

The Impact of Corporate Social and Environment Performance on Credit Rating Prediction: North America versus Europe

Gregor Dorfleitner^{*,a}, Johannes Grebler^a, Sebastian Utz^b

^a*Department of Finance, University of Regensburg*

^b*School of Finance, University of St.Gallen*

Abstract

We quantify to what extent the quality of credit rating predictions improves through integrating measures of corporate social performance (CSP) in an established credit risk model. We provide comprehensive evidence of the comparative informational advantage of considering CSP in predicting credit ratings of North American and European firms. In the North American sample both environmental and social performance have an explanatory impact. The out-of-sample prediction quality is improved by more than 0.8%. By contrast, only the social performance increases the explanatory power in the European sample while environment performance does not. Overall, we show that CSP is a relevant variable for predicting credit ratings. In general, our findings support the risk mitigation view of CSP indicating that firms with high CSP are less risky and thus have better credit ratings. However, obviously the quality of the relationship depends on the socio-economical and cultural environment as well, as can be seen from the differing results in North America and Europa.

Keywords: Credit risk, Credit ratings prediction, Corporate social performance, Risk mitigation

1. Introduction

Highlights:

^{*}Corresponding author, gregor.dorfleitner@ur.de

2019-04-17

- Measures of corporate social responsibility improve credit rating prediction quality
- Firms with high corporate social responsibility have better credit ratings
- Impact on credit ratings is more pronounced in North America than in Europe

This paper analyzes whether, how, and to what extent the prediction quality of a firm's credit rating can be improved by integrating its corporate social performance (CSP) in the forecasting model. To capture regional differences, we analyze two samples of firm-level data, one of North American and one of European firms. We unify the work of Jiraporn et al. (2014) for North America and Stellner et al. (2015) for Europe by providing a framework in which results for the two regions can validly be compared in an established credit risk model. We use CSP measures of the globally available Asset4 framework and investigate the impact of these CSP measures on both the explanation and the prediction of credit ratings. We capture region specific differences by estimating models for both North America and Europe separately and for a merged data set. In particular, we apply a two-stage approach with an estimation of credit risk models including CSP variables in a first stage and an out-of-sample analysis using the estimate of stage one to predict the credit ratings in a second stage. Finally, we measure the prediction quality by comparing predicted and actual credit ratings. In North America, the environmental and social CSP show explanatory impact on credit ratings while only the social CSP is relevant in Europe. Further, we find improved predicting power of the credit risk model using CSP scores for the North American sample while we document no improvement in the European sample.

In theory, corporate social responsibility, which is constantly addressed in this paper through the more narrow, however measurable concept of CSP, can co-exist with both better and worse credit ratings, providing there is evidence of an impact. According to Goss and Roberts (2011), there are two contrary views for the impact of CSP: The risk-mitigation view and the over-investment view. According to the risk-mitigation view, a firm with high CSP faces lower risks than a firm with a low CSP if all other aspects of these firms are comparable. High CSP preserves firms from legal, reputational, and regulatory risks (Bauer and Hann, 2010), allows firms to hire better

qualified employees (Turban and Greening, 1997), and lowers agency risks (Oikonomou et al., 2014). The opposite, the over-investment view regards investments in CSP as a waste of scarce resources. An increase in fixed costs related to sustainable investments in CSP increases the volatility of earnings and thus the default risk (Frooman et al., 2008). With the exception of the environment score in Europe, credit ratings are significant positively correlated with the CSP scores and thus our findings are consistent with the risk-mitigation view.

Whether CSP adds informational power to the explanation of credit ratings has been the subject of two recent studies, namely Jiraporn et al. (2014) analyzing North America and Stellner et al. (2015) analyzing Europe. These studies provide inconsistent evidence that could be either due to the different methodological designs or the differing regional focus in their samples. To be more precise, Stellner et al. (2015) show that CSP has no impact on credit ratings in Europe while Jiraporn et al. (2014) conclude that CSP does, in fact, have an impact on credit ratings in North America. Nevertheless, these results are inappropriate for concluding that regional differences exist since both studies use different model specifications and concepts for measuring CSP (data provider Asset4 resp. Kinder Lydenberg Domini (KLD)). In particular, Asset4 and KLD CSP data show major differences, even after adjustment for different CSP definitions (Chatterji et al., 2016; Dorfleitner et al., 2015). Asset4 provides a comprehensive calculation of the scores based on more than 750 indicators which are ordinal or metric compared with KLD, with only using a binary rating system reflecting CSP strengths and concerns for U.S. firms (Humphrey et al., 2012). We use CSP measures of Asset4 due to their global coverage that allows for a consistent estimation of an well-established credit risk model for North America and Europe. The major limitation of the existing studies in the CSP–credit rating context concerns the retrospective contemplation by measuring the correlation of credit ratings and lagged CSP which lacks of out-of-sample predictions. In our study, we predict the next period’s credit rating based on all available information at a certain point in time. Finally the prediction quality is determined by comparing actual and predicted credit ratings.

Our data set includes S&P counterparty ratings, which we match with Asset4 CSP scores and a set of control variables. We estimate several versions of an established ordered probit credit risk model, which is able to handle rating migrations. To be specific, we estimate one baseline model without CSP factors and three CSP model for each sample (i.e., the North American

sample, the European sample, and the merged sample comprising both), resulting in a total of twelve model specifications. The three CSP model specifications comprise one aggregated CSP measure (ES) model, a model specification including a score for the environmental dimension of CSP, and a model specification with a score for the social dimension of CSP. The in-sample period for determining the models' coefficients ranges for credit ratings from 2003 to 2013. Subsequently, we predict credit ratings on the data set covering the years from 2014 to 2017. This two-step-process ensures that only the available level of information is used to predict the following periods' credit ratings.

We find that the integration of measures for CSP in the credit risk model increase its explanatory power in the North American sample. The quality of out-of-sample credit rating predictions is improved by 0.8%. For the European firms, only high social performance is correlated to better credit ratings. The prediction quality experiences no improvement. Distinct findings for North America and Europe result from the geographical, social, and political environment of the two regions, which is reflected in the Asset4 scores. The average level of CSP scores of North American firms is lower and their variance is higher than that of European firms. This is one possible reason for CSP scores having a higher explanatory power in predicting credit ratings in North America since the explanatory variables show a certain degree of variance there. In a nutshell, CSP has an impact on credit ratings in both regions although to a different extent. For North America, our findings are consistent with those of Jiraporn et al. (2014), while our results suggest contrasting implications to those of Stellner et al. (2015).

Our findings reveal valuable insights for researchers, debt holders, and debt issuers. Based on our approach of incorporating CSP into credit risk models, we find that debt holders experience higher accuracy in their credit rating predictions if they include CSP factors as explanatory variables in their credit risk models. With the higher prediction power they profit twofold by preventing misjudgments: In the case of overestimation of risk, they could lose business due to excessive price-setting. In the case of underestimation of risk, the applied pricing does not cover the anticipated risk. This leads to immediate losses in the risk-adjusted performance measurement and material losses when risks become imminent. Finally, debt issuers can improve their credit rating prediction and hence their cost of debt by increasing their CSP. For instance, an increase in the ES score by one standard deviation for a BBB-rated North American firm results in average savings of 14.5 basis points.

The remainder of the paper is organized as follows. We review the related literature and discuss theoretical concepts in Section 2. Section 3 describes the data set and Section 4 introduces the employed methodology. Section 5 presents the empirical results followed by Section 6 with a discussion of the findings and Section 7 with robustness tests. Finally, Section 8 concludes.

2. Theoretical considerations

Our study is based on two streams of literature: First, the one on the impact of CSP on credit ratings and, second, the strand on the regional differences regarding the attitude of firms towards CSP.

2.1. The impact of CSP on credit ratings

From a theoretical perspective, an indirect link between CSP and credit ratings in the context of financing cost and corporate financial performance (CFP) exists. There is evidence to suggest that firms with high CSP have lower financing costs in terms of both cost of equity and cost of debt (Goss and Roberts, 2011; Dhaliwal et al., 2011; El Ghouli et al., 2011). As creditworthiness is negatively related to interest rates payable on debt (Kisgen, 2006), we expect to observe a positive relationship between CSP and credit ratings based on this consideration. From this perspective, CSP can be seen, as being an underlying factor, having an impact on both financing cost and credit ratings. With respect to single CSP pillars, the negative relationship between CSP and financing cost is shown for environmental (Schneider, 2011; Sharfman and Fernando, 2008) and social performance (Chen et al., 2011). A similar path of argumentation can be formulated when considering CFP as opposed to financing costs. CSP is positively related to CFP (Kang et al., 2016; Dorfleitner et al., 2018; Von Arx and Ziegler, 2014) in the sense of sustainable future cash flows. Furthermore, CFP is positively related to creditworthiness (Standard&Poor's, 2013). Finally, firms with a high CSP tend to have a lower idiosyncratic risk due to the risk-mitigation effect of CSP, which corresponds to both lower financing costs and a higher CFP (Orlitzky, 2008). Therefore, we expect a positive relationship between CSP and creditworthiness. When re-examining the different pillars of CSP, we can expose the underlying mechanisms. Firms with a low level of environment performance face legal, reputation, and regulatory risks (Bauer and Hann, 2010). Moreover, a good social performance allows firms to hire

better qualified employees, thus being a key factor for future success (Turban and Greening, 1997). It should be noted that a contrasting view, the over-investment, exists, according to which CSP lowers CFP when costs exceed additional positive returns (Brammer and Millington, 2008; Cornell and Shapiro, 1987; Aupperle et al., 1985). However, there is less supporting evidence in favor of this view.

From the empirical perspective, a few studies examine the impact of CSP on credit ratings by approaching an ordered response credit risk model and show that CSP is positively related to (good) credit ratings. Stellner et al. (2015) find no significant relationship between CSP and credit ratings in the Eurozone based on the Asset4 equal weighted rating score. However, high (low) CSP results in better credit ratings if the country's sustainability performance is high (low). Jiraporn et al. (2014) use the KLD composite score and find that the CSP policies of U.S. firms are affected by other firms' policies, when located in the same three-digit zip codes area. Firms with high CSP have better credit ratings. A deeper look at single dimensions of CSP by utilizing KLD data shows that U.S. firms with high environmental and social performance have better credit ratings (Oikonomou et al., 2014; Attig et al., 2013; Bauer and Hann, 2010; Bauer et al., 2009).

Although there exists empirical evidence on the general CSP-credit rating link, it is still not clear, if there is an impact of CSP on the prediction quality. Additionally, there is no consistent evidence on the question whether the impact of CSP on credit ratings depends on regional differences, a matter which we treat in the next subsection.

2.2. CSP in North America and Europe

In general, CSP is higher in Europe compared with North America (or the U.S. as its main representative). The respective explanations include the legal origin (common law versus civil law), the deviating institutional and political set-up, the level of economic development, the historic tendency to liberal democracy, and the perception of stakeholders (Liang and Renneboog, 2017; Doh and Guay, 2006; Cai et al., 2016; Welford, 2005; Maignan, 2001). In particular, the stakeholder perception is linked to the differing ideologies as defined by Lodge (1990). European countries are more closely related to the communitarian ideology (pursuing the goal of common long-term good) while the U.S. are subject to the individualistic ideology (pursuing individual short-term improvements). The firms' motivation to act in a socially responsible

way differs between these two regions depending on firm size and financial performance (Sotorrió and Sánchez, 2008).

CSP in the U.S. is ingrained in society while CSP is more state-oriented in Europe. Although, historically, CSP has been driven more by concrete corporate policies and programs which contribute to social concerns in the U.S. compared to Europe, the rise of CSP in Europe results from incentives for corporate engagement provided by the European Union (Matten and Moon, 2008).

Empirical evidence shows that North America's CSP exceeds Europe only regarding rare aspects such as business communication (Maignan and Ralston, 2002). Regarding most aspects and measurement concepts, CSP is higher in Europe. We expect the regional differences in the CSP level to have an impact on credit ratings and their predictions in this study.

3. Data

We match S&P credit ratings of North American and European counterparties from Compustat with the rating universe of the sustainability rating agency Asset4 provided by Thomson Reuters Datastream. Moreover, we use firm-year financial and accounting data from Datastream and Worldscope to control for well-documented influence factors on credit risk. Financial counterparties are excluded based on the economic sector level of Thomson Reuters Business Classification. Our final data set comprises a panel including 724 North American firms (5,393 firm-year observations) and 218 European firms (1,712 firm-year observations). The North American panel and the European panel follow the region classification of Fama and French (2012).¹

3.1. *S&P Credit ratings and Asset4 CSP scores*

We use S&P long-term borrower credit ratings reflecting the obligors' creditworthiness over a long-term time horizon (greater than one year) as the independent variable. The S&P issuer credit rating is defined as being the

¹ The North American panel includes the United States and Canada while the European panel includes Austria, Belgium, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom according to the FTSE Country Segment classification.

current assessment of an obligor's overall financial capacity to serve its debt, i.e., its creditworthiness. The rating grades comprise AAA, AA, A, BBB, BB, B, CCC, CC, and D, where D is assigned to obligors that are overdue in either their interest or capital payments. Credit ratings of BBB or better are often referred to as 'investment grade' while credit ratings below this threshold are often called 'non-investment' or 'speculative grade'. Standard&Poor's (2017) contains a detailed description of the rating methodology.

We estimate the CSP of a company through three different means: the equal weighted ES score and its two pillar scores from the Asset4 database, i.e., the environment score (ENV), and the social score (SOC). The Asset4 ES score reflects the overall score of a firm in a certain year. It is derived by aggregating the two pillar scores. Asset4 publishes scores acting as external measures for sustainable business models (Ioannou and Serafeim, 2012; Chatterji et al., 2016; Humphrey et al., 2012). These scores are based on publicly available and traceable information, for instance websites, SEC filings such as 10-K, DEF 14A, and 10-Q, sustainability reports, media sources, and NGO reports. To guarantee a high level of integrity of the data, every entry is cross-checked by at least one additional analyst and by further analyses through statistical tools. Therefore, using the Asset4 scores ensures the elimination of weaknesses such as the lack of transparency in the KLD, FTSE4Good, and Dow Jones-rating approaches (Chatterji and Levine, 2006) as far as possible. This is due to the fact that Asset4 evaluates more than 750 individual data points. Every data point matches a single question concerning the fulfillment of a specific item according to economic, environmental, social, and governance issues. The information gathered by the answers is aggregated in several stages to indicators, these again to pillars, and finally to the average CSP rating. The scores are updated on an annual basis and range from zero to 100 with a higher score indicating a higher level of CSP. The rating universe of Asset4 even includes a firm post-bankruptcy, a merger, and other causes of delisting. Thus the data set is free from survivorship bias.

3.2. Control variables

To capture well-documented effects of predicting credit ratings, we control for several variables. The detailed description of all of these variables is provided in Table A.7.

Following Standard&Poor's (2013); Merton (1974) we include three-year averages of the operating margin, the long-term debt, the total debt, and the interest coverage ratios. The interest coverage ratio is transformed as

suggested by Blume et al. (1998). Since negative values can be caused either by low interest payments or by high negative earnings, the magnitude of negative values for interest coverage is not meaningful and therefore these values are set to zero. The distribution of the interest coverage ratio is heavily skewed and the marginal effect of changes may be small if interest coverage is already on a high level. Accordingly, we cap the three-year average at 100. To capture the non-linear shape of interest coverage C_{it} for company i in year t , we apply the decomposition to four sub-variables c_{it}^A , c_{it}^B , c_{it}^C , c_{it}^D as defined by:

	c_{it}^A	c_{it}^B	c_{it}^C	c_{it}^D
if $C_{it} \in [0, 5)$	C_{it}	0	0	0
if $C_{it} \in [5, 10)$	5	$C_{it} - 5$	0	0
if $C_{it} \in [10, 20)$	5	5	$C_{it} - 10$	0
if $C_{it} \in [20, 100)$	5	5	10	$C_{it} - 20$.

Furthermore, we also include market capitalization because bigger firms tend to have superior credit ratings (Altman et al., 1977) and because Asset4 scores show a market cap dependence. Moreover, since all claims on assets must earn the same compensation per unit of risk (Merton, 1974; Campbell et al., 2008; Friewald et al., 2014), we also control for systematic risk (market model beta) and idiosyncratic risk. Additionally, the dividend policy of a firm also has an impact on credit risk (Hoberg and Prabhala, 2009). As profitable firms are less likely to default, we expect a positive correlation between the market-to-book ratio and credit ratings and thus include market-to-book ratio (Pástor and Pietro, 2003). Following DeAngelo et al. (2006), retained earnings can be used as a proxy of a company's life cycle phase. Mature, stable firms are generally awarded a better rating (Fons, 1994). Thus we also include retained earnings as control. Additionally, Tang (2009) finds that upgraded firms have more capital expenditure than downgraded firms. Hence, we expect a positive correlation between capital expenditure and credit ratings. Moreover, firms with a poor credit risk profile tend to have precautionary savings (Acharya et al., 2012), which is why we also use the cash balance as a control. Earlier studies (Rampini and Viswanathan, 2013) document an impact of tangibility on credit risk. Furthermore, Bangia et al. (2002) find evidence that S&P credit ratings change pro-cyclically. Hence, we control for business-cycle effects as depicted by the gross domestic product-growth rate. We further use year dummies to control for remaining systematic effects (Elton et al., 2001). We follow Dimitrov et al. (2015) on the

Table 1: This table reports on the total number of firms and observations per rating class including the partial quantity of rating upgrades and downgrades for the samples of North America and Europe. Both panels cover credit ratings from 2003 to 2013 for the coefficient estimation and from 2014 to 2017 for the out-of-sample prediction. Independent variables are lagged by one year compared with the credit ratings. We use S&P long-term borrower credit ratings reflecting the obligors' creditworthiness over a long-term time horizon (greater than one year).

		North America			Europe		
		Total	Upgr.	Downgr.	Total	Upgr.	Downgr.
AAA	2003-2013	49	4	0	6	1	0
	2014-2017	7	3	0	0	0	0
AA	2003-2013	109	12	0	92	12	0
	2014-2017	32	2	0	9	3	0
A	2003-2013	906	61	6	443	45	1
	2014-2017	265	28	2	123	13	0
BBB	2003-2013	1600	62	46	535	35	18
	2014-2017	599	33	11	204	6	5
BB	2003-2013	850	54	63	142	8	16
	2014-2017	462	34	17	73	6	5
B	2003-2013	257	5	48	35	1	8
	2014-2017	208	11	31	38	1	3
C	2003-2013	14	0	5	5	0	2
	2014-2017	35	0	8	7	0	1
Total	2003-2013	3785	198	168	1258	102	45
	2014-2017	1608	111	69	454	29	14
# Firms		724			218		

connection that the standard model in the literature comes without industry fixed effects in our main models. The underlying idea assumes industry-specific influence factors so far considered in the other controls. However, we consider industry effects in the robustness checks of Section 7.

3.3. Descriptive statistics

Compared to credit ratings, ES variables and controls are lagged by one period. The estimation set contains credit ratings covering the years from 2003 through 2013 and independent variables between 2002 and 2012. Out-of-sample predictions for credit ratings in the time period from 2014 to 2017 are based on independent variables from 2013 to 2016. Table 1 presents descriptive statistics of the credit rating variable, sorted by region (North America vs. Europe) and sub-period (2003-2013 vs. 2014-2017). Rating class BBB shows the largest number on observations in both regions.

For the ES score as well as for the ENV and SOC scores, both mean and median are lower in North America than in Europe indicating weaker overall

CSP. The standard deviation in the ES, ENV, and SOC scores is higher in North America than in Europe. Thus, CSP shows higher variability in North America than in Europe. The distributions of the measures for CSP are skewed to the left. A reason for this phenomenon may be found in the fact that companies with weak CSP ratings are less likely to provide the data required to obtain an ES rating. Hence, the proportion of weak performance companies in the database is less than in the basic population causing this skewness.

The most significant regional difference referring to descriptive statistics is that the mean size of firms in the European sample is bigger than in North America. This can be explained by the broader availability of credit ratings of smaller firms in North America. Moreover, the macroeconomic situation measured by the GDP growth rate shows a high degree of deviation. While the level of GDP growth in the estimation period and the out-of-sample period in North America is 1.6% and 2.2% respectively, the respective numbers in Europe are 0.4% and 0.8% lower.

4. Methodology

Based on the approach of Kaplan and Urwitz (1979), its continuation by Blume et al. (1998), and its application in many studies (e.g., Dimitrov et al., 2015; Baghai et al., 2014; Alp, 2013; Jiang et al., 2012; Becker and Milbourn, 2011), we estimate a threshold model based on an unobserved linking variable y_{it}^* , which represents the creditworthiness of a firm i and year t

$$y_{it}^* = \mathbf{x}_{i,t-1}'\boldsymbol{\beta} + \epsilon_{it}, \quad (1)$$

where $\mathbf{x}_{i,t-1}$ represents the vector of observed explanatory variables of firm i in year $t - 1$ and $\boldsymbol{\beta}$ is a vector of slope coefficients. The variable R_{it} is the rating category of firm i and year t . The linking variable y_{it}^* is continuous and its range comprises the set of real numbers. In our study, we consider seven different levels of credit ratings (i.e., AAA, AA, A, BBB, BB, B and C). The variable R_{it} is assigned to 7 if firm i has a rating of AAA, 6 if AA, 5 if A, 4 if BBB, 3 if BB, 2 if B, and 1 if C in year t . Thus, the first stage of our estimation maps the credit ratings into a partition of the unobserved

Table 2: This table reports on descriptive statistics for the Asset4 scores and control variables for the North American and the European sample of explanatory variables covering the estimation period of 2002 to 2012 and the out-of-sample prediction period of 2013 to 2016.

Asset4 Variables	Period	North America					Europe				
		Mean	25%	Median	75%	Std. Dev.	Mean	25%	Median	75%	Std. Dev.
ES Score (%)	2002-2012	51.516	24.120	48.850	80.090	28.093	81.668	75.840	88.440	92.979	15.778
ENV Score (%)	2013-2016	57.699	32.270	60.333	84.480	27.458	85.885	84.907	91.410	93.653	13.237
	2002-2012	50.538	18.820	48.070	82.850	31.098	81.390	76.817	89.805	93.287	17.912
	2013-2016	56.660	26.307	61.740	87.862	30.464	85.989	84.867	91.965	93.990	14.280
SOC Score (%)	2002-2012	52.437	25.210	51.790	80.820	28.726	82.024	74.972	89.075	94.648	17.024
	2013-2016	58.673	34.505	63.615	84.347	27.642	85.722	85.177	91.155	94.138	14.363
	2002-2012	11.108	2.640	5.510	12.080	16.972	8.668	2.870	4.920	8.588	13.388
Interest Coverage (%)	2013-2016	8.648	1.800	4.290	9.592	14.894	8.160	1.820	4.285	8.325	14.108
Operating Margin (%)	2002-2012	13.805	7.410	12.320	19.010	8.518	11.924	5.332	10.490	16.398	8.055
	2013-2016	13.259	6.397	11.870	18.605	8.801	10.830	4.500	8.550	15.428	8.154
	2002-2012	39.532	24.970	37.430	51.680	19.200	39.736	26.797	38.610	51.028	16.842
Long Term Debt (%)	2013-2016	46.004	31.887	44.665	57.555	19.497	41.544	28.895	38.995	53.940	16.364
	2002-2012	42.702	28.110	40.720	55.230	19.013	45.764	33.315	44.535	57.545	16.103
	2013-2016	48.518	33.778	47.400	60.540	19.518	46.829	33.733	44.515	59.460	15.818
Total Debt (%)	2002-2012	18.141	3.407	6.993	16.525	36.779	26.484	5.466	11.166	28.862	39.369
	2013-2016	21.981	2.924	7.774	20.054	45.205	27.660	5.379	13.592	31.325	38.744
	2002-2012	1.057	0.732	1.011	1.320	0.412	0.936	0.689	0.904	1.161	0.330
Beta	2013-2016	1.143	0.850	1.089	1.394	0.412	0.978	0.784	0.966	1.145	0.284
	2002-2012	1.677	1.116	1.500	2.038	0.743	1.477	1.063	1.325	1.758	0.564
	2013-2016	1.632	1.031	1.373	1.980	0.807	1.376	0.947	1.175	1.611	0.575
Idiosyncratic Risk (%)	2002-2012	0.755	1.000	1.000	1.000	0.430	0.910	1.000	1.000	1.000	0.286
	2013-2016	0.727	0.000	1.000	1.000	0.446	0.833	1.000	1.000	1.000	0.374
	2002-2012	2.768	1.430	2.180	3.450	1.961	2.342	1.260	1.960	3.050	1.425
Dividend Payer (%)	2013-2016	2.867	1.330	2.170	3.572	2.260	2.113	1.040	1.655	2.820	1.436
	2002-2012	0.259	0.094	0.258	0.425	0.256	0.163	0.046	0.143	0.275	0.165
	2013-2016	0.207	0.009	0.195	0.405	0.293	0.206	0.076	0.198	0.348	0.180
Retained Earnings (%)	2002-2012	5.136	2.510	4.550	7.630	3.058	5.076	2.910	4.680	6.928	2.673
	2013-2016	5.105	2.290	4.400	7.940	3.145	4.342	2.433	3.725	5.860	2.473
	2002-2012	0.090	0.023	0.063	0.137	0.081	0.093	0.046	0.078	0.124	0.060
Capital Expense (%)	2013-2016	0.086	0.024	0.064	0.132	0.076	0.106	0.063	0.097	0.137	0.059
	2002-2012	0.358	0.143	0.296	0.557	0.247	0.326	0.164	0.312	0.470	0.194
	2013-2016	0.386	0.129	0.319	0.651	0.280	0.309	0.144	0.272	0.479	0.196
Cash Holdings (%)	2002-2012	0.016	0.016	0.023	0.029	0.018	0.012	0.002	0.017	0.026	0.022
	2013-2016	0.022	0.018	0.022	0.025	0.005	0.014	0.008	0.014	0.022	0.011
	2002-2012	3785					1258				
# Observations	2013-2016	1608					454				

linking variable y_{it}^* as follows:

$$R_{it} = \begin{cases} 7 & \text{if } y_{it}^* \in [\mu_6, \infty) \\ 6 & \text{if } y_{it}^* \in [\mu_5, \mu_6) \\ 5 & \text{if } y_{it}^* \in [\mu_4, \mu_5) \\ 4 & \text{if } y_{it}^* \in [\mu_3, \mu_4) \\ 3 & \text{if } y_{it}^* \in [\mu_2, \mu_3) \\ 2 & \text{if } y_{it}^* \in [\mu_1, \mu_2) \\ 1 & \text{if } y_{it}^* \in (-\infty, \mu_1), \end{cases} \quad (2)$$

where μ_j are partition points independent of time t . Thresholds are not given ex ante, but instead determined in the statistical procedure of estimating the model.

Following the assumption that ϵ_{it} is normally and independently distributed with a mean of 0 and a variance of 1, which is ensured in the estimation procedure, the probabilities for the different rating classes (given \mathbf{x}_{t-1}) can be calculated according to:

$$P(R_{it} = j | \mathbf{x}_{i,t-1}) = \Phi(\mu_j - \mathbf{x}'_{i,t-1}\boldsymbol{\beta}) - \Phi(\mu_{j-1} - \mathbf{x}'_{i,t-1}\boldsymbol{\beta}), \quad j = 1, \dots, 7, \quad (3)$$

with $\mu_0 = -\infty$ and $\mu_7 = \infty$.

We utilize the panel structure of the data for the model estimation. Both a certain rating (i.e., a realization of R_{it}) and realizations of the input variables are ascribed to each company and each year during the observation period. In order to represent the state of information when predictions for the following period are calculated, all influence factors are lagged by one period. Table 3 provides an overview of the input factors, boundaries, and outputs of the estimated models. We estimate models for the North American and European sample as well as on their merged dataset. We focus on three different specifications of the CSP model, namely the ES model, the ENV model, and the SOC model. Each variant includes the corresponding Asset4 score as indicated by their name plus control variables. In the merged estimation, also a region dummy and an interaction term between region and CSP is further considered. The estimation is carried out by utilizing the maximum likelihood method referring to ordered probit models (Venables and Ripley, 2002; McKelvey and Zavoina, 1975). To account for the panel structure, we pool the observations and cluster standard errors on firm level, which is appropriate for short panels. Wald p -values are calculated following

Table 3: This table gives an overview of the estimated model specifications. The CSP Models for isolated estimation of the regions North America and Europe include Asset4 and control variables while the Base Model includes only controls among their independent variables. The boundaries that are required to assign rating classes depending on the linear predictor are output of the regression. The Asset4 score represents either the equal weighted ES rating, the environment, or the social score.

Variable Category	Variable	Regional Models		Merged
		Base Model	CSP Model	CSP Model
CSP Variables	Asset4 Score		x_0	x_0
	Interaction North America & Asset4 Score			x_1
Region Variable	North America Dummy			x_2
Control Variables	Interest Coverage A	x_1	x_1	x_3
	Interest Coverage B	x_2	x_2	x_4
	Interest Coverage C	x_3	x_3	x_5
	Interest Coverage D	x_4	x_4	x_6
	Operating Margin	x_5	x_5	x_7
	Long Term Debt	x_6	x_6	x_8
	Total Debt	x_7	x_7	x_9
	Market Capitalisation	x_8	x_8	x_{10}
	Beta	x_9	x_9	x_{11}
	Idiosyncratic Risk	x_{10}	x_{10}	x_{12}
	Dividend Payer Dummy	x_{11}	x_{11}	x_{13}
	Market/Book	x_{12}	x_{12}	x_{14}
	Retained Earnings	x_{13}	x_{13}	x_{15}
	Capital Expense	x_{14}	x_{14}	x_{16}
	Cash Balance	x_{15}	x_{15}	x_{17}
	Tangibility	x_{16}	x_{16}	x_{18}
	Gross Domestic Product Growth	x_{17}	x_{17}	x_{19}
	Dummy for Year 1	x_{18}	x_{18}	x_{20}
Boundaries	(following Years analogue)			
	Lower Boundary for Rating AAA	μ_6	μ_6	μ_6
	Lower Boundary for Rating AA	μ_5	μ_5	μ_5
	Lower Boundary for Rating A	μ_4	μ_4	μ_4
	Lower Boundary for Rating BBB	μ_3	μ_3	μ_3
	Lower Boundary for Rating BB	μ_2	μ_2	μ_2
Output	Lower Boundary for Rating B	μ_1	μ_1	μ_1
	Linear Predictor	y^*	y^*	y^*
	Rating Class	R	R	R

the approach of Huber (1967) to reveal coefficient significance. Moreover, we include estimated thresholds for the various rating levels and the goodness of fit statistics McFadden R^2 . As the link function in ordered probit models limits the interpretation of the estimated coefficients, we also calculate marginal effects at means to capture the ceteris paribus impact of a marginal change in the ES score on the credit rating prediction. The calculation is based on all independent variables fixed at their means. To computationally derive the marginal effects, we follow Greene (2011).² One objective of credit portfolio models is to predict future credit ratings appropriately. In order to determine the quality of the credit rating prediction of our model specifications, we calculate the Somers' D values as a measure of correlation between

²Let \mathbf{x} be the matrix of the independent variables. Then, the marginal effects give an indication of to which extent the probability of a firm being classified to a certain rating classes changes based on the first derivative of the probabilities in Equation (3).

$$\begin{aligned}\frac{\partial P(R_{it} = 1|\mathbf{x}_{i,t-1})}{\partial \mathbf{x}_{i,t-1}} &= -\phi(\mu_1 - \mathbf{x}'_{i,t-1}\boldsymbol{\beta})\boldsymbol{\beta} \\ \frac{\partial P(R_{it} = j|\mathbf{x}_{i,t-1})}{\partial \mathbf{x}_{i,t-1}} &= \left[\phi(\mu_{j-1} - \mathbf{x}'_{i,t-1}\boldsymbol{\beta}) - \phi(\mu_j - \mathbf{x}'_{it}\boldsymbol{\beta})\right]\boldsymbol{\beta}, \forall j \in \{2, 3, 4, 5, 6\} \\ \frac{\partial P(R_{it} = 7|\mathbf{x}_{i,t-1})}{\partial \mathbf{x}_{i,t-1}} &= \phi(\mu_6 - \mathbf{x}'_{i,t-1}\boldsymbol{\beta})\boldsymbol{\beta}\end{aligned}$$

actual and predicted ratings.³

5. Empirical tests

The results of the regional and the merged models are reported in Table 4. The estimation window for credit ratings in all of these probit models ranges from 2003 to 2013. For the North American sample each of the three CSP measures has a significantly positive coefficient. Thus, all else equal, firms with a high level of CSP in the significant specifications have a higher probability of obtaining better credit ratings than firms with a low level of CSP. In the European sample only the two specifications of the model with the ES and SOC score manifest a significance of the respective CSP measure at 1% level. The ENV pillar provides no significant explanatory benefit for Europe in contrast to North America. The control variables display reasonable signs in the regressions that are consistent with the findings of the literature displayed in Section 3.2. To rule out the regionally different relations between CSP and credit ratings, we also provide a merged model of the North American and the European sample with an interaction term between region and CSP. The interpretation of an interaction effect is more difficult as the marginal effect can even be of the opposite sign (Karaca-Mandic et al., 2012; Ai and Norton, 2003). The results of our merged estimation are consistent

³According to Somers (1962), Somers' D is a measure of ordinal association. For actual ratings Z and predicted ratings Y , Newson (2001) define Somers' D as following:

$$D_{YZ} = \frac{\tau(Z, Y)}{\tau(Z, Z)} \quad (4)$$

with Kendall rank correlation coefficient τ :

$$\tau = \frac{N_C - N_D}{n(n-1)/2} \quad (5)$$

Kendall's τ is calculated by taking the difference of the number of concordant pairs N_C and the number of discordant pairs N_D as well as the sample size n . Two pairs (z_i, y_i) and (z_j, y_j) are called concordant if the ranks of both elements agree, such as $z_i > z_j$ and $y_i > y_j$ or if $z_i < z_j$ and $y_i < y_j$. By contrast, two pairs are determined as being discordant if $z_i > z_j$ and $y_i < y_j$ or if $z_i < z_j$ and $y_i > y_j$. Somers' D can take values from -1 (only disagreeing pairs) to $+1$ (only agreeing pairs). In this context of measuring how predicted and actual credit ratings are associated, the Somers' D value of $+1$ expresses the optimal case in which all predictions are actually confirmed.

Table 4: This table displays the estimation results of the ordered probit model for the North American and European data sets covering the years from 2003 to 2013. Estimation is carried out by utilizing the maximum likelihood method referring to ordered probit models. Coefficients of all variables are displayed including the significance level marked by stars. They are regarded as being significant on the level of 1% (***), 5% (**), or 10% (*) when the p -value is below these levels. The lower boundaries of the rating categories according to Sec. 4 are also displayed.

	North America					Europe					Merged		
	Base	ES	ENV	SOC	Base	ES	ENV	SOC	ES	ENV	ES	ENV	SOC
Coefficients													
Asset4 Variable													
Asset4 * North Am.		0.014***	0.011***	0.012***		0.015***	0.006	0.018***	0.010***	0.005*	0.010***	0.005*	0.012***
North Am. Dummy													0.001
Interest Coverage A	0.209***	0.212***	0.213***	0.210***	0.286***	0.294***	0.289***	0.296***	0.239***	0.239***	0.239***	0.239***	0.236***
Interest Coverage B	0.045	0.053*	0.051*	0.053*	0.065	0.061	0.061	0.068	0.052**	0.050**	0.052**	0.050**	0.054**
Interest Coverage C	0.078***	0.084***	0.082***	0.083***	0.103***	0.109***	0.105***	0.111***	0.088***	0.086***	0.088***	0.086***	0.088***
Interest Coverage D	0.005	0.005	0.005	0.005	-0.012	-0.013*	-0.012*	-0.013*	0.002	0.002	0.002	0.002	0.002
Operating Margin	0.000	0.008	0.007	0.006	0.031***	0.035***	0.033***	0.034***	0.014***	0.013***	0.014***	0.013***	0.012***
Long Term Debt	-0.050***	-0.044***	-0.045***	-0.045***	0.000	0.002	0.001	0.001	-0.024***	-0.024***	-0.024***	-0.024***	-0.025***
Total Debt	0.030**	0.024**	0.024**	0.025**	-0.017	-0.019	-0.018	-0.018	0.005	0.005	0.005	0.005	0.006
Size	0.074***	0.060*	0.063*	0.061***	0.079**	0.070*	0.075***	0.069***	0.059***	0.062***	0.059***	0.062***	0.059***
Beta	-0.179	-0.189*	-0.201*	-0.173	-1.518***	-1.524***	-1.510***	-1.561***	-0.473***	-0.483***	-0.473***	-0.483***	-0.467***
Idiosyncratic Risk	-1.646***	-1.631***	-1.619***	-1.647***	-1.499***	-1.515***	-1.509***	-1.502***	-1.578***	-1.568***	-1.578***	-1.568***	-1.587***
Dividend Payer	1.080***	0.998***	1.006***	1.017***	1.204***	1.176***	1.190***	1.168***	0.989***	0.996***	0.989***	0.996***	1.007***
Market/Book	0.090***	0.081***	0.084***	0.081***	0.089*	0.090*	0.091*	0.082	0.082***	0.085***	0.082***	0.085***	0.081***
Retained Earnings	1.493***	1.481***	1.506***	1.456***	1.675***	1.572***	1.614***	1.606***	1.479***	1.511***	1.479***	1.511***	1.462***
Capital Expense	-0.016	-0.010	-0.010	-0.013	-0.002	0.001	-0.001	0.003	-0.016	-0.016	-0.016	-0.016	-0.019
Cash Holdings	-0.345	-0.866	-0.862	-0.703	1.182	0.934	1.100	0.834	-0.378	-0.353	-0.378	-0.353	-0.267
Tangibility	0.337	0.182	0.155	0.264	1.463***	1.315***	1.397***	1.296***	0.560***	0.547**	0.560***	0.547**	0.637***
GDP Growth	5.681***	0.841***	1.036***	2.186***	19.238***	20.911***	19.851***	21.326***	8.534***	7.733***	8.534***	7.733***	8.760***
Time Dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Lower Boundaries													
AAA—AA	12.488	12.208	12.284	12.212	14.388	14.854	14.541	15.045	11.988	11.805	11.988	11.805	12.110
AA—A	10.957	10.638	10.721	10.647	10.861	11.313	11.012	11.490	10.014	9.837	10.014	9.837	10.141
A—BBB	7.584	7.171	7.261	7.203	7.502	7.935	7.649	8.095	6.658	6.489	6.658	6.489	6.797
BBB—BB	3.924	3.435	3.539	3.476	3.640	4.047	3.781	4.187	2.960	2.802	2.960	2.802	3.105
BB—B	0.431	-0.058	0.052	-0.025	0.323	0.727	0.464	0.860	-0.505	-0.654	-0.505	-0.654	-0.369
B—C	-4.065	-4.490	-4.401	-4.453	-2.714	-2.277	-2.557	-2.156	-4.655	-4.825	-4.655	-4.825	-4.516
Goodness of fit													
McFadden R ²	0.367	0.375	0.374	0.373	0.334	0.337	0.335	0.340	0.362	0.361	0.362	0.361	0.361
# Observations	3785	3785	3785	3785	1258	1258	1258	1258	5043	5043	5043	5043	5043

with the isolated regional estimations. ES and SOC score are relevant on the same level in both regions as the coefficients of CSP scores are significant while the interaction term between North America and CSP scores are not. The ENV score, in contrast, is only relevant in North America as we document no significant coefficient for ENV in general but a significant interaction term between the North America dummy and the ENV score.

As the interpretation of coefficients is limited in terms of their magnitude, we estimate the marginal effect on the credit rating prediction of a change in the CSP scores and present the results in Table 5. For North America, we observe significant marginal effects for all three CSP scores. The lower and the upper triangular matrix for each score in Table 5 show a clear pattern indicating that an increase in CSP scores significantly increases the probability of firms being rated at a higher rating level and reduces the probability of the firms experiencing a rating downgrade. In particular, the marginal effects represent the difference in predicted probabilities for each rating class if *ceteris paribus* the mean CSP scores increase by 1%.

A detailed consideration of the diagonals shows that firms which are currently rated in rating levels AAA, AA and A also show a significantly higher probability of being classified in the current rating level again, while it is less likely for firms which are currently rated in rating level BBB and worse to remain in the current rating level if the respective score is increased. For the ES measure, for instance, the predicted probability of an A-rated North American firm remaining in the A rating category is 0.277% higher for a firm which has a 1% higher ES score compared with a firm that is otherwise identical regarding the levels of the control variables. The probability of a BBB-rated firm to obtain a rating upgrade increases per 1% ES score by 0.330% (equaling the sum of probabilities to obtain the AAA, AA, or A grade). In terms of absolute values, this emerges as an average saving of approximately USD 1.45 with respect to Basel III economic capital per loan nominal of USD 1,000 based on the ES score's change of one standard deviation (i.e., a saving of 14.5 basis points).⁴ For the European sample,

⁴This calculation is based on the internal ratings-based approach for general corporates described in Basel Committee on Banking Supervision (2017) on the time horizon of one year. We assume the loss given default (LGD) rate to be 40% analogously to the supervisory LGD for unsecured corporate exposure referring to the foundation approach. Required probabilities of default are provided by U.S. average historic one-year corporate rating transition rates (1981-2016) according to Standard&Poor's (2017).

Table 5: This table provides marginal effects at means based on the estimation for panels North America and Europe covering the years from 2003 to 2013. The marginal effects are listed in per mille and show the impact of an increase in the ES scores by one percentage point ceteris paribus on the predicted probabilities of occurrence for the various rating classes. The row sum for each panel is zero due to the fact that the sum of the predicted probabilities across all rating classes equals 1 and changes in probabilities hence equal 0.

predicted															
North America															
Europe															
	AAA	AA	A	BBB	BB	B	C	AAA	AA	A	BBB	BB	B	C	
ES															
AAA	1.321**	1.840**	-2.517***	-0.627***	-0.016***	-0.001**	0.000**	1.785*	0.080	-1.778***	-0.085**	-0.002*	0.000	0.000	
AA	0.691**	1.595***	-1.044***	-1.205***	-0.035***	-0.001**	0.000**	0.145*	2.668***	-1.621***	-1.163***	-0.028***	-0.001**	0.000	
A	0.118***	0.413**	2.771**	-3.074**	-0.221**	-0.007**	0.000**	0.026*	0.775**	2.568***	-3.206**	-0.157***	-0.006**	0.000	
BBB	0.016**	0.062**	1.718**	-0.440**	-1.304**	-0.051**	-0.001**	0.004*	0.141**	2.439**	-1.693**	-0.854**	-0.035*	-0.002	
BB	0.001**	0.004**	0.137**	2.749**	-2.085**	-0.795**	-0.011**	0.000*	0.010*	0.274**	3.385**	-3.168**	-0.475**	-0.027*	
B	0.000**	0.000**	0.013**	0.501**	2.791**	-3.189**	-0.117**	0.000	0.001**	0.019**	0.846**	2.473**	-2.968**	-0.372*	
C	0.001**	0.002**	0.074**	2.038**	-0.766**	-1.328**	-0.020**	0.000*	0.005*	0.149**	3.157**	-2.425**	-0.837**	-0.050*	
ENV															
AAA	1.035**	1.474**	-1.973**	-0.522**	-0.014**	0.000**	0.000**	0.663	0.094	-0.720	-0.036	-0.001	0.000	0.000	
AA	0.552**	1.269**	-0.812**	-0.979**	-0.029**	-0.001**	0.000**	0.057	1.028	-0.616	-0.457	-0.011	0.000	0.000	
A	0.096***	0.334**	2.214**	-2.459**	-0.179**	-0.006**	0.000**	0.010	0.300	0.997	-1.243	-0.063	-0.002	0.000	
BBB	0.013**	0.050**	1.374**	-0.336**	-1.058**	-0.042**	0.000**	0.002	0.056	0.948	-0.658	-0.333	-0.014	-0.001	
BB	0.001**	0.003**	0.113**	2.213**	-1.689**	-0.632**	-0.008**	0.000	0.004	0.110	1.305	-1.229	-0.181	-0.010	
B	0.000**	0.000**	0.010**	0.399**	2.243**	-2.561**	-0.093**	0.000	0.000	0.008	0.323	0.968	-1.155	-0.144	
C	0.000**	0.002**	0.061**	1.642**	-0.628**	-1.061**	-0.016**	0.000	0.002	0.059	1.224	-0.946	-0.321	-0.019	
SOC															
AAA	1.178**	1.625**	-2.222**	-0.566**	-0.015**	0.000**	0.000**	2.228**	-0.009	-2.119**	-0.098**	-0.002**	0.000	0.000	
AA	0.607**	1.401**	-0.874**	-1.100**	-0.033**	-0.001**	0.000**	0.176**	3.274**	-2.014**	-1.401**	-0.033**	-0.001**	0.000	
A	0.107**	0.370**	2.454**	-2.725**	-0.199**	-0.006**	0.000**	0.031**	0.944**	3.136**	-3.918**	-0.187**	-0.007**	0.000	
BBB	0.015**	0.056**	1.531**	-0.400**	-1.157**	-0.045**	-0.001**	0.005**	0.168**	2.979**	-2.074**	-1.034**	-0.042**	-0.002*	
BB	0.001**	0.003**	0.122**	2.434**	-1.850**	-0.701**	-0.010**	0.000*	0.011**	0.325**	4.157**	-3.873**	-0.588**	-0.033*	
B	0.000**	0.000**	0.012**	0.445**	2.469**	-2.823**	-0.103**	0.000*	0.001**	0.023**	1.006**	3.092**	-3.660**	-0.462**	
C	0.000**	0.002**	0.066**	1.805**	-0.684**	-1.171**	-0.018**	0.000*	0.006**	0.176**	3.849**	-2.934**	-1.036**	-0.061*	

*** significant on 1% level, ** significant on 5% level, * significant on 10% level

the results for the ES score reveal that firms with a higher score have *ceteris paribus* a higher probability of remaining in the current credit rating class (firms with current credit ratings AAA, AA and A), or even have an increased probability of a rating migration into a better credit rating class (firms across all current credit ratings). European firms with a current credit rating of BBB, BB, B, or C are less likely to remain in the current credit rating class. For instance, firms with current credit rating BBB exhibit a decrease in the probability of remaining in credit rating BBB by 0.169% and a decrease in the probability of experiencing a downgrade by one notch to non-investment grade by 0.085% as well as an increase in the probability of experiencing an upgrade to level A by 0.244%.

Table 6: This table reports on Somers' D values for panels of North America and Europe for predictions in the period from 2014 to 2017. Somers' D is used in order to measure correlation between predicted ratings and actual ratings. It can take values from -1 to $+1$ while the latter is the optimal case in which all predictions are actually confirmed. We display differences between ES models' Somers' D and those of the base models in order to show the improvement ascribed to CSP. The Wilcoxon-Mann-Whitney test (WMW) provides p -values in order to evaluate whether the probabilities of correct predictions are significantly higher in the CSP models than in the benchmark model.

	North America			Europe		
	WMW p -value	Somers' D	ref. Delta	WMW p -value	Somers' D	ref. Delta
Base Model		0.5968			0.5695	
ES Model	0.0001	0.6050	0.0082	0.9778	0.5705	0.0010
ENV Model	0.0005	0.6048	0.0080	0.9530	0.5705	0.0009
SOC Model	0.0001	0.6027	0.0059	0.9396	0.5703	0.0008
# Observations	1608			454		

Table 6 shows the results of the analysis of the prediction quality. A positive value in the Delta columns in Table 6 indicates that incorporating the respective CSP scores into the baseline model increases the prediction quality. For instance, considering the ES score in our credit risk model increases Somers' D of the North American sample by 0.82%. We apply the Wilcoxon-Mann-Whitney (WMW) test on the probabilities that the actual credit rating is predicted and find a p -value of 0.01% to show that the increase is different from zero. Regarding the two single pillar scores, we find likewise reasonable increases in Somers' D. In particular for the ENV score, the increases in Somers' D of 0.80% is higher compared to the SOC model (with 0.59%). In the European sample, none of the CSP models show any relevant improvement in Somers' D compared to the base model. Ensuing

from the North American improvement in prediction quality for ES and SOC and the significant coefficients for both measures for both regions, one would expect similar results in Europe, but finds lower ones. This may depend on the differing CSP distribution amongst firms in these regions. According to Table 3, the mean level of the ES score is, at 51.5%, lower in North America than the European sample with 81.7%. Its standard deviation of 28.1% is higher than its counterpart 15.8% in Europe. The median of 88.4%, compared with the lower mean, indicates an ES distribution in Europe which is skewed to the left. The SOC distribution is similar. In order to explain regional deviations in the predictive performance, the variability of the single CSP dimensions requires further considerations which we present in the following section.

6. Discussion

As the ES score is an aggregation over the dimensions of ENV and SOC, we next take a more thorough look at the effects of these two dimensions.

6.1. *Environmental regulation and credit ratings*

The explanatory power of environment performance for credit ratings is clearly shown for North America. Firms with good environment performance are more likely to be rewarded with a better credit rating (Tables 4 to 5). The prediction quality also increases by 0.8% (Table 6). In contrast, there is no observable relevance of environment performance for Europe (Tables 4 to 6). To explain the observed difference, we consider the structure of the ENV score. It consists of three categories, namely resource reduction, emission reduction, and product innovation. The awareness of resource and emission reduction may be influenced by geological conditions. North America is endowed with a broad abundance of natural resources. In contrast, Europe lacks of this variety on commodities. As a result, the necessity to use resources economically is higher in Europe. Further differences are reflected in the legislation of the two regions. Environmental regulations are weaker in the U.S. than in many European countries such as Denmark and Sweden (Johnstone et al., 2012). The willingness to accept binding environmental protection agreements can be exemplified by considering the Kyoto Protocol of 1997. It states reduction goals for greenhouse gases of 8% for the European community, but less for the United States (namely 7%), which have, however, never ratified it (Kyoto Protocol, 1998). The Doha Amendment to

the Kyoto Protocol has not yet been attended by the United States. The different geographical and political circumstances between North America and Europe are reflected in the different levels of the average ENV scores. Dorfleitner et al. (2018) show that for the U.S., high ENV scores can predict positive earnings surprises in later periods, which can partially explain the positive effect on creditworthiness that we find for North America. Finally firms in North America have a higher degree of freedom in order to differentiate themselves from their peers with respect to environmental issues compared to Europe, which results in explanatory and prediction power improvements only for North America.

6.2. *The benefits of social politics*

In the credit rating regressions (Tables 4), the coefficient for social performance is significant in both North America and Europe on a 1% level. Moreover, the marginal effects analysis reveals a distinct increase in the probability of both regions either maintaining their current rating or even migrating to a better rating class (Table 5). As firms profit from high CSP by being able to hire better qualified employees (Turban and Greening, 1997), it seems intuitive that this is true in both North America and Europe. Besides improvements in explanatory power in North America, good social performance results in significantly better credit rating predictions. The referring increase of prediction quality amounts to 0.6% while the increase in Europe is only 0.1% (Table 6). Although prediction quality improves only in North America, the impact of social performance is still confirmed for the European sample based on the coefficient estimation and the marginal effects. The differing prediction quality relevance of the SOC score between North America and Europe is underpinned by the varying score levels between the two regions. Referring to the descriptive statistics of our sample, the SOC score in North America is distinctly lower than it is in Europe (58.7 vs. 85.7) while the variance is higher (standard deviation 27.6 vs. 14.4). Previous literature has, up to now, identified differences between both regions regarding many aspects of the SOC score, such as employment quality, health and safety, training and development, diversity and opportunity, community, and product responsibility. An important indicator is the social expense of a country. The U.S. and Canada spent 19.2% and 17.0% of their GDP in 2014, respectively. European countries such as France and Germany tend towards a higher level of expenditure (31.9% resp. 25.8%) (OECD, 2016). North America, respectively the U.S., lacks a comprehensive labor market policy as opposed to Europe (Mat-

ten and Moon, 2008) as well as labor unions with strong negotiating power (Du Caju et al., 2009), good employment protection, and mandatory health protection (Pfeffer, 2010). Furthermore, the gender gap report of the World Economic Forum WEF (2017), which analyzes the emancipation of women and men regarding economic participation and opportunity, educational attainment, health and survival, and political empowerment, shows that the U.S. ranked 49th (worldwide), while many European countries such as Germany (rank 12) do better. Strategies for human rights protection are more common in Europe than in North America (Welford, 2005). Firms in North America have a higher degree of freedom in order to differentiate themselves from their peers with respect to social issues. Furthermore the diversification of the European firms in terms of social performance decreased between the in-sample period and the out-of-sample period (standard deviation 17.0 vs. 14.4). The high basic level and decreasing variability of social scores in Europe lead to a certain degree of similarity among firms, which explains insignificant results in our prediction models despite increased explanatory power.

7. Robustness checks

This section contains several robustness checks to rule out the fact that our results may be driven by our methodological framework, the sample selection, a missing data bias, a local bias, or a time period bias.

We analyze whether the general risk-level of firms affects the impact of CSP ratings on credit ratings and cluster firms into an investment grade and a non-investment grade (speculative) group. The marginal effects in Table 5 show that the impact of CSP on credit rating predictions differs across rating classes. The better the initial rating class of a firm is, the smaller the conditional probability by which an increase in the ES score predicts a better credit rating (except for rating class C). The results for single pillars (ENV and SOC) show similar patterns to the ES results. In Europe, we also find comparable evidence as in North America for the SOC pillar score which alone appears significant there. Also previous studies show differences between the impact of CSP on the cluster of investment grade ratings and on non-investment grade ratings. In the U.S., for specific CSP factors (such as the percentage of independent directors from the corporate governance dimension) the significance of the impact on credit ratings diminishes when restricting the sample to investment grade bonds only (Ashbaugh-Skaife et al.,

2006). Therefore, we re-estimate our models based on the investment grade sub-sample of our analysis. The results are contained in Tables B.9–B.10 in the appendix. The levels of significance in the investment grade subset decrease, certainly to some extent due to the smaller number of observations. We document lacking improvements of prediction quality in North America compared with the full sample (Table B.10, Panel A), while the improvement from the entire sample is 0.82% (see Table 6).

Since endogenous CSP ratings may generate reverse causality issues, we replace firm ES ratings by industry-based ES ratings ranks and rerun our analyses. The overview of industry classes of our data is presented in Table B.8. The absolute range of possible CSP activities can vary across industries. In order to address systematic differences across industries, we follow Utz (2018) and modify the Asset4 CSP ratings. First, firms in each industry are ranked by their Asset4 score. Second, each firm is assigned its percentile score within the respective industry. Hence the best CSP performing firm in each industry holds a value of 1 and the lowest CSP performing firm a zero. Overall, the credit-rating explanation quality in terms of significance of explanatory variables remains the same (Table B.9) and has thus compatible implications as our main results. The Somers' D improvement has a lower magnitude (Table B.10). However, the decrease compared to the normal case amounts only to 0.13% for ES when utilities are excluded as the model cannot capture all relevant effects of this industry class. The results show that the loss of information (the original distance between firms with respect to their CSP ratings) results in lower improvement of the prediction quality.

Furthermore, the standard model in literature does not contain industry fixed effects (Dimitrov et al., 2015). After adding industry fixed effects to our main model, coefficient significance levels for North America and Europe remain unchanged. The Somers' D values decrease at first glance but remain again comparable if we exclude utilities (see Table B.10).

We capture a possible effect on the results of excluded observations due to the lacking data of control variables. Instead of discarding these observations, we substitute lacking values by the mean according to the mean imputation method (Schafer, 1997). When rerunning our regressions, we find the same significance of the Asset4 scores (Table B.9) and even higher Somers' D improvements (Table B.10).

Moreover, we restrict our North American sample (5,393 observations) to a U.S. sample (4,849 observations) with respect to the assumption that the underlying drivers for credit rating predictions are more homogeneous inside

the domestic market compared to the region. Again, we observe no relevant changes in the significance of the CSP scores (Table B.9). The impact of CSP on the credit-rating prediction quality even increases to 1.0% for ENV and to 0.7% for SOC (Table B.10).

Finally, we check the robustness of results according to changes in the observation period for the parameter estimation panel and the prediction data set. The main results are based on a period covering observations between 2003 and 2013 while the prediction data set includes observations from 2014 to 2017. In order to increase the number of observations in the prediction data set, we reduce the considered years in the estimation data set in favor of the prediction data set. We choose the estimation data set lasting until 2011 and 2012 with the prediction data set commencing accordingly in 2012 and 2013, respectively. Overall there is no major difference regarding the significances of the scores (Table B.9) and the improvement of credit-rating prediction quality by considering CSP (Table B.10).

8. Conclusion

One central question in the finance literature addresses the prediction quality of credit ratings (Blume et al., 1998; Kisgen, 2006). CSP is an additional informational proxy for factors that reduce firm risk, as shown in several studies such as Kim et al. (2014); Utz (2018). Therefore, we investigate whether, how, and to what extent the integration of CSP measures in credit rating predictions improves their quality. The relationship between CSP and credit ratings is significantly positive increases in North America, i.e., high CSP performance goes along with better credit ratings. Additionally, out-of-sample predictions are improved by 0.8%. In Europe, the social score adds informational power to a basic prediction model while the environmental performance does not. In general, the regional differing impact of the environment performance presumably is due to regional differences in both economic areas, such as stronger existing legal and cultural frameworks. To embed our results in a theoretical framework, our findings show supporting evidence of the risk mitigation view of high CSP.

We resolve contrary results of earlier studies by generating comparable findings for an international sample consisting of North America and Europe. Our results are based on a consistent identification of the explanatory power and the quantification of the prediction quality of CSP dimensions for both regions. In particular, our study contributes an analysis of single

dimensions environment, social performance and their aggregate to the findings of Stellner et al. (2015). While Stellner et al. (2015) show that (aggregated) CSP has no impact on credit ratings of European firms, we ascertain that the social performance is a significant explanatory factor for the credit ratings. Moreover, we confirm the findings of Jiraporn et al. (2014) that North American firms with high CSP obtain better credit ratings although our study distinguishes itself by using the methodological more sophisticated Asset4 scores (cf. Humphrey et al., 2012). We complement this study by quantifying the improvement of the prediction quality, i.e., particularly 0.8% points for North America, which is economically significant with respect to less required economic capital. Overall, we find supporting evidence that the impact of CSP performance on credit ratings is independent of the sustainability rating agency in North America and Europe.

Since the country level of CSP is of importance in the relationship between CSP and creditworthiness (Stellner et al., 2015), future research may extend our study to different regions. This is particularly interesting, since Utz (2018) finds evidence for crash risk evidence of the fact that – consistent with our results – the risk mitigation view holds in North America and Europe, however the over-investment hypothesis applies in the region of Asia Pacific. As this study focuses rather on the technical effect for credit risk models, a further extension of this research could be to dig deeper into the economic channels through which the observed effects causally emerge.

References

- Acharya, V., Davydenko, S.A., Strebulaev, I.A., 2012. Cash holdings and credit risk. *Review of Financial Studies* 25, 3572–3609.
- Ai, C., Norton, E.C., 2003. Interaction terms in logit and probit models. *Economics Letters* 80, 123–129.
- Alp, A., 2013. Structural shifts in credit rating standards. *Journal of Finance* 68, 2435–2470.
- Altman, E.I., Haldeman, R.G., Narayanan, P., 1977. Zeta analysis: A new model to identify bankruptcy risk of corporations. *Journal of Banking & Finance* 1, 29–54.

- Ashbaugh-Skaife, H., Collins, D.W., LaFond, R., 2006. The effects of corporate governance on firms' credit ratings. *Journal of Accounting and Economics* 42, 203–243.
- Attig, N., El Ghouli, S., Guedhami, O., Suh, J., 2013. Corporate social responsibility and credit ratings. *Journal of Business Ethics* 117, 679–694.
- Aupperle, K., Carroll, A., Hatfield, J.D., 1985. An empirical examination of the relationship between corporate social responsibility and profitability. *Academy of Management Journal* 28, 446–463.
- Baghai, R.P., Servaes, H., Tamayo, A., 2014. Have rating agencies become more conservative? Implications for capital structure and debt pricing. *Journal of Finance* 69, 1961–2005.
- Bangia, A., Diebold, F.X., Kronimus, A., Schagen, C., Schuermann, T., 2002. Ratings migration and the business cycle, with application to credit portfolio stress testing. *Journal of Banking & Finance* 26, 445–474.
- Basel Committee on Banking Supervision, 2017. Basel III: Finalising post-crisis reforms. Bank for International Settlements, Basel, Switzerland.
- Bauer, R., Derwall, J., Hann, D., 2009. Employee relations and credit risk. Working Paper. Maastricht University, European Centre for Corporate Engagement (ECCE).
- Bauer, R., Hann, D., 2010. Corporate environmental management and credit risk. Working Paper. Maastricht University, European Centre for Corporate Engagement (ECCE).
- Becker, B., Milbourn, T., 2011. How did increased competition affect credit ratings? *Journal of Financial Economics* 101, 493–514.
- Blume, M.E., Lim, F., MacKinlay, A.C., 1998. The declining credit quality of US corporate debt: Myth or reality? *Journal of Finance* 53, 1389–1413.
- Brammer, S., Millington, A., 2008. Does it pay to be different? An analysis of the relationship between corporate social and financial performance. *Strategic Management Journal* 29, 1325–1343.
- Cai, Y., Pan, C.H., Statman, M., 2016. Why do countries matter so much in corporate social performance? *Journal of Corporate Finance* 41, 591–609.

- Campbell, J.Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. *Journal of Finance* 63, 2899–2939.
- Chatterji, A., Levine, D., 2006. Breaking down the wall of codes: Evaluating non-financial performance measurement. *California Management Review* 48, 29–51.
- Chatterji, A.K., Durand, R., Levine, D.I., Touboul, S., 2016. Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal* 37, 1597–1614.
- Chen, H., Kacperczyk, M., Ortiz-Molina, H., 2011. Do nonfinancial stakeholders affect the pricing of risky debt? Evidence from unionized workers. *Review of Finance* 16, 347–383.
- Cornell, B., Shapiro, A.C., 1987. Corporate stakeholders and corporate finance. *Financial Management* , 5–14.
- DeAngelo, H., DeAngelo, L., Stulz, R.M., 2006. Dividend policy and the earned/contributed capital mix: A test of the life-cycle theory. *Journal of Financial Economics* 81, 227–254.
- Dhaliwal, D.S., Li, O.Z., Tsang, A., Yang, Y.G., 2011. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *Accounting Review* 86, 59–100.
- Dimitrov, V., Palia, D., Tang, L., 2015. Impact of the Dodd-Frank act on credit ratings. *Journal of Financial Economics* 115, 505–520.
- Doh, J.P., Guay, T.R., 2006. Corporate social responsibility, public policy, and NGO activism in Europe and the United States: An institutional-stakeholder perspective. *Journal of Management Studies* 43, 47–73.
- Dorfleitner, G., Halbritter, G., Nguyen, M., 2015. Measuring the level and risk of corporate responsibility—an empirical comparison of different ESG rating approaches. *Journal of Asset Management* 16, 450–466.
- Dorfleitner, G., Utz, S., Wimmer, M., 2018. Patience pays off—corporate social responsibility and long-term stock returns. *Journal of Sustainable Finance & Investment* 8, 132–157.

- Du Caju, P., Gautier, E., Momferatou, D., Ward-Warmedinger, M.E., 2009. Institutional features of wage bargaining in 23 European countries, the US and Japan, in: European Central Bank Working Paper Series. European Central Bank, Frankfurt am Main, Germany. 974.
- El Ghouli, S., Guedhami, O., Kwok, C.C., Mishra, D.R., 2011. Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance* 35, 2388–2406.
- Elton, E.J., Gruber, M.J., Agrawal, D., Mann, C., 2001. Explaining the rate spread on corporate bonds. *Journal of Finance* 56, 247–277.
- Fama, E.F., French, K.R., 2012. Size, value, and momentum in international stock returns. *Journal of Financial Economics* 105, 457–472.
- Fons, J.S., 1994. Using default rates to model the term structure of credit risk. *Financial Analysts Journal* 50, 25–32.
- Friewald, N., Wagner, C., Zechner, J., 2014. The cross-section of credit risk premia and equity returns. *Journal of Finance* 69, 2419–2469.
- Frooman, J., Zietsma, C., McKnight, B., 2008. There is no good reason not to be good. ASAC2008, Halifax, Nova Scotia.
- Goss, A., Roberts, G.S., 2011. The impact of corporate social responsibility on the cost of bank loans. *Journal of Banking & Finance* 35, 1794–1810.
- Greene, W.H., 2011. *Econometric analysis*. Pearson Higher Education, Essex, UK.
- Hoberg, G., Prabhala, N.R., 2009. Disappearing dividends, catering, and risk. *Review of Financial Studies* 22, 79–116.
- Huber, P.J., 1967. The behavior of maximum likelihood estimates under nonstandard conditions, in: *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Berkeley, CA, USA. pp. 221–233.
- Humphrey, J.E., Lee, D.D., Shen, Y., 2012. Does it cost to be sustainable? *Journal of Corporate Finance* 18, 626–639.

- Ioannou, I., Serafeim, G., 2012. What drives corporate social performance? The role of nation-level institutions. *Journal of International Business Studies* 43, 834–864.
- Jiang, J.X., Stanford, M.H., Xie, Y., 2012. Does it matter who pays for bond ratings? Historical evidence. *Journal of Financial Economics* 105, 607–621.
- Jiraporn, P., Jiraporn, N., Boeprasert, A., Chang, K., 2014. Does corporate social responsibility (CSR) improve credit ratings? Evidence from geographic identification. *Financial Management* 43, 505–531.
- Johnstone, N., Haščič, I., Poirier, J., Hemar, M., Michel, C., 2012. Environmental policy stringency and technological innovation: Evidence from survey data and patent counts. *Applied Economics* 44, 2157–2170.
- Kang, C., Germann, F., Grewal, R., 2016. Washing away your sins? Corporate social responsibility, corporate social irresponsibility, and firm performance. *Journal of Marketing* 80, 59–79.
- Kaplan, R.S., Urwitz, G., 1979. Statistical models of bond ratings: A methodological inquiry. *Journal of Business* 52, 231–261.
- Karaca-Mandic, P., Norton, E.C., Dowd, B., 2012. Interaction terms in nonlinear models. *Health services research* 47, 255–274.
- Kim, Y., Li, H., Li, S., 2014. Corporate social responsibility and stock price crash risk. *Journal of Banking & Finance* 43, 1–13.
- Kisgen, D.J., 2006. Credit ratings and capital structure. *Journal of Finance* 61, 1035–1072.
- Kyoto Protocol, 1998. United nations framework convention on climate change. Kyoto, Japan.
- Liang, H., Renneboog, L., 2017. On the foundations of corporate social responsibility. *Journal of Finance* 72, 853–910.
- Lodge, G.C., 1990. Comparative business-government relations. Prentice Hall, Englewood Cliffs, NJ, USA.

- Maignan, I., 2001. Consumers' perceptions of corporate social responsibilities: A cross-cultural comparison. *Journal of Business Ethics* 30, 57–72.
- Maignan, I., Ralston, D.A., 2002. Corporate social responsibility in Europe and the US: Insights from businesses' self-presentations. *Journal of International Business Studies* 33, 497–514.
- Matten, D., Moon, J., 2008. 'Implicit' and 'explicit' CSR: A conceptual framework for a comparative understanding of corporate social responsibility. *Academy of Management Review* 33, 404–424.
- McKelvey, R.D., Zavoina, W., 1975. A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology* 4, 103–120.
- Merton, R.C., 1974. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance* 29, 449–470.
- Newson, R., 2001. Parameters behind 'nonparametric' statistics: Kendall's tau, Somers' D and median differences. *Stata Journal* 1, 1–20.
- OECD, 2016. OECD Factbook 2015-2016: Economic, Environmental and Social Statistics. OECD, Organisation for Economic Co-operation and Development.
- Oikonomou, I., Brooks, C., Pavelin, S., 2014. The effects of corporate social performance on the cost of corporate debt and credit ratings. *Financial Review* 49, 49–75.
- Orlitzky, M., 2008. Corporate social performance and financial performance: A research synthesis, in: Crane, A., McWilliams, A., Matten, D., Moon, J., Siegel, D.S. (Eds.), *The Oxford Handbook of Corporate Social Responsibility*. Oxford University Press, pp. 113–134.
- Pástor, L., Pietro, V., 2003. Stock valuation and learning about profitability. *Journal of Finance* 58, 1749–1789.
- Pfeffer, J., 2010. Building sustainable organizations: The human factor. *Academy of Management Perspectives* 24, 34–45.
- Rampini, A.A., Viswanathan, S., 2013. Collateral and capital structure. *Journal of Financial Economics* 109, 466–492.

- Schafer, J.L., 1997. Analysis of incomplete multivariate data. CRC press, Boca Raton, FL, USA.
- Schneider, T.E., 2011. Is environmental performance a determinant of bond pricing? Evidence from the US pulp and paper and chemical industries. *Contemporary Accounting Research* 28, 1537–1561.
- Sharfman, M.P., Fernando, C.S., 2008. Environmental risk management and the cost of capital. *Strategic Management Journal* 29, 569–592.
- Somers, R.H., 1962. A new asymmetric measure of association for ordinal variables. *American Sociological Review* 27, 799–811.
- Sotorrió, L.L., Sánchez, J.L.F., 2008. Corporate social responsibility of the most highly reputed European and North American firms. *Journal of Business Ethics* 82, 379–390.
- Standard&Poor's, 2013. Corporate criteria: Ratios and adjustments.
- Standard&Poor's, 2017. 2016 annual global corporate default study and rating transitions.
- Stellner, C., Klein, C., Zwergel, B., 2015. Corporate social responsibility and Eurozone corporate bonds: The moderating role of country sustainability. *Journal of Banking & Finance* 59, 538–549.
- Tang, T.T., 2009. Information asymmetry and firms' credit market access: Evidence from Moody's credit rating format refinement. *Journal of Financial Economics* 93, 325–351.
- Turban, D.B., Greening, D.W., 1997. Corporate social performance and organizational attractiveness to prospective employees. *Academy of Management Journal* 40, 658–672.
- Utz, S., 2018. Over-investment or risk mitigation? corporate social responsibility in Asia-Pacific, Europe, Japan, and the United States. *Review Financial Economics* 36, 167–193.
- Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S. Fourth ed., Springer, New York, USA.

- Von Arx, U., Ziegler, A., 2014. The effect of corporate social responsibility on stock performance: New evidence for the USA and Europe. *Quantitative Finance* 14, 977–991.
- Welford, R., 2005. Corporate social responsibility in Europe, North America and Asia: 2004 survey results. *Journal of Corporate Citizenship* Spring 2005, 33–52.
- World Economic Forum WEF, 2017. Global gender gap report 2017.

Appendix A. Information on input variables

Table A.7: Description of control variables. Source: Worldscope and Thomson Reuters Datastream

Variable	Definition
Interest coverage	Earnings before interest and taxes divided by interest expense on debt. Negative values of this ratio are floored at zero. Any 3-year average above 100 is capped at 100. The nonlinear form of interest coverage in this model is taken into account by categorizing the ratio according to the interval of (0-5) in sub-variable A, (5-10) in sub-variable B, (10-20) in sub-variable C, and (20-100) in sub-variable D.
Operating margin	3-year averages of operating income divided by net sales or revenues
Long term debt leverage	3-year averages of long-term debt divided by total capital
Total debt leverage	3-year averages of the sum of long-term and short-term debt (including current portion of long-term debt) divided by the sum of total capital and short-term debt
Market capitalization	Percentile of the referring company's market capitalization among those of NYSE listed companies
Idiosyncratic risk	Root mean squared error from a regression of a company's stock returns with the local market index returns as a benchmark. The regression made for each firm at the time horizon of one year is based on daily stock, respectively index returns. At least 50 observations per year are required to be made available for this calculation.
Beta - systematic risk	Market model beta from the above described market model regression to calculate idiosyncratic risk
Dividend payer	Dummy variable which takes 1 if a company has positive dividends per share in the referring year and 0 otherwise
Market-to-book	Ordinary (common) equity divided by the balance sheet value of the ordinary (common) equity in the company
Retained earnings	Retained earnings divided by total assets. Retained earnings reflect the accumulated after-tax earnings of the company which have not been distributed as dividends to shareholders or allocated to a reserve account.
Capital expenditures	Capital expenditures divided by last year's total assets

Cash balances	Sum of cash and short-term investments divided by total assets
Tangibility	Net property, plant and equipment (gross property, plant, and equipment less accumulated reserves for depreciation, depletion, and amortization) divided by total assets.

Appendix B. Additional tables

Table B.8: This table reports on industry classes. Although we follow Dimitrov et al. (2015) by omitting explicit industry effects, we cover this topic in the discussion section. We run two respective models. The first model includes the rank percentiles of the Asset4 score within each industry while the second includes industry fixed effects.

	North America		Europe	
	2003-2013	2014-2017	2003-2013	2014-2017
Basic Materials	271	163	152	67
Consumer Goods	520	234	112	56
Consumer Services	680	256	184	53
Healthcare	317	103	57	16
Oil and Gas	478	272	58	36
Technology	279	99	35	9
Utilities	330	128	208	66
Telecommunications	109	33	100	34
Industry	801	320	352	117
Total	3785	1608	1258	454

Table B.9: This table displays the estimation results of the ordered probit models in the discussion section for the panels of North America and Europe. Estimation is carried out by utilizing the maximum likelihood method referring to ordered probit models. Coefficients of all variables are displayed including the significance level marked by stars. Coefficients are regarded as being significant on the level of 1%, 5%, or 10% if the p -value is below these levels. The Hosmer-Lemeshow Test and MacFadden R^2 are calculated in order to evaluate the models' goodness of fit. High p -values indicate a sufficient fit because the null hypothesis stating that the model's fit cannot be rejected. MacFadden R^2 can take values between 0 and 1 while the latter indicates perfect model fit.

Model	Coefficient	Base	ES	ENV	SOC
<i>Panel A: Models including only investment grade</i>					
Panel North America	Asset4 Coefficient		0.001	0.002	−0.001
Panel Europe	Asset4 Coefficient		0.011	0.004	0.012*
Merged Estimation	Asset4 Coefficient		0.007	0.004	0.006
	Asset4 * North America		−0.005	−0.002	−0.006
	North America	−0.687***	−0.214	−0.485	−0.183
	# Observations	3726			
<i>Panel B: Models with Asset4 industry percentiles</i>					
Panel North America	Asset4 Coefficient		0.024***	0.017***	0.023***
Panel Europe	Asset4 Coefficient		0.027***	0.011	0.036***
Merged Estimation	Asset4 Coefficient		−0.016**	−0.024***	−0.003
	Asset4 * North America		0.018***	0.018***	0.013***
	North America	−0.655***	−1.848***	−1.888***	−1.433***
	# Observations	5043			
<i>Panel C: Models with Asset4 industry percentiles excl. utilities</i>					
Panel North America	Asset4 Coefficient		0.031***	0.024***	0.030***
Panel Europe	Asset4 Coefficient		0.016	0.003	0.025**
Merged Estimation	Asset4 Coefficient		−0.018**	−0.022***	−0.009
	Asset4 * North America		0.020***	0.018***	0.016***
	North America	−0.500***	−1.824***	−1.719***	−1.513***
	# Observations	4505			
<i>Panel D: Models with industry fixed effects</i>					
Panel North America	Asset4 Coefficient		0.013***	0.010***	0.011***
Panel Europe	Asset4 Coefficient		0.015***	0.007	0.017***
Merged Estimation	Asset4 Coefficient		0.007**	0.003	0.008**
	Asset4 * North America		0.007*	0.007**	0.005
	North America	−0.522***	−0.767***	−0.900***	−0.631**
	# Observations	5043			
<i>Panel E: Models with industry fixed effects excl. utilities</i>					
Panel North America	Asset4 Coefficient		0.015***	0.011***	0.013***
Panel Europe	Asset4 Coefficient		0.007	0.000	0.012**
Merged Estimation	Asset4 Coefficient		0.003	0.001	0.004
	Asset4 * North America		0.012***	0.011***	0.010***
	North America	−0.352***	−0.965***	−0.984***	−0.836***
	# Observations	4505			
<i>Panel F: Models with controls imputed</i>					
Panel North America	Asset4 Coefficient		0.014***	0.012***	0.012***
Panel Europe	Asset4 Coefficient		0.013***	0.004	0.017***
Merged Estimation	Asset4 Coefficient		0.010***	0.004*	0.011***
	Asset4 * North America		0.004	0.007***	0.001
	North America	−0.659***	−0.704***	−0.976***	−0.474*
	# Observations	5409			
<i>Panel G: Models with the North American observations restricted to U.S.</i>					
Panel North America	Asset4 Coefficient		0.013***	0.010***	0.012***
Merged Estimation	Asset4 Coefficient		0.010***	0.004	0.011***
	Asset4 * North America		0.004	0.007**	0.002
	North America	−0.667***	−0.709**	−0.952***	−0.524*
	# Observations	4701			
<i>Panel H: Models with shorter estimation period 2002-2012 and longer prediction period 2013-2016</i>					
Panel North America	Asset4 Coefficient		0.014***	0.011***	0.012***
Panel Europe	Asset4 Coefficient		0.013***	0.005	0.015***
Merged Estimation	Asset4 Coefficient		0.011***	0.006**	0.011***
	Asset4 * North America		0.003	0.005*	0.001
	North America	−0.633***	−0.572**	−0.784***	−0.452
	# Observations	4501			
<i>Panel I: Models with shorter estimation period 2002-2011 and longer prediction period 2012-2016</i>					
Panel North America	Asset4 Coefficient		0.014***	0.012***	0.012***
Panel Europe	Asset4 Coefficient		0.012**	0.005	0.014***
Merged Estimation	Asset4 Coefficient	36	0.011***	0.007**	0.011***
	Asset4 * North America		0.002	0.004	0.000
	North America	−0.596***	−0.480	−0.707***	−0.370
	# Observations	3958			

*** significant on 1% level, ** significant on 5% level, * significant on 10% level

Table B.10: This table reports on Somers' D values as a measure of the correlation between predicted ratings and actual ratings for panels of North America and Europe covering the years from 2014 to 2017 (if not explicitly noted otherwise). Somers' D is used in order to measure the correlation between predicted ratings and actual ratings. It can take values from -1 to $+1$ while the latter is the optimal case in which all predictions are actually confirmed. We display differences between ES models' Somers' D and those of the base models in order to exhibit the improvement ascribed to CSP.

	North America			Europe		
	WMW p-Value	Somers' D	Delta Somers' D	WMW p-Value	Somers' D	Delta Somers' D
<i>Panel A: Models including only investment grade</i>						
Base Model		0.3715			0.4180	
ES Model	0.4415	0.3714	-0.0001	1.0000	0.4133	-0.0048
ENV Model	0.2200	0.3716	0.0002	1.0000	0.4168	-0.0012
SOC Model	0.2158	0.3716	0.0001	0.9999	0.4118	-0.0062
# Observations	888			335		
<i>Panel B: Models with Asset4 industry percentiles</i>						
Base Model		0.5968			0.5695	
ES Model	0.0005	0.6008	0.0040	0.9981	0.5691	-0.0005
ENV Model	0.0165	0.6002	0.0034	0.9950	0.5698	0.0003
SOC Model	0.0000	0.6002	0.0034	0.9877	0.5683	-0.0012
# Observations	1608			454		
<i>Panel C: Models with Asset4 industry percentiles excl. utilities</i>						
Base Model		0.6267			0.6278	
ES Model	0.0001	0.6336	0.0069	1.0000	0.6273	-0.0005
ENV Model	0.0048	0.6328	0.0061	0.9996	0.6278	0.0001
SOC Model	0.0000	0.6325	0.0058	0.9997	0.6264	-0.0014
# Observations	1480			388		
<i>Panel D: Models with industry fixed effects</i>						
Base Model		0.6157			0.5935	
ES Model	0.0002	0.6221	0.0064	1.0000	0.5922	-0.0013
ENV Model	0.0053	0.6214	0.0057	0.9999	0.5935	0.0001
SOC Model	0.0001	0.6208	0.0051	0.9997	0.5916	-0.0019
# Observations	1608			454		
<i>Panel E: Models with industry fixed effects excl. utilities</i>						
Base Model		0.6339			0.6379	
ES Model	0.0001	0.6425	0.0086	1.0000	0.6374	-0.0004
ENV Model	0.0038	0.6416	0.0077	0.9978	0.6379	0.0000
SOC Model	0.0000	0.6408	0.0069	0.9998	0.6365	-0.0014
# Observations	1480			388		
<i>Panel F: Models with controls imputed</i>						
Base Model		0.5903			0.5672	
ES Model	0.0000	0.5995	0.0092	0.9983	0.5672	-0.0001
ENV Model	0.0001	0.5999	0.0095	0.9876	0.5676	0.0003
SOC Model	0.0000	0.5965	0.0062	0.9871	0.5669	-0.0004
# Observations	1698			485		
<i>Panel G: Models with the North American observations restricted to U.S.</i>						
Base Model		0.5590				
ES Model	0.0000	0.5687	0.0097			
ENV Model	0.0000	0.5693	0.0103			
SOC Model	0.0003	0.5661	0.0071			
# Observations	1406					
<i>Panel H: Models with shorter estimation period 2002-2012 and longer prediction period 2013-2016</i>						
Base Model		0.5938			0.5759	
ES Model	0.0000	0.6019	0.0081	0.1640	0.5757	-0.0001
ENV Model	0.0000	0.6018	0.0080	0.4880	0.5760	0.0002
SOC Model	0.0000	0.5995	0.0057	0.0175	0.5761	0.0002
# Observations	2024			580		
<i>Panel I: Models with shorter estimation period 2002-2011 and longer prediction period 2012-2016</i>						
Base Model		0.5915			0.5591	
ES Model	0.0007	0.5998	0.0082	0.7589	0.5579	-0.0012
ENV Model	0.0041	0.5998	0.0083	0.9585	0.5587	-0.0004
SOC Model	0.0003	0.5973	0.0057	0.1433	0.5584	-0.0007
# Observations	2441			706		