The Journal of Risk Finance

CDS spreads as an independent measure of credit risk

Florian Kiesel Jonathan Spohnholtz

CDS spreads as an independent measure of credit risk

Abstract

Purpose: The creditworthiness of corporates is most visible in credit ratings. This paper presents an alternative credit rating measure independently of credit rating agencies. The credit rating score is based on the CDS market trading.

Design/methodology/approach: A credit rating score is developed which is a linear function of logarithmized credit default swap (CDS) spreads. This new credit rating score is the first one completely independent of rating agency. The estimated ratings are compared with ratings provided by Fitch Ratings for 310 European and US non-financial corporates.

Findings: The empirical analysis shows that logarithmized CDS spreads and issuer credit ratings by agencies have a linear relationship. The new credit rating score provides market participants with an alternative risk assessment, which is solely based on market factors, and does not rely on credit rating analysts. The results indicate that our credit rating score is able to anticipate agency ratings in advance. Moreover, the analysis demonstrates that the trading volume has only limited influence in the anticipation of rating changes.

Originality/value: This study shows a new approach to measure the creditworthiness of firms by analyzing CDS spreads. This is highly relevant for regulation, firm monitoring, and investors.

Article classification: Research paper

Keywords: Credit Ratings, Credit default swaps, Credit risk, Credit rating agency, Implicit Ratings

1. INTRODUCTION

In 1909, John Moody established the first credit rating agency (CRA) by classifying bonds of the railway industry. He introduced simple rating symbols, from Aaa to C, which are still in use today. This simple letter scale allows investors to quickly judge and compare the default risk of investment opportunities. CRAs were considered important actors in financial markets. Moreover, regulators, e.g. the Securities and Exchange Commission (SEC) or the Basel Committee of the Bank for International Settlement, incorporated the requirement of credit ratings for investment decisions. Before the financial crisis started in 2007, investors relied on the accuracy of credit ratings and the evaluation procedures of CRAs (Kiesel, 2016). Prior research shows that downgrades have a negative effect on stock prices and increase CDS spreads, whereas rating upgrades only have a weak positive effect on stock returns and a weak negative or non-significant effect on CDS spreads. However, on September 15, 2008, the North American investment bank Lehman Brothers defaulted and in succession filed for bankruptcy. Prior to its default Lehman Brothers did not face a downgrade of its issuer rating (Flannery et al., 2010). The rating indicated that the creditworthiness of Lehman Brothers is of good quality and it was not changed until a few days prior to the default.

The power of CRAs in combination with the ability of CDS spreads reflecting credit default risks has lead research to investigate how raw market data information can be transformed into rating scales comparable to those used by CRAs. Hart and Zingales (2011) presented a new capital regulation based on the CDS price movement. Research about market-based risk measures increased in the last decade. Breger et al. (2002) introduced a method for translating spreads into so called implied ratings (IRs) using bond spreads. Their model and evaluation have been extended by Kou and Varotto (2005, 2008). Reyngold et al. (2007) are among the first who used CDS spreads as an indicator for a firm’s creditworthiness and applied a methodology with a broad sample of reference entities. Castellano and Giacometti (2012) made further refinements to it and compared their IRs to agency ratings from all three major CRAs. They all applied the same basic method for determining IRs. The developed method of IRs provides results with high accuracy. However, it requires prior knowledge of CRAs for each rating grade, in order to calibrate the model. Consequently, the existing IRs are still dependent on CRAs. This raises the question whether an alternative rating model without prior knowledge of credit rating grades is feasible. The aim of the paper is not to replace

CRAs. Boot et al. (2006) show that beside the measuring of corporate risk, CRAs have a monitoring role in financial markets, in particular with the rating review process. We argue that besides CRAs the CDS market has the ability to complement the risk assessment of CRAs. The objective of this work is the design and evaluation of an alternative market IR model, which is completely independent of opinions of CRAs. This study is the first one that implements a credit rating scale solely based on CDS spreads and which is therefore independent of CRA ratings.

The remainder of this paper is organized as follows. Section 2 presents the literature review. Section 3 reports the data set and Section 4 the applied methodology. Section 5 reports the results and Section 6 analyses the model fit. Section 7 concludes the paper.

2. LITERATURE REVIEW

Altman (1968) introduced a scoring method for the discrimination between bankrupt and non-bankrupt corporates. He uses a multiple discriminant analysis (MDA) with five ratios obtained from the balance sheets and income statements. The MDA is a method to distinguish between two or more mutually exclusive groups – in this case bankrupt and non-bankrupt corporates – by evaluating multiple variables. A simple linear combination of the variables forms an individual discriminant score that is denominated Z-score by Altman (1968). Based on a sample consisting of 33 bankrupt and 33 non-bankrupt manufacturing corporations he performs a MDA with the following five ratios, i.e. variables: Working Capital/Total Assets, Retained Earnings/Total assets, EBIT/Total assets, Market value equity/Book value of total debt, Sales/Total assets. In order to measure the accuracy of the model, Altman (1968) uses a matrix containing hits and misclassifications. The model classifies the corporates between the two groups with an accuracy of 95% in the sample used to derive the MDA-function. In order to provide a Z-score for privately owned companies Altman (2000) replaced the variable equity market value with book values. The boundaries of the revised MDA model were recalculated and proved slightly less reliant than the original one. This refined model is called Z'-score.

Another possibility to evaluate the creditworthiness is implied by spreads, which can be derived either from decomposed bond or from CDS spreads. Breger et al. (2002) introduced IRs using bond spreads. Their model has been extended by Kou and Varotto (2005, 2008). In this paper, we use the corporate CDS market instead of the bond market because CDS spreads provide a better measure for credit quality (Hull et al., 2004).

Moreover, in several studies, the CDS market has been evidenced to have lead on the bond market in the price discovery process (e.g. Forte & Peña, 2009; Hull, 2004; Norden & Weber, 2004; Blanco et al. 2005; Zhu, 2006). Since CDS spreads reflect credit risk in a gradual way, meaning the higher the credit risk, the higher the spread, they facilitate a distinction not only between bankrupt and non-bankrupt corporates, but also between different rating grades. Flannery et al. (2010) show that the CDS spread incorporated new information faster than rating agencies and therefore evidence that CDS spreads are more accurate compared to rating agencies. In contrast to the interpretation of the ordinal ranks by CRAs, however, raw CDS spreads are difficult to interpret without a comparison. A simple mapping model has dominated previous research concerning spread-IRs.

Fitch Solutions, a subsidiary of Fitch Ratings, reports their IR methodology in Reyngold et al. (2007). They were among the first who used CDS spreads instead of bond spreads to calculate boundaries for each rating grade. Fitch has its own pricing source, the Fitch CDS Pricing Service, which consists of more than 2,500 reference entities. The model is limited to five-year maturity CDS contracts, which are written on senior unsecured debt. Credit ratings are focused on providing an indication of the long-term creditworthiness of a corporate. CDS spreads, however, are subject to substantial short-term volatility, because they respond to news and market conditions (Reyngold et al., 2007). Therefore, Reyngold et al. (2007) introduce a preparation of the data. After testing a variety of statistical smoothing methods, they arrive at the application of an exponentially weighted moving average (EWMA).

After preparing the data by smoothing, a non-parametric mapping is applied in order to obtain boundaries to differentiate between rating grades. The basis for the non-parametric mapping process is set by Breger et al. (2002). They created a matched sample of agency ratings and credit spreads derived from bond spreads. This combination provides a data point for the upcoming non-parametric mapping process. Consequently, a set of agency ratings is required for deriving IRs and therefore have to be known a priori. At this point, the existing IR methodology falls short, since they cannot provide an IR, which is entirely independent of CRAs.

Figure 1 provides a schematic exemplary depiction of the boundary setting mechanism. The boundary, expressed as a spread, divides two adjacent rating grades. The points indicate the match of spreads on the horizontal axis and related rating grade on the vertical axis. The outliers, represented by the dotted line around the observations, result from overlapping of CDS spread observations in different agency rating categories, so that the lowest (highest) observation within a rating category cannot be used to indicate the lower (upper) boundary for the IR grade (Breger et al., 2002).

[Place Figure 1 around here]

The boundaries are set using a penalty function, which penalizes misclassified spreads. The basic penalty function is introduced by Breger et al. (2002). Reyngold et al. (2007) adjusted the model in order to avoid unintuitive crossing of boundaries. They used the following penalty function to obtain the dividing boundaries:

$$\min_{i} \sum_{r} \sum_{i}^{n_r} \begin{cases} s_i > b_r^+ & \dfrac{(s_i - b_r^+)^2}{n_r} \\ s_i < b_r^- & \dfrac{(b_r^- - s_i)^2}{n_r} \\ else & 0 \end{cases} \tag{1}$$

$s_i = spread\ observation\ i$

$b_r^+ = upper\ boundary\ for\ rating\ grade\ r$

$b_r^- = lower\ boundary\ for\ rating\ grade\ r$

$n_r = number\ of\ observations\ in\ rating\ grade\ r$

As shown in Figure 1, the upper boundary of a higher rating grade, e.g. A, equals the lower boundary of the adjacent lower rating grade, e.g. BBB. The process of penalizing works as follows: Every spread that is observed outside upper and lower boundaries of a rating grade, although it is assigned to this rating grade, is misclassified. If the spread lies within the upper and lower boundaries, it is not being penalized. The squared distance between misclassified spreads and boundary is calculated and normalized by dividing by the number of observations in the rating grade. This process is applied to all rating grades. By minimizing the sum over all misclassified observations in all rating grades, the boundaries are set optimally (Reyngold et al., 2007). Subsequently, these boundaries can be expressed as a discrete function mapping a CDS spread to an IR grade and could be used out of sample for deriving IRs for corporates, which are not included in the sample. The prior studies, applied the same mapping process, but vary the penalty functions.

After presenting the requirements for the data and introducing the non-parametric mapping process, a third property of the IR model emerges: When and how often is the mapping process performed, in other words, how often is the model calibrated? Whereas

Reyngold et al. (2007) process their model on a daily basis, Castellano and Giacometti (2012) limited the calibration to a year or semiannual. The latter implies that boundaries maintain for at least half a year. However, CDS spreads vary over time due to liquidity issues. Therefore, if the general CDS spread level increases, while boundaries remain unchanged, corporates could be reclassified arbitrarily. A daily or weekly calibration avoids the arbitrary reclassification by incorporating the overall level into the derived boundaries (Kou & Varotto, 2008).

Summarizing, previous research has already identified models for the interpretation and translation of credit spreads into IRs. However, the method requires a sample of agency ratings matched with CDS spreads for the purpose of deriving the boundaries. Consequently, on the one hand, the rating uses market-based information, on the other hand, the rating is based on a sample of corporates, which have to be rated by a CRA before performing the non-parametric mapping process. Therefore, the model is still dependent on CRAs. In Section 4, we provide a methodology which is independent of CRAs.

3. DATA SELECTION

In order to analyze the relation of CDS spreads and ratings, we use empirical data. The CDS spread data is obtained from the Thomson Reuters EOD. Thomson Reuters collects CDS quotes from multiple sources each day and computes a composite quote allowing a comprehensive market view. As they are obtained from different sources, Thomson Reuters however cannot ensure that the prices are from actual trades or from construction and quoting (Mayordomo et al., 2014). Nonetheless, CDS mid-spreads from Thomson Reuters are used by a variety of researchers (e.g. Longstaff et al., 2005) and conceded as viable representation. The sample contains corporates headquartered either in the United States or in Europe. CDS contracts are OTC derivatives and, therefore, may vary in terms and conditions. All contracts refer to senior unsecured debt of the reference entity. We only consider CDS contracts with a maturity of five years, because they are the most liquid ones.

In addition, we obtain Fitch ratings from the Fitch website from June 30, 2009 to July 1, 2014. For the analysis, a matched data sample is needed containing both information, the CDS spread as well as the credit rating for each day and corporate for at least some time of the investigation period. This process results in a sample consisting of 310, whereby

103 are European and 207 are from the United States. More detailed descriptive statistics are enclosed in appendix A. By analyzing the history of credit ratings of each corporate of the matched data sample, dates and absolute frequency of rating changes are obtained. 157 upgrades and 174 downgrades are found between week 24 in 2009 and week 27 in 2014. In order to classify the ratings, we transfer the rating classes into numbers. Thus, agency ratings are initially assigned to a numerical integer value as follows, AAA=1, AA+=2, and so on.

Figure 2 reports the 0.1 quantile, median and the 0.9 quantile of each trading day in the observation period.

[Place Figure 2 around here]

The level of CDS spreads vary heavily over time throughout the entire sample. Beside the 0.9 quantile, where an increase is most likely in times of financial turmoil due to a decreasing creditworthiness, but also for the median and 0.1 quantile a surge is recorded. In order to address this CDS spread variation over time, we recalculate the CDS spread boundaries on a weekly basis and do not assume a static relation over time.

4. METHDOLOGY

4.1 CDS spread smoothing

Following the methodology of Reyngold et al. (2007), an averaging process is applied to the CDS data in terms of smoothing the individual time series. As smoothing results into aggregation of data and loss of information, the rivaling interests of providing a long-term signal and timely reaction on changes in the creditworthiness have to be taken into consideration: Since agency ratings reflect the long-term creditworthiness of an underlying entity, IRs should provide a long-term signal as well. Löffler (2004) deeply analyze the anatomy of ratings through the cycle. He shows that ratings are not perfectly correlated with actual default risk. He also shows that rating stability is significantly higher than with a current-condition approach. In the light of timeliness however, smoothing may impede the ability of implied ratings to quickly react on changes in creditworthiness. Nonetheless, regulators and other market participants are primarily interested in the general outlook of creditworthiness rather than on short-term volatility. Additionally, adjusting the methodology by shortening time intervals leaves room for further evaluation.

In order to achieve the extraction of the long-term signal from the time series, the data has to be prepared by smoothing (Reyngold et al., 2007). Smoothing can be achieved by various statistical methods, e.g. different forms of moving averages. A moving average (MA) calculates arithmetic means consisting of a specified quantity of observations. Therefore, the MA averages over a certain time window. Building on the MA, the EWMA extends the model by introducing weights to the calculation of the arithmetic mean. Eventually, the EWMA assigns greater weights to observations that are more recent and weights are decaying exponentially going back in time (Hunter, 1986). The following formula constitutes the relation between observations and calculated values for the EWMA:

$$EWMA = \widehat{y_t} + \lambda * (y_t - \widehat{y_t}) \tag{2}$$

Extending this formula unveils the procedure of time decaying weights. Thus, the EWMA equals the predicted value plus the weighted error, whereby *EWMA* refers to the predicted value for the period (Hunter, 1986):

$$EWMA = \lambda * (y_1 + (1-\lambda) * y_2 + (1-\lambda)^2 * y_3 + \cdots) \tag{3}$$

λ is a constant that is defined between 0 and 1. With the alteration of lambda the memory of the EWMA is determined, i.e. small values of lambda lead to more historic values which are included in the EWMA, whereas lambda close to one solely includes the most recent observations (Hunter, 1986). The determination of lambda is made with N denoting the number of days and therefore setting the time window (Castellano & Giacometti, 2012). This relation is derived by asymptotical assumptions of a time series with infinite length and is therefore representative for large numbers of N:

$$\lambda = 2/(N+1) \tag{4}$$

Figure 3 provides an example time series from the sample. The blue line refers to the exponentially weighted MA with a time window of 250 trading days. Reyngold et al. (2007) tested several time windows for smoothing and they find that 250 trading days provide the best extraction of the long-term signal.

[Place Figure 3 around here]

These graphs illustrate how the series is smoothed and just the long-term signal is extracted. The mechanism of the EWMA leads to the disappearance of the first 250 data points.

4.2 Rating distribution and credit score

The general idea of the alternative credit rating score is based on assumptions about the distribution of credit ratings. Therefore, the following preliminary assumptions are supposed. First, credit ratings follow a certain distribution and, second, this distribution is not subject to significant variation over time (Löffler, 2004).

Figure 4 provides the distribution of sample credit ratings as of July 1, 2014. Visually, it can be assumed that credit ratings are Gaussian, i.e. normally, distributed. We use a Jarque-Bera-Test to test whether the empirical distribution is normally distributed. Preforming the Jarque-Bera-Test for the rating data of July 1, 2014 rejects the null hypothesis, that the ratings are normally distributed. Hence, the Gaussian distribution is no basis for the development of an alternative credit rating scale, but relevant descriptive statistics of a data set can be used if the individual distribution does not vary over time.

[Place Figure 4 around here]

Since agency ratings are merely changing due to long-term orientation and stability, the distribution of ratings is expected to be constant over time (Löffler, 2004). For providing evidence of an equal distribution over time, the Kolmogorov-Smirnov test is applied to the rating data set.

The Kolmogorov-Smirnov test is a goodness-of-fit test and assesses, whether the empirical distribution is not significantly different to the distribution stated under the null hypothesis. Therefore, the test is capable to compare the empirical distribution either to a reference distribution function (one-sample test) or to another empirical distribution function (two-sample test). However, the Kolmogorov-Smirnov test is implemented under the assumption of a continuous distribution. Since ratings are ordinal and discrete, an adjusted Kolmogorov-Smirnov test has to be applied to the observations of ratings. The adjusted Kolmogorov-Smirnov test is performed for the first half-year (June 30, 2009) with every following half-year under the hypothesis that there is no statistical significant difference in empirical distributions. The null hypothesis cannot be rejected at least on the 10 percent level of significance. Consequently, the ratings of the sample follow a constant distribution supporting the idea of developing IRs based on assumptions about the general distribution of credit ratings.

By assessing the Fitch Ratings Global Corporate Finance Transition and Default Studies of the years 2009 to 2014, the findings on the distribution of corporate credit ratings of the obtained sample can be transferred to the distribution of credit ratings among all corporates rated by Fitch Ratings. Therefore, Figure 5 reports the distribution of all corporate credit ratings of Fitch Ratings for the years 2009 to 2014. By comparing the distribution illustrated in Figure 4 with the distribution depicted in Figure 5, it can be assumed that the sample is representative in terms of the distribution of credit ratings. Furthermore, it is shown that the median is not changing. The relative frequency of corporates rated AAA is around 1% for every year. Similar results are observed for the combined lowest rating grade of CCC to C. The relative frequency accounts between 2% and 3% and therefore does not fluctuate heavily. As a result, the distribution provides general information for the formulation of an alternative credit rating score.

[Place Figure 5 around here]

Credit ratings use letter ratings in terms of ease of interpretation making bonds' creditworthiness comparable to others. In the last decades, market participants demanded for a finer distinction in rating classifications (Bannier & Hirsch, 2010). The rating scales were transformed from single letter ratings to multiple letter ratings. The next extension has been reached by introducing signs, which indicate the position within a specific credit rating grade. By adding rating outlooks and rating reviews, CRAs increased the informational content of their scales (Bannier & Hirsch, 2010). If a corporate is rated BBB- and is put under review for downgrade, the CRA implicitly classifies it among the corporates with lower creditworthiness in the rating grade BBB-, because it could be downgraded to BB+. Since there are three different reviews, namely reviewing for possible upgrade, downgrade or with direction uncertain, the rating grades are further divided by three gradations. In particular, the high volatility in credit profile within the "speculative" grade demands for rating watches in order to be information efficient. This development leads to the following introduction of a continuous credit rating scale instead of a discrete one. In addition, ordinal credit ratings lack of relevant information by assigning specific rating grades and impede mathematical analysis. Many statistical methods have to be transformed to fit the needs of ordinal data. Since the simple linear regression model requires continuous predictands, an ordinal score cannot be applied. In terms of regression models, the ordered probit model has been developed to circumvent this particular problem. In this case, however, an ordered probit cannot fit the purpose of estimating the relationship on sole basis of very few observations (Cheung, 1996). The method requires observations for every rating grade in order to derive probabilities for the assignment to a distinct class and are, therefore, again dependent on agency ratings.

Ultimately, a continuous credit rating scale not only provides additional and more precise information, but also enables the application of a linear regression model. For the upcoming analysis and development, Table 4 reports the employed scoring. Besides, the agency ratings of the sample are converted into numerical values as reported in Table 1. In terms of better comparability and the test of model fitness, ranges of scoring values are assigned to corresponding letter ratings as published by Fitch Ratings. These scoring values are named credit rating scores (CRS).

[Place Table 1 around here]

4.3 Relation between ratings and CDS spreads

As rating scales of CRAs are widely known and highly recognized in the market for providing risk assessment information, regulators or other private institutions can easily understand an independent measure for credit risk if it provides the same representation. For translating CDS spreads into the same scale, existing agency ratings are taken as a starting point for further analysis. Figure 6 illustrates a scatter plot of CDS spreads matched to the corresponding agency rating as of July 1, 2014. The plot shows that the CDS spreads and agency ratings are highly correlated with each another, because the observations are not distributed randomly. Moreover, the numerical agency rating and CDS spreads are positively correlated. This means high CDS spreads are observed for high credit ratings, indicating a worse creditworthiness. However, the plot also shows that the ranges of CDS spreads assigned to different rating grades overlap. As a result, the capability of CDS spreads predicting credit ratings seems to be limited. For example, a CDS spread of 100 bps is observed in seven different rating grades and therefore cannot be assigned to an agency rating unambiguously.

[Place Figure 6 around here]

Two conclusions result from the raw data: First, CDS spreads are not expected to mimic agency ratings, because they cannot be assigned unambiguously. Second, the distribution follows a pattern, suggesting that there is a relationship between CDS spreads and credit ratings. In order to investigate the relationship more deeply, the method of Reyngold et al. (2007) is applied to obtain an understanding how boundaries are derived for adjacent rating grades. Subsequently, a mathematical analysis is performed, which proves the relationship of boundaries and credit rating grades mathematically.

By applying the existing method introduced by Reyngold et al. (2007), assumptions about the distribution of determined boundaries can be made and lead eventually to an alternative method for setting the boundaries without the prior knowledge of agency ratings. Hence, the question arouses, whether there is a functional relationship between determined boundaries and the corresponding rating grade. The method of Reyngold et al. (2007) has been applied to the data, so that boundaries are obtained for every week in the observation period. In terms of better comparability, we use the mid-CDS spread[1] to average the upper and lower boundary of each rating grade. Figure 7 reports the distribution of these averages of CDS spread boundaries.

[Place Figure 7 around here]

It is visually evident that the boundaries follow a functional relationship. Furthermore, by analyzing the plot, a logarithmic relationship between CDS spreads and the corresponding rating grades can be found. First mathematical proof for a strong linear (positive correlation) relationship is provided by calculating the correlation coefficient for the logarithmized mid boundaries of CDS spreads and the corresponding rating grade. The correlation coefficient for weekly CDS spread mid boundaries and the corresponding rating grades ranges between 0.9686 and 0.9905.

A high correlation throughout all weeks implies that the relationship can be estimated by using a linear model. A linear regression model is applied to the data set in order to obtain evidence for this specific property. Based on the visually obtained findings, we model the rating grade as follows:

$$rating\ grade_i = \beta_1 + \beta_2 * \ln(CDSspread_i) \tag{5}$$

The function consists of a constant (β_1) representing the intercept with the ordinate and a second regressor estimating the slope (β_2) of the linear function. Hence the result of the function is the credit rating score assigned to a specific logarithmized CDS spread. In Figure 8 the linear regression model is fitted to the data points utilizing an ordinary least squares (OLS) estimator. The green line indicates the linear relationship between logarithmized CDS spread mid boundaries and the corresponding rating grade. The coefficient of determination of the linear regression applied to every week in the observation period ranges between 0.937 and 0.98 and therefore proves a good fit.

[Place Figure 8 around here]

Our analysis provides evidence for a functional relationship between logarithmized CDS spreads and the credit rating score. If the logarithmized CDS spread increases, the credit rating score is increasing linearly, which indicates a decline in creditworthiness. The upward sloping relationship can now be used in order to derive credit rating scores from an individual CDS spread. Furthermore, the application of the linear regression seems reasonable due to the obtained results.

5. EMPIRICIAL ANALYSIS OF CREDIT RATINGS AND CDS SPREAD

5.1 Mathematical model

After presenting evidence of a functional relationship between CDS spreads and rating grades, the question raises, whether this function could be derived without the application of the existing model and the derivation of boundaries, so that prior knowledge of agency ratings is not necessary.

The calculation of a linear regression is based on a data set, which consist of a dependent variable, in this case the IR, and one or more explanatory variables, here the logarithmized CDS spread. We already have shown that agency ratings follow a certain distribution and that this distribution is not subject to significant variation over time. Thus, the observations, which are required for the estimation, can be derived from assumptions about the matching of logarithmized CDS spreads to descriptive values of the distribution. Among many descriptive values of a data set the minimum (1) as well as maximum value (2) and the median (3) are most prominent. At this point, existing models are dependent on agency ratings, because they require observation of agency ratings for every rating grade. This model does not require any specific agency ratings for a sample of observations. However, the idea of estimating the credit rating score with the following subsamples rests upon the assumption that the CDS spreads of the sample follows a similar distribution in comparison to the distribution of agency ratings. This assumption can only be made if the CDS sample is large and contains observations for all levels of CDS spreads.

5.2 Further assumptions

We set the minimum as the lowest value of the data set. Above, the inverse relationship of logarithmized CDS spreads and credit ratings has been found, which means that the lowest CDS spread indicates the highest creditworthiness. The highest corresponding rating grade is AAA, which is transferred into the range between 0 and 1.49. Two different approaches emerge for the choice of the CDS spread: Either the CDS spreads of sovereigns or the minimum CDS spread of the corporate sample can be chosen. If sovereigns are rated with the highest rating grade, such as Germany or the United States, their CDS spread should represent the minimum value of every CDS sample. However, the minimum CDS spread from our corporate sample is lower than the sovereign CDS spreads of Germany and the United States. This suggests that CDS spreads with different types of reference entity cannot be compared. Actually, previous researches have decomposed CDS spreads of both sovereigns and corporates and find a different share of pure credit risk of the CDS spread. Badaoui et al. (2013) find that the pure credit risk premium only accounts for 54.5% of the spread of sovereign CDS. In contrast, Bühler and Trapp (2009) decomposed corporate CDS spreads and reported that 95% of the CDS premium are due to pure credit risk. Because the rest of this work is limited to corporate CDS spreads, only the minimum value of the corporate sample is used as reference point for the estimation of the linear regression. Therefore, the first subsample, in this case, data point, is obtained by assigning the minimum value of the corporate sample to the credit rating score of 1.

Moreover, the minimum value is considered to reflect substantial information for determining the position of the linear function. This arises the question, how the highest values of CDS spreads can contribute to the accuracy of the function. CDS spreads face a rapid increase prior to a triggering event. The data sample contains CDS contracts that have been triggered in course of history. Thus, although they represent the highest CDS spreads, they cannot be assumed to convey any helpful information. The triggered CDS contracts vary in level of their last official quote, making them inadequate for estimating the relationship between CDS spreads and credit rating grades. We therefore eliminated outliers either with a suddenly spread increase or a CDS spread that is remarkably higher than usually can be observed.

Based on this adjusted sample further assumptions can be made. Referring to the investigation of the distribution of agency ratings, an additional constant characteristic can be found: The cumulative relative frequency of the lowest credit ratings, namely CCC, CC and C, accounts on average for two percent of all ratings. Hence, by assigning the highest two percent of observed CDS spreads to the corresponding mid credit rating score of 17, a second subsample of data is derived for the estimation.

The third subsample determining the position of the linear function is set by matching the median CDS spreads with the median rating grade. The median rating grade is BBB, corresponding to a credit score value of 9. Because the Kolmogorov-Smirnov-test proved that the distribution of credit ratings is not statistically significant different over time, the median remains constant as well. Therefore, matching the median value of CDS spreads to the median rating grade provides a further subsample for the estimation. Since the median is a single value of a data set, the subsample is extended by adding up values that are located besides by four notches in order to consider more than one value.

6. TEST OF MODEL FITNESS

6.1 Hit-miss-match analysis

For validating the assumptions and provide proof that IR are a valid substitution for institutional judgement ratings of CRAs, the behavior of IR is compared to agency ratings. The analysis will provide proof that IR are at least as good as agency ratings for assessing creditworthiness and reflect changes in credit risk in an even timelier manner by anticipating ratings changes prior to the actual rating events. For the evaluation of hits, i.e. the IR equals the agency rating, and misses, i.e. IR and agency rating non-similar either in positive or negative direction, the hit-miss-match matrix provides information about distribution and extent of fitness. With denomination of agency rating on the abscissa and the according IR on the ordinate, the diagonal elements represent the percentage of hits for each rating grade and the off-diagonal elements indicate the percentage of reclassified entities within each rating grade (Castellano & Giacometti, 2012). Reclassification means that corporates' IR differ from the agency rating.

Table 2 and Table 3 provide a hit-miss-match matrix for the beginning of the sample period, week 29 in 2010, and the end of the sample period, week 27 in 2014. As the IR is calculated on a weekly basis, the hit-miss-match matrix represents rating alignment within this individual week. The green cells represent the diagonal elements, whereas reclassification by one and two notches is colored orange and red respectively. As of week 29 in 2010, 33.56% of ratings are matched excellently, whereas 72.55% and 92.73% are reclassified by one or two notches respectively. Similar results are found for week 27 in 2014. Here 29.39% of agency ratings are hit and 67.24% and 91.56% are reclassified.

[Place Table 2 around here]

[Place Table 3 around here]

However, the hit-miss-match matrix not only reports the share of reclassified ratings, but also unveils to which IR the reclassifying occurs. Whereas the reclassifying within lower rating grades is mainly limited to one or two notches, corporations with higher ratings are mostly reclassified by four notches in 2010. The impact of reclassification shrinks in 2014, where the reclassification of higher rated corporations is limited to three notches. This indicates potential limitations and weaknesses of the derived function for the credit rating score. On the one hand, this could be a result of a weak data basis of the sample for higher rated corporations. Only one corporation is assigned to the highest credit rating of AAA in the sample as of week 27 in 2014. On the other hand, it could be due to a potential misclassification of the derived logarithmic relationship.

Nevertheless, on average, the capacity of the derived model in predicting agency ratings is high and comparable to the results of previous studies. Reyngold et al. (2007), whose model is the basis for the model, report with 30.9% (71.7% and 89.0%) for America and Oceania and with 33.0% (78.6% and 93.8%) for Europe similar results for agency ratings hit (reclassified by one or two notches, respectively).

In order to gain an overview of the average hit-miss-match capacity of the derived IR for every year in the sample, Table 4 reports an aggregation of individual hit-miss-match matrices.[2] Therefore, the percentages of IR either hitting the agency rating or reclassifying within one, two or more than two notches are averaged annually. The reclassification is calculated by subtracting the numerical IR from the numerical agency rating. Thus, a negative notch reclassification indicates that the credit rating score assigned a better rating and therefore suggests higher creditworthiness.

Approximately 90% of the agency ratings are either hit or reclassified by one or two notches for every observed year. The majority of reclassifications is observed in negative direction meaning that the credit rating score is on average less strict and suggests a higher credit quality as the CRA.

[Place Table 4 around here]

6.2 Forward analysis

Since it is not necessary, that IR and agency rating equal each other, we test whether IR anticipate the future change of agency ratings. This supports the hypothesis that CDS spreads incorporate information about a change of creditworthiness of the underlying corporate faster than CRAs do. By constructing an event study around rating changes of an individual corporate, it can be tested whether the IR leads the agency rating and therefore anticipates the upcoming rating change. This test methodology follows Kou and Varotto (2008). Therefore, average differences between IR and agency rating are calculated prior to all observed rating changes. This calculation is made for downgrades and upgrades. If the mean average over the difference is significantly different from zero, the IR leads the agency rating. Given the mean and standard deviation over all averaged IR-AR differences for the observed rating changes, a *t*-test is performed. Note that in contrast to Hull et al. (2004) and Kou and Varotto (2008) the bootstrapping method is not necessary due to a higher sample size (at least 160 observations). The results are reported in Table 5.

[Place Table 5 around here]

The columns of Table 5 contain the mean averaged IR-AR difference, whereby the agency is subtracted from the IR. If the IR is higher (lower), i.e. reflecting a worse (better) rating grade, as the agency rating, the difference is positive (negative). A positive mean is observed for downgrades and is statistical significant at the 0.01 percent level of significance throughout all intervals. This indicates that the IR anticipates the future downgrade of a corporate at least 16 weeks prior to the announcement date. In terms of rating upgrades, the mean differences are negative and statistically significant throughout the periods, suggesting that IR also anticipates upgrades. Furthermore, it is noticeable that the mean average difference exceeds one for both upgrades and downgrades, which shows that the IR indicates on average a better (worse) creditworthiness by more than one notch. The mean difference generally increases by approaching the actual rating event.

For the relationship between IR and agency ratings, four different developments are possible: The IR either leads or is led by the agency rating. Furthermore, a divergent and converged behavior of the ratings is conceivable (Reyngold et al., 2007). The ability of IR anticipating the future movement of agency ratings, however, is most interesting, since this implies that IR reflect changes in creditworthiness timelier.

Whilst the study conducted above, only delivers evidence for anticipation power on average, in the following, the individual anticipation power of IR is investigated. A convergence analysis provides evidence for individual rating change anticipation. Whereas the analysis can be made for all four different patterns of implied behavior, this study is limited to the most relevant one: Do IR lead the agency ratings' behavior? Therefore, the following regression analysis is performed around every observed agency rating change.

$$implied\,rating_t = a + b * t + e_t \qquad (6)$$

Whereby *implied rating* denotes the derived IR at time *t* in weeks prior to the rating change, *a* and *b* are the regression parameters and *e* represents the error term. By comparing intercept (a) and slope (b) of the linear regression with the agency rating change, the relative behavior is evaluated (Kou & Varotto, 2008).

Figure 9 depicts the schematic behavior of an IR leading an agency downgrade, i.e. an increase in the numerical agency rating. Similar can be derived for agency upgrades (Kou & Varotto, 2008). Two possible cases for IRs leading agency ratings are investigated. Case one refers to the linear regression, where the estimate for *b* is not statistically significant different from zero. As a result, the IR equals the intercept, *a*, and is leading the agency downgrade if $\hat{a} \geq agency\,rating_2$. In the second case, $\hat{b}$ is statistically significant and the IR leads the agency downgrade if $\hat{a} + \hat{b} * n \geq agency\,rating_2$ (Kou & Varotto, 2008).

[Place Figure 9 around here]

Table 6 shows the results of the comparison between IR before the rating change and agency rating after the rating change. The columns report the individual anticipation power by calculating the percentage of rating changes that are successfully anticipated. Panel A contains the results for the derived IRs without the restriction concerning the volume. The IR anticipates at least 60% of the 173 downgrades and 163 upgrades.

Hence, further evidence is provided that IR lead agency ratings and, therefore, reflect changes in creditworthiness timelier.

[Place Table 6 around here]

Kou and Varotto (2008) conducted a similar study by deriving IR from bond spreads and Reyngold et al. (2007) as well as Castellano and Giacometti (2012) performed the non-parametric method for CDS spreads. We compare their results with our results. Although the method for deriving independent credit rating scores is developed by using an estimation method, the accuracy and anticipating power, which is found here, is comparable with the findings of the previous studies. Reyngold et al. (2007) found a similar anticipation power of 64.1% of the IR leading the agency rating. However, another result of their analysis is an increasing share of successfully anticipated rating changes approaching the actual announcement date. This behavior can only be observed for upgrades in the analysis of this work. However, Castellano and Giacometti (2012) did not observed this particular characteristic.

6.3 Liquidity limitations

Liquidity of CDS spreads is an important factor to consider. We therefore take the number of transactions that have been settled within a certain time interval as a proxy for CDS liquidity. The clearinghouse Depository Trust & Clearing Corporation (DTCC) aggregates CDS trading data of a variety of traders and issuers. The weekly reports include both gross and net notional positions as well as the total of settled contracts by each reference entity. The data sample contains the number of transaction, which have been made for every covered corporate in every week. This methodology is in analogy to Kapadia and Pu (2012) who use the number of contributors as a measure of liquidity. The sample is obtained from section IV of DTCC's Trade Information Warehouse Data.

For assessing the robustness of the model, we perform the same hit-miss-match analysis with a subsample only considering reference entities that had a weekly trading volume of more than 16 settled transactions, which is the median weekly transaction volume in the sample.

Table 7 reports the aggregated reclassification results for the analysis under the constraint that the IR is solely considered if at least 16 new trades, the median over all observations, have been made per week. Even though the share of IR hitting the agency ratings is at least equal and sometimes higher than without consideration of transaction volume, the difference is not significant from zero. As a result, the consideration of the trade volume as a proxy for the CDS liquidity has no significant influence on the capacity that the IR hit the agency rating. However, the distribution of the reclassification is different from Table 4. Whereas Table 4 shows that IRs are less strict the IR-AR-differences are distributed more equally under the restriction of the volume.

[Place Table 7 around here]

The results for the anticipation analysis under the restriction of volume are reported in Table 8. Again, the median volume of 16 trades is chosen as the minimum value. The share of anticipated downgrades increased significantly to 68.2% for one to four weeks prior to the rating change announcement. In contrast, this particular increase is not observed for upgrades. The share of successfully anticipated upgrades declines if transaction volume is taken into consideration. The analysis unveils that the number of contracts as proxy for liquidity has an influence on the anticipation power of IR. However, it is not observed in general that the anticipation power increases significantly.

[Place Table 8 around here]

Overall, the results provided in this section indicate that there is no statistical significant difference to our previous results with a liquidity restriction. Thus, liquidity has no immediate influence on the robustness of our model.

6.4 Example Best Buy

The evidence provided above shows that IRs can anticipate future rating changes. Figure 10 illustrates the movements of IR and agency rating of Best Buy. The green line outlines the development of the agency rating of Best Buy and the blue line indicate the IR. Best Buy has been downgraded three times in the observed period. Every downgrade is anticipated by the IR, because the numerical IR increases earlier than the agency rating. At this point, the time lag between agency rating and IR is visible. The rating adjustment of the CRA is delayed.

Furthermore, BestBuy did not only face agency rating downgrades by one notch, but also by two notches. Contrary, the IR adjusts the rating grade in smaller steps, i.e. only by one notch. Thus, the IR is able to predict rating changes by more than one notch as well. This confirms the hypothesis that IR reflects changes in creditworthiness faster than CRAs. However, the observed IR has a greater volatility and is less stable than agency ratings.

[Place Figure 10 around here]

Overall, this section provides evidence that the IR derived from the credit rating score is able to anticipate the agency ratings. Therefore, it is not only as accurate as agency ratings in determining the creditworthiness of an issuer, but also reflects changes in creditworthiness timelier. This anticipation is observed for both, upgrades and downgrades, already 16 weeks prior to the ratings change. Although the credit rating score is based on an estimation method, the results are comparable to those obtained by previous researchers who applied the non-parametric mapping process.

7. CONCLUSION

Despite issues concerning the timeliness and accuracy of credit ratings as well as a subjective rating methodology, market participants are highly dependent on CRAs and their opinions. With the introduction of credit derivatives and the rising prominence of CDS a new way of quantifying credit risk with market indicators emerged. Prior research has shown that CDS spreads are a viable substitution for credit ratings or at least providing a supplementary risk assessment. However, for example, a CDS spread of 100 bps is observed in seven different rating grades and therefore cannot be assigned to an agency rating unambiguously. The prevailing CDS-based model was relying on ratings of CRAs as they were needed for calibration purposes in the first place.

We have shown that the existing method derives IRs on the basis of CDS spreads and agency ratings by performing a non-parametric mapping process, which places boundaries optimally between two adjacent rating grades. Thereupon, we derived the general mathematical relationship between the established agency ratings and CDS-based boundaries. Combining this logarithmic relationship with assumptions about the distribution of credit ratings, CDS mid-spreads are translated into a rating scale, which market participants are familiar with and can easily being interpreted.

The provided model bears potential for implementation into risk monitoring processes of private institutions, such as lenders and asset managers, as well as regulators. The model extends the mechanism of Hart and Zingales (2011) to use CDS spreads also for capital requirements. Besides credit rating agencies CDS spreads can be used to measure the creditworthiness of corporates. Moreover, we show that IR can anticipate agency ratings and that changes in the creditworthiness of corporates are incorporated faster in the CDS market. The information in CDS spreads is not related to opinions of single rating agency analysts and incorporates only public available information.

ENDNOTES

[1] The resulting marginal shift of the curve has no influence on the general mathematical relationship. Instead of the average CDS spread, we could also use the lower border or upper border for each CDS spread boundary.

[2] The method is adapted from Castellano and Giacometti (2012).

References

Altman, E. I. (1968), "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance*, Vol 23 No. 4, pp. 589-609.

Altman, E. I. (2000), "Predicting financial distress of companies: Revisiting the Z-score and ZETA models", Stern School of Business, New York University.

Badaoui, S., Cathcart, L., and El-Jahel, L. (2013), "Do sovereign credit default swaps represent a clean measure of sovereign default risk? A factor model approach", *Journal of Banking & Finance*, Vol. 37 No 7, pp. 2392-2407.

Bannier, C. E., & Hirsch, C. W. (2010), "The economic function of credit rating agencies – What does the watchlist tell us?", *Journal of Banking & Finance*, Vol. 34 No. 12, pp. 3037-3049.

Blanco, R., Brennan, S., & Marsh, I. W. (2005), "An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps", *The Journal of Finance*, Vol. 60 No 5, pp. 2255-2281.

Boot, A. W., Milbourn, T. T., & Schmeits, A. (2006), "Credit ratings as coordination mechanisms", *Review of Financial Studies*, Vol. 19 No 1, pp. 81-118.

Breger, L. L., Goldberg, L., and Cheyette, O. (2003), "Market implied ratings: A simple approach to improve the classification of bonds and produce superior risk forecasts", *Risk Magazine*, July 2003.

Bühler, W., and Trapp, M. (2009), "Time-varying credit risk and liquidity premia in bond and CDS markets", Center for Financial Research, Working Paper No. 09-13.

Castellano, R., and Giacometti, R. (2012), "Credit default swaps: implied ratings versus official ones", *A Quaterly Journal of Operations Research*, Vol. 10 No. 2, pp. 163-180.

Cheung, S. (1996), "Provincial Credit Ratings in Canada: An Ordered Probit Analysis", Working paper 96-6, retrieved from Bank of Canada's website http://www.bankofcanada.ca/1996/04/working-paper-1996-6/.

Flannery, M. J., Houston, J. F., and Partnoy, F. (2010), "Credit default swap spreads as viable substitutes for credit ratings", *University of Pennsylvania Law Review*, pp. 2085-2123.

Forte, S., & Peña, J. I. (2009), "Credit spreads: An empirical analysis on the informational content of stocks, bonds, and CDS" *Journal of Banking & Finance*, Vol. 33 No 11, pp. 2013-2025.

Hart, O., & Zingales, L. (2011), "A new capital regulation for large financial institutions", *American Law and Economics Review*, Vol. 13 No 2, pp. 453-490.

Hull, J., Predescu, M., and White, A. (2004), "The relationship between credit default swap spreads, bond yields, and credit rating announcements", *Journal of Banking & Finance*, Vol. 28 No. 11, pp. 2789-2811.

Hunter, J. S. (1986), "The exponential weighted moving average", *Journal of Quality Technology*, Vol. 18 No. 4, pp. 203-210.

Jacobs, M., Karagozoglu, A.K. and Layish, D.N. (2016), "Credit risk signals in CDS market vs agency ratings", *The Journal of Risk Finance*, Vol. 17 No. 2, pp. 194-217.

Kapadia, N., & Pu, X. (2012), "Limited arbitrage between equity and credit markets", *Journal of Financial Economics*, Vol. 105 No 3, pp. 542-564.

Kiesel, F. (2016), "Do investors still rely on credit rating agencies? Evidence from the financial crisis.", *The Journal of Fixed Income*, Vol. 25 No. 4, pp. 20-31.

Kou, J., and Varotto, S. (2005), "Predicting agency rating movements with spread implied ratings", *ISMA Centre Discussion Papers in Finance, DP2005-06, ISMA Centre, University of Reading, UK*.

Kou, J., and Varotto, S. (2008), "Timeliness of spread implied ratings", *European Financial Management,* Vol. 14 No. 3, pp. 503-527.

Longstaff, F. A., Mithal, S., & Neis, E. (2005). Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. The Journal of Finance, 60(5), 2213-2253.

Löffler, G. (2004), "An anatomy of rating through the cycle". *Journal of Banking & Finance*, Vol. 28 No. 3, pp. 695-720.

Mayordomo, S., Peña, J. I., & Schwartz, E. S. (2014). Are all credit default swap databases equal?. *European Financial Management*, *20*(4), 677-713.

Norden, L., & Weber, M. (2004), "Informational efficiency of credit default swap and stock markets: The impact of credit rating announcements", *Journal of Banking & Finance*, Vol. 28 No 11, pp. 2813-2843.

Reyngold, A., Kocagil, A. E., Gupton, G. M., and DiGiambattista, J. (2007), "Fitch CDS Implied Ratings (CDS-IR) Model", *Quantitative Financial Research Special Reports, Fitch Ratings*.

Zhu, H. (2006), "An empirical comparison of credit spreads between the bond market and the credit default swap market", *Journal of Financial Services Research*, Vol 29 No 3, pp. 211-235.

About the Author

FLORIAN KIESEL is a postdoctor- al researcher at the Institute of Corporate Finance at Technische Universität Darmstadt. His primary research interest is in the area of corporate finance and empirical finance with a special focus on credit default swaps and credit ratings.

Table 1: Agency rating grades assigned to a numerical integer and the a discrete credit rating score

Rating grade	Numerical	Credit rating score
AAA	1	0...1.49
AA+	2	1.5...2.49
AA	3	2.5...3.49
AA-	4	3.5...4.49
A+	5	4.5...5.49
A	6	5.5...6.49
A-	7	6.5...7.49
BBB+	8	7.5...8.49
BBB	9	8.5...9.49
BBB-	10	9.5...10.49
BB+	11	10.5...11.49
BB	12	11.5...12.49
BB-	13	12.5...13.49
B+	14	13.5...14.49
B	15	14.5...15.49
B-	16	15.5...16.49
CCC/C	17	>16.5

Note: This table shows the relation between the Fitch rating and the discrete numerical rating.

Table 2: Hit-miss-match matrix as of week 29 in 2010.

		Implied Rating																
		AAA	AA+	AA	AA-	A+	A	A-	BBB+	BBB	BBB-	BB+	BB	BB-	B+	B	B-	C/ CCC
AGENCY RATING	AAA	0.00%	0.00%	0.00%	0.00%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	AA+	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	AA	0.00%	0.00%	0.00%	0.00%	0.34%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	AA-	0.00%	0.00%	0.00%	0.00%	0.00%	0.68%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	A+	0.00%	0.00%	0.00%	0.00%	2.74%	1.71%	1.03%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	A	0.00%	0.00%	0.00%	0.00%	1.71%	3.42%	2.05%	1.03%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	A-	0.00%	0.00%	0.00%	0.00%	0.34%	3.42%	6.51%	1.71%	1.71%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	BBB+	0.00%	0.00%	0.00%	0.00%	0.00%	4.45%	3.42%	3.42%	3.08%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	BBB	0.00%	0.00%	0.00%	0.00%	0.00%	1.71%	4.11%	3.08%	6.51%	1.71%	0.00%	0.68%	0.00%	0.00%	0.00%	0.00%	0.00%
	BBB-	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	1.03%	1.71%	4.45%	2.74%	1.71%	0.00%	0.68%	0.00%	0.00%	0.00%	0.00%
	BB+	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	1.71%	2.74%	1.71%	1.03%	0.00%	0.00%	0.00%	0.00%
	BB	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.68%	1.03%	1.03%	1.37%	0.68%	0.34%	0.00%	0.00%	0.00%
	BB-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.71%	1.03%	0.00%	0.00%	0.00%	0.00%
	B+	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.68%	1.37%	0.00%	0.34%	0.00%	0.00%
	B	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.68%	0.68%	0.68%	0.00%	0.00%
	B-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.68%	1.03%	1.37%	0.34%
	C/ CCC	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	1.03%

Table 3: Hit-miss-match matrix as of week 27 in 2014.

		Implied Rating																
		AAA	AA+	AA	AA-	A+	A	A-	BBB+	BBB	BBB-	BB+	BB	BB-	B+	B	B-	C/ CCC
AGENCY RATING	AAA	0.00%	0.00%	0.00%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	AA+	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	AA	0.00%	0.00%	0.00%	0.34%	0.00%	0.68%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	AA-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	A+	0.00%	0.00%	0.00%	0.00%	1.69%	1.01%	0.68%	0.34%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	A	0.00%	0.00%	0.34%	1.01%	0.68%	1.69%	3.72%	1.35%	0.68%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	A-	0.00%	0.00%	0.00%	0.00%	1.01%	1.69%	4.05%	3.04%	2.03%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	BBB+	0.00%	0.00%	0.00%	0.34%	0.68%	2.70%	2.36%	4.39%	2.03%	2.70%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	BBB	0.00%	0.00%	0.00%	0.00%	0.00%	0.68%	1.35%	6.76%	7.09%	1.35%	1.35%	0.34%	0.00%	0.00%	0.00%	0.00%	0.00%
	BBB-	0.00%	0.00%	0.00%	0.00%	0.34%	0.00%	0.00%	2.36%	5.07%	3.04%	2.36%	2.03%	0.34%	0.00%	0.34%	0.00%	0.00%
	BB+	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	1.01%	3.38%	1.69%	0.34%	0.00%	0.00%	0.00%	0.00%
	BB	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.68%	0.34%	0.34%	1.69%	0.68%	0.00%	0.00%	0.00%	0.00%
	BB-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	0.34%	1.35%	1.69%	0.34%	0.00%	0.00%	0.00%
	B+	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	1.35%	1.69%	0.00%	0.00%	0.00%	0.00%
	B	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	1.35%	0.34%	0.34%	0.00%	0.00%
	B-	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	1.35%	0.00%	0.00%	0.00%
	C/ CCC	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.34%	0.00%	0.34%

Table 4: Deviation of implied ratings from agency ratings as average for the years 2010-2014.

	Percentages of ratings in percent				
notches	2010	2011	2012	2013	2014
<-2	5	4	5	6	5
-2	12	10	15	13	13
-1	24	24	19	22	22
0	31	28	29	22	25
1	15	20	18	22	20
2	8	10	9	9	9
>2	4	5	7	6	6

Note. Notches mean a negative or positive deviation (called reclassification) whereby negative indicates that IR model indicates a higher creditworthiness than the agency rating suggested

Table 5: Mean difference between IR and AR prior to rating changes.

	Time interval prior to the AR change in weeks				
	[-1,- 4]	[-5,-8]	[-9,-12]	[-13,-16]	*n*
Downgrades	1.05**	1.08**	1.05**	1.03**	173
Upgrades	-1.10**	-1.05**	-1.02**	-0.97**	163

Note. ** p<.01 indicates a statistical significance of 0.01% level. The mean difference between IR and AR is calculated as average notches of deviation prior to an upcoming rating change.

Table 6: Lead analysis for implied ratings.

Panel A. Without volume restriction					
	Time interval prior to the AR change in weeks				
Leading ratings in %	[-1,- 4]	[-5,-8]	[-9,-12]	[-13,-16]	*n*
Downgrades	60.3	61.5	62.6	60.4	173
Upgrades	64.2	61.8	61.8	60.5	163

Note. The share ranges between 0% indicating that no rating change has been anticipated and 100% indicating that every rating change has been anticipated by the implied rating.

Table 7: Deviation of implied ratings from agency ratings as average for the years 2010-2014 with liquidity restriction.

	Percentages of ratings in percent				
notches	2010	2011	2012	2013	2014
<-2	3	2	3	4	6
-2	11	7	11	11	13
-1	23	21	15	19	21
0	31	28	27	25	27
1	18	23	21	21	20
2	11	12	11	10	9
>2	3	8	12	9	4

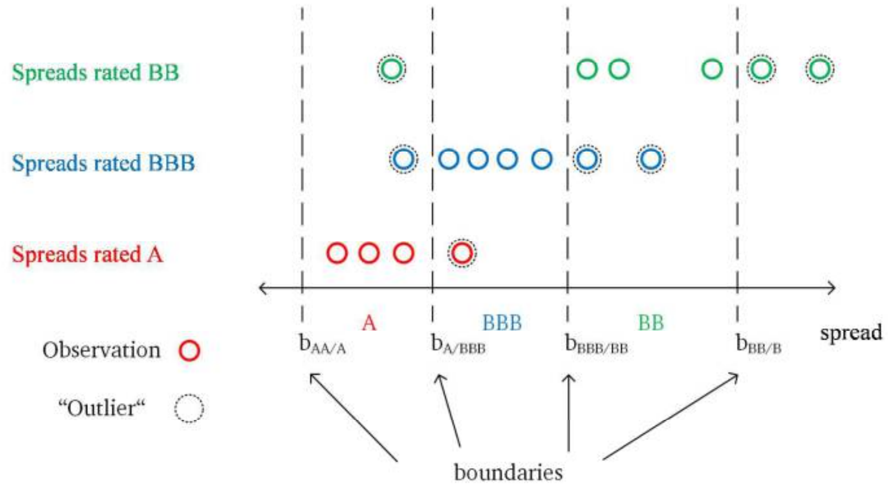
Note. Notches mean a negative or positive deviation (called reclassification) whereby negative indicates that IR model indicates a higher creditworthiness than the agency rating suggested

Table 8: Lead analysis for implied ratings with volume restrictions as proxy for liquidity.

	Time interval prior to the AR change in weeks				
Leading ratings in %	[-1,- 4]	[-5,-8]	[-9,-12]	[-13,-16]	*n*
Downgrades	68.2	64.5	69.0	65.5	110
Upgrades	62.1	62.7	60.9	54.4	87

Note. The share ranges between 0% indicating that no rating change has been anticipated and 100% indicating that every rating change has been anticipated by the implied rating.

Figure 1: Schematic setting of boundaries for rating grades.



Note: This figure shows the schematic exemplary depiction of the boundary setting mechanism. Each point represents a CDS spread and the corresponding spread rating.

Figure 2: Development of descriptive statistics of the CDS sample data.

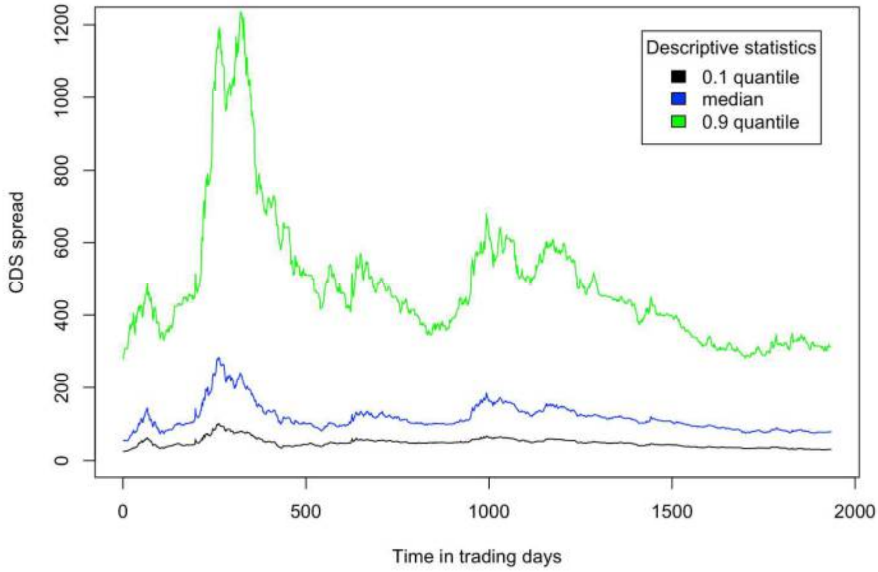


Figure 3: CDS spread of Johnson & Johnson with EWMA for the time window of 250 trading days.

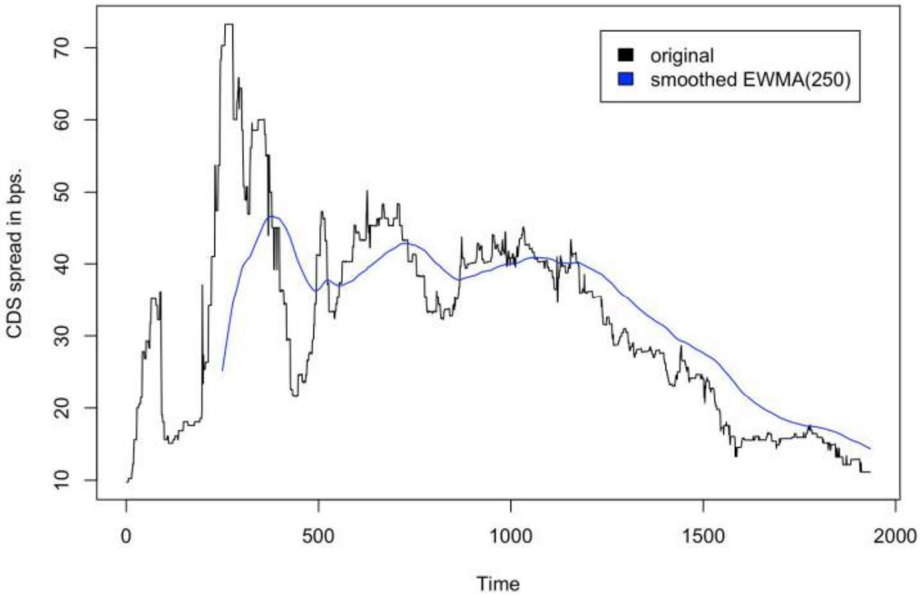
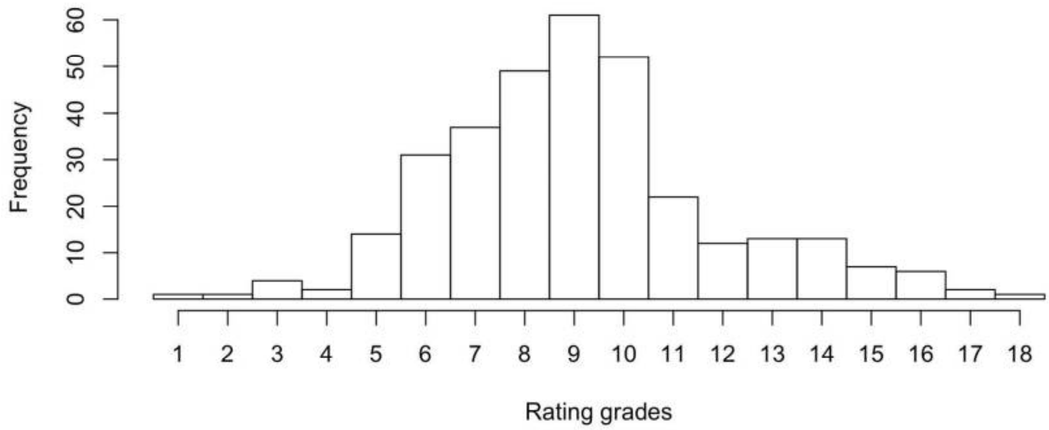
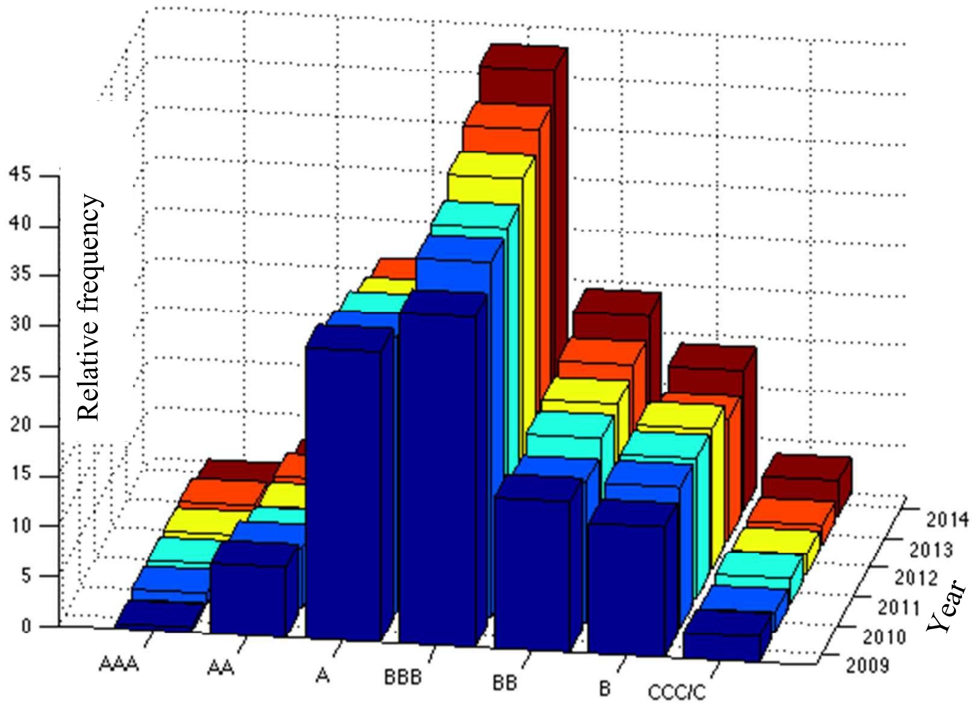


Figure 4: Absolute frequency of rating grades as of July 1, 2014.



Note. Absolute frequency is plotted over numerical integer ratings grades, whereby 1 is the numerical representation of AAA, 2 means AA+ etc.

Figure 5: Historical distribution of agency ratings by year from 2009 to 2014.[1]



Note. By plotting relative frequency in per cent over the respective rating grade category and adding the time dimension, the changes in distribution are visualized.

[1] The data is obtained from the Needham et al. database.

Figure 6: Agency ratings over corresponding CDS spreads.

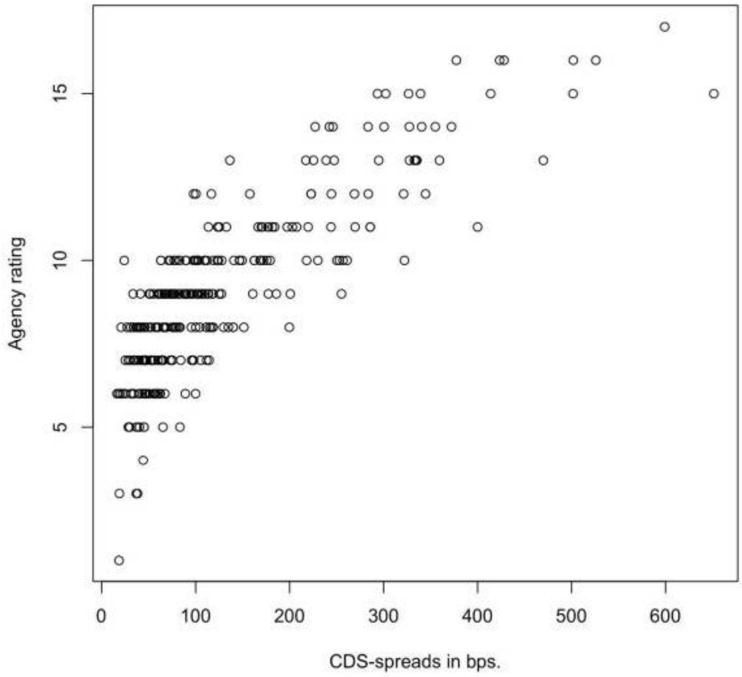
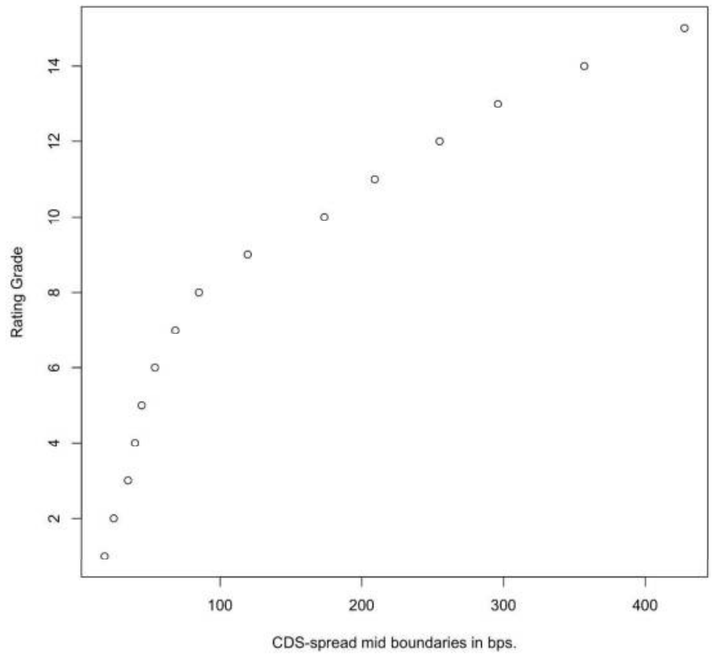


Figure 7: Rating grades over CDS spread-mid boundaries as of week 27 in 2014.



Note. CDS spread boundaries are calculated using the method of Reyngold et al. (2007)

Figure 8: Linear regression analysis of CDS spread boundaries and rating grades.

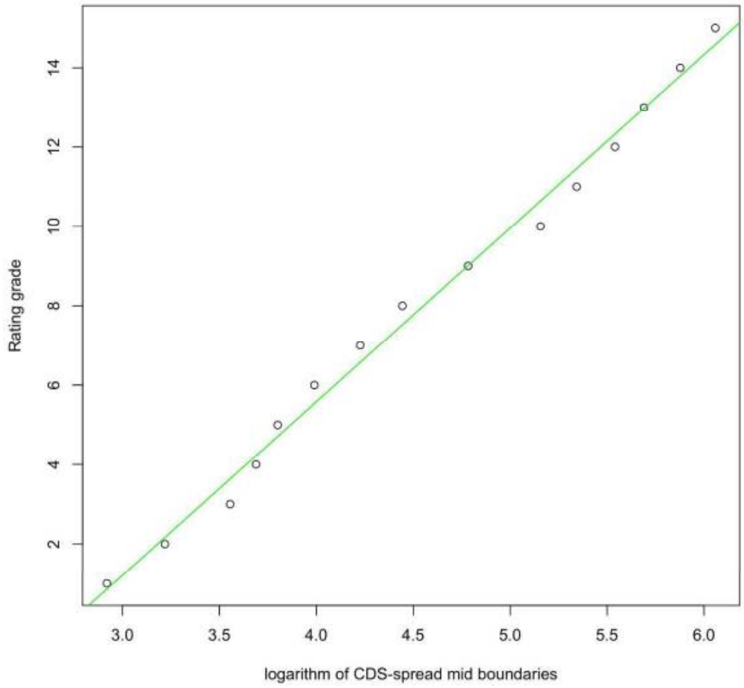
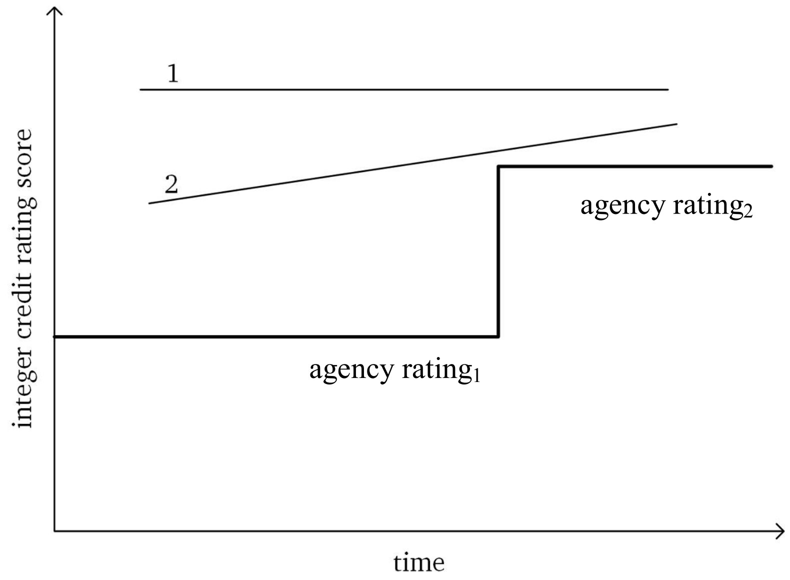
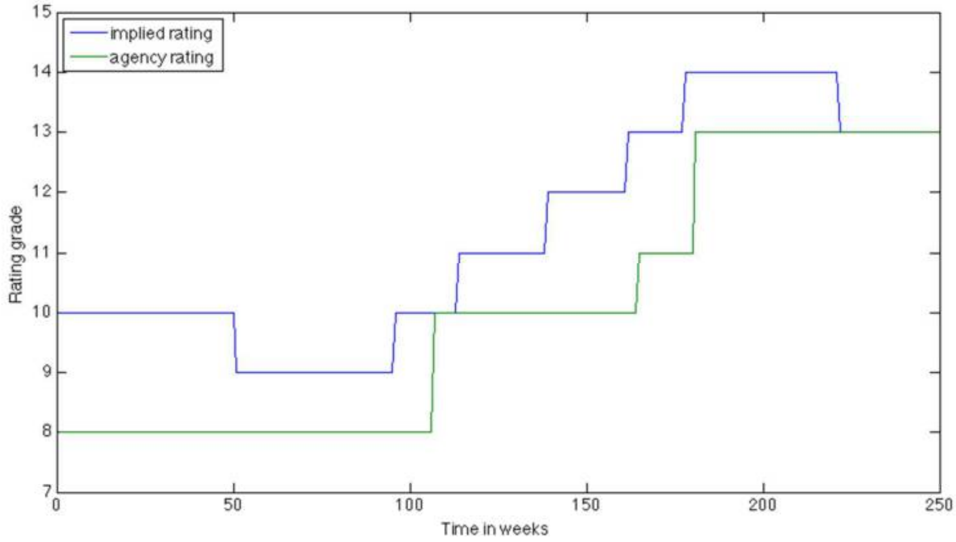


Figure 9: Schematic lead behavior of implied ratings.



Note. Adapted from Kou and Varotto (2008). The visualization shows an exemplary downgrade and two possible implied rating directions. (1) IR is already indicating a worse rating grade or (2) IR shows a trend for a downgrade coming from a better rating grade.

Figure 10: Relative behavior of implied rating and agency rating for Best Buy.



Note. This figure shows the implied rating and the Fitch rating of Best Buy. The figure shows that the CDS spread implied rating reacts earlier than the Fitch rating.

Appendix A - Descriptive statistics of the CDS data sample

Rating grade	Frequency in %
AAA	0.3
AA+	0.0
AA	1.0
AA-	0.3
A+	4.4
A	9.4
A-	11.7
BBB+	15.4
BBB	18.8
BBB-	15.8
BB+	6.0
BB	4.4
BB-	4.0
B+	3.7
B	2.3
B-	1.3
CCC/C	1.0

Note. Agency ratings of reference entities as of week 23 of 2014

Country of origin	Number of reference entities
Belgium	1
Denmark	2
Finland	4
France	15
Germany	18
Ireland	4
Italy	5
Luxembourg	2
The Netherlands	6
Portugal	2
Spain	5
Sweden	2
Switzerland	1
UK (incl. Bermuda and Jersey)	36
US	207