Tightened standards or business as usual?

- A study of European corporate credit rating cyclicality

by

Simen Christoffer Gunnerud

Simen Granviken Woxmyhr

Advisor

Francisco Santos

Master Thesis in Financial Economics

Norwegian School of Economics

Bergen, June 2014

Abstract

This paper examines which variables are statistically significant in a corporate credit rating process, and if there has been a tightening in rating practice during and after the financial crisis hit Europe in 2008. We investigate whether we can argue that the possible tightening of standards can be seen as a procyclical way of assessing credit ratings. By procyclical ratings we refer to CRAs' tendency to excessive downgrade in recession, leading to higher capital cost for the downgraded companies, thus leading to even poorer results and an intensified recession.

Utilizing annual fundamental values from firms in the Eurozone and the UK from 2004-2012 with a Moody's rating, we find indications for such a tightening of standards, but the evidence is not strong enough to draw any rigid conclusions. All our different models over predict post-crisis ratings, with significant difference of means. Further we find a general downgrading post-crisis of approximately 0.75, where a value of 1 equals one sub-rank rating, adjusted for the effects from corporate fundamental variables. Breaking it down into cross-section analysis, we find strong evidence that some industries and regions may have experienced stricter standards than others.

Our analysis is based on two different methods; (1) building models based on pre-crisis data using corporate financials and macroeconomic variables to compare an out-of- sample estimation with the actual post-crisis ratings, and (2) using propensity score matching also using corporate financials.

Keywords: Corporate Credit Rating, Procyclicality, CRAs, Eurozone, Propensity Score Matching, Panel Data, Pre-crisis vs. Post-crisis.

Preface

This paper concludes five years of study at the Norwegian School of Economics (NHH) and by that our Master of Science and Business Administration. The paper is written as a part of our major in financial economics.

Our main motivation for writing about corporate credit ratings was the desire to learn more about how ratings are assessed, as they have been frequently debated in the media the recent years. The motivation grew as we found that no comprehensive study had been made on the European market after the financial crisis, giving us an opportunity to explore a field until now somewhat unknown. Our research questions were developed along the process after we acknowledged our limitations - or in our opinion opportunities - with the data set we constructed. Our focus has been on generating results that could fill parts of the gap in the research made on cyclicality in corporate credit ratings in Europe.

We would like to thank our advisor, Francisco Santos (NHH), for professional and productive guidance. We also thank him for showing interest in our work and motivating us along the process.

We also need to thank our friends and families for supporting us in this anchor leg in our studies at NHH.

Bergen, June 20th 2014

Simen Christoffer Gunnerud   Simen Granviken Woxmyhr

Contents

List of Tables

List of Figures

1 Introduction

In February 2013 America's Department of Justice sued Standard and Poor's for $5 billion, accusing the credit rating agency (CRA) for having knowingly issued overgenerous ratings. In light of the financial crisis the credit rating agencies in general were blamed by the public to be inaccurate in their ratings and to have been too optimistic in their assessment of credit risk in both corporations and nations. History shows that they failed to downgrade the large investment banks with high- risk structured financial instruments in their balance sheets.

There has been a lot of focus on the U.S. market and the European sovereign debt crisis in recent time. Far less focus has been set on the European corporate bond market. In fact, we have not found any comprehensive studies on this topic since the outburst of the financial crisis. This paper wants to fill that gap by analyzing credit ratings on corporate bonds in the Eurozone in the time just before and just after the financial crisis.

We construct a data set that includes long-term corporate credit ratings from Moody's Investor Service, accessed from their own database. All corporations except banks and insurance companies with a rating at a time between 2004 and 2012 are included in the data set. Financial variables have been exported from Orbis [1] database, which covers most financial variables mentioned in the credit rating agencies' criteria guides and other empirical studies. The sample is limited to corporations within the Eurozone to magnify the possible effects of the financial crisis, which indeed hit the Eurozone hard. However, to increase the number of observations the U.K. has been included.

The goal of this paper is two-fold. On one hand, we want to enlighten the process in which these ratings are assessed. We look at which financial variables are included in the rating process and the importance of quantitative variables. On the other hand, we want to analyze whether there has been a tightening in the rating practice. We will refer to this possible tightening of standards and the CRAs' tendency to excessive downgrade in recession as a procyclical rating anatomy. This because downgrades lead to higher capital cost for the affected companies, thus leading to even poorer results and an intensified recession.

[1] Bureau Van Dijk - Orbis database

The results are, on average, that our models over estimate ratings for the post-crisis period. The results are robust with a significant difference of means and suggest that corporates with the same financial fundamentals would receive a lower credit rating after the crisis. Further we find that the general downgrade post-crisis is approximately 0.75, where a value of 1 equals one sub-rank rating, adjusted for the effects from corporate fundamental variables.

The variables found significant in our empirical model are: *interest cover, credit period, sales*, and *solvency ratio asset based*. These are measures of respectively *operational, size* and *structure*. The significant variables are in line with what Koller *et al.* (2010), Amato and Furfine (2004) and Blume *et al.* (1998) find in their studies regarding credit ratings in the U.S., yet there are some differences.

Our analysis is general in its nature, by accumulating observations from countries and industries. Due to sample size, cross-section analysis is a challenge, but by carefully choosing sections we end up with some interesting results for further research. When analyzing the changes in ratings from pre-crisis to post-crisis period, between speculative- and investment rated corporates, we cannot find any clear difference. There is however, a tendency of a smaller downgrade for corporates within Aaa-rated countries compared to corporates with domicile in the other countries, suggesting that the CRAs has a higher threshold for downgrades within more robust economies.

We also find support for the CRAs being reactive, or at least not proactive, in their rating behavior. This is shown through consistently larger difference in ratings for all significant matching methods when changing the post-crisis period start from 2008 to 2009. It seems, in other words, that the rating change is lagging compared to market fluctuations. It additionally supports our previous results that indicate a general tightening of rating policy.

This paper takes on several different methods to analyze the possible procyclicality of credit ratings. Initially we regress all variables included in the data set to find any statistically significant relationship with credit ratings. Then we build two credit rating benchmark models and one empirical model to conduct an out-of-sample estimation comparing with actual ratings. As a final stage in the analysis we use propensity score matching.

When building the benchmark models, we look at Standard and Poor's and Moody's Investor Service's corporate criteria methods. We use these general guidelines to create an unrestricted benchmark model built on an ordinary least squares platform by using the data from 2004-2007. By restricting this model and taking out those variables showing as insignificant, we find that only a few of those variables listed in the criteria methods remain in the model. Taking it a step further, we create an empirical model based on findings in relevant studies built on a panel data regression platform with fixed effects for the same time period. We use the benchmark models and the empirical model to make an out-of-sample estimation of ratings through 2008-2012.

We also use propensity score matching. By using propensity score matching, we are comparing corporate rating observations between 2004-2007 and 2008-2012, by pairwise matching observations from each period with similar corporate fundamental characteristics such as *size* and *profitability.* These are variables that fluctuated under the financial crisis and had an effect on credit ratings by itself. By comparing corporate rating observations with similar characteristics, we are removing these effects. The resulting difference between the observations from each period, after the propensity score matching, is therefore easier to attribute to changes in the CRAs' rating practice.

The nature of the analysis is a comparison of the conditions before and after the financial crisis. Due to annual observations of credit ratings and variables, our chance of splitting of the two periods is limited. The pre-crisis period is defined as 2004 through 2007, while the post-crisis is defined as 2008 through 2012. To justify our decision to use 2008 as a starting point of the financial crisis in Europe, we look at both the aggregated GDP for the Euro Area as well as FTSEEurofirst 300 Index, which measures the performance of Europe's 300 larges companies measured by market capitalization (Figure 1).

The GDP growth changes from positive to negative in mid 2008, thus showing signs of recession. The equity index shows signs of weakness already in 2007, but the real collapse strikes in 2008. Based on these observations, we think it is reasonable to start the post-crisis period in 2008.

This paper is structured as follows: Section 2 provides a background and literature review summarizing relevant information and findings in related studies. Section 3 develops our research questions. Section 4 explains how the data set was constructed with descriptive statistics. Section 5 describes the methodology used in this paper and our approaches to the empirical analysis. Section 6 presents our results from the empirical analysis, while Section 7 presents the additional cross sectional analysis. Section 8 provides discussion and limitations regarding our results, while section 9 concludes.

Figure 1: GDP growth and equity index for the Euro area

Note: Figure 1 shows the aggregated GDP growth for the Euro Area (red) as well as the FTSEurofirst 300 index (blue). Data Source: Macrobond, (accessed 8 June 2014).

2 Background and literature review

Credit ratings cover a big field of securities. This study focuses on long-term credit ratings on corporate bonds. Whenever ratings of a different kind are discussed, it will be clarified.

Today there are over 70 credit rating agencies worldwide (DefaultRisk, 2014). They differ in size, where they operate geographically, what kind of instruments they rate and what kind of method they use in their rating process. The Securities and Exchange Commission (SEC) in the U.S. announces a list of those CRAs that are Nationally Recognized Statistical Rating Organizations (NRSROs), which then means that SEC and other financial institutions legally can use their ratings (U.S. Securities and Exchange Commission, 2014). Today this list is compiled by nine CRAs, including the three largest; Standard and Poor's (S&P), Mood's Investor Service (Moody's) and Fitch Ratings (Fitch). 'The big three' were the first on the initial list first made in 1975, and they are still the leading agencies. S&P and Moody's have the largest market shares, and when including Fitch, the three CRAs have a combined market share of about 95%. This has been stable over the last decade (The Economist, 2007) (Alessi *et al.*, 2013).

The established dominant positions of S&P, Moody's and Fitch have led to a credit rating market where rating changes from one CRA may affect the others' ratings. One might look at it as a kind of interaction when one of them changes a rating. Güttler and Wahrenburg (2007) find evidence suggesting there is a pattern when looking at S&P and Moody's ratings. They see that a change from one agency is likely to cause the other to follow. This effect seems larger when the change is a downgrade. Becker and Milbourn (2011) are discussing what happened in the rating market when Fitch took a big chunk of the market in the early 2000s. They saw that ratings in general went up, but the correlation between the ratings and marked- implied yields fell, implying lower overall quality of the ratings. Their conclusion is that there is a negative link between competition and quality. These findings are relevant to this paper because we then know that whenever credit ratings change more frequently, as in a recession, this may trigger more rating changes from other companies that again may lead to further downgrades.

In this paper we are interested in understanding how ratings are designed and assessed. The different CRAs provide the market with extensive guides with criteria they use in their rating methods. However, these all differ. We focus on S&P and Moody's. For S&P they have a criteria guide for all corporate bonds in general. This makes it easy for users to understand their methodology, though the level of details lack. They state in their guide that: "Note that we do not have any predetermined weights for these categories. The significance of specific factors varies from situation to situation." (Standard & Poor's, 2008). S&P divides their assessment of a credit rating into two different measures of risk; (1) Business Risk with sub-categories country risk, industry factors, competitive position, profitability/peer group comparison, and (2) Financial Risk; country risk, industry factors, competitive position, profitability/peer group comparison. They have an extended guide discussing the importance of each of these variables, but, as mentioned, the weighing of each depends solely on the situation.

In the case for Moody's, they provide guides for different industries. This makes the guides more detailed in terms of variables and weighing, but it is not easily accumulated in a single rating model. By looking at a sample of the different industries (global chemical 2013, global surface transportation and logistics 2013, regulated electric and gas utilities, global steel 2012, global retail 2011, global telecommunications 2010, global automobile 2007, global tobacco 2010) we see that the overall measures can be generalized. We see that they use mostly measures of size/scale, profitability, leverage, coverage, financial policies and business profile. Further some more specific industries also measure diversification, franchise strength, growth potential, operating environment and market presence/market share.

One of the objectives of this paper is to determine which variables are explanatory in a rating process based on quantitative methods. We then look at other studies that determine a set of significant variables for model making. Blume *et al.* (1998) and Amato and Furfine (2004) both build rating models looking at the U.S. corporate bond market using S&P ratings. Blume et al find *book debt to assets, net income to assets, market equity, market model beta, interest coverage* and *book assets* as significant variables. Amato et al find *long-term debt, total debt, operating margin, market value, market model beta* and *interest coverage*. Considering natural differences in the data, these two models are quite a like. Pettit *et al.* (2004) also looks at the U.S. market and S&P ratings. They do not use the same method, but they also try to find explanatory variables. They find several variables with statistical significance; *market capitalization, sales, assets, market leverage, debt/EBITDA, EBITDA coverage* and *EBIT coverage*. A more recent study from Hung *et al.* (2013) looking at the U.S. market and S&P ratings shows that variables in all categories for both Business and Financial Risk, mentioned by S&P and Moody's, are significant. As a more conservative view, Koller *et al.* (2010) state that the only two variables explaining the credit ratings are a measure of size (*market capitalization, total sales, total assets* etc.) and a form of interest cover (Koller *et al.*, 2010). Even though these studies are based on the U.S. market we assume the results are representative for Europe as well.

Another aspect we need to consider is whether to use ratings from only one or more CRAs in our analysis. Moon and Stotsky (1993) find that S&P and Moody's use different determinants, and also weigh them differently. They also find that if the rating is self-selection or not is important in the case for Moody's ratings. A more recent study by Krämer and Güttler (2008) propose that ratings from S&P and Moody's contain different information, which suggests they set the ratings differently. While Livingston *et al.* (2010) do not look at the determinants, this study also concludes that the ratings are different to some extent. There is evidence that Moody's ratings are more conservative than S&P, and that investors also assign more reliance to these more conservative ratings. Güttler and Wahrenburg (2007) further state that Moody's are timelier than S&P in their rating changes. These studies conclude that ratings from different CRAs are to some extent unlike. Optimally we would then use ratings from S&P and Moody's, as we might get different results. However, due to accessibility and the width of this paper limit us to only using ratings from one CRA, Moody's.

We want to focus on long-term credit ratings. Compared to short-term ratings, they have a goal of assessing the long-term credit risk. By doing this, the CRAs must focus more on structures that are affected over time rather than measures that can fluctuate over shorter periods. Therefore the CRAs want to obtain stable ratings and not be disturbed by the anatomy of business cycles. 'Rating through' the cycle is a term used in empirical studies on this subject. As an example, this is what S&P has to say about ratings and business cycles: "We do not - and cannot - aim to 'rate through the cycle' entirely. Rating through the cycle requires an ability to predict the cyclical pattern - usually extremely difficult to do." (Standard & Poor's, 2008). So even though the CRAs state they try to rate without dealing with the business cycles, they also acknowledge that this is a very hard thing to do in practice.

Altman and Rijken (2004) tries to explain how the CRAs obtain these stable ratings by breaking it down into two sources of stability; (1) the 'through-the-rating' methodology and (2) the migration policy of the agencies. The study concluded that both sources of stability are contributing equally to the rating stability. In details, they find that the credit quality needs to exceed the level of rating class by 1.25 before any changes are done, and when triggered ratings are only partly adjusted. This tells us a lot about the migration policy. After the dot com-bubble in 2002, Moody's announced that they would try to enhance the timeliness and quality of their ratings: "More aggressive ratings changes - such as downgrading a rating by several notches immediately in reaction to adverse news rather than slowly reducing the rating over a period of time - as well as shortening the rating review cycle to a period of weeks from the current period of several months." (The Financial Times, 2002).

Many studies look at the impact of business cycles on credit ratings. Amato and Furfine (2004) do not find any evidence that credit ratings in general are unduly influenced by the business cycle. However, they find evidence of procyclicality in two special cases; when the sample is limited to only investment grade or only initial ratings. They define procyclicality by having their null-hypothesis: "business cycle variables should not have a marginal effect on the ratings assigned to a firm". This definition is not directly comparable to the one of this thesis, but they examine this questions looking at firms' underlying financial and business characteristics, making the studies like to some degree. Löffler has done several studies on the subject. First he shows that ratings indeed are very stable, but at a cost of a lower default prediction power (Löffler, 2004). In a later study, he finds that there might be a vital information loss when the CRAs are not taking action when a possible change in credit rating likely is reversed shortly (Löffler, 2005). He argues that the market's information preferences are not homogeneous, and therefore that this conclusion is debatable. To further analyze the problem, he tests if the CRAs are able to see through the cycle and rate correctly, as they claim. The results are that they truly have this ability, and that their long-term rating is correct (Löffler, 2006). His findings are not related to the time after the financial crisis, thus his conclusions are not representative for our results. However, he uses a broad set of data from 1980-2005, making his results a good indication of how the CRAs have performed up until his last study.

Other studies that look at the impact of a crisis on ratings are done by looking at national economies and sovereign ratings. This is not the topic in this paper, therefore the results are not discussed in detail. However, because of similarity in methods, we mention a study from the East Asian crisis in the 1990s. Ferri *et al.* (1999) demonstrate that CRAs failed to predict the crisis, and also became excessively conservative and downgraded the affected countries more than their economies' fundamentals would justify afterwards, making the crisis worse. This is an interesting find, but as mentioned this is a case with sovereign ratings, not companies. The method they use is similar to one we are using making a model based on a time series and then making out-of-sample estimations, comparing with actual ratings.

Manso (2013) relates the stability in ratings to feedback effects. He argues that rather than focusing solely on the accuracy of the ratings, the CRAs should also look at the effects of downgrades. Knowing that downgrades from one agency affect ratings from other, Manso states that the result of this might be increasing default frequency. His argument is that the CRAs by changing the rating, also changes the cost of capital for the bond issuer, and therefore this can be seen as a self-fulfilling spiral, where a downgrade leads to higher cost of capital that again can lead to further downgrade and so on. As an alternative point of view, Blume *et al.* (1998) tries to look at stricter rating policy as a reason for lower ratings over time. Using data from 1973-1992, he concludes, with caution, that stricter policy explains the decline in ratings at that given time rather than a decline in the debt quality itself. He further finds that accounting ratios and market based risk measures might be more informative for larger companies.

Another important aspect to consider in this paper is the effect of new regulations. One can divide regulations into two different grounds; (1) regulations on the financial market and the use of credit ratings, and (2) regulations on the CRAs themselves. The latter part needs evaluation because of new regulations that came right after the financial crisis, which may be an important factor explaining a possible tightening in rating standards. The CRAs were first 'self-regulated'. Prior to the crisis all the European CRAs had to do was to send an annual letter to the supervisory authorities, and also have an annual meeting discussing the Code of Conduct, at that time given by the Committee of European Securities Regulators. They only had to follow the IOSCO Code (The Technical

Committee of The International Organization of Securities Commissions, 2004), which is a statement of principles regarding the activities of credit rating agencies. Now measures have been taken to limit the way credit ratings can be assessed. Utzig (2010) discusses how the different regulations were set in place in the U.S. and Europe. The process started after the G-20 summit in November 2008. At that time, many European countries wanted to regulate the CRAs and ensure better quality ratings. Both the European Commission and Parliament were quick to act and made future plans in terms of a stand-alone European regulation, with the goal of later end up with a globally consistent regulation. The initial regulations considered transparency of rating methods, the number of expertise staff at CRAs, regular monitoring and potential conflicts of interests (Utzig, 2010). Further the IOSCO Code was set as a minimum standard. The European Commission provides a time line of regulations, discussions and working papers regarding this (European Commision, 2014). The first regulations regarding CRAs' technical standards came out in May 2012: "The regulatory technical standards will ensure a level of playing field, transparency and adequate protection of investors across the Union and contribute to the creation of a single rulebook for financial services" (European Commision, 2014). Measures to ensure certain standards and transparency are also taken in the U.S., but those will not be discussed.

There has been massive public critique after different economical crisis, complaining that the CRAs do not generate qualified ratings. In hindsight of the financial crisis, the focus has been on corporate credit ratings in the U.S. and sovereign ratings in Europe. Very little research has been made on the accuracy of credit ratings related to European corporates. Looking back, Partnoy (2002) highlights many examples of failure to predict defaults from the 1970s and onwards. His claim is that credit ratings are a paradox - agencies possess so much power, but they do not predict defaults, as they should. He further states that the CRAs exist mostly because of numerous legal rules and regulations, and that regulatory dependence then explains this paradox. When comparing the performance of the two leading agencies, Krämer and Güttler (2008) find evidence that ratings from one company contain other information than from the other, concluding that none of the agencies uniformly outperform the other. A common conclusion is that CRAs' task of assessing credit risk is complex and difficult when you cannot predict cyclical patterns.

3 Research question development

In this paper, we aim to analyze long-term credit ratings for European corporates in the time just before, during and just after the recent financial crisis starting in 2008. We investigate which explanatory factors are included in the rating process, and check the dynamics of these ratings through the financial crisis and see if ratings standards have changed.

The CRAs were subject to intense critique during and after the financial crisis. The criticism involved the CRAs' rating practice in which one claimed that rating agencies were too relaxed in the rating process and initially gave too high ratings, especially on financial instruments such as CDOs and sovereign ratings. The risk of these financial instruments was allegedly underestimated by the rating agencies either because of a lack of competition, poor accountability, or most likely the complexity of the instruments resulting in difficulty in assessing the risk (Taylor, 2009). One observation is that corporate credit ratings are rarely mentioned in this context, and when they are, it is usually in situations where structured financial instruments have major impact on the companies involved. The most common example is the major global investment banks that got into problems due to large amounts of investment-rated, structured financial instruments in their balance sheets with potential high risk. The investment banks had solid credit ratings, along with their structured products, but history show that those were not accurate.

CDOs and similar structured finance instruments are beyond the scope of this paper and we focus on corporate credit ratings. We focus on European companies. The focus in Europe the recent years has to a greater extent been on the sovereign ratings and the leverage in European countries. This limited research on the corporate ratings leaves us an impeccable opportunity to explore something that until know is fairly unknown.

The credit rating process is complex and includes quantitative as well as qualitative factors. The process can be seen as the CRAs' own business secrets, although regulations demand that there is transparency in the rating decisions. Our focus is on the quantitative variables due to data accessibility and objectivity, since there is much more consistent way to obtain fundamental corporate ratios rather than more subjective, qualitative variables. By quantitative variables we mean corporate fundamental values and macro economic indicators.

More specifically we want to investigate the factors that affect a credit rating in order to gain understanding of how and why credit rating changes. This is also important for later analysis, to be able to construct prediction models to examine rating dynamics. Thus, our first research question is:

Which quantitative variables are statistically significant in a corporate credit rating process?

The previous question lead to this paper's main focus, which is whether the CRAs in general tightened their standards in credit ratings after the crisis. We can see if lower credit ratings are justified by lower corporate fundamentals, or if there is a mis-match. If they actually downgraded corporates excessively, we may have a tendency of procyclical behavior, which we define as significant underrating after the crisis.

A general overrating during economic expansions can be argued to enhance economic booms with higher levels of risk taking, potentially accelerate into recession. Especially when taking into account the CRAs strong position within the bond market and the power of influence they possess.

A general overestimation of credit ratings in expansion periods will, all else equal, lead to a decrease in company's cost of capital, which will result in more accessible funding and hence opportunity for greater investment and growth for the companies.

The same rationale applies for underratings in recessions. If the CRAs systematically underrate corporates more than what fundamentals justify, we could argue that corporate funding costs rise excessively during the financial crisis, leading to less accessible funding and thus a larger reduction in investment activity. Our second and main question is therefore:

Has there been a tightening in corporate rating practice during and after the financial crisis?

4 Data description

In this section, the collection processes and data set used are described. The data set used in this study is the result of combining two main sources; long term credit ratings from Moody's and financial ratios from the Orbis database.

4.1 Ratings

Our data set consists of historical annual long-term ratings from 2004 through 2012. In order to find a complete list of corporate ratings in our defined period, we contacted both Moody's and S&P directly. S&P's policy is not to give out their ratings for academic purposes. We managed, however, to obtain an academic subscription from Moody's.

One obstacle was to find complete historical time series. Moody's responded that they did not have historical time series data and that their only solution was to put an analyst to create the data set, which we later could buy.

Our response was to manually collect each observation, by going into each company profile, retrieving historical ratings for each year for the total sample of 1226 corporates. We established a set of rules we applied when collecting all the rating-observations. These were:

- Not register the specific date the rating is assigned/upgraded/downgraded. Only on an annual basis.
- If there exist two ratings for the same year, due to change in ratings, the most recent rating will apply, i.e. the current rating 31/Dec each year.

Because all ratings are manually retrieved, some errors may occur. We have however tried to reduce this risk to a minimum by going through all ratings, verifying that the ratings are consistent with the database.

The ratings exported are the long-term ratings, which are defined by Moody's as:
"Ratings assigned to issuers or obligations with an original maturity of one year or more and reflect both the likelihood of a default on contractually payments and the expected financial loss suffered in the event of default" (Moody's Investor Service, 2013).

We have limited the observations to all corporate companies in the Eurozone. Due to a limited sample size, we also included the UK due to UK's tight relationship with the European market and the number of companies registered in the UK. Knowing that the UK had a negative GDP growth during the crisis and their sovereign rating was recently downgraded in 2013, makes us confident that this country can be compared to the rest of the Eurozone (Moody's Investors Service, 2013).

All financial institutions such as banks and insurance companies are excluded from this data set, because of their complex capital-structure that makes financial structural ratios more difficult to use. Furthermore there has been implementation of the Basel Accord II and preparations to implement Basel III in the time period we look at, making a time comparison harder to analyze alongside with all other industries (Basel Committee on Banking Supervision, 2013).

In order to decrease the possibility of survivorship bias, we have also included all withdrawn ratings from the same period.

4.2 Financial ratios

All financial ratios are exported from the Orbis database. The database contains information on both listed and unlisted companies and includes, among others, company financials, financial strength indicators, stock data and industry research.

One challenge with the actual export was to identify the same company in the Orbis database as registered at Moody's, since Moody's only register the rated companies with their own company identification number. This was resolved by running an open Internet search to check the company's history, possible acquisitions, name changes etc.

The ratios selected are based on the credit ratings criteria from both S&P and Moody's. S&P separate the risk factors in two main categories; Business Risk and Financial Risk. The relationship between the two risk types is illustrated in Table 1.

Table 1: Business risk / Financial risk

	Financial Risk Profile				
Business Risk Profile	**Minimal**	**Modest**	**Intermediate**	**Aggressive**	**Highly Leveraged**
Excellent	AAA	AA	A	BBB	BB
Strong	AA	A	A-	BBB-	BB-
Satisfactory	A	BBB+	BBB	BB+	B+
Weak	BBB	BBB-	BB+	BB-	B
Vulnerable	BB	B+	B+	B	B-

Note: Table 1 shows the relationship of business risk and financial risk in relation to credit rating. S&P do not use the lowest ratings in the matrix. This is because those ratings always reflect some impending crisis or extraordinary vulnerability. Data source: (Standard & Poor's, 2008)

As we base this paper on quantitative data, we do not have a particular focus on Business Risk, which implies more qualitative characteristics such as industry factors and competitive position. The only exception is business profitability. The Financial Risk however is easier to quantify, thus constituting a larger focus in our analysis. S&P divides Financial Risk into several sub-categories; Liquidity/Short-term factors, Capital structure, Cash Flow adequacy, Accounting and Financial policies/Risk tolerance and Governance. Pursuant to S&P, the single most critical aspect of credit rating decisions is cash flow analysis. This is because interests and principals are not serviced from earnings, but from cash. This is even more important for speculative-graded companies due to the lack of flexibility to access external financing.

Another important part of S&P's financial review is the company's capital structure. An analysis of the capital structure is used to establish the company's financial flexibility and how leveraged it is. (Standard & Poor's, 2008). Liquidity and short-term factors are also listed as a category due to the influence on the financial flexibility of the firm.

In contrast to S&P's general criteria for all industries, Moody's only operates with specific ratings criteria for each industry. In general, in line with S&P, they divide the risk measures into several 'rating factors' . These factors vary with industry, but we can observe certain similarities in the segmentation of risk that recur in all industry criteria. Scale recur in all industries, often as total sales, and with somewhat different annotations. This also applies to *leverage* and *coverage* that appear under names such as 'Financial Strength', 'Cash Flow and Debt Service', and 'Financial Strength, Key Financial Metrics'. Profitability is also pervasive in all industries. Like S&P, Moody's also operates with Business Profile and Financial Policy, which are largely qualitative in their characteristics.

To summarize, we see from both S&P's- and Moody's rating criteria that they to a certain degree segmenting the quantified risk relatively equal, focusing on both return on capital- and profitability measures, as well as more structural measures (such as leverage), and short term measures (such as liquidity and coverage). Based on these rating criteria we have selected general ratios that are easy to compare over time, industries and countries, segmented into the following categories: Profitability and Efficiency, Operational, Structure and Size.

Note that both CRAs also emphasize cash flow adequacy as an important factor. We have included some variables that include cash flow (*CF/Operating revenue, EBITDA margin*, all within the Efficiency category), but optimally we should have more related to this factor. The reason for the lack of such ratios is the data-coverage in Orbis, where they do not have any other ratios covering cash flow. We could have exported such ratios from other sources, but decided not to, due to uncertainty factors related to incorporating more than one accounting database (with individual sets of accounting standards etc.) in to the data set.

Another important aspect is that both S&P and Moody's have a much more thorough approach in these rating processes, than looking at general ratios. Because of the requirement for objectivity and general ratios, in order to achieve consistency in our analysis, our paper faces natural limitations on this point. Our analysis is only scratching the surface of a comprehensive rating process, but nevertheless illustrates many of the main factors.

In addition to corporate fundamentals, we have included some of the most used macroeconomic indicators from each country together with sovereign ratings. All macroeconomic data is exported from The World Bank (2014). Macroeconomic variables have been used in previous studies as time trend variables or to measure the business cycle. These measures are first of all used to assess sovereign ratings, but we want to include them to see the statistical fit for corporate ratings. Country risk and environment are two factors mentioned by S&P and Moody's. They further state that this is not directly comparable to the variable sovereign rating, but maybe some of these macroeconomic indicators can help explaining this type of risk.

See table 2 for statistical description of the selected variables. The total list of ratios and macroeconomic indicators, together with definitions, are summarized in Appendix A.

4.3 Data refinements

The ratings exported are from Moody's, based on Moody's own 'Long term Rating Scale'. In order to perform analysis, all ratings have been assigned to a numerical rating ranging from 1-21, starting with Aaa assigned as 1. Note that we are assigning numerical ratings to all sub ranks as well, such as Aa1 = 2, and Aa2 =3 in order to achieve a more detailed analysis. This way of assigning numerical values has been used in all empirical studies discussed in the literature review. See table 3 for a complete ranking chart.

An important observation is that relatively few companies are listed. Out of 1226 initial companies, 883 are unlisted. This causes relative few observations for all ratios involving market cap. This restricts the observations to a larger extent than we find satisfactory for our analysis. Thus, we ignore market cap and all related ratios in our analysis. As substitutes, we measure firm size in two ways; by total sales and by total assets. We have altered the scale for both *sales* and *assets* by including both variables as natural logarithms. By doing this, we achieve a more even distribution across companies that are easier to include in the analysis.

The *interest cover ratio* range from -95.6 to 839.9 in our data set. A negative interest cover implies that the company does not have enough operating revenue to pay its interests. Such a situation is not sustainable in the long-term, and therefore implies that something has to change. Blume *et al.* (1998) suggests replacing negative interest coverage ratios with the value of zero. On the other side of the scale, at high interest coverage ratios, it is natural to assume that the difference in the ability to fulfill debt obligations is marginal. Blume *et al.* (1998) thus set any interest rate coverage greater than 10.0 to the value of 10.0 based on an arbitrary choice. However, in our subsequent regression analysis we found that there were positive changes in results from setting a maximum interest coverage ratio equal to 15 compared to 10, related to coefficients and explanatory power of the regressions. The interest coverage ratio used in this data set is therefore set with a maximum ratio of 15. We also see the use of limits on interest coverage in Amato and Furfine (2004) where the same arguments apply. They divide the variable into four different scales; [0,5], [5,10], [10,20] and [20,100], where the lower and upper

Table 2: Variables before and after crisis

	Variable	Prior to the financial crisis (2004-2007) Mean	st.dev	During/after the financial crisis (2008-2012) Mean	st.dev	Difference of mean
Profitability	ROE using P/L b. tax, %	22.05	78.41	14.24	67.69	7.81**
	ROCE using P/L b. tax, %	18.55	57.01	11.43	47.79	7.13**
	ROA using P/L b. tax, %	3.9	12.45	2.58	10.7	1.32**
	ROE using net income, %	16.55	70.77	8.67	58.33	7.88**
	ROCE using net income, %	15.32	57.14	9.68	45.03	5.64**
	ROA using net income, %	2.92	11.87	1.64	10.66	1.28**
Efficiency	Profit Margin, %	9.66	18.8	7.79	21.11	1.87**
	Gross Margin, %	43.07	22.25	41.61	25.56	1.47
	EBITDA margin, %	17.89	19.34	17.82	21.32	0.06
	EBIT margin, %	11.66	21.11	12.18	27.5	-0.52
	CF / Operating revenue, %	13.67	16.75	12.5	17.64	1.17*
Operational	Net assets turnover, x	2.84	22.06	1.96	11.24	0.88*
	Interest cover, x	7.52	42.9	4.29	20.61	3.24**
	Stock turnover, x	41.12	95	47.43	104.88	-6.32*
	Collection period, days	68.77	90.22	58.53	87.87	10.24**
	Credit period, days	46.17	52.31	44.29	60.38	1.88
Structure	Current ratio, x	4.53	12.49	3.93	10.65	0.60*
	Liquidity ratio, x	4.3	12.39	3.77	10.68	0.53*
	Shareholders liquidity r, x	8.02	54.87	9.34	64.97	-1.32
	Solvency (Asset based), %	29.72	30.5	30.37	33.53	-0.65
	Solvency (Liability based), %	32.14	27.42	31.68	27.69	0.46
	Gearing, %	181.22	183.63	178.55	190.37	2.66
Size	Sales, mil EUR	1470.17	928.33	1475.3	928.6	-5.13
	Total Assets, mil EUR	3286.95	1947.11	3401.39	1953.12	-114.44**
Macro	Unemployment, %	6.12	2.38	7.52	4.29	-1.40**
	Inflation, %	2.02	0.7	2.29	1.3	-0.27**
	GDP growth, %	3.22	1.23	-0.29	2.54	3.51**
	Current Balance, %	1.88	6.2	1.91	5.28	-0.02
	Sovereign Rating	1.17	0.71	1.64	2.21	-0.47**

Note: Table 2 Provides descriptive statistics for company rating before the financial crisis (2004-2007) versus after and during the financial crisis (2008-2012). The difference in means are analyzed by a two-sided t-test. Symbols * and ** denote significance levels of 5% and 1% respectively. Data source: Bureau Van Dijk Orbis Database (accessed 30 March 2014) and World Bank World Development Indicators (accessed 7 April 2014).

Table 3: Convertion table for numerical ratings

Global Long Term Rating Scale			
Aaa	Aaa	1	Investment Grade
Aa	Aa1	2	
	Aa2	3	
	Aa3	4	
A	A1	5	
	A2	6	
	A3	7	
Baa	Baa1	8	
	Baa2	9	
	Baa3	10	
Ba	Ba1	11	Speculative Grade
	Ba2	12	
	Ba3	13	
B	B1	14	
	B2	15	
	B3	16	
Caa	Caa1	17	
	Caa2	18	
	Caa3	19	
Ca	Ca	20	
C	C	21	

Note: Moody's add numerical modifiers 1,2 and 3 to each generic rating class from Aa including Caa. The modifier 1 indicates that the obligation ranks in the higher end of its generic rating class; 2 indicates a mid-range ranking, while 3 indicates a ranking in the lower end of that generic rating class. (Moody's Investor Service, 2013) values are assigned to 0 and 100. However, this method is more complicated and we continue working with the method used by Blume *et al.* (1998).

As we conducted our analysis, we tried to refine the data set by winsorizing the variables to eliminate the chance of spurious outliers creating complications. We find, however, that we do not generate any other results by this form of censoring. Therefore we proceeded without refining the possible outliers in any way.

4.4 Final dataset

The final data set consists of 1226 companies, constituting 4270 firm-year observations from 28 different industries. See Figure 3 for a detailed overview. Our sample includes all firms in the Eurozone including UK and spanning the entire ratings spectrum, including both investment and speculative grade firms. The vast majority of the companies in our sample are however in the rating categories of A to B, as seen in the annual rating distribution in Figure 4. This may suggest that the range from A to B is an effective rating level for most companies. In terms of leverage, Koller *et al.* (2010) states that few companies have very high ratings, since a very low leverage would cost much in terms of tax savings and management discipline. The same argument can be used for the lower part of the rating scale, where high leverage will lead to very high interest ratios and thus a more costly and less accessible funding.

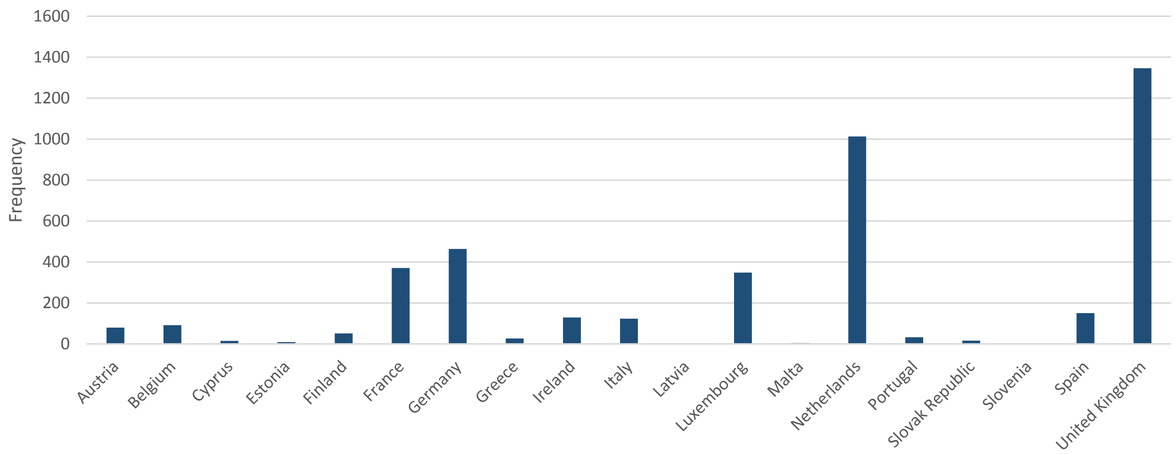
All financial ratios used and their annual mean values are listed in Table 2. During and after the financial crisis (Table 2), the general tendency was a worsening in corporate fundamentals, with profitability ratios and macroeconomic variables as the most significant. When analyzing the ratings later in this paper, we therefore have to consider this observation, especially in the question about whether the CRAs downgraded excessively during and after the crisis.

As mentioned earlier, the ratios are grouped in the following categories; Profitability, Efficiency, Operational, Structure, Size and Macro.

5 Methodology

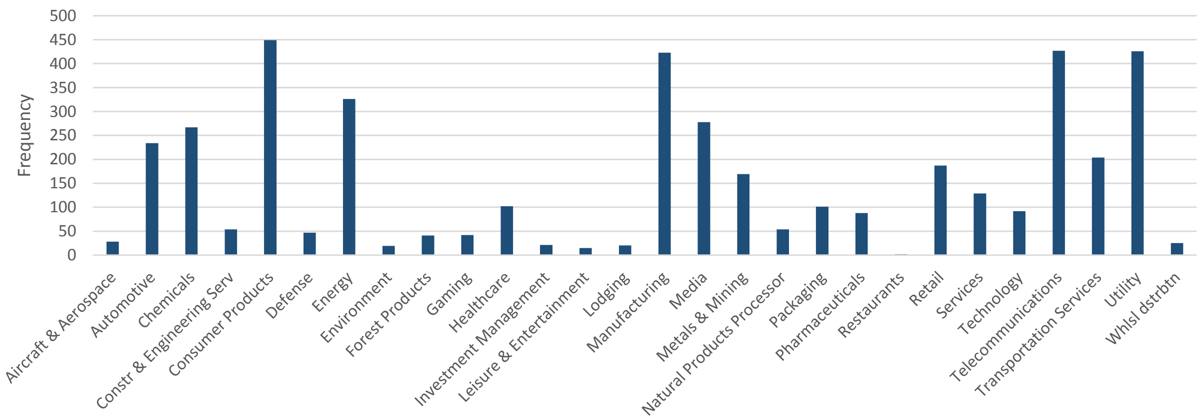
Different methods of analysis are used in this thesis. We perform numerous regression analyses to identify the most important quantitative ratios in the credit rating process. Matching propensity score is utilized in order to compare long-term credit ratings be-

Figure 2: Number of long term rating observations for each country



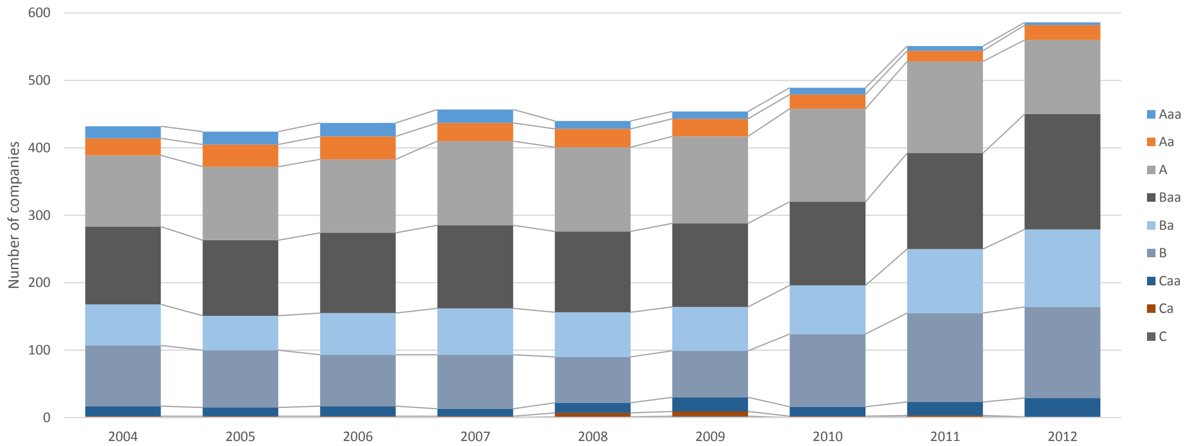
Note: Figure 2 provides a graphical illustration of number of rating observations within each country. Data Source: Moody's Investor Service (accessed 16 March 2014).

Figure 3: Number of long term rating observations within each industry



Note: Figure 3 provides a graphical illustration of number of rating observations within each industry. Data Source: Moody's Investor Service (accessed 16 March 2014).

Figure 4: Number of companies within each rating class each year



Note: Figure 4 provides a graphical illustration of number of rating observations within each rating class for each year. Data Source: Moody's Investor Service (accessed 16 March 2014).

fore and after the crisis, and in that way investigate whether the credit rating agencies tightened their credit rating policy in the years following the financial crisis.

In this part, we are giving an account for the methods used and the most central assumptions and conditions associated with each method.

5.1 Regression models

As mentioned in the introduction, we use regression models to determine which variables are explanatory for the credit ratings using the data from 2004 to 2007. We build two types of models; benchmark models and an empirical model.

Benchmark models
We will build two benchmark models using ordinary least square (OLS) regression derived from the CRAs' own criteria.

The first model will be an unrestricted model, which includes variables from all categories mentioned by the CRAs in their criteria guides. The models will be built by using variables from all categories from our variable list, which all are relevant looking at the criteria guides. We can further break down these categories by using the clusters of correlating variables shown in the correlation matrix in Appendix B. This model will from this point be referred to as *OLS unrestricted*.

The second model will be a restricted model, which is a statistical restriction of the first model. This model will only include those variables from the unrestricted model, which show up as significant from our sample data set, at a 5% significance level. This model will be referred to as *OLS restricted*.

Empirical model
The empirical model is build by taking into account all those variables shown as significant in previous studies. Further this model will be a panel data regression, which uses other properties of the data set. This model will be derived in the same manner as the OLS, only with the empirically shown statistical variables mentioned in the literature review. This model will be referred to as *Panel data regression FE*.

We will then use these models to make an out-of-sample estimation, comparing the estimations to the actual ratings from 2008 - 2012 to see if the models over- or under predicts the ratings.

Panel data regression

Panel data models examine fixed and/or random effects of the entity (Park, 2005). The core difference between fixed and random effects models lies in the role of the dummy variables. If dummies are considered a part of the intercept, this is a fixed effects model. In a random effects model, the dummies act as an error term. Statistically, fixed effects always give consistent results, but they might not be the most efficient model to run (Princeton Data and Statistical Services, 2014). Random effects will give higher p-values, as they are more efficient estimators.

The generally accepted way of determining whether you are dealing with random of fixed effects, is to run a Hausman-test (Hauser, 2013). This tests the more efficient random effects model to a less efficient but consistent fixed effects model to check if the more efficient model also gives consistent results. The null hypothesis is that the coefficients estimated by both models are the same. An insignificant p-value suggests that it is safe to use random effects.

From the Hausman-test in our case, we get a low and significant p-value, suggesting we should continue using fixed effects.

5.2 Prediction

For the prediction we will use the three different models and estimate ratings out-of-sample for 2008-2012. The predicted numerical ratings are converted to the alphabetical rating scale according to Moody's rating chart (Table 3 - Ranking chart for Moody's ratings). The limit of each rating category is set to the numerical values in the table (i.e. the category Aa includes all estimated ratings equal or greater than 2 and less than 5).

These estimations will then be compared to actual ratings. We will not be able to look at the comparison of each company because of the vast number of observations, but we will accumulate both the estimated and the actual ratings and illustrate the results in a table for each model.

The goal by doing this out-of-sample prediction is not to check the performance of the different models. The comparison will function as an indication on whether the CRAs have a stricter (less strict) rating policy after the crisis if the models over predict (under predict).

5.3 Propensity score matching

When analyzing whether the credit rating agencies actually tightened their credit rating policies, we investigate whether the credit ratings have changed significantly from pre-crisis (2004-2007) to post-crisis (2008-2012). In this case, our results may be attributed to changes in corporate fundamental values, rather than rating policy decisions.

We use propensity score matching (hereby PSM) to adjust for such fundamental fluctuations, and check whether a general corporate downgrading after the financial crisis can be attributed to tightening of corporate credit rating policies by the CRAs. Propensity Score matching were first introduced by Rosenbaum and Rubin (1983) as a method to match two groups based on different characteristics. This type of matching method tries to regenerate the characteristics to a 'treated' group by constructing a control group that fits the characteristics of the treatment group as much as possible. In this way we can find an estimate for the counter factual difference, and thus calculate the mean difference between the control- and treated group. In our analysis; we measure the mean difference of long term ratings before (control group = 'untreated') and after the financial crisis (treatment group), matching on corporate fundamental characteristics. Compared to matching directly on covariates, PSM reduce the dimensionality of matching to a single dimension, the propensity score (Caliendo and Kopeinig, 2008).

5.3.1 Propensity score estimation

The propensity score estimation in this analysis includes a binary treatment variable, before and after crisis previously defined as respectively 'untreated' (treatment=0) and 'treated' (treatment=1). Propensity scores are estimated using a Logit regression model with treatment as the dependent variable.

The explanatory variables are selected based on company characteristics rather than focusing on estimating the true propensity score as accurately as possible. The most important aspect is that the variables contribute to estimates of the propensity score that can match covariates between treated and control subgroups. This is more important than that the model estimates the true propensity score as accurately as possible (Grilli and Rampichini, 2011).

5.3.2 Propensity score matching

There are several matching methods, in this thesis we are testing four of them: Nearest-Neighbor (NN) matching, Kernel matching, Radius matching and Mahalanobis-metric matching.

6 Empirical analysis and results

6.1 Correlation

As an introduction to the results, we simply illustrate the coefficients from the correlation between the ratings and the given set of variables. By splitting the data we can see which financial variables that are statistically correlating with the rating process in the two different periods.

What we see is an increasing number of significant correlations from pre-crisis to post-crisis period. (Table 4).

The results are that, on average, financial variables and therefore quantitative methods have become a more reliable tool when giving a company a credit rating.

6.2 Regression models

As mentioned in the introduction and method, we use regression models to determine which variables are explanatory for the credit ratings in order to answer our first research question in this thesis - Which quantitative variables are statistically significant in a corporate credit rating process? First we build two multiple ordinary least squares

Table 4: Variable correlation with long-term rating

	Variable	Correlation 2004-2007	Correlation 2008-2012
Profitability	ROE using P/L b. tax, %	-0.05	-0.13*
	ROCE using P/L b. tax, %	-0.08*	-0.12*
	ROA using P/L b. tax, %	-0.16*	-0.28*
	ROE using net income, %	-0.04	-0.16*
	ROCE using net income, %	-0.05	-0.12*
	ROA using net income, %	-0.13*	-0.24*
Efficiency	Profit Margin, %	-0.19*	-0.29*
	Gross Margin, %	-0.04	-0.09*
	EBITDA margin, %	-0.07	-0.02
	EBIT margin, %	-0.07*	-0.05
	CF / Operating revenue, %	-0.15*	-0.16*
Operational	Net assets turnover, x	-0.08*	-0.01
	Interest cover, x	-0.26*	-0.30*
	Stock turnover, x	-0.00	0.01
	Collection period, days	-0.03	0.04
	Credit period, days	0.02	0.07*
Structure	Current ratio, x	0.02	0.02
	Liquidity ratio, x	0.02	0.01
	Shareholders liquidity r, x	-0.06*	0.02
	Solvency r (Asset based), %	-0.08*	-0.11*
	Solvency r (Liability based), %	-0.00	0.00
	Gearing, %	0.10*	0.16*
Size	logSales, mil EUR	-0.26*	-0.35*
	logTotal Assets, mil EUR	-0.28*	-0.33*
Macro	Current Balance, %	0.11*	0.07*
	GDP growth, %	0.09*	0.01
	Inflation, %	0.00	-0.00
	Unemployment, %	-0.06*	0.01
	Sovereign Rating	-0.02	0.11*

Note: Table 4 displays coefficients for pairwise correlation between variable and rating for company ratings before the financial crisis (2004-2007) versus after and during the financial crisis (2008-2012). Symbol * denote significance level of 5%. Data source: Bureau Van Dijk Orbis Database (accessed 30 March 2014) and World Bank World Development Indicators (accessed 7 April 2014).

benchmark models based on S&P's and Moody's own criteria. Then we build a panel data regression model that utilizes more of the information of the data set. Both types of models are built on data from 2004-2007.

The models are used to make an out-of-sample prediction for the ratings, comparing the estimated ratings to the actual ratings from 2008-2012. The results of the comparison will indicate whether the ratings are set justified to the financials of the companies after the crisis, and therefore help us answer our second research question - Has there been a tightening in corporate rating practice during and after the financial crisis?

6.2.1 Significant ratios

OLS-unrestricted

The results from the OLS unrestricted are presented in Table 5. We see that four of the variables show as insignificant.

OLS-restricted

For the OLS restricted, the model is refined into a statistical accepted model, see table Table 6. The list of variables has been shortened down to five significant determinants, which is far less than what the CRAs list as their criteria.

Panel data, fixed effects

The results from the panel data regression FE can be seen in Table 7. We see that this model to some extent look like the OLS restricted.

The overall results from the models tell us that there are few variables that show as significant. Even if the interpretations of the coefficients within a model does not automatically say which variables that are most explanatory, the vast difference between sales and interest cover compared to the rest of the variables gives an indication that these two variables are heavily weighted in the rating process. This result gives the same conclusion as the variable correlation as they are the top two variables in the statistical correlation between variables and ratings in Table 4. These two variables are also the same as Koller *et al.* (2010) argue are the most important variables. All the significant variables from the models will all be discussed under discussions and limitations. Further

Table 5: OLS regression unrestricted. 2004-2007

LTRating	**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
ROCEusingPLbeforetax	-0.030	0.005	-6.56	0.000	-0.039	-0.021
ProfitMargin	-0.020	0.011	-1.80	0.072	-0.042	0.002
Interestcover15	-0.188	0.041	-4.58	0.000	-0.268	-0.107
CreditPeriod	-0.009	0.003	-3.11	0.002	-0.015	-0.003
Collectionperiod	0.001	0.002	0.44	0.660	-0.003	0.005
Liquidityratio	-0.036	0.039	-0.92	0.356	-0.113	0.041
SolvencyratioAsset	-0.019	0.008	-2.41	0.016	-0.035	-0.003
logSales	-1.820	0.170	-10.71	0.000	-2.154	-1.486
SovereignRating	0.115	0.188	0.61	0.540	-0.254	0.485
_cons	17.812	0.761	23.39	0.000	16.316	19.307
$Prob > F$	0					
R-squared	0.35					
Adj R-squared	0.34					
Number of obs	560					

Note: Table 5 provides descriptive statistics from OLS regression based on pre-crisis observations (2004-2007).
ROCE using PL before tax:(Profit before tax + Interest paid) / (Shareholders funds + Non current liabilities)*100
Profit margin: (Profit before tax / Operating revenue)*100
Interest cover: Operating profit / Interest paid
Credit period: (Creditors / Operating revenue)*360
Collection period: (Debtors / Operating revenue)*360
Liquidity ratio: (Current assets Stocks) / Current liabilities
Solvency Asset: (Shareholders funds / Total assets)*100
logSales: Mil EUR
Data source: Bureau Van Dijk - Orbis Database, World Bank - World Development Indicators and Moody's Investor Service.

Table 6: OLS regression restricted. 2004-2007

LTRating	**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
ROCEusingPLbeforetax	-0.030	0.005	-6.51	0.00	-0.039	-0.021
ProfitMargin	-0.021	0.011	-1.97	0.05	-0.042	-0.000
Interestcover15	-0.180	0.040	-4.43	0.00	-0.259	-0.100
SolvencyratioAsset	-0.016	0.008	-2.03	0.04	-0.031	-0.000
logSales	-1.796	0.162	-11.06	0.00	-2.114	-1.477
_cons	17.313	0.632	27.39	0.00	16.072	18.555
$Prob > F$	0					
R-squared	0.34					
Adj R-squared	0.33					
Number of obs	563					

Note: Table 6 provides descriptive statistics from OLS regression restricted based on pre-crisis observations (2004-2007).
ROCE using PL before tax:(Profit before tax + Interest paid) / (Shareholders funds + Non current liabilities)*100
Profit margin: (Profit before tax / Operating revenue)*100
Interest cover: Operating profit / Interest paid
Solvency Asset: (Shareholders funds / Total assets)*100
logSales: Mil EUR
Data source: Bureau Van Dijk - Orbis Database, World Bank - World Development Indicators and Moody's Investor Service.

Table 7: Panel data regression, fixed effects. 2004-2007

LTRating		**Coef.**	**Std. Err.**	**t**	**P>t**	**[95% Conf.**	**Interval]**
Interestcover15		-0.059	0.017	-3.54	0.00	-0.091	-0.026
logSales		-0.704	0.357	-1.97	0.05	-1.406	-0.003
Creditperiod		-0.005	0.002	-2.95	0.00	-0.008	-0.002
SolvencyratioAsset		-0.016	0.005	-3.44	0.00	-0.025	-0.007
_cons		12.233	1.249	9.79	0.00	9.777	14.689
$Prob > F$	0						
R-squared	within: 0.1183						
	between: 0.1661						
	overall: 0.1308						
Number of obs	584						
Number of groups	193						

Note: Table 7 provides descriptive statistics from panel data regression with fixed effects based on pre-crisis observations (2004-2007).
Interest cover: Operating profit / Interest paid
logSales: Mil EUR
Credit period: (Creditors / Operating revenue)*360
Solvency Asset: (Shareholders funds / Total assets)*100
Data source: Bureau Van Dijk - Orbis Database, World Bank - World Development Indicators and Moody's Investor Service.

the R^2 is moderate in the benchmark models, but low in the empirical model, suggesting this quantitative method alone is not a good measure for ratings.

6.2.2 Out-of-sample test

The estimations for the three different models are presented in Table 8 as Panel A, B and C. The left column states the actual rating, while the matrix shows how the model performed. The diagonal path going from upper left to lower right represents a match of actual and estimated rating, while over (under) estimations are shown to the left (right) for the diagonal.

The results are to some degree unambiguous. There is a pattern for over-predicting ratings in all three models, which is supported by the paired mean-comparison tests ranging from 0.6 to 0.8 for the three models. We notice that for the higher ratings, the models under-predicts, but in total the models tend to over-predict. By looking at the prediction success rate for each rating category we observe that all models tends to predict better within mid-range rating categories. This coincide with the data description where the vast majority of the observations in our sample are in the rating categories of A to B, as seen in the annual rating distribution in Figure 4.

This result of over predictions is a small indication that the CRAs actually rated corporates stricter after the crisis. Based on these quantitative models alone we can say that companies need better financials, all others a like, to receive the same rating after the crisis compared to before.

6.3 Propensity scores

Propensity scores were estimated by a probit regression including the crisis dummy as a dependent variable (0 = company ratings before crisis, 2004-2007 and 1 = company ratings during crisis, 2008-2012).

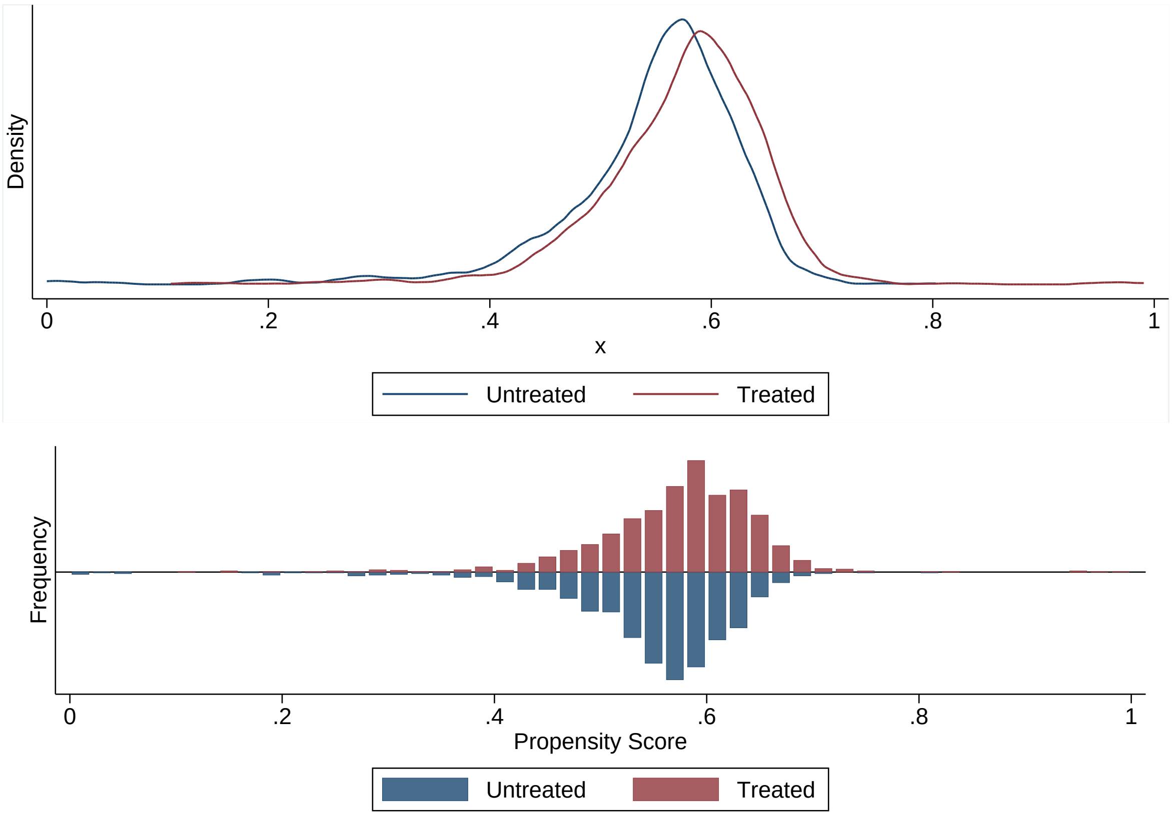
The ratios we ended up with in our final propensity score probit model are: *ROCE using net income, interest cover, solvency ratio asset based, logsales* and *gross margin.* As we see, the nature of the ratios coincide to a large degree with the ratios used in the OLS restricted and are reflecting company characteristics within both profitability, efficiency, structure and operation. We have chosen ratios that we know affect credit rating based on rating methodology, and the estimated regressions, hence reduced the emphasis on finding a model that estimate the true propensity score as accurately as possible.

Table 8: Actual ratings vs predicted ratings

Actual rating	Predicted Rating Aaa	Aa	A	Baa	Ba	B	Caa	Ca,C	Total	% Correct prediction
Panel A: OLS unrestricted										
Aaa	0	4	1	9	6	0	0	0	20	0.0 %
Aa	0	4	15	10	1	0	0	0	30	13.3 %
A	0	11	134	88	7	0	2	0	242	55.4 %
Baa	0	8	106	160	3	0	0	0	277	57.8 %
Ba	0	0	16	152	20	4	0	0	192	10.4 %
B	0	0	1	56	48	13	1	0	119	10.9 %
Caa	0	0	0	7	11	5	1	1	25	4.0 %
Ca and C	0	0	0	0	0	1	0	0	1	0.0 %
Total	0	27	273	482	96	23	4	1	906	
Over-estimation frequency		568								
Under estimation frequency		338								
Difference of mean		0.62**								
Panel B: OLS restricted										
Aaa	1	3	1	10	5	0	0	0	20	5.0 %
Aa	0	4	14	12	0	0	0	0	30	13.3 %
A	0	14	130	89	7	2	0	0	242	53.7 %
Baa	0	3	113	157	4	0	0	0	277	56.7 %
Ba	0	0	21	157	10	4	0	0	192	5.2 %
B	0	0	1	63	48	5	3	0	120	4.2 %
Caa	0	0	0	8	12	5	1	1	27	3.7 %
Ca and C	0	0	0	0	0	1	0	0	1	0.0 %
Total	1	24	280	496	86	17	4	1	909	
Over-estimation frequency		579								
Under-estimation frequency		330								
Difference of mean		0.70**								
Panel C: Panel data FE regression										
Aaa	0	0	0	21	0	0	0	0	21	0
Aa	0	0	11	20	3	0	0	0	34	0.0 %
A	0	0	90	157	12	0	0	0	259	34.7 %
Baa	0	0	50	230	1	0	0	0	281	81.9 %
Ba	0	0	7	184	5	0	0	0	196	2.6 %
B	0	0	0	123	5	0	0	0	128	0.0 %
Caa	0	0	0	24	7	0	0	0	31	0.0 %
Ca and C	0	0	0	0	4	0	0	0	4	0.0 %
Total	0	0	158	759	37	0	0	0	954	
Over-estimation frequency		558								
Under-estimation frequency		396								
Difference of mean		0.83**								

Note: Table 8 provides descriptive statistics of the regression models prediction power. The output in each panel shows a comparison of the predicted ratings for the post-crisis period (2008-2012) with the actual credit ratings in the same period. Panel A reflects the estimated coefficients in table 5. Panel B corresponds to the estimated coefficients in table 6. Panel C reflects the estimated coefficients in table 7. The difference in means are analyzed by a two-sided t-test (Observation-Prediction). Symbols * and ** denote significance levels of 5% and 1% respectively.

Figure 5: Propensity score distribution



Note: Figure 5 illustrates the propensity score distribution for untreated (period 2004-2007) and treated (period 2008-2012). The upper graph are showing the density plot, while the lower graph shows the histogram. The common support is analyzed by using program 'psmatch2' in Stata 13. Data Source: Bureau Van Dijk - Orbis Database (accessed 30 March 2014) and Moody's Investor Service (accessed 16 March 2014).

6.4 Common support

To ensure that the matched observations from both the control group and treatment group fell into the PS region of common support, we graphed the propensity score distribution for both groups. (see Figure 5). By looking at the distribution of propensity scores for both the treated observations and the control group, we see a good overlap between the two distributions. This applies to both the histogram and the density plot. We therefore conclude that the common support condition is satisfied.

6.5 Propensity score matching differences

By means of the estimated propensity scores, ratings before the financial crisis (control group) are matched with ratings after and during the financial crisis (treatment group), using the mentioned matching methods.

The difference we are analyzing is the average treatment effect of the treated (hereby ATT). ATT is the difference between the outcomes of treated and control observations, for the observations within the treated group (Caliendo and Kopeinig, 2008). The ATT thus focuses directly on the credit ratings after the financial crisis and calculates the gross effect the crisis had on the ratings. ATT is expressed analytical as:

$$ATT = E[Y(1) - Y(0)|W = 1] \tag{1}$$

All four matching methods used in this analysis gave significant ATT [2]. However, all methods except Mahalanobis-metric matching, gave absolute percentage bias over 5% level. The only matching method that gave satisfactory matching quality was Mahalanobis matching within calipers by the propensity score. See appendix C for a summary of all matching methods and their results.

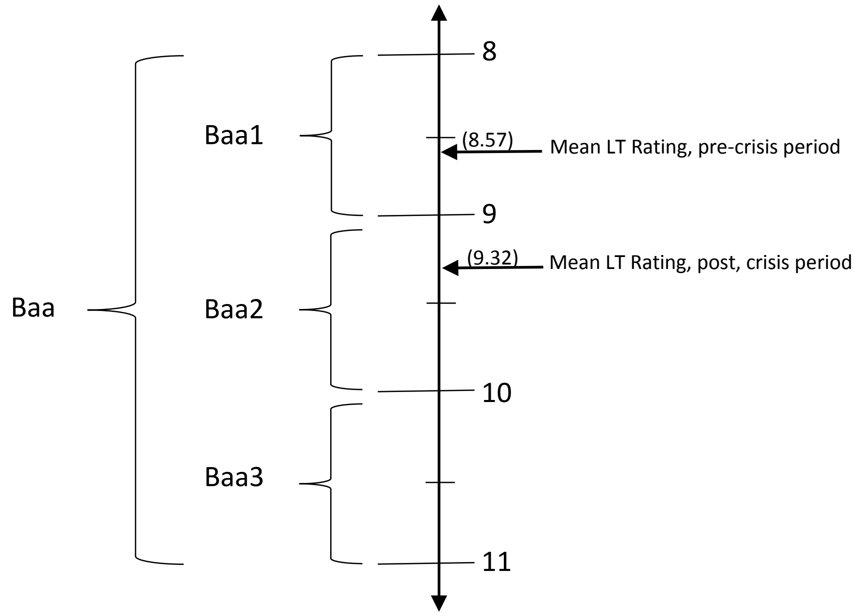
The resulting ATT for the Mahalanobis-metric matching was approximately 0.75. This implies that the average numerical rating has increased by 0.75 from the pre-crisis period to the post-crisis period, after controlling for fluctuations in corporate fundamentals. By converting the numerical difference using Moody's Long Term Rating Scale (Table 3) we see that the rating has gone from an average around Baa1 (8.57 in numeric) to Baa2 (9.32 in numeric). See Figure 6 for a graphical illustration.

7 Cross section analysis

Our analysis is to this point based on an accumulation of observations across countries and industries. To further explore the data set we want to break down the sample into several different subsets, so see if there is more and stronger evidence of excessive downgrading in some parts. We will look closer at investment- and speculative-ratings, Aaa-rated countries compared to others, differences in industries and also delaying the post-crisis period by one year to see if this affects the results.

[2] Significant difference implies a t-stat of more than 1.96 at a 5% significance level

Figure 6: Estimated treatment effect for long term credit rating

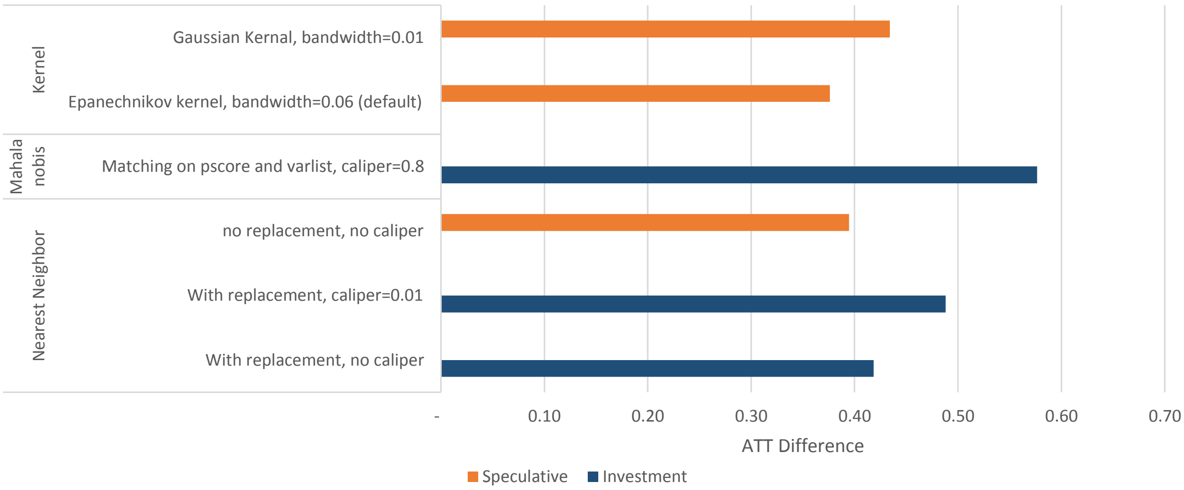


Investment versus speculative

Since credit ratings are divided into investment- and speculative-rated, it is interesting to look at differences between these two main groups. By separating the two groups and perform the same PS-analysis on each group, we find that there is a slightly higher difference for investment grade than for speculative, as seen in Figure 7.

The difference between the two groups is, however, not large enough to conclude that investment grade companies were hit harder by a tightening in rating policy, but may show a tendency. It is also important to note that the number of matched observations has gone down compared to the analysis for the whole sample, leading to a less robust result. We also perform out-of sample predictions for 2008-2012 for both investment- and speculative graded observations based on the three different regression models. When comparing with the actual ratings we see the same tendency as in the PSM, with a higher over-estimation frequency within investment-graded observations for all three models than within speculative-graded. This implies that CRAs actually rated investment-graded corporates stricter than companies with a speculative grade after the crisis. The observations are also supported by the mean-comparison tests between predictions and observations, which shows significant differences ranging from 0.52 to 0.95 within investment. For speculative, the mean difference is insignificant across all three models. Compared to the PSM results, the predictions show a somehow stronger result in the same direction, indicating that investment grade companies were evaluated stricter after

Figure 7: ATT difference for speculative vs investment graded companies



Note: Figure 7 shows the ATT difference in ratings between pre crisis period (2004-2007) post crisis period (2008-2012). The differences is calculated with the listed matching methods in Stata 13. Only included matching methods with an ATT Difference t-stat within 10% significance level. Data Source: Bureau Van Dijk - Orbis Database (accessed 30 March 2014) and Moody's Investor Service (accessed 16 March 2014).

the financial crisis compared to speculative grade companies.

Country domicile

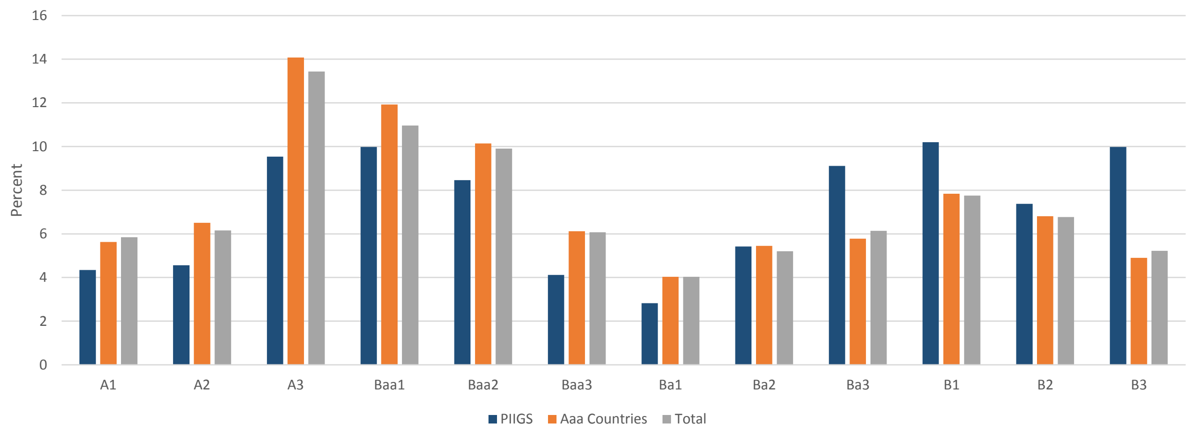
The same analysis can be carried out for companies within Aaa-rated countries compared to all the other companies. We define Aaa-rated countries as countries that had an Aaa rating through the whole period of analysis. The selection includes Germany, Finland, Netherlands, UK, Austria and Luxembourg. If we observe the rating distribution in companies with domicile in PIIGS[3] countries versus companies in Aaa countries (Figure 8), we see a clear tendency of higher share of investment-grade ratings within Aaa Countries.

An important note is that companies within PIIGS countries only constitute 461 observations. This is too few observations to perform a proper PS-analysis. We have however, compared the effect on ratings within Aaa-rated countries with all other countries. Even all other countries except Aaa-rated countries give quite few observations with a total of 967, but enough to give one satisfactory PS-matching model (See figure 9).

From Figure 9 we see that the difference in rating before and after the financial crisis increase dramatically if we remove companies within Aaa-rated countries. This may indicate that the credit rating agencies in general have a larger focus on companies within

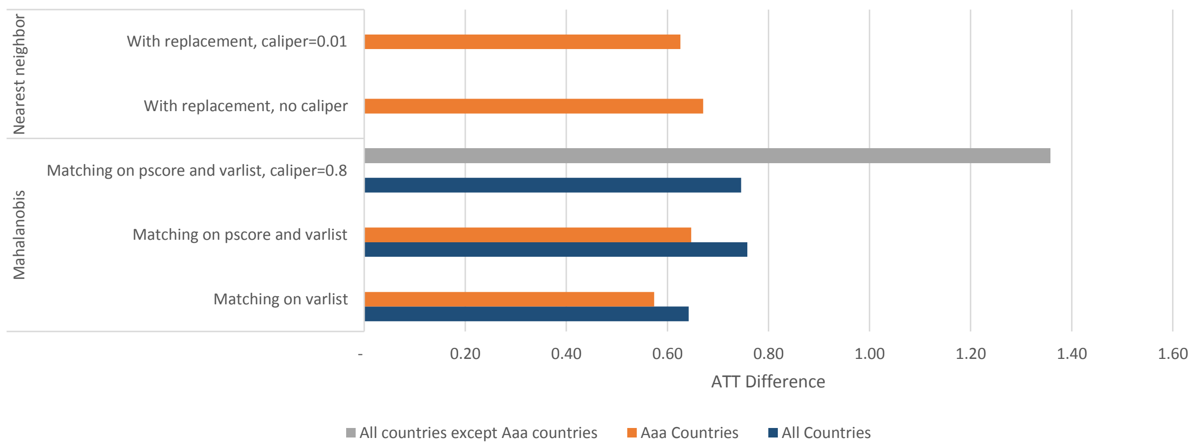
[3]PIIGS Countries: Portugal, Italy, Ireland, Greece and Spain

Figure 8: Rating distribution within PIIGS and Aaa rated countries (Only the rating classes with the largest number of observations is shown)



Note: Figure 8 provides a graphical illustration of all ratings within PIIGS countries and Aaa countries. The rating classes Aaa, Aa, Caa, Ca, C is leaved out for illustrational purposes. Data Source: Moody's Investor Service (accessed 16 March 2014).

Figure 9: ATT difference for Aaa countries, all countries except Aaa countries and total sample



Note: Figure 9 shows the ATT difference in ratings between pre crisis period (2004-2007) post crisis period (2008-2012). The differences is calculated with the listed matching methods in Stata 13. Only included matching methods with an ATT Difference t-stat within 10% significance level. Data Source: Bureau Van Dijk - Orbis Database (accessed 30 March 2014) and Moody's Investor Service (accessed 16 March 2014).

less robust countries, and that the threshold for a downgrade in rating is lower within these countries.

When performing out-of-sample predictions with our three regression models we find the same tendency. Both OLS-restricted and Panel data FE over-predicts more for observations within PIIGS countries and all countries except Aaa, than for observations within Aaa rated countries. More statistically we see a mean difference between predicted- and observed observations around 0.8 and 1.0 for OLS-restricted and Panel data FE respectively. However, OLS-unrestricted gives very inconsistent answers suggesting a significant under-prediction of 1.34 within PIIGS countries, measured as difference in means between predicted and actual observations.

If we focus on the OLS-restricted and Panel data FE we see the same tendency as for the PSM results; a smaller tightening of credit rating policy after the crisis for companies within Aaa-rated countries.

However, it is important to acknowledge that both the regression models and the PSM are based on quite few observations and may not capture all other effects on ratings. There is always a risk that it may be other effects on the ratings, that has nothing to do with rating policy, which trigger the result.

Ratings across industries

Another interesting aspect is whether the results vary between different industries. Due to a small sample of observations within each industry, it is problematic to use propensity score matching to analyze across industries. The same problem arise when using the regression models to generate out-of-sample predictions within the different industries. We have however looked at rating statistics within selected industries, as showed in Table 9.

From the statistics, we see that there are significant differences in both rating variations across industries, as well as in changes in ratings between the two periods. This implies that there may be significant differences in rating policy across industries, but it is hard to detect without further analysis.

Ratings as lagging indicators

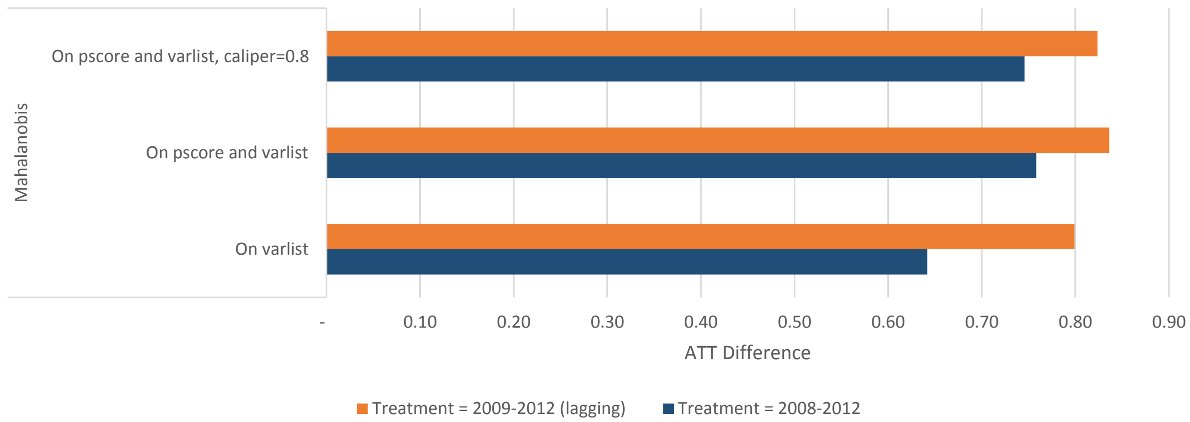
One might argue that ratings are functioning as lagging indicators, because they are reactive by nature. Based on this, it is natural to assume that the effects of the financial

Table 9: Rating statistics within each industry

Industry	# of obs	Prior to the financial crisis (2004-2007) Mean	std dev	During/after the financial crisis (2008-2012) Mean	std dev	Difference of mean
Total sample	4 270					0.69**
Utility	426	6.39	1.61	7.23	2.03	0.84**
Technology	92	9.24	4.14	11.07	4.42	1.83**
Transportation Services	204	3.82	4.67	6.25	4.76	2.43**
Whlsl dstrbtn	25	14.85	1.91	12.25	1.06	-2.60**
Investment Management	21	1	0	8.88	4.03	7.88**
Constr & Engineering Serv	54	10.14	4.75	13.78	4.04	3.64**
Lodging	20	11.22	1.79	13.73	2.33	2.51*
Gaming	42	13.76	1.78	14.11	2.21	0.74
Consumer Products	449	9.51	3.74	9.12	3.88	-0.39
Forest Products	41	12.73	3.58	11.93	2.08	-0.79
Energy	326	7.93	5.5	8.26	4.56	0.33
Telecommunications	427	10.51	3.85	10.71	3.68	0.21
Pharmaceuticals	88	7.49	5.2	8.04	4	0.55
Leisure & Entertainment	15	9.67	4.13	10.89	4.62	1.22
Packaging	101	14.38	1.63	14.43	2.25	0.05
Defense	47	7.55	3.98	7.44	5.17	-0.11
Aircraft and Aerospace	28	7.67	4.5	7.31	2.18	-0.36
Automotive	234	8.78	3.87	9.46	3.44	0.68
Chemicals	267	10.62	3.89	9.96	3.63	-0.66
Environment	19	11	4.14	8.36	2.8	-2.64
Healthcare	102	8.59	4.57	12	2.63	3.41**
Manufactoring	423	10.4	3.41	10.91	3.44	0.51
Media	278	10.86	3.31	11.35	3.47	0.49
Metals and mining	169	10.77	3.48	10.97	3.7	0.20
Natural Products Processor	54	9.56	2.25	11.89	2.99	2.33**
Retail	187	10.29	3.71	10.9	3.34	0.60
Services	129	10.71	4.32	13.3	2.9	2.59**

Note: Table 9 provides descriptive statistics for company rating within each industry. The difference in means is analyzed by a two-sided t-test. Symbols * and ** denote significance levels of 5% and 1% respectively. Data source: Moody's Investor Service (accessed 16 March 2014).

Figure 10: ATT difference for treatment(2009-2012) vs treatment(2008-2012)



Note: Figure 10 shows the ATT difference in ratings between pre-crisis period and post-crisis period. The post-crisis period is defined respectively as 2009-2012 (orange) and 2008-2012 (blue) The differences is calculated with the listed matching methods in Stata 13. Only included matching methods with an ATT Difference t-stat within 5% significance level. Data Source: Bureau Van Dijk - Orbis Database (accessed 30 March 2014) and Moody's Investor Service (accessed 16 March 2014).

crisis will not affect the ratings before some time. To include such a conception in the analysis, we redefine the post-crisis period as 2009-2012, in other words, delaying the effect with one year.

The results from the propensity score matching show a general increase in the difference before and after the crisis compared to the initial analysis. The results are shown graphically in Figure 10.

As for the out-of-sample predictions, using the regression models, we find significant mean-differences between predictions and observations of 0.77, 0.78 and 0.82 for OLS-unrestricted, OLS-restricted and Panel data FE respectively. These differences are slightly higher than the differences we found in the initial analysis with time period 2008-2012, (0.62, 0.70 and 0.83). The out-of-sample predictions thus points in the same direction as the PSM results.

The results from both PSM and the regression explicitly show that the difference in ratings is consistently larger when looking at the effect during and after 2009, instead of 2008. This implies that the CRAs may behave reactive, or at least not proactive. It additionally supports our previous results that indicate a general tightening of rating policy.

The differences of means between out-of-sample predictions and observations for all the analysis in this section are listed in Appendix D.

8 Discussions and limitations

8.1 Significant variables

The variables found significant in the OLS restricted regression were *ROCE using profit/loss before tax, profit margin, interest cover, solvency ratio asset based* and *sales*. Both *interest cover, sales* and *solvency ratio* were also significant in the panel data regression with fixed effect.

ROCE using profit/loss before tax gives a measure of the return the firm has earned on the capital employed. Because the ratio includes liabilities in the equation, it is useful when comparing companies in different industries with varying capital requirement. A high ROCE indicates that the company is able to generate profits with the capital employed, thus using their capital efficiently. We should therefore expect a positive correlation between *ROCE* and *rating*, which our results support.

The *profit margin* is here specified as profit before tax/operating revenue. By using pre-tax profit, you remove the fact that companies use different tax optimization techniques, thus also the possibility to manipulate earnings by manipulating the timing and size of the taxable income. However, profit margin contains financial items such as return from stock investments, which is unpredictable and has nothing to do with the underlying operational robustness (Loth, 2014). Both the intuition and our results show a positive correlation also for this ratio.

Interest cover ratio considers operating profit as a multiple of the firm's interest expenses. A high interest cover indicates therefore robustness, thus we should expect a positive correlation with rating. Our findings indicate that interest coverage has great importance in CRAs' rating methodology. Koller *et al.* (2010) underpins this result by finding that a limited number of credit ratios explained credit ratings fairly well, with interest coverage as the single most important indicator. The data set they used included all U.S. and European companies rated by S&P excluding financial institutions. They found that *interest coverage* explained more than 45 percent of rating differences. They also found that coverage ratios differed across sectors for a given credit rating. In other words that one sector could have better credit ratings than others at the same level of interest coverage. One explanation for this is the earnings volatility. Industries with volatile earnings must compensate with higher coverage in order to maintain a given rating. Volatile earnings increase the probability for not fulfilling the company's interest obligations due to lack of sufficient cash flow (Koller *et al.*, 2010).

Solvency ratio is here specified as shareholders' funds as a percentage of total assets. Solvency ratios in general measures the company's ability to meet its long-term obligations, and give insight in how much shareholders would expect to receive if the company liquidates. One important aspect with this ratio is that it fails to capture the company's ability to fulfill its debt obligations either short or long term. The reason is that market-to-book ratios can vary significantly across sectors and over time (Koller *et al.*, 2010). However, solvency is much more relevant in extreme scenarios where the company is under financial distress and creditors want to have a rough measure of available collateral. Solvency's relevance in a credit rating process is thus possible to defend.

Sales functions as a proxy of growth, but also reflects company size. *Sales* as an explaining factor in credit ratings, can be explained by the fact that larger companies naturally tend to be more diversified and therefore, all else equal, would tend to have lower business risk (Amato and Furfine, 2004). Note however that size is not a decision variable because it is mostly not within the management control. Another important aspect is that it usually only makes a difference for very large or very small companies.

One variable we expected to be relevant was *leverage*, in terms of *gearing*, defined in Orbis as non-current liabilities plus loans divided by shareholder funds. From general economic theory, we know that using external debt increases risk of bankruptcy and induces bankruptcy costs. In addition, high leverage increases the risk of business erosion and financial distress. Highly leveraged companies faces the risk of missing investment opportunities or reduce budgets in research in the future, since they need cash in order to repay debt obligations (Schindler and Schjelderup, 2012). This may lead to missing investment opportunities and hence opportunities to create more value for the company. Thus, we should expect a significant negative correlation with corporate ratings. On the other hand, replacing equity with debt reduces taxable income and thus increases the value of the firm. External creditors also impose management discipline that prevents corporate overinvestment (Koller *et al.*, 2010). There exist, in other words, a trade off which makes leverage not entirely negative. That may be the reason for why we do not observe gearing as a significant coefficient in the regression models.

None of the macroeconomic variables were shown significant. This is interesting, yet not too surprising. First, consider this citation from S&P's Corporate Ratings Criteria:
"The operating environment in the particular country - including, importantly, any sovereign related stress - can have an overwhelming impact upon company creditworthiness, both direct and indirect. Sovereign credit ratings suggest general risk faced by local entities, but they may not fully capture risk applicable to the private sector. As a result, when rating corporate or infrastructure companies or projects, we look beyond the

sovereign ratings to evaluate the specific economic or country risk that may impact the entity's creditworthiness. Such economic or country risk pertains to the impact of government policies upon the obliger's business and financial environment, and a company's ability to insulate itself from these risks." (Standard & Poor's, 2008)

Interpreting this would be that macroeconomic factors can explain the credit risk to some extent, but there are more qualitative measures that cannot easily be quantified.

As our results show, credit rating agencies tend to have largest focus on size and interest cover out of the quantitative factors when looking at both the regressions and correlation t-scores. This coincide well with the findings of Pettit *et al.* (2004) and Koller *et al.* (2010), who found that credit ratings are primarily related to size in terms of sales or market cap, and interest coverage terms.

8.2 Propensity matching results

The ATT from the Mahalanobis-metric matching for the total sample was approximately 0.75. This implies that the CRAs downgraded European corporates more than the worsening in corporate fundamentals would justify. However, the excessive downgrading is less than one subclass down, from a mean equivalent of Baa1 to Baa2. The results are hence not very large, and it is problematic to conclude on such a small difference. However, it may show a tendency toward a stricter rating policy after the financial crisis. It also points in the same direction as the regressions calculated from pre-crisis period sample, which in general overestimated the ratings compared to the actual ratings in the post-crisis period.

One rationale for rating agencies to become excessively conservative after the financial crisis is their lack of predicting the financial crisis. Much of the criticism after the financial crisis was in fact that CRAs had issued high ratings on investment products and corporates which later proved to involve great risk, where perhaps Lehman Brothers default remains as the most striking example[4] . Specially, rating agencies had an incentive to become more conservative in order to rebuild their own reputation (Ferri *et al.*, 1999).

On the other hand, the reason why we are not observing a larger ATT difference may be related to credit-cliff dynamic (Manso, 2013). A credit cliff implies that a poorly performing company will worsen its capital costs substantially by a rating downgrade. This can put pressure on the company's liquidity and its business (Gonzales *et al.*, 2004).

[4]Lehman Brothers' own debt still had an investment grade rating (A2) when it filed for bankruptcy protection 15 September 2008 (Krantz, 2013) (Moody's Investor Service, 2008)

Because of the major consequences such downgrades may have on businesses, CRAs may be hesitant to downgrade, at least to a certain extent.

The cross section analysis, with the same matching specification as for the total sample, showed an ATT difference of 1.36 for all companies that are were not rated Aaa, compared to a difference of approximately 0.63 for countries within Aaa rated countries. The reason for this result might be how the financial crisis affected Europe. As opposed to the financial crisis in USA, the financial crisis in Europe also evolved into a European sovereign crisis. This may led to a greater focus on company domicile for the CRAs, especially companies with domicile in less robust economies, thus leading to a lower threshold to downgrade companies in these economies compared to more robust economies with Aaa sovereign rating.

8.3 Procyclicality

From our results, with a certain tendency towards excessive downgrade during and after the financial crisis, we see that rating agencies may behave in a manner that potentially generates procyclical corporate ratings. By procyclical ratings we refer to CRAs' tendency to excessive downgrade in recession, leading to higher capital cost for the downgraded companies, thus leading to even poorer results and an intensified recession (same definition as in introduction).

One can alternatively argue that another aspect of procyclicality is the CRAs overgenerous ratings in time leading up to the crisis. Our results show a tightening in rating standards from 2004-2007 to 2008-2012. We cannot know if the pre-crisis period is representative for the 'normal' rating standard, or if the ratings in this period are generally too liberal. If the latter is true - that 2004-2007 represents a period of too high credit ratings according to fundamental values - maybe our results are a mere correction back to a level before the crisis. However, if the period represents a 'normal' period, we can argue that our results show a general tightening of standards that goes beyond the loosening before the financial crisis.

Ferri *et al.* (1999) argue that credit ratings are procyclical due to the reputation incentives faced by CRAs. They further elaborate that CRAs depend on what they refer to as 'reputation capital' and whether this capital fluctuates procyclically. They then may have an incentive to set ratings procyclically. We have already mentioned this argument for excessive downgrading, but a similar reasoning can also be used to explain a less conservative rating process during an expansionary period. Pursuant to Ferri *et al.* (1999), the CRAs' reputation capital is likely to be high during an expansionary period, thus leading to less focus on building reputation and hence a less strict rating process.

The cross section analysis shows signs of rating acting as a lagging indicator, meaning that the downgrades do not happen in the initial stage of the crisis. With respect to procyclicality this may argue against procyclicality due to the fact that lagging indicators will not be an accelerating factor, but rather behave in a reactive manner. El-Shagi (2010) supports this to some extent, although looking at sovereign ratings, stating that downgrading of sovereign ratings do not lead to an acceleration of the crisis, due to the lagging behavior. However, a procyclical behavior can also be defined as an effect that makes the crisis deeper and longer than initially expected, in this case the feedback effect from downgrades. Both Gärtner *et al.* (2011) and Manso (2013) mention such a feedback effect, stating that downgrades might actually trigger a self-fulfilling prophecy, making them part of a procyclical economy. Pursuant to Manso, small shocks to corporate fundamentals due to the financial crisis, may lead to multi-notch downgrades, making the business environment for the corporates even harder, leading to an additional reduction in business activity.

For the first time in recent history comprehensive regulations have been added on the CRAs' methods in Europe. They are to have more openness and transparency in their process to ensure more accurate ratings. Utzig (2010) argues that these regulations probably will help with the corporate governance of the CRAs. However, he addresses that the regulations lack the power to improve the competition to the oligopolistic market of CRAs and that they are not made more liable for their ratings. More relevant for this study is the immediate effects of the regulations, and whether they contributed to tighten the standards. Our results show stricter standards after the crisis, and one might look at this as a one-time drastic remedy to correct the levels of ratings, and hereby use these new harsher standards. In this way, the change we observe might not be explained by procyclical anatomy of credit ratings, but rather by a single event of tougher standards due to regulations.

8.4 Limitations

As seen from the cross section analysis, there are significant differences between industries. Optimally we should have included an industry-variable in our PSM in order to capture other industry characteristics, such as business cycle, competition, number of competitors etc. These are characteristics not reflected in the company's financial ratios. However, many of such characteristics are of more qualitative nature, thus much more difficult to quantify in a model.

Since we are doing PSM on firm characteristics, the same limitation applies to macro-indicators. The lack of macro-indicators in the PSM implies that we are not picking up changes in macro variables.

Another risk is the possibility of not matching on the most important variables, meaning that there could exist other and more relevant variables we that we could have used in the PSM replacing the other variables used.

Our study is based on financials that are considered quantitative variables. We acknowledged that our analysis lacked qualitative variables early in the research. Even though the CRAs themselves state that they use qualitative measures on a widely bases, this would be hard to implement in our models. However, some of our results highlight the problem of missing qualitative variables and we must therefore urge this limitation once again. The regression models have low to moderate explanatory power, suggesting quantitative variables do not explain the whole assessment of a credit rating.

We also need to express once more that all financial institutions such as banks and insurance companies are not included in our analysis. This is due to their complex capital structure and new regulations that are exclusively for this industry. We do therefore have no knowledge of what the effects would be if we implemented these companies into our models.

Ratings are not perfect, but constitute the best tool we have for quantification of risk. The alternative to traditional rating is to let market prices quantify the risk for us. Moody's have developed 'Market Implied Ratings' (MIR) based on prices from the CDS[5] , bond and equity markets (Moody's Analytics, 2010). By using MIR, you achieve ratings more in line with the market and its free flow of information and transparency. However, such ratings are also much more volatile and are not equally foresighted as traditional ratings. We do not want to go deeper into this field, but rather acknowledge that there exist alternatives to additional ratings with their pros and cons, constituting an interesting subject for further research.

9 Conclusion

This study addresses the influence of financial crisis on corporate credit ratings, investigating European corporate credit ratings in two different periods: before the financial

[5]Credit Default Swap

crisis (2004-2007) and after/during the financial crisis (2008-2012). The main focus is to analyze whether there has been a general tightening in rating practice during and after the financial crisis.

By using regressions, we find that rating agencies seem to focus on variables within *profitability, efficiency, operational, structure* and *size* in line with their own stated rating methodology. However, when looking at the regression coefficients and the statistical correlation we find that the CRAs tend to have largest focus on *size* and *interest coverage* out of the quantitative factors.

Next, using a two-sample t-test to test the difference of mean for all ratios, ascertain a general decrease in ratios after 2008, thus indicating that the ratings are expected to have declined somewhat even with unchanged rating policies. The predicted ratings for the period 2008-2012, using the estimated regressions based on 2004-2007, tend to be generally higher than the actual ratings in all three regression-models used. Based on the quantitative models alone we can say that companies need better financials, all others alike, to receive the same rating after the crisis compared to before. This result of over predictions is a small indication that the CRAs actually rated corporates stricter after the crisis.

The findings from the propensity score matching points in the same direction. By looking at the ATT difference between the two periods, we find a general downgrading of approximately 0.75, where a value of 1 equals one sub-rank rating, adjusted for the effects from corporate fundamental variables. Although the difference is not large, it indicates that the CRAs actually may have a tendency of lowering their ratings after the financial crisis.

When comparing the ATT difference between speculative- and investment rated corporates we cannot find any clear difference. The regression models show a weak tendency towards higher differences within investment, but not enough to conclude. There is however, a tendency of a smaller ATT difference for corporates within Aaa-rated countries compared to corporates with domicile in the other countries. The results from the other cross-section analysis are also too vague to draw any strong conclusions, but they give indications that could be an interesting approach for future research. We see considerable differences in rating fluctuations when comparing industries. Also, we find some evidence saying credit ratings are lagging indicators, as the ATT difference grows as we set our post-crisis period to a later point in time. This is also supported by the regression models.

References

Alessi, C., Wolverson, R. and Sergie, M. 2013. *The credit rating controversy*, [Online]. Available at: `http://www.cfr.org/financial-crises/credit-rating-controversy/p22328`. Accessed 16.05.2014.

Altman, E. I. and Rijken, H. A. 2004. How rating agencies achieve rating stability. *Journal of Banking & Finance* 28(11), pp. 2679–2714.

Amato, J. D. and Furfine, C. H. 2004. Are credit ratings procyclical? *Journal of Banking & Finance* 28(11), pp. 2641–2677.

Basel Committee on Banking Supervision. 2013. A brief history of the Basel Committee. Report, Bank for Internatioal Settlements.

Becker, B. and Milbourn, T. 2011. How did increased competition affect credit ratings? *Journal of Financial Economics* 101(3), pp. 493–514.

Blume, M. E., Lim, F. and MacKinlay, A. C. 1998. The declining credit quality of US corporate debt: Myth or reality? *The journal of finance* 53(4), pp. 1389–1413.

Caliendo, M. and Kopeinig, S. 2008. Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys* 22(1), pp. 31–72.

DefaultRisk. 2014. *Credit rating agencies - globally*, [Online]. Available at: `http://www.defaultrisk.com/rating_agencies.htm`. Accessed 07.05.2014.

El-Shagi, M. 2010. The role of rating agencies in financial crises: event studies from the asian flu. *Cambridge journal of economics* 34(4), pp. 671–685.

European Commision. 2014. *Rating agencies*, [Online]. Available at: `http://ec.europa.eu/internal_market/rating-agencies/index_en.htm`. Accessed 05.05.2014.

Ferri, G., Liu, L. G. and Stiglitz, J. E. 1999. The procyclical role of rating agencies: Evidence from the east asian crisis. *Economic Notes* 28, pp. 335–355.

Gärtner, M., Griesbach, B. and Jung, F. 2011. PIGS or lambs? The European sovereign debt crisis and the role of rating agencies. *International advances in economic research* 17(3), pp. 288–299.

Gonzales, F., Haas, F., Johannes, R., Persson, M., Toledo, L., Violi, R., Wieland, M. and Zins, C. 2004. *Market dynamics associated with credit ratings - a litterature review*, [Online].

Grilli, L. and Rampichini, C. 2011. Propensity scores for the estimation of average treatment effects in observational studies. Training sessions on causal inference, bristol - june 28-29, Departimento de Statistica "Giuseppe Parenti" Universit de Firenze.

Güttler, A. and Wahrenburg, M. 2007. The adjustment of credit ratings in advance of defaults. *Journal of Banking & Finance* 31(3), pp. 751–767.

Hauser, M. 2013. *Model specification, endogenity, causality and iv estimation*, [Online]. Available at: `http://statmath.wu.ac.at/~hauser/LVs/FinEtricsQF/FEtrics_Chp1.pdf`. Accessed 12.05.2014.

Hung, K., Cheng, H. W., Chen, S.-S. and Huang, Y.-C. 2013. Factors that affect credit rating: An application of ordered probit models. *Journal for Economic Forecasting* (4), pp. 94–108.

Koller, T., Goedhart, M. and Wessels, D. 2010. *Valuation: Measuring and managing the value of companies*, vol. 499. John Wiley and Sons.

Krämer, W. and Güttler, A. 2008. On comparing the accuracy of default predictions in the rating industry. *Empirical Economics* 34(2), pp. 343–356.

Krantz, M. 2013. *2008 crisis still hangs over credit-rating firms*, [Online]. Available at: `http://www.usatoday.com/story/money/business/2013/09/13/credit-rating-agencies-2008-financial-crisis-lehman/2759025/`. Accessed 12.05.2014.

Livingston, M., Wei, J. D. and Zhou, L. 2010. Moody's and S&P Ratings: Are They Equivalent? Conservative Ratings and Split Rated Bond Yields. *Journal of Money, Credit and Banking* 42(7), pp. 1267–1293.

Löffler, G. 2004. An anatomy of rating through the cycle. *Journal of Banking & Finance* 28(3), pp. 695–720.

Löffler, G. 2005. Avoiding the rating bounce: Why rating agencies are slow to react to new information. *Journal of Economic Behavior & Organization* 56(3), pp. 365–381.

Löffler, G. 2006. Can rating agencies look through the cycle. Working paper, University of Ulm.

Loth, R. 2014. *Financial ratio tutorial*, [Online]. Available at: `http://www.investopedia.com/university/ratios/`. Accessed 12.05.2014.

Manso, G. 2013. Feedback effects of credit ratings. *Journal of Financial Economics* 109, pp. 535–548.

Moody's Analytics. 2010. Market Implies Ratings FAQ. Report, Moody's Analytics.

Moody's Investor Service. 2008. *Lehman Brothers Holding Inc. Long Term Credit Rating*, [Online]. Available at: https://www.moodys.com/credit-ratings/Lehman-Brothers-Holdings-Inc-credit-rating-674575. Accessed 12.05.2014.

Moody's Investor Service. 2013. Rating symbols and definitions. Report, Moody's.

Moody's Investors Service. 2013. *Rating Action: Moody's downgrades UK's government bond rating to Aa1 from Aaa; outlook is now stable*, [Online]. Available at: https://www.moodys.com/research/Moodys-downgrades-UKs-government-bond-rating-to-Aa1-from-Aaa--PR_266844. Accessed 12.05.2014.

Moon, C.-G. and Stotsky, J. G. 1993. Testing the differences between the determinants of moody's and standard & poor's ratings an application of smooth simulated maximum likelihood estimation. *Journal of Applied Econometrics* 8(1), pp. 51–69.

Park, H. M. 2005. Linear regression models for panel data using SAS, Stata, LIMDEP, and SPSS. Working paper, The University Information Technology Services (UITS) Center for Statistical and Mathematical Computing, Indiana University.

Partnoy, F. 2002. The paradox of credit ratings. In: *Ratings, rating agencies and the global financial system*, Springer, pp. 65–84.

Pettit, J., Fitt, C., Orlov, S. and Kalsekar, A. 2004. The new world of credit ratings. Report, UBS Investment Bank.

Princeton Data and Statistical Services. 2014. *Data and stastical services - panel data*, [Online]. Available at: http://dss.princeton.edu/online_help/stats_packages/stata/panel.htm#models. Accessed 12.05.2014.

Rosenbaum, P. R. and Rubin, D. B. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1), pp. 41–55.

Schindler, D. and Schjelderup, G. 2012. Debt shifting and ownership structure. *European Economic Review* 56(4), pp. 635–647.

Standard & Poor's. 2008. Corporate ratings criteria. Report, Standard & Poor's.

Taylor, J. B. 2009. The financial crisis and the policy responses: An empirical analysis of what went wrong. Report, National Bureau of Economic Research.

The Economist. 2007. *Rating Agencies - Measuring the measurers*, [Online]. Available at: http://www.economist.com/node/9267952. Accessed 12.05.2014.

The Financial Times. 2002. Moody's mulls changes to its ratings process .

The Technical Committee of The International Organization of Securities Commissions. 2004. Code of conduct fundamentals for credit rating agencies. Report, The International Organization of Securities Commissions.

The World Bank. 2014. *World Development Indicators*, [Online]. Available at: https://secure.nhh.no/data/views/reports/,DanaInfo=databank.worldbank.org+tableview.aspx. Accessed 07.02.2014.

U.S. Securities and Exchange Commission. 2014. *Credit rating agencies and nationally recognized statistical rating organizations*, [Online]. Available at: http://www.sec.gov/answers/nrsro.htm. Accessed 12.05.2014.

Utzig, S. 2010. The financial crisis and the regulation of credit rating agencies: A European banking perspective. ADBI working paper series 188.

Appendices

A Ratios used in the data set

Ratios	Definition
Profitability / Capital efficiency	
ROE using P/L before tax (%)	(Profit before tax / Shareholders funds)*100
ROCE using P/L before tax (%)	(Profit before tax + Interest paid) / (Shareholders funds + Non current liabilities)*100
ROA using P/L before tax (%)	(Profit before tax / Total assets)*100
ROE using Net income (%)	(Net income / Shareholder funds)*100
ROCE using Net income (%)	(Net income + Interest paid) / (Shareholder funds + Non current liabilities)*100
ROA using Net income	(Net income / Total Assets)*100
Profit Margin (%)	(Profit before tax / Operating revenue)*100
Gross margin (%)	(Gross profit / Operating revenue)*100
EBITDA margin (%)	(EBITDA / Operating revenue)*100
EBIT margin (%)	(EBIT / Operating revenue)*100
Cash flow / Operating revenue (%)	(Cash flow / Operating revenue)*100
Enterprise value / EBITDA (x)	Enterprise value / (Operating profit (loss) + Depreciation)
Market Cap / Cash flow from operations (x)	Annual Market Cap / 315514 (non-US comps) or OTLO (US comps)
Operational Ratios	
Net asset turnover (x)	Operating revenue / (Shareholders funds + Non current liabilities)
Interest cover (x)	Operating profit / Interest paid
Stock turnover (x)	Operating revenue / Stocks
Collection period (days)	(Debtors / Operating revenue)*360
Credit period (days)	(Creditors / Operating revenue)*360
Export revenue / Operating revenue (%)	(Exports / Operating revenue)*100
R&D expenses / Operating revenue (%)	(Research & development / Operating revenue)*100
Structural ratios	
Current ratio (x)	Current assets / Current liabilities
Liquidity ratio (x)	(Current assets – Stocks) / Current liabilities
Shareholders liquidity ratio (x)	Shareholders funds / Non current liabilities
Solvency ratio (Asset based) (%)	(Shareholders funds / Total assets)*100
Solvency ratio (Liability based) (%)	(Shareholders funds / (Non current liabilities + Current liabilities))*100
Gearing (%)	((Non current liabilities + Loans) / Shareholders funds)*100
Size	
Sales	Mil EUR
Total assets	Mil EUR
Macroeconomic	
Unemployment	The share of the labor force that is without work but available for and seeking employment. (% of total labor force)
Inflation	Measured by the consumer price index (annual %)
GDP Growth	Annual percentage growth rate of GDP at market prices based on constant local currency
Current Balance	The sum of net exports of goods and services, net primary income, and net secondary income. (%of GDP)

Note: Table provides definitions of all financial ratios exported from Orbis and World Bank. Source: Orbis User guide Definitions and World Bank - World Development Indicators.

B Correlation matrix

	ROEusi~x	ROCEus~x	ROAusi~x	ROEusi~e	ROCEus~e	ROAusi~e	Profit~n	GrossM~n	EBITDA~n	EBITma~n	Cashfl~e	Netass~r	Inter~15	Stockt~r	Collec~d	Credit~d	Curren~o	Liquid~o	Shareh~o	Solven~t	Solven~y	Gearing	logSales	logTot~s
ROEusingPL~x	1.0																							
ROCEusingP~x	0.8	1.0																						
ROAusingPL~x	0.8	0.9	1.0																					
ROEusingne~e	0.8	0.7	0.6	1.0																				
ROCEusingn~e	0.7	0.8	0.8	0.8	1.0																			
ROAusingne~e	0.7	0.8	0.8	0.8	0.9	1.0																		
ProfitMargin	0.6	0.7	0.8	0.5	0.5	0.6	1.0																	
GrossMargin	0.1	0.2	0.2	0.2	0.2	0.2	0.4	1.0																
EBITDAmargin	0.3	0.3	0.4	0.2	0.2	0.3	0.7	0.7	1.0															
EBITmargin	0.5	0.6	0.7	0.5	0.5	0.6	0.9	0.5	0.8	1.0														
CashflowOp~e	0.2	0.3	0.4	0.3	0.4	0.5	0.7	0.6	0.9	0.7	1.0													
Netassetst~r	0.0	0.1	-0.1	-0.0	0.0	-0.1	-0.3	-0.5	-0.5	-0.4	-0.5	1.0												
Interestc~15	0.5	0.7	0.7	0.4	0.5	0.6	0.5	0.1	0.2	0.4	0.2	0.0	1.0											
Stockturno~r	0.0	-0.0	-0.0	0.1	0.0	0.0	0.1	0.2	0.2	0.1	0.3	0.1	-0.1	1.0										
Collection~d	-0.0	-0.1	-0.1	-0.1	-0.0	-0.1	-0.0	0.2	0.1	0.0	0.1	-0.3	-0.1	-0.0	1.0									
Creditperiod	-0.1	-0.1	-0.2	-0.1	-0.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.2	0.1	-0.1	-0.2	0.2	1.0								
Currentratio	0.0	0.0	0.1	-0.0	-0.0	0.1	-0.0	-0.2	-0.1	-0.0	-0.2	-0.1	0.1	-0.2	0.2	-0.2	1.0							
Liquidityr~o	0.0	0.0	0.1	0.0	0.0	0.1	0.0	-0.1	0.0	0.0	0.0	-0.2	0.1	0.0	0.3	-0.2	0.8	1.0						
Shareholde~o	-0.1	0.1	0.2	-0.0	0.1	0.1	0.0	-0.1	-0.1	-0.0	-0.1	0.1	0.3	-0.1	-0.1	0.1	0.2	0.0	1.0					
Solvencyra~t	-0.1	0.0	0.2	-0.1	0.0	0.2	0.2	0.1	0.1	0.2	0.1	-0.3	0.3	-0.2	-0.1	-0.2	0.4	0.2	0.7	1.0				
Solvency_r~y	-0.1	0.0	0.2	-0.1	0.0	0.2	0.2	0.1	0.1	0.1	0.1	-0.3	0.3	-0.2	-0.1	-0.2	0.4	0.2	0.7	1.0	1.0			
Gearing	0.1	-0.1	-0.2	0.1	-0.1	-0.1	-0.1	0.1	0.1	-0.0	0.1	-0.0	-0.3	0.1	0.1	0.0	-0.2	-0.1	-0.7	-0.8	-0.7	1.0		
logSales	-0.0	-0.0	-0.0	-0.0	-0.0	-0.0	-0.1	-0.2	-0.2	-0.2	-0.2	0.1	0.1	-0.2	-0.1	0.1	-0.3	-0.2	-0.1	-0.1	-0.1	-0.0	1.0	
logTotalAs~s	-0.0	-0.1	-0.0	-0.0	-0.1	-0.0	0.0	-0.0	0.0	0.0	0.0	-0.2	0.0	-0.2	-0.0	0.1	-0.3	-0.2	-0.1	-0.1	-0.1	0.0	0.9	1.0

Note: Table shows pairwise correlations for all variables. Data Source: Bureau Van Dijk - Orbis Database.

C Propensity score matching

			ATT Difference	t-stat
Nearest	Neighbor	no replacement, no caliper	0.656	2.8
		no replacement, caliper=0.01	0.738	3.12
		With replacement, no caliper	0.638**	2.22
		With replacement, caliper=0.01	0.601	2.1
1-to-1	matching	10 neighbor, no caliper	0.531**	2.2
		20 neighbor, no caliper	0.594**	2.52
		10 neighbor, caliper=0.01	0.512**	2.15
		20 neighbor, caliper=0.01	0.541**	2.32
Kernel		Epanechnikov Kernel, bandwidth=0.06 (default)	0.498	2.28
		Gaussian Kernel, bandwidth=0.06 (default)	0.407	1.88
		Epanechnikov Kernel, bandwidth=0.01	0.529**	2.27
		Gaussian Kernel, bandwidth=0.01	0.585**	2.54
Mahalanobis-metric	matching	Matching on varlist	0.642**	2.17
		Matching on pscore and varlist	0.758**	2.59
		Matching on pscore and varlist, caliper=0.1	0.456*	0.85
		Matching on pscore and varlist, caliper=0.8	0.746*	2.52
Radius	matching	no caliper	0.421	3.46
		caliper=0.01	0.514	2.23

Note: Table shows the ATT difference in ratings between pre crisis period (2004-2007) post crisis period (2008-2012) for all corporates. Symbols * and ** denote absolute bias levels between treated and control group within each financial ratio of 5% and 10%, respectively. Data Source: Bureau Van Dijk Orbis Database and Moody's Investor Service.

D Mean comparisons, out-of-sample predictions

	OLS unrestricted	OLS restricted	Panel data FE
Total sample	0.62**	0.70**	0.83**
Lagged (2009-2012)	0.77**	0.78**	0.82**
Investment	0.95**	0.52**	0.56**
Speculative	-0.01	0.13	0.22*
Aaa countries	0.68**	0.67**	0.66**
All other countries	0.02	0.81**	1.12**
PIIGS countries	-1.34**	0.83**	1.04**

Note: Table shows the pairwise comparisons of means between predictions and observations in the post crisis period (2008-2012). The listed differences are defined as: Observation - Prediction. Symbols * and ** denote significance levels of 5% and 1%, respectively. Data Source: Bureau Van Dijk Orbis Database and Moody's Investor Service.