The effect of insider trading on stock prices: A multiverse Event study

MASTER THESIS

Submitted in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Banking and Finance

HOLZMEISTER Felix Ph.D.
Department of Banking and Finance
The University of Innsbruck School of Management

Submitted by
LÖSCH Martina

Innsbruck, September 2021
Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.

Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Magister-/Master-/Diplomarbeit/Dissertation eingereicht.

__________________________  __________________________
29.09.2021                        Unterschrift
My special thanks go to my master’s thesis supervisor Ph.D. Felix Holzmeister for his support throughout the whole process. He always motivated me, even unintentionally, and gave my best help I could have gotten.

Looking back at the last two years, I want to thank also all my professors for the interesting insights in the various topics and a great time at the University of Innsbruck.

My biggest thank belongs to my family, who supported me in all my plans. They always stood by my side, tolerated my whims and motivated me even in difficult times.
Abstract

This thesis analyses the effect of insider trading of stock prices among the DAX 30 companies. Rather than conducting a common event study as in prior literature, it is extended by a multiverse analysis to investigate whether changes in the estimation window length, its distance to the event day and the pre-event window length have an impact on the outcomes of the event study. It is revealed that the variation in these parameters indeed cause different results and answers to the research question. Therefore, with some observations the null hypothesis could have been rejected, while in other cases a rejection was not possible. Regressions show that these differences indeed are explainable to a large extend by the variation in the parameters. This suggests, that the multiverse analyses give useful additional insights and should be considered in future researches.

Keywords: Insider trading, multiverse analysis, event study,
Contents

List of Abbreviations i

List of Figures ii

List of Tables iii

1 Introduction & Motivation 1

2 Literature 6

3 Methodology 10
   3.1 Event study 11
   3.1.1 Shortfalls 14
   3.2 Multiverse Analysis 16

4 Data and Model 19
   4.1 Data 19
   4.2 Model 20

5 Empirical Results 26
   5.1 Results from the multiverse event study 26
   5.2 Effects from the parameter selection 30

6 Discussion of this thesis’ multiverse analysis 36

7 Conclusion 38

References 46
List of abbreviations

AR Abnormal return

AAR Average abnormal return

APT Arbitrage Pricing Theory

OLS Ordinary least square

CAAR Cumulative average abnormal return

SEC Security and Exchange Commission

DAX Deutscher Aktien Index

CAPM Capital Asset Pricing Model

NYSE New York Stock Exchange

AMEX American Stock Exchange

Nasdaq National Association of Securities Dealers Automated Quotations
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Timeline of an event study</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>Comparison between the traditional and multiverse analysis</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>Distribution of insider trades among the DAX30 companies</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>Cumulative average abnormal returns for all observations</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>Cumulative average abnormal returns for all observations</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>Frequency distribution of p-values for all observations</td>
<td>29</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Estimation window properties</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>Event window properties</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>Classification of p-values in ranges</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>Results regression 1</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>Results regression 2</td>
<td>32</td>
</tr>
<tr>
<td>6</td>
<td>Results regression 3</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>Results regression 4</td>
<td>34</td>
</tr>
<tr>
<td>8</td>
<td>Impact of the explanatory variables on the p-value</td>
<td>35</td>
</tr>
</tbody>
</table>
1 Introduction & Motivation

The issue of impacts of insider trading on stock prices is an ongoing debate among the literature. Already in 1992, Meulbroek tries to answer the question whether illegal insider trading is harmful by measuring the abnormal returns. She finds that the price movement on insider trading days is about 40% to 50% larger than on days without news announcements or public news announcements. These results are also validated in the study of the Boesky’s insider trading activities around the acquisition of Carnation Company by Nestlé S.A. in 1984, where Chakravarty & McConnell (1997) regresses the effects on the stock prices. A third important study, whose results are in line with the ones previously explained, is the investigation of insider transactions around the Anheuser-Busch’s 1982 tender offer for Campbell Taggart, whereby a group of directors revealed non-public information to a small group of outsiders (Cornell & Sirri 1992). The consideration of these papers could attempt one to conclude, that the effects of insider trades on stock prices are obvious and do not need further investigation. But Chakravarty & McConnell (1999) argue, that these approaches have little contribution to the understanding of insider activity impacts. Their argument is that the right question is whether insider trading has a diverse effect than non-insider trading. This is the starting point for the re-analyzation of the papers from Meulbroek (1992), Cornell & Sirri (1992) and Chakravarty & McConnell (1997) with the conclusion, that the trades with non-public information have no distinguishable impacts on stock prices, compared the trades with public information. This result is also supported by Lakonishok & Lee (1998), who found that prices incorporate the information over a longer time horizon, which indicates a market’s underreaction to insider trades. It seems that insiders invest according to the contrarian investment strategy and are able to time the market better than other contrarian investors. The contrary conclusions of these studies raise the question why these differences occur. All authors use similar assumptions and methodologies to analyze their research question. While Chakravarty & McConnell (1997, 1999), Lakonishok & Lee (1998) and Cornell & Sirri (1992) set up different regressions,
Meulbroek (1992) conducts an event study that was used more frequently in the last decades (see e.g. Degryse, de Jong & Lefebvre (2009)). The latter follows the standard event study outlined in Campbell, Lo & MacKinlay (1997) and MacKinlay (1997), which are two of the essential literatures regarding this methodology. An event study focuses on measuring abnormal returns (AR) as the difference of expected returns, given that any event happened, and the actual return observed at the event. Even if the model seems straightforward in the first moment, there are several parameters that have to be arbitrarily selected by the researchers. First, the event itself can be a certain day or a period of a few days. The estimation window, that is used for the normal return calculation, can be chosen individually and differs among the literature. Generally, a period of 150 (Meulbroek 1992) to 250 (Degryse et al. 2009) trading days is taken into account, but the optimal length depends on whether a single security or more are observed. Additionally, various statistical or economic models can be adopted to estimate the normal returns. The most used are the constant-mean-return and the market model, but many others are applicable too. Campbell et al. (1997) By choosing the estimation window length and its distance to the event day as well as the event window length it is, even unintentionally, possible to generate results that were completely different. (Thompson 1995)

To solve this problem, a broader approach to conclude about the effects of a certain event is to extent the event study by a multiverse analysis. Multiverse analyses are a rather new approach and investigate all the reasonable scenarios across the whole dataset, which avoids to arbitrarily choose parameters. In this way, this methodology can increase the transparency and enables to report a wider range of results than the simple event study. Steegen et al. (2016) is one of the most important research in conjunction with multiverse analysis. It is based on two psychological studies by Durante et al. (2013) that examine the effect of fertility and the relationship status on religiosity, fiscal and social political attitudes and on voting and donation preferences. The new method led to an increase in the
robustness of the results and makes it possible to identify the most consequential choices in the original works. In this way, effects of parameter selection can easily be shown and interpreted. Furthermore, it allows a replication and extension of prior researches. \cite{Saraceno et al. 2020} The innovation of this approach justifies the rare literature regarding multiverse analysis. Although some studies in the psychological field were conducted that reveal lots of advantages and support the application of this type of study, in the financial or even economical context no multiverse analysis could have been found. In this manner, this thesis contributes to the literature as the first multiverse event study and the first multiverse analysis in regards to a financial topic.

The aim of this thesis is to answer the question whether insider trading does affect stock prices and contribute to the ongoing debate. Moreover, the multiverse analysis is used to point out the advantages of this approach and to test whether the parametric decisions that are made in single event studies have an impact on the results and the general view on the topic. More specifically, in this thesis are examined insider trades among the DAX 30 companies between February 2017 and February 2021. In this time period were detected 554 insider trades among 27 companies. Instead of conducting a common event study and setting the parameters arbitrarily, the model in this study is based on varying all periods that could influence the outcome. Thus, the length and the starting time of the estimation period, the length of the event window and the post-event window are changed and combined in each reasonable way, such that 4,200 combinations are run for 554 different events. This enables to identify possible patterns from changes in the variables. Conceivable patterns could be a higher statistical power or higher abnormal returns for some combinations of the periods. Although an alternation of the model could provide further informative insights, this is out the scope of this thesis. Also, as the dataset spans a long time period and contains many events, the variation of the periods is done on a non-daily basis, but in a 2-day or rather 5-day time lag to
simplify and speed up the computation process. Even though the multiverse in this way does not consider all the possible combinations, these steps should not bias the results to a large extend and are therefore regarded as appropriate. The inspection of the p-values of the different observations serves to get an answer regarding the research question whether insider trading does have an effect on stock prices or not. Afterwards, regressions are conducted to check the effect of the parameters on the p-values.

The results reveal interesting and informational insights. Although the cumulative average abnormal returns (CAAR) for all observations evolve in a similar way, there are noticeable differences in their extent. As already a graphical representation depicts a sharp decline in the CAAR before the event day, it indicates an effect of the insider trades on the stock prices. The test of this hypothesis finds strong variations in the p-values across the observation, wherefore a general answer to the research question cannot be given. In 482 cases the null hypothesis is rejected, while in 3,718 cases the null cannot be rejected. The regressions show that the variation in the p-values is at a high percentage (72.7%) explainable with the change in the parameters. The extend of the model by the interaction between the changes estimation window length and its distance to the event day as well as the interaction between the distance to the event day and the pre-event window length again slightly increase this explainable part. Moreover, the regressions indicated a positive relationship of the distance of the event window to the event day on the p-value, while all other parameters affected the p-valued in a negative direction. The large explanatory power of the parameters in the regression and the differences in the p-values suggest the usage of the multiverse analysis to improve the robustness of the results and to decreases the biases that come along with the conclusion drawing based on single statistics.

Section 2 gives a short overview of the existing literature concerning insider
trading. Afterwards the event study methodology and its shortfalls, as well as the concept of multiverse analyses are described in section 3. Section 4 depicts the model used in this study and all specifications of the parameters that are varied in the contexts of the multiverse. Further, some sample characteristics of the underlying data are given. In section 5, the results of the empirical analysis are interpreted and the effects of the parameter selection are tested. Section 6 discusses the insights from the multiverse before the main points are summarized in section 7.
2 Literature

The discussion of insider trading effects on stock prices has a long history and has provoked two controversial sights on it: while one group of economists argue that insider trades can increase market efficiency by releasing information to the market and do not harm other investors, another group sees the problem in the profits that insiders could make by trading on non-public information. Manne (1966) mainly started this discussion, when he published the book “Insider trading and the stock market”, where he outlines the stabilizing effect on the market and strengthens the argument that insiders can be helpful in foreshadowing stock prices in the short- as well as in the long-run. Manne (1966) criticizes the regulation of insider trading by the Security and Exchange Commission (SEC) as he suggests that corporate traders do not harm other shareholders and, instead, let them profit from faster rising prices. As an answer to these conclusions, Schotland (1967) states that even if insiders would not negatively affect stock prices and markets, it remains still morally wrong to trade based upon non-public information as it reduced the fiduciary trust of market participants. He refuted Manne’s arguments with the help of studies regarding the influence of insider trades on stock prices, which had been rare at this time: Smith (1941) analyzed such trades in NYSE (New York Stock Exchange)-listed companies with data from March 1935 to February 1939 and concluded that the impact on the market could not be confirmed. Fifty NYSE-listed corporations from 1957 to 1961 were investigated by Wu (1963) with the aim to find how regularly insider trades take place and how they affected the stock prices. The results suggested that only 1% of the total NYSE volume was explainable by corporate traders, wherefore the impact on prices should not be significant. Few years later, Lorie & Niederhoffer (1968) were able to gain contrary results by including the size of the trades, distinguishing between purchases and sales and examining trades around large price changes more deeply. They found that accumulated insider trades were able to outperform the market by 9.51% and that insiders are more likely to buy


6
before large price increases and sell before large price decreases. According to them, insider purchases are much more frequent than Wu (1963) observed. Additionally their results suggested that purchases are usually followed by other purchases and sales by sales. In the following centuries, insights in the topic rose as the effect of insider trades on the pricing efficiency, as well as their effects on financial markets was treated in different studies. Givoly & Palmon (1985) analyzed whether abnormal returns of corporate trades in 68 randomly chosen companies from the AMEX (American Stock Exchange) were caused by changes in the price after the non-public information was disclosed or whether they were a product of the trades themselves. Results showed that the abnormal returns were not explainable by the market’s reaction to the news and were rather induced by the insider trades. This conclusion is in line with Jaffe (1974) and Finnerty (1976), who measured abnormal returns with variations of the Capital Asset Pricing Model (CAPM) and noticed between 2% and 8.3% abnormal returns, depending on the considered length of the period after the trades. Moreover, Givoly & Palmon (1985) found evidence that outsider market participants follow insiders and imitate their trades. A further analysis that was build up on 60,000 insider transactions from 1975 to 1981 expands these insights by the finding that outsiders cannot earn abnormal returns net of trading costs by imitating insiders. The study also took into consideration the types of insiders and concluded, that corporate traders with more knowledge of the firm, such as members of the board of director or office-directors, are better predictors of future price evolvements than other insiders. (Seyhun 1986) The former seemed to act mostly based on firm-specific information, while the latter traded according the economywide condition, which is also in line with the outcomes from Seyhun (1988).

Lakonishok & Lee (1998) re-examined Seyhun’s (1986, 1988) finding of the insider’s predictability of future stock prices by running regressions based on an independent variable that measures the insider trading activity, with the aim to differ whether abnormal returns were a result of insiders’ trades or trades from contrar-
ians. Empirical analysis indicated that aggregated corporate traders were able to predict market price evolvement over a period of 12 months. This was explicable with the contrarian strategy that insiders followed and indicated a probable usefulness for timing the market. Even if manager’s trades were more informative than shareholder’s ones, their predictive power was not consistent over the whole observed period. In addition to that, insiders in smaller firms were better in forecasting than the ones in larger firms. Although insiders were successful predictors, Lakonishok & Lee (1998) underline that drawing trading strategies out of their actions might not be effective as trading in large stocks has a limited impact and trading in small stocks is costly. Their study investigated insider trades from all companies traded on the NYSE, AMEX and Nasdaq from 1975 to 1995. Thus, they had a sample of over one million corporate trades, wherefore it has been one of the most comprehensive studies at this time. Their aim was to measure the magnitude of the trading activity, whether insiders are able to forecast market evolvement and if these trades “explain the cross-sectional variation of individual stock returns”. In half of the examined companies, insider trading took place at least once a year, of which sales outweighed purchases. The event study showed that trades were more informative over a longer time horizon, while immediately after the trades, no large stock price movements took place. This indicated that the market underreacted to the trades. This was contrary to Meulbroek (1992), Cornell & Sirri (1992) and Chakravarty & McConnell (1997), who observed stock price reactions right at the events where insider traded based on information regarding company takeovers. Meulbroek (1992) took a sample of 218 different companies and found significant price movements on insider trading days that was up to 50% larger than on days without news announcement or public news announcements. An immediate price reaction to the corporate trades let her conclude that the stock price accuracy increased with insider trades. Cornell & Sirri (1992) studied in detail insider transactions around the Anheuser Busch’s acquisition of Campell Taggart and supported the findings of abnormal returns around such trades. Chakravarty & McConnell (1997) inspected
deeply Ivan Boesky’s trades, which were based on non-public information regard-
ing Nestlé’s acquisition of Carnation Company and wanted to regress the effects on stock prices, the bid/ask spreads and the market depth. For the bid/ask-spread and the market depth, no negative effects were noticeable. Price run ups in conjuncture with Boesky’s trades were confirmed, even after correction for other variables that could have affected the price.

Although the results of these three researches have been commonly accepted as valid and cited by others, Chakravarty & McConnell (1999) criticized the research question. They stated that instead of exploring whether corporate trading did affect stock prices, it would have been more correct to investigate whether the effects of insider trades were significantly different to the non-insider ones. Resampling their study from 1997 and conducting a $X^2$-test for the independent variables for the Boesky trades and non-Boesky trades revealed that, despite the fact that Boesky’s trade impacted stock prices, no difference to the non-Boesky trades was observable.

Nonetheless, the initial studies stayed fundamental for further research. For example, Degryse et al. (2009) build their examination upon Meulbroek (1992). They referred to a long discussion of insider trading regulation by setting up an event study to test short-term impacts of insider trades on stock prices in the Netherlands and the possible changes through regulations. Concluding that insiders revealed information to the market and were able to gain abnormal returns of up to 2% in one month and a half, they supported previous results (e.g. Meulbroek (1992), Cornell & Sirri (1992), Chakravarty & McConnell (1997)). Furthermore, abnormal returns of top executives were larger than the ones from other insiders, which is in line with Seyhun (1985). Also, Lakonishok & Lee (1998) outcomes that insider in small companies benefit more from their trades and that corporate traders act like contrarians was confirmed by the findings of Degryse et al. (2009) as they observed corporate purchases to take place more frequently after prices fell and sales after prices rose. Degryse et al. (2009) state that, although insiders possess non-public information
about their company, not every trade is based on such information. The authors saw an opportunity for outsiders to learn from insiders about the fundamental value of companies, but as regulatory changes caused insider trades to be less information driven, this opportunity shrank.

In 2017, Hoang et al. published a study with the aim to re-examine prior literature regarding insider activity and its impact on market efficiency and stock prices. They make usage of thirty insider purchases and thirty insider sales in S&P500 stocks. In order to check whether corporate traders were able to outperform the average market, their mean returns were calculated for periods of 1 month, 6 months and 12 months. Despite first results, that revealed that the outperformance of purchases was statistically significant over all periods, the abnormal returns of sales were only significant during a 1 month and 6 month period. Referring to Malkiel (1973), who states that higher profits could also be earned if insider manipulate the crowd such that stock prices evolve in a favorable direction, Hoang et al. (2017) concluded that it was not able to gain evidence that the abnormal returns that had been found, were a product of trading on non-public information.

To summarize, insider trading was broadly investigated throughout the literature but a general conclusion about the impact on stock prices as well as the informational value for markets is still not given. While some are convinced that insider trades may be harmful to markets as these traders are able to profit from non-public information but do not contribute to the market at all, others found evidence that the information is revealed to the market as the trades take place and thus, outsiders can derive strategies from insider’s transactions.

3 Methodology

This study is based on the event study methodology. To overcome possible parameter selection biases, it is extended by a multiverse analysis, wherefore these methods
are described in the following section.

3.1 Event study

Event studies have generally been used as a method to measure abnormal returns around a certain event and conclude whether an event has an impact on a securities price. MacKinlay (1997) provides a detailed overview of the methodology and its applications.

The aim of an event study is to examine whether the returns around an event do significantly differ from some expected returns without the event. Such events can be earnings or takeover announcements (Campbell et al. 1997), stock splits (Fama et al. 1969), dividend payments or tax law changes (Peterson 1989) or insider trades, whereas the latter ones can occur in connection with the former (e.g. Cornell & Sirri 1992, Chakravarty & McConnell 1997). It is required to identify the event of interest and the period that will be suggested to be influenced by the event. This is called the event window, which can have different lengths. Usually, not only the day at which the event occurred, but also a period after and sometimes before the event is taken into account. Especially, in the case of announcements, at least the day after the announcement is in the event period. The inclusion of days prior to the event could be needed if the market could be able to gather information about the event prior to the announcement, like it is possible for earnings announcements. As at some point the market should not anymore be influenced by the event and should go on like before, a post-event window can be introduced to examine if this comes to pass. (MacKinlay 1997) The lengths of the event and post-event windows vary a lot among the existing literature. Meulbroek (1992) uses an event window that starts 20 days prior to the event and ends with the day of the event itself. Same is valid for Degryse et al. (2009) but they add a post-event period of 30 days. Pettengill & Clark (2001) use a similar pre-event window with 25 days but a much longer post-event window with 125 days, while Ahern (2009) builds a model with
just 5 pre- and 5 post-event days. Kim et al. (2019) instead set the event window as the day of the event itself and one day prior.

As the method is build upon measuring the abnormal returns of these periods, an estimate of the normal returns is needed. This requires two definitions made by the researcher: a period that is used to calculate the expected returns and an appropriate model. The period, which is known as estimation period, is usually a period prior to the event-window that should not include the day of the event as it could bias the outcomes. (MacKinlay 1997) When choosing the length of the window, it should be reflected if parameters had been stable over time. Instable parameters could cause a researcher to set a shorter estimation period. This, on the contrary, may demand a prediction error correction. Moreover, the length depends on the data available. For daily data the usual length is between 100 and 300 days, for monthly data it is between 24 and 60 months. (Peterson 1989) Although it is common to use an estimation period prior to the event, it is possible to estimate the normal returns with data after the event or a combination of both (e.g. Mandelker (1974), Copeland & Mayers (1982)). Even the distance of the estimation window to the event can vary. For example, Kim et al. (2019) define an estimation period that ends 21 days before the event day such that the days till the event are not considered, neither in the event window nor in the estimation window. Figure 1 summarizes the time horizon that is looked at if an event study is applied:

![Figure 1: Timeline of an event study](image)

Figure 1: Timeline of an event study
By applying an event study, time is measured in event days rather than calendar days, where the event day is presented as $\tau=0$. As the event window can contain some days prior and after the event it is given by $\tau = T_1 + 1$ to $\tau = T_2$. The post-event window, that serves to check whether the market normalizes at some point in time, goes from $\tau = T_2 + 1$ to $\tau = T_3$. The estimation window, here starts and ends before the event and is depicted by $\tau = T_0 + 1$ to $\tau = T_1$. Thus, the event, post-event and estimation windows have lengths of $T_2 - T_1$, $T_3 - T_2$ and $T_1 - T_0$ respectively. (MacKinlay 1997)

The models to calculate normal returns can either be economic or statistical. Statistical models do not include economic arguments and strongly expect asset returns to behave according to statistical assumptions, while economic models rely on investors’ behaviors and, in practice, need to be expanded by statistical assumptions. (MacKinlay 1997) Often used statistical models in conjunction with event studies are the market model (e.g. Fama et al. (1969)), the market adjusted model (e.g. Brown & Warner (1980)), the constant-mean return model or multifactor models. Common economic models are the Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT) (e.g. Brown & Weinstein (1985)). While the CAPM was frequently applied for event studies in the 1970s, later on it was replaced by the market model as it enabled to overcome some of the CAPM’s restrictions. The APT is considered to be an appropriate model, but is more difficult in the application. Thus, economic models are less often used with event studies. (MacKinlay 1997) A detailed description of the model used in this thesis is given in section 4.2.

Knowing the estimation period and an appropriate model, the expected returns can be calculated. The abnormal returns are simply the difference of the observed returns in the event period and the expected returns. To draw general conclusions out of the analyzed abnormal returns of an event, an aggregation is needed. Abnormal returns can be aggregated through time or across securities or both. Generally,
as more than one events are regarded in an event study, aggregation is done in both ways, whereby the order does not change the results. Furthermore, the possible methods are the same, no matter which model was used to calculate the abnormal returns. One common method was shown by MacKinlay (1997). As it is used in the context of this thesis, it will be shown in the model specification in section 4.2.

The aggregation yields us the cumulative average abnormal returns (CAAR). In a last step, statistical testing of these is necessary to check whether the abnormal returns are statistically significant. Under the null hypothesis, which is that the event has no effect on returns, the CAAR must be normally distributed. Applicable statistical tests depend on the type of the data and whether single or multiple event studies are conducted and can be chosen by the researcher. The statistical test that is applied in the context of this thesis is described in section 4.2.

3.1.1 Shortfalls

Henderson (1990) depicts the difficulties that arise by conducting an event study. Defining the event date may seem straightforward, but as the length of the event should be chosen as the time that is expected for the market to incorporate the event, an inappropriate event window length could cause imprecise results. An accurate specification of the event date and, thus, making use of daily data, can improve the statistical power of an event study (Brown & Warner 1985). Not also the event period but also the estimation period can influence the outcomes of the event study. As mentioned above, a general approach is to use a period before the event. Although, including data after the event could decrease forecasting errors (Ahern 2009).

Mandelker (1974) examines the cumulative average abnormal returns among various estimation windows, whereby some included just periods before the event and some also included periods after the event. He argues that the CAAR may vary as
the beta depends on the estimation period considered. Moreover, returns in the estimation window could be influenced by other events. As it is not obvious to measure the impact of confounding events, a short period to estimate the normal returns can be applied to reduce the probability, that such events are included. \cite{McWilliams et al. 1999} On the other hand, longer periods can improve the parameter stability, whereas the method enforces the researcher to outweigh the pros and cons of the possible lengths and decide for an adequate estimation window \cite{Peterson 1989}.

A further consideration for the return calculation is the model selection. The debate over which model is the most appropriate for the event study methodology is still not solved. \cite{Brown & Warner 1980} compared the Mean adjusted return, the Market adjusted return and the Market and risk adjusted return model for monthly data. They found that the results in the case of non clustering, which means that event periods did not overlap, were similar for all models. For clustered events, the performance of the Mean adjusted return model was worse than the one of the other models and discovered significant abnormal returns only in 35% of the cases. Multifactor models or other more complicated approaches did not outperform the simple one-factor models. As properties of daily stock returns distinguish from monthly data, the study was repeated by \cite{Brown & Warner 1985} with daily data. It was concluded that the Market model and the Market adjusted return model did outperform the Mean adjusted return model. Thus, there does not exist one right model, but depending on the data set, one could perform better than others.

To sum up, the event study methodology is commonly used to measure effects of various events on stock prices. It is build up on several parameters, such as the length of the event window and the estimation window, as well as the timing of the estimation window prior or after the event and the model, that is used to estimate expected returns. These parameters can be individually chosen by the researcher and therefore could influence the outcomes and their statistical power.
3.2 Multiverse Analysis

To overcome the shortfalls of parameter selection, expanding a single analysis by a multiverse can be helpful. The idea behind multiverse analyses is to “perform all analyses across the whole set of alternatively processed data sets corresponding to a large set of reasonable scenarios” (Steegen, Tuerlinckx, Gelman & Vanpaeme 2016). In this way, multiverse analyses enable deeper insights in statistical results and eliminate possible biases that may result from the arbitrarily chosen parameters (Dragicevic et al. 2019). Moreover, key choices that lead to different results among various studies can be identified and the robustness of results will increase. Instead of providing one outcome and justify that, the multiverse allows to show a set of possible results, which states that there may be various correct outcomes for one dataset. By applying a multiverse analysis it can be possible that the null hypothesis can neither be rejected, nor not rejected unambiguously. Following that, the outcome of a multiverse analysis might not be one clear answer to the research question but rather a general description of the topic and the insights from the investigations. Although this can be considered as a weakness of the approach, it is also its main strength. It allows the researcher to examine the whole data set and to publish all of the outcomes. (Steegen et al. 2016) This is the main feature that it differentiates from a traditional analysis. Dragicevic et al. (2019) illustrates the advantage of a multiverse analysis in a straightforward manner. While in a traditional analysis only a small part from all the data set is analyzed and even a smaller part is reported, conducting a multiverse analysis means analyzing and reporting all possibilities. (see figure 2)

The first step in the application of this analysis is to identify the variables that have to be set in the one-dimensional model. Afterwards, there are specified all reasonable values that the identified variables can take with the underlying set of data. Finally, the analysis is executed with all possible combinations of the vari-
ables, which leads to a multiverse of results that must be interpreted to get overall inferences of the initial research question. (Simonsohn, Simmons & Nelson 2015) Steegen et al (2016) performed a multivariate analysis that is one of the most cited literature in this conjuncture. It was based on data of two prior studies that covered the effect of fertility and relationship status on the religiosity, fiscal and social political attitudes and on voting and donation preferences (Durante et al. 2013), to evaluate the robustness of the prior results. They build up a multiverse of data sets and examine each data set individually, which yield a set of statistical results for the whole multiverse. In this manner, 120 selection combinations for the rebuild of the first study and 210 for the second were calculated. An investigation of the p-values made visible, that most of the combinations indicate a statistically significant effect. Furthermore, the method increased transparency as all results were published and allowed to draw conclusions about the entire data instead of relying on one determined data set.

The multiverse approach is quite new, wherefore literature is limited. While in the field of psychology some studies were conducted, in the economy no examples
could have been found. Beside the known study from Steegen et al. (2016), Harder (2020) set up a multiverse analysis to re-examine prior studies regarding shooting biases with contradictory outcomes to provide a better general picture and combine all investigated single parts in one study. Saraceno et al. (2020) got the task to evaluate the correlation between the size of LGB population in American districts and the sponsoring for gay rights from their representatives in Congress. Replicating similar studies, they were able to show that the results were strongly influenced from the construction of the models, that were done by the researchers.

An application of multiverse analysis on event studies is feasible too. As described above, for the event study approach, researchers have to arbitrarily choose parameters like the event length, the pre- and post-event window length and the estimation window length as well the distance in time of the estimation window length to the event. Thus, a multiverse event study requires to define these variables with all possible and reasonable preferences. The variables are then combined such that for every possible combination an event study is carried out. A further implication could involve changing the model used to measure the expected and abnormal returns. Prior literature gives a fundamental idea of reasonable parameter choices. There already exist some attempts to vary the parameters as in Kim et al. (2019), where the cumulative average abnormal returns were calculated for a range from 5 to 120 days after the insider transaction. They found that in most of the cases, abnormal returns have been larger over longer periods. Already these results indicate that the extension of event studies by a multiverse could reveal interesting additional finding. Nonetheless, despite careful literature review, no multiverse event study could be found.
4 Data and Model

In the following chapter, the underlying data for this study and the model are explained in more detail.

4.1 Data

This study investigates the effects of insider trading in the DAX 30 companies on the corresponding stock prices. The underlying data used for this purpose are daily returns of the DAX 30 since February 2017 until February 2021. As the composition of the DAX changed over the years and for some of the terminated companies the necessary data was not available anymore, the current DAX-companies, as of February 2021, have been examined over the mentioned time horizon. They represent the thirty largest and most liquid companies in Germany from the automobile, chemical, insurance, pharma, software and other sectors. ([Deutsche Börse Group](https://www.deutscheboerse.com) 2021)

As the data was retrieved from Bloomberg, its definition of insider trading, is used. This is inclined towards the SEC Insider trading policy and specifies insider trading as the violation of the fiduciary duty and trust by using nonpublic information to trade securities. The SEC points out the shrinking integrity and fairness which goes on with such trades. Within this definition, officers, directors, stockholders or employees can act as insiders. ([SEC](https://www.sec.gov) 2013)

For this research, the volume of the trades was not considered. Furthermore, no differentiation between buying and selling was done. For some events, no company or market return on the event day was contained in the data, for others were recorded insider trades right at the beginning at their listing at the stock exchange, such that an insufficient amount of returns prior to the event was given. This required the exclusion of these event. According to this event specification, 554 events were detected since February 2017. A first analysis of the data set shows an unequal distribution of the events over the companies and over time. At SAP, with 63, were
recorded the most insider trades, while at Volkswagen and Continental only one took place. (see figure 3) Moreover, in three companies, no insider trades were found since February 2017. These are Linde, Munich Re and RWE.

In the sample data, insiders from different companies partially traded on the same day, which means that events were clustered in time. The most trades took place in the first quarter of 2020 as the Coronavirus spread all over the world and an economic crisis was foreseeable.

4.2 Model

This study is based on the event study methodology according to MacKinlay (1997). To overcome possible parameter selection biases, it is extended by a multiverse analysis as in Steegen et al. (2016). This enables a new perspective on the outcomes and a complete analysis of the data set rather than a single statistical analysis. The methodologies were described in section 3, whereas here, the applied model is
explained in more detail.

As mentioned above, in the context of this study, the events are specified as insider trades according to the SEC’s (2013) definition. To explore their impact on stock prices, the market model was used for the calculation of normal and abnormal returns. It is a statistical model that is often adopted in event studies as it is easy in its application and reliable for most of the data. Applying this model requires the assumption “that asset returns are jointly multivariate normal and independently and identically distributed through time”, which is unproblematic in practice. Under the market model, the expected returns for any security \( i \) at any point in time \( t \) is given by:

\[
R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it}
\]  

with \( \mathbb{E}(\epsilon_{it}) = 0 \) and \( \text{var}(\epsilon_{it}) = \sigma^2_{it} \),

where \( R_{it} \) and \( R_{mt} \) represent the return on security \( i \) and the market portfolio at period \( t \). \( \epsilon_{it} \) is the disturbance term that is expected to be zero.

The parameters \( \alpha_i, \beta_i \) and \( \sigma^2_{it} \) have to be estimated with a regression. As, under the assumption that the criteria for the statistical models regarding the distribution of the returns is fullfilled, ordinary least square is generally an adequate method to estimate the parameters (MacKinlay 1997), also this study makes usages of it. Apart from the appropriate regression method, a selection of the period that is used to estimate the normal returns, which is called the estimation window, is needed. By applying a multiverse analysis, the length of this estimation window varies. In this study, the estimation window goes back to between 370 and 125 trading days before the event. Rather than changing these by single days, the length are changed in a 5 days-rhythm, such that the possible lengths are 370, 365, \ldots, 125.
Moreover, as depicted in section 3.1, the distance of the estimation window to the event day can be chosen, which means that not all the returns till the event are taken into account for the normal return estimation by the OLS method, but only the returns up to a certain point in time before the event. Here, possible end days for the estimation window are the event day -40 days, or in intervals of 3 days, backwards to up to event day -1. Adding this variation, also the minimum length of the estimation window changes to 85 (minimum length of the estimation window of 125 minus the end of the estimation window at event day -40) trading days. Table 1 summarizes the possible estimation window lengths and end days. In total, there are 50 possible estimation window lengths and 14 possible distances to the event day.

Table 1: Estimation window properties

<table>
<thead>
<tr>
<th>Estimation window length</th>
<th>End day (in distance to event day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>370, 365, 360, ..., 125 trading days</td>
<td>-40, -37, -34,..., -1</td>
</tr>
</tbody>
</table>

Knowing the periods, the OLS, with the DAX 30 returns and an added constant as the independent variable and the stock returns as the dependent variable, can be regressed for every estimation period and every event. Inserting the estimated $\alpha_i$, $\beta_i$ and $\sigma^2_{\epsilon}$ and the market return $R_m$ in formula 1, the expected returns for the period around and after the event can be computed. As the matter of interest in an event study are the abnormal returns around an event, the expected returns are subtracted from the actually observed company returns by:

$$AR_{i\tau} = R_{i\tau} - \alpha_i - \beta_i R_{m\tau}$$  \hspace{1cm} (2)  

The periods for which these abnormal returns are calculated depend on the defined event window and post event window. Recalling figure the event window may cover some days before and after the event $\tau$ (see chapter 3.1). In this study, it starts prior to the event and ends with the event, as the timing of the trade is
clear and it is assumed that they should be incorporated immediately in the stock prices. The period prior to the event, again, is differed in the run-throughs of the model and can start from event day -30 days up to -20 days in a rhythm of 2 days. Hence, the lengths of the pre-event window to the event depend between 30 and 20 trading days. The length of the post-event window is helpful to check whether the market normalized after the event in a certain time horizon. For reasons of speed and simplicity, in this study, the post-event window is not varied and is fixed at 10 days. The variation of this window could be an further extension in later studies. Table 2 sums up the 10 possible lengths of the pre-event window.

<table>
<thead>
<tr>
<th>Table 2: Event window properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event window lengths</td>
</tr>
<tr>
<td>20, 22, 24, 26, 28, 30 trading days</td>
</tr>
</tbody>
</table>

The calculated abnormal returns for the mentioned parameters, are assumed to have certain characters: Conditional on the market return over the event window, under the null hypothesis, the abnormal returns "will be jointly distributed with a conditional mean of zero and a conditional variance \( \sigma^2(AR_{it}) \), where

\[
\sigma^2(AR_{it}) = \sigma^2_\epsilon + \frac{1}{L_1} \left[ 1 + \left( \frac{R_{m\tau} - \hat{\mu}_m}{\hat{\sigma}_m^2} \right)^2 \right]
\]

\[(3)\]

\( (L_1 \text{ refers to the length of the estimation window}) \]

The conditional variance in formula \( 3 \) is composed of the disturbance variance and a variance due to the sampling error of \( \alpha_i \) and \( \beta_i \), which approaches zero as the length of the estimation window \( L_1 \) gets larger. (MacKinlay 1997) In this thesis, for reasons of simplicity, it is assumed, that the estimation window is large enough, wherefore the second term is approximately zero from which we get:
As conducting a multiverse analysis means to vary all the parameters and run the model for all possible and reasonable combinations of these, these steps are carried out \(50 \times 14 \times 6 = 4,200\) times for every event. In this way, in a further step, the 4,200 observations can be compared to get a complete understanding.

To draw general conclusions, an aggregation must follow. Although this aggregation could be done also just through time or across securities, in this thesis, aggregation is carried out through both dimensions according to MacKinlay (1997). Specifically, this means that for each variation of the parameters, the abnormal returns are aggregated across all events and afterwards through time. Aggregation of the abnormal returns across events \(N\) is done by:

\[
AAR_{\tau} = \frac{1}{N} \sum_{i=1}^{N} AR_{i\tau}
\]

which yields an abnormal average return (AAR) for each period. For a large estimation window, its variance is given by

\[
var(AAR_{\tau}) = \frac{1}{N^2} \sum_{i=1}^{N} \sigma^2_{\epsilon}
\]

More specifically, as the output from the abnormal return calculation were 4,200 observations for each event, this step aggregates the results across the events such that an overall of 4,200 observations remain. This allows to calculate the cumulative average abnormal return (CAAR) for any interval in the event window as
CAAR(τ_1, τ_2) = \sum_{τ=\tau_1}^{\tau_2} AAR_τ \quad (7)

and its variance as

\text{var}(CAAR(τ_1, τ_2)) = \sum_{τ=\tau_1}^{\tau_2} \text{var}(AAR_τ) \quad (8)

The aggregation of abnormal returns in this way requires the assumption of non-clustered events, which means that the event windows do not overlap in calendar time. Clustering would require the consideration of the correlation among the securities and subsequently more complex aggregation methods (Campbell et al. 1997). Even though some events in the sample took place at the same days, and thus, events were clustered, for sake of simplicity, here it is assumed that the abnormal returns of the identified events are uncorrelated and not clustered.

After the aggregation, which yields the cumulative average abnormal returns and the calculation of its variance, the H₀, which is that insider’s trades have no effects on the stock price needs to be tested. While different test statistics are possible, this thesis uses MacKinlay’s 1997 approach:

θ₁ = \frac{CAAR(τ_1, τ_2)}{\text{var}(CAAR(τ_1, τ_2))^{\frac{1}{2}}} \sim N(0, 1) \quad (9)

In this manner, the cumulative average abnormal returns are standardized and, under the null, these should follow a normal distribution. If the CAAR are significantly different from zero, the H₀ can be rejected. As the computation of the abnormal returns and the aggregation is done for every possible and reasonable combination of the parameters, the H₀ can be tested for all of these specifications. A first comparison of the different outcomes of the 4,200 observation can be done by
plotting their p-values as in Steegen et al. (2016). This is done also in this study. Further investigation of the differences that may result from the parameters are done by making usage of regressions. Regressions are build up in two ways: The first regression combines all parameters, which are the estimation window length, its distance to the event day and the pre-event window length, as the independent variable and the p-value as the dependent variable, while in the second, the independent variable is always a combination of two parameters that were changed, such that three combined regressions are conducted.

5 Empirical Results

This section is divided in two subsections. The first one deals with the outcomes of the multiverse event study and answers the research question whether insider trading in the data sample influences the stock prices. The second tries to demonstrate the additional information that were gathered from the multiverse analysis and investigates the parameter effects on the p-values.

5.1 Results from the multiverse event study

The study examined the impacts of insider trades among DAX 30 companies on the referring stock prices. The variation of the parameters led to a total of 4,200 observations and, thus, as many single event studies. Indeed, although the sample data was the same for all observations, the empirical analysis of these showed quite different results.

Figure 4 shows the cumulative average abnormal returns for each of the 4,200 observations. Each observation is depicted by one line. The x-axis ranges from -30 to +10 event days as the longest observed pre-event window started at event day -30 days and the post-event window was fixed at +10 days. As the length of the pre-event window ranged between -30 and -20 event days, the CAAR-lines start at these days respectively. Day 0 is the event day, or rather the day of the insider
Figure 4: Cumulative average abnormal returns for all observations

transaction. The figure shows overlapping lines in a darker color than individual lines. As visible from the figure the cumulative average abnormal returns exhibits a similar evolvement but the extent of the CAAR differed across the observations with the variation in the parameters. In general, the CAAR decreases till the event day. Particularly in the 10 days before the event, a large reduction is noticeable from the figure. As the in this time period are darker it can be interpreted as a uniform flow across the observations. After the event, the cumulative average abnormal returns remain stable or slightly increase. In this period, the largest differences across the observations are visible as the lines vary strongly in their extent.

A cutout of figure 4 which considers just the CAAR from five days prior to the event to 2 days after the event, simplifies the interpretation of the effect at the event day (see figure 5). It clarifies that the changes of the CAAR in all observations at the event day were quite small. While in the days before the event a negative increment in the CAAR is given, the CAAR reaches a trough at the event day and remains nearly stable for a few days. As the aim of an event study is to measure
the abnormal returns around the event day, the general negative increment of the average abnormal returns before the event day indicates that the examined insider trades had impacts on the stock prices. The stability of the abnormal returns after the event is expected as the market should quickly incorporate the information and learn from the news. MacKinlay (1997)

Already these figures support Steegen et al. (2016) in her statement that multiverse analyses increase the robustness of the findings and decrease parameter selection biases as the dimensions of the CAAR are different among the observations. Although the figures above indicate stock price changes in connection with insider trades, the significance of the cumulative abnormal returns are uncertain. Until now it cannot be concluded, whether the abnormal returns in the sample are significant or not. Therefore, this is tested with formula 9 mentioned in the model description. For every parameter combination the p-values are calculated and tested with a normaltest at a significance level of 95%. The null hypothesis was not rejected in 3,718 cases. In 482 instead, the null was rejected. 76 observations were even significant at
a confidence level of 99%. Figure 6 depicts the p-values for every observation and outlines the differences. The distribution shows clearly a positive skewness and is not describable with a familiar distribution like the normal or student-t. Interestingly, no constant but rather a wave-shaped reduction in the frequency of observation with the increase of the p-value is given. The p-value distribution states the importance of the multiverse as with a single analysis the null would have been simply rejected or not rejected, while it becomes visible here that the same data set can lead to many different conclusions.

![P-values for all parameter combinations](image)

**Figure 6: Frequency distribution of p-values for all observations**

A classification of the p-values in ranges shows the distribution of the p-values and clarifies the differences across the individual observations. Considering a confidence level of 90% even 875 observations would lead to a rejection of the null hypothesis. (table 3)

As these values vary that much, it requires some more insights to understand which factors drive these variations. These will be covered in the next subsection.
Table 3: Classification of p-values in ranges

<table>
<thead>
<tr>
<th>Interval</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 0.01</td>
<td>76</td>
</tr>
<tr>
<td>0.01 - 0.05</td>
<td>406</td>
</tr>
<tr>
<td>0.05 - 0.1</td>
<td>393</td>
</tr>
<tr>
<td>0.1 - 0.2</td>
<td>557</td>
</tr>
<tr>
<td>0.2 - 0.3</td>
<td>875</td>
</tr>
<tr>
<td>0.3 - 0.4</td>
<td>682</td>
</tr>
<tr>
<td>0.4 - 0.5</td>
<td>254</td>
</tr>
<tr>
<td>0.5 - 0.6</td>
<td>334</td>
</tr>
<tr>
<td>0.6 - 0.7</td>
<td>193</td>
</tr>
<tr>
<td>0.7 - 0.8</td>
<td>118</td>
</tr>
<tr>
<td>0.8 - 0.9</td>
<td>310</td>
</tr>
<tr>
<td>0.9 - 1</td>
<td>2</td>
</tr>
</tbody>
</table>

5.2 Effects from the parameter selection

To find the reasons for the variation in the p-values, and thus, also possible biases in the conclusion if parameters are just arbitrarily chosen, different regressions have been set up. The first regression had the length of the estimation window, the distance of the estimation window to the pre-event window as well as the length of the latter as the independent variable, and the p-value as the dependent variable $y$. In the following regressions, the independent variable is given by a combination of two of the parameters and, therefore was 1) the estimation window length and the distance of the estimation window to the event day and the pre-event window length, 2) the estimation window length and the pre-event window length and 3) the distance of the estimation window to the pre-event window and the pre-event window length. In this way, if $x_1$ is defined as the length of the estimation window, $x_2$ as the distance of the estimation window to the event day and $x_3$ as the pre-event window length, the regressions can be formulated as:
\[ y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \]  

(10)

\[ y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 (x_1 \ast x_2) \]  

(11)

\[ y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 (x_1 \ast x_3) \]  

(12)

\[ y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 (x_2 \ast x_3) \]  

(13)

The ordinary least square method is used for the regressions. To relax the assumptions of independent and identically distributed errors that comes along with the OLS, heteroscedasticity-consistent standard errors according to MacKinnon & White (1985) are applied.

The regression outcomes reveal informational insights. All four models explain the p-values in a good way and are highly significant. In the following, the regressions are described separately before a general conclusion is given.

<table>
<thead>
<tr>
<th>Adj. R-squared</th>
<th>0.727</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>3202***</td>
</tr>
<tr>
<td>coefficient</td>
<td>z-statistic</td>
</tr>
<tr>
<td>Intercept</td>
<td>1.2707</td>
</tr>
<tr>
<td>$x_1$</td>
<td>$-8.912e^{-5}$</td>
</tr>
<tr>
<td>$x_2$</td>
<td>73.611</td>
</tr>
<tr>
<td>$x_3$</td>
<td>0.0459</td>
</tr>
</tbody>
</table>

*** Significant at a confidence level of 99%
The first regression analyses the effect of all parameters, namely the estimation window length, the distance of the estimation window to the event day and the pre-event window length, on the p-value and is the fundamental regression for all of the following. The OLS shows that already 72.2% of the p-values can be explained with this model. With an F-statistic of 3202, the model is statistically significant even on a confidence level of 99%. These facts suggest that already the fundamental regression model describes the variation in the p-values in an appropriate way. Table 4 summarizes the results of the OLS regression.

The second regression considers additionally the impacts of the interaction between the variation of the estimation window length and its distance to the event day. The extension of the regression model leads to a slightly higher explanation in the p-value of 72.9%. Again, the F-statistic is highly significant with a value of 2506. All single variables are significant too. Although, the coefficient of the estimation window and the considered interaction is almost zero and therefore do have little impact on the whole function (see table 5).

Table 5: Results regression 2

| Adj. R-squared | 0.729 |
| F-statistic     | 2506*** |

<table>
<thead>
<tr>
<th>coefficient</th>
<th>z-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.2128</td>
</tr>
<tr>
<td>$x_1$</td>
<td>0.0001</td>
</tr>
<tr>
<td>$x_2$</td>
<td>-0.0136</td>
</tr>
<tr>
<td>$x_3$</td>
<td>0.0459</td>
</tr>
<tr>
<td>$x_1 : x_2$</td>
<td>$1.142 e^{-5}$</td>
</tr>
</tbody>
</table>

Regression three extends the fundamental regression by the interaction of the estimation window length and the pre-event window length as independent variables. In comparison to the first regression, no higher Adjusted R-squared is noticable, which indicates, that the interaction of the estimation window and pre-event window lengths does not add explanatory value to the regression model. However, the model
is still highly statistically significant and explains the variation in the p-values with 72.7%. Considering the variables individually, in this model, the length of the coefficient of the estimation window length and the coefficient of the interaction between the estimation window and pre-event window lengths are not statistically significant anymore, neither at a confidence level of 99%, nor of 95%. The regression output is given in table 6.

Table 6: Results regression 3

<table>
<thead>
<tr>
<th>Adj. R-squared</th>
<th>0.727</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>2452***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>coefficient</th>
<th>z-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.3229</td>
</tr>
<tr>
<td>$x_1$</td>
<td>-0.0003</td>
</tr>
<tr>
<td>$x_2$</td>
<td>-0.0108</td>
</tr>
<tr>
<td>$x_3$</td>
<td>0.0480</td>
</tr>
<tr>
<td>$x_1 : x_3$</td>
<td>-8.427e-06</td>
</tr>
</tbody>
</table>

The last regression tries to explain the p-value as the dependent variable with the three varying parameters and the interaction between the pre-event window length and the distance of the estimation window to the event day. This model reveals the highest degree of explanation of the p-value with an Adjusted R-squared of 73.3%. Also the F-Statistic is highly significant with a value of 3594. Even the consideration of each individual independent variables shows significant statistics at a confidence level of 99%. In this way, it can be concluded that this model is, among the examined opportunities, the best to explain the variations in the p-values even though the coefficient of the interaction is quite small. (table 7)

Generally, all models described were able to explain the variation in the p-value with around 73%, which leads to the conclusion, that the parameter selection has indeed an effect on the outcome of the event study results. Despite the expectations
Table 7: Results regression 4

<table>
<thead>
<tr>
<th>Adj. R-squared</th>
<th>0.733</th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>3594***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>coefficient</th>
<th>z-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.0303</td>
</tr>
<tr>
<td>$x_1$</td>
<td>-8.912e-05</td>
</tr>
<tr>
<td>$x_2$</td>
<td>-0.0225</td>
</tr>
<tr>
<td>$x_3$</td>
<td>0.0363</td>
</tr>
<tr>
<td>$x_2 : x_3$</td>
<td>-0.0005</td>
</tr>
</tbody>
</table>

of the largest impact of the length of the estimation window on the p-values, this variable shows coefficients of nearly 0 in all models and is even insignificant in the third regression (see table 6). The distance of the estimation window to the event day instead is significant in all of the four models and has negative coefficients. As the input variables of the distance in the regression were negative too, the relationship becomes positive and the model suggests, that the longer the distance of the estimation window to the event day, the larger the increase in the p-value. Referring to figure 1, the further away $T_1$ is from $\tau$, the higher the p-value. The largest coefficients in all regression models were detected in the pre-event window length with around 0.05. Also the pre-event window length has negative input variables in the regression, which leads to the finding that, in conjunction with the positive coefficient, a longer pre-event window, ceteris paribus, causes the p-value to decline. The interaction coefficients are nearly 0 in all regressions. On top, the third regression in general, and also the interaction variable there, have not been statistically significant. Although coefficients were small, the last regression, that examined additionally to the fundamental model the effect of the interaction of the pre-event window length and the distance of the estimation window to the event day on the p-value, caused the highest Adjusted R-squared and was therefore able to explain the variation in the p-values in the best way. The interaction of these variables have a negative relationship to the variation of the p-value. Also the interaction
of the estimation window length and its distance to the event day is statistically significant. Again, as the distance to the event day has negative input variables in the regression model, the interaction of these variables has a negative relationship on the extent of the p-value. Table 8 summarizes the impact of the explanatory variables on the p-value:

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Relationship to p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation window length</td>
<td>-(^2)</td>
</tr>
<tr>
<td>Distance of the estimation window to the event day</td>
<td>+</td>
</tr>
<tr>
<td>Event window length</td>
<td>-</td>
</tr>
<tr>
<td>Estimation window length : Distance of the estimation window to the event day</td>
<td>-</td>
</tr>
<tr>
<td>Estimation window length : Pre-event window length</td>
<td>+(^3)</td>
</tr>
<tr>
<td>Distance of the estimation window to the event day : Pre-event window length</td>
<td>-</td>
</tr>
</tbody>
</table>

\(^2\) Regression 2  
\(^3\) not significant
6 Discussion of this thesis’ multiverse analysis

The results from the previous section make clear that various selected parameters may lead to various results. They support Steegen et al.’s 2016 findings, that p-values vary with the input variables and that the overall conclusion can be completely different if a multiverse analysis is conducted. This is underlined with the most important insight from the results:

Within this data set, researchers could have chosen 482 different parameter combinations and gotten significant p-values that would have enabled them to reject the null hypothesis (at a significance level of 95%). In all these cases it would be concluded that insider trading clearly impacts stock prices. On the other hand, 3,718 parameter combinations would lead to the result that insider trading does not affect stock prices as the null could not be rejected.

Whether intentionally or unintentionally, this could bring the desired outcome and biases the whole research and is in line with Thompson (1995). As in single analysis only one outcome is published, biased outcomes from the parameter selection, may distort the external perception in regards to the topic (Steegen et al. 2016). Especially the distribution of the p-values can solve this problem as a reader can see the frequencies of the rejection/non rejection of the null hypothesis at the first sight. Additionally, regressions can bring further insights to the importance of the individual parameters and improve the transparency of the results. Even though the multiverse analysis does not allow to answer the research question unambiguously, it enables to get a much more concrete view of the general topic and the factors that influence it. Rather than just saying whether there is an observable effect or not it makes it possible to describe the effect depending on the parameters that were selected or other choices like the exclusion of some variables that were made. It enables to draw a broader picture and reduces the possibility of outliers, such that the overall conclusion is more reliable. Furthermore, it increases the transparency in conjunction with the model’s set up and allows other researchers to verify the
results in a simple way. If event studies or other empirical approaches would always be built up on the basis of a multiverse, comparison between different studies could be simplified and, probably, drawing general conclusions would be possible in some cases.

As a multiverse analysis has the aim to provide results for all possible and reasonable combinations of the parameters, it could be criticized, that even in this study were made choices of these variables. For reasons of simplicity and speed of computation, the parameters were varied in intervals of some days. Nonetheless, it is expected that results will not change much if the estimation window length, the distance of the estimation window to the pre-event window and the pre-event window length are altered on a daily basis. Moreover, in this study, the length of the event itself was limited to one single day, the day of the insider trade itself. In a further step this could be varied too. Also, the post-event window length is often part of an event study but was fixed to 10 days in this approach. The discussion about the most appropriate model takes place since decades and could impact the results of the event study too. Therefore, a more complete view could be gathered if, additionally to the parameters that were modified in this thesis, the length of the post-event window and different models would be included. Regardless these extensions, in this thesis was shown that the chosen parameters can make a large impact on a single statistical result and that a multiverse can improve the robustness of the inferences, which points out that there are numerous advantages of multiverse analyses that outweigh some potential shortfalls like a larger effort in computing, which can be solved with modern software tools and the appropriate consideration of all possible and reasonable parameter combinations. Investing time in the model preparation and thinking through the special characteristics and requirements of the data set and the model enables here to get valuable information and open new perspectives of empirical insights, wherefore the multiverse approach should be considered as a powerful tool in the empirical analysis. Moreover, as outlined in section 2, the debate about the effect of insider trading among the literate endures already
decades and studies revealed various results, a re-examination of these studies could lead to additional insights although the data sets remain the same. In this way, the multiverse analysis could help to get a more uniform picture of the topic.

7 Conclusion

The purpose of this thesis was to answer the question whether insider trading does affect stock prices. Moreover, a multiverse approach should point out the advantages of this approach and to test whether arbitrarily parameter selection in traditional analyses can impact the outcomes. This approach was used with regard to prior studies that used similar methods but revealed contrary results. In this manner, the multiverse approach was an attempt to get more reliable and transparent results. For the conduction of the thesis, insider trades among the DAX 30 companies from February 2017 to February 2021 and the corresponding index and company returns were used. 554 insider trades among 27 companies were identified in this time horizon. The extension of the event study approach by a multiverse analysis, made it possible to change all input parameters that generally are chosen arbitrarily and thus, to increase the reliability and the power of the results. The parameters that have been changed are the estimation window length, the distance of the estimation window to the event day and the pre-event window length. In this manner, these parameters could have been combined such that $50 \times 14 \times 6 = 4,200$ different observations (and also event studies) were conducted.

As the aim of an event study is to measure the abnormal returns around the event, i.e. insider trade, the cumulative average abnormal returns gave first insights. The development of the CAAR across the observations was the same, but their extent differed. A sharp negative increase in the CAAR a few days prior to the event indicated an effect of the insider trades on the stock prices. This effect was tested to get answers to the research question. The results clearly showed variations in the p-values across the observations with different parameter combinations. The
frequency-distribution of the p-values was strongly positive skewed and did not follow a familiar distribution. Instead, the frequencies of the observed p-values were wave-shaped. At a significance level of 95%, the null hypothesis was rejected in 482 cases. 76 of these were even significant at a significance level of 99%. For 3,718 observations instead, the null hypothesis could not have been rejected at these significance levels. As the sample data was the same for all observations, it might seem straightforward that the differences must be an effect of the variations in the parameters. Therefore, regressions were made to check, whether the sizes of the p-values can be explained by the variables or rather by a certain variable. For this aim, a first regression checked the relationship between the p-value and the overall of the parameters which were the estimation window length, the distance of the estimation window to the pre-event window and the pre-event window length. Afterwards, regressions with the p-value as dependent variable and each combination of two parameters as the independent variable were set up. The analysis of the regressions showed clearly that a large part of the variation in the p-values can be explained by the parameters. The extension of the independent variables by the interaction of the estimation window length and the distance of the estimation window to the event day as well as the interaction of the latter and the pre-event window length increased the explainable part of the p-value. In this manner, the thesis states that the parameter selection indeed has an impact on the outcome of the event study results, the extent of the p-value and the conclusion that researchers draw out of it. Thus, it supports the relatively new research method of the multiverse analysis that was introduced by [Steegen et al., 2016] and recommends this approach for future studies. In this event study, the multiverse analysis enabled to gather more precise insights in the effects of the parameter selection and depicted significant differences in the results that could lead to various inferences, either the rejection of the null or the non-rejection. Moreover, it becomes clear that a multiverse analysis can prevent such biased conclusions and opens new possibilities in the research. Based on these insights, a re-examination of prior literature regarding insider trading could provide
further information and could lead to a different general view on this topic.
References

URL: https://www.sciencedirect.com/science/article/pii/S092753980900005X

URL: https://www.sciencedirect.com/science/article/pii/0304405X85900108


URL: http://www.jstor.org/stable/j.ctt7skm5

URL: http://www.jstor.org/stable/3666164


URL: https://www.sciencedirect.com/science/article/pii/0304405X82900046

URL: http://www.jstor.org/stable/2328976


URL:
https://ideas.repec.org/p/tiu/tiucen/0d74f29a-ed4f-4c26-9c55-9eac788ab845.html


URL: https://deutsche-boerse.com/dbg-de/media/deutsche-boerse-spotlights/spotlight/DAX-Index-Benchmark-und-Barometer-f-r-die-deutsche-Wirtschaft-148654


URL: https://doi.org/10.1145/3290605.3300295


URL: https://doi.org/10.1177/0956797612466416


URL: http://www.jstor.org/stable/2525569


URL: https://doi.org/10.1177/1745691620917678

URL: http://www.jstor.org/stable/253304


URL: http://www.nber.org/papers/w6656
URL: http://www.jstor.org/stable/724969

URL: http://www.jstor.org/stable/2729691

URL: https://www.sciencedirect.com/science/article/pii/0304407685901587


URL: https://doi.org/10.1177/109442819924002


URL: http://www.jstor.org/stable/40472954

44


URL: [https://doi.org/10.1086/712139](https://doi.org/10.1086/712139)


URL: [https://www.sec.gov/Archives/edgar/data/25743/000138713113000737/ex1402.htm](https://www.sec.gov/Archives/edgar/data/25743/000138713113000737/ex1402.htm)


**Significance Tests for Event Studies** (2021).

URL: [https://www.eventstudytools.com/significance-tests](https://www.eventstudytools.com/significance-tests)


Smith, F. P. (1941), ‘Management-trading and the stock-market profits’.


URL: https://doi.org/10.1177/1745691616658637


URL: https://www.sciencedirect.com/science/article/pii/S0927050705800738