Erasing Blind Spots

A data-driven evaluation of model overrides in case of corporate events

by

R.B. Steenbergen

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday 12 November 2020 at 3:00 PM.

Student number:	4377389	
Project duration:	1 March 2020 – 12 November 2020	
Thesis committee:	Prof. dr. ir. C. W. Oosterlee,	TU Delft
	Dr. P. Cirillo,	DIAM
	Msc. S. Schneider,	Robeco

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Abstract

Currently, quantitative asset pricing models are often not equipped to deal with merger and acquisition events. In such cases, portfolio managers make the assumption that the model is not working and they override its decisions for an entire year. This thesis studies the performance of quantitative models after these events and provides research based guidelines for future decisions on model overrides.

A selection of investment factors representing the model will be used to explain the post-event abnormal returns. In general, the data consists of monthly observations on one cross-section of firms over a time period of several years. Consequently, it may contain unobserved firm or time effects. Ignoring such effects leads to inefficient estimates and biased results. Hence, robust inference is conducted by adjusting the standard errors of a pooled regression, and through modelling the effects in a panel regression.

Multiple approaches show that the separate factors are not sufficiently able to explain the abnormal returns after an event, if compared to the model's performance in regular market circumstances. Consistent results for the post-event performance are particularly hard to find. Equally weighting all factors in a single regressor, leads to model performance being equivalent to that in non-event times after 9 months, which indicates that a override period of equal length is sufficient. Therefore, the current assumption that the model is not working properly after a merger or acquisition, is correct. The robust pooled model is favoured over the panel model in this research, due to its low complexity and its straightforward results.

Acknowledgements

First of all, I wish to express my appreciation towards my supervisors at Robeco, Sebastian Schneider, Iman Honarvar and Matthias Hanauer, for their great guidance during the 6 months of my internship there. Even though the circumstances were not optimal, they were there for me whenever support was needed. I am grateful for the opportunity to have been a part of the Quant Research team at Robeco.

I must convey my gratitude towards Dr. P. Cirillo for his instructions, time and feedback during the first part of my thesis. Besides, I am very thankful towards Prof. Dr. Ir. C.W. Oosterlee, who offered to take over the academic supervision halfway through my research. I want to express my gratitude for his guidance, for reviewing my work and for sending me valuable feedback.

Special thanks to my parents, brother and girlfriend for always believing in me and creating an ideal environment for me to pursue my goals. Without your unconditional support it would have been a lot harder to accomplish this.

Finally, I would also like to thank my dear friend Sam de Boer, for being such a great study companion. Thank you for always challenging me to get the best out of myself and providing help whenever needed, during high school and foremost, during my entire period at the TU Delft.

Contents

1
Introduction

Over the course of the past decades, researchers have been attempting to capture the return of stock prices in a wide range of models. The classical Capital Asset Pricing model (CAPM) of Sharpe [48] and Lintner [37], which linearly relates returns to market risk, has been replaced by models that incorporate additional risk factors that relate to stock returns. These factors are firm characteristics related to a group of stocks that have been proven to be drivers of excess returns in comparison to the market. Based on such quantitative asset pricing models, asset managers decide which stocks to buy or sell and herewith try to harvest the factors' risk premia. As an innovative dutch asset manager, Robeco started developing its quantitative models over 25 years ago. These models are based on in-house research and hold a scientific approach towards investing.

Quantitative models are developed to predict stock returns, but are not necessarily build to take into account all situations that might occur in the market. Think of regulatory events that cause shocks in the cross-section of stocks, or government bailouts in times of crises. In corporation with Robeco's portfolio managers (PMs), it has been decided to focus on events of the merger and acquisition (M&A) type. In a merger, two different companies or legal entities are consolidating into a single firm or entity. An acquisition is the type of event where one company is taking over the ownership of another firm. Since this generally leads to the consolidation of the firms' components as well, the two types of events are bundled into the single designation of M&A. These M&A events are the most interesting ones, due to their frequency and impact on the market.

In case of mergers and company take-overs stock returns might be driven by completely different processes than the well-known risk factors. For this reason, the Quant Research team at Robeco has always assumed that their model was blind towards such M&A events, meaning that the factors insufficiently explain the realized returns. This assumption ultimately led PMs to manually override the model's decisions. From the moment that an M&A deal is completed, a neutral position in the firms involved in the deal will be taken for a period of 12 months subsequent to the event. Herewith, buying these stocks up till benchmark weight. In case the factors retain their explanatory power, it could imply that Robeco has missed out on some valuable opportunities over the past years.

This thesis will form the foundation for future decisions regarding the procedures on firms involved in an M&A process. We will consider a simplified version of the quantitative model that Robeco's Quant Research department is using, by selecting a number of important and widely used investment factors, henceforth referred to as 'the model'. The aforementioned factors will be used in a regression model in which they are linearly related to firm specific returns. For each firm in our investable universe, defined as the combined MSCI World Index and MSCI Emerging Markets Index, stock and news data is gathered over a period spanning January 2000 to December 2018. Even though firms may enter or exit the universe, the data is generally collected over the same firms across time. Combined, the time series and cross-sectional data form a data set, called a panel. In general, parameters of models using panel data can be estimated with a pooled regression or a panel regression. However, collecting data of the same firm over a time span can lead to heterogeneity in the data. These so-called unobserved time and firm effects are not uncommon in financial panel data, and may lead to biases in the standard errors, and additionally inefficient estimates of pooled regressions. Nonetheless, researchers often choose to ignore the unobserved effects or do not correctly adjust the standard errors for the effects that are present. In this thesis, two main methods are used that do account for this heterogeneity. These methods are pooled regression with robust standard errors, and panel regression modelling the unobserved effects when present.

This thesis aims to provide Robeco's Quant Research team with a robust and research based answer to the following question:

- Does the model still work in case of an M&A event?

Naturally, a negative outcome would give rise to the question:

- How long should the override period be?

The structure of the thesis is as follows. Chapter 2 describes the data that is used to assess whether the model works properly after an M&A event. First, the universe of stocks is defined, after which an elaboration on the firm characteristics and explanatory variables follows. The second part of this chapter provides information on the news data. As the RavenPack news analytics database is used for the identification of events, a comprehensive introduction to the filters and data selection process is provided. Hereafter, Chapter 3 treats the abnormal performance framework that is typically part of the event study methodology. In Chapter 4, the mathematical tools are provided that will be used for our research. Among other things, the pooled model and panel model will be presented. In Chapter 5, the results of both regression methods are provided and compared. Finally, the conclusions and suggestions for further research are stated in Chapter 6.

2
Data

The data used in this research can be categorized into news data and stock data. In this chapter, we will start by briefly covering the universe of stocks that we are studying. Hereafter, the stock data and the explanatory variables of the regression models will be treated. The second section will provide information on RavenPack, which is the news analytics platform used for the collection of M&A events.

2.1. Investable universe

This research is restricted to companies in the MSCI World Index and the MSCI Emerging Markets Index in the period from January 2000 until December 2018. Both indices capture the large and mid-cap funds of over 20 countries and are set up to cover approximately 85% of the free-floating market capitalization in each of these. The MSCI World Index, henceforth referred to as DM, has approximately 1,600[1] constituents from 23 developed countries, including Australia, Canada, Hong Kong, Israel, Japan, New Zealand, Singapore, the US and Western European countries [43]. The US is by far the biggest contributor to the index, making up about 65% of its weight. Japan and the UK follow with a weight of 8% and 5%, respectively.

The MSCI Emerging Markets Index, hereafter referred to as EM, has over 1,100[2] constituents from countries that are transforming from a less developed economy towards a more industrial one with high living standards and accessibility for foreign investors. Among these are countries from Asia, Eastern Europe, Latin America and the Middle-East [44]. The largest contributors to the EM index are China, Taiwan and South Korea, with a share of approximately 40%, 13% and 12%, respectively. The indices are rebalanced semi-annually, meaning that our universe slightly changes twice a year, by firms that exit or enter. For each firm in our universe, monthly return data and firm characteristics have been retrieved from financial data provider FactSet. All monetary values in the data are expressed in US dollars.

2.2. Stock data

In this thesis, we are studying whether proven drivers of excess returns are still able to explain the abnormal returns after an M&A event. For this we need to identify what the common drivers of these returns are. In the 1960s, the CAPM was developed by Sharpe [48] and Lintner [37], which is an example of a classical asset pricing model. It linearly relates the expected return with market risk. The expected return of a stock is calculated as:

$$E[R_i] = R_f + \beta_i (E[R_m] - R_f), \tag{2.1}$$

where R_f is the risk-free rate, $E[R_m]$ the expected return of the market portfolio and β_{it} is a metric for the riskiness of an asset, as it measures how much risk a stock will add to a portfolio which is similar to the total market. When the stock has a higher risk than the market it will yield a $\beta_i > 1$ and vice versa for funds less risky than the market. Therefore, CAPM establishes a positive relationship between risk and return. This model has been debated as market risk turned out not to be the only factor driving stock returns.

Factors are characteristics of groups of stocks that are related to their risk and return. Fama and French [22] showed with their three-factor model, which will be more extensively treated in Chapter 3, that returns

[1] Numbers are as of April 2020, and are used to give a representation of how the index looks like in general [43, 44].

[2] See footnote 1.

are additionally driven by firm size and book-to-price ratio. Over the past decades, many more factors have been 'discovered' and the total set is sometimes mockingly denoted as the 'factor zoo'. Since it is not desirable to include all of these, we identify six important and well-known equity factors: Momentum, Quality, Reversal, Risk, Size and Value. The following subsection will provide a brief overview of each factor with a description of the corresponding strategy or rationale, the expected signs in a linear regression and the metrics that are used to capture the factor performance of a company.

2.2.1. Firm characteristics

Momentum strategies try to capture excess returns by buying past winners and selling past losers. Differently stated, a company with a higher Momentum score is expected to generate higher abnormal returns. As a consequence, the anticipated sign of the regression coefficient is positive. This factor can be split up into two different types, namely Price Momentum and Earnings Momentum. The well-known paper of Jegadeesh and Titman [32] studies Price Momentum and defines the winners as those companies with strong past returns and the losers as those companies with poor past returns. They form the Momentum variable by taking a firm's returns over the past 12 months and excluding the most recent one for short-term reversal effects. Carhart [17] expanded the Fama-French model to a four-factor model [3], by including a variable based on their measure for Momentum as an additional explanatory variable. Although found to be less strong, Earnings Momentum is still significant in cross-sectional regressions of future returns [18]. The effect is based on earnings surprises, which can be measured by means of analyst revisions of earnings forecasts.

The Quality factor is relatively new in the field of quantitative investing [3]. It is aimed at capturing excess returns of stocks with high quality relative to the broad market. Consequently, the expected sign of the regression coefficient is positive. There is a wide range of variables that can be used to measure the quality of a stock, such as profitability, earnings stability or solvency [3]. Our focus is on the profitability of a company, which can be captured by a firm's return on equity (ROE) and gross profit. The ROE ratio shows how much money the company made over the fiscal period for each dollar invested by the shareholders. Gross profit is calculated by subtracting the cost of goods sold from the revenue. For financials this metric is not defined, as the cost of goods sold is not specified. For such companies, the Quality factor will only be measured by ROE.

Reversal, or as mentioned in the Momentum description short-term reversal, is the phenomenon of abnormal returns in the upcoming month by stocks that have performed relatively poor during the last one [33]. As a consequence, the regression coefficient is expected to be negative. This strategy has a relatively low persistence, since longer term past returns are expected to have a positive influence and, thereby, contribute to the Price Momentum anomaly. Thus, a Reversal strategy requires regular updating of the portfolios to harvest excess returns.

The Risk factor, also known as low volatility, is a somewhat counter intuitive one. One of the most regularly assumed principles in finance is that higher risk is associated with higher returns. CAPM, developed by Sharpe [48] and Lintner [37] and provided in Equation (2.1), is an example of an asset pricing model which linearly relates the expected return with market risk. Among others, Blitz & van Vliet [12] found that low risk stocks earn higher risk-adjusted returns than the market portfolio and the effects of the Risk factor are of comparable magnitude to the classical ones in for example Carhart's four-factor model [17]. The metrics that we will be using to identify the low risk stocks are the historical volatility with 1-year and 3-year look back periods, and CAPM β measured against the corresponding index (either EM or DM). Since a lower risk will lead to higher returns, the anticipated sign of the regression coefficient is negative.

The last two factors are the ones that, besides the market risk premium, constitute the Fama-French three-factor model [22]. First discovered by Banz [7], and subsequently confirmed by other researchers, the Size factor captures a small cap premium. That is to say, firms with relatively small market capitalization earn excess returns relative to their large counterparts and, thus, the expected sign of the regression coefficient is negative. The measure for this factor is the market capitalization in the indices[4], which are in the case of EM and DM free float-adjusted. Common practice is to take the logarithm of the market capitalization, which is possible as this measure is positive by definition, and so shall we precede in this thesis as well.

The final factor is Value. Value strategies consist of buying stocks which are cheap compared to their fundamental value and selling stocks which are relatively expensive. The value of a company can be captured with multiple fundamental variables, such as book value, sales or earnings. The 'cheapness' is measured relative to the stock's market price. Naturally, it arises that the anticipated sign of the regression coefficient

[3] More extensively treated in Section 3.1.2.
[4] Measured in thousands of dollars.

is positive. Variables that are commonly used and have been found to work are book-to-price, used by Fama and French [22] in their three-factor model, and earnings-to-price ratios [9, 47].

2.2.2. Explanatory variables

The linear models, treated in Chapter 4, use the factors in a regression framework and linearly relate them to the abnormal returns. Part of the previously treated factors are formed by the average of multiple firm characteristics, for example, Value consists of book-to-price and earnings-to-price. These metrics are not necessarily in the same order of magnitude, which is why taking equally weighted averages possibly erases the implication of one characteristic. Furthermore, the observations of firm characteristics are spread over the entire time period. Pooling firm characteristics from different parts of the economical cycle in a single model can yield undesirable situations. An example can be given in terms of the Momentum factor. Returns in periods of crises can be very low, compared to the average in times of economic prosperity, resulting in moderate Price Momentum in absolute terms. In such cases, firms that are outperforming their peers on this factor might still be attractive to investors. When pooled, a raw Momentum score will not be able to capture these relative winners.

A general procedure for standardization and labelling outliers is to compute z-scores. This value subtracts the mean of a random variable and hereafter divides by its standard deviation. However, this z-score might be heavily affected by outliers. Iglewicz and Hoaglin [30] show that in a worst case scenario, the maximum z-score does not depend on the data values anymore, but merely on the number of observations. Therefore, a robust method of standardisation is preferred in order to safely compare and average the firm characteristics. This is done by taking modified z-scores, computed with the following formula [30]:

$$z_{it} = \frac{x_{it} - \tilde{x}}{1.483 * median_{it}(|x_{it} - \tilde{x}|)}, \tag{2.2}$$

where $\tilde{x}$ is the median. The modified z-score is a so called resistant estimator. By taking medians rather than means, it will take up to 50% of the data to be replaced by any randomly selected number until the estimated value can become infinite [30]. The second term of the numerator is called the median absolute deviation or MAD, which is the median of the absolute deviations from again the median. The z-score shows how much a characteristic differs from a typical value. Since we will relate the firm characteristics or factors to the abnormal returns, it is crucial that the medians in the computation of the z-scores are computed for the same benchmark. A typical rule is to label observations with a z-score larger than three in absolute value as an outlier. By company policy, z-scores that are larger (lower) than three (minus three) are set to three (minus three).

Besides the separate factors, a combined Quant Research (QR) factor will be used to explain the abnormal returns. The factor is computed using the following formula:

$$QR_{it} = \frac{1}{3}(z_{Momentum,it} + z_{Quality,it} + z_{Value,it}) - \frac{1}{3}(z_{Reversal,it} + z_{Risk,it} + z_{Size,it}) \tag{2.3}$$

In this equation, the values of the z-scores are the averages of the metrics corresponding to the factors. The signs are determined by the expected effect of the single factor on the abnormal returns, as treated in Section 2.2.1. As a result of the cap and floor of the individual z-scores, the composite factor will be between minus three and three. Equivalently to a stock ranking model, the QR factor will sort the stocks from best (highest QR) to worst (lowest QR) based on their average performance on the selection of factors above.

2.3. News data

For the identification of events, we are using the RavenPack database. RavenPack is a leading data analytics provider with a News Analytics platform that gathers news from over 22,000 sources, including premium news providers, regulatory newswires, press releases and social media. Its services are able to detect about 98% of the companies in the entire investable universe and their database consists of news articles from January 2000 onward. RavenPack tags relevant information in the article such as topics, company or entity and assigns scores all in a matter of milliseconds, enabling users to access the processed news data in real-time.

Articles on M&A almost always include multiple entities. RavenPack's classification system uses a role detection step, which deduces the role of each of the entities in the articles. As such, both sides of a transaction will be tagged in separate categories[5]. Each company has a unique ID consisting of a six digit alphanumeric

[5] For example: "Company A completed acquisition of Company B". This headline will yield a tag of Company A in the category acqui-

combination mapping into commonly used identifiers. In this study we use SEDOL (Stock Exchange Daily Official List) codes to link the news data from RavenPack to the companies in our universe. This mapping is not one-to-one, as multiple SEDOL codes can map into the same RavenPack identifier. Such cases occur for certain subsidiaries of large corporations or Class A and C shares of the same company. The irrelevant entities are filtered out by the fact that we only focus on companies in the DM or EM index and through the application of relevance scores in the initial data retrieval.

2.3.1. Category Selection

RavenPack uses a four-layer taxonomy tree to identify each event, starting with five classifiers at the least granular 'Topic' level: Business, Economy, Environment, Politics and Society. From these five we will merely focus on Business, as this topic maps into 26 'Groups' among which the one of our interest: 'Acquisitions Mergers'. Going further down the tree there is first the 'Type' layer, separating mergers, acquisitions and unit acquisitions for example. The final branch of the tree, under the 'Acquisitions Mergers' group consists of 32 'Categories'.

For this study, it is essential to select the categories that capture the exact event in which a merger or acquisition is announced. This moment will mark the point in time where the PMs override the model. That is to say, rumors, opinions, regulatory scrutiny, etc. should be excluded as PMs will not undertake any action yet. Furthermore, unit acquisitions are treated differently for either side of the deal. For acquirers of business units PMs will override the model just as with usual acquisitions, while this will only happen for acquirees if more than 50% of the shares are being sold. Unfortunately, RavenPack does not gather data on the amount of shares included in a transaction. Hence, unit acquisitions are only considered from the acquirer's point of view.

On the preceding criteria, the following categories have been selected: acquisition approved acquirer, acquisition approved acquiree, acquisition completed acquirer, acquisition completed acquiree, unit acquisition approved acquirer, unit acquisition completed acquirer, merger approved and merger completed.

2.3.2. RavenPack filters

By selecting the categories as specified above, RavenPack is able to identify a total of 824,563 events over our 18-year time window. This data is refined by applying the right filters.

First of all, some obvious filters need to be set. The entity type is set to company and the Fact Level is set to Fact. Secondly, we filter on relevance. This stage can be broken down into two components: Event Relevance and Entity Relevance. Event Relevance is reflected by an integer, increasing in relevance, ranging from 0-100. The earlier the first mention of the event and the higher the number of mentions, the higher the score. If the event is mentioned in the headline it will receive a score between 90-100, depending on the rank of the mention. When the event is mentioned in the first two paragraphs it will receive a score ranging from 80-89. Events in the remaining part of the article will receive a score lower than 80, accordingly. The Entity Relevance score is constructed in a similar fashion. An entity gets a relevance score of 100 if it is found to play a key role in the first event of the headline in the article and if that role is not a role that is of lower relevance, such as a rater. Otherwise, depending on the importance of the reference, its location and the number of mentions in the remaining part of the article, an entity can be rewarded a score ranging from 0-99. By requiring both relevance scores to be greater or equal than 90, we are left with news articles where the event and the entity are at least mentioned in the headline and the tagged entity is highly likely to play a key role in the first event detected.

The remaining data set includes similar articles on the same entity from different news sources. By applying RavenPack's novelty filter, called Event Similarity Days (ESD), these should be excluded. The ESD score is a number ranging from 0.00000 [6] up to and including 365, indicating the amount of days between the article that has been graded and an article with a similar event for the same entity. A score of 365 indicates that at least one year has past since such an article was detected. Using the maximum score of 365 days, it is more likely that all the articles contain a completely new event.

2.3.3. Post processing

The entity filter in RavenPack is inconvenient to filter out the events for the firms in our universe, since we would have to specify the period that they are included for each firm separately. Therefore, we match the news

sition completed acquirer, while Company B will be tagged in the category acquisition completed acquiree. Both records will have the same underlying article.

[6]Including five decimals to show that the time between similar articles can be measured up to milliseconds.

data with the companies in EM and DM afterwards and drop all other observations. Although RavenPack is one of the leading news analytics providers, its algorithms are not 100% waterproof and similar events on the same company may still occur in the data set. Additionally, some of the events are put in the wrong category. Therefore, a final evaluation of the remaining data set is done by hand. All headlines have been evaluated on their tagged category, tagged firm and similarity. In some cases the category and firm are incorrectly tagged and adjusted accordingly. If necessary, the observation is removed, for example when the observation is already in the data set or has nothing to do with the tagged firm/category. Each remaining news headline has been transformed into a data point, contained in an event matrix denoted by ED. The rows represent the firms and the columns denote the months, spanning January 2000 to December 2018. Each entry of this matrix is defined as follows:

$$ED_{it} = \begin{cases} 1, & \text{if an event happened for firm } i \text{ in month } t \\ 0, & \text{otherwise.} \end{cases} \tag{2.4}$$

2.3.4. Filtered data

The final news data set consists of 8,170 events, with 7,362 (89%) of those happening in DM and only 802 (11%) in EM. Over the course of the years some countries have entered and left the indices, making a total of 26 countries in DM and 27 in EM. Table 2.1 shows the top 10 most frequently occurring countries in either index. Interesting to notice here is the overweighting of anglophone countries compared to their weights in the indices. In DM, Canada and Great-Britain overtook Japan, which is the index' runner up in terms of size. We observe a similar phenomenon in the EM index. India, where English is still one of the official languages, caught up on China and also South-Africa entered the top five countries with most events. This indicates the tilt of RavenPack towards English new sources. Furthermore, the total number of events is spread over 2,132 firms of which nine out of the top 10, by occurrence, come from the US. The most frequently observed entities are General Electric, IBM and Cisco.

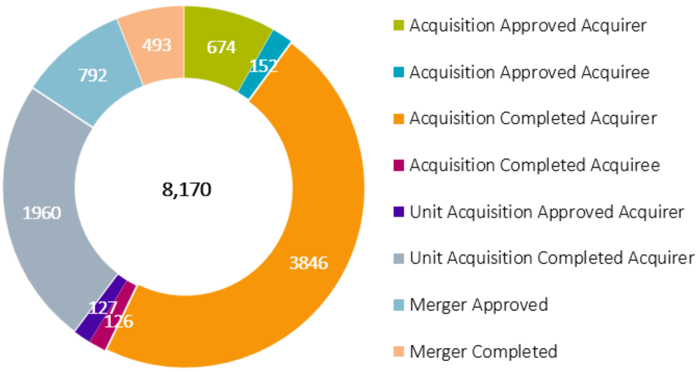
Table 2.1: Number of events per country

	DM		EM	
1.	US	4,019	IN	169
2.	CA	574	BR	82
3.	GB	491	CN	73
4.	JP	415	ZA	63
5.	FR	362	KR	60
6.	DE	270	MX	46
7.	AU	193	MY	43
8.	SE	153	RU	41
9.	CH	141	TW	37
10.	NL	128	IL	37

The table contains the top 10 countries in terms of M&A events over the period 2000-2018 in the DM and EM indices.

Figure 2.1 contains the number of news events per category. We observe that the categories from the acquirer's perspective are significantly larger than those from the acquiree's perspective, having 7,892 and 278 observations, respectively. Relating this to the role detection system applied by RavenPack, it might be that RavenPack is better at detecting acquirers rather than acquirees. Ordered by size, the categories acquisition completed acquirer, unit acquisition completed acquirer and merger approved are the largest, together accounting for approximately 80% of the events.

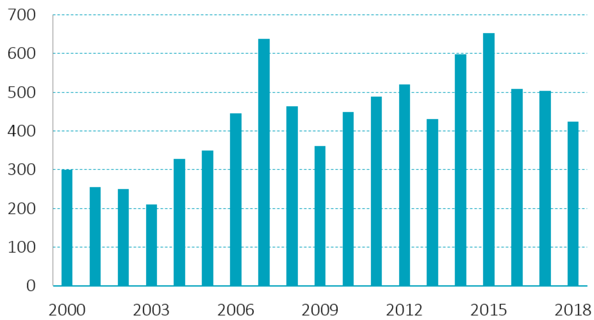
Figure 2.1: Events per category



The figures graphs the number of events in each RavenPack category.

The distribution of events over the years is plotted in Figure 2.2. It is interesting to notice that during times of financial turmoil, the M&A activity seems to be relatively high. The graph shows peaks during the dot.com bubble in the early 2000's and during the late zeros global financial crises. In general the number of events has increased as well, which has two main reasons. First of all, RavenPack improved its data collection systems and secondly an increasing number of articles have become available online [20].

Figure 2.2: Distribution of events over time



The figures contains the distribution of events over the years.

3
Abnormal returns framework

Over the past decades, the event study approach has become a popular method to measure security price reactions after corporate events. The battery of tests that is available, to assess the significance of this reaction, will not be treated in this thesis. Yet another application is to use firm characteristics in a regression framework to explain this price reaction [35]. This testing of firm characteristics is even relevant in cases where the average move in stock prices is zero, which makes its application particularly suited for this thesis. Our focus is on this part of the event study. To that end, this chapter is dedicated to the construction of the abnormal returns, by providing the generally used framework in the event study methodology literature.

3.1. General framework

A vast body of literature has been published on event study methodology. The proposed structures in the methodology literature are quite similar, with at the basis of every framework the construction of an abnormal return. In his work Thompson [50] distinguishes between two forms of abnormal performance phenomena: the announcement effect and the valuation effect. The first can be seen as a price change resulting from the market's expectation of the change in value of the company, while the valuation effect is the actual price change of the security due to the event happening. Based on the efficient market hypothesis, the first effect should be captured in the first days or even hours after an event. The valuation effect, which is what we are focusing on, is measurable on a longer term. The methodology literature distinguishes between short and long-horizon studies for the computation of abnormal returns. In the following subsections both will be treated separately.

3.1.1. Short-term returns

Short term abnormal results, applicable for periods of several days to maximally a year after or centered around the event, can be computed with the well-established methods of Brown and Warner [13]. Herein, the abnormal return is the return in excess of the normal or benchmark return or differently formulated, the abnormal return is the return conditional on the event minus the expected return unconditional on the event. The model for the normal or expected return should be constructed in such a way that the expected value is zero and the variance of the abnormal return should be reduced. Three models are proposed that satisfy the first condition of unbiasedness, being: mean-adjusted returns, market-adjusted returns and market and risk-adjusted returns. In terms of variance in the abnormal returns, the mean-adjusted model underperforms the market model and is therefore out of scope for the remaining part of this section [11]. The two remaining models both form a collection of approaches. In the market and risk-adjusted return models, the abnormal return is given by the residual that results from the regression of a specific model. A frequently used approach is the market model, given by:

$$R_{it} = \alpha_i + \beta_i R_{mt} + \epsilon_{it}, \tag{3.1}$$

such that,

$$AR_{it} = \epsilon_{it} = R_{it} - \alpha_i - \beta_i R_{mt}, \tag{3.2}$$

where AR_{it} is the abnormal return of stock i in month t, R_{it} is the raw return, R_{mt} the return of the market or benchmark portfolio, and $\hat{\alpha}_i$ and $\hat{\beta}_i$ are the OLS estimates of the regression coefficients. By substituting

$\hat{\alpha}_i = 0$ and $\hat{\beta}_i = 1$, one obtains the market-adjusted return model as the simple difference between the raw return and the market return. According to Chandra, Moriarty and Willinger [19] the market-adjusted return model and the market and risk-adjusted model perform similarly, since the latter's greater precision will be compensated for due to the errors in the model parameter estimation. Notice that by extending the market model with additional factors, one can obtain the abnormal returns with a more complicated model like the Fama-French three-factor model [22]. Their model, is given by:

$$R_{it} - R_{ft} = \alpha + \beta_1(R_{mt} - R_{ft}) + \beta_2 SMB_t + \beta_3 HML_t + \epsilon_{it}, \tag{3.3}$$

where R_{it} is the raw return, R_{ft} the risk-free rate, α the regression intercept, β_i the OLS regression coefficients, R_{mt} the market return, SMB_t (small minus big) the Size premium , HML_t (high minus low) the Value premium and ϵ_{it} the remainder disturbance. These factor premia are calculated by double sorting the stocks independently in the cross-section. The small cap premium is calculated as the difference between the averages of three small (S) stock portfolios and three big (B) stock portfolios, for which the second sorting is given by growth (G), neutral (N) and value (V) stocks:

$$SMB_t = \frac{1}{3}(SG_t + SN_t + SV_t) - \frac{1}{3}(BG_t + BN_t + BV_t). \tag{3.4}$$

Equivalently, the Value premium is the difference between the averages of two value portfolios and two growth portfolios, where the second sorting is given by small and big stocks:

$$HML_t = \frac{1}{2}(SV_t + BV_t) - \frac{1}{2}(SG_t + BG_t). \tag{3.5}$$

Fama and French [22] use the 30th and 70th percentile as breaking points for the sorting of stocks on value. An advantage of this method over a more simple market model is that it will isolate the firm specific returns better, because proven risk premia are removed from the raw return [22]. However, the three-factor model also has its drawbacks. First of all, the model still does not include all widely excepted drivers of returns. For example, one could think of the Momentum factor, which was added to the three factors in the highly appreciated model of Carhart [17]. Additionally, the model assumes a positive relationship between market β and return, which is in contrast with the low risk premium, treated in Section 2.2.1. Furthermore, the market and risk-adjusted return models require an estimation window around the event, while our specific interest is from the exact moment the event happens.

Often, researchers are interested in the average effects of a certain event on shareholders wealth and, therefore, the cross-sectional mean of abnormal returns is computed by:

$$AAR_t = \sum_{i=1}^{N_t} \frac{AR_{it}}{N_t}, \tag{3.6}$$

where N_t is the number of observations in month t. Additionally, aggregating the abnormal returns over time allows researchers to draw inferences on multiple time period event windows. These cumulative abnormal returns are given by:

$$CAR_{i(t_1,t_2)} = \sum_{t=t_1}^{t_2} AR_{it}. \tag{3.7}$$

Next to averaging or aggregating, one might also sum the mean of abnormal returns over a number of months in order to study the average cumulative effects:

$$CAAR_{t_1,t_2} = \sum_{t=t_1}^{t_2} AAR_t. \tag{3.8}$$

3.1.2. Long-term returns

For event windows longer than a year it is typical to apply long-term return methods. Unlike short-term approaches, it is of critical importance to adjust for risk in long-horizon tests since even small miscalculations or errors in the risk adjustments might cause large differences [35]. Two main methods have been proposed in the literature, being the buy-and-hold abnormal returns (BHAR) and the calendar-time portfolio approach, also known as Jensen's alpha. In the BHAR approach, one uses the characteristics of the event firm to match it with a non-event firm or benchmark portfolio having return equal to $E[R_{i,t}]$. The matching procedure is based on the researcher's believes on the characteristics that best describe stock returns. Matches can be on Size, book-to-price, earnings-to-price or Momentum for example. Following Barber & Lyon [8](p.344), the T-month BHAR is given by:

$$BHAR_{iT} = \prod_{t=1}^{T}(1+R_{it}) - \prod_{t=1}^{T}(1+E[R_{it}]), \tag{3.9}$$

with $R_{i,t}$ the return of an event firm in month t. Presence of abnormal performance is tested using the sample average BHAR. This method has been criticized because of multiple biases, which is why Lyon et al. [38] proposed a bootstrapped skewness adjusted t-statistic and a pseudoportfolio-based test statistic. The two approaches are designed to overcome the biases caused by new listing, skewness and rebalancing [1]. On the contrary, the authors, little promising, conclude that the alternatives are more prone to cross-sectional dependence and thus the test statistics will merely be well-specified in random samples.

According to Mitchell & Stafford [39], event study samples are not random as firms actively decide or decide not to engage in an event. They express their concern on cross-correlation. Greatly overstated test statistics can arise in large data sets with relatively small positive average correlations[2]. Therefore, Mitchell & Stafford [39] and Fama [21] advocate for a calendar-time portfolio approach or Jensen's alpha, which control for the problems arising from cross-sectional dependence by forming portfolios. Each month of the sample period an event portfolio return, R_{pt} is constructed that includes those firms that had an event within the previous T months. Here T is the period over which the abnormal performance is measured. R_{pt}, is either value-weighted or equal-weighted. After subtraction of the risk-free rate, the time series of portfolio returns are regressed on a market and risk-adjusted model such as the Carhart four-factor model, giving the following equation:

$$R_{pt} - R_{ft} = \alpha_p + b_p(R_{mt} - R_{ft}) + s_p SMB_t + h_p HML_t + m_p UMD_t + \epsilon_{pt}. \tag{3.10}$$

Here, α_p is the average monthly abnormal return on the portfolio of event stocks. SMB, HML and UMD are the differences in return on portfolios of 'small' and 'big' stocks, 'high' and 'low' book-to-price stocks and past one-year 'winners' and 'losers' respectively. Although not particularly clear on the reason, Lyon et al. [38] state that this approach results in misspecification of test statistics in nonrandom samples as well and, as a consequence, their conclusion is indecisive in the model of choice.

3.1.3. Market return

A critical part in the construction of normal returns is the specification of a market return, regardless of the time-horizon. The methodology literature is mainly focused on single-markets, whilst our universe comprises both the MSCI World Index and the MSCI Emerging Markets Index. This means that our benchmark should be suitable for multi-country samples. Campbell et al. [16] conducted empirical research of the performance of event studies in (non-US) multi-country samples. They document that market-adjusted returns are commonly used by researchers and also find that simple market-model methods with country wide market indices used as market returns work well. No specific conclusions are stated on the preference of a value or equal weighted index, since the latter is not included in the used database. Other researchers do study the differences between using an equal-weighted index over a value-weighted index, but most of the times these results are focused on test performance regarding the detection of abnormal results and varying conclusions are being drawn [13, 40]. More interesting are the findings of Krueger & Johnson [36]. They found that value or equally weighting stocks in the specification of the market return does not influence the results when testing for market anomalies. Therefore, we can be indifferent in our choice, since we are testing whether these market anomalies still work in case of an M&A event.

[1] See [38] (p.165-166) for a detailed explanation on these biases.
[2] See [35] (p.29) for an example.

4
Linear Models

The main focus in this research is to study the ability of quant models to explain abnormal returns after M&A events. For this, we apply linear models to test whether the factors from Section 2.2.1 are significantly related to the post-event abnormal returns. The data that we are using in our analysis is collected for all firms in the investable universe over the period spanning January 2000 to December 2018. By combining the time series data of a cross-section of firms, a panel data set is formed. Since the universe is rebalanced semi-annually, observations are missing due to firms exiting or entering at some point in time. In terms of panel data, we call this an unbalanced panel [26].

For financial panels, two general forms of dependence are most common [46]. A firm effect is present if the residuals of a certain firm are correlated across different months. Alternatively, dependence among residuals of different firms in the same month is called a time effect. The firm effect can be thought of as an unobserved ability or characteristic of a certain firm, while a time effect could represent an unobserved shock that occurs during a particular month affecting the market as a whole [4]. Including an unobserved effect in the general regression framework, yields a panel model of the following form:

$$y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + u_{it}, \quad i = 1, \dots, N; \; t = 1, \dots, T_i. \tag{4.1}$$

In this equation, y_{it} is the dependent variable, α is a scalar also known as the intercept, $\mathbf{x}'_{it}$ consists of K regressors, $\boldsymbol{\beta}$ is a $K \times 1$ vector with the regression coefficients that need to be estimated and u_{it} consist of a remainder error disturbance and an unobserved effect. Furthermore, we have N equal to the total number of companies and T_i the maximum number of months that firm i is part of our universe. In panel data it is important to know the form of the unobserved effect and whether it is present, as omission or incorrect treatment of it causes biases in the standard errors which might yield spurious relationships due to inflated t-statistics [46].

This chapter treats a variety of regression approaches available for modelling panel data. The first section treats the pooled regression, which yields consistent and efficient estimates when no unobserved effects are present and if the remaining assumptions hold [26]. Multiple approaches will be provided that can be used if these assumptions are violated. The second section treats the regression approaches that model the unobserved effects, and will be called the panel models. A detailed explanation of these models will be provided, as well as the tools that can be used to construct the most appropriate one. The final section treats the methods that will be used to compare the post-event performance of the model to that in regular market circumstances, and discusses the caveats of these.

4.1. Pooled regression methodology

In case that the unobserved effect in Equation 4.1 is absent or ignored, the pooled model is denoted as:

$$y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + \epsilon_{it}, \quad i = 1, \dots, N; \; t = 1, \dots, T_i \tag{4.2}$$

The general approach to solve a model of the above form is the ordinary least squares (OLS) regression. The least squares coefficient vector $\boldsymbol{\beta}$ is obtained by minimizing the sum of residuals ($\mathbf{e}$) squared, which is in matrix notation given by:

$$RSS = \mathbf{e}'\mathbf{e} = (\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}). \tag{4.3}$$

If the following assumptions are met [26]:

1. Strict exogeneity: $E[\epsilon_{it}|\mathbf{x}_{i1},\mathbf{x}_{i2},\ldots,\mathbf{x}_{iT_i}] = 0, \quad i = 1,\ldots,N;\ t = 1,\ldots,T_i,$
2. Homoskedasticity: $Var(\epsilon_{it}|\mathbf{x}_{i1},\mathbf{x}_{i2},\ldots,\mathbf{x}_{iT_i}) = \sigma^2_\epsilon, \quad i = 1,\ldots,N;\ t = 1,\ldots,T_i,$
3. Independence: $cov(\epsilon_{it},\epsilon_{js}|\mathbf{x}_{i1},\mathbf{x}_{i2},\ldots,\mathbf{x}_{iT_i}) = 0, \quad i \neq j;\ t \neq s,$

the OLS estimates are consistent, efficient, and given by:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{4.4}$$

The t-statistic is used to test whether there exists a statistically significant relationship between the dependent variables and each of the separate regressors. Under the null of no relationship the t-statistic, given by:

$$t_{\hat{\beta}_k} = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)}, \tag{4.5}$$

will have a t-distribution with $n-K$ degrees of freedom, where n the total number of observations and K the number of regressors, which is asymptotically standard normal [31]. Typically, one rejects if $|t_{\hat{\beta}_k}| > 1.96$ or equivalently for a p-value < 0.05. In this case the regressor under scrutiny is significantly related to the dependent variable. The standard errors in the pooled regression setting are given by:

$$SE(\hat{\boldsymbol{\beta}}) = \sqrt{s^2(\mathbf{X}'\mathbf{X})^{-1}} = \sqrt{\frac{\mathbf{e}'\mathbf{e}}{n-K}(\mathbf{X}'\mathbf{X})^{-1}}. \tag{4.6}$$

Greene [26] states that in the case of panel data the previously mentioned assumptions are not likely to be met. Ignoring an unobserved effect while it is present, renders the independence assumption invalid. If so, the coefficient estimates can have heavily biased standard errors, over or underestimating the true variability [46]. Underestimating the standard errors might have big consequences, if it yields spurious relationships between the factors and the returns. It is thus of utmost importance to correctly deal with transgressions of these assumptions. Therefore, the following subsections will provide robust alternatives. Generally, the Fama-MacBeth approach is not placed in the regular pooled OLS framework. Nonetheless, it will be treated in this section as it does not separately model the unobserved effect, but its application is validated when a specific assumption of the pooled model is not met.

4.1.1. Standard error clustering

When researchers choose to consciously omit the unobserved effect, the bias introduced in the OLS standard standard errors can be derived through a simple example. Assuming a fixed firm effect, we can decompose the terms of 4.2 into $x_{k,it} = \gamma_{k,i} + \eta_{k,it}$ and $\epsilon_{it} = \mu_i + \nu_{it}$. In such cases:

$$corr(x_{k,it}, x_{k,js}) = \begin{cases} 1, & \text{for } i = j \text{ and } t = s \\ \rho_{x_k} = \frac{\sigma^2_{\gamma_k}}{\sigma^2_{x_k}}, & \text{for } i = j \text{ and } t \neq s \\ 0, & \text{otherwise,} \end{cases} \tag{4.7}$$

$$corr(\epsilon_{it}, \epsilon_{js}) = \begin{cases} 1, & \text{for } i = j \text{ and } t = s \\ \rho_\epsilon = \frac{\sigma^2_\mu}{\sigma^2_\epsilon}, & \text{for } i = j \text{ and } t \neq s \\ 0, & \text{otherwise,} \end{cases} \tag{4.8}$$

represent the correlations among the k explanatory variables and among the errors, respectively [46]. Here $\sigma^2_{\gamma_k}$ and σ^2_μ are the firm specific variance components and $\sigma^2_{x_k}$ and σ^2_ϵ the remainder variance components. Assuming independence across firms, Petersen [46] derived the asymptotic variance of the OLS coefficient estimates to be:

$$AVar[\hat{\beta}_k - \beta_k] = \frac{\sigma^2_\epsilon}{\sigma^2_{x_k} NT}(1 + (T-1)\rho_{x_k}\rho_\epsilon), \tag{4.9}$$

with n the total number of observations[1]. In the absence of a fixed firm effect, the asymptotic variance of the coefficient would be given by the same formula, without the term between parentheses. Hence, the bias is growing in the number of time periods in the panel. Furthermore, one can observe from (4.9), that there will be a bias in the standard errors if and only if both the regressors and errors are correlated within firms, and the bias is increasing in magnitude of these correlations. In case of an unbalanced panel, the bias of the standard errors must be determined using another formula. When assuming equicorrelated errors within clusters [41]:

$$\tau_k \approx 1 + \rho_{x_k}\rho_u\Big(\frac{Var(N_g)}{\bar{N}_g} + \bar{N}_g - 1\Big), \tag{4.10}$$

where ρ_u is the within cluster error correlation, N_g the size of cluster g and ρ_{x_k} the within cluster correlation of the regressors. The latter is given by:

$$\rho_{x_k} = \frac{\sum_i \sum_{j\neq l}(x_{k_{ij}} - \bar{x}_k)(x_{k_{il}} - \bar{x}_k)}{Var(x_k)\sum_i N_i(N_i - 1)}, \tag{4.11}$$

where $x_{k_{il}}$ is the l-th value of i-th cluster of the k-th regressor. Again, this shows that both regressors and the residuals should be correlated for the standard errors to be biased. Strong dependencies in the data and both high variance in cluster size may lead to strong biases in the standard errors of unbalanced panels [42].

Clustering of standard errors is a remedy to biases caused by correlation of the residuals. The state-of-the-art paper by Petersen [46], extensively treats the applications of this method on financial panel data. Clustering is an approach in which the data is grouped on the dimension in which the correlation is observed. Within these groups or clusters correlation of arbitrary form is allowed, meaning that the unobserved effect does not have to be permanent for clustered standard errors to be unbiased. On the contrary, independence is must hold among observations from different clusters. A side advantage of clustering is that it yields additional robustness against violations of the homoskedasticity assumption [46].

The most important result in the paper of Petersen [46] is two-fold, namely correct clustering will give unbiased standard errors and clustered standard errors can provide guidance in improving the model. That is to say, by comparing the clustered standard errors to the heteroskedasticity robust, or White, standard errors, one can check whether a time and/or firm effect is present. This is the case when the clustered standard errors are at least twice as large as the White ones.

Notice that a firm and time effect can coexist in panel data. On finding the most important dimension to cluster on, there are some contradictions in the literature between Petersen [46] and Thompson [51]. The first argues that one should cluster on the dimension that has the most clusters to obtain the least biased standard errors, while the latter states that clustering on the dimension with the smallest number of clusters will yield a bigger bias reduction. However, consensus is found in simultaneously clustering on firm and time, even if the precise form of the dependence is unknown [46]. Thompson [51] proposes to compute the covariance matrix as:

$$V(\hat{\boldsymbol{\beta}}) = V_{Firm}(\hat{\boldsymbol{\beta}}) + V_{Time}(\hat{\boldsymbol{\beta}}) - V_{White}(\hat{\boldsymbol{\beta}}), \tag{4.12}$$

with:

$$\begin{aligned}
V_{Firm}(\hat{\boldsymbol{\beta}}) &= (\boldsymbol{X}\boldsymbol{X}')^{-1}\sum_{i=1}^{N}\Big((\sum_{t=1}^{T}\mathbf{x}_{it}e_{it})(\sum_{t=1}^{T}\mathbf{x}_{it}e_{it})'\Big)(\boldsymbol{X}\boldsymbol{X}')^{-1},\\
V_{Time}(\hat{\boldsymbol{\beta}}) &= (\boldsymbol{X}\boldsymbol{X}')^{-1}\sum_{t=1}^{T}\Big((\sum_{n=1}^{N}\mathbf{x}_{it}e_{it})(\sum_{n=1}^{N}\mathbf{x}_{it}e_{it})'\Big)(\boldsymbol{X}\boldsymbol{X}')^{-1},\\
V_{White}(\hat{\boldsymbol{\beta}}) &= (\boldsymbol{X}\boldsymbol{X}')^{-1}\sum_{t=1}^{T}\sum_{i=1}^{N}\Big((\mathbf{x}_{it}e_{it})(\mathbf{x}_{it}e_{it})'\Big)(\boldsymbol{X}\boldsymbol{X}')^{-1},
\end{aligned}$$

which are the covariance matrices corresponding to cluster robust errors with respect to firm, time and heteroskedasticity, respectively. In these equations, $\mathbf{x}_{it}$ denotes the vector with regressors and e_{it} denotes the OLS residual.

According to Thompson [51] this method yields unbiased standard errors as long as there are enough clusters for both dimensions. Herein, 'enough' is depending on the situation. For example Cameron and

[1]See [46] p.439-440 for a complete derivation.

Miller [14] mention that in some cases with balanced clusters, more than 20-50 clusters is enough, while this number increases for unbalanced cases (where cluster size is not constant). Thompson [51] mentions that if there are no persistent common shocks, 25 clusters on each dimension will be sufficient, while Cameron et al. [15] provide Monte-Carlo experiment results where both firm and time effects and a heteroskedasticity component are present and show that 100 clusters on each dimension provides satisfying rejection probabilities of a true null hypothesis.

4.1.2. Fama-MacBeth regression

In asset pricing the Fama-MacBeth regression is often used for the estimation of parameters. Petersen [46] found that this method is the second most applied method in recently published papers to control for biases in standard errors. With T the number of time periods, for which we use months, the approach comprises of T cross-sectional OLS regressions which each yield a regression coefficient for all regressors [23]. From this time series of regression coefficients, the actual coefficient estimates and standard errors can be derived:

$$\hat{\beta}_k = \sum_{t=1}^{T} \frac{\hat{\beta}_{kt}}{T}, \tag{4.13}$$

$$SE(\beta_k) = \frac{\sigma(\hat{\beta}_i)}{\sqrt{T}} = \sqrt{\frac{1}{T}\sum_{t=1}^{T} \frac{(\hat{\beta}_{kt}-\hat{\beta}_k)^2}{T-1}}. \tag{4.14}$$

The formula for the standard errors will give unbiased results only if the coefficients in the time series are independent of one another. If not, i.e. there is a firm effect in the data, the variance and consequently the standard errors will be too small[2]. One must be extremely careful in applying this procedure in such cases, since an increase in the firm effect will increase the bias in the Fama-MacBeth standard errors. Petersen [46] shows that this is caused by the fact that the firm effect does not appear in the estimated variance. An increase in the correlation will, therefore, reduce the estimated standard errors while the true standard error actually increases. Hence, the procedure is designed for situations with merely cross-sectional correlations, in which it will result in unbiased standard errors [46].

Notice that the regular formulas for the coefficient estimates and the standard errors, given in (4.13) and (4.14), equally weight each of the coefficients. Our study consists of a heavily unbalanced panel, since the indices are both rebalanced twice a year. Therefore, we propose the following formula for the coefficient estimates:

$$\hat{\beta}_k = \sum_{t=1}^{T} w_t \hat{\beta}_{kt} = \sum_{t=1}^{T} \frac{N_t}{n} \hat{\beta}_{kt}, \tag{4.15}$$

with N_t the number of observations at time t and n the total number of observations. The sum of the weights w_t equals one, hence following Kirchner [34] the weighted standard errors are computed by:

$$SE(\hat{\beta}_k) = \frac{\sigma(\hat{\beta}_{kt})}{\sqrt{T_{eff}}} = \sqrt{\sum_{t=1}^{T} \frac{w_t(\hat{\beta}_{kt}-\hat{\beta}_k)^2}{T_{eff}-1}}, \tag{4.16}$$

where:

$$T_{eff} = \frac{1}{\sum_{t=1}^{T}(w_t^2)}, \tag{4.17}$$

the number of efficient months. It can be shown that this number converges towards one if all the observations are concentrated in a single month, while it will converge to the total number of months if all observations are equally distributed. Similarly to the clustered standard errors, the Fama-MacBeth approach has the convenient property of robustness to heteroskedasticity [46].

4.1.3. Heteroskedasticity tests

According to Binder [11], the residuals of a regression of abnormal returns on firm characteristics can expected to be heteroskedastic. The abnormal returns often do not have equal variance for different firms or the variance changes during the period of an event. When neither the clustering approach, nor the Fama-MacBeth regression is used, one needs to test whether the data departs from the homoskedasticity assumption and robust inference can be conducted using the White standard errors.

[2] See Petersen [46] p.446-447 for a derivation of the variance of the coefficient estimates in this case.

Two commonly used methods are the White test, and the Breusch-Pagan test. Under homoskedasticity the variance is constant, and testing for violations can be done with the following hypothesis:

$$H_0 : Var(\boldsymbol{\epsilon}|\boldsymbol{X}) = \sigma^2, \tag{4.18}$$

where $\boldsymbol{\epsilon}$ the disturbance term in the OLS framework, and $\boldsymbol{X}$ the vector of regressors. Following Wooldridge [55], we require the assumption of strict exogeneity to hold, such that $Var(\boldsymbol{\epsilon}|\boldsymbol{X}) = E[\boldsymbol{\epsilon}^2|\boldsymbol{X}]$, and this hypothesis boils down to:

$$H_0 : E[\boldsymbol{\epsilon}^2|\boldsymbol{X}] = E[\boldsymbol{\epsilon}^2] = \sigma^2. \tag{4.19}$$

showing that if the data is heteroskedastic the errors will depend on the value of the regressors. Assuming a particular function, testing requires an auxiliary regression of the regressors on the squared errors. Since the true disturbances are unknown, these are replaced by the OLS residuals $\boldsymbol{e}$. The White test and Breusch-Pagan test each use a different parameterization of the residuals:

$$\text{BP}: \quad \boldsymbol{e}^2 = \delta_0 + \delta_1 \mathbf{x}_1 + \delta_2 \mathbf{x}_2 + \ldots + \delta_k \mathbf{x}_k + \boldsymbol{\nu}, \tag{4.20}$$

$$\text{White}: \quad \boldsymbol{e}^2 = \delta_0 + \delta_1 \hat{\mathbf{y}}_1 + \delta_2 \hat{\mathbf{y}}^2 + \boldsymbol{\nu}, \tag{4.21}$$

where $\mathbf{x}_k$ the vector of the k-th regressors, $\hat{\boldsymbol{y}}$ the fitted values of the original OLS regression and $\boldsymbol{\nu}$ the errors of the auxiliary regression. The Breusch-Pagan and White test statistics are given by:

$$LM = n * R^2_{e^2}, \tag{4.22}$$

with n the sample size and $R^2_{e^2}$ the R-squared of the corresponding auxiliary regression, (4.20) or (4.21). The R-squared can be interpreted as the ratio of variation that is explained by the model over its total variation. Under the null, LM follows a χ^2_k or χ^2_2 for the Breusch-Pagan and White test, respectively [55].

Notice that the White test provided here is a special case, put forward by Wooldridge in [55]. The regular White test uses all original regressors, their squares and their interactions as independent variables. With six regressors in the original model this means that 27 regressors would be included in the auxiliary regression, which yields a relatively large loss of degrees of freedom. The alternative preserves the degrees of freedom and is easier to implement. The White test does not make any assumptions about the form of heteroskedasticity. Simultaneously, this generality has its advantages and potential shortcomings. A rejection of the null hypothesis can be due to another specification error and does not provide information about the subsequent steps [26].

4.2. Panel regression methodology

If an unobserved effect is present, one can also use the panel regression approach to account for it using a fixed effects or random effects model. The general unbalanced panel model is denoted by,

$$y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + u_{it}, \quad i = 1, \ldots, N; \ t = 1, \ldots, T_i, \tag{4.23}$$

where y_{it}, α and x_{it} are defined as in (4.1). In panel data modelling, the random disturbance term is replaced by u_{it} which additionally captures the firm and/or time effect. Although both can be present, most of the panel data regressions include a one-way error component [4]. Depending on the type of effect, the error term is given by:

$$u_{it} = \begin{cases} \mu_i + \nu_{it}, & \text{if there is a firm effect} \\ \lambda_t + \nu_{it}, & \text{if there is a time effect,} \end{cases} \tag{4.24}$$

where μ_i is the unobservable firm effect, λ_t is the unobservable time effect and ν_{it} is the remainder disturbance.

Two important assumptions of the panel model are that the disturbances are homoskedastic and the unobserved effect is constant. The latter means that a firm effect needs to permanent, and a time effect needs to be equal for all individuals observed in a certain time period. Violations of these assumptions will yield biased standard errors and inefficient coefficient estimates of both panel models [4].

The two general approaches, being the fixed effects model and the random effects model, will be treated in Section 4.2.1 and 4.2.2 respectively. For the purpose of introducing the one-way error component model, we focus on the case with a firm effect in the data.

4.2.1. Fixed effects model

The main assumption of the fixed effects model is that the error component or omitted effect is allowed to be correlated with the independent variables. Furthermore, the unobserved effect is fixed and does not vary across the opposite dimension in the panel. As a consequence, a firm effect can be modelled as an individual intercept, denoted by the μ_i term, yielding:

$$y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + \mu_i + \nu_{it}, \quad i = 1,\dots,N;\ t = 1,\dots,T_i, \tag{4.25}$$

where $\nu_{it} \sim IID(0,\sigma_\nu^2)$ [4]. A generally used approach, to model the individual intercepts, is the least squares dummy variable (LSDV) model. In matrix form this is given by:

$$\mathbf{y} = \alpha\mathbf{1}_n + \mathbf{X}\boldsymbol{\beta} + \mathbf{D}_\mu\boldsymbol{\mu} + \boldsymbol{\nu}. \tag{4.26}$$

Here, $\mathbf{1}_n$ is a unit vector with length equal to the total number of observations, and $\mathbf{D}_\mu = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_N]$ with $\mathbf{d}_i$ a $T \times 1$ vector of ones for each time period that the firm is observed and zeros elsewhere. Estimates for $\alpha, \boldsymbol{\beta}$ and $\boldsymbol{\mu}$ are obtained by OLS regression. The restriction $\sum_{i=1}^N \mu_i = 0$ is imposed on the individual intercepts to prevent the model from perfect multicollinearity [4].

The number of parameters to be estimated grows with the number of firms and is likely to exceed computational power for large panels [26]. In that case an alternative to the LSDV model is provided by the within regression. Following Baltagi [4] on the within transformation, the average of Equation (4.25) over time is given by[3]:

$$\bar{y}_{i.} = \alpha + \bar{\mathbf{x}}'_{i.}\boldsymbol{\beta} + \mu_i + \bar{\nu}_i, \tag{4.27}$$

where α and μ_i are time-invariant and thus equal their time-averaged equivalents. Running a regression over (4.27) is called a between regression. Subtracting (4.27) from (4.25) yields:

$$y_{it} - \bar{y}_{i}. = (\mathbf{x}'_{it} - \bar{\mathbf{x}}'_{i.})\boldsymbol{\beta} + (\nu_{it} - \bar{\nu}_{i.}), \tag{4.28}$$

which is called the within transformed model of (4.25). Notice that the within transformation will wipe out all time-invariant variables, including the firm effect. The estimates of the regression coefficients, henceforth referred to as $\tilde{\boldsymbol{\beta}}$, are obtained by an OLS regression. Hereafter, the restriction on the sum of the individual effects allows one to compute the remaining coefficients. First,

$$\tilde{\alpha} = \bar{y}_{..} - \tilde{\boldsymbol{\beta}}\bar{\mathbf{x}}_{..}, \tag{4.29}$$

where the double dots indicate the average across all observations. Subsequently,

$$\tilde{\boldsymbol{\mu}} = \boldsymbol{y} - \tilde{\alpha}\mathbf{1}_n - \mathbf{X}\tilde{\boldsymbol{\beta}}. \tag{4.30}$$

In case of two-way error component disturbances, the fixed effects model is given by:

$$y_{it} = \alpha + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \mu_i + \lambda_t + \nu_{it}, \quad i = 1,\dots,N; \quad t = 1,\dots,T_i. \tag{4.31}$$

The first approach to solve this problem is running an LSDV regression on (4.31), but now including dummies for both the time and entity dimension. Once more, this method is most likely infeasible for problems with a large number of firms and time periods in the data which is why the within transformation provides a solution [4]. This method allows one to use OLS regression for the coefficient estimates on the transformed data:

$$\tilde{\boldsymbol{\beta}} = (\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{y}}, \tag{4.32}$$

with the regressors $\tilde{\boldsymbol{X}} = \boldsymbol{Q}_{[\boldsymbol{\Delta}]}$ and the dependent variables $\tilde{\boldsymbol{y}} = \boldsymbol{Q}_{[\boldsymbol{\Delta}]}\boldsymbol{y}$. Following Baltagi [4], the within transformator is given by:

$$\begin{aligned} \boldsymbol{Q}_{[\boldsymbol{\Delta}]} &= \boldsymbol{I_n} - \boldsymbol{P}_{[\boldsymbol{\Delta}]}, \\ \boldsymbol{P}_{[\boldsymbol{\Delta}]} &= \boldsymbol{\Delta}(\boldsymbol{\Delta}'\boldsymbol{\Delta})^{-1}\boldsymbol{\Delta}', \quad \boldsymbol{\Delta} = (\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2) \\ \boldsymbol{\Delta}_1 &= (\boldsymbol{D}'_1, \dots, \boldsymbol{D}'_T), \quad \boldsymbol{\Delta}_2 = diag[\boldsymbol{D}_t\mathbf{1}_N], \end{aligned} \tag{4.33}$$

where $\boldsymbol{D}_t$ an $N_t \times N$ matrix of ones, omitting the rows for individuals that are not observed in year t.

[3] Note that for each firm the average is taken over the number of months it is observed in our universe.

4.2.2. Random effects model

The main assumption in the random effects model is that the unobserved effect is uncorrelated with the independent variables and may be modelled in the disturbance term [26]. This means that we can use:

$$y_{it} = \alpha + \mathbf{x}'_{it}\boldsymbol{\beta} + \mu_i + \nu_{it}, \quad i = 1,\ldots,N;\ t = 1,\ldots,T_i, \tag{4.34}$$

but now $\mu_i \sim \text{IID}(0,\sigma_\mu^2)$ and $\nu_{it} \sim \text{IID}(0,\sigma_\nu^2)$, such that μ_i independent of ν_{it} [4]. As a result, $Var(u_{it}) = \sigma_\mu^2 + \sigma_\nu^2$ for all i and t, so:

$$cov(u_{it}, u_{js}) = \begin{cases} \sigma_\mu^2, & i = j, t \neq s \\ \sigma_\mu^2 + \sigma_\nu^2, & i = j, t = s \\ 0, & \text{otherwise.} \end{cases} \tag{4.35}$$

From which follows:

$$corr(u_{it}, u_{js}) = \begin{cases} \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\nu^2}, & i = j, t \neq s \\ 1, & i = j, t = s \\ 0, & \text{otherwise.} \end{cases} \tag{4.36}$$

One can observe that the error term is no longer independent for observations on the same firm in different months. Therefore, the OLS regression does no longer yield unbiased and efficient estimates. Instead, generalized least squares (GLS) can be used [4].

Following Baltagi [4], a one-way error random effects model, for further derivations more practically written in matrix form, is given by:

$$\boldsymbol{y} = \alpha \mathbf{1}_n + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\mu} + \boldsymbol{\nu} = \boldsymbol{Z}\delta + \boldsymbol{Z_\mu}\boldsymbol{\mu} + \boldsymbol{\nu}, \tag{4.37}$$

where $\boldsymbol{Z} = (\mathbf{1}_n, \boldsymbol{X})$, $n = \sum_{i=1}^N T_i$, $\boldsymbol{\delta}' = (\alpha, \boldsymbol{\beta}')$, $\boldsymbol{Z_\mu} = diag(\mathbf{1}_{T_i})$ and $\mathbf{1}_{T_i}$ the vector of ones with length T_i. Note that T_i is equal to the number of months that company i is observed. The GLS estimator is now given by:

$$\hat{\boldsymbol{\delta}}_{GLS} = (\boldsymbol{Z}'\boldsymbol{\Omega}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{\Omega}^{-1}\boldsymbol{y}, \tag{4.38}$$

where the covariance matrix is:

$$\boldsymbol{\Omega} = E(\boldsymbol{u}\boldsymbol{u}') = \sigma_\nu^2(\boldsymbol{I}_n + \rho \boldsymbol{Z_\mu}\boldsymbol{Z_\mu}') = \sigma_\nu^2 \boldsymbol{I}_n + \sigma_\mu^2 \boldsymbol{Z_\mu}\boldsymbol{Z_\mu}', \tag{4.39}$$

in which we use $\rho = \frac{\sigma_\mu^2}{\sigma_\nu^2}$ in the last step [4]. $\boldsymbol{\Omega}$ is a matrix of size $n \times n$, where n is again equal to the total number of observations. For large panels, the inversion of such a matrix might be infeasible. Rather than inverting this matrix, Fuller and Battese [24] suggest a transformation to the original system of equations, which allows one to use OLS yielding equivalent estimates. They derived that:

$$\sigma_\nu \boldsymbol{\Omega}_j^{-1/2} = \boldsymbol{I}_{T_j} - \theta_j \bar{\boldsymbol{J}}_{T_j}, \tag{4.40}$$

where $\boldsymbol{I}_{T_j}$ is the identity matrix of size T_j, $\theta_j = 1 - \frac{\sigma_\nu}{w_j}$ with $w_j^2 = T_j\sigma_\mu^2 + \sigma_\nu^2$ and $\bar{\boldsymbol{J}}_{T_j} = \frac{\boldsymbol{J}_{T_j}}{T_j}$ with $\boldsymbol{J}_{T_j}$ a unit matrix of size $T_j \times T_j$[4]. By premultiplying the system with $\sigma_\nu \boldsymbol{\Omega}^{-1/2}$, giving $\boldsymbol{y}^* = \sigma_\nu \boldsymbol{\Omega}^{-1/2}\boldsymbol{y}$ and $\boldsymbol{Z}^* = \sigma_\nu \boldsymbol{\Omega}^{-1/2}\boldsymbol{Z}$, the GLS estimator is obtained by:

$$\hat{\boldsymbol{\delta}} = (\boldsymbol{Z}^{*'}\boldsymbol{Z}^*)^{-1}\boldsymbol{Z}^{*'}\boldsymbol{y}^*, \tag{4.41}$$

which is the OLS estimate of the transformed system. Wooldridge [54] calls this transformation quasi-time-demeaning. In equational form, the model can be written as:

$$y_{it} - \theta_i \bar{y}_i = (1 - \theta_i)\alpha + (\mathbf{x}_{it} - \theta_i \bar{\mathbf{x}}_i)\boldsymbol{\beta} + (u_{it} - \theta_i \bar{u}_i), \tag{4.42}$$

where the bar represents the average of the parameters over time. An advantage of using OLS over GLS is that in the case of disturbances which are heteroskedastic or serially correlated, i.e. the random effect is not permanent, robust covariance matrices are easily obtained [54].

Regardless of the estimation procedure, the variance components σ_ν^2 and σ_μ^2 are required. Baltagi [4] puts forward multiple ANOVA methods to estimate these, as they are among the most popular methods. These

[4] See [4] Chapter 9 for the complete derivations.

estimators are best quadratic unbiased (BQU) in complete panels, but due to the unbalancedness of the data other optimality properties are lost [4]. The BQU estimators,

$$\hat{w}^2 = \frac{\boldsymbol{u}'\boldsymbol{P}\boldsymbol{u}}{tr(\boldsymbol{P})}, \quad \hat{\sigma}_v^2 = \frac{\boldsymbol{u}'\boldsymbol{Q}\boldsymbol{u}}{tr(\boldsymbol{Q})}, \tag{4.43}$$

with $\boldsymbol{P} = diag[\bar{\boldsymbol{J}}_{T_i}]$ and $\boldsymbol{Q} = diag[\boldsymbol{E}_{T_i}]$ where $\boldsymbol{E}_{T_i} = \boldsymbol{I}_{T_i} - \bar{\boldsymbol{J}}_{T_i}$, are not feasible as their computation requires the true disturbances to be known. Wallace and Hussain [53] suggested to substitute the true disturbances for the OLS residuals. Amemiya [1] suggests an approach which replaces the disturbances for the within residuals, $\tilde{\boldsymbol{u}}$ obtained by the regression in (4.28). This yields:

$$\begin{aligned} \hat{\sigma}_v^2 &= \frac{\tilde{\boldsymbol{u}}'\boldsymbol{Q}\tilde{\boldsymbol{u}}}{(n-N-K+1)}, \\ \hat{\sigma}_\mu^2 &= \frac{\tilde{\boldsymbol{u}}'\boldsymbol{P}\tilde{\boldsymbol{u}} - [N-1+tr\{(\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{P}\boldsymbol{X}\} - tr\{(\boldsymbol{X}'\boldsymbol{Q}\boldsymbol{X})^{-1}\boldsymbol{X}'\bar{\boldsymbol{J}}_n\boldsymbol{X}\}]\hat{\sigma}_v^2}{n - \sum_{i=1}^N T_i^2/n}. \end{aligned} \tag{4.44}$$

Swamy and Arora [49] use the between and within residuals as a replacement for $\boldsymbol{u}$ in the estimation of $\hat{w}^2$ and $\hat{\sigma}_v^2$ in (4.43), respectively. The between residuals follow from the regression given by (4.27). Therefore, the variance of the remainder error component, σ_v^2, remains the same as in (4.44) and:

$$\hat{\sigma}_\mu^2 = \frac{\hat{\boldsymbol{u}}^{b'}\boldsymbol{P}\hat{\boldsymbol{u}}^b - (N-K)\hat{\sigma}_v^2}{n - tr[(\boldsymbol{Z}'\boldsymbol{P}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{Z}_\mu\boldsymbol{Z}_\mu'\boldsymbol{Z}]}, \tag{4.45}$$

with $\hat{\boldsymbol{u}}^b$ the between residuals.

The simple ANOVA-type feasible GLS estimators will perform similarly to more advanced methods concerning the estimation of regression coefficients and the remainder variance component. However, unbalancedness of the data might yield a relatively poor estimate of the individual specific variance component compared to that of more complicated methods. Baltagi [4] found that in this situation the Minimum Variance Quadratic Unbiased Estimator (MIVQUE), given by iteration of the following system:

$$\begin{bmatrix} \sigma_\mu^2 \\ \sigma_v^2 \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{bmatrix}^{-1} \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}, \tag{4.46}$$

performs better. In this equation, we have:

$$\begin{aligned} &\gamma_{11} = tr(\boldsymbol{Z_\mu Z_\mu' R Z_\mu Z_\mu' R}), \quad \gamma_{12} = tr(\boldsymbol{Z_\mu Z_\mu' R R}), \quad \gamma_{22} = tr(\boldsymbol{RR}), \\ &\delta_1 = \boldsymbol{y}'\boldsymbol{R Z_\mu Z_\mu' R y}, \quad \delta_2 = \boldsymbol{y}'\boldsymbol{RRy}, \\ &R = (\boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1}\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{\Sigma}^{-1})/\sigma_v^2. \end{aligned} \tag{4.47}$$

In (4.47), $\boldsymbol{\Sigma} = \boldsymbol{I}_n + \rho \boldsymbol{Z_\mu Z_\mu'}$, and all other parameters are as before. This method requires a priori values of the variance components, for which Baltagi [4] states that the ANOVA estimators of Swamy and Arora [49] perform best. Although this might result in better estimation of the variance components, it does not guarantee better coefficient estimates.

4.2.3. Error component tests

The first step in the construction of a panel model is to identify which unobserved effect is significantly present. The literature on significance tests of time and firm effects in panel data is extensive. First we will treat the ones that are suitable for a fixed effects model. Hereafter, the tests for random effects will be provided.

Fixed effects

When one models the data using a fixed effects model, the F-test can be used to test the significance of the unobserved effects. This procedure tests whether the effects among all groups, months or firms, are jointly significant. Under the null hypothesis, the effects are assumed to be zero so that pooled OLS can be used if the remaining assumptions hold. Rejection of the F-test implies that the unobserved effect is significant, meaning that the independence assumption of the pooled regression is violated and its coefficient estimates are inefficient and standard errors biased. Consequently, the fixed effects model is preferred [26].

When testing for a firm effect in unbalanced panel data, the statistic is given by:

$$F = \frac{(R^2_{pooled} - R^2_{FE})/(N-1)}{R^2_{FE}/(n-N-K)} \overset{H_0}{\sim} F_{N-1,\, n-N-K}, \tag{4.48}$$

where R^2_{pooled} squared sum of residuals of the restricted, pooled, model, R^2_{FE} the squared sum of residuals of the fixed effects model, N the number of firms, n equal to the total number of observation, and K the number of explanatory variables [4]. A similar statistic can be obtained for testing the significance of time effects, by replacing the number of firms (N) with the number of months (T). For the two-way fixed effects model, the test is extended to:

$$F = \frac{(R^2_{pooled} - R^2_{FE})/(N+T-2)}{R^2_{FE}/(n-N-T+1-K)} \overset{H_0}{\sim} F_{N+T-2,\, n-N-T+1-K}. \tag{4.49}$$

Baltagi [4] notes that the F-test is robust against nonnormality, which is a desirable property as the normality assumption of the disturbance terms may not be holding in practice. Additionally, the test is found to have good size performance.

Random effects

A battery of Lagrange Multiplier (LM) tests is available for random effects. The general idea behind all of these, is to test whether the variance component of the unobserved effect is significantly different from zero. Baltagi [4] summarizes a number of LM tests that are suitable for unbalanced panel data. The Breusch-Pagan (BP) test for testing a two-way error component with $H_0 : \sigma^2_\mu = \sigma^2_\lambda = 0$, can be denoted by:

$$\begin{aligned} LM &= \frac{1}{2} n^2 \Big[\frac{A_1^2}{M_{11} - n} + \frac{A_2^2}{M_{22} - n} \Big], \\ A_1 &= \frac{\sum_{i=1}^N (\sum_{t=1}^{T_i} e_{it})^2}{\mathbf{e}'\mathbf{e}} - 1, \quad A_2 = \frac{\sum_{t=1}^T (\sum_{i=1}^{N_t} e_{it})^2}{\mathbf{e}'\mathbf{e}} - 1, \\ M_{11} &= \sum_{i=1}^N T_i^2, \quad M_{22} = \sum_{t=1}^T N_t^2, \quad n = \sum_{i=1}^N T_i, \end{aligned} \tag{4.50}$$

in which $\mathbf{e}$ the vector with the OLS residuals from the pooled regression. Under the null hypothesis this statistic is asymptotically distributed as χ^2_2. The LM statistic in (4.50) is additive. By reducing it to the first or second term within parentheses one can test the separate null hypotheses, i.e. $H_0 : \sigma^2_\mu = 0$ and $H_0 : \sigma^2_\lambda = 0$, against a two-sided alternative. Under the null, both test statistics are asymptotically distributed as χ^2_1.

Moulton and Randolph (MR) [42] derived two LM statistics, which equal the square root of the separate terms in (4.50), that are used as the basis for other error component tests. The statistics are given by:

$$\begin{aligned} LM_1 &= n[2(M_{11} - n)]^{-\frac{1}{2}} A_1, \\ LM_2 &= n[2(M_{22} - n)]^{-\frac{1}{2}} A_2, \end{aligned} \tag{4.51}$$

where all terms are as defined in (4.50). The tests use the same null hypothesis as the BP tests, against their one-sided alternatives: $H_1 : \sigma^2_\mu > 0$ or $H_1 : \sigma^2_\lambda > 0$ for LM_1 and LM_2, respectively. Under the null, both follow an asymptotic $N(0,1)$ distribution as $n \to \infty$ and $N \to \infty$.

Baltagi [4] shows that a uniformly most powerful test, originally derived by Honda (HO) [29], can be given by:

$$HO = \frac{LM_1 + LM_2}{\sqrt{2}}. \tag{4.52}$$

A convenience of this statistic is that it is robust to nonnormality of the error component [29]. Furthermore, Baltagi [4] provides the unbalanced version of the King and Wu (KW) test which is the locally mean most powerful one-sided test for the two-way error component model:

$$KW = \frac{\sqrt{M_{11} - n}}{\sqrt{M_{11} + M_{22} - 2n}} LM_1 + \frac{\sqrt{M_{22} - n}}{\sqrt{M_{11} + M_{22} - 2n}} LM_2, \tag{4.53}$$

where LM_1 and LM_2 are defined in 4.51 and the other parameters as in 4.50. Both HO and KW are asymptotically distributed as a standard normal under H_0, and both are one-sided tests for the two-way model.

Unfortunately, unlike the BP test, Baltagi [4] does not mention anything about the additivity of the two statistics and their application to the marginal tests in unbalanced panels.

Gourieroux et al. [25], henceforth referred to as GHM, derived a test which is immune to negative values of LM_1 or LM_2. Such situations tend to occur if the variance components are close to zero [4]. The statistic for unbalanced panel data is given by:

$$\chi_m^2 = \begin{cases} LM_1^2 + LM_2^2, & \text{if } LM_1 > 0, LM_2 > 0, \\ LM_1^2, & \text{if } LM_1 > 0, LM_2 \leq 0, \\ LM_2^2, & \text{if } LM_1 \leq 0, LM_2 > 0, \\ 0, & \text{if } LM_1 \leq 0, LM_2 \leq 0. \end{cases} \tag{4.54}$$

Under the null hypothesis the GHM statistic is distributed as $\chi_m^2 \sim \frac{1}{4}\chi^2(0) + \frac{1}{2}\chi^2(1) + \frac{1}{4}\chi^2(2)$.

Baltagi et al. [6] study the performance of the joint tests with a Monte Carlo experiment. They find that the nominal sizes of the HO and KW tests are not accurate for different patterns of the error component, but their standardized counterparts are. However, due to the size of our panel, the standardized counterparts are not feasible as it includes the calculation of moments of the residuals[5]. Furthermore, BP and GHM have better size performance than the other two tests mentioned, but the BP-test is vulnerable to negative values of LM_1 and LM_2. In over 50% of the cases where either σ_μ^2 or σ_λ^2 is close to zero, it occurs that either of these takes on a negative value [6]. In this specific setting, the BP test often rejects the null hypothesis as it takes the square of the LM statistics. With a Monte Carlo experiment, Baltagi et al. [6] showed that rejection of the null due to negative values of LM_2, is not uncommon with empirical data. Therefore, the GHM test performs better than the BP test in cases where one or both of the variance components are small, and is recommended for unbalanced panels. Finally, the power of each test increases if the panel consists of more firms and when the variance components become larger [6].

Wooldridge [54] provides a test for random firm effects, which is robust to heteroskedasticity and departures from normality. It comes down to testing $H_0 : \sigma_\mu^2 = 0$, against its one-sided alternative with the following test statistic:

$$Z = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T_i-1} \sum_{s=t+1}^{T_i} e_{it} e_{is}}{\left[\sum_{i=1}^{N} \left(\sum_{t=1}^{T_i-1} \sum_{s=t+1}^{T_i} e_{it} e_{is}\right)^2\right]^{\frac{1}{2}}} \overset{H_0}{\sim} N(0,1), \tag{4.55}$$

where e_{it} are the residuals from the pooled OLS regression. In the remainder of this thesis we will refer to (4.55) as the Wooldridge test. A disadvantage of this test is that rejection is not necessarily caused by a firm effect, but might follow from any form of serial correlation [54].

4.2.4. Serial correlation test

A panel model assumes that the correlation among the disturbances of a single firm is solely due to the presence of a permanent effect. In our research we are studying the returns of stocks, which may observe temporary shocks caused by events in the market. Therefore, the assumption of a permanent firm effect may be untenable. Estimates obtained with panel models in the presence of other forms of serial correlation are inefficient and the corresponding standard errors are biased [4].

In the case of first-order serial correlation, the errors in the panel model can be expressed as:

$$\nu_{it} = \rho \nu_{it-1} + \eta_{it}, \tag{4.56}$$

with ν_{it} the idiosyncratic error component of the panel model, η_{it} denoting the part of the error that is IID across firms and time, and $|\rho| < 1$ [52]. Under $H_0 : \rho = 0$, the errors are not dependent across months. On the contrary, if $\rho \neq 0$, the errors of a firm are linearly dependent over time and, thus, serially correlated. In this section we provide a test which can be used in the panel model, and treat an approach for robust inference if serial correlation is present.

Durbin-Watson test

In time series analysis, a popular method to test for first-order serial correlation is given by the Durbin-Watson test. Bhargava et al. [10] modified this test, so that it is applicable in a fixed effects model:

$$d = \frac{\sum_{i=1}^{N} \sum_{t=2}^{T} (\tilde{u}_{it} - \tilde{u}_{it-1})^2}{\sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{u}_{it}^2}, \tag{4.57}$$

[5] The number of observations restricts us, as calculations of the moments of the regression residuals would result in matrices of size $400{,}000 \times 400{,}000$ approximately.

where $\tilde{u}$ the residuals of the aforementioned model. Testing for positive serial correlation comprises of checking where the statistic is located with respect to the interval (d_l, d_u). The null of serial independence, is rejected if $d < d_l$. If $d > d_p$ the null is accepted. However, if the statistic falls within the interval, the test is inconclusive. The critical values of this test are particularly hard to obtain, since it involves the calculation of eigenvalues of a matrix with size $n \times n$. Experiments showed that for panels with large N and T, in our case the number of firms and months, respectively, the interval becomes very tight. As such, it is sufficient to check whether $d < 2$. Equivalently, one can test for negative serial correlation using $4 - d$ [10].

For unbalanced panels, a modification of this test was presented by Baltagi & Wu [5]. With the same underlying test criterion, the statistic is given by:

$$d = \frac{\sum_{i=1}^{N} \sum_{j=1}^{n_i} (\tilde{u}_{it_{ij}} - \tilde{u}_{it_{ij-1}} I[t_{ij} - t_{ij-1} = 1])^2}{\sum_{i=1}^{N} \sum_{j=1}^{n_i} \tilde{u}_{it_{ij}}^2}, \tag{4.58}$$

where $\tilde{u}$ as in (4.57) and n_i is specified such that each firm observes data at times t_{ij} for $j = 1, \ldots, n_i$. The indicator function, I, prevents the test from including observations with gaps larger than a month.

For the random effects model, a wide range of tests are available. However, their applicability to unbalanced panels is limited. For example, Baltagi and Wu [5] developed a locally best invariant test for serial correlation which degenerates to two if some observations in the panel are more than two periods apart. Nevertheless, the fixed effects estimator remains consistent in the random effects model. Hence, the modified Durbin-Watson test can be used for both specifications of the panel model [52].

Robust inference

As the fixed effects model and the random effects model can both be estimated using an OLS regression, robust standard errors are easily obtained. Wooldridge [54] suggests the use of the covariance matrix put forward by Arellano [2]. This matrix is robust to serial correlation and heteroskedasticity of any form in cases where T is relatively small compared to N.

Alternatively, Greene [26] suggests cluster robust covariance matrices for panel models, as a solution to situations where unobserved effects are not permanent. An important statement that he makes is that if the standard errors of the robust covariance matrix and the original one differ a lot, the model is probably not appropriately specified in the first place. This is follows from the fact that correctly modelling the unobserved effect should yield homoskedastic, uncorrelated errors.

4.2.5. Specification tests

Baltagi [4] and Greene [26] agree that there are several shortcomings in both the fixed and random effects model. For the fixed effects model this comes down to the proliferation of parameters in the LSDV approach, and incompetence of accommodating time-invariant independent variables in the within transformation, while for the random effects model it is the mean independence assumption of the unobserved effect which is likely to be untenable [26]. A critical difference between the random effects model and the fixed effects model is the assumption whether or not the unobserved effect is correlated with the regressors. Two tests based on this assumption are treated here.

Hausman's specification test

Hausman [28] devised a specification test based on testing $H_0 : E[u_{it}|\mathbf{X}_{it}] = 0$ against $H_a : E[u_{it}|\mathbf{X}_{it}] \neq 0$. In the fixed effects model, the unobserved effect is captured in a time or firm-invariant component. As we saw in 4.28, the within transformation of the original system wipes out these terms. As a consequence, the fixed effects estimator, $\tilde{\boldsymbol{\beta}}_{FE}$, remains a consistent and unbiased estimator, whether the null hypothesis holds or not. On the contrary, the random effects estimator will merely be consistent when no correlation is present between the error term and the regressors, which is the case under the null. Hence, rejection of the null means that the fixed effects model is preferred [4].

The idea behind the test is to verify whether the estimates of the fixed effects model and random effects model are systematically different. If both methods are consistent, such that there is no misspecification in the random effects model, this should not be the case [26]. For that reason the test statistic is based on the difference of both estimators. In general, this comes down to a simple Wald test which is used assess whether the two estimates are significantly different.

Two numerically identical approaches are treated. Under the null, both yield a test statistic which is asymptotically distributed as χ^2_K, with K the number of regressors [4]. In the first alternative the statistic is given by:

$$m = (\hat{\boldsymbol{\beta}}_{FGLS} - \tilde{\boldsymbol{\beta}}_{within})'[Var(\hat{\boldsymbol{\beta}}_{FGLS} - \tilde{\boldsymbol{\beta}}_{within})]^{-1}(\hat{\boldsymbol{\beta}}_{FGLS} - \tilde{\boldsymbol{\beta}}_{within}), \tag{4.59}$$

where $\hat{\boldsymbol{\beta}}_{FGLS}$ is obtained by means of FGLS or the quasi-time demeaned equivalent, and $\tilde{\boldsymbol{\beta}}_{within}$ is representing the within estimator. Baltagi [4] shows that $Var(\hat{\boldsymbol{\beta}}_{FGLS} - \tilde{\boldsymbol{\beta}}_{within}) = Var(\tilde{\boldsymbol{\beta}}_{within}) - Var(\hat{\boldsymbol{\beta}}_{FGLS})$. A drawback is that the difference of the variance matrices is not necessarily positive definite, possibly yielding a negative value of m. Since $m \overset{H_0}{\sim} \chi^2_K$ implies $m \overset{H_0}{>} 0$, we will automatically, and possibly falsely, accept the null [26]. The second alternative, using the between estimator, has the convenient property:

$$Var(\tilde{\boldsymbol{\beta}}_{within} - \hat{\boldsymbol{\beta}}_{between}) = Var(\tilde{\boldsymbol{\beta}}_{within}) + Var(\hat{\boldsymbol{\beta}}_{between}). \tag{4.60}$$

Here, the variance is defined as the sum of two terms which are by definition larger than zero, such that:

$$m = (\tilde{\boldsymbol{\beta}}_{within} - \hat{\boldsymbol{\beta}}_{between})'[Var(\tilde{\boldsymbol{\beta}}_{within}) + Var(\hat{\boldsymbol{\beta}}_{between})]^{-1}(\tilde{\boldsymbol{\beta}}_{within} - \hat{\boldsymbol{\beta}}_{between}), \tag{4.61}$$

and we do not encounter the issue of non positive definiteness.

Mundlak's approach

To elaborate on Mundlak's approach of the specification test, we shall assume a firm effect. The random effects model with a firm effect assumes mean independence, i.e. $E[\mu_i|\boldsymbol{X}_i] = 0$ [26]. Mundlak [45] devised a test, based on a regression of the random effects model, where he replaces this untenable assumption for:

$$E[\mu_i|\boldsymbol{X}_i] = \bar{\mathbf{x}}'_{i.}\boldsymbol{\gamma}, \tag{4.62}$$

where $\bar{\boldsymbol{x}}_{i.}$ are the group means. If substituted in the random effects model, one can derive the following equation:

$$\begin{aligned} y_{it} &= \mathbf{x}'_{it}\boldsymbol{\beta} + \nu_{it} + \mu_i \\ &= \mathbf{x}'_{it}\boldsymbol{\beta} + \bar{\mathbf{x}}'_{i.}\boldsymbol{\gamma} + \nu_{it} + (\mu_i - E[\mu_i|\boldsymbol{X}_i]) \\ &= \mathbf{x}'_{it}\boldsymbol{\beta} + \bar{\mathbf{x}}'_{i.}\boldsymbol{\gamma} + \nu_{it} + \eta_i, \end{aligned} \tag{4.63}$$

where $\boldsymbol{\gamma}$ are the regression coefficients corresponding to the time averages. One can observe that the specification of the model remains the same as (4.34), by replacing $\nu_{it} - E[\mu_i|\boldsymbol{X}_i]$ with the η_i term, and merely requires the additional estimation of $\boldsymbol{\gamma}$. Under the null, we assume that mean independence holds, thus, $\boldsymbol{\gamma} = 0$. A Wald test,

$$W = \hat{\boldsymbol{\gamma}}Var(\hat{\boldsymbol{\gamma}})^{-1}\hat{\boldsymbol{\gamma}}, \tag{4.64}$$

which is distributed as χ^2_K, with K the number of regressors, under H_0, is used to check whether $\hat{\boldsymbol{\gamma}}$ significantly differs from zero. If it does, i.e. $E[\mu_i|\boldsymbol{X}_i] = \bar{\mathbf{x}}'_{i.}\boldsymbol{\gamma} \neq 0$, this violates the independence assumption of the unobserved effect. Therefore, rejection of the null means that the fixed effects model would be more appropriate [54].

One of the advantages of this regression based approach is that a robust covariance matrix can be specified in case of heteroskedasticity and remainder serial correlation. Omission of these while present, will yield a Hausman test which does not have the correct size [54].

4.3. Separation methods

In this research, we focus on the ability of investment factors to explain post-event abnormal returns. The performance of these factors in non-event periods are required as a benchmark, so that we can tell whether the model works again. The two approaches that can be used to separate the two sets of coefficient estimates, will be treated here.

A first approach comprises of two separate regressions, for the event and non-event observations. For this, one can follow the frameworks for the pooled regression and panel regression as provided in Section 4.1 and Section 4.2, respectively. Even though this is feasible with a pooled regression, application in a panel regression is rather complicated. Due to firms entering and leaving the universe the overall data set is already unbalanced, let alone the data without the event sample. Additionally, the event data by itself does not form a panel.

An alternative approach, using both types of observations in a single regression, is preferable in the panel model. Event dummies are frequently used in the estimation frameworks of event studies [11]. By including interaction terms between the regressors and event dummies (ED), defined in the matrix (2.4), the panel model is given by:

$$y_{it} = \alpha + \alpha_{ED} ED_{it} + \mathbf{x}'_{it}\boldsymbol{\beta} + ED_{it}\mathbf{x}'_{it}\boldsymbol{\gamma} + u_{it} \quad i = 1,\dots,N;\ t = 1,\dots,T_i. \tag{4.65}$$

In (4.65), α_{ED} is the event specific intercept, $\boldsymbol{\gamma}$ denotes the regression coefficient of the interaction term, and the remaining variables are as in (4.1). In (4.65), $\boldsymbol{\gamma}$ captures the difference between the coefficient estimates of the non-event sample, $\boldsymbol{\beta}$, and that of the event sample. Hence, the coefficient estimates for the event data are:

$$\boldsymbol{\beta}_{event} = \boldsymbol{\beta} - \boldsymbol{\gamma}, \tag{4.66}$$

after which the standard errors can be calculated with the formula for the sum of variances:

$$\text{SE}(\boldsymbol{\beta}_{event}) = \sqrt{\text{SE}(\boldsymbol{\beta})^2 + \text{SE}(\boldsymbol{\gamma})^2 + 2cov(\boldsymbol{\beta},\boldsymbol{\gamma})}. \tag{4.67}$$

For the pooled regression the coefficient estimates obtained with the interaction model are numerically equivalent to that of the separate regressions.

The interaction model using panel regression has one major drawback. Estimation of the fixed effects model with the within estimator requires the system to be transformed using (4.28). For the random effects a quasi-time demeaning transformation, given by (4.42), needs to be performed. Whereas the interaction term in (4.65) is merely included in the pooled model if $ED_{it} = 1$, the transformations yield a slightly different equation.

This is illustrated by assuming a fixed firm effect, such that the within transformation of (4.65) becomes:

$$y_{it} - \bar{y}_{i\cdot} = \alpha_{ED}(ED_{it} - \overline{ED}_{i.}) + (\mathbf{x}'_{it} - \bar{\mathbf{x}}'_{i.})\boldsymbol{\beta} + (\mathbf{x}_{it}ED_{it} - \overline{\mathbf{x}_{it}ED_{it}})\boldsymbol{\gamma} + (v_{it} - \bar{v}_{i.}). \tag{4.68}$$

Notice that the time-invariant variables, including the intercept, are wiped out. Zooming in on the interaction term, $(\mathbf{x}_{it}ED_{it} - \overline{\mathbf{x}_{it}ED_{it}})$, we note that if firm i is involved in at least one M&A event for $t = 1,\dots,T_i$, and at least one explanatory variable is non-zero, we will have that $\overline{\mathbf{x}_{it}ED_{it}} \neq 0$. As a consequence, $(\mathbf{x}_{it}ED_{it} - \overline{\mathbf{x}_{it}ED_{it}}) \neq 0$, regardless of the type of observation.

Similarly, the transformed interaction term in the random effects model, given by $(\mathbf{x}_{it}ED_{it} - \theta_i\overline{\mathbf{x}_{it}ED_{it}})$, will also be non-zero for each observation of a firm that is involved in at least one event during the entire period.

As a results, the event regression coefficient of the separate regression and that of the interaction model will not be the same. Moreover, the coefficient estimates of the interaction term in the panel model, denoted by $\boldsymbol{\gamma}$, are not based on merely event observations. One must keep in mind that applying (4.66) will merely provide an approximation of the event coefficient.

5 Results

For each month in the period from January 2000 to December 2018 the firm characteristics and abnormal returns are observed for the companies in our universe. Each observation is defined as either an event observation or non-event observation, depending on whether or not an M&A event manifested for a firm in that particular month. The observations are labelled as such by an indicator, hereafter referred to as the event dummy, in the matrix ED which was defined in (2.4).

In this chapter, the z-scores of the factors described in Section 2.2.1, i.e. Momentum, Quality, Reversal, Risk, Size and Value, are used in a linear regression model to explain the abnormal returns realized by these event or non-event firms. The model in which the factors are used as individual regressors, will be referred to as the factor model. Similarly, the QR model represents the model in which the abnormal returns are regressed on the QR factor.

To omit a forward looking bias, it is required that the regressors are computed in the month prior to the start of the window over which we are considering the abnormal returns. Furthermore, if event windows are overlapping with one another, then only the last event is used in the regression to prevent contaminated event observations. For example, if Company X has events in January and March 2018 and we are looking at a 3-month event window, the first event contaminates the second and will be dropped. Additionally, non-event observations are removed in case the return window overlaps with an event. Finally, a restriction is made on the observations at the end of the time dimension in the panel. Observations are merely included in the regressions if the returns can be considered over the full extent of the return window under scrutiny. For example, if we are aggregating returns over 3 months, the observations with return windows starting in November and December 2018 will be removed from the data. The aforementioned restrictions will hold for all the results in the remainder of this chapter, unless specifically stated otherwise.

This chapter is divided into three sections. Firstly, we will motivate the chosen model for the computation of abnormal returns. Secondly, we will construct the pooled regression models, using robust standard errors to correct for possible biases, and provide the results. Finally, the panel regression approach will be treated.

5.1. Abnormal returns

The main focus of this research is to find out whether the model is able to explain the post-event abnormal returns, within the limit of the current override period of 12 months. In the framework of abnormal returns, this means that the time horizon is ranging from short to mid-term. From the event study methodology literature, the preferred methods are the market-adjusted return model and the market and risk-adjusted return model. Both perform similarly in event studies, as stated in Section 3.1.1. Since the exact months of the M&A events are known, we can focus on the time frame after an event. From an investment perspective the market-adjusted return model poses an advantage over the risk-adjusted alternative. It does not require an estimation window and provides insight in what happens from the exact moment in which we can decide to either invest or not invest. Another advantage of the market-adjusted returns is that these are easily obtained. This is relevant, since we want to benchmark the post-event performance of the firm characteristics to that of non-event observations. Hence, calculation of the abnormal returns is required for the whole sample. For these reasons the market-adjusted return approach is chosen, yielding:

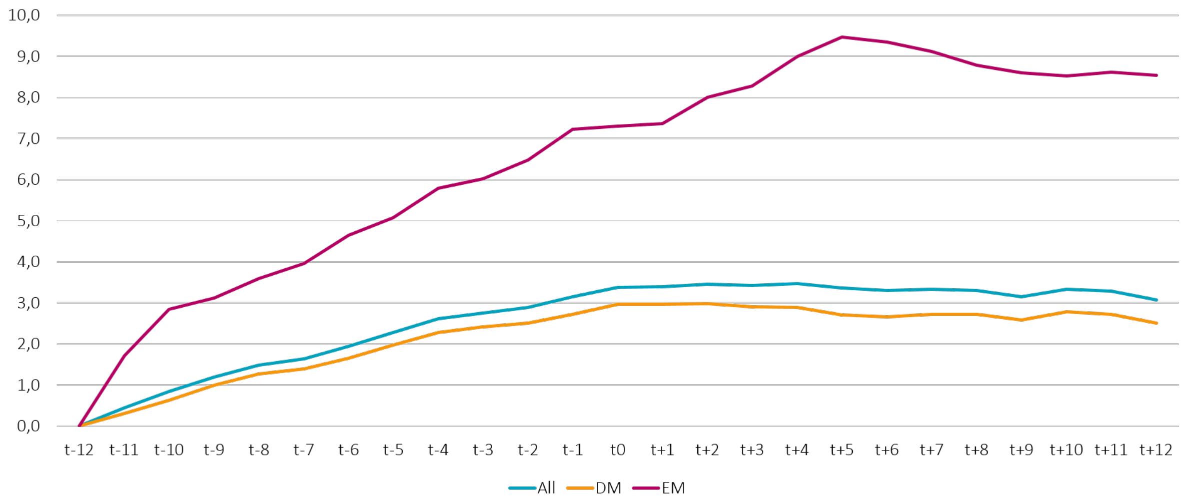
$$AR_{it} = R_{it} - R_{mt}, \tag{5.1}$$

with R_{it} the raw returns of firm i in month t, and R_{mt} the market return during the same month. Subtracting a market return should prevent the abnormal returns of different firms from common effects.

The raw returns are observed and, thus, the specification of a benchmark return remains. In line with Campbell et al. [16], a country index will be used for firms in EM. For the companies in DM, the benchmark is specified as the sector-region combination, to get a more refined abnormal return. The sector definitions follow the MSCI Global Industry Classification Standards[1]. The different regions in DM are: America, Europe, and Japan-Pacific. The literature review of Section 3.1.3 did not result in any concrete preference between an equally or value-weighted benchmark. For ease of computations, the individual returns in the benchmark will be equally weighted.

Even though we are not studying the significance of the mean price effect after an M&A event, it is interesting to provide a brief insight. In Figure 5.1 the cumulative average abnormal returns, given by (3.8), are plotted for a 2-year event window. With the event manifesting in t_0, one can observe that the abnormal returns for the companies in DM are realized more or less entirely prior to that moment. In the period after the event, some slightly negative abnormal returns are realized. On the contrary, companies in EM are subject to positive abnormal returns in this period. From half a year after the event, the returns become negative as well. The overall sample behaves quite similar to the EM index, as almost 90% of the events take place there.

Figure 5.1: Cumulative average abnormal returns



The figure graphs the cumulative average abnormal returns over an event window of 2 years centered around the event.

[1] For more information, see `https://www.msci.com/gics`.

5.2. Pooled model

We start our analysis with a pooled regression model. For the construction of the pooled model, we do not fully ignore the unobserved effects. In order to conduct robust inference, the standard errors will be adjusted accordingly. Therefore, this section will start with the analysis of the standard errors, hereby verifying whether any unobserved effects are present, before going to the actual results. Consecutively, this framework will be used for the pooled model using cumulative abnormal returns and single month returns as dependent variables.

5.2.1. Cumulative abnormal returns

In a pooled model, a separate regression for the event and non-event observations is easily performed. However, by including interaction terms, of the event dummies with the regressors, the differences between the coefficient estimates are observed in a single regression. With all the factors included as z-scores and aggregating the abnormal returns over a certain amount of months s, the cumulative abnormal returns (CAR) are modelled as:

$$CAR_{i,(t+1,t+s+1)} = \alpha + \alpha_{ED} ED_{it} + \sum_{k=1}^{K} \beta_k z_{k,it} + \sum_{k=1}^{K} \gamma_k ED_{it} z_{k,it} + \epsilon_{it}, \quad i = 1,\dots,N;\ t = 1,\dots,T_i, \tag{5.2}$$

where α the intercept, β_k and γ_k the regression coefficients, $z_{k,it}$ the z-score of the k-th regressor, ϵ_{it} the remainder error term and

$$ED_{it} = \begin{cases} 1, & \text{if an event happened for company } i \text{ in month } t \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

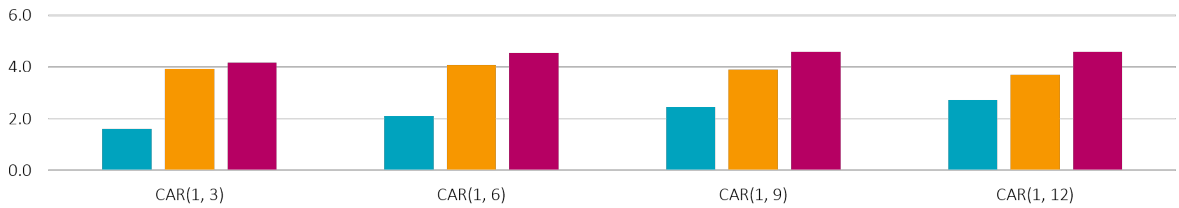
Notice that if the QR factor is used as a single regressor, i.e. $K = 1$, the sums in (5.2) degenerate to a single term. The pooled model uses OLS regression to obtain the coefficient estimates. Since the abnormal returns are aggregated over multiple months and have overlap in the non-event sample it is not likely that the time serial independence assumption is met, resulting in biased standard errors. Time effects are less expected as we are subtracting the market return from the raw returns, which should prevent common shocks to be transferred to the abnormal returns. For the event sample, we do not expect to observe any time or firm effects, since the events are distributed over the months and firms. However, as mentioned in Section 4.1.3, it is unlikely that the data is homoskedastic. In order to analyse the dependencies in the data, we are using the clustering approach which was proposed in Section 4.1.1. The amount of clusters naturally arises from the number of unique components of the dimensions in our panel, being firms and months. For the former dimension this leads to $5,750$ clusters and for the latter this results in 228 clusters.

Standard error analysis

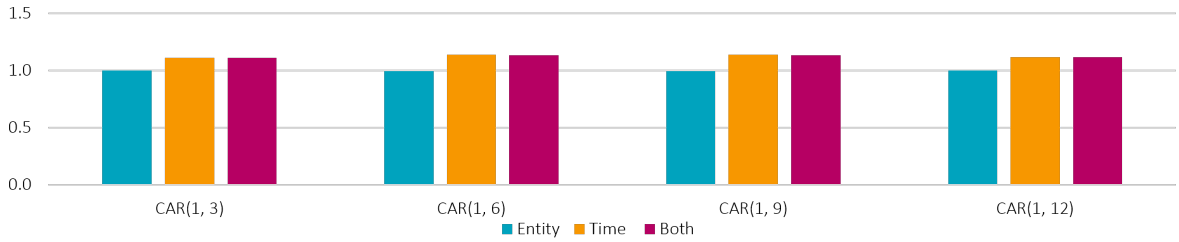
Figure 5.2a depicts the ratios of the clustered standard errors with the heteroskedasticity robust ones for the non-event firms, using the factors as regressors. Note that we are considering the average of the standard errors over the β estimates only. Furthermore, the returns are cumulated over an increasing period, namely from 3 up until 12 months. Petersen [46] states that clustered standard errors indicate that there is a time or firm effect in the data, if this ratio is at least in between two and four. Following this approach, one can observe that the time clustered standard errors are indicating a time effect, as these are on average approximately four times the size of the heteroskedasticity robust ones. This might come rather unexpected, since we used the abnormal returns instead of raw returns to exclude market movements. However, looking at (4.10), the large ratio between the time clustered errors and the White ones might be caused by the large number of observations we have each month. In such cases, low correlation may still cause a significant bias.

Besides a time effect, the results show that clustering on entity inflates the errors by a factor of at least 1.5, increasing over the length of the event window under scrutiny. Clustering on both dimensions results in significantly higher standard errors compared to the firm clustered ones. For longer return periods the difference with the time clustered ones grows as well. Together, this indicates that there might be both a time and a firm effect in the data [46]. From a practical point of view, it is more desirable to be on the conservative side since type I errors will be more influential regarding the implications of this research than type II errors. Therefore, clustering on both dimensions will be used for further analysis of the non-event firms.

Figure 5.2: Ratio of clustered standard errors - factor model



(a) Non-event observations



(b) Event observations

The figures contain the ratios of the clustered standard errors with the heteroskedasticity robust ones for the factor model. The ratios are averaged over the βs.

The standard errors of the event firms are analyzed following the same method. Figure 5.2b shows that there are no indications of any dependencies in the data that should be taken care of, as all ratios are more or less equal to one, thus, time or firm effects are unlikely to be present. Nonetheless, for the standard errors to be unbiased, homoskedasticity is required. In Section 4.1.3, two different tests were proposed. The Breusch-Pagan test and White test both have their advantages. The former is used, since the latter might be rejected due to a misspecification of the model. In such cases, the White test does not indicate what steps are to be taken next [26].

Table 5.1: Breusch-Pagan test

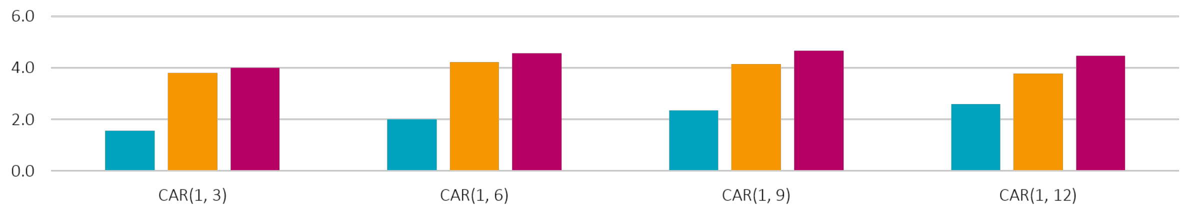
	CAR(1, 3)	CAR(1, 6)	CAR(1, 9)	CAR(1, 12)
Panel A: tests for the factor model				
LM	148.09	148.21	152.60	118.20
p-value	<2.20e-16	<2.20e-16	<2.20e-16	<2.20e-16
Panel B: tests for the QR factor model				
LM	13.33	4.62	0.18	4.40
p-value	2.00e-04	0.03	0.67	0.04

The table contains the outcomes of the Breusch-Pagan test for heteroskedasticity on multiple event windows.

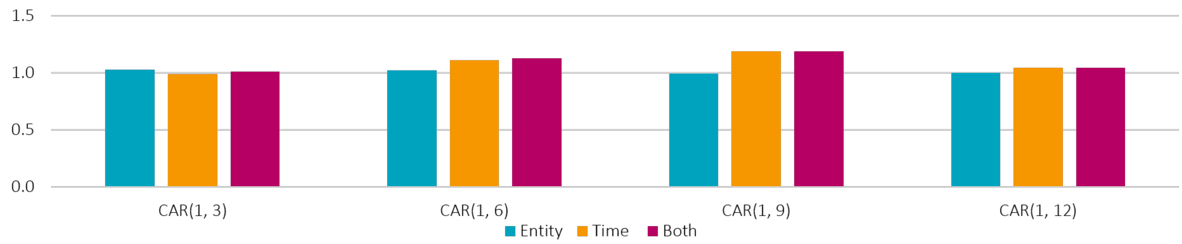
Panel A of Table 5.1 contains the outcomes of this test on the factor model for various event windows. One can observe that each test is highly rejecting the null hypothesis on a 5% significance level, since the p-values are smaller than 0.05. As expected, we can conclude that the data is not homoskedastic and pooled OLS will result in biased standard errors and inefficient estimates. Nevertheless, the White standard errors are used, which do not require any particular assumption about the form of heteroskedasticity for robust inference [26].

As we saw in (4.10), the bias in the standard errors depends on the correlation in the disturbance term and in the regressors. Hence, for the QR factor a similar, but separate, analysis on the standard errors must be conducted. Regarding the event observations, for which the ratios are given in Figure 5.3b, the results do not indicate any need to correct for either a time or firm effect as the clustered standard errors are more or less equal to the heteroskedasticity robust ones. In Panel B of Table 5.1, we observe that the Breusch-Pagan LM statistic for the CAR(1, 9) does not reject the null of homoskedasticity. However, for all other event windows it does. Hence, following the same approach as the factor model, the heteroskedasticity robust standard errors are used in the regression of event observations.

Figure 5.3: Ratio of clustered standard errors - QR model



(a) Non-event observations



(b) Event observations

The figures contain the ratios of the clustered standard errors with the heteroskedasticity robust ones for the QR model.

On the contrary, considering the ratios in Figure 5.3a, the non-event observations are clearly not independent. Whereas the time effect is present for all event windows, an additional firm effect is more likely for longer windows. The latter is presumably due to the fact that a large part of the cumulative abnormal returns are equal within firms where the aggregation periods overlap. When looking at the longer time horizons in Figure 5.3a, we observe that separately clustering on each dimension yields the standard errors to more than double. However, when clustering on both dimensions the standard errors do not significantly increase compared to the time clustered ones. Nevertheless, the standard errors will be of crucial importance in deciding whether the model works or not. Additionally, the methods provided by Petersen [46] are merely guidelines and he specifically states that the ratios of standard errors provide indications and not solid evidence for any effect. Therefore, the most conservative choice is picked, i.e. clustering on both dimensions for non-event firms.

Factor model results

The regression coefficients in the pooled modelled can be estimated with OLS. Using the two-way clustered standard errors as specified above, the results of the factor model (5.2) with the CAR(1, 3) as dependent variable, are given in Table 5.2. The total number of observations can be divided into 5,837 event observations and 461,473 non-event observations.

We clearly observe that the majority of the factors are highly relevant for the non-event firms. As expected, there is a positive relationship between Momentum and the abnormal returns of a firm, meaning that if the average of the past returns and analyst revisions becomes larger, higher cumulative abnormal returns are realized in the following 3 months. For the Quality coefficient we also observe what was anticipated. The relationship is significant so that the Quality metrics, being ROE and gross-profit, have a positive impact on the abnormal returns. Unfortunately, the coefficient estimate of Reversal does not have the expected sign. This might be due to the fact that we are aggregating returns over several months. As mentioned in Section 2.2.1, the factor is effective on a short-term, with last months underperformers expected to outperform next month. Therefore, aggregation might erase its implications, yielding an insignificant factor for the non-event sample. Nevertheless, even for the non-event sample no significant relationship is found. Besides Reversal, Risk is the one factor that is not significant on a 5% level, but we do find the expected negative sign. On the contrary, a strong negative relationship is found between Size and the cumulative abnormal returns. The negative Size coefficient suggests that abnormal returns are smaller for large firms. Of all factors, Value has the largest impact on the cumulative abnormal returns, as its regression coefficient is the highest one found for the non-event data. As expected, firms with high book-to-price and earnings-to-price ratios are expected to generate larger abnormal returns.

The interaction terms denote the differences between the coefficient estimates of the non-event firms and the event firms. This shows that the Momentum and Risk factor have a higher influence on the post-event abnormal returns, than on the abnormal returns of non-event firms. For Quality the increase in the estimate is somewhat smaller, while the effects of Size and Value are weaker after an event. The effects of Value notably changed, and are almost halved for the event firms.

Table 5.2: Interaction factor model for CAR(1, 3)

	Non-event	Interaction
Constant	0.09	-0.20
	(2.25)	(-1.01)
Momentum	0.57	0.11
	(4.37)	(0.51)
Quality	0.11	0.02
	(1.97)	(0.12)
Reversal	-0.03	0.04
	(-0.48)	(0.23)
Risk	-0.32	-0.07
	(-1.68)	(-0.34)
Size	-0.35	0.06
	(-6.34)	(0.44)
Value	0.61	-0.29
	(5.96)	(-1.28)
Observations	467,310	

The table contains the OLS estimates of (5.2) and the t-statistics, within parentheses.

Since we want to draw conclusions about the ability of the factors to explain the post-event abnormal returns we need the t-statistics of the event coefficients. In Section 4.3, we noted that for the pooled OLS model one can obtain the same estimates with two separate regressions. The t-statistics of the separate types of firms, in the interaction model, can be obtained using 4.67. However, as we saw in Figure 5.2a and 5.2b, clustering on two dimension was needed for the non-event sample, while the heteroskedasticity robust standard errors are needed for the event sample. Hence, we prefer to perform the regression for the event and non-event data separately.

The results of the separate regression for the event sample are given in the CAR(1, 3) column of Table 5.3. Despite that the Value coefficient was the only one that negatively differed a lot in magnitude from the non-event sample, one can observe that Momentum and Size are the sole factors that are significantly related to the post-event returns. The coefficients have a t-statistic of 2.80 and 1.97, respectively. Compared the non-event sample, for which the Value and Quality factors had explanatory power as well, these results imply that the factor model performance is worse in the period directly after an M&A event.

By introducing skip months, we aim to capture the post-event model performance for a certain time frame. The model, with an event occurring in t and the skipped periods, is given by:

$$CAR_{i,(t+s+1,t+s+3)} = \alpha + \sum_{k=1}^{K} \beta_k z_{k,it+s} + \epsilon_{it+s}, \tag{5.4}$$

where s denotes the number of skipping months. Note that we are still using the firm characteristics from the month previous to the return window. As for the dependencies in the data, no indications were present for any time or firm effects using the regressions with the returns aggregated up till a period of 12 months. Hereby, motivating the use of the heteroskedasticity robust standard errors for the calculation of t-statistics. The number of skipping months ranges from 3 up till 12, with intervals of 3 months.

The regression results for the factor model are given in the four rightmost columns of Table 5.3. Regarding the signs of the estimates for the factors, we observe that Reversal and Momentum have the incorrect sign for the CAR(4, 6). For this return window, none of the coefficients are significant on either a 5% or 10% level. Performance for this period is even worse than during the initial 3 post-event months. Considering longer skipping periods, i.e. 6, 9 or 12 months, all relationships become somewhat stronger and Size is even consistently significantly related to the post-event CAR as its t-statistics are -2.67, -2.60 and -3.59, respectively. However, some instability in the other relationships remains. For example, there is a highly positive relationship between Value and the abnormal returns aggregated over the seventh until ninth month after an event, indicated by a t-statistic of 2.56 and coefficient estimate of 0.75, while this decreases in both magnitude and significance for the CAR after larger skipping periods. For a skipping period of 1 year, we find that Momentum, Size and Value have significant explanatory power. The remaining coefficient estimates are of similar magnitude as that of the non-event observations in Table 5.2, but remain insignificant. This might have to do with the sample size, which is significantly smaller for the event observations.

Table 5.3: Pooled regression for 3-month CAR - Factor model

	CAR(1,3)	CAR(4, 6)	CAR(7, 9)	CAR(10, 12)	CAR(13,15)
Constant	-0.11	-0.08	-0.10	0.04	0.45
	(-0.53)	(-0.37)	(-0.45)	(0.17)	(1.73)
Momentum	0.68	-0.08	0.26	0.87	0.65
	(3.10)	(-0.37)	(0.98)	(3.36)	(2.08)
Quality	0.13	0.06	0.33	0.33	0.27
	(0.69)	(0.29)	(1.52)	(1.45)	(0.94)
Reversal	0.01	0.05	-0.08	-0.27	-0.15
	(0.05)	(0.29)	(-0.39)	(-1.27)	(-0.53)
Risk	-0.39	-0.36	-0.31	-0.45	-0.22
	(-1.82)	(-1.56)	(-1.20)	(-1.73)	(-0.70)
Size	-0.29	-0.11	-0.45	-0.47	-0.71
	(-1.97)	(-0.70)	(-2.67)	(-2.60)	(-3.59)
Value	0.32	0.27	0.75	0.40	0.63
	(1.51)	(1.12)	(2.56)	(1.51)	(2.01)
Observations	5,837	5,218	4,600	4,188	3,772

The table presents the coefficients and the t-statistics, in parentheses, for the pooled regression of the post-event CAR on the factors.

QR model results

We repeat the previous analysis with the QR model. The regression of (5.2), given in Table 5.4, yields the anticipated sign for the non-event coefficient estimate. A higher QR factor, which means that on average a company scores relatively well on the firm characteristics, leads to a higher return. Furthermore, the outcomes verify what we have found for the separate factors. The regression coefficient of the interaction term is negative, implying that the relationship between the factors and the cumulative abnormal returns is weaker in the event sample.

Table 5.4: Interaction QR model for CAR(1, 3)

	Non-event	Interaction
Constant	0.11	-0.20
	(3.57)	(-1.27)
QR	1.68	-0.23
	(7.16)	(-0.47)
Observations	467,310	

The table contains the OLS estimates of (5.2) and the t-statistics, within parentheses.

By replacing the factors for a combined QR factor in (5.4) and repeating the regressions with skip months, we aim to get more consistent results. The results of this are given in Table 5.5. For the CAR(1, 3), the regression coefficient is 1.45, and highly significant with a t-statistic of 2.89. This is a signal that the QR model will still be able to pick the right stocks based on their average factor performance, in case these are involved in an M&A deal. However, an initial skipping period of 3 months affects the OLS estimates significantly. Despite the correct sign, the slope estimate heavily decreased from 1.45 to 0.62. Consequently, the relationship between the QR factor and the abnormal return is no longer significant on a 5% level. Hence, during the fourth to sixth month after the M&A event, the model is not able to explain the post-event abnormal returns. This is confirming what we found in Table 5.3, where none of the factors were significant during this period. Nevertheless, one can observe that skipping at least 6 months improves the ability of the model to explain the cumulative abnormal returns, as we observe that the QR factor has a strong positive relationship with both the CAR(10, 12) and the CAR(13, 15).

Table 5.5: Pooled regression for 3-month CAR - QR model

	CAR(1,3)	CAR(4, 6)	CAR(7, 9)	CAR(10, 12)	CAR(13,15)
Constant	-0.08	-0.09	-0.09	0.00	0.32
	(-0.52)	(-0.51)	(-0.46)	(-0.01)	(1.46)
QR	1.45	0.62	2.07	2.53	2.34
	(2.89)	(1.17)	(3.64)	(3.98)	(3.36)
Observations	5,837	5,218	4,600	4,188	3,772

The table presents the coefficients and the t-statistics, in parentheses , for the pooled regression of the post-event CAR on the QR factor.

Summarizing, it appears that the majority of the factors are not working in the first months after an event, which would validate the current assumption, that the model is not working. However, using an average score, one would be able to pick the right stocks over the first 3 months following the event. Hereafter, a period follows where neither of the regressions led to significance of any of the coefficient estimates, implying that a model override is necessary. Nonetheless, after half a year or more, the explanatory power of the factors becomes stronger and the QR factor works again in explaining the abnormal returns. Thus, the override period, now set to be 12 months, might be somewhat overstated. It is therefore interesting to look at the exact months in which the model is not able to capture the winners and losers, which is motivating the use of single month returns as dependent variables. We continue by performing an analysis on this model in Section 5.2.2.

5.2.2. Single month returns

In order to verify our previous findings, we proceed the analysis with single month returns. Additionally, we aim to provide a better insight regarding the required length of the override period, by pinning down the exact months in which the model is not able to pick the winning stocks. The same independent variables are used as before, i.e. the factors and the combined QR factor, but the cumulative abnormal returns are substituted for single month abnormal returns. Therefore, a separate analysis on the pooled regression's standard errors must be conducted to assess whether any unobserved effects are present. Since the pooled regression allows one to conduct individual regressions for the event and non-event observations, the model can be denoted by:

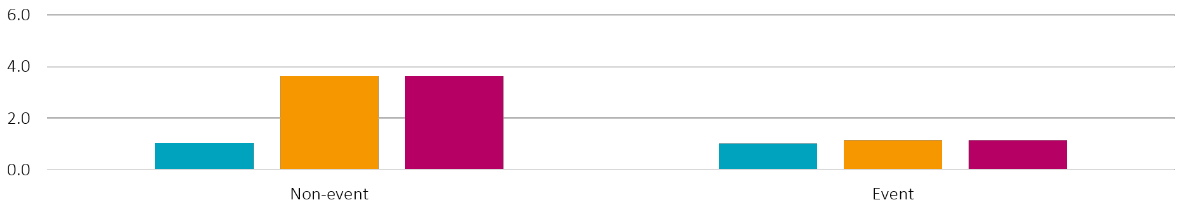
$$AR_{i,t+1} = \alpha + \sum_{k=1}^{K} \beta_k z_{k,it} + \epsilon_{it}, \tag{5.5}$$

where α is the intercept, β_k the regression coefficients, $z_{k,it}$ the z-scores for factor k, or, if $K = 1$, the QR factor, and ϵ_{it} the remainder error term. Although we are not aggregating the returns, an event window of multiple months can still be specified by transferring observations, in the event matrix (2.4), from the non-event data set to the event data set.

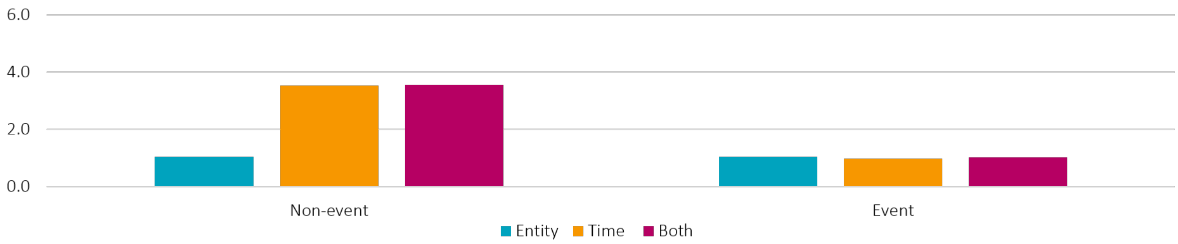
Standard error analysis

As stated in Section 4.1, the assumptions that are needed for OLS regression to yield unbiased standard errors and efficient estimates are unlikely to be met with panel data [26]. Again, violations of the independence assumptions are analysed using the framework provided by Petersen [46].

Figure 5.4: Ratio of clustered standard errors - AR(1)



(a) Factor model



(b) QR model

The figures contain the ratios of the clustered standard errors with the heteroskedasticity robust ones. In Figure 5.4a the ratios are averaged over the βs.

Note that the ratios for the factor model, in Figure 5.4a, are averaged over the β's only, thus, excluding that of the intercept. One can observe that clustering on the entity dimension more or less yields the same standard errors as the White ones for both the non-event and event observations, which does not come as a surprise. Since we are no longer aggregating the returns, no overlap exists in the dependent variables. On the contrary, the ratio of the White errors with the time clustered ones is on average approximately 3.6 for the non-event firms. Being well above two, this indicates that there are correlations among the observations of different firms in the same month [46]. Similar behaviour can be observed in Figure 5.4b, in which the ratios of the standard errors of the QR model are depicted. As seen in Section 4.1.2, the Fama-MacBeth regression is designed for this exact setting. In the presence of a time effect this model produces unbiased standard errors. Although there are no indications of such an effect in the event data, it is not likely to yield biases in the standard errors as a firm effect does not appear to be present. Additionally, the resulting standard errors are heteroskedasticity robust, so there is no need to specify a robust covariance matrix [46]. Thus, the Fama-MacBeth approach can be applied for the regressions of both data sets.

Factor model results

The Fama-MacBeth approach can be split up into two steps. The first step consists of a set of monthly cross-sectional OLS regressions, creating a time series of coefficients. Recall that our panel is heavily unbalanced in general. Running the regression of (5.5) for the non-event firms and event firms separately, yields even more unbalanced data. As a consequence, averaging the time series of regression coefficients overweights those months with less observations. One needs to control for this using the correct formulas. Therefore, in the second step, the coefficient estimates and standard error are computed with (4.15) and (4.16), respectively. These formulas take the weighted averages of the time series of regression coefficients and correct the standard errors accordingly.

Table 5.6: Fama-MacBeth results for AR on factors

	AR(1, 3)	AR(1, 6)	AR(1, 9)	AR(1, 12)	AR(1, 18)	AR(1, 24)
Panel A: non-event regression						
Constant	0.02	0.02	0.02	0.02	0.02	0.02
	(1.39)	(1.42)	(1.56)	(1.65)	(1.42)	(1.19)
Momentum	0.29	0.29	0.30	0.30	0.30	0.31
	(4.78)	(4.91)	(4.96)	(4.95)	(5.06)	(5.19)
Quality	0.04	0.04	0.04	0.04	0.03	0.03
	(1.77)	(1.73)	(1.67)	(1.58)	(1.30)	(1.26)
Reversal	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15
	(-4.07)	(-4.07)	(-4.10)	(-4.00)	(-3.86)	(-3.90)
Risk	-0.16	-0.16	-0.17	-0.17	-0.17	-0.18
	(-1.93)	(-1.93)	(-1.95)	(-1.95)	(-2.04)	(-2.12)
Size	-0.10	-0.10	-0.10	-0.10	-0.10	-0.10
	(-4.02)	(-4.03)	(-3.99)	(-3.90)	(-3.77)	(-3.71)
Value	0.19	0.19	0.19	0.19	0.19	0.20
	(5.22)	(5.29)	(5.18)	(5.23)	(5.35)	(5.38)
Observations	474,217	457,816	443,633	430,915	420,881	391,903
Panel B: event regression						
Constant	0.01	-0.02	-0.04	-0.05	-0.03	-0.03
	(0.20)	(-0.27)	(-0.72)	(-1.02)	(-0.65)	(-0.89)
Momentum	0.20	0.12	0.13	0.18	0.18	0.20
	(1.99)	(1.29)	(1.60)	(2.26)	(2.40)	(2.72)
Quality	0.03	0.03	0.04	0.07	0.09	0.09
	(0.49)	(0.60)	(0.85)	(1.48)	(2.10)	(2.02)
Reversal	-0.19	-0.17	-0.14	-0.16	-0.18	-0.16
	(-2.52)	(-2.75)	(-2.36)	(-2.87)	(-3.39)	(-3.00)
Risk	-0.15	-0.15	-0.17	-0.17	-0.15	-0.16
	(-1.30)	(-1.35)	(-1.63)	(-1.70)	(-1.46)	(-1.60)
Size	-0.08	-0.04	-0.07	-0.09	-0.10	-0.10
	(-1.22)	(-0.89)	(-1.56)	(-2.13)	(-2.40)	(-2.31)
Value	0.09	0.07	0.12	0.10	0.10	0.12
	(1.04)	(1.04)	(1.86)	(1.80)	(1.74)	(2.18)
Observations	18,477	31,993	42,886	52,334	65,638	74,384

The table contains the coefficients and the t-statistics, in parentheses, of the Fama-MacBeth regression. The results are provided for the regression using multiple event windows, ranging from 3 up till 24 months. Panel A present the results of non-event sample and Panel B of the even.

The estimates from the factor model for multiple event windows are given in Table 5.6. In Panel A, one can observe the results of the non-event regressions and Panel B contains the results of the event regressions. Notice that the length of the event window merely affects the value of the events dummies, which are used to transfer observations from the non-event set to the event set. Although not all event coefficients are significantly different from zero, the signs in Panel A and B are all as hypothesized in Section 2.2.1 for each of the windows. Momentum has the strongest positive relationship with the abnormal returns for an event window of 3 months, with a regression coefficient of 0.20 and t-statistic of 1.99 in the first column, and those windows longer than a year. For these, we perceive a regression coefficient of at least 0.18 and t-statistics higher than 2.26 in the three rightmost columns. Unfortunately, in between the aforementioned time windows, the relationship becomes much weaker indicated by the small coefficients and t-statistics. Reversal is the single factor that has a significant relationship with the post-event abnormal returns for all of the event window lengths. Its coefficient estimates range from −0.19, for the AR(1, 3), up till −0.14 for the AR(1, 9) and all t-statistics are well below −1.96. In contrast to the pooled regression on aggregated returns, see Table 5.2, the Fama-MacBeth approach with single month returns seems to be a good method for studying the effects of less persistent factors. The regression approach requires monthly updating of the independent variables and, thus, short-term effects become more important. A marked difference between the non-event and event coefficient for the Quality factor can be observed. The relevance strengthens as longer event windows are being considered for the event observations, while it weakens for the overall sample. This can be induced from the decreasing and increasing coefficients and t-stats in Panel A and Panel B, respectively. Regarding the Risk factor, there is not much difference between the magnitude of the coefficients after an event and for the non-event sample. Nevertheless, post-event significance is not found for any of the event window lengths. Remarkably the factors Size and Value are performing worse in this regression, than in the ones with the cumulative abnormal returns. Although firm size has the expected negative implication on the abnormal returns, it only yields regression coefficients that are significant for event windows longer than a year. Furthermore, Value has much higher explanatory power for the non-event firms than for the event firms, as its coefficient estimates for the event sample are a lot lower and do not yield any relationships that are significantly different from zero. Only for the AR(1, 24) we find that Value is significantly related, with a t-statistic of 2.18 and a regression coefficient of 0.12.

Since we merely find somewhat consistent relationships for event windows that extent a period of longer than a year, we do not expect to find strong explanatory power using the skipping periods.

Table 5.7: Fama-MacBeth regression with skip months

	AR(4, 6)	AR(7, 9)	AR(10, 12)	AR(13, 15)
Constant	-0.01	-0.05	-0.02	0.06
	(-0.12)	(-0.65)	(-0.25)	(0.58)
Momentum	0.03	0.06	0.38	0.24
	(0.29)	(0.56)	(3.08)	(1.86)
Quality	0.03	0.11	0.08	0.09
	(0.40)	(1.26)	(0.96)	(0.97)
Reversal	-0.11	-0.10	-0.28	-0.23
	(-1.37)	(-1.13)	(-3.01)	(-2.16)
Risk	-0.11	-0.08	-0.17	-0.11
	(-0.90)	(-0.64)	(-1.30)	(-0.74)
Size	0.00	-0.12	-0.11	-0.13
	(-0.02)	(-1.69)	(-1.45)	(-1.71)
Value	0.01	0.28	0.10	0.25
	(0.13)	(2.39)	(0.96)	(2.10)
Observations	15,761	13,758	12,413	11,104

The table contains the regression results for a Fama-MacBeth regression of the factors over the abnormal returns. The header contains the amount of months that we skip after an event. Notice that we have specified a 3-month return window for each of the regressions.

In Table 5.7, we are using 3-month event windows. Indeed, we find similar results to that of Table 5.3. We observe that in the period from 4 to 6 months after an event, coefficients in the AR(4, 6) column, none of the factors are significant. Coefficient estimates of Momentum and Reversal, which were the factors that worked for the first 3 months, are significantly affected by the shift with t-statistics of 0.29 and −1.37, respectively. Interesting to notice is that the factors Quality, Size, and Value, are are very close to zero. However, in the period ranging from the 7 to 9 months after an event, the aforementioned factors work relatively well. Value even becomes significant with a coefficient of 0.28 and t-statistic of 2.39. Unfortunately, we do not observe any consistent pattern in the results of Table 5.7.

Skipping 9 or 12 months illustrates the fluctuating significance of factors. For the former period, Reversal and Momentum are significant, while for the latter Reversal and Value are. Interesting to note is all of the coefficients, with the exception of Risk and Momentum, are larger in magnitude than that of the non-event observations of Table 5.6. Therefore, the insignificant t-statistic might be due to a lack of power, possibly caused by the number of observations.

QR model results

Since the results of the regression using the factor model and skip months are not consistent, we are not able to assess the exact months in which the model is not working. Therefore, similar regressions are performed with the QR model. This method has proven to be more consistent in Section 5.2.1 as well. The results can be found in Table 5.8. Note that we expand the event window up to a year, which is sufficient to find consistent results.

Table 5.8: Fama-MacBeth results for AR on QR factor

	AR(1, 3)	AR(1, 6)	AR(1, 9)	AR(1, 12)
Panel A: non-event regression				
Constant	0.02	0.02	0.02	0.02
	(2.93)	(2.73)	(2.90)	(3.03)
QR	0.76	0.77	0.77	0.76
	(6.73)	(6.80)	(6.76)	(6.70)
Observations	474,217	457,816	443,633	430,915
Panel B: event regression				
Constant	0.00	-0.02	-0.04	-0.06
	(-0.03)	(-0.47)	(-1.00)	(-1.52)
QR	0.71	0.52	0.58	0.67
	(3.52)	(3.04)	(3.89)	(4.68)
Observations	18,477	31,993	42,886	52,334

The table contains the coefficients and the t-statistics, in parentheses, for the Fama-MacBeth regression where the AR is used as independent variable. Notice that the standard errors are robust to cross-correlation.

One can observe that the event coefficient of the QR factor, in Panel B, is somewhat lower than the one of the non-event sample, in Panel A. For the latter, the slope estimate is rather stable around 0.76 and highly significant with t-statistics above 6.70. Nevertheless, in both panels we observe highly significant positive relationships, indicating that a higher QR factor yields higher returns for a firm. According to the results, next month's ideal company, regardless of the occurrence of an event, would have a z-score of three for Momentum, Quality and Value and a z-score of minus three for Reversal, Risk and Size. This shows that the QR factor is working directly after an event, as it is able to distinct the future winners and losers based on their average factor performance. Despite the fact that we find a statistically significant coefficient for each of the event windows, the relationship becomes markedly weaker for periods longer than 3 months. A possible explanation might be given by the behaviour of the model found in Section 5.2.1 for the separate factors and the QR factor. There, the model clearly underperformed in the period ranging from 4 to 6 months after an event.

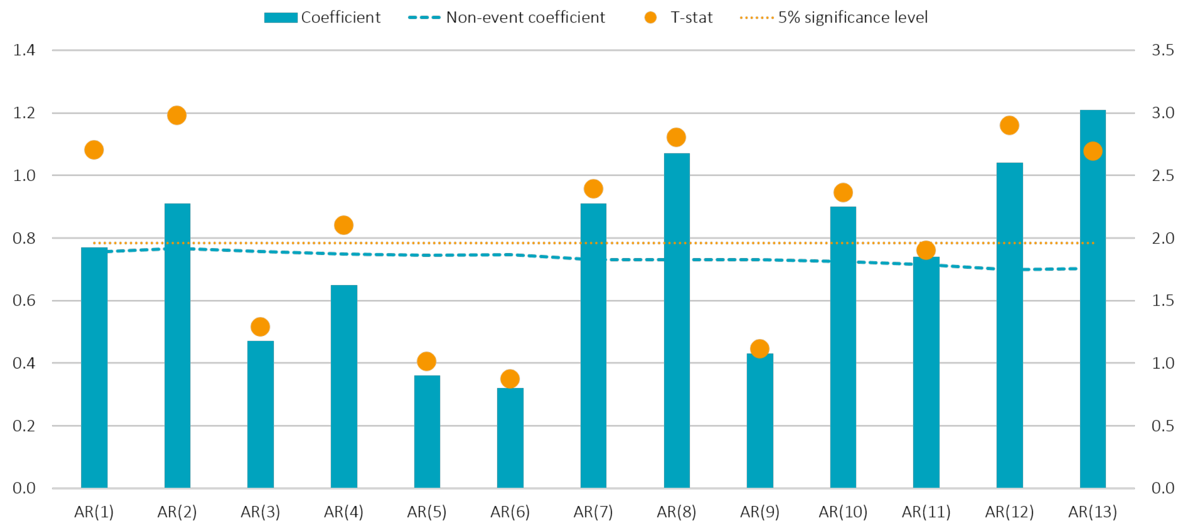
Once again, an initial amount of skipping months is introduced to get an insight in the periods following an M&A event. Since the QR factor yields consistent results, the event windows are reduced to single months, so that:

$$AR_{i\,t+s+1} = \alpha + \beta_{QR} QR_{i\,t+s} + \epsilon_{i\,t+s}, \tag{5.6}$$

where s is equal to the amount of skipping months and the other parameters are as in the Equation (5.5). The observations that fall in between the event month t and the month after skipping $t+s$ are not considered non-event observations and are removed. We proceed with the Fama-MacBeth regression to find the exact post-event period for which the model does not work.

Figure 5.5 shows the results for the event and non-event firms separately. Notice that the regression coefficient for the non-event set decreases from approximately 0.77, in the month directly after the event, to 0.70, when skipping a year after the event. Due to the different number of observations, this slightly deviates from what can be found in Table 5.8. Although not depicted in Figure 5.5, the coefficient estimates for these firms are highly significant, with t-statistics ranging from 6.13 to 6.86. Regarding the event data, the QR factor is able to explain the abnormal returns in the first and second month after an event. The coefficient estimates of the AR(1) and AR(2) are 0.77 and 0.91 with t-statistics of 2.70 and 2.98, respectively, indicating that the relationships are statistically significant. Hereafter, a major drop can be observed in the coefficients, which explains the behaviour of the model as found in Table 5.8. Combining the results for the first three estimates yields a significant positive relationship between the QR factor and the abnormal returns for AR(1, 3). The first 2 months seem to compensate for the lacking performance in the third month. However, if we are looking at the coefficient estimates of the AR(4) up until the AR(6), we observe that merely the first coefficient is significant. Defining a 3-month event window, including these months, does not yield a significant relationship. Consequently, the deterioration in the relationship for the 6-month event window, in comparison to 3 months, can be explained by the model being blind in the fifth and sixth month after the event. With the exception of a skipping period of 8 and 10 months, denoted by the coefficients of the AR(9) and AR(11), all regression coefficients are significant on a 5% level if we are skipping more than half a year. For the AR(9), the coefficient estimate is drawn down by a single outlier in the time series of regression coefficients. The first coefficient in the time series is almost twice as negative as the second lowest coefficient estimate, and, additionally, has a moderate weight. Leaving the observation out of the time series will cause the coefficient to increase to 0.53 with a t-statistic of 1.50. Therefore, the outlier is not treated for, as removing it would not change the implications of that month's coefficient. For the skipping period of 10 months we observe a t-statistic of 1.90, which is significant on a 10% level. On average, the slope coefficients of the final seven observations are even higher than that of the overall sample, being over 0.90.

Figure 5.5: FMB regression results for single month event windows with skip months



The figure plots the regression coefficients in blue. The bars represent those of the event sample and, equivalently, the dotted line those of the non-event sample. Values can be read on the primary (left) y-axis. In orange, the event coefficients' t-statistics are depicted by dots and the 5% significance level by the dotted line. The values can be read from the secondary (right) y-axis.

Summarizing, the results in this section confirm the conjecture that the factors are not equally capable of explaining the abnormal returns after an M&A event as they do in regular circumstances, which is not in conflict with the current assumption that the model is blind. If compared to the overall sample, the factors that are incorporated in the model are not similarly related to the abnormal returns that are being realized. For the shorter event windows, this might have to do with the sample size, as the magnitude of the coefficient of for example Risk and Size are similar to that of the non-event observations, but no significance is found. Altogether, the results for the QR model, in Figure 5.5, imply that the model is in not working in the period between 3 and 6 months after an event, as the abnormal returns can no longer be explained by the QR factor. Hereafter, with exception of the ninth month, the QR factor does work again as a monthly significant positive relationship is found. Next months winners can be recognized by high values of the QR factor and the losers are those with low values. Regarding the second research question, the results show that an override period of 12 months might be too pessimistic and valuable gains can be missed out on. Our findings show that the QR factor is capable of explaining the returns from 9 months after the event onward. The relationship between the QR factor and the post-event abnormal returns is then comparable to that in regular market circumstances. Therefore, the pooled model implies that the override period can be reduced to 9 months.

5.3. Panel model

The analysis in Section 5.2 showed that the assumptions for pooled OLS to be efficient and unbiased are not met. As a consequence, the standard errors of the non-event coefficients were highly underestimated, and needed to be clustered for robust inference. In this section, we will use panel regression to model the unobserved effects. Since the panel model accounts for correlation induced among the residuals by permanent firm and/or time effects, the expectation is that we will be able to capture the true variability of the coefficients estimates more closely. Additionally, if the unobserved effects are modelled correctly, the panel model should result in more efficient estimates of the coefficients. For the event coefficients, we do not expect considerably different implications, as the pooled model did not indicate any unobserved effects in this part of the data. We will include an interaction term in our model to capture the difference in performance for the event and non-event observations.

The panel model is constructed using the series of tests that has been put forward in Section 4.2. The resulting analysis is more comprehensive and more complex than the construction of the pooled model. To keep things clear, the factor model and the QR model will be treated separately.

5.3.1. Cumulative abnormal returns - Factor model

We start our analysis using cumulative abnormal returns as dependent variables, such that the panel model is given by:

$$CAR_{i,(t+1,t+s+1)} = \alpha + \alpha_{ED} ED_{it} + \sum_{k=1}^{K} \beta_k z_{k,it} + \sum_{k=1}^{K} \gamma_k ED_{it} z_{k,it} + u_{it}, \quad i = 1,\dots,N;\ t = 1,\dots,T_i, \tag{5.7}$$

where all the terms, but u_{it}, are as in (5.2). u_{it} is the disturbance term, given in (4.24), containing a firm or time effect and an idiosyncratic part. Recall that six regressors are included in the factor model, namely Momentum, Quality, Reversal, Risk, Size and Value. For model construction we aggregate the returns over 3 months, meaning that $s = 3$.

Error component tests

First of all, we must assess which of the unobserved effects are significant, which can be done along the lines developed in Section 4.2.3. The clustering approach of Petersen [46] indicated the presence of a time and firm effect in the data, motivating us to test for both. The LM tests and F-tests are used for random effects and fixed effects, respectively. Since this is the first panel model that we are developing, we shall use each of the tests proposed in the previously mentioned section. Hereby, we aim to obtain a clear picture of their applicability to our data.

Table 5.9: Random effects tests

H_0 :	$\sigma^2_\mu = 0$	$\sigma^2_\lambda = 0$	$\sigma^2_\mu = \sigma^2_\lambda = 0$
BP	6053.3	19.4	6072.6
p-value	<2.2e-16	1.1e-05	<2.2e-16
MR	77.8	-4.4	-
p-value	<2.2e-16	<2.2e-16	-
HO	-	-	51.9
p-value	-	-	<2.2e-16
KW	-	-	15.7
p-value	-	-	<2.2e-16
GHM	-	-	6053.3
p-value	-	-	<2.2e-16
Wooldridge	22.6	-	-
p-value	<2.2e-16	-	-

This table presents the outcomes of the LM tests, and Wooldridge test, for the random error components.

All tests in Table 5.9 reject the null of no firm effect, indicated by the p-values in the first column being well below 0.05. The abnormal returns are aggregated over 3 months for both the non-event and the event observations. Although, all non-event observations are removed that conflict with an event, the non-event sample will contain partial overlap in the dependent variables. As a consequence, a firm effect is not unexpected.

In the second column, the outcomes are contradictory. The BP test is rejecting the null of no random time effect, while the MR test does not. As discussed in Section 4.2.3, this is not uncommon and can be explained by a brief breakdown of the tests. The BP statistic is the square of the LM_2 term, given by (4.51), in the MR test for time effects. Negative values of this statistic may arise when the variance component is small and close to zero [4]. Since $LM_2^2 = (-4.4)^2 = 19.4$, the critical value of the χ_1^2 distribution is highly exceeded and most likely falsely rejecting the null in favour of a time effect. Therefore, we ignore the outcome of the BP test.

For the joint test, Baltagi [4] recommends using the GHM test. Since LM_2 is negative, the outcome of this test is equal to the square of LM_1, hereby rejecting the joint null based on the presence of a firm effect. This is in line with what is found using the HO and KW tests, which both reject the null of no unobserved effects. Combined the MR and GHM statistics are sufficient to test for random effects and imply the presence of a firm effect.

The F-tests, results in Table 5.10, are used to test for the presence of fixed effects. In the first column, significance is found for the firm effect. Again, we find no indications of time effects in the data. The corresponding statistic is 0.6 with a p-value of 1.0 and, thus, accepts the null that $\boldsymbol{\lambda} = 0$. The joint hypothesis is rejected, due to the significance of the firm effect.

Table 5.10: F-test

H_0 :	$\boldsymbol{\mu} = 0$	$\boldsymbol{\lambda} = 0$	$\boldsymbol{\mu} = \boldsymbol{\lambda} = 0$
F	6.0	0.6	5.9
p-value	<2.2e-16	1.0	<2.2e-16

This table presents the outcomes of the F-tests.

Specification tests

The presence of a firm effect confirms our statement in the introduction of this section; the independence assumption for pooled OLS is not met. Both specifications of the panel model are preferred over the pooled model, so the next step is to define which variant is more appropriate.

For the fixed effects model we prefer the within estimator over the LSDV regression, since the latter would include approximately 5000 firm dummies. Besides, our model does not contain any time-invariant regressors which would be wiped out by the within transformation.

For the random effects model, Swamy and Arora's [49] ANOVA estimators are used for the variance components. This model employs a two-step procedure as described in Section 4.2.2. In the first step, the within and between residuals are obtained by means of OLS, of the corresponding transformed systems, and used for calculation of σ_μ^2 and σ_ν^2. These variance components are used in the second step, where the quasi-time-demeaned transformation is applied to the data and estimates are obtained by OLS. Unfortunately, more involved methods like MIVQUE are infeasible, as they require the inversion of a matrix of size $n \times n$, which is in the order of $400,000 \times 400,000$[2]. In Section 4.2.2, we already mentioned that the random effects model with ANOVA estimators for the variance components performs similarly to more advanced methods, in the estimation of the regression coefficients and remainder error term [4].

Since we need the coefficient estimates and standard errors of both the fixed effects model and random effects model, these are presented in Table 5.11. Notice that the intercept, α in (5.7), of the fixed effects model is missing, as the within transformation wipes out all time-invariant parameters. An intercept can be calculated with the formulas provided in Section 4.2.1, however, we are merely interested in the coefficient estimates of the factors and the corresponding interaction terms so this step has been omitted.

Some comments can be made regarding the estimates of the panel model. First of all, note that the sign of the Risk coefficient differs among the models. In Section 2.2.1, it was hypothesized that the sign should be negative. Modelling the data with a fixed firm effect yields a regression coefficient of 0.04, while this is −0.35 in the random effects case. Moreover, for the Quality factor we observe a negative sign in both models, which is also not what we expected from an economical point of view. However, in the random effects model, using Equation (4.5), this coefficient has a t-statistic of −1.50 and is not significant. On the contrary, it can be seen

[2] See (4.47) for the complete formulas.

that in the fixed effects model, the coefficient is significant. Fortunately, the other coefficients do have the anticipated signs.

Table 5.11: Panel model estimates

	FE		RE	
	Coefficient	standard error	Coefficient	standard error
Constant	-	-	-0.15	0.03
ED	-0.27	0.20	-0.21	0.20
Momentum	0.35	0.02	0.44	0.02
	0.25	0.22	0.16	0.22
Quality	-0.30	0.03	-0.03	0.02
	-0.02	0.19	0.01	0.19
Reversal	-0.16	0.02	-0.08	0.02
	0.18	0.16	0.09	0.16
Risk	0.04	0.04	-0.35	0.03
	-0.08	0.20	-0.08	0.20
Size	-3.35	0.05	-0.76	0.03
	0.08	0.16	0.09	0.16
Value	0.98	0.04	0.75	0.03
	-0.19	0.21	-0.24	0.22
Observations	467,310		467,310	

The table contains the estimates for the fixed effects (FE) and random effects (RE) models. The interaction terms are in the rows without index.

Since the fixed firm effect and random firm effect were found to be significant, the specification test will be used to find which model better suits the data. Following Section 4.2.5, we can distinguish between the models by performing a test on the consistency of the random effects model, based on the assumption of the firm effect being independent of the regressors. Next to the Hausman test, we perform Mundlak's approach. This method introduces an extra set of explanatory variables to the model which represent the conditional expectation of the firm effect with respect to the regressors. If the random effects models the data appropriately there should be mean independence of the firm effect, and the coefficient estimates of the newly introduced variables should not be significantly different from zero. Hence, a regular Wald test on these estimates is sufficient. One advantage of this model over the Hausman test is that, using the Arellano [2] covariance matrix treated in Section 4.2.4, it can be made robust to heteroskedasticity and serial correlation. This is required in cases where the firm effect is not permanent.

Including the interaction terms and the dummy variable, the regression model counts 13 regressors, yielding a χ^2 distribution with 13 degrees of freedom under the null hypothesis of both tests. From Table 5.12, we observe that both the Hausman test and Mundlak's approach reject the null hypothesis of both models being consistent, indicating that the mean independence assumption is violated and, thus, the fixed effects model is more appropriate.

Table 5.12: Specification tests

	Hausman Test	Mundlak's approach
χ^2_{13}	142.2	19365.0
p-value	<2.2e-16	<2.2e-16

The table contains the outcomes of the specification tests.

Serial correlation test

Before we proceed to the results, it is important that one is aware of the potential shortcomings in the fixed effects model. Whenever the firm effect is not constant over time, the within regression will over or underestimate the standard errors making them biased [46]. However, the fixed effects model is solved by an OLS regression of the within transformed system, so robust estimates of the standard errors are easily obtained. Wooldridge [54] even recommends robust inference, whenever this is possible. In order to assess whether any serial correlation is present in the disturbances of the fixed effects model, we apply the modified Durbin-Watson test as treated in Section 4.2.4. Under the null hypothesis, the residuals are assumed to have no positive first-order serial correlation. Using (4.58), which is suitable for unbalanced panels, we obtain a test statistic of $d = 0.98$. Bhargava et al. [10] showed that the critical value is close to two for large panels. Since the exact bounds are particularly hard to obtain in that setting, they remark that it is sufficient to check whether $d < 2$ when testing for positive serial correlation. Clearly, the test statistic indicates the presence of positive first-order serial correlation. The Arellano covariance matrix, given in Table A.2 of Appendix A, is used for robust inference. The standard errors resulting from this are robust to serial correlation and, additionally, heteroskedasticity, of any form [2].

Remarks

The coefficient estimates and the corresponding standard errors of the event sample can be calculated with (4.66) and (4.67), respectively. We must emphasize that we are dealing with approximations of the actual coefficients estimates for the event firms. As described in Section 4.3, this is best shown with a short example. Assume that $ED_{it} = 0$, while $ED_{is} = 1$ for some $s \neq t$ in (5.7). The within transformed interaction term of some regressor k becomes:

$$ED_{it} z_{k,it} - \overline{ED_{it} z_{k,it}} = -\overline{ED_{it} z_{k,it}} \neq 0. \tag{5.8}$$

The interaction term will, thus, be non-zero for all observation of firms that are involved in at least one event. Keeping this in mind, we shall proceed our analysis accordingly.

Factor model results

Table 5.13 presents the results for the fixed effects model. As expected, the Momentum factor is positively related with the abnormal returns. The coefficient is significant and has high explanatory power for both event and non-event observations. The relationship between the CAR(1, 3) and Quality has a somewhat strange interpretation in both samples. The non-event coefficient of −0.30 is highly significant with a t-statistic of −4.28, which suggest that the abnormal returns shrink as companies of higher quality are considered. With Quality being measured by two metrics that capture a firm's profitability, this is not what one would expect. Regarding the event coefficient, we observe that Reversal does not have the expected sign as well. However, the factor does not have a significant impact on the post-event abnormal returns. This is similar to what has been found with the pooled regression in Table 5.3. In the non-event sample, we find that Risk has the lowest explanatory power of all factors, as it has the lowest coefficient estimate and is not significantly related to the CAR(1, 3). Additionally, it does not have the sign which was hypothesized in Section 2.2.1. In the event sample, the sign is opposite, but once again the factor is not significantly related to the cumulative abnormal returns. The Size coefficient has by far the highest explanatory power among the factors, this holds after an M&A event, as well as in regular circumstances. The relationship suggests, as anticipated, that a higher market capitalization is associated with lower returns. The final factor, Value, has the expected effect on the cumulative abnormal returns. In both samples, a positive and highly significant coefficient is found.

Compared to the pooled regression, found in Table 5.3, we observe that the coefficient estimates for Size and Value have increased notably. Consequently, Value is the sole factor in the event sample that is significant in the panel model while it was not in the pooled model. Moreover, multiple coefficients yielded an unexpected sign in the fixed effects model. In the pooled model, we only observed such behaviour for the event coefficient of Reversal which was not even significant. Still, the main implications of the two models are consistent. That is to say, the factors do not have the same explanatory power after an M&A event, as in non-event periods.

Table 5.13: Fixed effects model estimates

	Event	Non-event
Momentum	0.60	0.35
	(2.73)	(7.95)
Quality	-0.32	-0.30
	(-1.62)	(-4.28)
Reversal	0.02	-0.16
	(0.12)	(-7.29)
Risk	-0.04	0.04
	(-0.18)	(0.53)
Size	-3.27	-3.35
	(-17.18)	(-29.15)
Value	0.79	0.98
	(3.62)	(12.63)
Observations	5,837	461,473

The table contains the estimates and t-statistics, in parentheses, for the CAR(1, 3).

As mentioned in the introduction of this section, we expected the fixed effects model to capture the true variability of the coefficient estimates more closely. Table 5.14 presents the (robust) standard errors of the fixed effects model. It is difficult to make statements whether this goal has been reached in the non-event sample. Compared to the fixed effects standard errors, the OLS ones are higher. Nevertheless, this does not hold with the robust fixed effects model's standard errors. Unfortunately, these are inflated by a factor of two, for all coefficients except Reversal. In Section 4.2.4, we mentioned that in a correctly specified model, the robust standard errors should not be considerably different from the original ones. An incorrectly specified fixed effects model does not yield efficient estimates, likewise did the pooled model [26]. Regarding the standard errors of the event observations, the models produce more or less equal results. The robust standard errors do not differ dramatically from the fixed effects standard errors nor from the heteroskedasticity robust OLS errors. Consequently, the violations of the fixed effects model's assumptions, as found in the pooled model, are mainly present in the non-event sample. Moreover, the pooled model seems to be equally well specified, which is why one could argue that it is preferred over the panel model with its additional complexity and mere approximations of the coefficient estimates of the event sample.

Table 5.14: Standard errors of the fixed effects model

	Non-event			Event		
	OLS	FE	FE (robust)	OLS	FE	FE (robust)
Momentum	0.13	0.02	0.04	0.22	0.21	0.22
Quality	0.06	0.03	0.07	0.19	0.19	0.20
Reversal	0.06	0.02	0.02	0.20	0.16	0.17
Risk	0.19	0.04	0.08	0.21	0.20	0.22
Size	0.06	0.05	0.11	0.15	0.16	0.19
Value	0.10	0.04	0.08	0.21	0.21	0.22

This table presents the standard errors from the, robust pooled (OLS) model, fixed effects model (FE) and the robust ones.

Despite, the question remains whether the panel model is able to capture the moment from which the factors are able to explain the post-event returns. We will follow the same procedure as with the pooled model, i.e. introducing skip months after the event. In the panel model (5.7), this boils down to adjusting the event dummies:

$$ED_{it} = \begin{cases} 1, \text{if an event happened for firm } i \text{ at time } t-s \\ 0, \text{otherwise,} \end{cases} \tag{5.9}$$

with s denoting the amount of months that we are skipping after an event. The observations that are in between the event and the return window under scrutiny are removed from the data. Hereby, only a small part of the data is adjusted, which is why the complete construction of a new panel model is omitted. Henceforth, we will use the fixed effects model with robust standard errors. Once more, the event coefficients and corresponding standard errors are calculated using (4.66) and (4.67), respectively. The corresponding non-event estimates and Arellano variance matrices can be found in Appendix A.

The results, presented in Table 5.15, show that all the factors except Quality consistently have the anticipated sign. An exception is made for the CAR(4, 6) where the signs of the Momentum and Reversal factor are reversed, which was observed for the pooled factor model as well. Remarkably, the factors are doing relatively well during that period, having three factors that are highly related to the abnormal returns, i.e. Quality, Size and Value. This is in contrast with results in Section 5.2.1, where this period was performing the worst in multiple approaches. However, this is the only period where the signs of two factors are flipped, indicating that the model is performing differently compared to other periods.

It appears that the fixed effects model is performing the worst in explaining the returns aggregated over the seventh until ninth month after an event, as merely two factors, i.e. Size and Value, are able two explain the cumulative abnormal returns. Overall, these two factors are the only ones which have a strong relationship with the post-event abnormal returns throughout all of the first 15 months. Similar behaviour was found with the pooled regression, where the Size factor had a strong relationship and Value was found to work in half of the periods under consideration. This is showing that these two factors are performing relatively well in explaining the cumulative abnormal returns. However, with maximally three factors that are found to be significantly related, the moment in which model performs equally well as in non-event circumstances is not found.

Table 5.15: Fixed effects model with skip months

	CAR(4, 6)	CAR(7, 9)	CAR(10, 12)	CAR(13, 15)
Momentum	-0.05	0.43	1.13	1.13
	(-0.23)	(1.67)	(4.38)	(3.68)
Quality	-0.47	-0.29	-0.29	-0.31
	(-2.29)	(-1.28)	(-1.25)	(-1.07)
Reversal	0.02	-0.02	-0.30	-0.07
	(0.09)	(-0.11)	(-1.48)	(-0.28)
Risk	-0.03	-0.02	-0.21	-0.03
	(-0.14)	(-0.08)	(-0.75)	(-0.09)
Size	-3.04	-3.35	-3.33	-3.62
	(-16.82)	(-16.78)	(-15.53)	(-16.44)
Value	0.71	1.11	0.72	1.11
	(2.96)	(3.63)	(2.71)	(3.66)
Observations	5,218	4,600	4,188	3,772

This table presents the estimates for the fixed effects model and the robust t-statistics in parentheses.

To conclude, modelling the data with a firm effect did not affect the implications of the results, found earlier with the pooled model. Slight differences occurred with respect to the signs of some factors. Whereas the pooled model mainly yielded the anticipated signs, this was not the case for the panel model. Furthermore, there is some controversy surrounding the correct implementation of an interaction term under the within transformation, posing a disadvantage over the straightforward interpretation in the pooled model. Additionally, modelling of a fixed firm effect is not sufficient to obtain efficient and unbiased estimates. Therewith, one could argue that modelling the unobserved effect in a fixed effects model does not have the desired effect, and its additional complexity does not outweigh its possible advantages over the pooled model.

5.3.2. Cumulative abnormal returns - QR model

Our analysis of the cumulative abnormal returns continues with the QR factor as a single regressor in the panel model. The QR factor is again an average of the six factors, as defined in (2.3). In the panel setting, this yields the following equation:

$$CAR_{i(t+1,t+s+1)} = \alpha + \alpha_{ED} ED_{it} + \beta_{QR} QR_{it} + \gamma_{QR} ED_{it} QR_{it} + u_{it}, \quad i = 1,\dots,N;\ t = 1,\dots,T_i, \tag{5.10}$$

where $CAR_{i(t+1,t+s+1)}$ is the abnormal return of firm i, aggregated from $t+1$ up until $t+s+1$, α and α_{ED} are the intercept for the non-event and event observations, respectively, β_{QR} and γ_{QR} are the regression coefficients, QR_{it} the QR factor, and u_{it} the error component. Note that the independent variables are measured in the month prior to the return window under scrutiny. For construction of the panel model, we will use an event window of 3 months, again.

Error component tests

For the construction of our panel model it is essential to know which of the unobserved effects are significant. In Section 5.3.1, we applied six significance tests for random effects. The results of these tests indicated that it is sufficient to apply the joint GHM test complemented by the MR test. In Table 5.16, the outcomes of these tests on (5.10) are presented. As with the factor model, the null of no firm effect is rejected in favour of $\sigma^2_\mu > 0$. On the other hand, the negative outcome for the test on time effects is most likely indicating that σ^2_λ is close to zero [4]. As a consequence of the negative value of LM_2, the GHM statistic is given by the square of the one-sided MR test for firm effects, LM_1. With a test statistic of 6060.0 the critical value of the mixed χ^2 distribution, which the GHM statistic has under the null of no unobserved effects, is highly exceeded. Combined, the tests suggest the presence of a firm effect.

Table 5.16: Lagrange Multiplier tests

$H_0:$	$\sigma^2_\mu = 0$	$\sigma^2_\lambda = 0$	$\sigma^2_\mu = \sigma^2_\lambda = 0$
MR	77.8	-4.4	-
p-value	<2.2e-19	1.0	-
GHM	-	-	6060.0
p-value	-	-	<2.2e-19

The table contains the values of the test statistics for several LM tests for firm and/or time effects.

The F-test has been performed to test the assumption of no unobserved fixed effects. A rejection of the null means that a fixed effects model is preferred over a pooled model. In Table 5.17, the outcomes are presented. The small p-value of the test for $H_0 : \boldsymbol{\mu} = 0$ suggests that the firm effect is significant. However, there are no indications for fixed time effects.

Table 5.17: F-test

$H_0:$	$\boldsymbol{\mu} = 0$	$\boldsymbol{\lambda} = 0$	$\boldsymbol{\mu} = \boldsymbol{\lambda} = 0$
F	4.7	0.6	4.6
p-value	<2.2e-16	1.0	<2.2e-16

The table presents the outcomes of the F-tests.

The previous findings are partially in line with our expectations. Although a panel model with a firm effect is favoured over the pooled OLS, the clustered standard errors of the non-event data in the pooled regression indicated a pronounced time effect as well. No significant random or fixed effects were found in this dimension of the panel, suggesting that clustering on the time dimension was unnecessary. As a result, the true standard errors of the non-event sample might be overstated in the pooled model. Nevertheless, the estimate of the QR factor's coefficient, found in Table 5.4, was highly significant so a lower estimate of the standard error will not adjust the implications of these results.

Specification tests

As both variants of the panel model are favoured over the pooled model, we need to apply a specification test that distinguishes the two models based on their consistency. Before we apply this test, we will present the estimates of both models. Using the same motivation as for the factor analysis, the within transformation is employed for the fixed effects model, and estimates of the random effects model are obtained with the quasi-time demeaning transformation, with Swamy and Arora's [49] ANOVA estimators for the variance components of the error terms.

Table 5.18: Panel regression estimates for CAR

	FE		RE	
	Coefficient	Standard error	Coefficient	Standard error
Constant	-	-	-0.49	0.05
ED	-0.42	0.18	-0.35	0.18
QR	2.11	0.06	2.01	0.06
	-0.24	0.47	-0.24	0.47
Observations	467,310		467,310	

This table contains the coefficient estimates and standard errors for the fixed effects (FE) model and random effects (RE) model. The interaction terms are in the rows without index.

Table 5.18 contains the regression results for the aforementioned fixed effects model and random effects model. Regarding the non-event coefficient estimates of the QR factor, we observe that both models yield the expected positive relationship with cumulative abnormal returns. For the fixed effects model we observe a coefficient estimate of 2.11 and for the random effects model this is 2.01. Firms with a high QR factor realize higher returns in the following 3 months. Both interaction terms are negative, implying that the impact of the QR factor on the post-event abnormal returns is smaller.

Remark that the residuals of the fixed effects factor model were not serially independent. Hereby, there is no reason to think that the residuals in the QR model are. Whereas the Hausman test is sensitive to such a model misspecification, as it uses the variance of the coefficient estimates, Mundlak's approach can be made robust to it [54]. This test comprises of an auxiliary regression including an additional set of explanatory variables, which represent the conditional expectation of the time effect with respect to the regressors. Under the null of mean independence, the regression coefficients of the extra terms are zero and the random effects model is consistent. By using the Arellano [2] covariance matrix for the coefficients, allowing for any form of heteroskedasticity or serial correlation, we obtain a robust test [54].

Table 5.19: Specification tests

	Mundlak's approach
χ^2_3	2.93
p-value	0.40

The table contains the outcomes of Mundlak's variant of the specification tests on panel data.

Table 5.19 presents the outcome of the Wald test, which tests for the significance of the additional term. With a p-value of 0.40 we accept $H_0 : \hat{\boldsymbol{\gamma}} = \mathbf{0}$, meaning that the main assumption of the random effects model is not violated. We conclude that the random effects model is the preferred specification.

Serial correlation test
In order to do robust inference, we must assess whether there is any autocorrelation in the remainder error term, v_{it}. If there is, then the firm effect is not permanent and the random effects model will yield biased standard errors. For testing the presence of first-order serial correlation we can use the modified Durbin-Watson test. Even though this test is designed for the fixed effects model, it can be applied in the random effects model as well, as the fixed effects estimator will still be consistent [52]. The tests yields a statistic of $d = 0.98$ which is significantly lower than the critical value of two, hence, the null of no positive first-order serial correlation is rejected.

A useful property of the quasi-time-demeaned formulation of the random effects model, (4.42), is that robust covariance matrices can easily be obtained, due to the fact that OLS regression is used for estimation. The Arellano [2] covariance matrix has a perfectly general structure such that it is robust in the presence of serial correlation and heteroskedasticity, provided that T is relatively small in comparison with N [54].

Shortly summarizing, we will use a random effects model with a firm effect and robust standard errors to generate the results of the QR model. For the estimation of the variance components in the model we will use the approach by Swamy and Arora [49].

QR model results
Table 5.20a contains the coefficient estimates and the corresponding t-statistics for the CAR(1, 3). The results of the interaction model need to be split into separate estimates for event and non-event observations, to accommodate comparison of the model's performance in the different samples. Due to the quasi-time demeaning transformation, the interpretation of the interaction term has changed in a similar way as the in the factor panel model[3]. The term does not solely capture the difference of the coefficient estimates, but represents an approximation. With this is mind, the event coefficients are calculated with (4.66) and the standard errors are obtained accordingly, using (4.67). The corresponding t-statistics are calculated with (4.5).

Table 5.20

(a) Random effects model

	CAR(1, 3)
Panel A: non-event coefficients	
QR	2.01
	(20.16)
Observations	461,473
Panel B: event coefficients	
QR	1.77
	(3.40)
Observations	5,837

The table contains the coefficient estimates and the corresponding t-statistics, in parentheses.

(b) Standard errors of the random effects model

	Non-event		Event	
	RE	RE (robust)	RE	RE (robust)
QR	0.06	0.10	0.47	0.52

The table presents the (robust) standard errors of the random effects (RE) model.

The coefficient estimates for the QR factor are positive and highly significant for both event and non-event firms. In Panel A of Table 5.20a, the coefficient is 2.01 with a t-statistic of 20.16. The event coefficient, in Panel B, is somewhat lower but still a strong relationship is found with the cumulative abnormal returns. Thus, stocks that perform well on the average of the selected factors are expected to generate higher cumulative abnormal returns in the following 3 months. This holds for firms that had an M&A event in the month prior to the return window and for the remaining firms in our universe.

If the random effects model is correctly specified, the omitted heterogeneity should be captured by the firm effect and the residuals should be homoskedastic and independent. In Table 5.20b we observe that the relative increase of the robust standard errors, with respect to the regular ones, is rather small for the event coefficient. Using this as a rough interpretation of the model specification, one could argue that the main problems are found in the non-event observations. For this, the robust standard errors almost double. Apparently, the firm effect is not sufficiently capturing the correlation among the residuals in this part of the

[3] See Section 4.3 for a detailed explanation.

data and without robust covariance matrix we would still have highly biased standard errors. These findings, and those of the coefficient estimates, are in accordance with our expectations and the results of the pooled OLS in Section 5.2.1.

Next to the performance of the random effects model directly after an M&A event, we are interested in the months that follow. For this, we need to adjust the event dummies in the random effects model, as done in (5.9).

Table 5.21: Random effects model with skip months

	CAR(4, 6)	CAR(7, 9)	CAR(10, 12)	CAR(13, 15)
QR	0.89	2.32	2.77	2.65
	(1.69)	(4.05)	(4.42)	(3.95)
Observations	5,218	4,600	4,188	3,772

The table contains the coefficient estimates and the corresponding t-statistics, in parentheses, for the event observations.

The results are presented in Table 5.21. For the CAR(4, 6) we perceive a regression coefficient of 0.89, which is somewhat higher than the 0.62 for the pooled model in Table 5.5. Nevertheless, the QR factor is not significantly related to the CAR, as the t-statistic of variable in the random effects model is 1.69. Both regression methods have the same implication, i.e. the QR factor is not able to explain the CAR over the window spanning the fourth to sixth month after an event. For the other periods under scrutiny, the QR factor has a strong positive relationship with the CAR. Despite the tendency of the coefficient estimates to be slightly higher, the implications of the panel model do not deviate from those of the pooled model.

Summarizing, our expectations of the random effects model were not met. Due to the firm effect not being permanent, the random effects model is not efficient and biased. Still, the results of the QR panel model imply the same as the pooled model. In the 3 months directly after an event, the QR factor has a strong positive relationship with the abnormal returns. For a period of 3 months after this, no significant impact of the QR factor is found. In the pooled regression model, we observed that this was caused by the model being blind in the fifth and sixth month after an event. In the following section we aim to confirm this performance of the QR factor using single month abnormal returns as dependent variables.

5.3.3. Single month returns - QR model

In Section 5.2.2 we found that the independence assumption, required for pooled OLS to be efficient and unbiased, was violated for the single month returns by the presence of a time effect in the non-event data. The Fama-MacBeth regression was used to control for this effect. However, none of the tests in the preceding section implied the presence of a time effect. Moreover, the abnormal returns are computed as the difference of the raw return and a time market return, to exclude common market movements in particular. Applying different methods to assess the presence of the unobserved effects, and using a panel regression to model it, might be more appropriate.

In this section we will use an interaction model, rather than separately perform the event and non-event regressions as this would make the panel even more unbalanced. The factor model is not considered here, since the pooled factor model did yield any consistent results for event windows up to a year. Besides, we found in Section 5.3.1 that the results of the panel model were rather complex due to unanticipated signs in the regression coefficients.

The model considered in this section, uses the QR factor and an interaction term to explain the abnormal returns. Dummy variables are used to flag event observations, so that event windows can exist of multiple months and can be shifted. The panel model is given as follows:

$$AR_{it+1} = \alpha + \alpha_{ED} ED_{it} + \beta_{QR} QR_{it} + \gamma_{QR} ED_{it} QR_{it} + u_{it}, \quad i = 1, \dots, N; \; t = 1, \dots, T_i, \tag{5.11}$$

with AR_{it+1} the abnormal returns, β_{QR} and γ_{QR} the regression coefficients, QR_{it} the QR factor, ED_{it} the event dummy, and u_{it} the error component of the model. Again, the event dummy is defined by (5.3). For construction of the panel model, we will use an event window of 1 month.

Error component tests

Just as with the cumulative abnormal returns, our panel analysis starts with LM tests for random error components. Table 5.22, shows that all null hypotheses are accepted. The MR test yields a negative test statistic for testing a time effect. As mentioned in Section 4.2.3, this is not uncommon in situations where σ_λ^2 is close to zero. Also, the GHM test does not reject the joint null hypothesis of no unobserved effects. Remarkably, the LM tests do not indicate the presence of an unobserved random effect in the panel data. If no unobserved effect is present in our data, the Fama-MacBeth regression would have been redundant and pooled OLs would have been sufficient for efficient and unbiased estimates.

Table 5.22: LM tests for random effects

H_0 :	$\sigma_\mu^2 = 0$	$\sigma_\lambda^2 = 0$	$\sigma_\mu^2 = \sigma_\lambda^2 = 0$
MR	0.6	-7.2	-
p-value	0.3	1.0	-
GHM	-	-	0.4
p-value	-	-	0.5

The table presents the test statistics and p-values for two LM tests.

Nevertheless, one test in our arsenal remains. For the F-test, the formulation of (4.48) is used where the residual sums of squares of the within transformation and that of the pooled model are compared. Clearly, the tests, presented in Table 5.23, show that a firm effect might be present while the null of no time effects is not rejected. The joint F-statistic is significant as well. This is most likely caused by the significance of the fixed firm effect.

Merely the fixed effects model is deemed more appropriate than the pooled model. Moreover, the fixed effects model is consistent whether the mean independence assumption holds or not [4]. Thus, the specification tests are not necessary, so we will use a fixed effects model with a firm effect to model the data.

Table 5.23: F-test

H_0 :	$\boldsymbol{\mu} = 0$	$\boldsymbol{\lambda} = 0$	$\boldsymbol{\mu} = \boldsymbol{\lambda} = 0$
F	1.7	0.4	1.7
p-value	<2.2e-16	1.0	<2.2e-16

The table contains the outcome of the F-test.

Serial correlation test

For efficient and unbiased estimates it is required that the residuals of the fixed effects model are homoskedastic and independent. The modified Durbin-Watson test, treated in Section 4.2.4, has been carried out to check for positive first-order serial correlation. Using the fixed effects model with a firm effect, this test yields a static of $d = 0.98$. With a critical value of $d = 2$, the null hypothesis of no positive first-order serial correlation is highly rejected. Petersen [46] warns for this situation in his paper. Modelling the data using a permanent firm effect, while it actually varies over time results in biased standard errors in the fixed effects model. As a consequence, one needs to adjust the standard errors accordingly. For this, the Arellano [2] covariance matrix is used so that the standard errors are robust to autocorrelation. Since these standard errors are robust to heteroskedasticity as well, we need not test for that. Note that we are using (4.66) and (4.67) to calculate the coefficient estimates and the corresponding standard errors of the event sample.

QR model results

For comparison with the Fama-Macbeth regression, Table 5.24 presents the results for the fixed effects model for event windows with lengths of 3 months up until 12 months. The event dummies in (5.11) are adjusted according to the lengths of these. Hereby, the sample is not affected a lot, so we use the fixed effects model with a robust covariance matrix for our estimations.

Table 5.24: Fixed effects model

	AR(1, 3)	AR(1, 6)	AR(1, 9)	AR(1, 12)
Panel A: non-event coefficients				
QR	1.07	1.09	1.09	1.08
	(23.15)	(23.05)	(22.77)	(22.44)
Observations	474,217	457,816	443,633	430,915
Panel B: event coefficients				
QR	1.02	0.81	0.88	0.98
	(5.68)	(6.21)	(7.83)	(9.38)
Observations	18,477	31,993	42,886	52,334

This table presents the coefficient estimates and the corresponding t-stats in parentheses.

Again, a strong positive relationship between the QR factor and the abnormal returns is found in all models. The relationship is found to be much stronger for firms that are not involved in an M&A event, indicated by higher coefficients and t-statistics in Panel A. Compared to Table 5.8, we observe that the regression coefficients for both type of observations are higher when using the fixed effects model. On the other hand, the difference in magnitude of the coefficient estimates over the time windows is rather similar. The event coefficient for the CAR(1, 3) is found to be higher than that of longer windows. After this period, the relationship gradually strengthens. In the Fama-MacBeth regression, we found that this was caused by the factor not being able to explain the abnormal returns in some months, during the first half year, after an event. To pin down the reason for this behaviour of the fixed effects model, the results for a single month event window need to be studied.

Table 5.25: Standard errors - fixed effects model

	AR(1, 3)	AR(1, 6)	AR(1, 9)	AR(1, 12)
Non-event				
FE	0.037	0.038	0.038	0.039
FE(robust)	0.046	0.047	0.048	0.048
Event				
FE	0.164	0.127	0.110	0.101
FE(robust)	0.180	0.131	0.113	0.104

The table presents the (robust) standard errors for the regression coefficients of the fixed effects model.

By comparing the standard errors of the fixed effects model to their robust counterparts, we get a rough insight whether the model is correctly specified. In Table 5.25, the different standard errors are presented. As observed with the previous panel models, the relative increase for the robust standard errors is larger for the non-event coefficients. As a consequence of the firm effects not being able to fully capture the capture the omitted heterogeneity, serial correlation is found in the disturbances. For the event coefficients the robust standard errors do not yield dramatically different estimates. Considering we observed something similar in the pooled model, this is what we initially anticipated.

We need to study 1-month event windows to find out why the regression coefficients decrease for longer periods. For the introduction of skipping months we need to redefine the model. With the fixed effects model given by:

$$AR_{it+1} = \mu_i + \alpha_{ED} ED_{it} + \beta_{QR} QR_{it} + \gamma_{ED,QR} ED_{it} QR_{it} + u_{it}, \quad i = 1, \dots, N; \; t = 1, \dots, T_i, \tag{5.12}$$

but now:

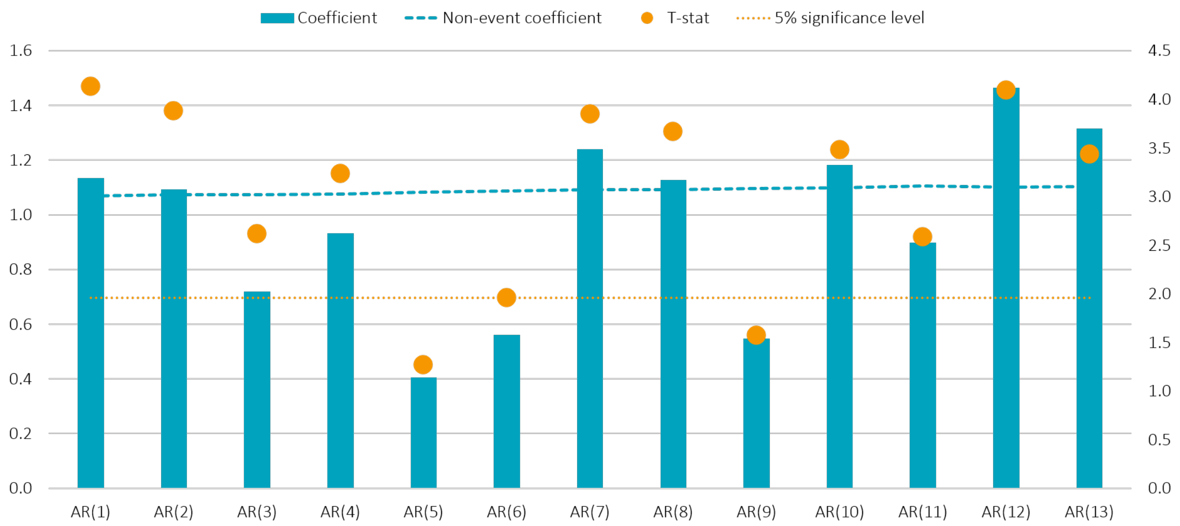
$$ED_{it} = \begin{cases} 1, \text{if an event happened for firm i at time t-s} \\ 0, \text{otherwise,} \end{cases} \tag{5.13}$$

with s is the amount of months that has been skipped after the event. The regression coefficients and corresponding statistics for event observations are extracted as usual. The results are in Figure 5.6.

The two leftmost bars indicate that the relationships between the QR factor and the AR(1) and AR(2) are of similar magnitude as in regular circumstances. The coefficient estimates, indicated by the blue bars, are even slightly higher than those of the non-event sample, which are indicated by the blue dotted line. For the AR(3) and AR(4) we observe that the coefficients are still highly significant with the t-statistics, being well above 1.96, but there is a drop in the coefficient estimates. During the following 2 months a phenomenon is observed that is similar to that in the Fama-MacBeth regression. The relationship deteriorates and yields an insignificant coefficient for the AR(5). In the subsequent month, the QR factor is significant again, but, the relationship is by far not as strong as in the non-event sample. Additionally, no significance is found after a skipping period of 8 months, denoted by the AR(9). For the remaining months, the post-event performance is more or less similar to that in regular circumstances. One could come to the same conclusion as with the Fama-MacBeth regression, meaning that the QR factor has a significant impact on the post-event abnormal returns from 9 months after an M&A event onward.

Petersen [46] studied the behaviour of the Fama-MacBeth approach in the presence of a time effect and came to the conclusion that in such cases, the approach will highly underestimate the true standard errors. Compared to the coefficient estimates of the Fama-MacBeth regression, we observe higher levels of significance for the event data in Figure 5.6. Additionally, the standard errors in Section 5.2.2, indicated that a time effect was merely present in the non-event data. In this section we found that the robust standard errors merely deviated from the standard ones in the non-event data as well. Together, this indicates that violations of the pooled regression's assumptions are primarily observed in the non-event sample.

Figure 5.6: Fixed effects model for single month event windows with skip months



The figure plots the regression coefficients in blue. The bars represent those of the event sample and, equivalently, the dotted line those of the non-event sample. Values can be read on the primary (left) y-axis. In orange, the event coefficients' t-statistics are depicted by dots and the 5% significance level by the dotted line. The values can be read from the secondary (right) y-axis.

To conclude, the implications for the QR model with single month returns are the same across the two regression approaches. We can state that after an M&A event, the QR factor is definitely not explaining the abnormal returns as well as in non-event circumstances. However, the situation does not seem to be as dramatically different as indicated by the Fama-MacBeth regression. Although a model override of 12 months might be a little overstated, it would not be inappropriate to follow a similar procedure involving a shorter period. The results of the fixed effects model suggest that 9 months would be enough. Furthermore, the fixed effects model does not generate unbiased and efficient estimates, as the fixed firm effect was not able to capture all dependence among the residuals of the same firms.

6
Conclusion

6.1. Conclusion

In this thesis we have studied whether quantitative investment factors are able to explain the abnormal returns realized by firms that have been involved in a merger and acquisition (M&A) event, using linear regression models. Additionally, we have examined the required duration of a model override after such events. Currently, the assumption is made that the factors are not able to explain the post-event returns, hence a model override of 12 months is put into practice.

Our data consists of two parts, namely the stock data and the news data. Both types of data are collected for all firms in the MSCI World Index or the MSCI Emerging Markets Index (EM), over a time window spanning from January 2000 to December 2018. Combining this time series of cross-sectional data, a panel data set is formed. Since the indices are rebalanced semi-annually, meaning that firms may enter or leave, the panel is unbalanced.

The event study is a popular method to measure security price reactions and to explain this phenomenon using firm characteristics. This thesis focused on the latter part of this approach. An event study requires the specification of a so-called abnormal returns framework. As the main interest in this research is on the period directly after an event, ranging from several months up till a year after an M&A deal, the focus is on mid-term time horizons. Moreover, the choice has been made to use the market-adjusted return model. An advantage of this approach is that it does not require an estimation window around the event and therewith provides insights into the returns from the exact moment one can decide to invest after an event. Additionally, it is easy to implement and can be used for the non-event observations as well. Market-adjusted returns are defined as the simple difference between the raw returns and a benchmark return. For firms in EM, this benchmark has been defined as an equally weighted average return of the sector, by MSCI GICS classification, and region combination, while for DM benchmarks country averages are used.

In our research we have studied the explanatory power of investment factors on these abnormal returns. The selected factors, being Momentum, Quality, Value, Reversal, Risk and Size, are proven drivers of excess returns. In regular market circumstances, the former three have a positive implication on the returns, meaning that a higher factor score is related to a higher return, while this is negative for the latter three. All of these are measured by one or more firm characteristics, assigned to capture the excess return generated by the factors. For standardization, we have used robust z-scores of these firm characteristics, and labelled all with an absolute value larger than three as an outlier. Outliers are assigned a value three or minus three accordingly. Furthermore, a QR factor, equally weighting each of the factors, was used to explain the abnormal returns. To omit forward looking biases in the regression models, the firm characteristics are related to abnormal returns in the following month.

The main approaches used to model the data are a pooled regression model, and the panel regression model. In panel data, the independence assumption underlying the pooled regression model is often violated by the presence of a time or firm effect [46]. These so-called unobserved effects cause the residuals in the aforementioned model to be dependent, leading to inefficient estimates and biased standard errors. For the pooled models, we considered the pooled OLS regression and the Fama-MacBeth regression. The panel model approaches, modelling the unobserved effect, can be divided into the fixed effects model and the random effects model. The main difference between these panel methods is the mean independence assumption of the regressors with the unobserved effect. The fixed effects model allows for correlation between the unobserved effect and the regressors, and assumes that unobserved effect is fixed. The random effects model does not allow for this correlation and assumes that the unobserved effect is part of the error term. By conducting a battery of tests, the most suitable panel model was selected for each regression.

Firstly, we focused on the pooled model regressing the cumulative abnormal returns and single month abnormal returns on the explanatory variables. In order to conduct robust inference, the approach by Petersen [46] was employed to assess whether any unobserved effects were present in the data. For the cumulative abnormal returns, the ratio of the clustered standard errors and the heteroskedasticity robust ones implied that both a firm and a time effect were present in the non-event observations. The former might have been induced by overlapping abnormal returns, while the latter is most likely a consequence of the large size of the cross-section of firms. No indications were found for any of these effects in the event data. Therefore, the regressions were conducted separately for the event and non-event data, using the heteroskedasticity robust and two-way clustered standard errors, respectively. Hereby, we could still conduct robust inference while omitting the unobserved effects.

By running the regression for both types of observations, we were able to compare the performance of the factors. For the non-event sample, merely Risk and Reversal were not significantly related to the cumulative abnormal returns. On the contrary, for an event window of 3 months, merely Momentum and Size were significantly related to the post-event abnormal returns. Using the same dependent variable in the QR model, a strong positive relationship was found for both samples, meaning that a better average performance on the factors yields higher abnormal returns. During the subsequent 3-month post-event period, the performance of the model worsened. Neither the factor model nor the QR model were able to explain the abnormal returns, implying that the models do not work properly during such periods. However, by performing similar regressions that included longer skipping periods after the event, the QR model was able to explain the abnormal returns correctly. The factor model did not yield similar results, with merely two factors consistently having significant impact on the returns.

Hereafter, single month returns were used to pin down the exact months in which the model does not work. Due to the presence of a time effect, the Fama-MacBeth regression was chosen to estimate the coefficients as it yields unbiased standard errors in this situation. The Fama-MacBeth regression confirmed the previous results, with the factor model not being able to explain the post-event abnormal returns as well as it does in regular circumstances. During the first half year after an M&A event, we found that the QR factor was able to explain the first two monthly abnormal returns. For the remaining months, the model significantly underperformed compared to the non-event QR model. Similarly, the QR model was unable to explain the abnormal returns in the ninth month after an event. Hereafter, the QR model reached similar performance to that in the non-event sample. Altogether, this implied that the model is not working correctly and portfolio managers are right in their decision to override the model. However, an override period of 12 months is likely unnecessarily large, as results showed that the QR model performs as intended after merely 9 months. The override period can thus be reduced by 3 months.

Next, we used the panel models to explain the post-event abnormal returns. Our initial expectations of the panel approach were that, by modelling the unobserved effects, we would obtain unbiased standard errors and more efficient estimates for the non-event data. However, the panel model did not yield results that lived up to these expectations. The residuals were serially correlated, yielding biased standard errors biased and inefficient coefficient estimates, requiring an adjusted covariance matrix for robust inference. In general, this is caused by the unobserved effect not being permanent [46]. We found that mainly the robust standard errors of the non-event coefficients were inflated compared to the regular ones. Using this as a rough indicator of appropriateness, it suggests that the panel models are still not correctly specified. Moreover, the construction of a panel model is more involved than that of a pooled model. This is because it is of utmost importance that one correctly implements and interprets the tests required to find the most appropriate model. Additionally, our main interest was in the post-event performance. The coefficient estimates for this sample needed to be approximated in the panel models, as we could not perform the regressions separately. Besides, the standard errors did not indicate the presence of unobserved effects or issues with serial correlation in this part of the data. Altogether, the shortcomings and complexity of the panel model do not outweigh the advantages.

Bearing this in mind, the panel model did not change any of the conclusions found with the pooled model. Again, we started our analysis with aggregated abnormal returns as dependent variables. In both the factor model and the QR model, multiple Lagrange Multiplier (LM) tests, for random effects, and an F-test for fixed effects, indicated the presence of a time effect. Hereafter, specification tests yielded that a fixed effects model was more appropriate for the factor model, while a random effects model was more appropriate for the QR model. For the estimation of these, we used the within transformation and the quasi-time demeaned transformation, with Swamy and Arora's [49] estimators for the variance components, respectively. Since serial correlation was found in the residuals, the Arellano [2] covariance matrix, allowing for any form of serial correlation and heteroskedasticity, was used for robust inference. Similar to the pooled regression, we found that less factors were able to explain the post-event abnormal returns than in non-event periods. Remarkably, a negative and significant relationship was found between the Quality factor and the returns aggregated over the fourth until the sixth month after an M&A event. This was rather unexpected, as Quality is measured by two metrics capturing the profitability of a firm. The relationship implies that lower profitability results higher returns in the previously mentioned period. In the QR model, we did not find anything unanticipated. Positive relationships were found between the QR factor and the abnormal returns, except for those aggregated over the fourth until sixth month after an event.

Due to the complexity of the factor model in the panel setting, we decided not to investigate its performance using single month returns. For this, merely the QR factor was used as a regressor. With the LM tests, no random effects were found. The F-test indicated a fixed firm effect. Robust covariance matrices were used in the calculations of standard errors. Similar findings were observed as in the pooled model. The relationship between the QR factor and the abnormal returns in the period ranging from the third up until the sixth month after an event, was markedly weaker than in non-event periods. After skipping 9 months, the QR factor was found to have a consistently positive effect on the post-event abnormal returns. Overall, the findings with the panel model were in line with that of the pooled model.

Summarizing, we can conclude that the models were not working properly for multiple months after an M&A event. Even though some factors are significant, the model's ability to explain the abnormal returns is rather poor compared to that in non-event periods. Similarly, the QR factor is not significantly related with the abnormal returns for several months of the first half year after an M&A event. These results were obtained with pooled regression models and confirmed by the outcomes of panel regression models. As a consequence, the assumption that the model is not working properly can be deemed valid. However, our findings imply that the current override period of 12 months is somewhat overstated. A positive and significant relationship is found by the QR model from 9 months after an M&A event onward, meaning that one can use the average factor performance to explain the monthly abnormal returns. Hence, a reduction of 3 months in the override period would be appropriate. Although these results follow from both the pooled regression model and the panel regression model, the latter one is found to be less suitable for this research.

6.2. Shortcomings

We have carried out this research as thoroughly as possible, but it is important to mention some limitations. A first limitation is posed in the method used for the selection of events. For this thesis, we are using the RavenPack News Analytics database to identify whether an M&A event has occurred for a particular observation. The events are extracted from news articles, and do not always include a correctly identified event. Even though the events were thoroughly reviewed, some erroneous observations may be present in the data. Furthermore, we saw that the composition of the news data is rather monotonous. Less than 5% of the total number of observations is for the acquiree side of an M&A deal. As a consequence, the acquirer side of the M&A deals is more highly represented.

A second limitation can be found in the panel model approach. The panel model is deployed to verify our findings with the pooled model, while modelling the unobserved effects. Due to some shortcomings and additional complexity, it is found that this approach is not the most suitable for our research. One of the limitations of the panel regression follows from the separation of the event and non-event coefficients. For the pooled regression, the data set could easily be split up. Using the right formulas, equal estimates could be obtained through ordinary least squares regression of an interaction model and by performing two separate regressions. Splitting the model for the event data and non-event data is more difficult for the panel model. The event observations are well-spread over time and firms, and do not actually form a panel on their own. Therefore, an interaction term must be used in the panel regression model. As a result of the transformation of the systems, required for both the fixed effects model and the random effects model, the interaction terms do not solely capture the difference between the coefficient estimates of the two samples. Consequently, the event coefficient can merely be approximated.

6.3. Future work

For further research it would be interesting to separate the models for acquirers and acquirees. The news data in this thesis is tilted towards the acquirers in an M&A deal, as it includes relatively few events for acquirees. If enough observations are available, one could separate the two parties and compare the explanatory power of the factors on post-event abnormal returns. Different outcomes might suggest a tailored length of override periods for either side of the deal. For this, we suggest the use of a different database for the identification of events, so that both acquirers and acquirees are more equally represented in the data. A database that has been used for the announcement of the events to portfolio managers is favourable, so that each event will be included which historically yielded a model override.

The panel model was implemented to include the unobserved effects. However, it appeared that these models were not able to capture all correlations induced by the firm effect. The residuals were found to be serially correlated, rendering the coefficient estimates inefficient and the standard errors biased. In such cases, Greene [26] mentions that one might want to use time series in the panel model, rather than using robust covariance matrices. Alternatively, an estimation procedure with less restrictive assumptions on the independence of residuals would be interesting, as the results of the pooled model implied the presence of correlation along both the time and firm dimension of the panel. One could think of a method like generalized estimating equations, which is extensively treated in Hardin & Hilbe [27]. Potentially, the use of these methods will yield other outcomes, and it would be interesting to compare these to what we have found in this thesis.

A

Fixed effects factor model

Table A.1 contains the regression coefficients for the non-event firms in the panel regression where the cumulative abnormal returns are used as dependent variable with skip months. One can observe that the coefficient estimates and t-statistics are more or less equal for the different skipping periods.

Table A.1: Fixed effects model for non-event observations

	CAR(4, 6)	CAR(7, 9)	CAR(10, 12)	CAR(13, 15)
Momentum	0.36	0.36	0.36	0.34
	8.06	8.00	7.84	7.18
Quality	-0.31	-0.32	-0.33	-0.34
	-4.33	-4.35	-4.43	-4.54
Reversal	-0.15	-0.16	-0.17	-0.18
	-6.91	-7.11	-7.32	-7.55
Risk	0.05	0.04	0.03	0.02
	0.59	0.53	0.34	0.21
Size	-3.36	-3.36	-3.34	-3.34
	-28.69	-28.40	-28.12	-27.72
Value	0.98	1.00	1.00	1.02
	12.48	12.39	12.19	12.17
Observations	450,276	435,503	422,383	410,564

The table contains the fixed effects model estimates with robust t-statistics for the non-event observations.

The tables below contain the robust covariance matrices that have been used for the fixed effects model of Section 5.3.1:

Table A.2: CAR(1, 3)

	Momentum	EventDummy	Quality	Reversal	Risk	Size	Value	ED*Momentum	ED*Quality	ED*Reversal	ED*Risk	ED*Size	ED*Value
Momentum	0.0019	0.0000	-0.0002	0.0002	-0.0002	-0.0001	0.0009	-0.0015	0.0002	0.0001	-0.0003	-0.0001	0.0000
EventDummy	0.0000	0.0420	-0.0002	0.0000	-0.0004	0.0004	0.0003	-0.0093	-0.0103	-0.0046	0.0069	-0.0161	-0.0079
Quality	-0.0002	-0.0002	0.0049	0.0000	0.0000	-0.0021	-0.0021	-0.0002	-0.0014	0.0003	0.0000	0.0006	-0.0001
Reversal	0.0002	0.0000	0.0000	0.0005	0.0000	-0.0001	0.0003	-0.0002	-0.0001	-0.0004	-0.0001	0.0001	-0.0003
Risk	-0.0002	-0.0004	0.0000	0.0000	0.0058	-0.0001	0.0003	0.0003	-0.0001	-0.0003	-0.0016	-0.0001	-0.0002
Size	-0.0001	0.0004	-0.0021	-0.0001	-0.0001	0.0132	0.0023	-0.0014	0.0002	0.0004	0.0007	-0.0006	-0.0006
Value	0.0009	0.0003	-0.0021	0.0003	0.0003	0.0023	0.0060	-0.0001	-0.0001	0.0000	0.0002	-0.0003	-0.0015
ED*Momentum	-0.0015	-0.0093	-0.0002	-0.0002	0.0003	-0.0014	-0.0001	0.0495	0.0056	-0.0007	-0.0053	-0.0015	0.0109
ED*Quality	0.0002	-0.0103	-0.0014	-0.0001	-0.0001	0.0002	-0.0001	0.0056	0.0369	0.0020	-0.0026	-0.0017	0.0076
ED*Reversal	0.0001	-0.0046	0.0003	-0.0004	-0.0003	0.0004	0.0000	-0.0007	0.0020	0.0298	-0.0060	-0.0003	0.0056
ED*Risk	-0.0003	0.0069	0.0000	-0.0001	-0.0016	0.0007	0.0002	-0.0053	-0.0026	-0.0060	0.0457	0.0051	-0.0096
ED*Size	-0.0001	-0.0161	0.0006	0.0001	-0.0001	-0.0006	-0.0003	-0.0015	-0.0017	-0.0003	0.0051	0.0243	-0.0028
ED*Value	0.0000	-0.0079	-0.0001	-0.0003	-0.0002	-0.0006	-0.0015	0.0109	0.0076	0.0056	-0.0096	-0.0028	0.0446

Table A.3: CAR(4, 6)

	Momentum	EventDummy	Quality	Reversal	Risk	Size	Value	ED*Momentum	ED*Quality	ED*Reversal	ED*Risk	ED*Size	ED*Value
Momentum	0.0020	-0.0001	-0.0002	0.0002	-0.0002	-0.0001	0.0009	-0.0008	0.0002	0.0002	0.0000	-0.0001	-0.0005
EventDummy	-0.0001	0.0384	0.0002	0.0000	-0.0007	0.0007	0.0004	-0.0058	-0.0049	-0.0012	0.0036	-0.0155	-0.0050
Quality	-0.0002	0.0002	0.0051	0.0000	0.0000	-0.0022	-0.0022	-0.0003	-0.0012	-0.0002	-0.0003	-0.0001	-0.0002
Reversal	0.0002	0.0000	0.0000	0.0005	0.0000	-0.0001	0.0003	-0.0003	-0.0001	-0.0001	0.0000	0.0000	0.0001
Risk	-0.0002	-0.0007	0.0000	0.0000	0.0059	-0.0002	0.0003	0.0003	0.0000	0.0000	-0.0007	0.0005	0.0006
Size	-0.0001	0.0007	-0.0022	-0.0001	-0.0002	0.0137	0.0025	0.0008	0.0004	0.0000	0.0003	-0.0011	-0.0004
Value	0.0009	0.0004	-0.0022	0.0003	0.0003	0.0025	0.0062	0.0000	0.0006	-0.0003	0.0005	-0.0002	-0.0007
ED*Momentum	-0.0008	-0.0058	-0.0003	-0.0003	0.0003	0.0008	0.0000	0.0496	-0.0013	-0.0026	-0.0080	0.0001	0.0052
ED*Quality	0.0002	-0.0049	-0.0012	-0.0001	0.0000	0.0004	0.0006	-0.0013	0.0400	-0.0001	0.0070	-0.0038	0.0078
ED*Reversal	0.0002	-0.0012	-0.0002	-0.0001	0.0000	0.0000	-0.0003	-0.0026	-0.0001	0.0293	0.0000	-0.0022	0.0035
ED*Risk	0.0000	0.0036	-0.0003	0.0000	-0.0007	0.0003	0.0005	-0.0080	0.0070	0.0000	0.0486	0.0055	-0.0019
ED*Size	-0.0001	-0.0155	-0.0001	0.0000	0.0005	-0.0011	-0.0002	0.0001	-0.0038	-0.0022	0.0055	0.0210	-0.0030
ED*Value	-0.0005	-0.0050	-0.0002	0.0001	0.0006	-0.0004	-0.0007	0.0052	0.0078	0.0035	-0.0019	-0.0030	0.0534

Table A.4: CAR(7, 9)

	Momentum	EventDummy	Quality	Reversal	Risk	Size	Value	ED*Momentum	ED*Quality	ED*Reversal	ED*Risk	ED*Size	ED*Value
Momentum	0.0021	0.0001	-0.0002	0.0002	-0.0002	-0.0001	0.0010	-0.0004	-0.0002	0.0000	0.0001	0.0000	0.0004
EventDummy	0.0001	0.0427	0.0002	-0.0001	0.0000	0.0003	-0.0001	-0.0019	-0.0016	-0.0051	0.0126	-0.0177	-0.0006
Quality	-0.0002	0.0002	0.0053	0.0000	0.0001	-0.0023	-0.0023	-0.0008	-0.0002	-0.0005	-0.0001	-0.0002	-0.0002
Reversal	0.0002	-0.0001	0.0000	0.0005	0.0000	-0.0001	0.0003	-0.0002	0.0000	-0.0002	0.0001	0.0002	-0.0001
Risk	-0.0002	0.0000	0.0001	0.0000	0.0062	-0.0002	0.0003	0.0000	0.0001	0.0005	-0.0015	0.0000	-0.0005
Size	-0.0001	0.0003	-0.0023	-0.0001	-0.0002	0.0140	0.0025	0.0003	0.0001	0.0010	0.0005	-0.0005	-0.0020
Value	0.0010	-0.0001	-0.0023	0.0003	0.0003	0.0025	0.0065	0.0005	-0.0009	0.0001	-0.0003	0.0005	-0.0014
ED*Momentum	-0.0004	-0.0019	-0.0008	-0.0002	0.0000	0.0003	0.0005	0.0652	0.0007	-0.0047	0.0037	-0.0032	0.0236
ED*Quality	-0.0002	-0.0016	-0.0002	0.0000	0.0001	0.0001	-0.0009	0.0007	0.0468	-0.0007	0.0118	-0.0047	0.0072
ED*Reversal	0.0000	-0.0051	-0.0005	-0.0002	0.0005	0.0010	0.0001	-0.0047	-0.0007	0.0368	-0.0045	0.0015	0.0037
ED*Risk	0.0001	0.0126	-0.0001	0.0001	-0.0015	0.0005	-0.0003	0.0037	0.0118	-0.0045	0.0633	-0.0057	0.0141
ED*Size	0.0000	-0.0177	-0.0002	0.0002	0.0000	-0.0005	0.0005	-0.0032	-0.0047	0.0015	-0.0057	0.0268	-0.0121
ED*Value	0.0004	-0.0006	-0.0002	-0.0001	-0.0005	-0.0020	-0.0014	0.0236	0.0072	0.0037	0.0141	-0.0121	0.0901

Table A.5: CAR(10, 12)

	Momentum	EventDummy	Quality	Reversal	Risk	Size	Value	ED*Momentum	ED*Quality	ED*Reversal	ED*Risk	ED*Size	ED*Value
Momentum	0.0021	-0.0001	-0.0002	0.0002	-0.0003	-0.0001	0.0010	-0.0010	0.0000	-0.0004	-0.0001	-0.0002	0.0001
EventDummy	-0.0001	0.0484	0.0005	0.0001	-0.0001	0.0001	0.0000	-0.0039	-0.0069	-0.0028	0.0108	-0.0185	-0.0071
Quality	-0.0002	0.0005	0.0055	0.0000	0.0000	-0.0023	-0.0023	0.0001	-0.0011	0.0001	-0.0001	0.0002	-0.0006
Reversal	0.0002	0.0001	0.0000	0.0005	0.0000	-0.0002	0.0003	-0.0001	-0.0001	-0.0003	0.0000	0.0000	0.0002
Risk	-0.0003	-0.0001	0.0000	0.0000	0.0065	-0.0003	0.0004	0.0002	0.0003	0.0002	-0.0017	-0.0005	-0.0001
Size	-0.0001	0.0001	-0.0023	-0.0002	-0.0003	0.0141	0.0026	-0.0013	0.0003	0.0004	0.0019	-0.0001	-0.0006
Value	0.0010	0.0000	-0.0023	0.0003	0.0004	0.0026	0.0068	0.0001	-0.0005	0.0000	-0.0003	-0.0001	-0.0010
ED*Momentum	-0.0010	-0.0039	0.0001	-0.0001	0.0002	-0.0013	0.0001	0.0670	0.0011	-0.0078	-0.0041	-0.0057	0.0111
ED*Quality	0.0000	-0.0069	-0.0011	-0.0001	0.0003	0.0003	-0.0005	0.0011	0.0494	-0.0033	0.0064	-0.0041	0.0029
ED*Reversal	-0.0004	-0.0028	0.0001	-0.0003	0.0002	0.0004	0.0000	-0.0078	-0.0033	0.0403	-0.0022	0.0020	0.0000
ED*Risk	-0.0001	0.0108	-0.0001	0.0000	-0.0017	0.0019	-0.0003	-0.0041	0.0064	-0.0022	0.0730	0.0022	-0.0004
ED*Size	-0.0002	-0.0185	0.0002	0.0000	-0.0005	-0.0001	-0.0001	-0.0057	-0.0041	0.0020	0.0022	0.0320	-0.0058
ED*Value	0.0001	-0.0071	-0.0006	0.0002	-0.0001	-0.0006	-0.0010	0.0111	0.0029	0.0000	-0.0004	-0.0058	0.0652

Table A.6: CAR(13, 15)

	Momentum	EventDummy	Quality	Reversal	Risk	Size	Value	ED*Momentum	ED*Quality	ED*Reversal	ED*Risk	ED*Size	ED*Value
Momentum	0.0022	0.0000	-0.0002	0.0002	-0.0003	-0.0001	0.0010	-0.0009	0.0003	-0.0001	-0.0002	-0.0002	-0.0002
EventDummy	0.0000	0.0618	0.0004	-0.0002	0.0003	-0.0006	-0.0002	-0.0085	-0.0148	-0.0135	0.0219	-0.0248	-0.0134
Quality	-0.0002	0.0004	0.0057	0.0000	0.0000	-0.0024	-0.0024	0.0005	-0.0010	-0.0002	0.0004	0.0001	-0.0002
Reversal	0.0002	-0.0002	0.0000	0.0005	0.0000	-0.0002	0.0003	-0.0001	0.0003	-0.0002	-0.0001	0.0001	-0.0003
Risk	-0.0003	0.0003	0.0000	0.0000	0.0067	-0.0002	0.0005	0.0001	-0.0006	0.0001	-0.0015	0.0001	-0.0014
Size	-0.0001	-0.0006	-0.0024	-0.0002	-0.0002	0.0145	0.0027	-0.0017	0.0006	0.0003	-0.0001	-0.0011	0.0009
Value	0.0010	-0.0002	-0.0024	0.0003	0.0005	0.0027	0.0070	-0.0003	-0.0008	0.0002	-0.0014	-0.0002	-0.0016
ED*Momentum	-0.0009	-0.0085	0.0005	-0.0001	0.0001	-0.0017	-0.0003	0.0937	0.0100	0.0088	-0.0194	-0.0043	0.0127
ED*Quality	0.0003	-0.0148	-0.0010	0.0003	-0.0006	0.0006	-0.0008	0.0100	0.0780	0.0154	-0.0030	0.0000	0.0126
ED*Reversal	-0.0001	-0.0135	-0.0002	-0.0002	0.0001	0.0003	0.0002	0.0088	0.0154	0.0679	-0.0141	0.0041	0.0110
ED*Risk	-0.0002	0.0219	0.0004	-0.0001	-0.0015	-0.0001	-0.0014	-0.0194	-0.0030	-0.0141	0.0866	-0.0050	-0.0061
ED*Size	-0.0002	-0.0248	0.0001	0.0001	0.0001	-0.0011	-0.0002	-0.0043	0.0000	0.0041	-0.0050	0.0362	0.0040
ED*Value	-0.0002	-0.0134	-0.0002	-0.0003	-0.0014	0.0009	-0.0016	0.0127	0.0126	0.0110	-0.0061	0.0040	0.0875

Bibliography

[1] Takeshi Amemiya. The estimation of the variance in a variance-components model. *International Economic Review*, 12(1):1–13, 1971. doi: 10.2307/2525492.

[2] Manuel Arellano. Computing robust standard errors for within-group estimators. *Oxford Bulletin of Economics and Statistics*, 49(4):431–434, 1987.

[3] Clifford S. Asness, Andrea Frazzini, and Lasse Heje Pedersen. Quality minus junk. *Review of Accounting Studies*, 24(1):34–112, 2019. doi: 10.1007/s11142-018-9470-2.

[4] Badi H. Baltagi. *Econometric Analysis of Panel Data*. John Wiley & Sons, Ltd, 2005. ISBN 9780470014561.

[5] Badi H. Baltagi and Ping X. Wu. Unequally spaced panel data regressions with ar(1) disturbances. *Econometric Theory*, 15(6):814–823, 1999.

[6] Badi H. Baltagi, Young-Jae Chang, and Qi Li. Testing for random individual and time effects using unbalanced panel data. *Advances in Econometrics*, 13(1):1–20, 1998. doi: 10.1108/S0731-9053(1999)0000013003.

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

[8] Brad M. Barber and John D. Lyon. Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics*, 43(3):341–372, 1997. doi: 10.1016/S0304-405X(96)00890-2.

[9] Sanjay Basu. Investment performance of common stocks in relation to their price-earnigs ratios: A test of the efficient market hypothesis. *Journal of Finance*, 32(3):663–682, 1977. doi: 10.1111/j.1540-6261.1977.tb01979.x.

[10] A. Bhargava, L. Franzini, and W. Narendranathan. Serial correlation and the fixed effects model. *The Review of Economic Studies*, 49(4):533–549, 1982. doi: 10.2307/2297285.

[11] John J. Binder. The event study methodology since 1969. *Review of Quantitative Finance and Accounting*, 11(2):111–137, 1998. doi: 10.1111/j.1540-6261.1996.tb05202.x.

[12] David Blitz and Pim van Vliet. The volatility effect: Low risk without lower return. *Journal of Portfolio Management*, 34(1):102–113, 2007. doi: 10.3905/jpm.2007.698039.

[13] Stephen J. Brown and Jerold B. Warner. Measuring security price performance. *Journal of Financial Economics*, 8(3):205–258, 1980. doi: 10.1016/0304-405X(80)90002-1.

[14] A Colin Cameron and Douglas L. Miller. A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372, 2015.

[15] A Colin Cameron, Jonah B. Gelbach, and Douglas L. Miller. Robust inference with multiway clustering. *Journal of Business Economic Statistics*, 29(2):238–249, 2011. doi: 10.1198/jbes.2010.07136.

[16] Cynthia J. Campbell, Arnold R. Cowan, and Valentina Salotti. Multi-country event-study methods. *Journal of Banking and Finance*, 34(12):3078–3090, 2010. doi: 10.1016/j.jbankfin.2010.07.016.

[17] Mark M. Carhart. On persistence of mutual fund performance. *Journal of Finance*, 52(1):57–82, 1997. doi: 10.1111/j.1540-6261.1997.tb03808.x.

[18] Louis K. C. Chan, Narasimhan Jegadeesh, and Josef Lakonishok. Momentum strategies. *Journal of Finance*, 51(5):1681–1713, 1996. doi: 10.1111/j.1540-6261.1996.tb05222.x.

[19] Ramesh Chandra, Shane Moriarty, and G. L. Willinger. A re-examination of the power alternative return-generating models and the effect of accounting for cross-sectional dependencies in event studies. *Journal of Accounting Research*, 28:398–408, 1990.

[20] Lili Dai, Rui Shen, and Bohui Zhang. Does the media spotlight burn or spur innovation?, 2017. URL `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3037838`.

[21] Eugene F. Fama. Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49(3):283–306, 1998. doi: 10.1016/S0304-405X(98)00026-9.

[22] Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, 1993. doi: 10.1016/0304-405X(93)90023-5.

[23] Eugene F. Fama and James D. MacBeth. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636, 1973.

[24] Wayne A. Fuller and George E. Battese. Estimation of linear models with crossed-error structure. *The Journal of Econometrics*, 2(1):67–78, 1974. doi: 10.1016/0304-4076(74)90030-X.

[25] Christian Gouriéroux, Alberto Holly, and Alain Monfort. Likelihood ratio test, wald test, and kuhn-tucker test in linear models with inequality constraints on the regression parameters. *Econometrica*, 50(1):63–80, 1982. doi: 10.2307/1912529.

[26] William H. Greene. *Econometric Analysis Seventh Edition*. Pearson Education, 2012. ISBN 9780273753568.

[27] James W. Hardin and Joseph M. Hilbe. *Generalized Estimating Equations*. CRC Press, 2013. ISBN 9781439881149.

[28] Jerry A. Hausman. Specification tests in econometrics. *Econometrica*, 46(6):1251–1271, 1978. doi: 10.2307/1913827.

[29] Yuzo Honda. Testing the error components model with non-normal disturbances. *The Review of Economic Studies*, 52(4):681–690, 1985. doi: 10.2307/2297739.

[30] Boris Iglewicz and David C. Hoaglin. *How to Detect and Handle Outliers*. ASQC, 1993. ISBN 087389247X.

[31] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning*. Springer, 2013. ISBN 9781461471370.

[32] Narashimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance*, 48(1):65–91, 1993. doi: 10.1111/j.1540-6261.1993.tb04702.x.

[33] Narashimhan Jegadeesh and Sheridan Titman. Short-horizon return reversals and the bid-ask spread. *Journal of Financial Intermediation*, 4(2):116–132, 1995. doi: 10.1006/jfin.1995.1006.

[34] James Kirchner. Toolkit 12: Weighted averages and their uncertainties, 2006. URL `http://seismo.berkeley.edu/~kirchner/Toolkits/Toolkit_12.pdf`.

[35] S. P. Kothari and Jerold B. Warner. Econometrics of event studies. *Handbook of Empirical Corporate Finance*, 1:3–36, 2007. doi: 10.1016/B978-0-444-53265-7.50015-9.

[36] Thomas M. Krueger and Keith H. Johnson. Parameter specifications that make little difference in anomaly studies. *Journal of Business Finance & Accounting*, 18(4):567–582, 1991. doi: 10.1111/j.1468-5957.1991.tb01123.x.

[37] John Lintner. Security prices, risk, and maximal gains from diversification. *Journal of Finance*, 20(4):587–615, 1965. doi: 10.2307/2977249.

[38] John D. Lyon, Brad M. Barber, and Chih-Ling Tsai. Improved methods for tests of long-run abnormal stock returns. *Journal of Finance*, 54(1):165–201, 1999. doi: 10.1111/0022-1082.00101.

[39] Mark L. Mitchell and Erik Stafford. Managerial decisions and long-term stock price performance. *Journal of Business*, 73(3):287–329, 2000. doi: 10.1086/209645.

[40] Sara B. Moeller and et al. Wealth destruction on a massive scale? a study of acquiring-firm returns in the recent merger wave. *Journal of Finance*, 60(2):757–782, 2005. doi: 10.1111/j.1540-6261.2005.00745.x.

[41] Brent R. Moulton. Random group effects and the precision of regression estimates. *Journal of Econometrics*, 32(3):385–397, 1986. doi: 10.1016/0304-4076(86)90021-7.

[42] Brent R. Moulton and William C. Randolph. Alternative tests of the error components model. *Econometrica*, 57(3):685–693, 1989. doi: 10.2307/1911059.

[43] MSCI. Msci world index (usd), 2020. URL `https://www.msci.com/documents/10199/149ed7bc-316e-4b4c-8ea4-43fcb5bd6523f`.

[44] MSCI. Msci emerging markets index (usd), 2020. URL `https://www.msci.com/documents/10199/c0db0a48-01f2-4ba9-ad01-226fd5678111`.

[45] Yair Mundlak. On the pooling of time series and cross section data. *Econometrica*, 46(1):69–85, 1978. doi: 10.2307/1913646.

[46] Mitchell A. Petersen. Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1):435–480, 2009. doi: 10.1093/rfs/hhn053.

[47] Barr Rosenberg, Kenneth Reid, and Ronald Lanstein. Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, 11(3):9–16, 1985. doi: 10.3905/jpm.1985.409007.

[48] William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964. doi: 10.1111/j.1540-6261.1964.tb02865.x.

[49] P.A.V.B. Swamy and S.S. Arora. The exact finite sample properties of the estimators of coefficients in the error components regression models. *Econometrica*, 40(2):261–275, 1972. doi: 10.2307/1909405.

[50] Rex Thompson. Empirical methods of event studies in corporate finance. *Handbooks in Operations Research and Management Science*, 9:963–992, 1995. doi: 10.1016/S0927-0507(05)80073-8.

[51] Samuel B. Thompson. Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, 99(1):1–10, 2011. doi: 10.1016/j.jfineco.2010.08.016.

[52] Marno Verbeek. *A Guide to Modern Econometrics.* John Wiley Sons, Ltd, 2004. ISBN 9780470857731.

[53] T.D. Wallace and Ashiq Hussein. The use of error components models in combining cross section with time series data. *Econometrica*, 37(1):55–72, 1969. doi: 10.2307/1909205.

[54] Jeffrey M. Wooldridge. *Econometric Analysis Of Cross Section and Panel Data.* The MIT Press, 2010. ISBN 9780262232586.

[55] Jeffrey M. Wooldridge. *Introductory Econometrics: A Modern Approach.* Southern-Western, 2012. ISBN 9781111531041.