108189, 8)
```

Step 4

```
#Read in Compustat check the data, shape and null values
df_compustat = pd.read_csv('Compustat_Full.csv', dtype={'GVKEY':object,
'LPERMNO':object, 'datadate':object, 'fyearq':object, 'fqtr':object,
'indfmt':object, 'consol':object,'popsrc':object, 'datafmt':object,
'tic':object, 'datacqtr':object, 'datafqtr':object, 'rdq':object, 'ibq':object,
'pstkq':object, 'saleq':object, 'seqq':object, 'txditcq':object,
'costat':object, 'loc':object, 'sic':object})

df_compustat

df_compustat.shape
#(863998, 23)

df_compustat.isnull().sum()
```

Step 5

```
# Drop the values that have no rdq
dfComp = df_compustat.dropna(subset = ['rdq'], how='any')

dfComp.isnull().sum()

dfComp.shape
#(672203,23)
```

Step 6

```
#Save cleaned data to CSV
dfCrsp.to_csv('dfCrsp_Cleaned.csv')

#Save cleaned data to CSV
dfComp.to_csv('dfComp_Cleaned.csv')
```

14.6 Appendix VI – Code for Data Preparation: Part II

```
# IN PART II THE CRSP AND COMPUSTAT DATA IS MERGED
# VARIABLES THAT ARE NEEDED FOR THE ANALYSIS ARE CREATED AND MODIFIED
```

Step 1

```
# Import pandas and numpy libraries
import pandas as pd
import numpy as np
```

Step 2

```
# Read in Comp
dfComp = pd.read_csv('dfComp_Cleaned.csv', dtype={'GVKEY':object,
'LPERMNO':object, 'datadate':object, 'fyearq':object, 'fqtr':object,
'indfmt':object, 'consol':object,'popsrc':object, 'datafmt':object,
'tic':object, 'datacqtr':object, 'datafqtr':object, 'rdq':object, 'ibq':object,
'pstkq':object, 'saleq':object, 'seqq':object, 'txditcq':object,
'costat':object, 'loc':object, 'sic':object})

dfComp

# Read in Crsp
dfCrsp = pd.read_csv('dfCrsp_Cleaned.csv', dtype={'PERMNO':object,
'date':object, 'EXCHCD': object, 'PRC':object, 'VOL':object, 'RET':object,
'SHROUT':object, 'vwretd':object})

dfCrsp
```

Step 3

```
# Double check that there are no nulls
dfCrsp.isnull().sum()

# Double check that there are no unexpected nulls
dfComp.isnull().sum()
```

Step 4

```
# Create PERMNODATE Variable for CRSP dataframe
dfCrsp["PERMNODATE"] = dfCrsp["PERMNO"].map(str) + dfCrsp["date"].map(str)

dfCrsp.head()


# Create PERMNODATE Variable for COMP dataframe
dfComp["PERMNODATE"] = dfComp["LPERMNO"] + dfComp["rdq"]

dfComp.head()
```

Step 5

```
#Merge data frames on PERMNODATE variable and create new data frame
dfFinal = pd.merge(dfCrsp, dfComp, how='left', on='PERMNODATE')

dfFinal.shape
# (52109068, 34)

# Save merged data frame as a CSV file
dfFinal.to_csv('dfFinalMerged.csv')
```

Step 6

```
# Close the kernel down to clear the RAM and read in the merged data frame
# Only read in the most necessary variables

dfMerged = pd.read_csv('dfFinalMerged.csv', usecols=['PERMNO', 'date', 'EXCHCD',
'PRC', 'VOL', 'RET', 'SHROUT', 'vwretd', 'PERMNODATE', 'GVKEY', 'datadate',
'fyearq', 'fqtr', 'tic', 'rdq', 'ibq', 'pstkq', 'saleq', 'seqq', 'txditcq',
'costat', 'loc', 'sic'], dtype={'PERMNO':object, 'date':object, 'EXCHCD':object,
'PRC':object, 'VOL':object, 'RET':object, 'SHROUT':object, 'vwretd':object,
'PERMNODATE':object, 'GVKEY':object, 'datadate':object, 'fyearq':object,
'fqtr':object, 'tic':object, 'rdq':object, 'ibq':object, 'pstkq':object,
'saleq':object, 'seqq':object, 'txditcq':object, 'costat':object, 'loc':object,
'sic':object})

dfMerged
dfMerged.shape
#(52109068,23)
```

Step 7

```
# Convert VOL into a numeric variable
dfMerged['VOL'] = dfMerged['VOL'].apply(pd.to_numeric, errors='coerce')

# Convert PRC and SHROUT into numeric variables
dfMerged['PRC'] = dfMerged['PRC'].apply(pd.to_numeric, errors='coerce')
dfMerged['SHROUT'] = dfMerged['SHROUT'].apply(pd.to_numeric, errors='coerce')

# Convert RETX and vwretx into numeric variables
dfMerged['RET'] = dfMerged['RET'].apply(pd.to_numeric, errors='coerce')
dfMerged['vwretd'] = dfMerged['vwretd'].apply(pd.to_numeric, errors='coerce')
```

Step 8

```
#Create a variable with the absolute value of price
dfMerged['Abs_Price'] = dfMerged['PRC'].abs()
```

Step 9

```
# Calculate Abnormal Return (aRET = RET - vwretd)
dfMerged['aRET'] = dfMerged.RET - dfMerged.vwretd
```

Step 10

```
# Logarithm of aRET, log(1+x)
dfMerged['Log_aRET'] = dfMerged['aRET'].apply(np.log1p)
```

Step 11

```
# Calculate the DollarVolume
dfMerged['DollarVolume'] = dfMerged['Abs_Price'] * dfMerged['VOL']
```

Step 12

```
# What number has the column of VOL and Log_aRET?
list(dfMerged)
```

Step 13

```
# Create the CAR3 variable
# Remember that iloc takes y-1 for columns
def ret_func(row):
    return dfMerged.iloc[row.name-1:row.name+2,24].sum()

dfMerged['CAR3'] = dfMerged.apply(lambda x: ret_func(x) if pd.notnull(x.rdq)
else np.nan,
                                  axis=1)
```

Step 14

```
# Control that CAR3 is calculated correctly and based on Log_aRET
# Save back up in case kernel dies from memory being full
dfMerged.to_csv('CAR3.csv')
```

Step 15

```
# Create the Dollar_Liquidity variable (total 66 days)

def dollar_func(row):
    return dfMerged.iloc[row.name-64:row.name+2,27].sum()

dfMerged['Dollar_Liquidity'] = dfMerged.apply(lambda x: dollar_func(x) if
pd.notnull(x.rdq) else np.nan,
                                  axis=1)
```

Step 16

```
# Create data frame with the rows that have CAR3
dfOnly_CAR3 = dfMerged.loc[pd.notnull(dfMerged.CAR3)]

dfOnly_CAR3.shape
#(658118,29)
```

14.7 Appendix VII – Code for Data Preparation: Part III

```
# IN PART III THE LAST VARIABLES ARE CREATED
# THE GROUPS ARE CREATED AND THEN FURTHER SPLIT UP INTO NEWS GROUPS
# TEXT-FILES ARE SAVED WHICH ARE THE BASIS FOR THE EVENT STUDIES ARE GENERATED
```

Step 1

```
# Import pandas and numpy libraries
import pandas as pd
import numpy as np
```

Step 2

```
# Read in the data
dfData = pd.read_csv('dfOnly_CAR3.csv')

dfData = dfData.drop(['Unnamed: 0'], axis =1)

dfData
```

Step 3

```
# Check the shape of the data and generate summary statistics

dfData.shape
#(658118,32)

dfData.describe()
```

Step 4

```
# Create MarketValue Variable
dfData['MarketValue'] = dfData['Abs_Price'] * dfData['SHROUT']

# Create BookValue

dfData['BookValue'] = dfData['seqq'] + dfData['txditcq'] - dfData['pstkq']


# Create Book-to-Market
dfData['BookMarket'] = dfData['MarketValue']/dfData['BookValue']

dfData.head()
```

Step 5

```
# LOAD IN THE S&P500 AND FLAG THE DATA

# Read in S&P500
dfSP500 = pd.read_csv('MyS&P500.csv')

dfSP500 = dfSP500.drop(['Unnamed: 0'], axis =1)

dfSP500
```

Step 6

```
# Flag if the permno was in S&P500 at that specific date
dfData['SP500flag'] = np.piecewise(np.zeros(dfData.count()[0]),
                                  [(dfData['PERMNO'].values == id) &
                                  (dfData['rdq'].values >= start) &
                                  (dfData['rdq'].values <= end) for id, start,
                                  end in zip(dfSP500['permno'].values,
                                  dfSP500['start'].values,
                                  dfSP500['ending'].values)],
                                  dfSP500['permno'].values).astype(int)

dfData
```

Step 7

```
# Turn the zeros into nulls
dfData['SP500flag'] = dfData['SP500flag'].replace(0, np.nan)

dfData

dfData.isnull().sum()
```

Step 8

```
# Generate variable that is only the year of the rdq date
import datetime
def to_year(row):
    return datetime.datetime.strptime(str(row.rdq),'%Y%m%d').year

dfData['year'] = dfData.apply(lambda x: to_year(x), axis=1)
```

Step 9

```
# CREATE GROUP 1
# Group 1 which is the ones that are flagged under the SP500 variable
dfGroup1 = dfData.loc[pd.notnull(dfData.SP500flag)]

dfGroup1
```

Step 10

```
# CREATE GROUP 2 AND GROUP 3

# Find the thresholds for MarketValue and Dollar_Liquidity for each year
# The threshold level is equal to the smallest of the S&P500 firms for each year
dfThreshold =
dfGroup1[['year','MarketValue','Dollar_Liquidity']].groupby(by=['year']).min().r
eset_index()

dfDataSPFlagNull = dfData.loc[pd.isnull(dfData['SP500flag'])]

# Create Group 2 and Group 3 based on the thresholds
# Group 2 has both MarketValue & Dollar_Liquidity larger than the threshold
# Group 2 fails to be above the threshold on at least one of the two

dfGroup2 = pd.DataFrame(columns=dfDataSPFlagNull.columns)
dfGroup3 = pd.DataFrame(columns=dfDataSPFlagNull.columns)

for x in dfDataSPFlagNull.year.unique():
    ThresholdMV = dfThreshold.MarketValue.loc[dfThreshold.year == x].values[0]
    ThresholdLiq = dfThreshold.Dollar_Liquidity.loc[dfThreshold.year ==
x].values[0]
    # Group 2
    tempdfGroup2 = dfDataSPFlagNull.loc[(dfDataSPFlagNull.year == x) &
                                        ((dfDataSPFlagNull.MarketValue >=
ThresholdMV) &
                                        (dfDataSPFlagNull.Dollar_Liquidity >=
ThresholdLiq))]
    dfGroup2 = pd.concat([dfGroup2, tempdfGroup2]).reset_index(drop=True)

    # Group 3
    tempdfGroup3 = dfDataSPFlagNull.loc[(dfDataSPFlagNull.year == x) &
                                        ((dfDataSPFlagNull.MarketValue <
ThresholdMV) |
                                        (dfDataSPFlagNull.Dollar_Liquidity <
ThresholdLiq))]
    dfGroup3 = pd.concat([dfGroup3, tempdfGroup3]).reset_index(drop=True)
```

Step 11

```
# Save down the threshold levels to a new data frame
dfThreshold_10 = dfThreshold

dfThreshold_10
```

Step 12

```
# Make sure that the split was correct by checking that the groups
# together have the same length as the full dataset

print(len(dfGroup1))
print(len(dfGroup2))
print(len(dfGroup3))
print(len(dfData))


print(len(dfGroup1)+len(dfGroup2)+len(dfGroup3))
```

Step 13

```
# Generate summary statistics for all groups
dfGroup1.describe()

dfGroup2.describe()

dfGroup3.describe()
```

Step 14

```
# Create Group 2.1 which is the top 1000 rows per year
# with the highest MarketValue
# NOTE that in the code Group 2.1 in referred to as Group2_TopMV

dfGroup2_TopMV = dfGroup2.groupby('year').apply(lambda x:
x.nlargest(1000,['MarketValue'])).reset_index(drop=True)

print(len(dfGroup2_TopMV))

# Generate summary statistics for Group 2.1
dfGroup2_TopMV.describe()
```

Step 15

```
# Drop the lowest values in Group 3 based on the thresholds we got from
# the firms with the lowest market value and dollar liquidity per year

# Current size of Group 3
dfGroup3.shape
#(471292,35)

# Take 10% of the saved thresholds for each year
dfThreshold_10.iloc[:,1:] = dfThreshold_10.iloc[:,1:] * 0.1

dfThreshold_10

# Generate the Group3_Final data frame based on the 10% thresholds

dfGroup3_Final = pd.DataFrame(columns=dfGroup3.columns)

for x in dfGroup3.year.unique():
    Threshold_MarketValue = dfThreshold_10.MarketValue.loc[dfThreshold_10.year
== x].values[0]

    # Group 3
    temp_dfGroup3 = dfGroup3.loc[(dfGroup3.year == x) &
                                 (dfGroup3.MarketValue >=
Threshold_MarketValue)]
    dfGroup3_Final = pd.concat([dfGroup3_Final,
temp_dfGroup3]).reset_index(drop=True)

# Final size of Group 3
dfGroup3_Final.shape
#(324716,35)

# Generate summary statistics for the final version of Group 3
dfGroup3_Final.describe()
```

Step 16

```
# SORT EACH GROUP BASED ON POSTIVE NEWS, NEUTRAL/NO NEWS, AND NEGATIVE NEWS

# Find boundaries for news from full dataset
Lower = dfData['CAR3'].quantile(0.10) #Lower 10%
Upper = dfData['CAR3'].quantile(0.90) #Upper 10%
print(Lower)
print(Upper)

# Lower = -0.0960135612921
# Upper = 0.088400776781
```

Step 17

```
# Split full dataset into news groups

# Neutral Earnings Surprises
dfNeutral_Data = dfData[(dfData.CAR3 >= Lower) & (dfData.CAR3 <= Upper)]

# Positive Earnings Surprises
dfPositive_Data = dfData[dfData.CAR3 > Upper]

# Negative Earnings Surprises
dfNegative_Data = dfData[dfData.CAR3 < Lower]

# Control the lengths of the new data frames
print(len(dfNeutral_Data))
print(len(dfPositive_Data))
print(len(dfNegative_Data))
print(len(dfData))

print(len(dfNeutral_Data)+len(dfPositive_Data)+len(dfNegative_Data))

# Generate summary statistics for the news groups
dfNeutral_Data.describe()


dfPositive_Data.describe()


dfNegative_Data.describe()
```

Step 18

```
# Split Group 1 into news groups

# Neutral Earnings Surprises
dfNeutral_Group1 = dfGroup1[(dfGroup1.CAR3 >= Lower) & (dfGroup1.CAR3 <= Upper)]

# Positive Earnings Surprises
dfPositive_Group1 = dfGroup1[dfGroup1.CAR3 > Upper]

# Negative Earnings Surprises
dfNegative_Group1 = dfGroup1[dfGroup1.CAR3 < Lower]

# Control the lengths of the new data frames
print(len(dfNeutral_Group1))
print(len(dfPositive_Group1))
print(len(dfNegative_Group1))
print(len(dfGroup1))

print(len(dfNeutral_Group1)+len(dfPositive_Group1)+len(dfNegative_Group1))
```

```
# Generate summary statistics for the news groups
dfNeutral_Group1.describe()

dfPositive_Group1.describe()

dfNegative_Group1.describe()
```

Step 19

```
# Split Group 2 and Group 2.1 into news groups

# Neutral Earnings Surprises
dfNeutral_Group2 = dfGroup2[(dfGroup2.CAR3 >= Lower) & (dfGroup2.CAR3 <= Upper)]
dfNeutral_Group2_TopMV = dfGroup2_TopMV[(dfGroup2_TopMV.CAR3 >= Lower) &
(dfGroup2_TopMV.CAR3 <= Upper)]

# Positive Earnings Surprises
dfPositive_Group2 = dfGroup2[dfGroup2.CAR3 > Upper]
dfPositive_Group2_TopMV = dfGroup2_TopMV[dfGroup2_TopMV.CAR3 > Upper]

# Negative Earnings Surprises
dfNegative_Group2 = dfGroup2[dfGroup2.CAR3 < Lower]
dfNegative_Group2_TopMV = dfGroup2_TopMV[dfGroup2_TopMV.CAR3 < Lower]

# Control the lengths of the news groups for Group 2
print(len(dfNeutral_Group2))
print(len(dfPositive_Group2))
print(len(dfNegative_Group2))
print(len(dfGroup2))
print(len(dfNeutral_Group2)+len(dfPositive_Group2)+len(dfNegative_Group2))

# Control the lengths of the news groups for Group 2
print(len(dfNeutral_Group2_TopMV))
print(len(dfPositive_Group2_TopMV))
print(len(dfNegative_Group2_TopMV))
print(len(dfGroup2_TopMV))
print(len(dfNeutral_Group2_TopMV)+len(dfPositive_Group2_TopMV)+len(dfNegative_Gr
oup2_TopMV))

# Generate summary statistics for news groups of Group 2
dfNeutral_Group2.describe()

dfPositive_Group2.describe()

dfNegative_Group2.describe()
```

```
# Generate summary statistics for news groups of Group 2.1
dfNeutral_Group2_TopMV.describe()


dfPositive_Group2_TopMV.describe()


dfNegative_Group2_TopMV.describe()
```

Step 20

```
# Split Group 3 into news groups

# Neutral Earnings Suprises
dfNeutral_Group3 = dfGroup3_Final[(dfGroup3_Final.CAR3 >= Lower) &
(dfGroup3_Final.CAR3 <= Upper)]

# Positive Earnings Suprises
dfPositive_Group3 = dfGroup3_Final[dfGroup3_Final.CAR3 > Upper]

# Negative Earnings Suprises
dfNegative_Group3 = dfGroup3_Final[dfGroup3_Final.CAR3 < Lower]

# Control the lengths of the new data frames
print(len(dfNeutral_Group3))
print(len(dfPositive_Group3))
print(len(dfNegative_Group3))
print(len(dfGroup3_Final))

print(len(dfNeutral_Group3)+len(dfPositive_Group3)+len(dfNegative_Group3))

# Generate summary statistics for the news groups
dfNeutral_Group3.describe()


dfPositive_Group3.describe()


dfNegative_Group3.describe()
```

Step 21

```
# Save CSV files with PERMNO & RDQ
# Full time period

# Text files for full dataset
dfESIPUT_Neutral_Data = dfNeutral_Data[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Data.to_csv('dfESIPUT_Neutral_Data.txt', header=False,
index=False, sep='  ')

dfESIPUT_Positive_Data = dfPositive_Data[['PERMNO', 'rdq']]
```

```
dfESIPUT_Positive_Data.to_csv('dfESIPUT_Positive_Data.txt', header=False,
index=False, sep='	')

dfESIPUT_Negative_Data = dfNegative_Data[['PERMNO', 'rdq']]
dfESIPUT_Negative_Data.to_csv('dfESIPUT_Negative_Data.txt', header=False,
index=False, sep='	')


# Text files for Group 1
dfESIPUT_Neutral_Group1 = dfNeutral_Group1[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group1.to_csv('ESINPUT_Neutral_Group1.txt', header=False,
index=False, sep='	')

dfESIPUT_Positive_Group1 = dfPositive_Group1[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group1.to_csv('ESINPUT_Positive_Group1.txt', header=False,
index=False, sep=' ')

dfESIPUT_Negative_Group1 = dfNegative_Group1[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group1.to_csv('ESINPUT_Negative_Group1.txt', header=False,
index=False, sep=' ')

# Text files for Group 2
dfESIPUT_Neutral_Group2 = dfNeutral_Group2[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group2.to_csv('ESINPUT_Neutral_Group2.txt', header=False,
index=False, sep='	')

dfESIPUT_Positive_Group2 = dfPositive_Group2[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group2.to_csv('ESINPUT_Positive_Group2.txt', header=False,
index=False, sep=' ')

dfESIPUT_Negative_Group2 = dfNegative_Group2[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group2.to_csv('ESINPUT_Negative_Group2.txt', header=False,
index=False, sep=' ')

# Text files for Group 2.1
dfESIPUT_Neutral_Group2_TopMV = dfNeutral_Group2_TopMV[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group2_TopMV.to_csv('ESINPUT_Neutral_Group2_TopMV.txt',
header=False, index=False, sep='	')

dfESIPUT_Positive_Group2_TopMV = dfPositive_Group2_TopMV[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group2_TopMV.to_csv('ESINPUT_Positive_Group2_TopMV.txt',
header=False, index=False, sep=' ')

dfESIPUT_Negative_Group2_TopMV = dfNegative_Group2_TopMV[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group2_TopMV.to_csv('ESINPUT_Negative_Group2_TopMV.txt',
header=False, index=False, sep=' ')
```

```python
# Text files for Group 3
dfESIPUT_Neutral_Group3 = dfNeutral_Group3[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group3.to_csv('ESINPUT_Neutral_Group3.txt', header=False,
index=False, sep='   ')

dfESIPUT_Positive_Group3 = dfPositive_Group3[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group3.to_csv('ESINPUT_Positive_Group3.txt', header=False,
index=False, sep=' ')

dfESIPUT_Negative_Group3 = dfNegative_Group3[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group3.to_csv('ESINPUT_Negative_Group3.txt', header=False,
index=False, sep=' ')

# Control that the sizes are correct
print('Year 89-18')
print('Full Dataset')
print((dfESIPUT_Neutral_Data).shape)
print((dfESIPUT_Positive_Data).shape)
print((dfESIPUT_Negative_Data).shape)


print('Group 1')
print((dfESIPUT_Neutral_Group1).shape)
print((dfESIPUT_Positive_Group1).shape)
print((dfESIPUT_Negative_Group1).shape)

print('Group 2')
print((dfESIPUT_Neutral_Group2).shape)
print((dfESIPUT_Positive_Group2).shape)
print((dfESIPUT_Negative_Group2).shape)

print('Group 2 TopMV')
print((dfESIPUT_Neutral_Group2_TopMV).shape)
print((dfESIPUT_Positive_Group2_TopMV).shape)
print((dfESIPUT_Negative_Group2_TopMV).shape)

print('Group 3')
print((dfESIPUT_Neutral_Group3).shape)
print((dfESIPUT_Positive_Group3).shape)
print((dfESIPUT_Negative_Group3).shape)
```

Step 22

```
# A special date-format is needed for the time-split script to work
# From YYYYMMDD to YYYY-MM-DD
# It can be created through this loop.
import datetime
def to_date(row):
    return datetime.datetime.strptime(str(row.rdq),'%Y%m%d')

dfData['date_final'] = dfData.apply(lambda x: to_date(x), axis=1)
```

Step 23

```
#1st January 1989-31st December 1998

split_date_1999 = pd.datetime(1999,1,1)

# Text files for Group 1
dfNeutral_Group1_89to98 = dfNeutral_Group1.loc[dfNeutral_Group1['date_final'] <
split_date_1999]
dfESIPUT_Neutral_Group1_89to98 = dfNeutral_Group1_89to98[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group1_89to98.to_csv('dfESIPUT_Neutral_Group1_89to98.txt',
header=False, index=False, sep='    ')

dfPositive_Group1_89to98 = dfPositive_Group1.loc[dfPositive_Group1['date_final']
< split_date_1999]
dfESIPUT_Positive_Group1_89to98 = dfPositive_Group1_89to98[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group1_89to98.to_csv('dfESIPUT_Positive_Group1_89to98.txt',
header=False, index=False, sep='  ')

dfNegative_Group1_89to98 = dfNegative_Group1.loc[dfNegative_Group1['date_final']
< split_date_1999]
dfESIPUT_Negative_Group1_89to98 = dfNegative_Group1_89to98[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group1_89to98.to_csv('dfESIPUT_Negative_Group1_89to98.txt',
header=False, index=False, sep='  ')

# Text files for Group 2
dfNeutral_Group2_89to98 = dfNeutral_Group2.loc[dfNeutral_Group2['date_final'] <
split_date_1999]
dfESIPUT_Neutral_Group2_89to98 = dfNeutral_Group2_89to98[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group2_89to98.to_csv('dfESIPUT_Neutral_Group2_89to98.txt',
header=False, index=False, sep='    ')

dfPositive_Group2_89to98 = dfPositive_Group2.loc[dfPositive_Group2['date_final']
< split_date_1999]
dfESIPUT_Positive_Group2_89to98 = dfPositive_Group2_89to98[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group2_89to98.to_csv('dfESIPUT_Positive_Group2_89to98.txt',
header=False, index=False, sep='  ')
```

```
dfNegative_Group2_89to98 = dfNegative_Group2.loc[dfNegative_Group2['date_final']
< split_date_1999]
dfESIPUT_Negative_Group2_89to98 = dfNegative_Group2_89to98[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group2_89to98.to_csv('dfESIPUT_Negative_Group2_89to98.txt',
header=False, index=False, sep='	')


# Text files for Group 2.1
dfNeutral_Group2_TopMV_89to98 =
dfNeutral_Group2_TopMV.loc[dfNeutral_Group2_TopMV['date_final'] <
split_date_1999]
dfESIPUT_Neutral_Group2_TopMV_89to98 = dfNeutral_Group2_TopMV_89to98[['PERMNO',
'rdq']]
dfESIPUT_Neutral_Group2_TopMV_89to98.to_csv('dfESIPUT_Neutral_Group2_TopMV_89to9
8.txt', header=False, index=False, sep='	')

dfPositive_Group2_TopMV_89to98 =
dfPositive_Group2_TopMV.loc[dfPositive_Group2_TopMV['date_final'] <
split_date_1999]
dfESIPUT_Positive_Group2_TopMV_89to98 =
dfPositive_Group2_TopMV_89to98[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group2_TopMV_89to98.to_csv('dfESIPUT_Positive_Group2_TopMV_89t
o98.txt', header=False, index=False, sep='	')

dfNegative_Group2_TopMV_89to98 =
dfNegative_Group2_TopMV.loc[dfNegative_Group2_TopMV['date_final'] <
split_date_1999]
dfESIPUT_Negative_Group2_TopMV_89to98 =
dfNegative_Group2_TopMV_89to98[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group2_TopMV_89to98.to_csv('dfESIPUT_Negative_Group2_TopMV_89t
o98.txt', header=False, index=False, sep='	')

# Text files for Group 3

dfNeutral_Group3_89to98 = dfNeutral_Group3.loc[dfNeutral_Group3['date_final'] <
split_date_1999]
dfESIPUT_Neutral_Group3_89to98 = dfNeutral_Group3_89to98[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group3_89to98.to_csv('dfESIPUT_Neutral_Group3_89to98.txt',
header=False, index=False, sep='	')

dfPositive_Group3_89to98 = dfPositive_Group3.loc[dfPositive_Group3['date_final']
< split_date_1999]
dfESIPUT_Positive_Group3_89to98 = dfPositive_Group3_89to98[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group3_89to98.to_csv('dfESIPUT_Positive_Group3_89to98.txt',
header=False, index=False, sep='	')

dfNegative_Group3_89to98 = dfNegative_Group3.loc[dfNegative_Group3['date_final']
< split_date_1999]
dfESIPUT_Negative_Group3_89to98 = dfNegative_Group3_89to98[['PERMNO', 'rdq']]
```

```
dfESIPUT_Negative_Group3_89to98.to_csv('dfESIPUT_Negative_Group3_89to98.txt',
header=False, index=False, sep='  ')


# Control that the sizes are correct
print('Year 89 to 98')
print('Group 1')
print((dfESIPUT_Neutral_Group1_89to98).shape)
print((dfESIPUT_Positive_Group1_89to98).shape)
print((dfESIPUT_Negative_Group1_89to98).shape)

print('Group 2')
print((dfESIPUT_Neutral_Group2_89to98).shape)
print((dfESIPUT_Positive_Group2_89to98).shape)
print((dfESIPUT_Negative_Group2_89to98).shape)

print('Group 2 TopMV')
print((dfESIPUT_Neutral_Group2_TopMV_89to98).shape)
print((dfESIPUT_Positive_Group2_TopMV_89to98).shape)
print((dfESIPUT_Negative_Group2_TopMV_89to98).shape)

print('Group 3')
print((dfESIPUT_Neutral_Group3_89to98).shape)
print((dfESIPUT_Positive_Group3_89to98).shape)
print((dfESIPUT_Negative_Group3_89to98).shape)
```

Step 24

```
#1st January 1999-31st December 2008

split_date_1999 = pd.datetime(1999,1,1)
split_date_2009 = pd.datetime(2009,1,1)


# Text files for Group 1
dfNeutral_Group1_99to08 = dfNeutral_Group1.loc[(dfNeutral_Group1['date_final']
>= split_date_1999) & (dfNeutral_Group1['date_final'] < split_date_2009)]
dfESIPUT_Neutral_Group1_99to08 = dfNeutral_Group1_99to08[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group1_99to08.to_csv('dfESIPUT_Neutral_Group1_99to08.txt',
header=False, index=False, sep='    ')

dfPositive_Group1_99to08 =
dfPositive_Group1.loc[(dfPositive_Group1['date_final'] >= split_date_1999) &
(dfPositive_Group1['date_final'] < split_date_2009)]
dfESIPUT_Positive_Group1_99to08 = dfPositive_Group1_99to08[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group1_99to08.to_csv('dfESIPUT_Positive_Group1_99to08.txt',
header=False, index=False, sep='  ')
```

```
dfNegative_Group1_99to08 =
dfNegative_Group1.loc[(dfNegative_Group1['date_final'] >= split_date_1999) &
(dfNegative_Group1['date_final'] < split_date_2009)]
dfESIPUT_Negative_Group1_99to08 = dfNegative_Group1_99to08[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group1_99to08.to_csv('dfESIPUT_Negative_Group1_99to08.txt',
header=False, index=False, sep='	')


# Text files for Group 2
dfNeutral_Group2_99to08 = dfNeutral_Group2.loc[(dfNeutral_Group2['date_final']
>= split_date_1999) & (dfNeutral_Group2['date_final'] < split_date_2009)]
dfESIPUT_Neutral_Group2_99to08 = dfNeutral_Group2_99to08[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group2_99to08.to_csv('dfESIPUT_Neutral_Group2_99to08.txt',
header=False, index=False, sep='	')

dfPositive_Group2_99to08 =
dfPositive_Group2.loc[(dfPositive_Group2['date_final'] >= split_date_1999) &
(dfPositive_Group2['date_final'] < split_date_2009)]
dfESIPUT_Positive_Group2_99to08 = dfPositive_Group2_99to08[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group2_99to08.to_csv('dfESIPUT_Positive_Group2_99to08.txt',
header=False, index=False, sep='	')

dfNegative_Group2_99to08 =
dfNegative_Group2.loc[(dfNegative_Group2['date_final'] >= split_date_1999) &
(dfNegative_Group2['date_final'] < split_date_2009)]
dfESIPUT_Negative_Group2_99to08 = dfNegative_Group2_99to08[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group2_99to08.to_csv('dfESIPUT_Negative_Group2_99to08.txt',
header=False, index=False, sep='	')
# Text files for Group 2.1
dfNeutral_Group2_TopMV_99to08 =
dfNeutral_Group2_TopMV.loc[(dfNeutral_Group2_TopMV['date_final'] >=
split_date_1999) & (dfNeutral_Group2_TopMV['date_final'] < split_date_2009)]
dfESIPUT_Neutral_Group2_TopMV_99to08 = dfNeutral_Group2_TopMV_99to08[['PERMNO',
'rdq']]
dfESIPUT_Neutral_Group2_TopMV_99to08.to_csv('dfESIPUT_Neutral_Group2_TopMV_99to0
8.txt', header=False, index=False, sep='	')

dfPositive_Group2_TopMV_99to08 =
dfPositive_Group2_TopMV.loc[(dfPositive_Group2_TopMV['date_final'] >=
split_date_1999) & (dfPositive_Group2_TopMV['date_final'] < split_date_2009)]
dfESIPUT_Positive_Group2_TopMV_99to08 =
dfPositive_Group2_TopMV_99to08[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group2_TopMV_99to08.to_csv('dfESIPUT_Positive_Group2_TopMV_99t
o08.txt', header=False, index=False, sep='	')

dfNegative_Group2_TopMV_99to08 =
dfNegative_Group2_TopMV.loc[(dfNegative_Group2_TopMV['date_final'] >=
split_date_1999) & (dfNegative_Group2_TopMV['date_final'] < split_date_2009)]
```

```
dfESIPUT_Negative_Group2_TopMV_99to08 =
dfNegative_Group2_TopMV_99to08[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group2_TopMV_99to08.to_csv('dfESIPUT_Negative_Group2_TopMV_99t
o08.txt', header=False, index=False, sep='  ')

# Text files for Group 3
dfNeutral_Group3_99to08 = dfNeutral_Group3.loc[(dfNeutral_Group3['date_final']
>= split_date_1999) & (dfNeutral_Group3['date_final'] < split_date_2009)]
dfESIPUT_Neutral_Group3_99to08 = dfNeutral_Group3_99to08[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group3_99to08.to_csv('dfESIPUT_Neutral_Group3_99to08.txt',
header=False, index=False, sep='    ')

dfPositive_Group3_99to08 =
dfPositive_Group3.loc[(dfPositive_Group3['date_final'] >= split_date_1999) &
(dfPositive_Group3['date_final'] < split_date_2009)]
dfESIPUT_Positive_Group3_99to08 = dfPositive_Group3_99to08[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group3_99to08.to_csv('dfESIPUT_Positive_Group3_99to08.txt',
header=False, index=False, sep='  ')

dfNegative_Group3_99to08 =
dfNegative_Group3.loc[(dfNegative_Group3['date_final'] >= split_date_1999) &
(dfNegative_Group3['date_final'] < split_date_2009)]
dfESIPUT_Negative_Group3_99to08 = dfNegative_Group3_99to08[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group3_99to08.to_csv('dfESIPUT_Negative_Group3_99to08.txt',
header=False, index=False, sep='  ')

# Control that the sizes are correct

print('Year 99 to 08')
print('Group 1')
print((dfESIPUT_Neutral_Group1_99to08).shape)
print((dfESIPUT_Positive_Group1_99to08).shape)
print((dfESIPUT_Negative_Group1_99to08).shape)

print('Group 2')
print((dfESIPUT_Neutral_Group2_99to08).shape)
print((dfESIPUT_Positive_Group2_99to08).shape)
print((dfESIPUT_Negative_Group2_99to08).shape)

print('Group 2 TopMV')
print((dfESIPUT_Neutral_Group2_TopMV_99to08).shape)
print((dfESIPUT_Positive_Group2_TopMV_99to08).shape)
print((dfESIPUT_Negative_Group2_TopMV_99to08).shape)

print('Group 3')
print((dfESIPUT_Neutral_Group3_99to08).shape)
print((dfESIPUT_Positive_Group3_99to08).shape)
print((dfESIPUT_Negative_Group3_99to08).shape)
```

Step 25

```
#1st January 2009-31st December 2018

split_date_2009 = pd.datetime(2009,1,1)

# Text files for Group 1
dfNeutral_Group1_09to18 = dfNeutral_Group1.loc[dfNeutral_Group1['date_final'] >=
split_date_2009]
dfESIPUT_Neutral_Group1_09to18 = dfNeutral_Group1_09to18[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group1_09to18.to_csv('dfESIPUT_Neutral_Group1_09to18.txt',
header=False, index=False, sep='	')

dfPositive_Group1_09to18 = dfPositive_Group1.loc[dfPositive_Group1['date_final']
>= split_date_2009]
dfESIPUT_Positive_Group1_09to18 = dfPositive_Group1_09to18[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group1_09to18.to_csv('dfESIPUT_Positive_Group1_09to18.txt',
header=False, index=False, sep='	')

dfNegative_Group1_09to18 = dfNegative_Group1.loc[dfNegative_Group1['date_final']
>= split_date_2009]
dfESIPUT_Negative_Group1_09to18 = dfNegative_Group1_09to18[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group1_09to18.to_csv('dfESIPUT_Negative_Group1_09to18.txt',
header=False, index=False, sep='	')


# Text files for Group 2
dfNeutral_Group2_09to18 = dfNeutral_Group2.loc[dfNeutral_Group2['date_final'] >=
split_date_2009]
dfESIPUT_Neutral_Group2_09to18 = dfNeutral_Group2_09to18[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group2_09to18.to_csv('dfESIPUT_Neutral_Group2_09to18.txt',
header=False, index=False, sep='	')

dfPositive_Group2_09to18 = dfPositive_Group2.loc[dfPositive_Group2['date_final']
>= split_date_2009]
dfESIPUT_Positive_Group2_09to18 = dfPositive_Group2_09to18[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group2_09to18.to_csv('dfESIPUT_Positive_Group2_09to18.txt',
header=False, index=False, sep='	')

dfNegative_Group2_09to18 = dfNegative_Group2.loc[dfNegative_Group2['date_final']
>= split_date_2009]
dfESIPUT_Negative_Group2_09to18 = dfNegative_Group2_09to18[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group2_09to18.to_csv('dfESIPUT_Negative_Group2_09to18.txt',
header=False, index=False, sep='	')
```

```
# Text files for Group 2.1
dfNeutral_Group2_TopMV_09to18 =
dfNeutral_Group2_TopMV.loc[dfNeutral_Group2_TopMV['date_final'] >=
split_date_2009]
dfESIPUT_Neutral_Group2_TopMV_09to18 = dfNeutral_Group2_TopMV_09to18[['PERMNO',
'rdq']]
dfESIPUT_Neutral_Group2_TopMV_09to18.to_csv('dfESIPUT_Neutral_Group2_TopMV_09to1
8.txt', header=False, index=False, sep='    ')

dfPositive_Group2_TopMV_09to18 =
dfPositive_Group2_TopMV.loc[dfPositive_Group2_TopMV['date_final'] >=
split_date_2009]
dfESIPUT_Positive_Group2_TopMV_09to18 =
dfPositive_Group2_TopMV_09to18[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group2_TopMV_09to18.to_csv('dfESIPUT_Positive_Group2_TopMV_09t
o18.txt', header=False, index=False, sep='  ')

dfNegative_Group2_TopMV_09to18 =
dfNegative_Group2_TopMV.loc[dfNegative_Group2_TopMV['date_final'] >=
split_date_2009]
dfESIPUT_Negative_Group2_TopMV_09to18 =
dfNegative_Group2_TopMV_09to18[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group2_TopMV_09to18.to_csv('dfESIPUT_Negative_Group2_TopMV_09t
o18.txt', header=False, index=False, sep='  ')

# Text files for Group 3
dfNeutral_Group3_09to18 = dfNeutral_Group3.loc[dfNeutral_Group3['date_final'] >=
split_date_2009]
dfESIPUT_Neutral_Group3_09to18 = dfNeutral_Group3_09to18[['PERMNO', 'rdq']]
dfESIPUT_Neutral_Group3_09to18.to_csv('dfESIPUT_Neutral_Group3_09to18.txt',
header=False, index=False, sep='    ')

dfPositive_Group3_09to18 = dfPositive_Group3.loc[dfPositive_Group3['date_final']
>= split_date_2009]
dfESIPUT_Positive_Group3_09to18 = dfPositive_Group3_09to18[['PERMNO', 'rdq']]
dfESIPUT_Positive_Group3_09to18.to_csv('dfESIPUT_Positive_Group3_09to18.txt',
header=False, index=False, sep='  ')

dfNegative_Group3_09to18 = dfNegative_Group3.loc[dfNegative_Group3['date_final']
>= split_date_2009]
dfESIPUT_Negative_Group3_09to18 = dfNegative_Group3_09to18[['PERMNO', 'rdq']]
dfESIPUT_Negative_Group3_09to18.to_csv('dfESIPUT_Negative_Group3_09to18.txt',
header=False, index=False, sep='  ')
```

```python
# Control that the sizes are correct
print('Year 09 to 18')
print('Group 1')
print((dfESIPUT_Neutral_Group1_09to18).shape)
print((dfESIPUT_Positive_Group1_09to18).shape)
print((dfESIPUT_Negative_Group1_09to18).shape)

print('Group 2')
print((dfESIPUT_Neutral_Group2_09to18).shape)
print((dfESIPUT_Positive_Group2_09to18).shape)
print((dfESIPUT_Negative_Group2_09to18).shape)

print('Group 2 TopMV')
print(dfESIPUT_Neutral_Group2_TopMV_09to18.shape)
print(dfESIPUT_Positive_Group2_TopMV_09to18.shape)
print(dfESIPUT_Negative_Group2_TopMV_09to18.shape)

print('Group 3')
print((dfESIPUT_Neutral_Group3_09to18).shape)
print((dfESIPUT_Positive_Group3_09to18).shape)
print((dfESIPUT_Negative_Group3_09to18).shape)
```

14.8 Appendix VIII – Code for S&P 500 Constituents List

```
# NOTE - This code should be run in R

getwd()

setwd("/Users/Alexandra/Dropbox/THESIS/Econometrics/S&P500"

# install.packages('RPostgres', dependencies = TRUE)

library(RPostgres)


# Connect to the database

wrds <- dbConnect(Postgres(),

                  host='wrds-pgdata.wharton.upenn.edu',

                  port=9737,

                  user='mape14aq',

                  password='Globalreptrak2011',

                  sslmode='require',

                  dbname='wrds')


# Test accessing the monthly stock file "msf" from CRSP. Get data for Google and Apple during 2017

res <- dbSendQuery(wrds,

                   "select date, permno,  prc, ret

                   from crsp.msf

                   where permno in ('14593', '14542')

                   and date between '2017-01-01' and '2017-12-31'

                   ")


aapl_goog_prc <- dbFetch(res)

dbClearResult(res)

aapl_goog_prc
```

```
# Get all constituents from the monthly S&P500 list
res <- dbSendQuery(wrds,
                   "select *
                   from crsp.msp500list
                   where start >= '1989-01-01'
                   ")

msp500list <- dbFetch(res)
dbClearResult(res)
msp500list

# msp500list permno start ending

# Create a CSV-file with the output
write.csv(msp500list, file="MyS&P500.csv")
```

14.9 Appendix IX – Code for Graph Generation

The code for generating the graphs for the event study filled over 20 pages so therefore we have provided the code for Test 2 - 1989-2018 and Test 1 - 1989-2018. Starting from these the code is easy to modify to change the event window or do FF3.

```
* NOTE this code should be run in STATA
********************************************************************************
* THESIS GRAPHS
* Alexandra Petersson
* Copenhagen Business School
********************************************************************************
clear all

capture log close
log using Stata_Graphs.log, replace
set more off
about
********************************************************************************
* First install blindschemes
ssc install blindschemes, replace all

* We are using the plotplain scheme
set scheme plotplain, permanently
********************************************************************************

*** TEST 2 *** EVENT WINDOW [-5;30] ***

********************************************************************************

* Read in data Test 2 - MM 89-18
use test2_MM_89_18.dta, clear

* Convert numbers into percentages
gen Good_Group_1 = Good_Group1 * 100
gen Good_Group_2 = Good_Group2 * 100
```

```
gen Good_Group_21 = Good_Group21 * 100

gen Good_Group_3 = Good_Group3 * 100

gen Neu_Group_1 = Neu_Group1 * 100

gen Neu_Group_2 = Neu_Group2 * 100

gen Neu_Group_21 = Neu_Group21 * 100

gen Neu_Group_3 = Neu_Group3 * 100

gen Bad_Group_1 = Bad_Group1 * 100

gen Bad_Group_2 = Bad_Group2 * 100

gen Bad_Group_21 = Bad_Group21 * 100

gen Bad_Group_3 = Bad_Group3 * 100


drop Good_Group1 Good_Group2 Good_Group21 Good_Group3 Neu_Group1 Neu_Group2
Neu_Group21 Neu_Group3 Bad_Group1 Bad_Group2 Bad_Group21 Bad_Group3


* Legend

label variable Good_Group_1 "Group 1"

label variable Good_Group_2 "Group 2"

label variable Good_Group_21 "Group 2.1"

label variable Good_Group_3 "Group 3"

label variable Neu_Group_1 "Group 1"

label variable Neu_Group_2 "Group 2"

label variable Neu_Group_21 "Group 2.1"

label variable Neu_Group_3 "Group 3"

label variable Bad_Group_1 "Group 1"

label variable Bad_Group_2 "Group 2"

label variable Bad_Group_21 "Group 2.1"

label variable Bad_Group_3 "Group 3"


* Construct graph

twoway line Good_Group_1 Good_Group_2 Good_Group_21 Good_Group_3 Neu_Group_1
Neu_Group_2 Neu_Group_21 Neu_Group_3 Bad_Group_1 Bad_Group_2 Bad_Group_21
Bad_Group_3 evttime, xline(0, lpattern(solid)) yline(0, lpattern(solid))
xtitle("Event Time") ytitle("CAAR (%)") xsc(lcolor(black)) ysc(lcolor(black))
xlabel(#10) ylabel(#10) title("Post-Earnings Announcement Drift [-5;30 Days]",
size(medlarge)) subtitle("1989-2018 US Equity Market", size(medium) color(gs6))
scale(0.7) aspect(0.7) legend(pos(3) height(70) order( - "Good News:" 1 2 3 4 -
" " "Neutral News:" 5 6 7 8 - " " "Bad News:" 9 10 11 12)) lp(solid solid solid
solid dash dash dash dash shortdash shortdash shortdash shortdash) lc(gs0 gs4
gs8 gs12 gs0 gs4 gs8 gs12 gs0 gs4 gs8 gs12)
```

```
* Save graph
graph save Graph_89to18_5to30
********************************************************************************

*** TEST 1 *** UNDERLYING TREND *** EVENT WINDOW [-5;30] ***

********************************************************************************
* Read in data Test 1 - 89-18
use test1_5to30.dta, clear

* Convert numbers into percentages
gen Good_News = Goodnews * 100
gen Neu_News = Neutralnews * 100
gen Bad_News = Badnews * 100

drop Goodnews Neutralnews Badnews

* Legend
label variable Good_News "Good News"
label variable Neu_News "Neutral News"
label variable Bad_News "Bad News"

* Construct graph
twoway line Good_News Neu_News Bad_News evttime, xline(0, lpattern(solid))
yline(0, lpattern(solid)) xtitle("Event Time") ytitle("CAAR (%)")
xsc(lcolor(black)) ysc(lcolor(black)) xlabel(#10) ylabel(#10) title("Post-
Earnings Announcement Drift [-5;30 Days] - Underlying Trend", size(medlarge))
subtitle("1989-2018 US Equity Market", size(medium) color(gs6)) scale(0.7)
aspect(0.7) legend(pos(3) height(70) order( - 1  - 2 - 3 -)) lp(solid dash
shortdash) lc(gs0 gs4 gs8)

* Save graph
graph save Graph_Test1_5to30

********************************************************************************
```