Navigating Financial Turmoil: Assessing Distress in Japanese Companies through Readability and Length Cues.

Noriyuki TSUNOGAYA,
Yoshikatsu SHINOZAWA
Graduate School of Business Administration
Hitotsubashi University, Tokyo, Japan

Abstract

Purpose: This research scrutinizes the narrative disclosure information, specifically "Management Discussion and Analysis (MD&A)," within the quarterly flash earnings reports submitted by bankrupt firms on the Japan Stock Exchange. The primary objective is to explore potential differences in readability and content between insolvent and solvent firms in Japan.

Design/Methodology/Approach: The sample comprises 80 bankrupted firms listed on the Exchange in Japan between 2009 and 2019, matched with 80 solvent firms of similar sales size and industry. Employing text mining techniques, we conduct both syntactic and content analyses. Syntactic analysis involves assessing readability through metrics such as sentence length, the proportion of verbs, and particle content. Content analysis involves classifying words in the text as positive or negative. Regression analyses, incorporating control variables (size, age, debt ratio, industries, and years), examine the relationship between readability, content, and the bankruptcy dummy variable. Additional analysis includes regressing readability and content measures against the dummy variable distinguishing "liquidation" and "reconstruction" firms within the bankrupted sample.

Findings: The results support our hypothesis, revealing that MD&A documents from bankrupt firms exhibit lower readability, characterized by shorter sentences and a more negative sentiment compared to their solvent counterparts. However, no significant differences emerge between "liquidation" and "reconstruction" firms within the bankrupted sample in terms of readability, sentence length, and sentiment.

Originality: This study addresses a gap in the existing literature by considering non-quantitative data in the analysis of financial statements, leveraging advancements in AI technology, particularly in text mining. Unlike previous studies employing univariate analyses without accounting for firm diversity, our research controls for certain firm characteristics while focusing on the unique context of Japanese firms.

Table 1: Descriptive Statistics of the dependent variables with T-test results

Panel A	Dependent Variable (1) Readability	Obs	Mean	Std. err.	Std. dev.	[95% conf.	interval]	T Test
	Solvent Firms	80	0.498	0.057	0.514	0.384	0.613	
	Bankrupt Firms	80	0.308	0.064	0.568	0.182	0.435	
	All Samples	160	0.403	0.043	0.549	0.318	0.489	
	Difference		0.190	0.086		0.021	0.359	t=2.2158 p=0.0281
Panel B	Dependent Variable (2) Length of Sentences (the decimal logarithm)							
	Solvent Firms	80	3.729	0.043	0.385	3.643	3.815	
	Bankrupt Firms	80	3.612	0.027	0.244	3.558	3.666	
	All Samples	160	3.670	0.026	0.327	3.619	3.721	
	Difference		0.117	0.051		0.016	0.218	t=2.2977 p=0.0229
Panel C	Dependent Variable (3) Sentiment							
	Solvent Firms	80	0.355	0.012	0.107	0.331	0.379	
	Bankrupt Firms	80	0.498	0.015	0.130	0.469	0.527	
	All Samples	160	0.426	0.011	0.139	0.404	0.448	
	Difference		-0.143	0.019		-0.180	-0.106	t=-7.5730 p=0.000

Table 2: Descriptive Statistics of the control variables.

Total Sample	Mean	SD	Maximum	Minimum
Z-score	-0.875	26.150	30.604	-258.524
Firm Size (the decimal logarithm)	9.486	0.768	11.462	7.584
Firm Age	38.856	25.950	106.000	-11.000
Debt Ratio (debts with interest over total asset)	34.063	76.332	869.231	0.000
Sub-sample (bankrupted firms)				
Z-score	-6.356	36.007	11.434	-258.524
Firm Size (the decimal logarithm)	9.173	0.691	11.462	7.584
Firm Age	34.438	26.031	106.000	-11.000
Debt Ratio (debts with interest over total asset)	55.888	102.390	869.231	0.000
Sub-sample (solvency firms)				
Z-score	4.607	4.355	30.604	0.468
Firm Size (the decimal logarithm)	9.798	0.716	11.448	8.508
Firm Age	43.275	25.263	99.000	7.000
Debt Ratio (debts with interest over total asset)	12.238	16.692	60.451	0.000

Table 3: The regression results (I)

VARIABLES	(Model 1) Readability	(Model 2) Readability	(Model 3) Sentences	(Model 4) Sentences	(Model 5) Sentiment	(Model 6) Sentiment
Bankruptcy Firm (dummy)	-0.206**	-0.162*	-0.162***	-0.115**	0.125***	0.115***
	(0.0983)	(0.0880)	(0.0558)	(0.0524)	(0.0210)	(0.0211)
Bankruptcy Year (dummy)	0.0316	0.0102	0.0130	-0.0189	-0.00851	0.00990
	(0.0865)	(0.0897)	(0.0484)	(0.0610)	(0.0188)	(0.0208)
FirmSize(decimal logarithm)	-0.0606	-0.0162	0.000265	0.0429	-0.0272*	-0.0368**
	(0.0661)	(0.0765)	(0.0312)	(0.0376)	(0.0139)	(0.0150)
Firm Age	-0.00132	0.00165	-0.00452***	-0.00289***	0.000588	0.000165
	(0.00158)	(0.00188)	(0.000920)	(0.00109)	(0.000447)	(0.000453)
Debt Ratio	-0.000754	-0.000476*	0.000129	-6.23e-06	0.000148***	8.89e-05
	(0.000486)	(0.000250)	(0.000254)	(0.000205)	(5.36e-05)	(6.88e-05)
Constant	1.142*	-0.0348	3.914***	3.298***	0.598***	0.797***
	(0.647)	(0.767)	(0.314)	(0.394)	(0.132)	(0.166)
Industry Sector Controlled	No	Yes	No	Yes	No	Yes
Year Controlled	No	Yes	No	Yes	No	Yes

Observations	160	160	160	160	160	160
R-squared	0.052	0.419	0.157	0.400	0.301	0.547

Table 4: The regression results (II)

	(Model 7)	(Model 8)	(Model 9)	(Model 10)	(Model 11)	(Model 12)
VARIABLES	Readability	Readability	Sentences	Sentences	Sentiment	Sentiment
Reconstruction Firm (dummy)	-0.0841	-0.0403	0.108**	0.0952	0.0175	0.0147
	(0.132)	(0.144)	(0.0534)	(0.0726)	(0.0318)	(0.0306)
Bankruptcy Year (dummy)						
	-0.0146	-0.0430	-0.000152	0.0279	0.00138	-0.00396
	(0.132)	(0.0956)	(0.0532)	(0.0404)	(0.0294)	(0.0230)
Firm Size (decimal logarithm)	-0.0400	-0.0122	0.0751**	0.111***	-0.0366*	-0.0324
	(0.104)	(0.126)	(0.0319)	(0.0410)	(0.0208)	(0.0195)
Firm Age	-0.00155	0.00140	-0.000213	-0.000463	-0.000492	-0.000818
	(0.00224)	(0.00273)	(0.000776)	(0.00118)	(0.000648)	(0.000890)

Debt Ratio	-0.000427	-0.000472*	-0.000147	-0.000115	0.000158**	0.000135
	(0.000381)	(0.000237)	(0.000209)	(0.000207)	(6.28e-05)	(0.000122)
Constant	0.803	-0.195	2.882***	2.440***	0.831***	0.906***
Industry Sector Controlled	No	Yes	No	Yes	No	Yes
Year Controlled	No	Yes	No	Yes	No	Yes
Observations						
R-squared	80	80	80	80	80	80
R-squared	0.022	0.735	0.118	0.665	0.060	0.697

Figure 1.

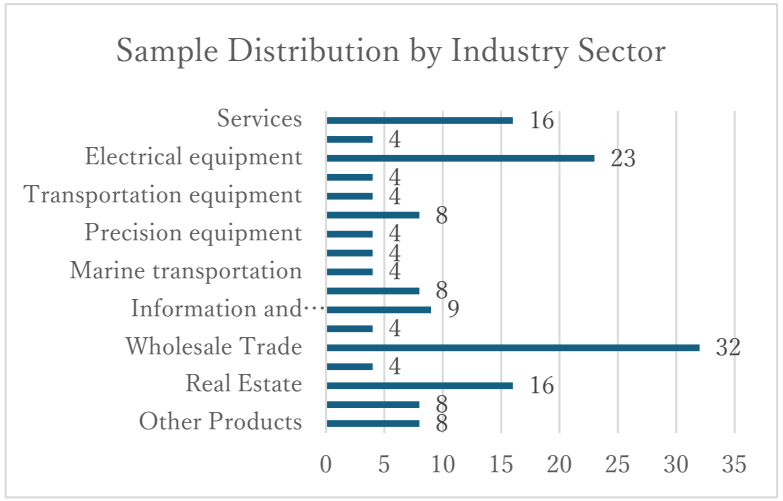

Figure 2