Utilizing news topics for credit risk management: the explanation of bank CDS spreads

Jan Roeder, Matthias Palmer, Jan Muntermann

Angaben zur Veröffentlichung / Publication details:

Roeder, Jan, Matthias Palmer, and Jan Muntermann. 2020. "Utilizing news topics for credit risk management: the explanation of bank CDS spreads." *Journal of Decision Systems* 29 (Suppl 1): 32–44. https://doi.org/10.1080/12460125.2020.1848384.

Utilizing News Topics for Credit Risk Management: The Explanation of Bank CDS Spreads

Jan Roeder, Matthias Palmer and Jan Muntermann

Chair of Electronic Finance and Digital Markets, University of Goettingen, Goettingen, Germany

ABSTRACT

Monitoring the default risk of banks is highly relevant for companies when they enter into derivative contracts with banks to hedge risks. These contracts are subject to the counterparty risk of the bank defaulting. In credit risk management, market-based measures such as credit default swap (CDS) spreads are used to continuously assess the credit risk of counterparties. At the same time, CDS spreads are difficult to interpret and explain. We investigate which financial news topics are positively or negatively related to bank CDS spreads. Our findings indicate that especially negative news topics are significantly associated with changes in CDS spreads. These include not only financial aspects such as a decrease in dividends or revenue guidance but also non-financial topics such as sanctions and legal issues. Thereby, our understanding of which topics addressed in financial news are relevant for supporting decision-making in the context of credit risk management is improved.

KEYWORDS

Credit risk management; CDS spreads; news topics; sentiment analysis

1. Introduction

Companies such as manufacturers or retailers use derivative contracts to hedge risks, e.g., the volatility of exchange rates or fluctuations in raw material prices. Banks, who act as buyers and sellers of these contracts, are subject to default risk themselves. With regards to these counterparties (i.e., banks), it is possible to obtain a timely risk assessment using specific market-based risk measures such as credit spreads or credit default swap (CDS) spreads (Van Gestel & Baesens, 2008). However, it can be difficult for managers, e.g., in treasury departments, to track and understand changes in these quantitative risk measures, especially when a broad portfolio of counterparties (i.e., banks) is monitored. This is particularly challenging if the credit risk exposure needs to be adjusted spontaneously due to deteriorations in the creditworthiness of the respective counterparties. Apart from that, an understanding of the reasons for changes in credit risk is relevant for reports to senior management.

While decision support systems (DSS) are already an essential part of credit risk assessment by companies and banks, the growing amount of unstructured data remains underutilised in many aspects. In addition to corporate publications, regulatory filings, analyst opinions, and social media data, news is supposedly the most easily accessible and

CONTACT Jan Roeder jan.roeder@uni-goettingen.de

often most up-to-date source of information (Mitra & Mitra, 2011; Tsai et al., 2016). Given the large amount of news and related sources, we consider it important to analyse and identify news topics that are of particular relevance. This is key to making a meaningful selection of those news topics that should be prepared for decision-makers and decision documentation.

Amongst others, a recent study indicates that CDS spreads of both European and US banks can be an indicator of financial distress (Avino et al., 2019). In a news context, Smales (2016) shows that there is an asymmetric negative correlation between news and CDS spreads of banks, with negative news having a stronger effect on CDS spreads. Yang et al. (2019) conduct their analysis in a similar setting but with a focus on corporate CDS spreads. In addition to finding that news reduces information asymmetries, especially in times of crisis, this study elaborates that the negative correlation between news sentiment and CDS spreads is particularly evident when there is negative news, news on firm fundamentals, and unexpected news. Nevertheless, none of the previous studies has focused on a broad range of news topics and on how they are associated with changes in bank CDS spreads although this is relevant for supporting decision-making in the domain of credit risk management. Against this background, we pose the following research question:

Which topics addressed in financial news are positively or negatively related to bank CDS spreads?

Considering the research question, our paper is structured as follows: Section 2 contains the theoretical background related to credit risk management and news analytics related to CDS spreads. Against this background, we formulate research hypotheses for the subsequent analysis. In section 3, we describe and discuss the research setup. This includes a first look at the data set. Furthermore, we present our analysis approach, in which we analyse the relationship between bank CDS spreads and news characterised by different topics. Section 4 provides the empirical results showing which news topics are particularly well suited for explaining returns of bank CDS spreads. The paper concludes with a discussion of the implications of the results and limitations of the research approach in section 5.

2. Theoretical background

2.1. Credit risk management and credit default swaps

Credit risk is the risk that an entity is unable or unwilling to repay an agreed upon amount of debt within a certain time period (Van Gestel & Baesens, 2008). To avoid putting a company's own business at risk, managing credit risks plays a decisive role for companies in all industry sectors. To hedge this risk, various derivative instruments can be used. Credit derivatives allow a credit risk to be passed on to other market participants, for example, in exchange for a premium that is paid (Hull et al., 2004). Since these instruments are traded mostly over-the-counter with banks and other financial institutions, counterparty risk exists in these business relationships (McNeil et al., 2015). CDS are a specific credit derivative and can be regarded as an insurance against the default or failure to pay of a reference entity (Lando, 2004). Since the CDS spread (the premium) can be understood as the cost of insuring against a credit event, CDS spreads and equity prices tend to be correlated negatively (Hull et al., 2004). Apart from the analysis of balance sheet ratios or the use of credit ratings, CDS spreads can represent a market-based component of a comprehensive credit risk modelling system. Moreover, studies suggest that in comparison to bond markets, CDS markets reflect the assessments of market participants more quickly (Blanco et al., 2005). Additionally, CDS are determined less by factors not directly related to credit risk such as interest rate risk or short-term illiquidity (Andres et al., 2016). Galil and Soffer (2011) examine the reaction of CDS spreads to rating changes and show that the reaction to rating downgrades is stronger than for rating upgrades. Overall, various determinants of CDS spreads have been analysed in prior research (Chiaramonte & Casu, 2013; Collin-Dufresn et al., 2001; Donovan et al., 2018; Galil et al., 2014).

2.2. News analytics for CDS spreads

In order to generate a comparative advantage, text data and especially news data are frequently utilised to forecast stock prices or exchange rates (Nassirtoussi et al., 2014; Xing et al., 2017). In contrast, forecasting credit risk using unstructured data plays a comparatively minor role in the academic literature thus far. In this respect, research is particularly promising which pursues long-term strategies and supplements existing valuations or risk profiles with alternative data, e.g., unstructured or satellite data, and puts them into a richer context (Monk et al., 2019). For news data, it can be difficult to determine the importance of news topics for credit risk. Due to the sheer amount of news data available, the problem of information overload becomes apparent. Therefore, it is important to carefully select relevant news items and topics.

For news analysis, sentiment analysis and topic modelling are methods of choice. By applying sentiment analysis, it has repeatedly been shown that particularly negative sentiment values can be associated with stronger capital market reactions (Kearney & Liu, 2014). By using advanced machine learning for short texts, financial news can nowadays be classified better and used for predictive modelling (Hagenau et al., 2013; Mitra & Mitra, 2011), and for predicting bank distress (Cerchiello et al., 2017). The advantage of topic modelling lies in the interpretability the output provides, as it is possible to examine which latent topics are discussed in the texts. This can increase the predictive power for models in the financial domain (Cerchiello & Nicola, 2018).

News data providers, e.g., RavenPack or Thomson Reuters, offer sentiment scores and topic classifications for individual financial news. Building on this kind of data, J. B. Kim et al. (2018) combine CDS spreads with news sentiment and analyse managers' voluntary disclosure. Additionally, existent studies explore the relationship between news data, corporate bond liquidity (Jiang & Sun, 2015) and stock prices (Jiang et al., 2019). The relationship between sovereign CDS spreads and financial news is analysed by several studies, including Erlwein-Sayer (2018), Apergis (2015), Chebbi and Sarraj (2017), and S.-J. Kim et al. (2015). Regarding CDS spreads of companies, Lu et al. (2012) successfully model changes in company creditworthiness using Wall Street Journal news and financial ratios. Tsai et al. (2016) show that higher media coverage and negative news sentiment increase corporate credit risk measured by CDS spreads. Liebmann et al. (2016) examine the different interpretations of financial news for stock and CDS traders. In this case, the news data sample is limited to news related to cash flow changes and default risk. Yang et al. (2019) use a dataset that is categorised by news types to show that there is a negative correlation between news sentiment and CDS spreads, especially for negative, fundamental, and unexpected news. Two papers analyse the correlation between bank CDS spreads and news in somewhat different contexts but both find a negative correlation between news and CDS spreads (Avino et al., 2019; Smales, 2016). However, to the best of our knowledge, there is no existing research on which news topics possess explanatory power regarding bank CDS spreads. To address this research gap, we propose the following two hypotheses:

H 1: *Topics of* ***positive*** *bank-related financial news are associated with a* ***decrease*** *in CDS spreads.*

H 2: *Topics of* ***negative*** *bank-related financial news are associated with an* ***increase*** *in CDS spreads.*

3. Research setup

3.1. Data and descriptive statistics

First, we collected CDS spreads from Thomson Reuters Eikon at a daily level. We use the *mid spread*, defined as the midpoint between bid and ask for the respective CDS. Our analysis covers a period ranging from 9/1/2009 to 5/31/2018. We focus on banks, based on their described relevance for a wide range of companies. The selected banks are constituents of the KBW Nasdaq Global Bank Index (GBKX). The index tracks systemically important banks as classified by the Financial Stability Board (FSB) and the Basel Committee on Banking Supervision (Nasdaq, 2019). Analogous to stock prices, where returns are calculated to obtain a stationary time series, returns are also used in the case of CDS spreads (Liebmann et al., 2016). Table 1 provides an overview of CDS spreads (in basis points) and CDS spread returns for our data set. Whether a credit event is formally triggered by a company restructuring depends on the regional regulation. *No restructuring* (X = R), which is standard in North America, does not consider restructuring to be

Table 1. Descriptive statistics per bank for news and CDS spreads for the period from 9/1/2009 to 5/31/2018.

Bank name	CDS spread mean	CDS spread return mean	Restructuring standard	Term	News count
BNP Paribas	105.566	0	M = R	5	2,381
Banco Santander	181.072	0	M = R	5	2,289
Bank of America	123.905	-0.001	X = R	5	8,042
Barclays	108.438	0	M = R	5	5,336
Citigroup	119.093	-0.001	X = R	5	7,676
Credit Suisse	98.38	0	M = R	5	4,859
Deutsche Bank	117.986	0	M = R	5	5,121
HSBC	88.617	0	M = R	5	4,980
ING Group	104.275	0	M = R	5	2,185
JPMorgan Chase	77.701	-0.001	X = R	5	9,196
Morgan Stanley	145.44	-0.001	X = R	5	5,505
Societe Generale	139.126	-0.001	M = R	5	1,446
Standard Chartered	107.117	0	M = R	5	1,805
UniCredit	217.256	0	M = R	5	1,592
Wells Fargo	70.841	0	X = R	5	6,237

a credit event. In Europe, *modified-modified restructuring* (M = R) is used, which includes a clause to limit opportunistic behaviour (Packer & Zhu, 2005). In all cases, 5-year CDS are examined, as these tend to be the most liquid (Breitenfellner & Wagner, 2012).

The financial news data is extracted from RavenPack News Analytics. This database condenses unstructured data from various financial news portals (e.g., Dow Jones Newswires, Reuters, Ticker Report) into structured data points. Among others, the following aspects are quantified: the individual news sentiment of specific news, an aggregated news sentiment over a 91-day window, the news volume, the news novelty, and a news impact projection on the stock market. To improve the signal-to-noise ratio, we have only selected news items that have a relevance of 100 (maximum) and a novelty value of at least 50 (range: 0 to 100). Furthermore, a classification according to a hierarchical taxonomy of news topics is carried out. This classification is central to our analysis, as it adds an additional degree of interpretability to the primarily quantitative variables. This taxonomy is composed of a set of news topics [T] and corresponding news sub-topics [ST]. Before carrying out the actual analysis, it is necessary to evaluate which of the topics and sub-topics can be included. Two aspects were considered: First, the content-related relevance of a specific topic must be apparent. For example, news reports on debt are relevant when assessing the risk of a bank. In a wider sense, a topic such as the excessive compensation of managers could be interpreted negatively by the market and act as a deterrent to customers. Therefore, we also consider such subjects. The second key aspect is data availability. To analyse a topic, a sufficient amount of data points is required across the different banks. For this purpose, we aggregated the topics and sub-topics. An excerpt is shown in Table 2. In many cases, the sub-topics have an adequate number of data points. However, the example [T] Business contract shows a case where the selection of the sub-topic [ST] would have led to a drastic reduction of data points, which is why the topic [T] is selected in such a case. A complete list of the selected topics and sub-topics is shown in the regression table (Table 4).

Another step towards understanding the relationship between quantitative financial news data and credit risk metrics in the form of CDS spreads is to visualise the relationship between CDS spread returns and the quantitative news variables we described above, provided by RavenPack News Analytics. The quantitative data was normalised using Z-scores, i.e., the mean is 0 and the standard deviation is 1. The scores were calculated per company to consider differences at the company level. Both scatterplots and histograms for selected variables are shown in Figure 1. The CDS spread returns and the individual news sentiment show a symmetrical distribution. In contrast, the news novelty shows a left-

Table 2. Illustration of the news topic hierarchy.

News topic [T]	News topic [T] count	News sub-topic [ST]	News sub-topic [ST] count
Stock	11,953	Gain	6,756
		Loss	5,197
Earnings	11,750	Up	2,287
		Down	1,684
		Above expectation	1,270
		Below expectations	491
		…	
…		…	
Business contract	7,562	Terminated	178
		…	
…		…	

skewed distribution (i.e., some news is repeated often) and the news impact projection a right-skewed distribution (i.e., some news shows an extraordinary impact on the market).

For further illustration, the occurrences of selected topics over time are shown for the example of Banco Santander. In Figure 2, the upper/lower graph shows the CDS spreads/ CDS spread returns (aggregated for two weeks) from 9/2009 to 5/2018. The vertical lines indicate the number of topic mentions, which are normalised per topic. For example, we see that the number of times the topic *acquisition* is mentioned increases between 2010 and 2012. Additionally, *regulatory investigations* appear more often in 2013. It becomes apparent that the topic occurrence over time shows significant variation depending on the topic.

3.2. Analysis

We perform a regression analysis to investigate the relationship between news topics and changes in CDS spreads. The dependent variable is the logarithmic daily return of the CDS spread (ΔCDS). Since the analysis is supposed to take place at the daily level, we are faced with challenges due to the nature of the CDS spread data. Compared to stock prices, the

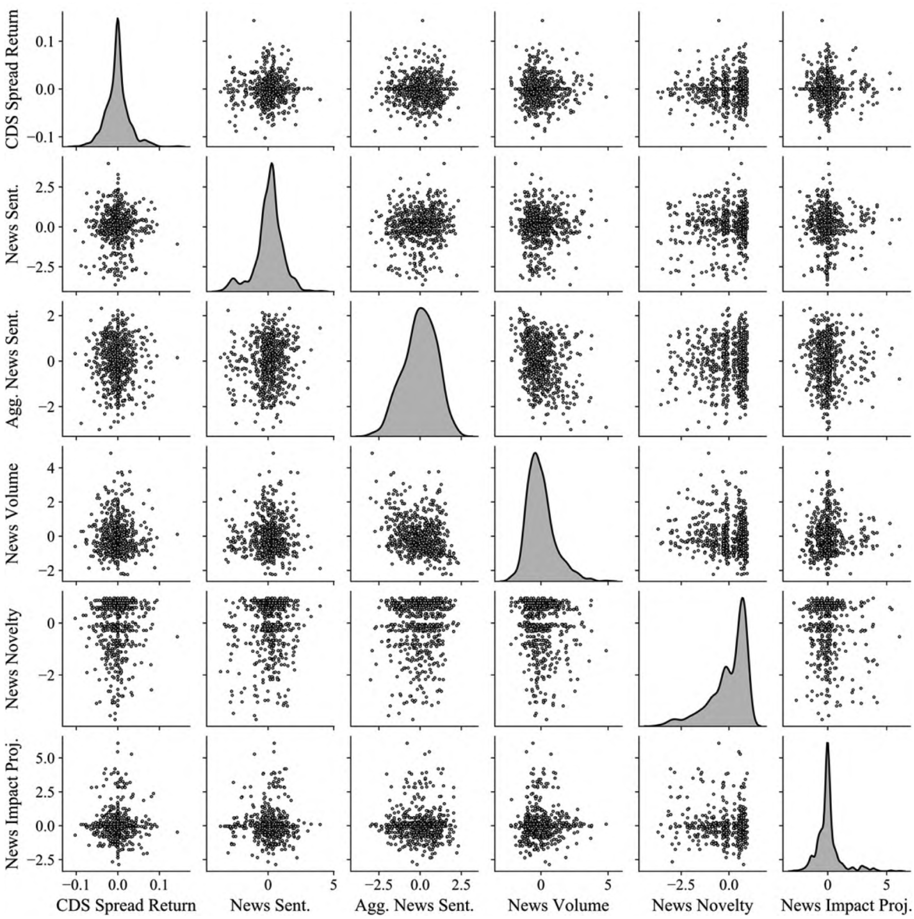

Figure 1. Relationships between CDS spread returns and quantitative news-related variables.

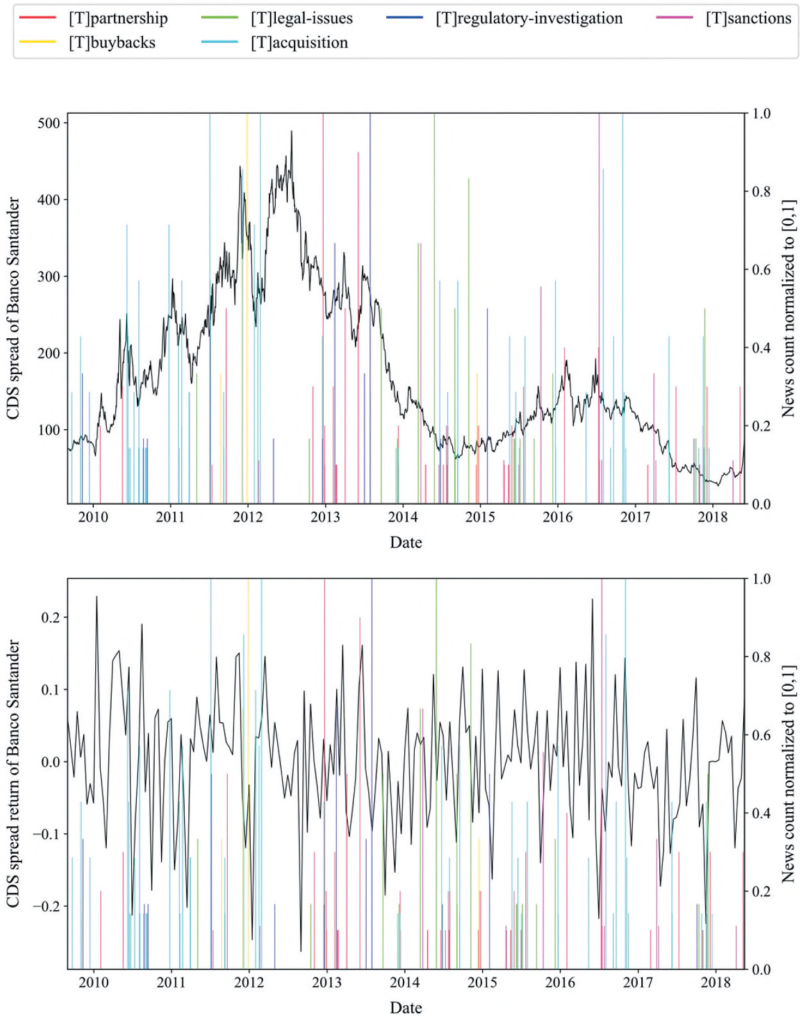

Figure 2. Occurrences of selected topics mentioned in financial news visualised for Banco Santander.

prevalence of stale values, i.e., values that remain constant for multiple days, is comparatively higher. In addition, in the early 2010s occasionally days with missing values occur. For this purpose, we set values that do not change for three days to none and then perform a linear interpolation where the timeframe is constrained to five periods. Table 3 contains the relevant variables ordered by subject area. A short definition is also included.

The CDS spreads for bank *i* on date *t* are regressed on the news-related variables. *NEWSTOPICS* represents the counts of all considered news topics. The related quantitative data such as the sentiment is represented via *NEWSQUANT*. The choice of variables related to equity markets is partially inspired by Yang et al. (2019) and subsumed under *MARKET*. Additionally, *CONTROLS* contains dummy-coded variables for each year and bank. Lagged instances of the dependent variable are included for *n* lags as *LAGS*. After assessing different lags with the autocorrelation function, we found lags of up to three days to be a good fit.

Table 3. Summary of variables and their definition.

Variable	Definition
Dependent variable	
CDS	Logarithmic return of the CDS spread
News variables quantitative	
News sentiment	Sentiment score of a specific news article
Aggregated news sentiment	Aggregated news sentiment over a 91-day window
News volume	Volume of news events for the bank over a rolling 91-day window
News novelty	Quantifies the novelty of a news article
News Impact Projection	Proprietary measure calibrated to predict the market reaction in the following two hours
Market-based variables	
Stock return	Logarithmic return of the return index of the bank
Volatility index	Change in implied volatility
Market to book	Logarithm of the ratio of market value and book value

$$\Delta CDS_{i,t+1} = \ln(CDS_{i,t+1} \quad - \ln(CDS_{i,t}$$

$$LAGS_{i,t} = [\Delta CDS_{i,t}, \Delta CDS_{i,t-1}, \ldots, \Delta CDS_{i,t-n}]$$

$$\begin{aligned} \Delta CDS_{i,t+1} &= \beta_0 + \beta_1^T NEWSQUANT_{i,t} + \\ &\beta_2^T NEWSTOPICS_{i,t} + \beta_3^T MARKET_{i,t} + \\ &\beta_4^T CONTROLS_{i,t} + \beta_5^T LAGS_{i,t} + \varepsilon_{i,t+1} \end{aligned}$$

4. Results

The results of the multivariate regression analysis are presented in Table 4. It is apparent that the topics associated with negative developments clearly dominate the statistically significant variables. Firstly, there are non-financial topics. *Sanctions* and *legal issues* are particularly important. Their positive coefficients indicate that an increase in news related to these topics is accompanied by a rise in the CDS spread return of the following day. Since the topic and sub-topic variables are measured by the number of topic-related news, the estimated parameter can be understood as follows: one additional news article regarding *sanctions* is associated with an increase in CDS spread return of 0.04 percentage points (%). Market participants perceive an increased risk of default and the price to insure against the bank's default rises, as the increased CDS spreads suggest. This observation seems to be evident across the broad range of banks analysed. Among the financial topics, *reduced dividends, downgraded price targets*, and *revenue guidance down* show statistically significant coefficients. Changes in these variables appear to correlate with market participants' risk perception of default risk.

A negative coefficient was mainly shown for news that report the *price target* as *set*. The negative coefficient means that more news on the subject is associated with a lowered perceived risk of default. While it is intuitive that setting the price target can help to reduce the uncertainty that the market participants experience, it would be expected that more positive news topics are significantly associated with CDS spread returns. Regarding *positive analyst rating change* and *earnings below expectation* we find coefficients with signs that are opposite to what would be assumed.

Table 4. Regression table with parameter estimates and standard deviation reported.

Covariates Topic [T]	Sub-topic [ST]	Estimate	Covariates Topic [T]	Sub-topic [ST]	Estimate
acquisition		-0.0349 (0.0264)	partnership		-0.0177 (0.0267)
analyst rating change	negative	0.0084 (0.0357)	pre-tax earnings	down	0.1429 (0.098)
	neutral	-0.057 (0.0709)		up	-0.0673 (0.1557)
	positive	0.0614* (0.0003)	price target	downgrade	0.223** (0.1078)
asset		0.0622 (0.0422)		set	-0.3871*** (0.1302)
business contract		0.025 (0.0181)		upgrade	0.0309 (0.0519)
buybacks		-0.0303 (0.0606)	product release		-0.0166 (0.0192)
credit rating change	affirmation	-0.0246 (0.016)	regulatory investigation		-0.0139 (0.0299)
	downgrade	0.0032 (0.0255)	revenue guidance	down	0.2756** (0.1304)
	set	0.0295 (0.0338)		up	-0.0023 (0.1045)
	upgrade	-0.0682 (0.0441)	revenue	down	0.0275 (0.0775)
debt	increase	0.0757 (0.0947)		up	-0.0221 (0.0774)
	reduction	-0.1237 (0.1646)	sanctions		0.0454** (0.0194)
dividend	down	0.4674** (0.2019)	stock	gain	0.0277 (0.0227)
	up	-0.0677 (0.0631)		loss	-0.0117 (0.0284)
earnings per share	above expect.	-0.005 (0.1132)	unit acquisition		-0.0194 (0.0186)
	below expect.	0.1856 (0.2984)	aggregated news sentiment		-0.0002 (0.0013)
earnings	above expect.	-0.0472 (0.0519)	news volume		-0.0 (0.0002)
	below expect.	-0.1687** (0.0823)	news novelty		-0.0044 (0.0163)
executive salary	cut	0.0363 (0.0591)	news sentiment		0.0001 (0.0005)
	increase	-0.0351 (0.0648)	news impact projection		0.0005 (0.0006)
expenses	charge	0.0314 (0.1651)	stock return		-0.1177*** (0.0116)
	down	-0.072 (0.1577)	volatility index		0.0108*** (0.0032)
legal issues		0.0741*** (0.018)	market to book		0.208*** (0.0794)
Observations		24,791	Residual Std. Error		2.5044
R^2		4.62 %	F Statistic		12.689***
Adj. R^2		4.33 %			

*p < 0.1; **p < 0.05; ***p < 0.01.
Control variables: Year, Firm.

As is to be expected, the established capital market variables show a considerable relationship, which has remained robust throughout different iterations of the research setup. We see that *stock price returns* are inversely correlated with CDS spread returns, which is shown by Smales (2016) for example. The *implied volatility* of the respective index (for Europe and North America) also shows a clear positive correlation with the CDS spreads. This is understandable, as the implied volatility is often seen as an indicator of fear or uncertainty. It appears to have a direct impact on the perceived default risk of banks. The *market to book* coefficient is also positive.

In addition, the regression model was also created using $\Delta CDS_{i,t}$as a dependent variable. As expected, a significantly larger R^2 of about 25 % is found, depending on the respective configuration. Thus, the CDS spread return was put into relation with the independent variables of the same day. Here, for the *news sentiment* score a negative relationship and for the news topic of *stock loss* a positive relationship was found. This relationship was also shown in other papers. However, for the case of $\Delta CDS_{i,t+1}$, this relationship does not seem to be evident.

For the model, we tried to account for the limitations inherent to the study design. First, the goal of this analysis to analyse the interrelations at the daily level, which can potentially enable a timely risk assessment. At the same time, this makes it more difficult to consider accounting measures, which were omitted in this analysis, as control variables. Since the basis of the analysis are time series (large *n*) for a moderately large number of banks, there are potential risks regarding heteroscedasticity and serial autocorrelation. To account for this effect, the covariance matrix is estimated using the Newey West estimator, similar to Breitenfellner and Wagner (2012).

In our analysis we find an R^2 of 4.62 % and an adjusted R^2 of 4.33 %. While this may seem small in some scenarios, it is not uncommon when taking a look at related literature that analyzes CDS spreads at the daily level. Bouzgarrou and Chebbi (2016) report an R^2 of 13.4% in the context of sovereign CDS, with a smaller number of entities and a shorter observation period. While there are differences in the study design, the R^2 of Liebmann et al. (2016) is on a comparable scale. Hence, we consider the reported R^2 to be acceptable.

5. Discussion and conclusion

With this paper, we contribute to research in the field of credit risk management by investigating which topics of news reports are related to CDS spread returns, which is a market-based measure for assessing company or bank credit risk. Regarding H1, the results of regressing changes in CDS spreads on news topics from the prior day do not show a clear picture. Primarily the topic *price target set* shows an association with CDS spread returns. The measure *analyst rating change positive* surprisingly shows a positive correlation, which is not consistent with the underlying expectation. For H2, we find that the non-financial topics *sanctions* and *legal issues* are associated with positive CDS spread returns, i.e., an increased risk of default. Regarding financial news topics, a *price target downgrade, decreased dividends*, and *revenue guidance down* play a significant role. Contrary to intuition, *earnings below expectation* is accompanied by a reduction in CDS spread returns. Although our analysis could provide first insights, we see the need for complementary investigations.

Furthermore, we can derive implications for the model component of a decision support system. Quantitative metrics such as sentiment or news volume should be supplemented with qualitative measures such as the topic under discussion. The rationale for this is that the present investigation has shown that some of the news topics are associated with CDS spread returns. They are a proxy for credit risk. Furthermore, there are also implications for the user interface component. In this case it is also possible to enrich the quantitative values with additional qualitative insights.

Our study is relevant for practitioners, as it contributes to a better understanding of the relationship between issues in financial news and CDS spreads. This is particularly crucial to enrich risk monitoring systems that use quantitative metrics. In this research context, the question arises, which topics from financial news related to changes in CDS spreads are particularly significant. For example, do news regarding the acquisition of another company lead to an increase or decrease in credit risk? The results of the analysis carried out indicate that both financial and non-financial topics may contain relevant signals.

At the same time, our study is subject to limitations. First, further additional control variables both at the company level (balance sheet ratios) and macroeconomic level could be utilised. However, the challenge here is that these figures are usually updated quarterly at most. The limitations of the statistical analysis were addressed, for example, by using the Newey West estimate of the covariance matrix. Nevertheless, there is further potential for refining the analysis, for example, by utilising more specialised time series analysis methods. Moreover, a more differentiated view on the different phases of crisis and economic recovery after the global financial crisis could be vital to uncover further insights. Finally, it should be noted that CDS spreads of banks of different restructuring standards were compared.

Several promising opportunities for subsequent research are evident. First, the analysis could be conducted using a larger selection of banks. Additionally, in the present analysis, the topics are compared with the following day. Even more diverse combinations of frequencies (weekly, monthly) and lags of CDS spreads and quantitative news variables might be considered and evaluated in future research. Another research opportunity is to validate the robustness of the results with out-of-sample forecasts. Finally, an in-depth focus on mixed-frequency regression offers great potential.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Andres, C., Betzer, A., & Doumet, M. (2016). *Measuring abnormal credit default swap spreads*. https://ssrn.com/abstract=2194320

Apergis, N. (2015). Forecasting credit default swaps (CDSs) spreads with newswire messages: evidence from European countries under financial distress. *Economics Letters*, *136*, 92–94. https://doi.org/10.1016/j.econlet.2015.08.032

Avino, D.E., Conlon, T., & Cotter, J. (2019). Credit default swaps as indicators of bank financial distress. *Journal of International Money and Finance*, *94*, 132–139. https://doi.org/10.1016/j.jimonfin.2019.03.001

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*, *60*(5), 2255–2281. https://doi.org/10.1111/j.1540-6261.2005.00798.x

Bouzgarrou, H., & Chebbi, T. (2016). The reaction of sovereign CDS spread volatilities to news announcements. *Journal of Asset Management*, *17*(5), 347–360. https://doi.org/10.1057/jam.2016.20

Breitenfellner, B., & Wagner, N. (2012). Explaining aggregate credit default swap spreads. *International Review of Financial Analysis*, *22*, 18–29. https://doi.org/10.1016/j.irfa.2012.02.002

Cerchiello, P., & Nicola, G. (2018). Assessing news contagion in finance. *Econometrics*, *6*(1), 1–19. https://doi.org/10.3390/econometrics6010005

Cerchiello, P., Nicola, G., Rönnqvist, S., & Sarlin, P. (2017). Deep learning bank distress from news and numerical financial data. In: *Proceedings of the DEM working paper series*. Department of Economics and Management, University of Pavia.

Chebbi, T., & Sarraj, M. (2017). News and sovereign CDS markets: Evidence from the Euro Area. *International Journal of Bonds and Derivatives*, *3*(4), 320–334. https://doi.org/10.1504/IJBD.2017.091611

Chiaramonte, L., & Casu, B. (2013). The determinants of bank CDS spreads: Evidence from the financial crisis. *The European Journal of Finance*, *19*(9), 861–887. https://doi.org/10.1080/1351847X.2011.636832

Collin-Dufresn, P., Goldstein, R.S., & Martin, J.S. (2001). The determinants of credit spread changes. *The Journal of Finance*, *56*(6), 2177–2207. https://doi.org/10.1111/0022-1082.00402

Donovan, J., Jennings, J., Koharki, K., & Lee, J. (2018). *Determining credit risk using qualitative disclosure*. https://ssrn.com/abstract=3149945

Erlwein-Sayer, C. (2018). Macroeconomic news sentiment: Enhanced risk assessment for sovereign bonds. *Risks*, *6*(4), 141. https://doi.org/10.3390/risks6040141

Galil, K., Shapir, O.M., Amiram, D., & Ben-Zion, U. (2014). The determinants of CDS spreads. *Journal of Banking & Finance*, *41*, 271–282. https://doi.org/10.1016/j.jbankfin.2013.12.005

Galil, K., & Soffer, G. (2011). Good news, bad news and rating announcements: An empirical investigation. *Journal of Banking & Finance*, *35*(11), 3101–3119. https://doi.org/10.1016/j.jbankfin.2011.04.010

Hagenau, M., Liebmann, M., & Neumann, D. (2013). Automated news reading: Stock price prediction based on financial news using context-capturing features. *Decision Support Systems*, *55*(3), 685–697. https://doi.org/10.1016/j.dss.2013.02.006

Hull, J., Predescu, M., & White, A. (2004). The relationship between credit default swap spreads, bond yields, and credit rating announcements. *Journal of Banking & Finance*, *28*(11), 2789–2811. https://doi.org/10.1016/j.jbankfin.2004.06.010

Jiang, H., Li, S.Z., & Wang, H. (2019). *News momentum*. https://ssrn.com/abstract=2679614

Jiang, H., & Sun, Z. (2015). *News and corporate bond liquidity*. https://ssrn.com/abstract=2437975

Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, *33*, 171–185. https://doi.org/10.1016/j.irfa.2014.02.006

Kim, J.B., Shroff, P.K., Vyas, D.U., & Moerman, R.W. (2018). Credit default swaps and managers' voluntary disclosure. *Journal of Accounting Research*, *56*(3), 953–988. https://doi.org/10.1111/1475-679X.12194

Kim, S.-J., Salem, L., & Wu, E. (2015). The role of macroeconomic news in sovereign CDS markets: domestic and spillover news effects from the US, the Eurozone and China. *Journal of Financial Stability*, *18*, 208–224. https://doi.org/10.1016/j.jfs.2015.04.008

Lando, D. (2004). *Credit risk modeling: Theory and applications*. Princeton University Press.

Liebmann, M., Orlov, A.G., & Neumann, D. (2016). The tone of financial news and the perceptions of stock and CDS traders. *International Review of Financial Analysis*, *46*, 159–175. https://doi.org/10.1016/j.irfa.2016.05.001

Lu, H.-M., Tsai, F.-T., Chen, H., Hung, M.-W., & Li, S.-H. (2012). Credit rating change modeling using news and financial ratios. *ACM Transactions on Management Information Systems*, *3*(3), 1–30. https://doi.org/10.1145/2361256.2361259

McNeil, A.J., Frey, R., & Embrechts, P. (2015). *Quantitative risk management: Concepts, Techniques and tools – revised edition*. Princeton University Press.

Mitra, G., & Mitra, L. (2011). *The handbook of news analytics in finance*. John Wiley & Sons.

Monk, A., Prins, M., & Rook, D. (Winter 2019). Rethinking alternative data in institutional investment. *The Journal of Financial Data Science*, *1*(1), 14–31. https://doi.org/10.3905/jfds.2019.1.1.014

Nasdaq. (2019). *KBW Nasdaq global bank index (GBKX)*. https://indexes.nasdaqomx.com/docs/Methodology_GBKX.pdf

Nassirtoussi, A.K., Aghabozorgi, S., Ying Wah, T., & Ngo, D.C.L. (2014). Text mining for market prediction: A systematic review. *Expert Systems with Applications*, *41*(16), 7653–7670. https://doi.org/10.1016/j.eswa.2014.06.009

Packer, F., & Zhu, H. (2005, March). *Contractual terms and CDS pricing*. BIS quarterly review. https://ssrn.com/abstract=1884901

Smales, L.A. (2016). News sentiment and bank credit risk. *Journal of Empirical Finance*, *38*, 37–61. https://doi.org/10.1016/j.jempfin.2016.05.002

Tsai, F.-T., Lu, H.-M., & Hung, M.-W. (2016). The Impact of news articles and corporate disclosure on credit risk valuation. *Journal of Banking & Finance*, *68*, 100–116. https://doi.org/10.1016/j.jbankfin.2016.03.018

Van Gestel, T., & Baesens, B. (2008). *Credit risk management: Basic concepts: financial risk components, rating analysis, models, economic and regulatory capital*. Oxford University Press.

Xing, F.Z., Cambria, E., & Welsch, R.E. (2017). Natural language based financial forecasting: A Survey. *Artificial Intelligence Review*, *50*(1), 49–73. https://doi.org/10.1007/s10462-017-9588-9

Yang, S., Liu, Z., & Wang, X. (2019). *News sentiment, credit spreads, and information asymmetry*. https://ssrn.com/abstract=3220291.