Turun yliopisto
University of Turku

INFORMATION CONTENT OF BANK CREDIT RATING CHANGES

Evidence from Europe

Master´s Thesis
in Accounting and Finance

Author:
Lauri Inkinen

Supervisors:
Prof. Hannu Schadewitz
Ph.D. Vesa Partanen

1.8.2016
Turku

Turun kauppakorkeakoulu • Turku School of Economics

Table of contents

Appendices

List of figures

List of tables

1 INTRODUCTION

1.1 Growing importance of credit ratings

Partnoy (2006, 61) compresses the paradox of the existence of credit ratings as follows:

> *"Rating changes are important, yet they possess little informational value. Credit ratings do not help parties manage risk, yet parties increasingly rely on ratings. Credit rating agencies are not widely respected among sophisticated market participants, yet their franchise is increasingly valuable. The agencies argue that they are merely financial journalists publishing opinions, yet ratings are far more valuable than the opinions of even the most prominent and respected financial publishers."*

Since the beginning of the rating industry, people have raised questions about the existence of credit ratings. This is certainly not surprising. In just one century the industry has managed to grow huge in size and has created a strong regulatory barrier for newcomers to enter into the market. This has kept the number of rating agencies rather small through years. The major participants dominating the industry with a market share of some 95 percent, Standard & Poor's (S&P), Moody's Investors Service and Fitch Ratings, have been around since the inception of the industry. (White 2010)

A matter of concern is the fact that the whole financial system is crucially dependent on credit ratings. Credit ratings, among other things, ameliorate information asymmetry between issuers and investors, as well as are heavily used in financial regulation (Rhee 2015, 162–163). Thus, although credit ratings are argued to be merely publishing opinions of financial journalists, they are not absolutely consumed as those (Partnoy 2006). As ratings are produced by humans, mistakes and conflicts of interest can never fully be avoided. This, and the oligopolistic industry created with the help of regulators, naturally raises problems and criticism.

The debate has been especially vigorous over the 21st century. There are several reasons explaining this. First, everyone remembers the worn-out examples of Enron and WorldCom which received fine ratings just before their collapses in the beginning of the century. Second, rating agencies are blamed for playing a key role in the 2007–2008 financial crisis due to their inadequate performance, mostly related to structured debt products (see e.g. Pagano & Volpino 2010). Besides, during the crisis even a larger set of investment grade companies were graded to junk shortly before defaulting.

Structured debt obligations are a good example of a new area in the world of finance that requires high expertise to understand the content. Thus, it is no wonder that a growing amount of income for rating agencies has recently come from this source (Pagano &

Volpino 2010, 407–408). However, as the case of collateral debt obligations during the crisis reveals, it was challenging for even experts of rating agencies do deal with the complex content of these securities.

After the financial crisis, began the European sovereign debt crisis of 2010–2012. During this time, European Union officials blamed rating downgrades for accelerating the crisis (CFR 2015). Sovereign ratings in Europe are surely not the only source of debate over the last years. Few would have guessed that it would be possible to downgrade the rating of the United States federal government, until S&P reduced the rating from "outstanding" to "excellent" in 2011. It is curious that in the current financial system this kind of power has been given mainly to three private companies. These same firms give ratings from small singular securities to entire nations.

The recent debate on credit ratings has been closely related to banks. After the financial crisis, popular indignation many times focused on credit ratings assigned to banks that enjoyed good ratings just before defaulting. Likewise, the problematic structured debt obligations were sold by banks. Credit ratings are also in the core of banking regulation as they can be used to determine how much capital banks need to hold in order to cover their credit risk. Basel II, the second and the recent implementation of the international bank regulatory framework reform by the Basel Committee on Banking Supervision, has since been criticized for increasing the reliance on external credit ratings. (Hau, Langfield & Marques-Ibanez 2013, 291–292, 296). In response, the still ongoing development of Basel III is aiming to reduce this reliance and has announced proposals for alternatives methods to measure bank credit risk (e.g. BCBS 2015). Anyhow, because of substantial negative feedback that some of these proposals have received from banks using external ratings, it seems that external credit ratings will also have a place in banking regulation in the future, at least to some extent.

As the dominance of the Big Three rating agencies and the systemic importance of credit ratings are likely to continue also in the future, the question whether rating agencies do in fact provide accurate and new information on the market is interesting. Due to the substantial amount of criticism ratings have attracted, there exists many research papers studying these questions in form or another. For example, rating agencies are criticized about moving too slowly and their ratings are said to be too inflexible (Altman & Saunders 2001, 6). A recent study of Hau et al. (2013) reveals that differential risk weights recommended by the Basel regulation have no significant relationship to empirical default probabilities in the case of investment grade banks. In other words, if two banks are given upscale ratings, such as AAA and BBB, they must hold different amounts of capital to cover their credit risk even though their credit ratings may not reveal any information of the probability of future default. Besides, the authors find some evidence that on average large banks receive more positive ratings compared to smaller ones. If these findings are true, the use of external credit ratings is obviously not the ideal way to calculate bank risk weights as it brings inequality and bias to the system.

One stream of credit rating studies investigates the market response to rating announcements. These studies generally concentrate on bond and stock prices and most recently on credit default swap spreads. Despite some of the more contradictory and mixed study results, the trend among these studies seems to be that the market reacts asymmetrically to the announcements of rating upgrades and downgrades. This is especially true in the case of stock prices. While rating downgrades are typically associated with significantly negative abnormal returns, the reaction among rating upgrades is often more limited or insignificant (e.g. Holthausen & Leftwich 1986; Hand, Holthausen & Leftwich 1992; Goh & Ederington 1993; Fieberg, Körner, Prokop & Varmaz 2015).

In the past, stock market studies have mainly concentrated on the United States market (e.g. Pinches & Singleton 1978; Holthausen & Leftwich 1986). Since the 1990s papers from other markets have also started to show up (e.g. Matolcsy & Lianto 1995; Elayan, Hsu & Meyer 2003). Still, there is a smaller amount of papers using European stock market data. Besides, only few papers focus directly on banking industry (see e.g. Gropp & Richards 2001; Fieberg et al. 2015). This is a somewhat surprising taking into account the central role of the industry. Thus, further research in this area is relevant, in order to explore whether the results match with studies using data from other markets.

1.2 Study objective

The main objective of this study is to examine the short-term effect of credit rating announcements on daily stock returns for European banks indexed in STOXX Europe 600 Banks. This is conducted by deriving relevant research hypotheses from previous literature. These hypotheses are formed to explore:

- whether rating announcements have an effect on stock returns
- when this possible reaction occurs
- whether the reaction is asymmetric between rating upgrades and downgrades.

These hypotheses are answered by conducting an event study, which is a common statistical methodology used in this research field. Besides the main hypotheses, a variety of additional tests are conducted to investigate other potential determinants of the returns surrounding rating changes. The set of additional tests includes typical issues and determinants studied in this research area. Credit rating announcements are gathered from the beginning of 2002 till the end of 2015 and represent issuer ratings produced by S&P, Moody’s and Fitch.

To the knowledge of the author, there have not been any recent published academic papers examining stock price effects of credit rating announcements particularly on the

European banking sector. Probably the latest paper with a similar focus was conducted by Gropp and Richards (2001). However, their sample size was smaller and they did not run as many additional tests as is done in this thesis. Anyhow, there is one very recent and comprehensive paper conducted by Fieberg et al. (2015) who used a broad banks-specific sample from 76 countries worldwide. Thus, their focus on the banking sector is much more universal. Besides, they mainly focus on how the size of a bank and the outbreak of the subprime crisis affect the results. Inspired by them, these two issues are also paid attention to in this thesis, although on a smaller scale.

All in all, the majority of event studies concerning stock price effects of credit rating changes focus on the United States market using data from several industries. Corresponding studies using European, and specifically industry-specific data, are less common. The banking sector is especially interesting due to its unique characteristics related to high regulation, opaqueness and global systemic importance.

This study unites both, Europe and the banking sector, providing evidence on whether bank stocks are affected by rating announcements in the European context. European economies provide a heterogeneous environment which is much more bank driven, if compared, for example, to the United States market. In the United States it is easier to get funding outside the commercial banking system leading to a situation where bonds, and thus credit ratings, play a more significant role in the financial system (Cancian 2016). These fundamental differences between the United States and Europe are likely to be seen in the stock price reaction of credit rating announcements.

Although this thesis has mainly theoretical contributions to the existing literature, it may provide some useful abnormal return information for investors interested in the European bank market. After all, if credit rating announcements do provide pricing pertinent information, it could be used in trading strategies. On the other hand, bank executives might be interested whether rating announcements have any short-term impact on bank market values.

1.3 Event study methodology

To explore whether credit rating announcements do produce new information on the market, a standard event study approach is used in this thesis. An event study is a statistical method used to study the effect of a specific event on the value of a company using financial market data. Event study methodology has been one of the most common used tools in financial research in the past decades due to its simplicity, although, there appears to be some criticism as well (Wells 2004, 61, 66).

Event studies have a long history in academic research. Presumably the first published event study was conducted by Dolley in 1933 who studied price effects of stock splits.

However, the first more well-known and more groundbreaking event studies in academia were conducted in the late 1960s by Ball and Brown (1968) and Fama, Fisher, Jensen & Roll (1969). Ball and Brown studied the information content of earnings and Fama et al. examined the effect of stock splits after removing dividend increases. These two research papers introduced the event study methodology which basically follows the same principles as studies conducted today. (MacKinlay 1997, 14)

Over the next decades, event studies became a common tool in various research areas and has had numerous methodological applications and extensions. Besides methodological development, a remarkable matter for event study methodology has been the increasing availability of daily data (and sometimes intraday) instead of monthly data. The methodology differs between short-term and long-term event studies. While short-time event studies try to explore the price behavior around the event days, long-term studies investigate how the events affect the prices over long periods of time, generally the event window being one year or more. Long-term event studies are shown to be less reliable in general due to the amount of their limitations. For instance, long-horizon tests are highly susceptible to the joint-test problem and have lower power. (Kothari & Warner 2004, 8–9).

Event studies dominate the empirical research especially in the area of corporate finance (MacKinlay 1997, 36). One of the most common examples is to study the effects of financial decisions on the value of the company that is often done by using returns. Besides returns, trading volume and volatility are sometimes used in the area of finance and accounting. Typical events include, for example, mergers and acquisitions activity, earnings announcements, capital structure changes, dividend changes and credit rating changes. Event studies also serve a significant purpose in capital market research as they are used to test market efficiency. Nonzero abnormal security returns are inconsistent with market efficiency if they systematically persist after a particular type of corporate event (Kothari and Warner 2004). Besides finance, accounting and economics, event studies are also conducted in disciplines, such as, management, marketing, law, history and political science (Corrado 2011, 207). Due to the versatile properties of the methodology, it can be easily modified to any research problem where the effects of a specified event are studied on time series.

In this thesis the event study is used to statistically test whether an economic event has an impact on the value of a firm using financial market data. This is done by separating the effect of the event from other market movements to get the pure influence of the event. This difference between the realized return and the expected return is called the abnormal return. The abnormal return may be either positive or negative and it simply implies how the security has performed over a given period of time. After determining the abnormal return, there are two ways to proceed with the study. The first method is more traditional in which the statistical significance of abnormal returns is tested individually each day or cumulative. The other approach is to identify variables that should explain the reaction. This is done by using a regression analysis to get the parameter values for each variable explaining the effect.

Conducting an event study includes the following steps:

1. defining the research hypotheses
2. defining the dates upon which the market has received the news associated with the events
3. collecting time series data
4. estimating the expected returns for each firm using the data outside the event days
5. calculating the abnormal returns
6. testing the statistical and economical significance for the specified hypotheses.

The more technical aspects of conducting an event study, including the methodological choices and the data used in this thesis, are further introduced in Chapter 4.

1.4 Thesis structure

The remainder of this paper is organized as follows. Chapter 2 will introduce the framework and previous literature on the subject to get an insight what is done in this study. This chapter is divided into three sections. The first one focuses on the credit rating industry and is divided into five subsections. The first subsection gives a brief introduction of the history and the current state of the credit rating industry. This information is given to better understand the nature of the industry and to perceive why the Big Three agencies have been chosen for this study. The second subsection is the most theoretical one as it explains reasons for the existence of credit ratings. The third subsection introduces rating types and scales. The information provided by this section is vital to understand the data used in this thesis. The next subsection describes the rating process. After all, to be able to question the quality of ratings, one must know the process every rating has to go through before publishing. As already presented in the introductory section, there exists plenty of criticism concerning ratings. The fifth and the final subsection opens up these aspects of criticism further. Previous research on the topic is often based on criticism about ratings and exists to increase information and to offer possible solutions to problems.

The second section of Chapter 2 introduces features of the banking sector and why credit ratings are closely related to banks. This section consists of three subsections. The first one specifies the unique characteristics of the banking sector. Taking into consideration the theme of this thesis, the results concerning regulated financial companies may be different compared with other industries. The idea of the second subsection is to describe how credit ratings are used in banking regulation. This includes the introduction of the Basel framework. As this study is conducted using bank data and ratings serve a special purpose in banking regulation, this subsection should give further motivation to study the features of credit ratings. The third subsection concentrates on the characteristics of the European banking sector as it is the matter of focus in this study.

Finally the third section of Chapter 2 introduces the previous research on the investigation of the market reaction of rating announcements. This section will begin with a brief introduction of the research field. The second subsection gives a few examples of studies concerning debt market reaction. Although these studies are not directly related to the topic of this thesis, a majority of previous research focuses on debt market reaction including corporate and sovereign debt as well as credit default swaps. Thus, it is of use to briefly introduce this topic. Besides, there are similarities between the responsiveness of debt and stock markets. The third and the fourth subsections focus on stock market reaction. The third subsection gives an overall view of what has been studied. In order to make this part more convenient to read, the study results are emphasized, instead of data and methodological choices. The last subsection concentrates on banks. The included study results are explained in more details depending on how relevant the concerned study is for this thesis.

Chapter 3 introduces the research hypotheses based on the literature used in research for this thesis. Besides the main hypotheses, this chapter explains what types of additional tests are conducted to investigate potential determinants surrounding rating changes. Chapter 4 describes the empirical methodology and choices as well as the data. The first part of the chapter includes the details of the conducted event study and cross-sectional regression analysis. Also, some problems and limitations related to this study and event studies in general are introduced. The end of the chapter gives details of the data.

The results of the research are presented in Chapter 5. The fifth chapter is further divided into two sections: credit rating changes and credit rating watches. The results concerning rating downgrades and upgrades are introduced separately in both sections. Chapter 6 summarizes and concludes the thesis.

2 THEORY AND LITERATURE

2.1 Credit rating industry

2.1.1 Development and current state of the industry

The credit rating industry is just about a century old, as the capital market was originally developed without rating agencies (see e.g. Sylla 2001). The rating industry has grown especially during the past few decades and has played a key role in the financial world, providing tools for investor to measure creditworthiness of various financial instruments.

The beginning of the industry originates in the United States in the early twentieth century. The prior small-scale credit reporting was not enough to satisfy the rising need for information that was derived from the expansive growth of the bond market. This expansion resulted mainly from the construction of extensive railroad systems which was primarily financed by corporate bonds. (Sylla 2001, 6–7)

The first security ratings were published by Moody's in 1909. Moody's provided a single rating symbol for each security to make complex data reports more user-friendly. In the next few years, Moody's encountered competition from Poor's Publishing Company in 1916, Standard Statistics Company in 1922 and the Fitch Publishing Company in 1924. Later Poor's Publishing Company and Standard Statistics Company merged to become Standard & Poor's in 1941. Although additional rating agencies were formed in the following years, these three original companies have dominated the industry to this date. (White 2010, 211)

During the industry's first decades, credit ratings were sold to investors. However, in the early 1970s this business model changed among the major rating agencies. In the new model, the issuer, instead of the investor, pays the rating agency for a rating. There have been multiple suggestions to explain this change. One plausible explanation could be the spreading of the high-speed photocopy machine. This caused the free rider problem, as it was easy to share the information once produced by a rating firm. Other explanations include, for example, the default of Penn-Central Railroad in 1970 that damaged the long period of economic stability. This shocked the bond markets and started a liquidity crisis that might have been a reason for debt issuers to want to pay for rating firms to ensure their ratings. Nevertheless, the introduction of the new "issuer pays" business model created potential conflicts of interest to the system as most agencies were, and still are, paid by the same firms they get ratings from. (White 2010, 214–215).

The usage of credit ratings in regulation began in the 1930s when bank regulators decided to encourage banks to invest only in safe bonds. In 1936 a decree was set that prohibited banks from investing below investment grade bonds. The following years also marked the beginning of the use of credit ratings in insurance regulation. Insurance companies got to face new minimum capital requirements in their investment strategies and these requirements were geared to credit ratings. Also federal pension regulators followed a similar strategy. The ratings used in financial regulation had to be determined by “recognized ratings manuals”, meaning only S&P, Moody’s and Fitch. (White 2010, 212–214)

The next target group to face rating-related minimum capital requirements were broker-dealers in 1970s. These included major investment banks and securities firms. However, the Securities and Exchange Commission (SEC) was not satisfied with the vague definition of “recognized ratings manuals” and created a new category – “nationally recognized statistical rating organization” (NRSRO). Originally the SEC gave the title only to S&P, Moody’s and Fitch and stated that NRSROs are the only companies to give valid ratings used in broker-dealers regulation. Soon the category of NRSRO was also adopted by other financial regulators. The creation of NRSROs has been seen as one of the major events influencing the industry. It created a regulatory barrier for newcomers to enter into the market, which has kept the number of rating agencies rather small through years. (White 2010, 212–214)

Besides the creation of the NRSRO category, there are also other reasons which explain the concentration of the industry, such as; economies of scale, the advantages of experience and the brand name reputation (White 2010, 217). These features altogether make it extremely difficult for a new rating firm to enter into the market. Most newcomers concentrate on a small market sector in order to be able to start building their own reputation in that specialized area.

The credit rating industry has remained oligopolistic in its entire existence and the current market structure seems to remain stable, also in the future. The total revenue by all of the NRSROs for their 2014 fiscal year was approximately $5.9 billion (SEC 2015, 17). The two biggest agencies today, Standard & Poor's and Moody's Investors Service, control about 40 percent of the global market share each. The combined market share when taking Fitch Ratings also into consideration has been around 95 percent (White 2010, 216–217). This leaves the role of other rating agencies very marginal in the global scale. There were ten credit rating agencies registered as NRSROs by the end of year 2015. The following table lists the names, origins and their primary business areas. Interestingly enough, the SEC has given NRSRO status to three foreign companies yet none of them are European.

Table 1 Nationally recognized statistical rating organizations in 2015 (SEC 2015)

Name	**Origin**	**Primary area of focus**
A.M. Best Company, Inc.	US	Insurance
DBRS, Inc.	Canada	Canada
Egan-Jones Ratings Company	US	US
Fitch Ratings, Inc.	US	Global
HR Ratings de México, S.A. de C.V	Mexico	Mexico
Japan Credit Rating Agency, Ltd.	Japan	Japan
Kroll Bond Rating Agency, Inc.	US	Financial companies
Moody’s Investors Service, Inc.	US	Global
Morningstar Credit Ratings, LLC	US	Structured finance
Standard & Poor’s Ratings Services	US	Global

There is also a great amount of credit rating agencies that are not registered as NRSROs. The total amount of rating agencies around the world in 2010 was approximately 150. There are several dimensions to how the agencies differ from each other: business focus, rating methodology, pricing model and the type of scale. The three major agencies follow a similar pattern as they all have a global focus and provide cross-industry, issuer and instrument specific ratings. They have an analytical approach with committee reporting, follow an issuer-pays business model and use ordinal scales. All three major agencies have a solid, strong reputation. Investors do not commonly make radical presumptions on which of the companies performs best and in some markets it is a common habit to automatically take ratings from two or three agencies. (Langohr & Langohr 2010, 384–389). As choosing a major agency is the rule rather than an exception, the focus will be on these three companies for the rest of the paper.

Standard & Poor’s Rating Services is the largest of the three major agencies today based on its market share. It is a part of Standard & Poor's Financial Services that is a subsidiary of McGraw Hill Financial. Moody's Investors Service was owned by Dun & Bradstreet for many decades and became a separate company in 2000. Moody’s Investors Service is now owned by Moody’s Corporation. Unlike S&P and Fitch, Moody’s Corporation is a free-standing company. Fitch Ratings is the smallest of the three major agencies when comparing their market shares. Fitch Ratings is a part of the Fitch Group that is nowadays a jointly owned subsidiary of Hearst Corporation and FIMALAC. Fitch Group is the result of a complex series of mergers and acquisitions unlike S&P and Moody’s that have grown more organically. (SEC 2015; White 2010, 211)

The percentage of outstanding credit ratings of total ratings issued by NRSROs was nearly 49 percent for S&P in 2014, the total amount of ratings being 2.4 million. As comparison, for Moody's the same number was about 35 percent and for Fitch 12 percent. The amount of financial institutions of the total ratings produced by NRSROs was roughly 8 percent. (SEC 2015, 10–12).

For the rest of the paper, names S&P, Moody's and Fitch are used to refer to the three major rating agencies, not to their holding companies. When referring to the firms together, the term "Big Three" is used.

2.1.2 Functions of credit ratings

Credit rating agencies try to make it clear that ratings are not buy or sell recommendations, nor are they a guarantee that default will not occur. Ratings always represent agency's own opinions of the credit quality and the investor should always make an independent analysis of the creditworthiness of the assets and not lean solely on the rating. (see e.g. Fitch Ratings 2016). Taking this into consideration, it is not surprising that since the inception of rating agencies, people have raised questions about the existence and the role of the industry.

There is a strong market rationale for the existence of rating agencies. Two standard theories state that (1) rating agencies ameliorate *information asymmetry* between issuers and investors and (2) rating agencies reduce the cost of regulation. Information asymmetry refers to a situation where one party has more or better information than the other. In the case of financial markets, borrowers know more about their creditworthiness than creditors. This results in a situation where creditors raise rates for protection against lower quality borrowers driving higher quality borrowers out of the market. In economic literature this is called the "lemon" problem. Rating agencies are said to correct the problem as they act as an information intermediary producing independent information on the borrower's creditworthiness. (Rhee 2015, 162–163).

The other common explanation for the existence of rating agencies is that they exist because they reduce the net cost of regulation as investors and regulators do not have to erect an analytic infrastructure to analyze bond investments (Rhee 2015, 164). The use of credit ratings in financial regulation has grown in importance since the introduction of NRSROs in the 1970s. Nowadays states worldwide use credit ratings for protection against systematic risk. By preventing the accumulation of too much risk at certain points in the financial system, states aim for financial stability in order to increase confidence between all participants. (Dittrich 2007, 15). Even though the use of ratings in regulation is prevalent around the world, it is still most widespread in the United Sates. Compared with the United States, the rating-based regulation is much less common on the European level. Although several European nations use credit rating standards for similar purposes as the United States, the number of means of use does not come anywhere close to what it is in the States. (Langohr & Langohr 2010, 431)

In his study, Dittrich (2007) gives four practical points of view on why credit ratings are ideal instruments in financial regulation:

1. Credit ratings have proven to be efficient in their high correlation between risk categories and default rates.
2. Credit ratings are readily available to all market participants at no direct cost.
3. The need for continued detailed oversight can be kept at a minimum by matching market recognition and regulatory recognition of rating agencies.
4. Ratings are based on reputation and thoroughness, making them an ideal instrument to increase confidence.

According to Dittrich, these reasons altogether make ratings a rather simple instrument to influence the behavior of financial market participants.

According to Adams, Mathieson and Schinasi (1999, 200) rating-based regulation can be divided into three areas that are *investment restrictions*, *disclosure requirements* and *capital requirements*. Probably the most extensive use of ratings has been the investment restrictions on regulated institutions. These restrictions include prohibiting investments on low-rated or unrated securities to reduce the riskiness of the overall portfolio. Ratings are also used to define disclosure and issuance requirements. An adequate credit rating may lower legislative obligations. Ratings may also be a requirement for issuing certain types of securities. The last category covers all rules concerning capital requirements that are given to financial institutions in order to prevent financial market instability.

According to Rhee (2015), the two common theories stating that rating agencies ameliorate information asymmetry and reduce the cost of regulation do not fully answer the question of why credit ratings exist. In his study he suggests an alternative explanation. Even though rating agencies produce little new information to the market, they are needed to sort out the large volume of information in the credit market. This sorting function facilitates better credit analysis and investment selection and it cannot easily be replicated.

Langohr and Langohr (2010, 89–91) for their part give three more functional purposes of what credit ratings are for. First, credit ratings objectively *measure the credit risk* of the issuer's business and its debt financing. This resolves the aforementioned problem of information asymmetry. The second function is to provide a *means of comparison* between all issues of the credit risks embedded in them. This gives the investors the opportunity to compare the credit risk between all possible types of issues. The third function is that ratings provide market participants with a *common standard to refer* to credit risk. This means that a rating is public and all parties in a contract can observe it, and acknowledge the level of it, in the same way at any point of time.

In general, credit ratings simply make it easier for lenders and borrowers to face each other. Investors need information on the quality of their investments and issuers for their part need access to funds. Credit ratings reduce the cost of information for investors, and the cost of market access for issuers. (Langohr and Langohr 2010, 89–91). It is also good to perceive that without rating agencies, small and less informed investors might be left out of the market as they may not have the skills and resources to come up with necessary information to make investment decisions.

2.1.3 Rating types and scales

A credit rating could be generally defined as a method of measuring the creditworthiness of a debt issuer. It can be assigned to any entity that seeks to borrow money. This could be an individual, corporation, state or provincial authority or a sovereign government. Besides external credit ratings that are provided by external rating agencies such as S&P, Moody's and Fitch, many financial institutions have their own internal rating systems to represent the creditworthiness of their clients. These internal ratings (a.k.a. credit score) are for institution's personal use only and should not be confused with regulated external ratings.

There are different types of credit ratings that rating agencies produce. First, ratings can be divided into short-term and long-term credit ratings. Ratings express the likelihood that the borrower will go into default within a given time period. A short-term rating reflects the likelihood of the rated party defaulting within a year and a long-term rating over longer term. In the past long-term ratings have been more preferred but recently short-term ratings have become the norm as well. (Investopedia 2016)

Second, agencies produce issuer-specific ratings and issue-specific ratings. An issuer rating rates the issuer as a whole – it provides an overall assessment of a firm's creditworthiness. In contrast, an issue rating, also known as an instrument rating, deals with the performance of one particular debt instrument. Naturally issuer and instrument ratings are never fully independent of each other. (Langohr & Langohr 2010, 42–43). When, for instance, news journalists give credit rating disclosures, it is not always clearly specified which type of a rating is in question. For example, it is not necessarily an issuer rating that is changed when a company's credit rating is informed to be changed in the news. This may cause confusion.

Third, ratings are divided into local and foreign currency ratings. This refers to considerations of country and currency risk. Local currency ratings measure the likelihood of repayment in the currency of the jurisdiction. Foreign currency rating additionally considers the profile of the issuer after taking into account transfer and convertibility risk.

Both foreign currency and local currency ratings are internationally comparable assessments. (Fitch Ratings 2016a).

Different rating agencies use their own variation of an alphabetical combination of lower and upper-case letters to summarize their opinions about obligors in ratings. Plus and minus signs or numbers are further added to tune the ratings. In this thesis the focus is on long-term issuer-specific ratings provided by S&P, Moody's and Fitch. They all use ordinal credit rating scales that are categorized in Table 2. Even though the scales of these three companies look somewhat similar, it is good to keep in mind that each rating agency has its own definition of a credit rating. Thus, absolute linkage between the symbols between the different companies cannot be defined because the symbols do not measure the same thing (Watson 2006, 187). Ratings produced by S&P and Fitch measure the probability of default although their methods of analysis differ from each other. Moody's, on the other hand, is interested in the expected financial loss, not in default probability per se. However, considering the aim of this study, there is no need to further distinguish between different rating scales.

Table 2 Long-term issuer-specific rating classes

	Moody's	S&P	Fitch	Rating description
Investment grade	Aaa	AAA	AAA	Prime
	Aa1	AA+	AA+	High grade
	Aa2	AA	AA	
	Aa3	AA−	AA−	
	A1	A+	A+	Upper medium grade
	A2	A	A	
	A3	A−	A−	
	Baa1	BBB+	BBB+	Lower medium grade
	Baa2	BBB	BBB	
	Baa3	BBB−	BBB−	
Non-investment grade (a.k.a. speculative grade)	Ba1	BB+	BB+	Non-investment grade speculative
	Ba2	BB	BB	
	Ba3	BB−	BB−	
	B1	B+	B+	Highly speculative
	B2	B	B	
	B3	B−	B−	
	Caa1	CCC+	CCC+	Substantial risks
	Caa2	CCC	CCC	
	Caa3	CCC−	CCC−	
	Ca	CC	CC	Extremely speculative
	Ca	C	C	Default imminent
	C	R	RD	In default
	/	SD	D	
	/	D	/	

The rating scales in Table 2 measure the credit risk on a risk index and not an absolute measure. This means that B rated instruments are more likely to default than the A rated ones, but it does not give information on how much more or how much less likely (ordinal scale). (Langohr & Langohr 2010, 44). The most visible threshold, which is also used in regulation, is the line between the investment grade ratings and the speculative grade ratings. Other sub-categories are seldom used and their verbal descriptions vary. A non-investment grade bond is often called a junk bond in spoken language to refer to its higher default risk.

Besides issuer and instrument ratings, rating agencies also publish *rating watches* and *outlooks* to indicate the likely direction of a rating in the future. When a company is on the watch, it indicates that there is a heightened probability of a rating change in the short-term. The direction of the possible change may also be informed but it is not necessary. After a rating has been upgraded, downgraded or confirmed, it is removed from the watch. It is good to notice that the rating does not have to be placed on the watch list before the change takes place. (Fitch Ratings 2016b). A rating watch (Fitch) is also known as a credit watch (S&P) or a rating review (Moody's) depending on the agency. In this study the term rating watch is used as a synonym to refer to all of these short-term direction reviews.

The outlook is somewhat similar to the rating watch as it also indicates the direction a rating is likely to move to in the future. The main difference, however, is the time-horizon as the outlook indicates the likelihood of a rating to move over a one- to two-year period. Thus, outlooks reflect financial and other trends that may affect the ratings in the future. The direction of the possible change is also usually stated. The majority of outlooks are generally designated as stable. Just like in the case of rating watches, any given outlook does not imply that the rating will necessarily change. (Fitch Ratings 2016b)

2.1.4 Rating process

Rating agencies put their reputation on the line every time they give a new rating. Thus a new rating action is a high-risk decision for an agency. To be able to give a high-quality and accurate rating, every rating goes through a comprehensive process that focuses on objectivity, diligence and transparency. The rating process itself tends to be quite similar between rating agencies. Also, the largest agencies seem to follow similar procedures for similar types of instruments. (Langohr & Langohr 2010, 161, 187). Nevertheless, the quantitative and qualitative factors vary between companies and debt instruments. Some rating agencies are quite open in their rating procedures. For example, Moody's publishes methodology descriptions for different types of ratings on their website, such as, a report setting out a methodology for determining their bank ratings. Besides the actual report, Moody's provides tutorials and videos as guidance through the methodology.

The methodology report by Moody's is A detailed report of more than a hundred pages for determining bank ratings (see Moody's 2016). According to them, their methodology incorporates and builds upon their own research, their experience of the recent financial crisis and academic literature. The overall approach to rating bank instruments comprises the following steps:

- analyzing a bank's financial and operating environment (macro profile as well as financial and qualitative factors) to capture its standalone probability of failure in the absence of external support
- capturing the likelihood of affiliate support, such as, a parent, group or co-operative structure
- capturing the risks different creditors are exposed to in the event of the bank's failure
- capturing the extent to which risk to each creditor class is mitigated by public support.

After these steps the final credit rating for each rated instrument is given as well as the counterparty risk assessment. A counterparty risk assessment is an assessment of an issuer's ability to avoid defaulting on its operating, such as, non-debt or non-deposit obligations.

Figure 1 illustrates the rating process for ratings in the case of S&P. The process begins when the issuer requests a rating and signs an engagement letter. After the issuer has delivered necessary documentation, a team of analysts will review pertinent information. In the third phase analysts meet the management of the issuing company to review and discuss information. Some of this information may be confidential and it will not be published along with the rating outcome. The next step is the actual analysis where analysts evaluate information before proposing the rating to a rating committee. At the core of the rating process is the rating committee that reviews the lead analyst's rating recommendation and then votes on the credit rating. The pre-publication rationale for the rating will then be disclosed to the issuer. At this point the issuer will have a chance to appeal that leads to a new meeting with the management. If the issuer is satisfied with a rating, the rating will be published and disseminated. Once the rating is published, the agency will manage surveillance on the issuer to keep the rating current. (S&P 2016).

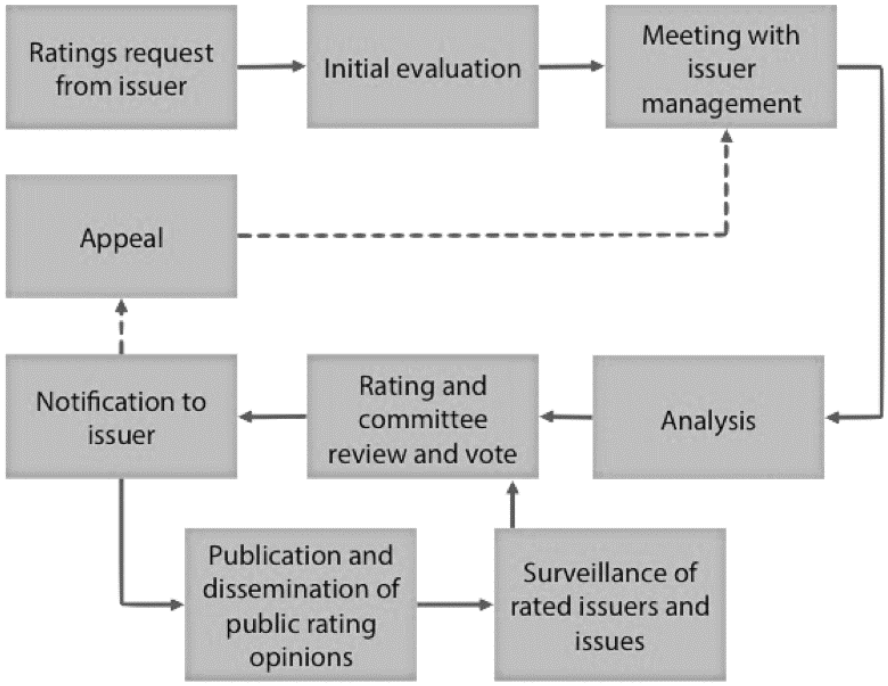

Figure 1 Rating process (S&P, 2015)

The figure of the rating process is not found on the website of Standard & Poor's anymore as such. The appealing part has been removed. This is an interesting finding and may implicate that agencies are trying to restrain so-called "rate shopping" which indicates that issuers are trying to buy the most favorable rating on the market they can get.

Some rating agencies produce unsolicited ratings – ratings that are not requested by the issuer. These ratings are usually based only on publicly available information and the issuer does not pay for these rating assessments. Unsolicited ratings are not as common as solicited ratings. As they just rely on publicly available information, their information content is questioned. Because issuers are encouraged to get as accurate a rating as possible, they are usually willing to pay for the rating. (Dittrich 2007, 111–113)

2.1.5 Criticism of rating agencies

The use of rating systems has received plenty of criticism in the past decades due to critical errors in ratings and poor operating practices. The criticism has increased especially after the 2007–2008 financial crisis, due to agencies inadequate performance, mostly related to structured debt products and to "serial downgrading" of securities during the crisis. During the European sovereign debt crisis of 2010–2012 European Union officials blamed rating downgrades for accelerating the crisis (CFR 2015). As a result, credit ratings have been placed under intense scrutiny and the regulation of the industry has increased. The dependence of the whole financial system on subjective credit ratings and their oligopolistic industry has been questioned. For example, one aim of Basel III is to reduce the reliance on external credit ratings in banking regulation (BCBS 2010).

Since credit rating agencies shifted from an “investor pays” to an “issuer pays” model in the 1970s, rating agencies have been criticized for their serious conflicts of interest for various reasons. When ratings are paid by the issuer, it creates an incentive for rating agencies to please their clients by distorting ratings to win further business or higher fees from them. Another problem is called *rate shopping*. Issuers are able to ask rating agencies how a certain financial instrument would be rated, or even how an instrument should be modified to obtain a certain rating. This kind of behavior leads to the situation where issuers choose the rating agency that gives the most favorable rating. Thirdly, credit ratings are largely immune to civil and criminal liability for malfeasance as they are considered as “journalists” by several United States court decisions. (Partnoy 2006)

Credit ratings had a significant role in the recent subprime mortgage crisis. The new complex securities used to finance subprime mortgages were rated by Big Three agencies. According to Pagano and Volpino (2010), a massive mispricing of risk of structured debt played a key role in the crisis. They state that the mispricing of risk is a result mainly from the aforementioned issues of conflicts of interest, such as rate shopping, creating rating inflation. Another major reason causing the crisis was that in the process of securitization, characteristics of the underlying assets were lost. Thus, investors did not understand the true nature of structured debt products. Pagano and Volpino argue that the reason why uninformative ratings were produced, was simply the aim to expand the primary market of those securities.

After the massive rating downgrades of structured products, rating accuracy, stability and forward-looking characteristics of ratings have been questioned. Big Three agencies also downgraded similar securities simultaneously which raises the question of the independence of ratings. (Baklanova 2009, 6–7). However, not all criticism comes from the failure of only structured products. For instance, famous past examples of rating failures, for being too slow to downgrade following credit quality deterioration, include companies such as Enron, WorldCom and Parmalat. S&P and Moody’s gave Enron a fine rating until four days before its collapse, WorldCom until three months before, and Parmalat until 45 days before (Langohr & Langohr 2010, 189).

Ratings are supposed to evaluate default risk over the economic cycle. Thus the recent period of massive downgrades of securities may not be fully justifiable (Hau et al. 2013, 291–292). Rating agencies tend to be quite careful when giving ratings so they prefer to downgrade gradually, one notch at a time, instead of decreasing a corporate’s rating many notches at once. Thus, during a period of recession, this phenomenon turns out in the form of serial downgrading.

Other much criticized areas include, for example, the subjective nature of ratings, the oligopolistic industry, regulation and supervisory practices, and recently the aggressiveness in rating Eurozone countries' creditworthiness. According to Hau et al. (2013, 296), high reliance on rating agencies increases the exposure of the financial system to the accuracy of credit ratings. Mistakes and biased forecasts that agencies make, have the potential to cause or exacerbate crises.

2.2 Features of the banking industry and the role of credit ratings

2.2.1 Overview of the industry

Bank stocks are among the hardest to analyze if compared with other industries. There are plenty of reasons to support this statement. Many banks hold billions of euros in current assets and have subsidiaries in other industries. The era of the Internet and a global trend of deregulation have opened up many new businesses to the banks, besides the fundamental role as a financial intermediary. The entire industry is extremely complex and requires high expertise to understand all the ins and outs of how banks function. (Investopedia 2014). Thus, information asymmetries between market participants are significant. The central role of the industry leads to high regulation by the society enabling regulatory interventions. Unlike in most other regulated industries, in the banking sector the main regulator, the central bank, is also a market participant.

All in all, banks are inherently opaque and exposed to a large scale of risks. The unique characteristics of the industry pose a particular challenge for external rating agencies. According to Morgan (2002), the uncertainty over the bank ratings stems from certain assets, loans and trading assets in particular – the risks of which are hard to observe or easy to change. In his study Morgan also finds that rating agencies disagree more often on ratings concerning banks and insurance companies compared with other industries. He concludes that uncertainty over banks seems inevitable and inherent to the business to some extent. Thus, it seems that the quality of a bank credit rating may be lower compared with other companies.

The central role of financial institutions also seems to affect credit ratings. This is especially true in the case of large banks that could be defined as systemically important financial institutions. This definition simply implies the importance of a particular bank to the financial system – an institution whose failure would be disastrous to the greater economic system. As governments try to avoid these failures, there has been lots of discussion about the so-called "too big to fail" (TBTF) subsidy on large banks. This subsidy may be seen, for example, in banks' borrowing costs, abnormal stock returns or risk-taking behavior. (Fieberg et al. 2015)

The TBTF assumption asserts that certain companies are so big and so interconnected that their failure might trigger a financial crisis. Thus, these companies are likely to be supported by the government more than the less systemically significant companies in the case of failure. In general, these companies are large financial institutions. The TBTF subsidy refers to the fact that investors view investments with these banks as a safer investment than deposits with smaller banks. There is evidence that credit rating agencies would assign more favorable ratings to larger banks than those of small banks (e.g. Hau et al. 2013). Besides, stock prices of large banks seem to be affected less by rating changes than those of small banks (e.g. Fieberg et al. 2015).

Just as for any other company, a credit rating acquired by a bank is significant for the bank's operations. Bank ratings are an especially relevant determinant of the issuance cost of senior unsecured debt. This debt type remains the largest source of long-term funding for banks. Thus, credit ratings of senior unsecured debt remain an important assessment of bank creditworthiness. (Hau 2013, 295). One interesting feature concerning bank credit ratings is that due to the global bank capital requirements, a bank's credit rating is often dependent on the credit ratings of some of its own assets. This is due to the fact that external credit ratings of a bank's assets may be used in calculation of bank's regulatory capital.

2.2.2 Rating-based regulation in banking

According to Liapis (2011), the banking industry is a highly regulated business for the following reasons:

- The monetary nature of bank liabilities.
- The role of banks as payment intermediaries and providers of credit in an economy.
- The information deficiencies that characterize the business of banking, including those related to historical cost accounting, bank secrecy and confidentiality.

One of the most relevant parties concerning global bank capital requirements is the Basel Committee on Banking Supervision (BCBS). It was established by the central bank governors of the G-10 countries in 1974. According to BCBS it provides a forum for regular cooperation on banking supervisory matters and its aim is to enhance understanding of key supervisory issues and improve the quality of banking supervision worldwide. In 1988 BCBS published an international regulatory framework for banks focusing primarily on credit risk and risk-weighting of assets. This was called the 1988 Basel Accord, also known as Basel I, and it was enforced by law in the G-10 countries in 1992. (BIS 2015)

The update with the intent to supersede the Basel I accords was initially published in 2004 and was implemented in most major economies by 2008. Basel II includes a more comprehensive set of rules (see BCBS 2006) and also takes into account the operational risk and market risk while Basel I focused almost entirely on credit risk. The Basel II framework sets out a range of methods that banks could use in order to measure risk to calculate regulatory capital. These include the *standardized approach* and the *internal rating-based (IRB) approach*. The standardized approach allows banks to measure credit risk in a standardized manner based on external credit ratings. In this approach external credit ratings are used to compute risk weights for bank assets. For instance, a security with a better credit rating would apply for a smaller risk weight in calculation of bank's regulatory capital. In many countries the standardized manner is the only approach used to approve Basel II Accord.

The alternative way to calculate regulatory capital, the IRB approach, is based on a bank's use of its own internal models. To be able to use the IRP approach, banks need to meet certain minimum conditions, disclosure requirements and approval from their national supervisors. In practice, only large banks are usually able to meet these requirements and also capable to develop their own advanced internal risk control systems. Overall, it can be said that Basel II increased regulatory reliance on credit ratings. There even exists anecdotal evidence that the internal models themselves often tend to rely heavily on credit ratings for actual or methodological input (Hau et al. 2013, 296).

After the global financial crisis of 2007–2008, it was widely considered that the deficiencies in financial regulation had led to the crisis, and thus, the third installment of the Basel Accords was proposed in 2010. The implementation of Basel III is supposed to take place by the end of 2010s. Basel III focuses primarily on the risk of a bank run and does not supersede the Basel I and Basel II guidelines for the most part and will rather work alongside them. (BIS 2015)

During the writing of this thesis, the development of Basel III is still going on. The BCBS seeks to significantly improve the standardized approach for credit risk in a variety of ways. These include, for example, reducing the reliance on external credit ratings and strengthening the link between the standardized approach and IRB approach (BCBS 2010). However, Stefan Ingves, the chairman of BCBS, told in his speech at the IIF Annual Membership Meeting in October 2015 that he expects supervisors to follow the path of simplification rather than increasing complexity and that this approach is likely to include a reintroduction of role for external credit ratings (BIS Speeches). Anyway, the development of Basel III has been rather slow and the proposals introduced by BCBS (see e.g. BCBS 2015) have received substantial negative feedback, delaying the process further.

2.2.3 Characteristics of the European banking sector

The European financial system is dominated by the banking sector in most European countries. The capital market represents the second main component of the system. (Sargu 2012, 70). If compared with the United States, banks are more important for the credit mechanism in the economy on the European level. In the United States the capital market dominates and it is easier to get funding outside of the commercial banking system. (Cancian 2016).

The current European landscape has been plagued by financial, economic and national debt crisis. The banking system of the European Union represents one of the most affected economic sectors by the global financial and economic crisis. A vivid debate at the policy maker level has been required for the huge budgetary effort to ensure the stability of the system to avoid bank failures. Bank failures have been avoided through public capitalizations and nationalization or mergers and acquisitions. (Sargu 2012, 68)

Figure 2 demonstrates the development of STOXX Europe 600 Banks Price Index between 2000 and 2015. The STOXX Europe 600 Banks Index contains of 47 banks from 16 European countries and is one of the most comprehensive European banking indices.

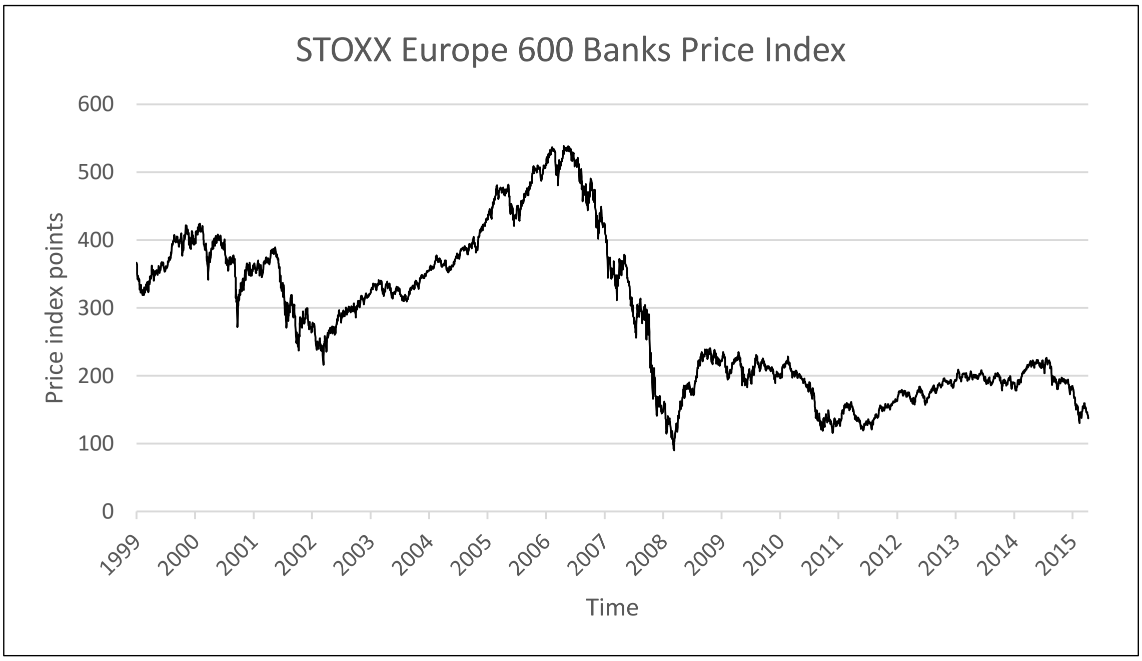

Figure 2 STOXX Europe 600 Banks Price Index (2000–2015)

The figure shows clearly how significantly the banks were affected by the financial crisis after a rapid growth period of nearly five years from 2002 till 2007. During the crisis the index dropped from over 500 index points to 100 points in less than two years. Also the volatile environment of the European sovereign and banking crisis starting just after the financial crisis can be seen in the figure.

Taking into account the number of countries in Europe, it is rather challenging to make strong generalizations of the characteristics of the banking sector. If considering only the member states of the European Union (28 states in 2016), those could be distinguished between two relatively homogeneous groups based on the main characteristics of the banking system. These are the EU-15 member countries and the new EU member states. The EU-15 countries have benefited from a long period of capitalism and free market and achieved a more sophisticated and developed financial system. The newer EU member states went through a process of transition to market economy prior to the European integration process and their financial system is thus less developed. (Sargu 2012)

There have been few studies examining whether the location of banks affects credit ratings. One recent study by van Loon and de Haan (2015), however, examines whether credit ratings of banks are related to their location inside and outside the euro area. They find evidence that banks located in euro area member countries receive a higher credit rating from Fitch than banks located outside the euro area in average. Their results also suggest that large banks in the euro area receive higher credit ratings if compared with their smaller competitors. Thus, evidence for the TBTF effect was found – large banks seem to have a competitive advantage over small banks.

2.3 Review of literature of credit rating announcements

2.3.1 Introduction to the research field

Poon and Chan (2008) divide the research of credit ratings into two strands. The first strand of literature that this thesis also belongs to, investigates whether initial ratings and rating changes instantaneously affect stock and bond prices. The direction of an effect as well as possible anticipation and delay are on the focus. The event study methodology is a common tool in this research area and it is often combined with single regression equation models. If credit ratings are useful, market participants should react to the new information in rating announcements. The second strand examines the determinants of credit ratings. This study field focuses on the predictability of credit rating changes based on the accounting information of the firms and capital market conditions. Deterministic models such as discriminant and cluster analysis are often used.

The aim of this study is to examine the effect of rating announcements on daily stock returns. Thus, the question of whether stock prices integrate all pertinent information to pricing is critical. In efficient markets, first introduced by Fama (1970), stock prices should reflect all information. There are three common versions of efficient-market hypothesis that implicate how markets work: weak, semi-strong and strong forms. If capital markets are weak-form efficient, future prices cannot be depending and determined on historical prices. A semi-strong form of market efficiency symbolizes that all publicly available information is mirrored in the prices. In strong-form efficiency all available information, both public and non-public, is reflected in security prices.

In the case of credit ratings, when a rating of a firm is announced, its effect should come in the firm's stock price. As the purpose of credit ratings is to measure credit risk, there are two alternative views on how rating agencies manage to produce this kind of information content (Gropp & Richards 2001, 7). One view is that rating agencies only have access to publicly available information and that there is a lag in processing it. If this is the case, ratings should not have an effect on market prices as prices already include all publicly available information (semi-strong market efficiency).

The other view, however, states that rating agencies are able to obtain private information in their rating process and thus produce new information affecting market prices. After all, rating agencies have discussions with the management of their target companies in their rating processes. Besides, many empirical studies provide evidence in favour of the second theory as significant price effects are observed on stock and bond markets especially after rating downgrades. In the light of these facts, the second view looks somewhat more potential. There are naturally many reasons which may affect stock and bond prices besides rating announcements. These could include determinants, such as, firm size, industry, market area, capital structure and the recent credit class, to mention some.

There has been a voluminous body of literature studying the effect of rating changes on stock and debt markets since the 1970s that this thesis is related to. Studies are usually conducted using data from S&P, Moody's and Fitch. The research has previously focused on corporate stocks and bonds. However, recently a substantial amount of research has also been done on credit default swaps and sovereign debt markets. Even though the information content of credit ratings is a widely investigated topic, prior literature proves a lack of consensus due to varied and mixed findings. This is not a surprise as the research field includes numerous studies with different data sets from different markets. Also, the used methodological design varies a lot among the studies. A challenging example of a methodological problem is called *data contamination*. A rating change is contaminated, if there are earnings announcements or other relevant news around the announcement, such as, other rating announcements. These *confounding events* are unwanted events that occur during the window of observation and have an impact on a company's stock price disturbing the true influence of a rating announcement.

2.3.2 Debt market reaction

There are several studies that examine the impact of credit ratings on bond prices. A considerable amount of first studies concerning rating change announcements were done on corporate bonds. For example, Hand et al. (1992) found evidence that daily bond returns are sensitive to bond rating changes. According to them, bond prices respond negatively to downgrades and positively to upgrades in the case of rating changes and credit watches. Thus, rating agencies seem to produce new information for the market. For instance, Wansley, Glascock and Clauretie (1992) confirm the significant negative effect of downgrades on bond returns using weekly data. However, in the case of upgrades this effect was not found.

There haven't been many studies related to corporate bonds recently. A number of first credit rating studies were conducted studying the effect on bond prices. Afterwards the research has been more focused on stock market response. For the last decade, there have been plenty of studies concerning credit default swaps (CDS). A CDS is a credit derivative contract between two parties where the buyer makes periodic payments to the seller in exchange for a commitment to a payoff if a third party defaults (Hull, Predescu & White 2004, 2790).

One of the first researches to study the relationship between CDS spreads and credit rating announcements were Hull et al. (2004). According to them, CDS spreads are an interesting alternative to bond prices in empirical research for a couple of reasons. They mention, for example, that CDS spreads are already credit spreads, so no adjustment is required. Bond yields, for their part, require an assumption about the appropriate benchmark risk-free rate before they can be converted into credit spreads. Overall, Hull et al. found that rating announcements for positive rating events are much less significant than the results for negative rating ones. Also, for example, Galil and Soffer (2011) report that the market reaction to downgrades is stronger than the reaction to upgrades. Study results of Daniels and Jensen (2004) reveal that the CDS market reacts faster and more significantly to changes in credit ratings if compared with the bond market.

Besides corporate bond and CDS research, there are studies examining the effect of sovereign credit rating announcements on sovereign bond yields spreads. Previous work in literature has focused on emerging and developing economies as debt crises have been more common in those environments. Nevertheless, after the 2007–2008 financial crisis, the research related to developed economies has rapidly increased. Likewise, in these studies rating downgrades are often more significant, while the reaction of spreads to positive rating events is more mitigated (see e.g. Afonso, Furceri & Gomes 2012).

2.3.3 Stock market reaction

A substantial amount of research has been completed to investigate the stock market reaction of credit rating announcements. As the data and methodological design used in these studies varies a lot, the results are often mixed and contradictory. This makes generalizing the results rather challenging although there seem to exist at least some trends in average.

The majority of the papers are conducted by using event studies that are often combined with single regression equation models. Event studies usually focus on the direction of stock price reaction and possible anticipation and delay. Besides rating announcements, some studies also take into consideration other information disclosures produced by agencies, such as, rating watches and outlooks. Regressions are used to investigate what characteristics of rating changes explain abnormal returns.

Most previous studies, especially the first ones, focused on the United States market (e.g. Weinstein 1977; Wakeman 1978; Pinches & Singleton 1978; Holthausen & Leftwich 1986). The early studies were often conducted using monthly data (e.g. Pinches & Singleton 1978). Since the adoption of daily data in 1980s, the amount of this research type increased (Corrado 2011, 213). According to Brown and Warner (1985), daily data generally presents few difficulties for event studies but can sometimes be advantageous. Since the 1990s, studies using European data increased (e.g. Barron, Clare & Thomas 1997; Linciano 2004; Li, Visaltanachoti & Kesayan 2004), and a bit later papers emerged focusing on other stock markets around the world (e.g. Matolcsy & Lianto 1995; Richards & Deddouche 1999; Elayan et al. 2003; Li, Shin & William 2006; Poon & Chan 2008; Habib, Nazir, Hashmi & Saeed 2015).

The trend in prior research seems to be that stock prices tend to react asymmetrically to announcements of upgrades and downgrades of a company’s credit rating. Rating downgrades are typically associated with significantly negative abnormal returns while the reaction among rating upgrades is often more limited or insignificant. Studies, such as, Holthausen and Leftwich (1986), Hand et al. (1992), Schweitzer, Szewczyk and Varma 1992, Goh and Ederington (1993), Matolcsy and Lianto (1995) and Fieberg et al. (2015) support this statement. However, few studies also find evidence of positive abnormal returns in the case of upgrades. It can be noticed that many of these studies use non-Unites States market data, such as, data from New Zealand (Elayan et al. 2003), Sweden (Li et al. 2004), Japan (Li et al. 2006) and data from emerging markets (Han, Shin, Reinhart & Moore 2009). There are also some particularly contradictory studies. For example, Richards and Deddouche (1999) studied emerging market banks and found significant negative market reaction in the case of rating upgrades. Abad-Romero and Roble-Fernandez (2006) got similar results from the Spanish stock market.

According to one view, the reasoning behind the trend of significant negative downgrades and insignificant upgrades comes from the idea that the expertise of rating agencies and their access to private information allows them to sometimes uncover information about target companies. As companies usually have a strong incentive to publish positive information about their prospects, market prices already reflect this favourable information in most cases. Thus in the case of rating upgrades, little or no impact in the stock prices is observed. Whereas in the case of downgrades, the impact is more significant as companies generally have no such incentive to publish unfavourable information that quickly. (Richards & Deddouche 1999, 5)

The evidence also suggests that the abnormal returns around announcements are small for downgrades and even smaller for upgrades, if even significant. The magnitude of the returns witnessed over the previous year or so is much greater. Hence, most of the information contained in rating changes is already reflected in stock prices prior to the actual change. For instance, Holthausen and Leftwich (1986) discovered cumulative abnormal returns in the 300 trading days before ratings changes to be 12-15 percent for upgrades and around -20 percent for downgrades. By contrast, during the announcement period the abnormal return for the downgrades stocks was just 1 percent. (Richards & Deddouche 1999, 5).

However, there seem to be many factors in which the strength and the direction of the effects observed are dependent on. These include, for instance, the rated firm’s former credit rating (e.g. Goh & Ederington 1999; Jorion & Zhang 2010), whether a rating agency is certified or not (Beaver, Shakespeare & Soliman 2006), whether a rating is solicited or not (Behr & Güttler 2008) and whether investors have free access to rating reports or not (Chan, Edwards & Walter 2009). There have also been comparison studies including firm specific factors, such as, the industry (Linciano 2004) and the size of the firm (Fieberg et al. 2015) as a data separator. Some papers focus on the impacts of crises – especially the recent financial crisis has received much attention. These studies investigate the effect of rating announcements, either during the crisis (e.g. Ghachem 2015), or using the crisis as a cut-off point to divide the data (e.g. Fieberg et al. 2015). It is also good to keep in mind that there are studies finding evidence that rating announcements do not have any impact at all. Especially the first papers (e.g. Weinstein 1977; Wakeman 1978) found no significant effect regarding rating announcements about stock returns. However, this may result from the use of weekly or monthly data (Habib et al. 2015).

There are also papers studying whether stock prices react differently to rating changes depending on the underlying reason. Goh and Ederington (1993) point out that while recent studies have found in average that the market reaction to downgrades is significantly negative, this should not be expected of all downgrades. They give two reasons for this. Firstly, some rating changes are anticipated by market participants as many of them follow news of an increase in the firm’s riskiness. Secondly, and more importantly, a rating downgrade is naturally bad news for bondholders but not necessarily for stockholders. If a downgrade results from an increase in leverage which will transfer wealth from bondholders to stockholders, this should have a positive effect on stock prices but a negative effect on bonds. In other words, if a downgrade is motivated by an increase in risk, rather than a deterioration of earnings, it could be associated with positive returns. The same theory could explain why upgrades do not have a significant stock price effect. If some upgrades are due to anticipated increases in earnings and others to anticipated declines in leverage, the average reaction may not be significant.

In their study Goh and Ederington (1993) received some modest support for their theory. They divided their uncontaminated sample into categories based on the reason of the rating change. The results show that the downgrades associated with deteriorating financial prospects convey new negative information to the capital market, but that the downgrades due to changes in firms' leverage do not. Their conclusion is that rating changes cannot be treated as homogeneously as the underlining reason matters. For instance, Gropp and Richards (2001) support this theory in their study of European banks.

2.3.4 Bank-related research

There appears not to exist that many event study-based papers concerning credit rating announcements which focus solely on the banking sector. There is no consensus on whether the banking sector gives any different results from other industries in this research area. Schweitzer et al. (1992) give two alternative hypothesis on whether rating actions affect differently for banks and for other corporates. The first hypothesis suggests that as banks are highly regulated entities, there may be more information available on the market compared with other corporates. Thus, the information content of rating announcements might be lower. According to the other view, rating actions matter more for banks. This is based on the idea that regulators might allow withholding of adverse information to preserve the stability of the banking system. For example, financially distressed banks would otherwise lose their access to the capital markets disrupting the financial stability. However, Schweitzer et al. did not find evidence in their study that the United States bank holding companies would be dampened relative to those for unregulated industrial firms.

There are also a few studies that have found evidence of a difference between financial and industrial firms. For example, in his study Linciano (2004) reports some modest empirical evidence from the Italian market that downgrades lead to stronger effect when involving industrial firms compared with banks. Abad-Romero and Roble-Fernandez (2006) on the other hand noticed that the main effects of downgrades do not depend on the company's industry, whereas in the case of upgrades there seems to be some differences. Yet, this topic clearly needs more investigation.

There are at least two United States studies showing that rating changes for bank debt affect their stock prices. The first one of these conducted by Schweitzer et al. (1992) focused on 32 different bank holding companies containing 95 rating changes obtained from S&P between 1977 and 1987. However, only 18 of the announcements were upgrades. The estimation period of [-200, -60] was used to estimate market model parameters. They also examined any firm-specific events for the [-1, 1] estimation window reported in the Wall Street Journal Index. An observation was classified as being contaminated if any events occurred. Besides, they segregated announcements by whether a change of a rating grade was across or within classes (see Table 2).

In overall, Schweitzer et al. (1992) report that downgrade announcements are associated with negative and statistically significant abnormal returns. This holds for grade changes within and across classes and for contaminated and uncontaminated ratings. An average abnormal return during the announcement window for downgrades was -1.5 percent and the previous 60 trading days represented abnormal returns of nearly -7 percent. Little evidence was found between contaminated and uncontaminated samples or whether the rating changes were across or within classes. The samples of 18 upgrades were associated with positive but only marginally significant abnormal returns of 1.1 percent for the announcement window. However, there was no significant stock price performance within the pre-announcement and post-announcement windows. Besides, Schweitzer et al. (1992) conducted a cross-sectional regression analysis of downgrades. The variables included the number of grades by which a rating is downgraded and whether a preceded credit watch or a rival rating from another agency matters. Only the variable concerning number of grades was significant but just barely.

The second United States study by Billett, Garfinkel and O'Neal (1998) examined the relationship between changes in bank credit risk and the use of insured deposit. Their initial samples gathered from Moody's consisted of rating announcements of 59 bank holding companies between 1990 and 1995. As a part of their study they confirm the negative announcement effect of downgrades. They reported a cumulative average abnormal return of -1.1 percent for the 3-day [-2, 0] event window. No data is provided on the pre-announcement abnormal returns. They argue that the share of insured deposits in total liabilities is the most significant variable in explaining abnormal returns. Thus, banks can shield themselves from the full costs of market discipline through increases in insured deposits.

Probably the first ones to study emerging markets were Richards and Deddouche (1999). They focused on banks in 15 countries, mostly located in Asia and South America, and their sample included 49 banks with 219 rating changes. They expected the data for emerging markets to yield larger abnormal returns following rating announcements compared with mature markets. That is, one would expect there to be less information provided to the public about banks in emerging markets, giving rating agencies an important role acquiring new information. They also assumed rating changes to have a greater effect on banks compared with other industries as banks' operating profits are highly dependent upon the cost of their funding.

Richards and Deddouche (1999) focused on rating changes that are not immediately proceeded by other changes, reducing the sample to 43 downgrades and 15 upgrades. They used weekly data between 1989 and 1998. The usefulness of weekly data is argued by non-trading biases and possible inefficiencies in emerging markets. Market adjusted returns were chosen instead of the market model. They used weeks -35 to -4 as the estimation window and weeks -3 to 2 as the event window.

The results of the study of Richards and Deddouche (1999) are rather surprising as they found negative abnormal returns immediately following upgrades and positive abnormal returns following downgrades. The reaction to upgrades was statistically quite strong but in the case of downgrades, the positive returns were not statistically significant. During the 35 weeks prior to rating upgrades, the cumulative abnormal return of -1 percent was shown. In the case of downgrades, the corresponding return was -13 percent. Richards and Deddouche also used a regression analysis to define whether a series of explanatory variables might have a possible influence on the magnitude of price reaction to rating changes. These variables include, for example, the number of rating grades and whether a rating is preceded by a credit watch. However, they did not get economically significant or meaningful results for their data set. Thus, they concluded that rating changes do not seem to convey valuable information for emerging equity markets in the way one would expect.

Habib et al. (2015) likewise found similar kinds of contradictory results from emerging markets. Their study examined the impact of credit rating announcements on stock returns of 22 Pakistani banks rated by the Pakistan Credit Rating Agency. Thus, this study is the only paper introduced in this subsection that does not include ratings from any Big Three companies. Habib et al. used daily stock returns from 2008 to 2014. Their results suggest that all banks individually generate insignificant average abnormal returns, thus rating announcements do not have an impact on stock returns. The study reveals a positive significant cumulative average abnormal return for rating downgrades but an insignificant cumulative average abnormal return for upgrades. The market also significantly reacts before rating downgrades so rating changes seem to be anticipated. As there are also significant results occurring after rating downgrades, it seems that the market is not efficient in managing and handling the rating downgrade information.

Probably the most relevant, although not very recent study, is the paper from Gropp and Richards (2001). They focused on rating changes for European banks using a sample of 32 banks with 163 rating changes during 1989 to 2000. The sample included banks from 15 European Union members. The sample size was rather small due the fact that there had not been that many rating changes, but also because the authors solely chose banks that had a major subordinated debt instrument outstanding as they used ratings of this debt type. In their study, Gropp and Richards analyzed rating watches together with rating changes.

Gropp and Richards (2001) used the estimation window of [-100, -6] and the event window of [-1, 1]. A standard market model was used for estimation window in which they used the respective stock market index of the country as the market indicator. To follow the previous literature to identify contaminated rating changes (e.g. Schweitzer et al. 1992), they scanned news stories in Bloomberg for the [-1, 1] window regarding the bank. If a text which revealed the reasoning for the rating change was found, they considered the event as contaminated. Furthermore, they also attempted to identify the reasons behind rating changes to follow Goh and Ederington's (1993) intuition that the reason may matter.

Besides, Gropp and Richards (1999) controlled for expected versus unexpected rating changes. To do this, they identified whether a rating change occurred after a rating watch in the same direction. For example, if a downgrade occurred after a negative watch signal, they pronounced the rating change as expected. One could anticipate expected rating changes to not necessarily be associated with a market reaction.

Overall, Gropp and Richards (2001) found that rating changes have statistically significant and economically substantial effects. The upgrades were associated with a positive abnormal return of 1.2 percent on the event day and 1.5 percent in the event window of [-1, 1]. For the downgrades, the abnormal return was -0.5 percent on the event day but insignificant in the [-1, 1] event window. Thus, Gropp and Richards did not find that prices react more strongly to downgrades. This is in contrast to some of the previous literature (e.g. Schweitzer et al. 1992).

However, Gropp and Richards (2001) noticed that in the case of uncontaminated downgrades, average abnormal returns still remained significant whereas in the case of uncontaminated upgrades, they did not. Thus, there appears to be information contained in the rating changes per se in the case of downgrades, when news stories and the release of information about the bank is somewhat excluded. Gropp and were also able to support Goh and Ederington's theory that the reason for the rating change may matter. If a downgrade is motivated by an increase in risk, rather than a deterioration of earnings, it is more likely associated with positive returns.

Gropp and Richards (2001) received evidence that unexpected ratings react more strongly to stock prices than expected ones. This is in accordance with their expectations. Besides, Gropp and Richard tested for pre-announcement and post-announcement effects but found little evidence of substantial drifts. The absence of pre-announcement drifts may provide that there is news in the rating announcements that was not yet in the public domain. Overall, the results of the study suggest that rating agencies may have a useful role in summarizing and obtaining non-public information on banks for stockholders. However, they noted that they faced very small data samples in their study which were likely to affect the test results.

One of the most relevant and recent studies examining stock price effects of bank credit rating changes was conducted by Fieberg et al. (2015). They studied the information content of about 3300 global bank rating changes in 154 countries between 2000 and 2012, making it probably the most comprehensive bank-related event study so far. Besides upgrades and downgrades, they also divided the sample into small and large banks to examine whether large banks benefit from the TBTF subsidy. This is interesting as large financial institutions often stand out in terms of their systemic relevance, complexity and opacity, compared with non-financial firms and small banks. Furthermore, Fieberg et al. studied the impact of rating changes before and after the Lehman bankruptcy in September 2008 to assess if differences in stock market reactions for small and big banks emerge. For example, Salvador et al. (2014) state that rating agencies have significantly tightened their bank rating policies since the financial crisis which may be seen in the results.

Fieberg et al. (2015) used the estimation period from 250 to 46 days prior to the event to estimate the parameters required by the Fama-French three-factor model which they used as a benchmark model to estimate expected returns. Their findings confirm the asymmetry in upgrades and downgrades proved by many prior studies. In their study, rating upgrades were not associated with significant abnormal returns, but downgrades had a significantly negative effect. This was the case for both small and large banks, although negative abnormal returns were substantially stronger for small banks. This is in line with the recent TBTF debate that potential negative effects on large banks' returns are dampened by an implicit TBTF bonus. Though, the main impact of rating downgrades for large banks was captured by negative abnormal returns well before the event and it was not significant for the [-1, 1] event window. On the contrary, banks in the small portfolio experienced cumulative average abnormal return of -11.3 percent for the [-1, 1] event window.

In addition, Fieberg et al. (2015) found that negative abnormal returns were significantly larger for small banks after Lehman, but almost all of the increase in negative abnormal returns occurred before the event. The similar pattern was seen in the case of large banks, but the negative effect was economically weaker. When observing the [-1, 1] event window, negative abnormal returns were significant and about the same size pre- and post-Lehman for small banks. For large banks, abnormal returns were not significantly different from zero in either periods. According to them, this may indicate that an implicit TBTF subsidy remains even after the financial crisis.

Table 3 summarizes the bank-related stock price studies introduced in this subsection. It includes both the technical brief and the study results in a compressed form:

Table 3 Summary of studies concerning bank stock prices

Technical aspect	Study results
Schweitzer et al. (1992)	
• USA • 32 banks, 95 rating changes • S&P • daily data 1977–1987 • market model • debt ratings	• upgrades: marginally significant positive returns • downgrades: significant negative returns The results hold even when confounding events are removed from the sample.
Billet et al. (1998)	
• USA • 59 banks, 233 rating changes • Moody's • daily data 1990–1995 • market model • debt ratings	• downgrades: significant negative returns The share of insured deposits in total liabilities is the most important variable in explaining abnormal returns.
Richards and Deddouche (1999)	
• 15 emerging countries • 49 banks, 219 rating changes • Big Three • weekly data 1989–1998 • market adjusted returns • long-term debt/financial strength	• upgrades: significant negative returns • downgrades: insignificant Bank stock prices do not respond to rating changes in the way that one would expect.
Gropp and Richards (2001)	
• 15 European countries • 32 banks, 163 rating changes • Big Three • daily data 1989–2000 • market model • subordinated bonds	• upgrades: significant positive returns • downgrades: significant negative returns (overall) Stock prices react differently to rating downgrades depending on the reason – positive reaction if motivated by an increase in risk.
Fieberg et al. (2015)	
• 154 countries • 3300 rating changes • Big Three • daily data 2000–2012 • Fama-French three-factor model • long-term issuer ratings	• upgrades: insignificant • downgrades: significant negative returns Downgrades affect small banks more negatively than large ones.
Habib et al. (2015)	
• Pakistan • 22 banks, 154 rating changes • Pakistan Credit Rating Agency • daily data 2008-2014 • CAPM • long-term issuer rating	upgrades: insignificant downgrades: positive significant results Bank stock prices do not respond to rating changes in the way that one would expect.

3 HYPOTHESES

3.1 Main hypotheses

The aim of this study is to examine the effect of credit rating changes on daily stock returns for a sample of European banks. Based on the efficient-market hypothesis, credit rating announcements should not affect stock prices as stock prices already reflect all available information (Fama 1970). Even though empirical results show that markets are not usually efficient, at least in the strong form, the efficient-market hypothesis is still an easy and suitable starting point to help to analyze deviations via market efficiency. Therefore, it is also used as a basis to form the research hypotheses of this study. According to the efficient-market hypothesis, the following hypothesis should hold true:

H1: Credit rating announcements have no significant effect on stock returns.

If *H1* holds true, it seems that stock prices already reflect all the information contained in the rating announcement. Rating agencies just process this information with a lag. In this case, there is no link between the rating announcement and stock prices. All price effects are observed early prior the event day.

However, a vast amount of studies report statistically significant abnormal returns associated with credit rating announcements. Besides the announcement time, the research focuses on the time periods prior and after the event. Whether rating agencies are able to obtain new information to the market, the price effect should occur after the announcement. In an efficient market, the reaction to news should be rapid and take place after the disclosure, unless information is leaked (Fama 1970). In the case of credit rating agencies, it is unlikely that a rating itself would leak to the market beforehand. However, the information included in a rating may be available to the public prior to the announcement, in which case the market "anticipates" the following rating action. Rating agencies also tend to give their own subjective signals about future ratings in the form of rating watches and outlooks.

Regardless, the only way to test whether rating announcements provide new information to the market, is to examine the abnormal returns around the event day. If rating announcements do provide real-time and pertinent new information (rejection of *H1*), the following hypothesis should hold true with regard to the efficient market hypothesis:

H2: Stock markets do not anticipate credit rating announcements, but react immediately afterwards.

The more anticipation there is, the less pertinent new information there is in the rating announcement. In the real world one could expect *H2* to hold true partially: some anticipation before the event and some reaction in stock prices afterwards, given that one can assume *H1* to be rejected. For example, Schweitzer et al. (1992) find anticipation of downgrades over the previous 60 trading days before the announcement, but they also discover significant abnormal returns during the announcement window. Fieberg et al. (2015) similarly report anticipation before the announcement and significant abnormal returns on the event window. Gropp and Richards (2001) for their part do not find evidence of abnormal returns over longer pre- or post-announcement periods. Thus, they conclude that rating announcements include news that was not already in the public domain.

Although there have been papers studying whether stock prices react differently to rating changes depending on the underlying reason (e.g. Goh and Ederington 1993), the majority of empirical research seems to concur that stock market effects in average are asymmetric for rating downgrades and upgrades (e.g. Hand et al. 1992; Matolcsy and Lianto 1995: Goh and Ederington 1993). More specifically, the average stock price reaction to rating downgrades is negative, and insignificant or marginal to rating upgrades. Anyhow, the majority of these studies focus on the United States markets. There are also studies reporting contradictory results especially from non-US markets, such as, from Sweden (Li et al. 2004), Japan (Li et al. 2006) and from emerging markets (Han et al. 2009). As there are not many studies concerning especially banks, it is not possible to state that the research trend would also hold with this industry. However, some bank studies, such as, Schweitzer et al. (1992) and Fieberg et al. (2015) provide evidence for the trend by finding negative abnormal returns for downgrades and insignificant or marginal returns for upgrades.

With regard to the study of Goh and Ederington (1993), even if the underlining reason triggering the rating event (an increase in risk versus a deterioration of earnings) matters, this may be unlikely with respect to banks. In their study Fieberg et al. (2015, 248) note that bank capital structure is typically heavily biased towards debt financing. This gives them a very limited leeway for changes in leverage which would benefit shareholders. On the other hand, Gropp and Richards (2001) do find evidence supporting Goh and Ederington in their study concerning European banks.

All in all, it is convenient to form the hypothesis concerning upgrades and downgrades according to the research trend. Thus the presumption follows the subsequent hypothesis that is divided into two sub-hypotheses:

H3: Rating announcements are asymmetric for downgrades and downgrades on stock prices around the announcement day.

H3a: Rating downgrades have a statistically significant (negative) effect.

H3b: Rating upgrades have no statistically significant (positive) effect.

Besides rating change announcements, agencies also produce rating watches to indicate the likely direction of a rating in the future. For instance, Holthausen and Leftwich (1986) and Hand et al. (1992) stress that it is important to also consider information contained in rating watches. They find evidence that significant abnormal returns are associated with rating watch announcements. Thus, the same hypotheses concerning rating changes are also tested with rating watches in this thesis. Typically, rating watches are less often studied compared to rating changes. Sometimes they are even used together with rating changes if there have not been enough observations (e.g. Gropp & Richards 2001). To the knowledge of the author, the information content of rating watch announcements is not comprehensively studied regarding banks. It is informing to see whether they produce any new information to the market and whether their impact on stock prices is similar to rating changes in the case of European banks.

3.2 Additional tests

Besides studying only stock price effects of rating announcements, some additional tests are conducted to investigate other potential determinants of the abnormal returns surrounding downgrades and upgrades. The determinants to be studied are as follows:

- market value of the corresponding firm
- number of notches by which each rating is changed
- whether a rating is preceded by a credit watch
- whether a rating was preceded by a rival rating in the near past
- whether a rating changes between or within classes
- whether a rating changes within the investment or below investment class
- whether a rating occurred before or after the beginning of the financial crisis.

The market value is used to control the firm size effect. According to, for example Fieberg et al. (2015), smaller and larger banks are affected differently. Large financial institutions often stand out in terms of their systemic relevance, complexity and opacity. When controlling firm size effect, large banks may be benefiting from the implicit public TBTF guarantee. Thus, their price effect is probably smaller if compared with smaller banks. Fieberg et al. had a large sample of 3200 observations which allowed them to compare the largest and the smallest banks, excluding medium sized banks entirely. However, in this thesis the sample size is much smaller and the procedure in question is not possible. Nevertheless, it is interesting to see whether the market value as a control variable still effects abnormal returns when the sample includes banks of all sizes.

One could expect the price effect to be stronger if a rating is changed many notches at once instead of just one. Rating agencies tend to give rating changes gradually, one notch at a time, so changing many notches at once is rather exceptional. For instance, Schweitzer et al. (1992) and Richards and Deddouche (1999) have studied this topic in their papers examining banks. Regression analysis were conducted in both studies but little evidence for a stronger price reaction in the case of many notches was found.

Besides studying whether rating watches contain market relevant information, those are also exploited in another way in this study. They are used as a means of distinguishing between expected and unexpected rating changes. For instance, Hand et al. (1992) and Gropp and Richards (2001) argue and got support in their studies that a rating change that is preceded by a ratings watch in the same direction, should not necessarily be associated with a reaction in market prices. In other words, unexpected rating changes should cause a stronger price reaction if compared to expected ones.

One could expect the abnormal returns to be stronger if there are no rival rating changes close before a rating announcement. If the rating announcement is preceded by another rating announcement of a rival agency in the same direction, it may be possible that investors react to the first rating more intensively. Thus, the second rating by another agency is most likely to be less unpredictable. There have been various ways how this variable is coded. For instance, Schweitzer et al. (1992) investigated whether it matters if the Wall Street Journal Index for the year of the downgrade reports that a bank's rating change by S&P is preceded by a similar rating change by Moody's. However, the regression coefficients for this variable are not statistically significant.

The next two determinants are related to rating classes that were described in Table 2. A rating change occurs between classes, for example, if a rating with a grade of A+ is upgraded into AA-. A rating change between classes is naturally expected to have a stronger impact compared with changes within classes. However, the threshold between investment and non-investment grade ratings is more significant as it is widely used in regulation unlike the other sub-categories. To the knowledge of the author, these topics are not investigated especially concerning banks. Anyway, there is evidence that announcements of downgrades across rating classes are associated with more negative abnormal returns compared to changes within classes (Holthausen & Leftwich 1986). Moreover, for downgrades, the average returns are demonstrated to be stronger for below investment grade stocks than for investment grade stocks (Hand et al. 1992).

Fieberg et al. (2015) studied the information content of rating changes before and after the bankruptcy of Lehman Brothers with their global sample of 3200 bank observations.

They were interested to see whether rating agencies had tightened their bank rating policies since the outbreak of the subprime crisis leading to potential changes in the information content of bank rating signals. The Lehman incident was chosen as a cut-off point as they regarded it as the crucial turning point of the financial crisis. Fieberg et al. find an increase in negative abnormal returns post-Lehman, but only for small banks. In this thesis the sample size is too small to sort out small and large banks. However, it is still of interest to test whether the outbreak of the financial crisis affects banks of all sizes specifically in the European context.

4 DATA AND METHODS

4.1 Methodological choices and dilemmas

4.1.1 Event study

An event study can be demonstrated graphically with a timeline divided into the estimation window, event window and the post announcement period. The estimation window is a period over which parameters are estimated for each company. The event window is the period over which the event occurs, K implying the actual event day (day 0) that is the rating announcement day in this study. The post announcement period simply implies the duration of the period that may be used if one wants to follow the reaction in the longer run after the announcement, often up to one calendar year. It is less often used except for studies in which it takes more time for an event to be seen in the market (Lim 2011, 161). Examples of these kinds of studies could include mergers and acquisitions, buyouts and IPOs. Post announcement effects are not examined in this study as rating changes have been shown to have a very short-term effects on stock prices.

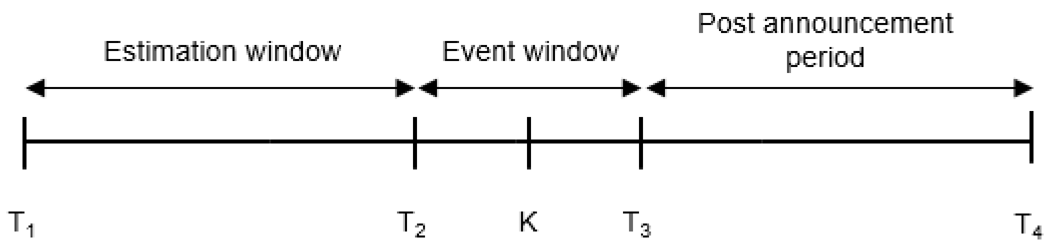

Figure 3 Event study timeline

There are no clear rules on how to choose the length for each window. The lengths of the event window and the post announcement period depend mainly on what the researcher aims to study. In some studies the event window stands for just the actual time the event takes place. In these cases the choice of the window is more crucial as the event day is not always easy to define and the main reaction may spread throughout on a couple of days. Thus, it may be occasionally better to use the window of [-1, 1] or [0, 1] instead of just the expected announcement day. Other studies, however, define the event window to include more days instead of just the actual event day or days. Then, the event window should be large enough to show any possible changes to returns due to the event (Lim 2011, 161). The typical length of the event window in many studies is two calendar weeks or 10 trading days before and after the announcement. This approach has also been chosen for this study. Thus, the length of the event window [T_2, T_3] is altogether 21 (10+1+10) trading days even though the actual event day still remains as day 0. However, a couple of cumulative average abnormal returns starting from day -20 are reported besides the main results just to make sure that there was not any anticipation prior to the event.

The choice of the estimation window is more problematic as it is used to estimate the parameters for each company and these parameters vary over time. A longer estimation period may result in an improved prediction model but may cause model parameter instability (Peterson 1989, 38). In environments where the market may be disruptive and beta may have changed over the calendar year, a shorter estimation period may be preferred (Lim 2011, 161). Typically the estimation period in event studies using daily data varies from 60 days to 300 days. On occasion, the post announcement data is included with the estimation window data to estimate the normal return model. This could be done to increase the robustness of the normal market return measure to gradual changes in its parameters (MacKinlay 1997, 20). In any case, it is important that the estimation window and the event window would not overlap. In this situation, both the normal returns and the abnormal returns would capture the impact of an event causing methodological problems (MacKinlay 1997, 20). In this paper the estimation period is chosen to start 260 trading days and ending 11 trading days prior to the rating announcement date of each bank. This gives 250 trading days in total which is a commonly used length in event studies.

The realized stock market returns are nowadays easily obtained from databases but there are various methods to estimate the expected returns for event studies. The most common way to calculate the expected return is to use the market model, which is also utilized in this study due to its simplicity and restrictedness (Corrado 2011, 225).

Besides the market model, there are also other ways to estimate the expected return, such as the capital asset pricing model (CAPM), the constant mean return model and various multifactor models, for example, the arbitrage pricing theory (APT). The choice of the model has been a widely discussed topic in event study literature. The choice depends pretty much on the situation but the market model has been shown to be good enough in most cases and the differences between models are not necessarily very large (see e.g. Brown and Warner 1985). The use of the CAPM was common in event studies in the 1970s until it was shown in many studies to be outperformed by the market model (Kothari and Warner 2004, 25; see e.g. Cable and Holland 1999).

The economic models such as the CAPM and APT impose additional restrictions on the statistical models. For instance, in the case of the CAPM, the intercept is set to equal the risk free rate making the variance of the error term larger than in the market model. With the APT, there usually exits one factor behaving like the market factor. Additional factors, however, provide very little explanatory power. (MacKinlay 1997, 19). Nevertheless, as many studies claim that the market model performs well enough in general, and it is rather straightforward to use, it has been chosen to estimate the expected returns in this paper. Besides, many studies report that the difference in results when using various models is not that notable in short-term event studies. In long-horizon studies the choice of the model is more crucial (Kothari and Warner 2004, 8).

The market model assumes a linear relation between the market and security returns. The equation of the market model to estimate the expected return is

$$E(R_{it}) = \alpha_i + \beta_i R_{mt} + \epsilon_{it} \tag{1}$$

where R_{mt} denotes the return of the market portfolio m on day t, ϵ_{it} is a random term and parameters alpha (α_i) and beta (β_i) are the ordinary least squares estimates obtained from an estimation period. Frequently, stock market indexes are used as market proxies depending on the research area. In this study STOXX Europe 600 Banks Index is used to measure the market portfolio R_{mt}. Another option would be to use country-specific market indexes (see e.g. Gropp and Richards 2001). However, as the banks to be studied are taken from STOXX Europe 600 Banks Index, it is reasonable to use this index as a market proxy.

The next step is to calculate the difference between the realized return and the expected return. The abnormal return (AR) for security i on day t is the deviation of security i's realized return R_{it} from an expected return $E(R_{it})$ generated by the market model.

$$AR_{it} = R_{it} - E(R_{it}) \tag{2}$$

After obtaining the abnormal returns, average abnormal returns (AARs) are calculated on day t. This is done by adding ARs up for each day and then dividing by the number of ARs (Formula 3). Logarithmic returns are used in this study due to their statistical properties for additive processes. The abnormal returns are trimmed at the data panel when calculating each AAR to avoid the averages being distorted by outliers. In the case of rating changes, the sample sizes are more than 100 observations and trimming is done at the 2nd and 98th percentiles. When the data was trimmed at the 5th and 95th percentiles, the results remained practically very close to the results of the 2nd and 98th percentiles. However, in smaller samples containing less than 100 observations, outliers have more effect on the results. In the case of rating watches, the data is trimmed at the 5th and 95th percentiles. All in all, trimming significantly affects the results so the untrimmed samples would be distorted by extreme outliers.

$$\mathrm{AAR}_t = \frac{1}{N}\sum_{i=1}^{N} AR_{it} \tag{3}$$

Besides testing only single-day returns, there is a need to scrutinize the behavior of returns within specified time intervals. Thus, the returns must be aggregated over time. Cumulative average abnormal returns (CAARs) are simply calculated over various event interval periods by adding AARs up.

$$CAAR_{t_1,t_2} = \sum_{i=t_1}^{i=t_2} AAR_t \tag{4}$$

To test the statistical significance of whether the abnormal returns differ from zero, test statistics need to be applied. There exist numerous test statistics in the event study literature for this purpose. These significant tests can be grouped in parametric and non-parametric tests. Unlike non-parametric tests, parametric tests assume that abnormal returns of an individual firm are normally distributed. Parametric tests are more traditional and are based on the standard t-test that follows a Student's t-distribution. (Dutta, 2014). When the sample size (N) increases, the test statistic approaches a normal distribution.

The test statistics used in this study adapts the framework of Lim (2011, 165–169). The following formulas are applied to calculate the parametric test statistics for average abnormal returns (5) and for cumulative average abnormal returns (6):

$$stat(AAR_t) = \frac{AAR_t}{\sqrt{var(AAR_t)}} \sim N(0,1) \tag{5}$$

$$stat\left(CAAR_{t1,t2}\right) = \frac{CAAR_{t_1,t_2}}{\sqrt{var(CAAR_{t_1,t_2})}} \sim N(0,1) \tag{6}$$

The variances that are used in the denominators of the formulas are calculated as follows:

$$var(AAR_t) = \frac{1}{N^2}\sum_{i=1}^{N} \sigma_i^2 \tag{7}$$

$$var(CAAR_{t_1,t_2}) = \frac{1}{N^2}\sum_{i=1}^{N} (t_2 - t_1 + 1)\,\sigma_i^2 \tag{8}$$

In some papers, the variance (σ_i^2) is defined as the variance of the estimation period of abnormal returns. Kothari and Warner (2004), however, state that this approach of using only historical time-series variability may understate the true variability of the abnormal performance of the event window. This, because the event period may be associated with increased uncertainty. Thus, using historical data may overstate the statistical significance of the event window. To overcome this problem, they suggest also using the event window variance. This approach is applied in this study and the variance for each security includes both the estimation window and the event window.

4.1.2 Cross-sectional regression analysis

To investigate other potential determinants of the abnormal returns surrounding downgrades and upgrades, a linear regression model is used. The aim of the regression analysis is to give an idea of potential determinants for further examining and to see whether the results differ from past research papers. The variables in the regression represent commonly used variables in this research field, even though, there have not been many papers which study the banking industry in particular. There are not any guidelines so show what variables to include in the regression nor how these variables should be coded. Thus, the regression is unique in this matter. The regression follows the subsequent formula:

$$\begin{aligned} AR_i = \beta_0 + \beta_1 \log(\mathrm{MV})_i + \beta_2 NOTCH_i + \beta_3 WATCH_i + \beta_4 CLEAN\_30_i \\ +\beta_5 CLASS_i + \beta_6 JUNK_i \ + \ \beta_7 LEHMAN_i \end{aligned} \tag{9}$$

where

AR_i	=	abnormal return for each individual bank on the announcement day;
$\log(MV)_i$	=	control variable calculated as the natural logarithm representing the market value of a company;
$NOTCH_i$	=	number of notches by which each rating is changed;
$WATCH_i$	=	dummy variable set equal to one if the rating was preceded by a credit watch and zero otherwise;
$CLEAN\_30_i$	=	dummy variable set equal to one if the rating is preceded by a rival rating occurring within 30 days prior the announcement;
$CLASS_i$	=	dummy variable set equal to one if the rating changes between classes and zero if the rating changes within a class;
$JUNK_i$	=	dummy variable set equal to one if the rating changes within a non-investment class and zero otherwise. In case that the rating cross the threshold between investment and non-investment classes, the new class determines the class the rating belongs to;
$LEHMAN_i$	=	dummy variable set equal to one if the rating occurred after the bankruptcy of Lehman (September 2008) and zero if it occurred before.

Thus, the ordinary least squares model includes seven variables in which five represent dichotomous variables. The market values for the first dependent variable are obtained from Datastream and represent the bank's market value at the end of the previous month before the event. The time period for the dummy variable concerning rival ratings is chosen to be 30 days in this study. This time period could also be much longer but if the variable has any reducing effect on stock prices, it should be more significant just before the rating.

4.1.3 Problems and limitations

Even though event studies are a common tool in finance, there are some problems and limitations that one must be aware of. These dilemmas cannot be solved as such, instead they can only be dealt with. There are problems, for example, related to the choice of the event date, calculating returns, statistical testing, as well as to the methodology in general.

As noted above, the choice of the event day is one of the most critical choices of the event study. In many cases the choice is rather challenging and one has to use uncertain event dates. For instance, in studies related to dividends or mergers and acquisitions one has to choose between the announcement day and the actual day the action takes place. In some cases there could be even more days to choose between or the information may be available gradually. Different trading times across the world may also cause problems. Besides, the information is often gathered from secondary sources, such as, from magazines or other media releases which means it comes with a delay. The more days one has to include in the event window because of uncertainty of the actual date, the lower the power is of the event study methodology (Brown and Warner 1980, 224–227). On the other hand, in the case of uncertain event dates, it is still sometimes possible to use accumulated abnormal returns over a bit of a longer period in order to detect events even though the power of the test statistic is lower.

The longer the event window, the more external confounding events it is likely to include. Confounding events are unwanted events that occur during a window of observation and have an impact on a company's stock. Examples of these kinds of events include negative or positive news to an individual firm and macroeconomic events. Especially financial companies are vulnerable to macroeconomic events, such as an increase in interest rates, and tend to all react to the news in the same way. Thus, focusing only on one industry may cause problems. Confounding events may be random or systematic. If they are random, and could then be considered noise, they make finding a true pattern less likely. However, if they are not random, they make finding a false pattern more likely. Nevertheless, the existence of confounding events is a problem when conducting an event study and there are different ways how researches control for these problems. The most common methods include simply eliminating contaminated firms from a sample. The biggest challenge, however, is to find the true events that have an impact on companies. (Kwok, Meznar & Nigh 1998, 717–718).

Taking into account this study, the choice of the event date is rather straightforward. The event date is simply the day the rating announcement is given by an agency. Of course it is possible that the data obtained from Bloomberg is not fully accurate or that some errors have been made while processing data. The event window is chosen to be larger than just the event day to see whether there exists any anticipation and whether the reaction occurs immediately or with a delay. Besides the AAR for each day in the event window, several CAARs are also reported. These include the windows of [-20, 10], [-20, -6], [-5, -1], [0, 1], [2, 5] and [6, 10]. The window of [0, 1] is reported just in case that the event date is not as accurate as expected.

The existence of confounding events is a bigger problem in this study. These are not taken into consideration in this thesis because of limited access to databases where this kind of news data is available and because of the sample size. Eliminating firms could create methodological problems associated with the sample size which is this small. Thus, in this study confounding events are not controlled except for rival rating announcement taking place before an event. Anyway, macroeconomic events and the fact that this study concentrates only on one industry are not likely to cause bigger problems. This, because firm-specific announcements occurred at all times of a year for a long time period. This reduces the impact of cyclical market trends. Thus, confounding events disrupting this study are most likely just random noise. In studies where the same event dates are used for several companies, macroeconomic events matter more.

Just as using test statistics in general, there are two types of errors of inference in event studies. A type I error occurs when the null hypothesis is falsely rejected. A type II error occurs if the null hypothesis is falsely accepted. The test statistic used in this study grounds on the traditional parametric t-test and is used to test whether the abnormal returns significantly differ from zero. Kothari and Warner (2004, 15–16) stress that economic interpretation of an event study test is not straightforward because all tests are joint tests. Thus, event study tests are well-specified only to the extent that the assumptions underlying their estimation are correct. Besides testing whether abnormal returns are zero, event study tests always test whether the assumed model of expected returns, such as a market model, is correct. Also, the assumptions concerning the statistical properties of the abnormal return measures must be correct. For instance, a standard t-test assumes that the mean abnormal performance for the cross-section of securities is normally distributed. However, for example, Brown and Warner (1980) state in their well-known study that the t-test works rather well in general, at least if compared with non-parametric tests such as a sign test or a Wilcoxon signed rank test.

As noted above, the market model has been shown to be good enough in most cases. Nonetheless, it still has its own drawbacks. For instance, Wells (2004) states that the market model depends on estimates on stocks' betas which measure the future's variability. In many studies, including this, the beta is estimated over the time period before the event and then used to compute returns around the event. Though, empirical tests suggest that the beta is not constant trough the time – the past is not a predictor of the future. Besides, a particular event could change the relationship between stock and market returns. Also macroeconomic variables such as interest rates and business cycles may cause alterations in betas.

The reliability of event studies may also suffer in the case of infrequently traded securities. In these cases stock prices may be biased if no trades have been made after an event. Infrequently trading may cause statistical tests to be poorly specified (Cowan & Sergeant 1996). Also, stock closing prices may be quoted in different times that may cause problems. Nevertheless, the companies used in this study are highly traded and located in nearby time-zones so these issues should not be of disturbance.

From a statistical point of view, one of the major concerns of this study is the rather small sample size. The full sample includes 321 rating announcement that consists of 200 rating downgrades and 121 upgrades. There are no clear rules of what an ideal sample size is, but for statistical generalization a large sample size is always more preferred. When a subsample contains more than 100 observations the sample size should be adequate. However, when dividing samples into even smaller subsets, let's say, samples containing less than 30 observations, the results should be taken with a grain of salt. In any case, sample sizes of even less than 20 observations are still often reported in papers in this research field. This is especially true with rating upgrades as the amount of those in many markets are rather scarce.

One problem related to credit rating changes, stems from the fact that rating agencies renew their rating methodologies every once in a while. This leads to situations when a lot of ratings are reviewed at the same time. For example, in March 2015 Moody's published its new global bank rating methodology which affected hundreds of banking entities (Moody's 2015). From the research point of view, these rating changes are problematic as they distort study results. Even though the reasons for these rating actions are merely methodological, the new ratings may still have an effect on the stock prices of individual banks. Besides, it is not always easy to find out whether a rating change of an individual bank stems from changes in the rating methodology or changes in its creditworthiness. In this thesis the study period is rather long and it would be challenging to become familiar and deal with every rating announcement concerning changes in agencies' methodological inputs. Thus, these ratings are not removed in this study and are likely to cause bias to the study results.

4.2 Data

The data set in this study is gathered from Bloomberg and it includes banks' solicited long-term issuer ratings from S&P, Moody's and Fitch. In this case, there are no differences between local and foreign currency ratings. The full sample consists of 41 banks from 14 European nations including altogether 321 credit rating changes from January 2002 to December 2015.

The sample of banks is chosen to contain the components included in the STOXX Europe 600 Banks Index. The STOXX Europe 600 Banks Index is derived from the STOXX Europe 600 Index that represents 600 large, mid and small capitalization companies across 18 countries of the European region. The STOXX Europe 600 Banks Index consists of 47 banks from 16 countries included in the STOXX Europe 600 Index. After restricting the data only to issuer ratings and filtering off companies' first rating announcements, the final sample includes the aforementioned 41 bank and 321 observations. All banks in the sample are publicly traded holding companies. The full list of the banks used in this study including their country and rating agency information can be found in Appendix 1. It must be emphasized that the data is restricted to only issuer ratings. This, because issuer ratings best reflect the issuers' overall credit quality, and thus, should cause a major stock price reaction.

The STOXX 600 Europe Banks Index is also used as a benchmark model to estimate the expected returns. Daily euro returns for the index and individual banks are provided by Thomson Reuters Financial Datastream. As the total return index for STOXX Europe 600 Banks was not allowed until 2001 and the estimation period to estimate market model parameters requires 260 preceding trading days, January 2002 is chosen as the starting point of this study. The sample includes only rating events for which there exist time-series of stock returns for the estimation period.

Table 3 demonstrates how the chosen banks and their ratings are distributed across nations. The sample is clearly biased towards certain countries. Among 14 nations, more than 60 percent of the total number of ratings are included in five countries (France, Greece, Italy, Spain and United Kingdom). Every state excluding the Czech Republic represent EU-15 countries. The Czech Republic, Denmark, Sweden and the United Kingdom are the only countries that do not belong to the monetary union.

Table 4 Distribution of ratings across nations

Country	Number of banks	Number of ratings	% of total
Austria	2	6	1.9 %
Belgium	1	10	3.1 %
Czech Republic	1	14	4.4 %
Denmark	2	12	3.7 %
France	4	34	10.6 %
Germany	2	23	7.2 %
Greece	2	32	10.0 %
Ireland	1	16	5.0 %
Italy	8	58	18.1 %
Netherlands	1	7	2.2 %
Portugal	1	8	2.5 %
Spain	7	45	14.0 %
Sweden	4	25	7.8 %
United Kingdom	5	31	9.7 %
Grand Total	**41**	**321**	**100.0 %**

Notice may also be given to how the banks are distributed over time. The time period contains 14 years for the time period of 2002–2015. Thus, economic fluctuations are included in the sample. It is especially informing to scrutinize how the 2007–2008 financial crisis and economic uncertainty over the following years affect the number of ratings. Table 4 shows how rating downgrades and upgrades are distributed over the years.

Table 5 Distribution of ratings across time

Year	Downgrades	% of total	Upgrades	% of total	Total	% of total
2002	13	6.5 %	4	3.3 %	17	5.3 %
2003	5	2.5 %	7	5.8 %	12	3.7 %
2004	0	0.0 %	12	9.9 %	12	3.7 %
2005	2	1.0 %	8	6.6 %	10	3.1 %
2006	0	0.0 %	22	18.2 %	22	6.9 %
2007	3	1.5 %	21	17.4 %	24	7.5 %
2008	24	12.0 %	2	1.7 %	26	8.1 %
2009	51	25.5 %	1	0.8 %	52	16.2 %
2010	23	11.5 %	1	0.8 %	24	7.5 %
2011	3	1.5 %	0	0.0 %	3	0.9 %
2012	0	0.0 %	0	0.0 %	0	0.0 %
2013	33	16.5 %	3	2.5 %	36	11.2 %
2014	11	5.5 %	21	17.4 %	32	10.0 %
2015	32	16.0 %	19	15.7 %	51	15.9 %
Grand Total	**200**	**100.0 %**	**121**	**100.0 %**	**321**	**100.0 %**

Figure 3 illustrates the same information more clearly in the form of a graph. Before the crisis there were not many rating changes and most of the rating announcement were positive. During the years of 2008-2010 a peak in the number of rating downgrades occurred. Almost no rating upgrades were given that time. After the crisis there were a couple of calm years until the amount of rating changes increased again in 2013. In 2014 and 2015 the amount of rating upgrades increased again after six years of keeping a low profile.

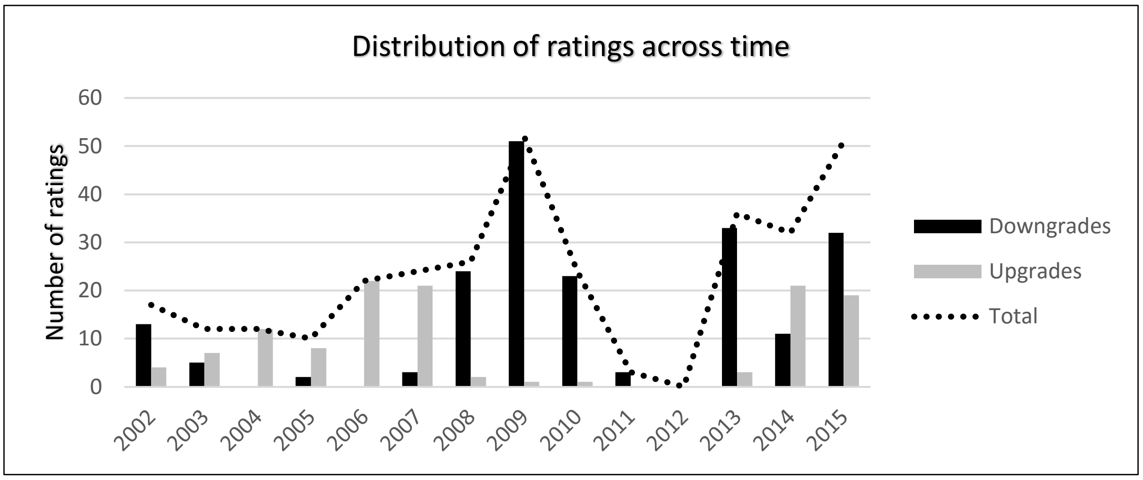

Figure 4 Distribution of ratings across time

Although rating announcements between agencies are seen as equals in this study, there might be differences in how they affect stock prices in reality. Table 5 describes how the ratings are divided between agencies. S&P represents about half of the total amount of events in the sample. Approximately one third of the events belong to Fitch and one fourth to Moody’s. It would be interesting to see whether an increase in sample size would compensate the distribution between agencies.

Table 6 Distribution of ratings between agencies

	Number of ratings	% of total
S&P	**152**	**47.4 %**
Downgrade	93	29.0 %
Upgrade	59	18.4 %
Moody’s	**69**	**21.5 %**
Downgrade	32	10.0 %
Upgrade	37	11.5 %
Fitch	**100**	**31.2 %**
Downgrade	75	23.4 %
Upgrade	25	7.8 %
Grand total	**321**	**100.0 %**

Various pieces of information about ratings in the sample are gathered in Table 6. This information is used to conduct a linear regression model to investigate potential determinants of the abnormal returns around rating downgrades and upgrades.

Rating agencies tend to give rating changes gradually, one notch at a time. This can be seen in the first category of Table 6. Only 15 percent of all rating changes move more than one notch at a time. Most of the rating changes included in this 15 percent are double-notch changes. Only 14 observations represent cases in which the rating is changed more than two notches at once.

The second category shows whether a rating change is predicted by the agency or not. If the rating change is preceded with a rating watch, it is named as an expected rating change in the table. Most rating changes in the sample were unexpected and only one fourth of the total amount of ratings was put on watch beforehand.

The third category shows whether there has been a rating change of any kind by a rival agency within 30 days prior the announcement. If all rating agencies functioned efficiently, they would change companies' ratings somewhat at the same time when needed. In this case, the first agency to update the rating would probably cause a more significant stock price reaction than the following ones. When removing rival announcements within a month prior the change, rating change clusters are avoided and the sample will probably be a bit "cleaner". In this study 44 rating changes are filtered off to test whether it has an impact on the results when conducting a linear regression.

The fourth category makes a distinction whether a rating change occurs between or within rating classes. As one rating class mainly includes three ratings, most of the ratings occur within classes. The fifth category, on the other hand, shows whether a rating change occurs within the investment grade or non-investment grade class. If a rating change occurs across these groups, the new rating grade determines which of the groups it belongs to in the table. As it can be seen in the table, less than one fifth of all ratings belong to the non-investment grade group.

The last category divides the ratings occurring before and after the bankruptcy of Lehman Brothers which can be regarded as the crucial turning point of the global financial crisis. The date of this cut-off point is 15 September 2008. As it was already seen in Figure 3, most rating changes have occurred after this turning point. However, most of the rating upgrades occurred before Lehman when the market was growing for many years.

Table 7 Descriptive information about the sample

	Downgrades	Upgrades	Grand total	% of total
1 notch at a time	163	110	273	85.0 %
> 1 notch at a time	37	11	48	15.0 %
Not-expected rating change	139	102	241	75.1 %
Expected rating change	61	19	80	24.9 %
No pre-rating from other agency	166	111	277	86.3 %
Pre-rating from other agency	34	10	44	13.7 %
Between classes	79	44	123	38.3 %
Within classes	121	77	198	61.7 %
Investment grade	160	103	263	81.9 %
Non-investment grade	40	18	58	18.1 %
Before Lehman	35	76	111	35.6 %
After Lehman	165	45	210	65.4 %

Besides credit rating changes, the impact of rating watches on stock prices is additionally studied in this paper. In total, the time period of 2002–2015 includes 80 negative but only 22 positive rating watch announcements. The sample does not include observations in which a company is put on watch concurrently with a rating change. Figure 5 illustrates the distribution of rating watches across time. Just as was the case with rating changes, most negative rating watches occur during and after the financial crisis. It can be pointed out that there have been only 22 positive rating watches and almost half of them occurred in 2015.

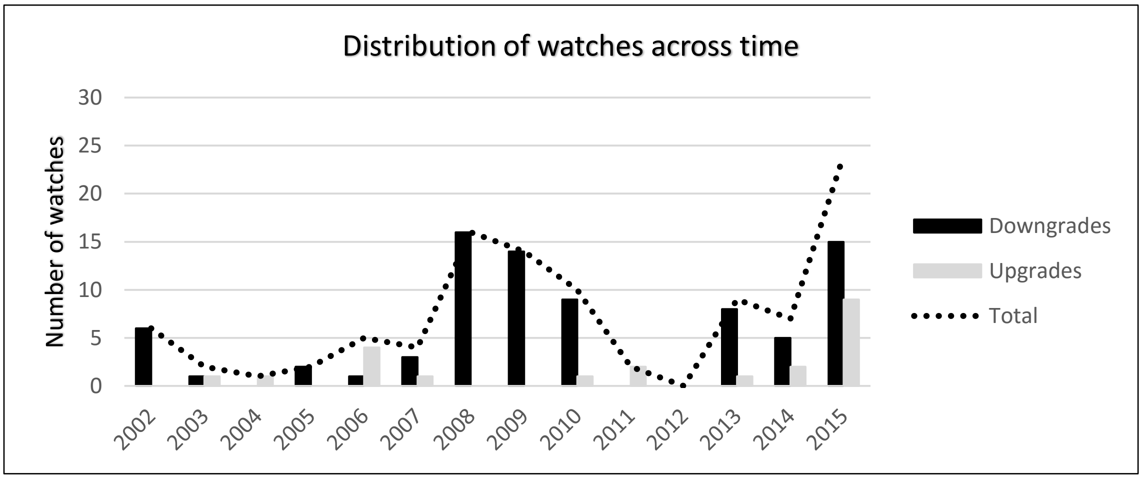

Figure 5 Distribution of rating watches across time

The impact of positive watches is reported in this paper but taking into consideration that it only includes 22 observations, the results should be taken as preliminary. Nevertheless, the sample sizes of either positive or negative watches are too small to further divide them into smaller subsets for additional tests.

5 EMPIRICAL RESULTS

5.1 Credit rating changes

5.1.1 Full sample and overview

Results of the full sample are displayed in Table 7. With respect to the first research hypothesis, the abnormal return of -0.67 percent on the event day already provides evidence that rating announcements would have a statistically significant effect on stock returns. Thus, *H1* can be rejected in the case of rating changes. As 62 percent of the rating changes in the sample are negative and stock prices are also expected to react more significantly to negative ratings, the direction of the return on the event day is not a surprise.

Table 8 Summary of abnormal returns of rating changes (full sample)

Day [t]	AAR
-10	0.18 %
-9	-0.02 %
-8	0.01 %
-7	-0.04 %
-6	0.14 %
-5	-0.17 %
-4	-0.13 %
-3	-0.19 %
-2	0.12 %
-1	-0.11 %
0	-0.67 % ***
1	-0.12 %
2	-0.05 %
3	0.14 %
4	0.03 %
5	-0.22 %
6	0.49 % ***
7	0.26 %
8	0.07 %
9	0.00 %
10	-0.11 %

Days [t1, t2]	CAAR
[-20, 10]	-0.66 %
[-20, -6]	-0.91 %
[-5, -1]	-0.48 %
[0, 1]	-0.79 % ***
[2, 5]	-0.11 %
[6, 10]	0.71 % *

Significance levels: *** 1%, ** 5% and * 10%.

There is also a small statistically significant positive abnormal return on day [6] which may imply that the market mitigates the negative stock price reaction occurring on the event day after a few days. However, many studies have already proven that one cannot be sure about the direction of stock prices caused by rating downgrades and upgrades. Because rating downgrades and upgrades may cause stock prices to move to opposite directions, there is no point to further analyze Table 7.

The analysis becomes more useful when the subgroups of downgrades and upgrades are separated from each other. Figure 6 demonstrates the CAARs for rating upgrades and downgrades over the [-20, 10] event window. The asymmetry in market reaction to rating upgrades and downgrades is clearly visible in the figure. While the graph of rating upgrades does not seem to fluctuate significantly from zero, the graph of rating downgrades seems to anticipate the rating change, at least to some extent. The reaction to the announcement also seems stronger if compared with rating upgrades. The stock returns also throwback a bit in the case of rating downgrades after day 5. The reaction to rating upgrades on the event day seems insignificant. According to the figure, it is challenging to further observe whether upgrades have impact on stock returns at all.

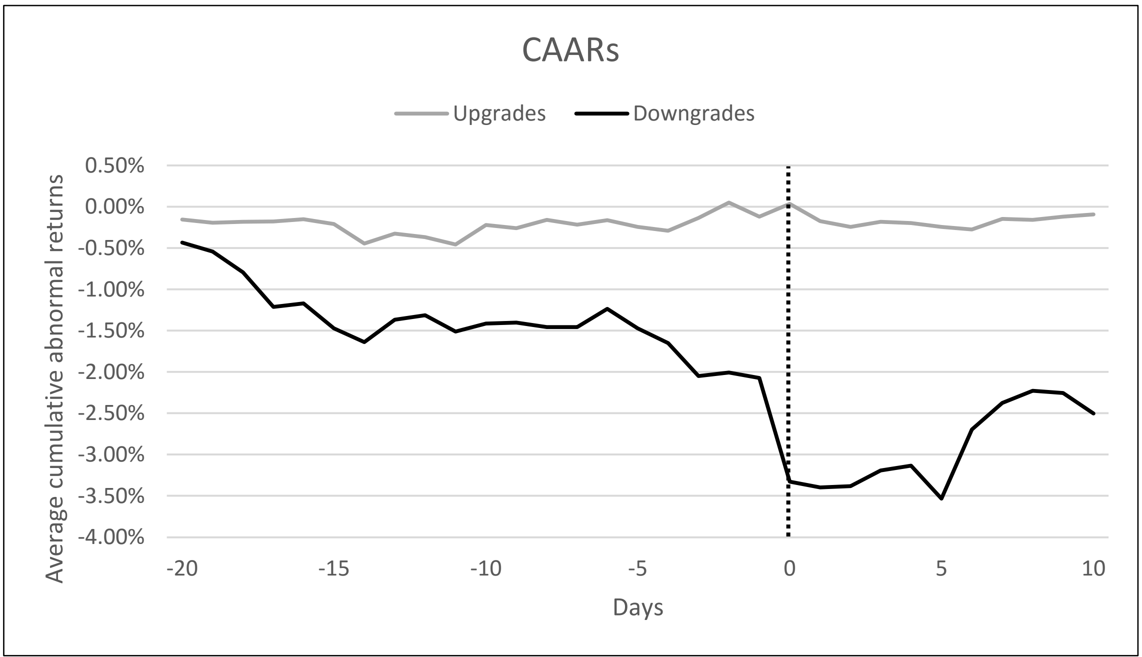

Figure 6 Cumulative average abnormal returns of rating changes

5.1.2 Downgrades

The AARs of rating downgrades and their statistical significance are described more comprehensively in Table 8. To give a more illustrative picture of the AARs, they are also displayed in a form of a graph in Figure 7.

Table 9 Summary of abnormal returns of rating changes (downgrades)

Day [t]	AAR
-10	0.10 %
-9	0.01 %
-8	-0.05 %
-7	0.00 %
-6	0.22 %
-5	-0.24 %
-4	-0.18 %
-3	-0.40 %
-2	0.04 %
-1	-0.07 %
0	-1.26 % ***
1	-0.07 %
2	0.02 %
3	0.19 %
4	0.06 %
5	-0.40 % *
6	0.84 % ***
7	0.32 %
8	0.15 %
9	-0.03 %
10	-0.25 %

Days [t1, t2]	CAAR
[-20, 10]	-1.03 %
[-20, -6]	-1.23 %
[-5, -1]	-0.84 %
[0, 1]	-1.33 % ***
[2, 5]	-0.14 %
[6, 10]	1.03 % **

Significance levels: *** 1%, ** 5% and * 10%.

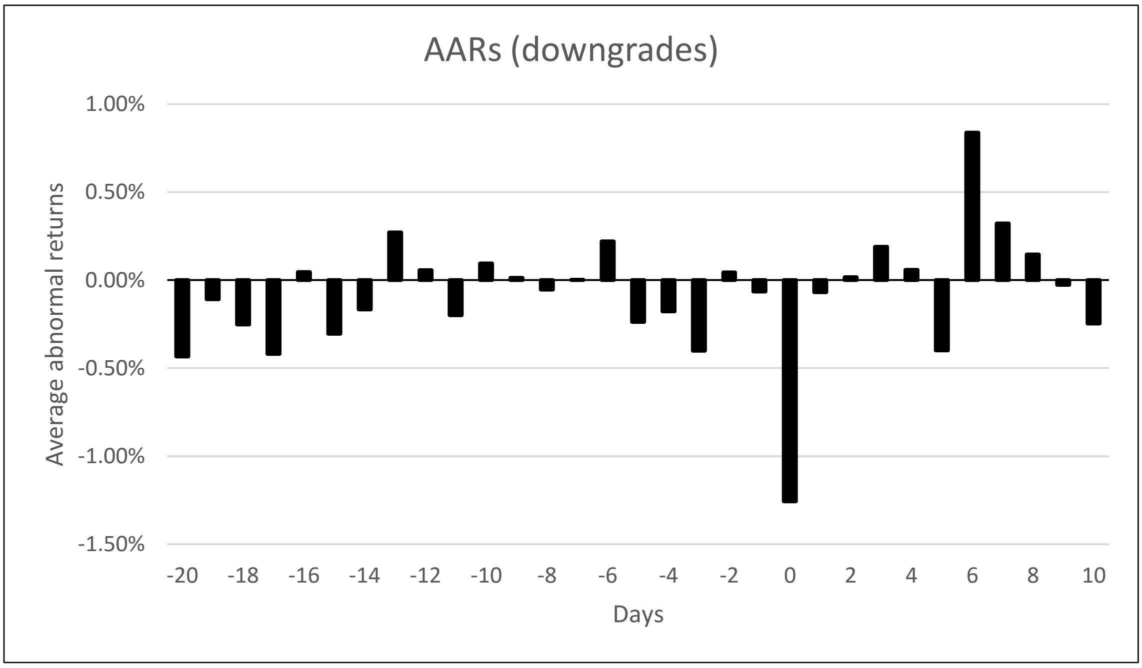

Figure 7 Average abnormal returns of rating changes (downgrades)

Table 8 supports the perceptions of what was made with regard to the rating downgrades in Figure 6. A statistically and economically significant abnormal return of -1.26 percent exists on the event day. However, any statistically significant anticipation is not observed over the period of [-20, -6] nor [-5, -1]. Thus, *H2* is accepted in the case of rating downgrades.

The period after the announcement, however, is more interesting. The first few days after the announcement are stable but the period of [6, 10] gives a CAAR of 1.03 percent, mitigating most of the downward reaction of the event day. Most of this throwback occurs on day [6] which gives a statistically significant abnormal return of 0.84 percent. Additional scrutinizing reveals that there are no particularly large negative abnormal returns of individual companies on that day. Only a bigger sample would reveal whether the abnormal return on this day is random or not. The CAAR for the whole estimation period of [-20, 10] is -1.03 percent even though it is statistically insignificant. Thus, it cannot be concluded that the event would affect the market more but slightly on the event day. All in all, *H3a* can be accepted.

5.1.3 Upgrades

The results concerning rating upgrades are shown in Table 9 and Figure 8. The statement that rating upgrades are not associated with significant abnormal returns is fully supported by this thesis. None of the AARs in the event window are statistically significant. The abnormal return on the announcement day is only 0.16 percent. Table 9 and Figure 8 show clearly that the variation in the event window also looks very stable and the absolute value does not exceed 0.25 percent for the whole period. According to these results, also *H3b* is accepted.

Table 10 Summary of abnormal returns of rating changes (upgrades)

Day [t]	AAR
-10	0.24 %
-9	-0.04 %
-8	0.10 %
-7	-0.06 %
-6	0.05 %
-5	-0.08 %
-4	-0.05 %
-3	0.15 %
-2	0.19 %
-1	-0.17 %
0	0.16 %
1	-0.21 %
2	-0.07 %
3	0.06 %
4	-0.01 %
5	-0.05 %
6	-0.03 %
7	0.13 %
8	-0.01 %
9	0.04 %
10	0.03 %

Days [t1, t2]	CAAR
[-20, 10]	-0.14 %
[-20, -6]	-0.16 %
[-5, -1]	0.04 %
[0, 1]	-0.05 %
[2, 5]	-0.07 %
[6, 10]	0.15 %

Significance levels: *** 1%, ** 5% and * 10%.

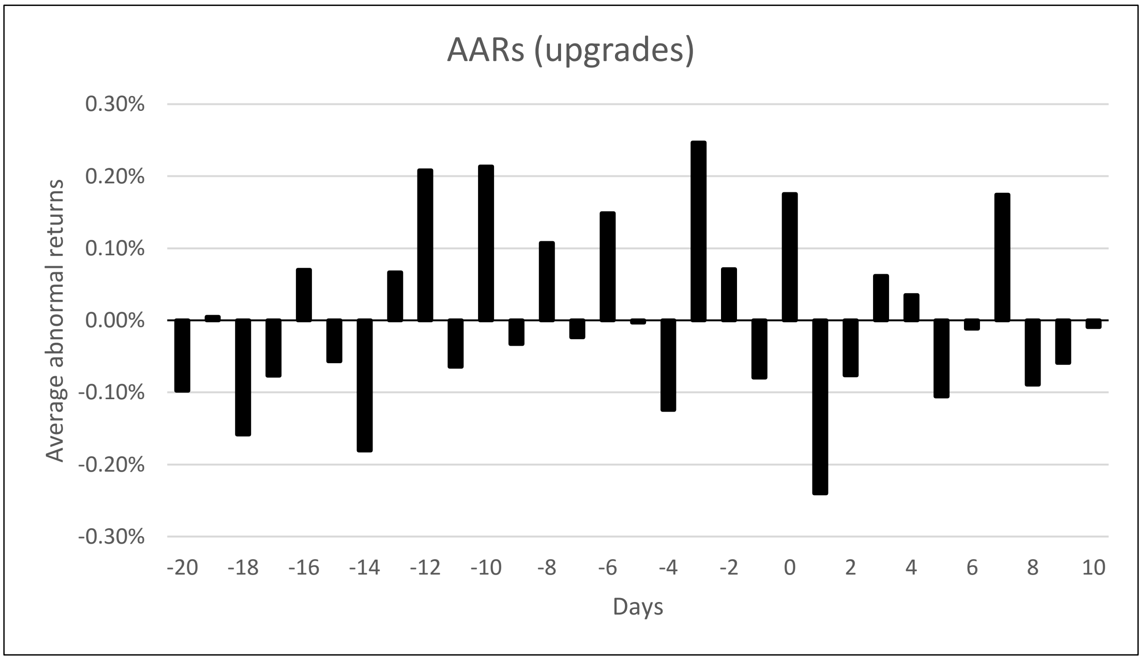

Figure 8 Average abnormal returns of rating changes (upgrades)

5.1.4 Regression results

Table 10 shows the results of the cross-sectional regression for both downgrades and upgrades. Besides the intercept, the regression model includes seven variables that are respectively: market value, number of notches, preceded credit watch, preceded rival rating, within versus between classes, investment versus non-investment grade and before versus after Lehman. The abnormal returns on the event day [0] are used as a dependent variable. As the results concerning rating downgrades reveal, the main impact is seen precisely on the event day. In the case of rating upgrades, no significant results are seen. However, it could still be possible that the cross-sectional regression revealed some determinants of the abnormal returns surrounding upgrades. As it can be seen from Table 10, this is not the case.

Table 11 Cross-sectional regression

This table provides the regression results based on equation 9. The abnormal returns on day [0] are used as a dependent variable. Besides the intercept, the regression model includes the following variables respectively: logarithmic market value, number of notches, preceded credit watch, preceded rival rating, within versus between classes, investment versus non-investment grade and before versus after Lehman.

	Downgrades	Upgrades
Variables	Coefficients	Coefficients
INTERCEPT	-0.0605 **	0.0021
log(MV)	0.0050 *	-0.0010
NOTCH	0.0049	0.0087
WATCH	0.0004	0.0002
CLEAN_30	0.0079	-0.0077
CLASS	0.0044	0.0024
JUNK	-0.0192 *	-0.0002
LEHMAN	-0.0068	0.0023
R-squared	0.0449	0.0285

Significance levels: *** 1%, ** 5% and * 10%.

The only statistically significant coefficients are related to rating downgrades. The logarithmic market value controlling firm size and the dummy variable distinguishing between investment and non-investment grades ratings are statistically significant, although only at 10 percent risk level. Some alternative variations to model the regression were also made but these two variables were the only ones to remain statistically significant in most cases.

Nevertheless, the model is rather simple and aims only to give a rough idea of potential determinants. For this purpose the model succeeded quite well and the results seem plausible. The direction of the effect seems to be what was expected for both significant variables for downgrades. The values of R-squared are small which is typical in this type of research.

The effect of the market value is positive in the case of rating downgrades. This means that the stock price reaction of larger banks is positively higher compared to smaller banks. Because the average price reaction is negative in the case of rating downgrades, the result suggests that the price reaction of larger banks is more mitigated compared to smaller banks. This, for example, gives support to the results of Fieberg et al. (2015) in the European context. It is not surprising that the coefficient is only barely statistically significant as the sample in this thesis includes medium size banks and not just small and large ones as the study of Fieberg et al. Nevertheless, the firm's size seems to matter.

The most significant variable is the one distinguishing between investment and non-investment grades for downgrades. Also, the direction of this coefficient is as expected. Taking into account how the dummy variable is coded, it being negative means that rating changes within the non-investment class react more significantly. This is rather obvious and supported by many researchers, such as Hand et al. (1992).

Statements and expectations behind other variables do not receive support in this case. There may be several reasons to explain this. First, the regression model is insufficient and thus unable to explain other variables. Second, the sample sizes for some categories are not large enough to get reliable results. For example, there are only 10 upgrades that are preceded by a rival rating within 30 days prior to the rating change. Third, the expectations concerning other variables may be wrong or the variables could be inadequate. For instance, Fieberg et al. (2015) observe an increase in negative abnormal returns post-Lehman only for small banks. If this is true, it is not surprising that the sample containing banks of all sizes is not statistically significant.

5.2 Credit rating watches

5.2.1 Overview

Figure 9 demonstrates the CAARs for positive and negative rating watches over the [-20, 10] event window. It is necessary to keep in mind that the overall amount of rating watches is much smaller compared to rating changes. The full sample includes 80 negative and only 22 positive watches announcements. The number of negative watches should be adequate but the results concerning positive ratings should be taken with caution. Due to the fact that the sample size is this small, a few observations with greater values may control the results even if some trimming is done from the tails of the distribution.

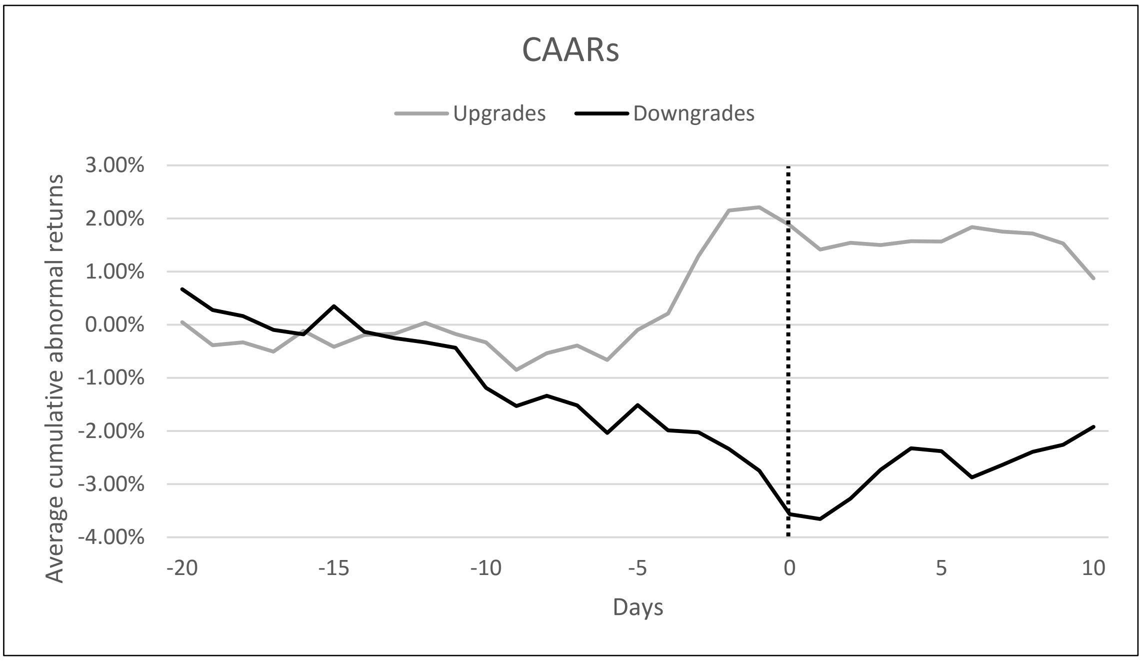

Figure 9 Cumulative average abnormal returns of rating watches

Unlike in the case of rating changes, the graphs concerning rating watches seem somewhat more symmetrical between upgrades and downgrades. There seems to be some anticipation in both cases. In the case of positive watches, the anticipation starts five days prior to the event and is stronger compared to negative watches in the same time window. In the case of negative watches, the trend is downward for the whole time period prior the event day.

The reactions on the event day, however, look rather slight in both cases. It is curious that the reaction is negative also in the case of positive watches, although, unlikely to be statistically significant. After the event day, the stock prices start to throwback in the case of negative watches. There is no clear trend that can be pointed out in the case of positive watches after the announcement.

5.2.2 Downgrades

The AARs of the negative watches and their statistical significance are described in Table 11. For a more illustrative picture of the AARs, they are also displayed in the form of a graph in Figure 10.

Table 12 Summary of abnormal returns of rating watches (downgrades)

Day [t]	AAR
-10	-0.75 % **
-9	-0.35 %
-8	0.20 %
-7	-0.18 %
-6	-0.52 %
-5	0.53 %
-4	-0.48 %
-3	-0.04 %
-2	-0.32 %
-1	-0.41 %
0	-0.82 % **
1	-0.09 %
2	0.39 %
3	0.54 %
4	0.40 %
5	-0.05 %
6	-0.50 %
7	0.24 %
8	0.25 %
9	0.13 %
10	0.34 %

Days [t1, t2]	CAAR
[-20, 10]	-3.88 % **
[-20, -6]	-2.04 %
[-5, -1]	-0.71 %
[0, 1]	-0.91 % *
[2, 5]	1.28 % *
[6, 10]	0.45 %

Significance levels: *** 1%, ** 5% and * 10 %.

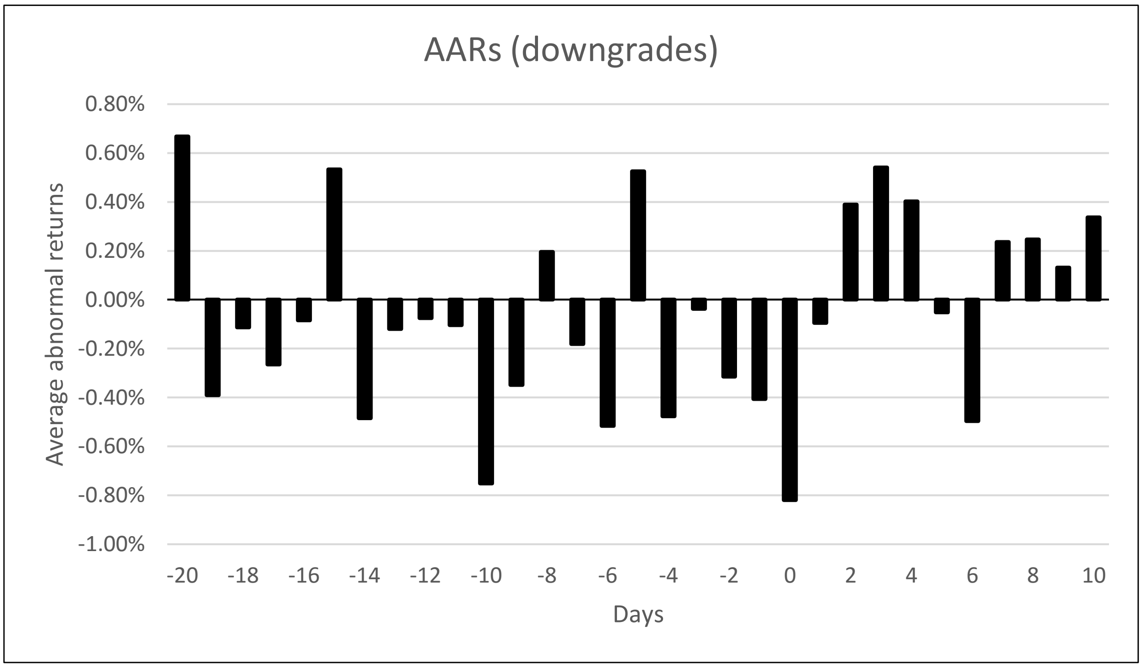

Figure 10 Average abnormal returns of rating watches (downgrades)

A return of -0.82 percent exists on the event day which is statistically significant at the 5 percent risk level. Thus, the reaction is weaker if compared with the corresponding AAR of the negative rating changes. Besides, there seems to be a random negative AAR on day [-10] which is also statistically significant at the same risk level. Interestingly this AAR is almost the same in magnitude as the AAR on the event day. It is difficult to point out any specific reason for this observation and it rather illustrates the randomness of stock returns in general. It is also good to keep in mind that the size of this sample is much smaller compared with rating changes. Thus, random patterns are much more likely to occur.

In the case of the negative rating changes, the CAAR of the [-20, 10] window is statistically insignificant. In this case the corresponding return is -3.88 percent and statistically significant. This implies that although some of the reaction was caused by rating agencies on the event day, most negative effects occurred on other days within the event window. Thus, rating watches seem to be announced during the times of critical changes in test companies and the market reacts to these changes before rating agencies.

Just as in the case of negative rating changes, positive returns appear after the event day. Most AARs after the announcement are positive in the following ten days, although economically weak. Most positive returns are seen on days [2], [3] and [4] which makes the reported CAAR of [2, 5] statistically significant at the 10 percent risk level. Overall, the results concerning negative rating watches are not as clear as the results concerning negative rating changes. With regard to the research hypothesis, *H1* can be rejected also in the case of rating watches. The statement of *H2* is only partially correct – most price reactions occur before the announcement but there is still a statistically significant abnormal return showing up just on the event day.

5.2.3 Upgrades

The AARs of positive watches and their statistical significance are described in Table 12 and graphed in Figure 11. The results concerning positive rating watches are curiously enough but not very reliable due to the sample size of only 22 observations. Additional scrutinizing reveals that although the data concerning rating watches is trimmed at the 5th and 95th percentiles, the AARs are still partially distorted by a few observations with higher values. Naturally, trimming small sample sizes is not very effective. However, it is not reasonable to remove more observations decreasing the sample size even further.

Table 13 Summary of abnormal returns of rating watches (upgrades)

Day [t]	AAR
-10	-0.16 %
-9	-0.52 %
-8	0.32 %
-7	0.14 %
-6	-0.27 %
-5	0.57 % *
-4	0.30 %
-3	1.09 % ***
-2	0.85 % ***
-1	0.06 %
0	-0.33 %
1	-0.46 %
2	0.12 %
3	-0.04 %
4	0.07 %
5	0.00 %
6	0.27 %
7	-0.09 %
8	-0.03 %
9	-0.19 %
10	-0.65 % **

Days [t1, t2]	CAAR
[-20, 10]	-1.25 %
[-20, -6]	-0.66 %
[-5, -1]	2.87 % ***
[0, 1]	-0.79 % *
[2, 5]	0.15 %
[6, 10]	-0.70 %

Significance levels: *** 1%, ** 5% and * 10 %.

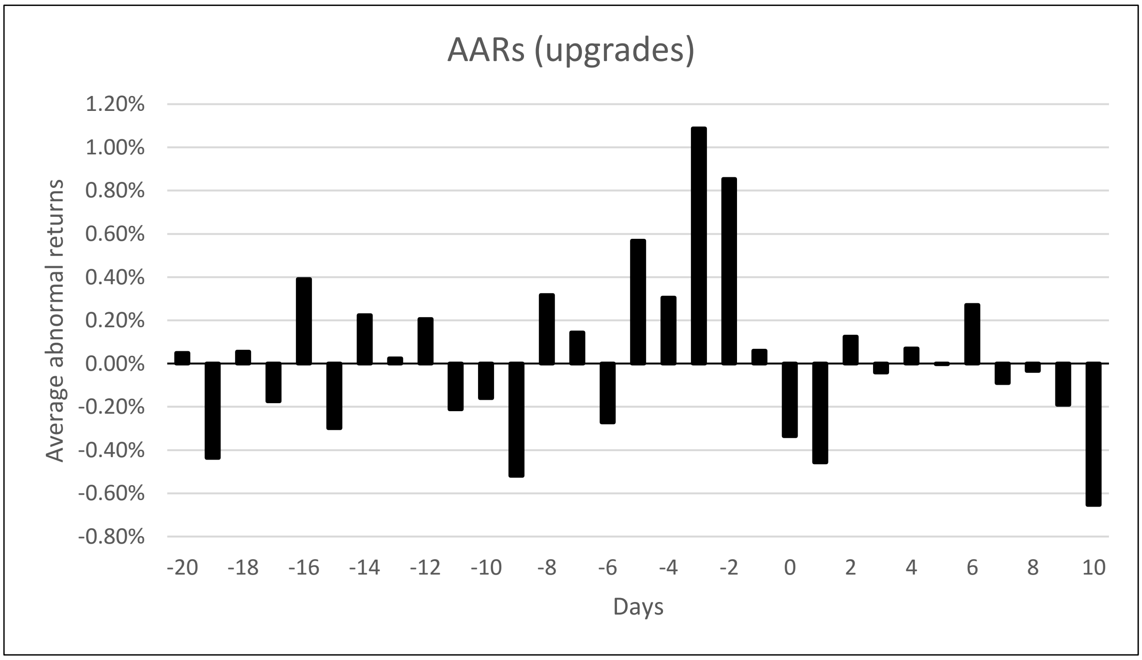

Figure 11 Average abnormal returns of rating watches (upgrades)

In any case, the negative AAR on the event day is not statistically significant unless the return on day [1] is also taken into account. Besides, it would still be only barely statistically significant. The most impact is seen in the window of [-5, -1] that produces the positive CAAR of 2.87 percent. This CAAR is also statistically significant at the 1 percent risk level. Thus, it seems that the market reacts to positive changes in test companies before rating agencies. However, regarding these results, it is rather challenging to say whether positive watch announcements have any true impact on stock returns. Additional scrutinizing reveals that the statistically significant positive reaction prior the announcement results at least partly from a group of observations with higher values. Thus, the results are not robust enough to make assumptions with regard to the research hypotheses in the case of positive rating watches.

6 SUMMARY AND CONCLUSIONS

Credit rating agencies play an important role in capital markets nowadays. Their systemic importance has grown especially during the last few decades. Besides the fact that credit ratings ameliorate information asymmetry between issuers and investors, they have a significant purpose in financial regulation. However, rating agencies and their credit ratings have received plenty of criticism over the last years. For example, credit ratings are criticized about moving too slowly and their ratings are said to be too inflexible (Altman & Saunders 2001, 6). Also, their ability to give any information of the probability of future default is questioned (see e.g. Hau et al 2013). Credit ratings are significant for firms' operation costs so it is interesting to investigate how investors react to changes in creditworthiness.

There is a voluminous body of literature studying the effect of rating changes on stock prices since the 1970s. Nevertheless, as the market data and methodological design used in this research field varies a lot, study results are often mixed and contradictory. However, there are two assumptions that a large part of researchers seem to agree with. First, a vast amount of studies report statistically significant abnormal returns on stock prices associated with credit rating announcements. Thus, rating announcements seem to provide pricing pertinent information. Second, there seems to be a trend concerning asymmetric reaction of rating downgrades and upgrades. While rating downgrades are typically associated with significantly negative abnormal returns, the reaction among rating upgrades is often more limited or insignificant.

Most previous research papers concentrate on the United States market in which the aforementioned assumptions are strongly supported. The results from other markets are often more contradictory. Even though the stock price reaction of rating announcements is a widely investigated topic, a very limited amount of papers exist, which focus solely on the banking industry. The banking sector is especially interesting due to its unique characteristics related to high regulation, opaqueness and global systemic importance. Although a couple of bank papers acknowledge the asymmetric reaction between downgrades and upgrades, it is not known whether it is a common trend also in this field.

The main objective of this thesis is to examine the short-term effect of credit rating announcements on daily stock returns for European banks indexed in STOXX Europe 600 Banks. In practice this includes using data of 41 exchange-listed banks from 14 European nations. The full sample consists of 321 credit rating changes and 102 rating watch announcements. Rating announcements are gathered from S&P, Moody's and Fitch and represent banks' solicited long-term issuer ratings. The investigation period of this study dates from the beginning of 2002 to the end of 2015.

The study objective is achieved by conducting an event study. The event study is extended with a cross-sectional linear regression analysis to investigate other potential determinants surrounding credit rating changes. The research hypotheses and the motivation for additional tests are derived from past research. The main hypotheses are formed in order to explore; whether rating announcements have an effect on stock prices, when does this possible reaction occur and whether it is asymmetric between rating upgrades and downgrades. The additional tests include investigating the following determinants: market value of a firm (cf. Fieberg et al. 2015), number of notches by which each rating is changed (cf. Schweitzer et al. 1992), whether a rating is preceded by a credit watch (cf. Gropp & Richards 2001), whether a rating was preceded by a rival rating within a prior month (cf. Schweitzer et al. 1992), whether a rating changes between or within classes (cf. Holthausen & Leftwich 1986), whether a rating changes within the investment or below investment class (cf. Hand et al. 1992) and whether a rating occurred before or after the collapse of Lehman Brothers (cf. Fieberg et al. (2015).

The findings of this thesis provide evidence that rating announcements have an impact on stock returns in the context of European banks. Thus rating announcements include information which has not before been known in the public domain. The results also support the existence of an asymmetry in capital market reactions to rating downgrades and upgrades. Rating downgrades are associated with significantly negative abnormal returns on the announcement day although the reaction is rather modest (-1.26%). No statistically significant reaction is found associated with rating upgrades on the event day. These results hold true both with rating changes and rating watches.

There is no statistically significant anticipation observed in the case of credit rating changes. Besides, the stock price reaction in the case of rating downgrades occurs precisely on the event day and not after. Thus, the market seems to be efficient in this matter. However, the price reaction of negative ratings is only temporary and it is compensated within the next few days. The results concerning rating changes are somewhat in line with the results of another European based bank study by Gropp and Richards (2001) who also find only a small abnormal return associated with rating downgrades on the event day. Besides they do not find evidence of abnormal returns over longer pre- or post-announcement periods. However, they do find small abnormal returns associated with rating upgrades which contradicts the results obtained in this study.

The results are not as clear concerning rating watches. This is not surprising as rating watches are not as followed by the public as the actual rating changes. Also, the sample size of rating watch announcements is much smaller compared with rating changes in this study so the results are less reliable. In the case of negative watches, there is a precise price reaction which occurs on the event day although the reaction is even smaller compared to rating downgrades (-0.82%). No statistically significant reaction associated with positive watches if found. It seems that the market anticipates the changes in test companies before rating agencies, as there is a statistically significant cumulative price reaction occurring before the event day for both negative and positive watches announcements. In the case of negative watches, the negative cumulative price reaction is divided into a longer time period prior to the announcement. For positive watches, the positive price reaction is observed on the very few days before the event day. Just as with negative rating changes, positive returns appear after the event day in the case of negative watches compensating some of the price reaction. A similar pattern is not observed with positive watches. Altogether, it seems that rating watches provide some new information to the European bank market in the case of negative watches even though the reaction is very limited.

The event study is extended with a cross-sectional regression analysis to investigate other potential determinants surrounding rating downgrades and upgrades on the event day. The only barely statistically significant variables are associated with the rating downgrades and represent the market value of a firm and the distinction between investment and below investment grade classes. The latter observation is intuitive as investors are more concerned about their investments in lower-rated companies. Thus, the stock price reaction after rating downgrades is stronger in the case of non-investment grade companies. This study result is in line with the results of, for instance, Hand et al. (1992).

A debate has emerged about the implicit public "too big to fail" guarantee benefiting large banks. The TBTF subsidy refers to a situation in which investors view investments with these banks as a safer investment than deposits with smaller banks. This subsidy may be seen, for example, in abnormal returns surrounding rating announcements. As the variable controlling firm size is statistically significant in this study, this means that the price reaction of larger banks is more mitigated compared to smaller banks in the case of rating downgrades. This supports, for example, the study results of a more comprehensive and universal bank study of Fieberg et al. (2015) and may provide evidence of the existence of the TBTF subsidy for larger banks. Of course it is good to keep in mind that larger banks are often more widely followed by the public which may also be seen in the results.

Overall, according to this thesis, credit rating announcements seem to provide some modest pricing pertinent information in the case of European banks. Besides supporting the common trend in this research field, the results of this study may provide some useful abnormal return information for investors interested in the European bank market.

However, there are some limitations that should be noticed when scrutinizing the results. First, the sample size is rather small with regard to some subgroups. Especially, the results concerning positive rating watches should be taken with a grain of salt as the sample size is only 22 observations. Second, the data is clearly biased towards certain countries. Thus, it cannot be said that the results could be generalized to take into account whole of Europe. Rather, the study focuses on certain EU-15 countries (and the Czech

Republic) in which some nations are more stressed than the others. As the legislative and financial environment varies among these countries, the results are likely to be biased.

Also, this study does not take into account the data contamination. A rating announcement is contaminated, if there are earnings announcements or other relevant disturbing news stories around the announcement, such as, other rating announcements. Only the preceding rival ratings within one month prior the event are controlled in this study in a form of a dummy variable to test whether this information matters. According to this study, it does not matter. However, the test method used in this study is rather simple. There exists more sophisticated methods and theories in literature on how rival ratings could be taken care of. All in all, other relevant news stories around the event day are more likely to cause biases in this thesis.

There have been studies arguing that the reason for the rating change may matter. For example, Goh and Ederington (1993) state that if a downgrade results from an increase in leverage that will transfer wealth from bondholders to stockholders, this should have a positive effect on stock prices but a negative effect on bonds. However, Fieberg et al. (2015) note that this may be unlikely with respect to banks. This, because a bank capital structure is typically heavily biased towards debt financing. This gives them a very limited leeway for changes in leverage which would benefit shareholders. Nevertheless, the theory of Goh and Ederington may still have an impact on the results.

The limitations related to this study naturally arise ideas for further research. Increasing the sample size alone would give more reliable results and allow further dividing into smaller subsamples. Besides, using also European banks not included in STOXX Europe 600 Banks Index could decrease the bias towards certain countries. Also, dealing with contaminated data by scanning relevant news around events and removing the confounding observations might improve the results. It is also good to keep in mind that comprehensive data cleaning is impossible in reality.

In this thesis the data is restricted to only issuer ratings as those best reflect the issuers' overall credit quality. However, rating agencies also provide a range of other ratings that could be used in this type of a study. Taking instrument ratings also into consideration, the amount of available ratings could be increased. For instance, senior unsecured debt remains the largest source of long-term funding for banks so its bond ratings could be considered as one possibility.

At least to the knowledge of the author, there have not been many academic papers comparing industries in this research field. A comprehensive industry comparison focusing on a specific market or even at a global scale would be interesting. For example, according to this study, the size of a bank matters which may give evidence for the existence of TBTF subsidy for larger banks. However, it would be useful to investigate whether price reactions of large banks differ from price reactions of other large companies in general.

REFERENCES

Adams, C. – Mathieson, D. – Schinasi G. (1999) *International capital markets: Developments, prospects, and key policy issues*. International Monetary Fund, Washington D.C, USA.

Afonso, A. – Furceri, D. – Gomes, P. (2012) Sovereign credit ratings and financial markets linkages: Application to European data. *Journal of International Money and Finance*, Vol. 31, 606–638.

Altman, E. – Saunders, A. (2001) An analysis and critique of the BIS proposal on capital adequacy and ratings. *Journal of Banking and Finance*, Vol. 25, 25–56.

Baklanova, V. (2009) Regulatory use of credit ratings: How it impacts the behavior of market constituents. *International Finance Review*, Vol. 10, 65–104.

Ball, R. – Brown, P. (1968) An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, Vol. 6, 159–178.

Barron, M. – Clare, A. – Thomas, S. (1997) The effect of bond rating changes and new ratings on UK stock returns. *Journal of Business Finance & Accounting*, Vol. 24 (3), 497–509.

Beaver, W. – Shakespeare, C. – Soliman, M. (2006) Differential properties in the ratings of certified versus non-certified bond-rating agencies. *Journal of Accounting and Economics*, Vol. 42 (3), 303–334.

Behr, P. – Güttler, A. (2008) The informational content of unsolicited ratings. *Journal of Banking & Finance*, Vol. 32 (4), 587–599.

BCBS (2006) International convergence of capital measurement and capital standards. Bank for International Settlements. <http://www.bis.org/publ/bcbs128.pdf>, retrieved 19.7.2016.

BCBS (2010) Basel III: A global regulatory framework for more resilient banks and banking systems. Bank for International Settlements. <http://www.bis.org/publ/bcbs189.pdf>, retrieved 8.6.2016.

BCBS (2015) Revisions to the standardised approach for credit risk. Consultative document. Bank for International Settlements. <http://www.bis.org/bcbs/publ/d307.pdf>, retrieved 8.6.2016.

Billett, M. – Garfinkel, J. – O'Neal, E. (1998) The cost of market versus regulatory discipline in banking. *Journal of Financial Economics*, Vol. 48, 333–358.

BIS (2015) History of the Basel Committee. <http://www.bis.org/bcbs/history.htm>, retrieved 27.3.2016.

BIS Speeches. Speech by Stefan Ingves at the IIF Annual Membership Meeting. <https://www.bis.org/speeches/sp151010.htm>, retrieved 2.11.2015.

Brown, S. – Warner, J. (1980) Measuring security price performance. *Journal of Financial Economics*, Vol. 8, 205–258.

Brown, S. – Warner, J. (1985) Using daily stock returns: The case of event studies. *Journal of Financial Economics*, Vol. 14, 3–31.

Cable, J. – Holland, K. (1999) Modelling normal returns in event studies: A model-selection approach and pilot study. *The European Journal of Finance*, Vol. 5, 331–341.

Cancian, D. (2016) European banking sector records worst performance since 2008 amid global turmoil. *International Business Times*. <http://www.ibtimes.co.uk/european-banking-sector-records-worst-performance-since-2008-amid-global-turmoil-1542643>, retrieved 17.7.2016.

CFR (2015) The credit rating controversy. <http://www.cfr.org/financial-crises/credit-rating-controversy/p22328>, retrieved 6.7.2016.

Chan, P. – Edwards, V. – Walter, T. (2009) The information content of Australian credit ratings: A comparison between subscription and non-subscription-based credit rating agencies. *Economic Systems*, Vol. 33 (1), 22–44.

Corrado, C. (2011) Event studies: A methodology review. *Accounting and Finance*, Vol. 51, 207–234.

Cowan, A. – Sergeant, A. (1996) Trading frequency and event study test specification. *Journal of Banking & Finance*, Vol. 20, 1731–1757.

Daniels, K. – Jensen, M. (2004) The effect of credit ratings on credit default swap spreads and credit spreads. *The Journal of Fixed income*, Vol. 15 (3), 16–33.

Dittrich F. (2007) The credit rating industry: Competition and regulation. Working Paper, University of Cologne, Department of Economics.

Dolley, J. (1933) Characteristics and procedure of common stock split-ups. *Harvard Business Review*, Vol. 11, 316–326.

Dutta, A. (2014) Parametric and nonparametric event study tests: A review. *International Business Research*, Vol. 7 (12), 136–142.

Elayan, F. – Hsu, W. – Meyer, T. (2003) The information effect of credit rating announcements on share prices in a small market. *Journal of Economics and Finance*, Vol. 27 (3), 337–356.

Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25 (2), 383–417.

Fama, E. – Fisher, L. – Jensen, M. – Roll, R. (1969) The adjustments of stock prices to new information. *International Economic Review*, Vol 10 (1), 1–21.

Fieberg, C. – Körner, F. – Prokop, J. – Varmaz, A. (2015) Big is beautiful: The information content of bank rating changes. *The journal of Risk Finance*, Vol. 16 (3), 233–252.

Fitch Ratings (2016a) Understanding credit ratings — Limitations and usage. <https://www.fitchratings.com/jsp/general/RatingsDefinitions.faces?context_ln=5&detail_ln=500&context=5&detail=509>, retrieved 27.3.2016

Fitch Ratings (2016b) Rating watches and rating outlooks. <https://www.fitchratings.com/jsp/general/RatingsDefinitions.faces?context=5&detail=518&context_ln=5&detail_ln=500>, retrieved 19.7.2016

Galil, K. – Soffer, G. (2011) Good news, bad news and rating announcements: An empirical investigation. *Journal of Banking and Finance*, Vol. 35, 3101–3119.

Ghachem, D. (2015) The determinants of rating announcements impact on stock markets during crisis periods: The case of the 2008 worldwide financial crisis. *International journal of economics and finance*, Vol. 7 (4), 159–175.

Goh, J. – Ederington, L. (1993) Is a bond rating downgrade bad news, good news, or no news for stockholders? *Journal of Finance*, Vol. 48 (5), 2001–2008.

Gropp, R. – Richards, A.J. (2001) Rating agency actions and the pricing of debt and equity of European banks: What can we infer about private sector monitoring of bank soundness? *European Central Bank Working Paper Series*. <http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp076.pdf>, retrieved 22.7.2016.

Habib, Y. – Nazir, M. – Hashmi, S. – Saeed, M. (2015) Credit rating announcements and stock returns: Evidence from the banking sector of Pakistan. *Journal of Business Studies Quarterly*, Vol. 7 (2), 61–84.

Han, S. – Shin, Y. – Reinhart, W. – Moore, W. (2009) Market segmentation effects in corporate credit rating changes: The case of emerging markets. *Journal of Financial Services Research*, Vol. 35 (2), 141–166.

Hand, J. – Holthausen, R. – Leftwich, R. (1992) The effect of bond rating agency announcements on bond and stock prices. *Journal of Finance*, Vol. 47, 733–752.

Hau, H. – Langfield, S. – Marques-Ibanez, D. (2013) Bank ratings: What determines their quality. *Economic Policy*, Vol. 28 (74), 289–333.

Holthausen, R. – Leftwich, R. (1986) The effect of bond rating changes on common stock prices. *Journal of Financial Economics*, Vol. 17 (1), 57–89.

Hull, J. – Predescu, M. – White, A. (2004) The relationship between credit default swap spreads, bond yields, and credit rating announcements. *Journal of Banking & Finance*, Vol. 28 (11), 2789–2811.

Investopedia (2014) The industry handbook: The banking industry. <http://www.investopedia.com/features/industryhandbook/banking.asp>, retrieved 14.7.2016.

Investopedia (2016) Credit ratings. <http://www.investopedia.com/terms/c/creditrating.asp>, retrieved 19.7.2016.

Jorion, P. – Zhang, G. (2010) Information transfer effects of bond rating downgrades. *Financial Review*, Vol. 45 (3), 683–706.

Kothari, S. – Warner, J. (2004), Econometrics of event studies. In: *Handbook of corporate finance: Empirical corporate finance*, B. Espen Eckbo (ed.), 3–36, Handbooks in Finance Series, Elsevier, North-Holland. <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=608601>, retrieved 5.6.2016.

Kwok, C. – Meznar, M. – Nigh, D. (1998) Announcement of withdrawal from South Africa revisited: Making sense of contradictory event study findings. *Academy of Management Journal*, Vol. 41 (6), 715–730.

Langohr, H – Langohr, P. (2010) *The rating agencies and their credit ratings: What they are, how they work, and why they are relevant*. John Wiley & Sons, Chichester, England.

Li, J. – Shin, Y. – William, T. (2006) Reactions of Japanese markets to changes in credit ratings by global and local agencies. *Journal of Banking & Finance*, Vol. 30 (3), 1007–1021.

Li, H. – Visaltanachoti, N. – Kesayan, P. (2004) Effects of credit rating announcements: Swedish stock market. *International Journal of Finance*, Vol. 16 (1), 2872–2891.

Liapis, K. (2011) The regulation framework for the banking sector: The EMU, European banks and rating agencies before and during the recent financial and debt crisis. *International Journal of Disclosure and Governance*, Vol. 9 (4), 301–330.

Lim, K – (2011) *Financial valuation and econometrics*. World Scientific Publishing, Singapore.

Linciano, N. (2004) The reaction of stock prices to rating changes. Quaderni di Finanza. English version, No. 57. Commissione Nazionale per le Società e la Borsa. <www.consob.it/documenti/quaderni/qdf57en.pdf>, retrieved 5.6.2016.

Loon, E., van – Haan, J., de (2015) Location of banks and their credit ratings. *The Journal of Risk Finance*, Vol. 16 (3), 220–232.

MacKinlay, A. (1997) Event studies in economics and finance. *Journal of Economic Literature*, Vol. 35, 13–39.

Matolcsy, Z. – Lianto, T. (1995) The incremental information content of bond rating revisions: The Australian evidence. *Journal of Banking & Finance*, Vol. 19 (5), 891–902.

Moody's (2015) Rating action: Moody's reviews global bank ratings. <https://www.moodys.com/research/Moodys-reviews-global-bank-ratings--PR_321005>, retrieved 18.7.2016.

Moody's (2016) Rating methodology – Banks. <https://www.moodys.com/researchdocumentcontentpage.aspx?docid=PBC_186998>, retrieved 27.3.2016.

Morgan, D. (2002) Rating banks: Risk and uncertainty in an opaque industry. *The American Economic Review*, Vol. 92 (4), 874–888.

Pagano, M. – Volpino, P. (2010) Credit ratings failures and policy options. *Economic Policy* Vol. 62, 401–431.

Partnoy, F (2006) How and why credit rating agencies are not like other gatekeepers. In: *Financial gatekeepers: Can they protect investors?*, Yasuyuki Fuchita & Robert Litan (eds.), 59–102, Brookings Institution Press. <http://ssrn.com/abstract=900257>, retrieved 5.6.2016.

Peterson, P. (1989) Event studies: A review of issues and methodology. *Quarterly Journal of Business and Economics*, Vol. 28 (3), 36–66.

Pinches, G. – Singleton, J. (1978) The Adjustment of stock prices to bond rating changes. *Journal of Finance*, Vol. 33 (1), 29–44.

Poon, W. – Chan, K. (2008) An empirical examination of the informational content of credit ratings in China. *Journal of Business Research*, Vol. 61 (7), 790–797.

Rhee, R. (2015) Why credit rating agencies exist. *Economic Notes*, Vol. 44 (2), 161–175.

Richards, A. – Deddouche, D. (1999). Bank rating changes and bank stock returns; puzzling evidence from the emerging markets. *Working Paper of the International Monetary Fund.*

Schweitzer, R. – Szewczyk, S. – Varma, R. (1992) Bond rating agencies and their role in bank market discipline. *Journal of Financial Services Research*, Vol. 6, 249–263.

SEC (2015) Annual report on nationally recognized statistical rating organizations. <https://www.sec.gov/ocr/reportspubs/annual-reports/2015-annual-report-on-nrsros.pdf>, retrieved 16.7.2016.

S&P (2015) Rating methodologies and approaches. <http://www.standardandpoors.com/aboutcreditratings/>, retrieved 27.10.2015.

S&P (2016) Understanding ratings. <http://www.spratings.com/en_US/understanding-ratings>, retrieved 27.3.2016.

Sylla, R. (2001) A historical primer on the business of credit ratings. *The Role of Credit Reporting Systems in the International Economy. Washington, DC: The World Bank.*

Wakeman, L. (1978) Bond rating agencies and the capital markets. Working Paper, University of Rochester N.Y.

Wansley, J. – Glascock, J. – Clauretie, T. (1992) Institutional bond pricing and information arrival: The case of bond rating changes. *Journal of Business Finance & Accounting*, Vol. 19 (5), 733–750.

Watson, R – Carter, J. (2006) *Asset securitisation and synthetic structures: Innovations in the European credit markets.* Euromoney Books, London.

Weinstein, M. (1977) The effect of a rating change announcement on bond price. *Journal of Financial Economics*, Vol. 5, 329–350.

Wells, W. (2004) A beginner's guide to event studies. *Journal of Insurance Regulations*, Vol. 22 (4), 61–70.

White, L. (2010) Markets: The credit rating agencies. *Journal of Economic Perspectives*, Vol. 24 (2), 211–226.

Appendix 1 List of banks used in the study

Name	Country	S&P	Moody's	Fitch	Issuer ratings available
Alpha Bank AE	Greek	9	-	6	2008–2015
Banca Monte dei Paschi di Siena SpA	Italy	2	-	3	2008–2015
Banca Popolare dell'Emilia Romagna SC	Italy	5	-	3	2004–2015
Banca Popolare di Milano Scarl	Italy	4	5	3	2002–2015
Banca Popolare di Sondrio Scarl	Italy	-	-	2	2005–2013
Banco Bilbao Vizcaya Argentaria SA	Spain	3	4	1	2007–2015
Banco Comercial Portugues SA	Portugal	5	-	3	2006–2015
Banco de Sabadell SA	Spain	3	-	1	2007–2014
Banco Popolare SC	Italy	4	2	2	2002–2015
Banco Popular Espanol SA	Spain	5	-	3	2008–2015
Banco Santander SA	Spain	6	6	4	2002–2015
Bank of Ireland	Ireland	5	7	4	2007–2015
Bankia SA	Spain	2	-	2	2013–2015
Bankinter SA	Spain	2	-	1	2010–2014
Barclays PLC	Great Britain	3	2	2	2008–2015
BNP Paribas SA	France	3	4	1	2002–2014
CaixaBank SA	Spain	1	1	-	2014–2015
Commerzbank AG	Germany	4	3	5	2002–2015
Credit Agricole SA	France	1	2	3	2003–2013
Danske Bank A/S	Denmark	3	6	1	2004–2015
Deutsche Bank AG	Germany	6	3	2	2002–2015
Erste Group Bank AG	Austria	2	-	1	2014–2015
Eurobank Ergasias SA	Greek	9	-	8	2005–2015
HSBC Holdings PLC	Great Britain	2	-	1	2004–2015
ING Groep NV	Netherlands	5	-	2	2002–2013
Intesa Sanpaolo SpA	Italy	7	3	2	2002–2015
Jyske Bank A/S	Denmark	2	-	-	2007–2009
KBC Groep NV	Belgium	4	3	3	2006–2015
Komercni banka as	Czech	7	-	7	2002–2013
Lloyds Banking Group PLC	Great Britain	4	-	3	2007–2015
Mediobanca SpA	Italy	3	-	-	2010–2014
Natixis SA	France	5	2	5	2002–2013
Nordea Bank AB	Sweden	1	-	-	2005
Raiffeisen Bank International AG	Austria	2	-	1	2014–2015
Royal Bank of Scotland Group PLC	Great Britain	5	-	4	2003–2015
Skandinaviska Enskilda Banken AB	Sweden	3	6	-	2003–2015
Societe Generale SA	France	3	-	5	2003–2013
Standard Chartered PLC	Great Britain	3	-	2	2006–2015
Swedbank AB	Sweden	3	7	1	2006–2015
Svenska Handelsbanken AB	Sweden	1	3	-	2004–2015
UniCredit SpA	Italy	5	-	3	2014