Financial Statement Dissimilarity Matters: Evidence from Private Debt Contracts

Huu Nhan Duong
Department of Banking and Finance
Monash University
huu.duong@monash.edu

Amanjot Singh
Faculty of Business
University of New Brunswick
amanjot.singh@unb.ca
+1 226 201 5119

Harminder Singh
Department of Finance
Deakin University
harminder.singh@deakin.edu.au

Abstract

We show that banks impose higher loan costs and more covenant restrictions for borrowers with higher year-over-year financial statement textual dissimilarity. Dissimilarity contains incremental information about weak firm performance and a higher probability of performance covenant violations. The relationship between dissimilarity and loan costs is stronger when financial statements have a more ambiguous tone, borrowers have an opaque information environment, have no prior relationship with banks, or are subject to weak corporate governance. Our findings address the SEC's concerns over users' attention to non-repeated disclosures, suggesting that, unlike equity market participants, banks understand the relevance of financial statement textual dissimilarity.

Keywords: Annual reports; financial statement dissimilarity; cost of debt; loan covenants; private debt

JEL codes: G21; G32; M41

We would like to thank the participants of the seminar at University of Victoria (2024), King's University College at Western University (2024). We thank Michael King, Vidhan Goyal, Grigori Erenburg, Isha Agarwal, and Li Liu, for their helpful discussions and suggestions.

Financial Statement Dissimilarity Matters: Evidence from Private Debt Contracts

Abstract

We show that banks impose higher loan costs and more covenant restrictions for borrowers with higher year-over-year financial statement textual dissimilarity. Dissimilarity contains incremental information about weak firm performance and a higher probability of performance covenant violations. The relationship between dissimilarity and loan costs is stronger when financial statements have a more ambiguous tone, borrowers have an opaque information environment, have no prior relationship with banks, or are subject to weak corporate governance. Our findings address the SEC's concerns over users' attention to non-repeated disclosures, suggesting that, unlike equity market participants, banks understand the relevance of financial statement textual dissimilarity.

Keywords: Annual reports; financial statement dissimilarity; cost of debt; loan covenants; private debt

JEL codes: G21; G32; M41

"What gets measured gets managed." – Peter Drucker.

1. Introduction

Financial reporting plays a vital role in the allocation of capital (Roychowdhury et al., 2019; Blankespoor et al., 2020). However, financial statements have become longer and less readable over the last two decades (Chakraborty et al., 2022), leading to a significant surge in repeated information between sections and across filings over time (e.g., Brown & Tucker, 2011; Cazier & Pfeiffer, 2017; Dyer et al., 2017a, 2017b; Li, 2019). Even though repeated disclosures, i.e., similarity in disclosures, could help financial statement users in processing value-relevant information (Brown et al., 2022); however, the Securities and Exchange Commission (SEC) has raised concerns over disclosure repetition, as it could inhibit users' attention and ability to process value-relevant non-repeated disclosures (SEC, 2018, 2019, 2020). In particular, users could remain inattentive to non-repeated material information, i.e., information contained in dissimilar financial statements. Indeed, Cohen et al. (2020) document that financial statement textual dissimilarity from one year to the next year contains useful information that is seemingly being ignored by equity market participants.

An important question remains unanswered: do other stakeholders of firms appreciate the value implications of financial statement textual dissimilarity? We answer this question by examining whether and how banks incorporate the information contained in year-over-year financial statement textual dissimilarity in the pricing and designing of private debt contracts. Bank loans act as a substantial source of financing for US firms. According to the flow of funds data from the Federal Reserve System, there have been around $4165 billion in net debt issuances since 1980. Among the debt issuances, bank loans constitute about 40% of the total debt. Given the significant role of bank debt, it

becomes imperative to understand the role of financial statement textual dissimilarity in affecting the pricing and design of private bank debt contracts.

Our analysis is further motivated by the important role that financial statements play in the lending process. Specifically, in a survey of commercial bank lenders, Donelson et al. (2017) find that nearly 88 percent of lenders consider financial statement disclosures to be an "important'' or ''very important'' factor in lending decisions. In addition, banks are sophisticated market participants, who are much more centralized in assessing their clients than other capital providers (Diamond, 1984; Diamond, 1991; Bharath et al., 2011; Vashishtha, 2014; Regenburg & Seitz, 2021; Ma et al., 2022; Hope et al., 2023). For instance, existing research suggests that banks have lower information processing costs than equity investors (Fama, 1985; Chava & Roberts, 2008; Nini et al., 2012; Chakraborty et al., 2022). Banks are, therefore, much more likely than equity investors to understand and incorporate the information in financial statements in their lending decisions. Indeed, while equity investors do not fully anticipate the implications of less reliable accruals (Richardson et al., 2005), banks charge higher interest costs for borrowers with poorer accounting quality, as defined by abnormal operating accruals (Bharath et al., 2008). Applying these arguments and empirical findings in our setting, we reason that it may be easier for commercial bank lenders to decipher the hidden information in dissimilar financial statements.

We postulate that higher changes or textual dissimilarity in financial statements will be viewed negatively by banks and are associated with higher loan costs and more restrictive loan terms for borrowers. The underlying year-over-year textual changes in financial statements are related to negative sentiments and weak firm performance (Cohen et al., 2020). As such, a higher level of dissimilarity reduces the cash flow available for debt repayment. Banks, while discounting this information, increase the costs of loans and make covenants more restrictive for borrowers with higher financial statement textual dissimilarity. The textual changes in

financial statements also increase the uncertainty for banks about their understanding of the borrowers' operations. This uncertainty increases the perceived information asymmetry between banks and borrowers. Banks, therefore, will ask for higher levels of compensation and protection in the form of higher costs of loans and more restrictive covenants when lending to borrowers with higher financial statement textual dissimilarity.

We examine the relationship between year-over-year financial statement textual dissimilarity and the cost of bank loans over the sample period between 1995 and 2017. We find that loan costs are higher for borrowers with high financial statement dissimilarity in the fiscal year before the loan starting date. The documented effect is non-trivial, a one-standard deviation increase in financial statement textual dissimilarity is associated with a 2.22% increase in loan spreads, which is equivalent to an increase of about $US 10 million dollars for an average loan size of $US 440 million dollars. Our results remain robust to several alternative measures, model specifications, or tests to mitigate confounding issues, such as the instrumental variable approach with heteroskedasticity-based instruments (Lewbel, 2012), the impact threshold for a confounding variable (ITCV) analysis, the coefficient stability test of Oster (2019), and the propensity score matching (PSM) based analysis. The relationship between financial statement textual dissimilarity and loan costs is also robust even after controlling for the effect of financial statement readability (Ertugrul et al., 2017) and financial statement comparability (Fang et al., 2016).

Having documented a significant and robust relationship between financial statement dissimilarity and loan costs, we explore the underlying channels for this relation. We argue that financial statement dissimilarity provides a signal about weak firm performance (Cohen et al., 2020), which in turn increases the cost of loan financing. We use the return on asset (ROA), net income volatility, and default risk as indicators of firm performance. Our findings suggest that firms with high financial statement dissimilarity have lower ROA, higher net income

volatility, and a higher likelihood of default. We supplement these results with an analysis of loan covenants. We find that financial statement dissimilarity is positively related to the probability of violation of performance covenants, rather than capital covenants, in the future.

We further examine how financial statement dissimilarity is related to covenant restrictions. This analysis is motivated by the role of covenants as an important monitoring mechanism for lenders (Diamond,1984; Jensen & Meckling, 1976; Myers, 1977; Smith & Warner, 1979; Rajan & Winton, 1995; Bradley & Roberts, 2015). We reason that if financial statement dissimilarity raises concerns over firm performance, banks would protect themselves from potential deterioration in borrowers' performance by imposing more covenants, especially performance-related covenants. Our results are supportive of this argument. Specifically, by employing the covenants intensity index from Bradley and Roberts (2015) as a measure of covenant restrictions, we find a higher level of covenants intensity index for borrowers with higher levels of financial statement dissimilarity. We further show that high financial statement dissimilarity is also related to a higher proportion of performance covenants relative to capital covenants.

In the final set of analyses, we examine the cross-sectional variations of the relationship between financial statement dissimilarity and loan costs. Since ambiguous tone in financial statements is related to potential accounting fraud and higher information risk (Loughran & McDonald, 2014; Ertugrul et al., 2017), we reason that financial statement dissimilarity exerts a stronger impact on loan costs when the financial statement contains more ambiguous tone. Supporting this reasoning, we find that the influence of financial statement dissimilarity on loan costs is restricted to the borrowers with high proportions of uncertain words and low proportions of strong modal words in their financial statements.

We further study the impact of firms' information environment on the relationship between financial statement dissimilarity and loan costs. If financial statement dissimilarity

influences loan costs because it provides information about firm performance, such information will be more valuable for borrowers with more opaque information environment. We use analyst coverage and the analysts' forecast dispersion as measures of borrowers' information environment (Linnainmaa et al., 2016; Chen et al., 2020; Jeon et al., 2022; Hallman et al., 2023). Our results show that the relationship between financial statement dissimilarity and loan costs is more pronounced for borrowers with low analyst coverage and high forecast dispersion. We also show that when banks do not have access to borrowers' valuable soft information through relationship lending, they charge higher loan costs to firms with dissimilar financial statements.

Finally, we explore the effect of firm-level corporate governance mechanisms. Prior studies show that monitoring from institutional investors and board of directors play an important role in improving firm performance (Shleifer & Vishny, 1986; Coles et al., 2014; Borochin & Yang, 2017; Masulis & Zhang, 2019; Liu et al., 2020; Kim et al., 2022). Drawing insights from these studies, we conjecture that if financial statement dissimilarity is an indicator of negative firm performance, it will be more prominent for borrowers subject to a weaker level of monitoring from institutional investors and the board of directors. Lending support to this conjecture, we find that the documented effect of financial statement dissimilarity on loan costs is concentrated among borrowers with lower levels of institutional ownership and higher levels of board co-option.

Our paper contributes to the prior literature in two important ways. First, we complement recent studies on how market participants incorporate soft information in financial statements. For example, less readable financial reports provide valuable information about earnings performance (Li, 2008; Loughran & McDonald, 2014; Loughran & McDonald, 2016) and are associated with greater analyst forecast dispersion and lower accuracy (Lehavy et al., 2011). Financial statement comparability reduces information acquisition costs and helps

improve analyst forecast accuracy and lowers analyst forecast dispersion (De Franco et al., 2011). We differ from these studies by focusing on the textual dissimilarity of 10-K filings. In this way, our paper is closely related to Cohen et al. (2020), who show that equity market investors remain inattentive to year-over-year textual dissimilarity and the construction of financial reports. In contrast to the findings for arm's length equity investors in Cohen et al. (2020), we show that financial statement textual dissimilarity plays an important role for banks – a more informed and sophisticated group of market participants – when designing loan contracts.

Second, we extend the literature on the pricing and design of loan contracts. An important line in this literature emphasizes the role of financial statements in driving loan contract terms. Bharath et al. (2008) and Graham et al. (2008) show that poorer accounting quality and financial restatement lead to higher loan costs. Banks also impose higher loan costs for borrowers with lower financial statement readability (Ertugrul et al., 2017), lower financial statement comparability (Fang et al., 2016), and higher financial statement complexity (Chakraborty et al., 2022). We contribute to this literature by showing that banks also incorporate the information contained in financial statement dissimilarity when designing loan contracts, with higher costs and more restrictive loan terms for borrowers with higher financial statement dissimilarity.

The remainder of the paper is organized as follows. Section 2 discusses data and variables. Section 3 contains the main results on the relationship between financial statement dissimilarity and the loan spreads and Section 4 discusses the potential channels underlying this relationship. Section 5 examines how the relationship between financial statement dissimilarity and loan spreads varies depending upon the tone of financial statements, the information transparency, lending relationships, and monitoring of the borrowers. Section 6 concludes the paper.

2. Data and Variables

We gather our data from a variety of sources. First, we use the WRDS SEC analytics database, which provides textual analysis for firm-level financial statements. We use the two measures of firm-level dissimilarity in 10-K reports to capture to what extent this year's 10-K filing is similar to the previous year's 10-K (excluding 10-K amendments) filing in terms of its text. The first measure is the Jaccard measure of textual changes. The Jaccard measure is binary (each word is counted only once as part of a given set) and is the size of the intersection divided by the size of the union of the two term frequency sets (Cohen et al., 2020). It is defined as the following:

$$Jaccard = | w_1^{TF} \cap w_2^{TF} | / |w_1^{TF} \cup w_2^{TF}| \quad (1)$$

Where, w_1^{TF} and w_2^{TF} are the term frequency vectors of the set of words occurring in documents D_1 and D_2. The second measure is the minimal edit distance (Mineditdist). Mineditdist is computed by counting the smallest number of operations or changes required to transform one document into the other (Cohen et al., 2020). The Mineditdist measure is defined as the following:

$$Mineditdist = \frac{\sum_{i=1}^{k} |d1i - d2i|}{max\{\sum_{i=1}^{k} d1i, \sum_{i=1}^{k} d2i\}} \quad (2)$$

Where, $d1i$ is the frequency of the set of words in document D_1, and $d2i$ is the frequency of the set of words in document D_2. In terms of interpretation, high values of Jaccard denote high financial statement similarity. On the other hand, high values of Mineditdist denote high financial statement dissimilarity. Therefore, we multiply our Jaccard measure with -1 to facilitate its interpretation alongside Mineditdist, i.e., higher values of Jaccard and Mineditdist capture higher financial statement dissimilarity.

We collect loan facility data from the Thomson Reuters' Dealscan database, available via WRDS. The Dealscan database provides information on the starting date and ending date of the loan facility, the interest costs, and covenant restrictions at the loan package level.

Following prior work (Graham et al., 2008; Hasan et al., 2014), we consider each loan facility as a separate unit of loan observation. We link the loan level data with the 10-Ks' dissimilarity based on the Dealscan-Compustat mapping file (Chava & Roberts, 2008).[1] We further gather firm-level financial data from the Compustat industrial annual database.

Our sample period starts from 1995, which is the first year of the WRDS SEC analytics database, and ends in 2017, which is the last year of the Dealscan-Compustat linking file. As a standard practice, our sample excludes financial and utility firms. We further remove observations with missing or negative values for total assets. Our final sample contains 16,647 loan facilities. We winsorized all the continuous variables at the 1st and 99th percentiles.

3. Results

3.1 Summary Statistics

Table 1 presents a summary statistic for our final dataset. The distribution of these measures varies across the respective measures. Jaccard and Mineditdist have a mean value of -0.66 and 0.43, and a standard deviation of 0.19 and 0.25, respectively. The average loan spread is 231 basis points with an average loan maturity of 50 months and a size of around $US 440 million dollars. Approximately 95% of the loans are syndicated. Firms in our sample, on average, have a size of $US 5991 million dollars, leverage of 31%, tangibility of 29%, cash holdings of 10%, ROA of 13%, market-to-book (MB) ratio of 1.76, sales growth of 13%, and Z-score of 1.43. These loan and firm characteristics are comparable to those reported in prior studies (e.g., Hasan et al., 2014; Ertugrul et al., 2017). All the variable definitions are provided in the appendix (Table A1).

[INSERT TABLE 1 HERE]

[1] We use the latest update of the Robert's Dealscan-Compustat Linking Database, which contains matched identifiers through the end of 2017. The file is available at the WRDS Support page.

3.2 Baseline Regression

To examine the relationship between the respective dissimilarity measures and the cost of bank loans, we use the following regression specification:

$$Loan_{spread_{i,j,t}} = f\left(Dissimilarity_{j,t-1}, FirmCtrl_{j,t-1}, LoanCtrl_{i,j,t}, FEs\right), \quad (3)$$

where $Loan_{spread_{i,j,t}}$ denotes the natural logarithm of the spread between the loan interest rate and the LIBOR for loan *i* granted to firm *j* in year *t*; $Dissimilarity_{j,t-1}$ denotes our main dissimilarity measures, i.e., Jaccard and Mineditdist, in the fiscal year prior to the loan origination year. $FirmCtrl_{j,t-1}$ denotes control variables to account for different firm-level characteristics that could be correlated with the bank loan spread. We use firm size, leverage, tangibility of assets, cash holdings, ROA, MB, sales growth, cash flow volatility, Z-score, and credit ratings as our control variables. $LoanCtrl_{i,j,t}$ denotes other loan-level variables, such as loan maturity, loan size, and loan syndication, that could also be correlated with the bank loan spread. *FEs* denote fixed effects related to 2-digit SIC industry codes, sample years, loan type, and the purpose of the loan to account for unobserved heterogeneity. We estimate this regression, with the standard error clustered at the borrowing firms.

Table 2 reports our baseline regression results on the impact of the respective dissimilarity measures on the bank loan spreads. Our coefficients for the dissimilarity measures are positive and statistically significant for both Jaccard and Mineditdist, capturing a significant role of financial statement dissimilarity in influencing the bank loan spreads. A positive coefficient implies that loan costs are higher for borrowers with higher financial statement dissimilarity in the fiscal year before the loan starting date. Our findings remain intact even after the inclusion of other firm-related and loan-related control variables. Our documented effect is also economically significant. For example, a one-standard deviation increase in financial statement dissimilarity (i.e., 0.19 – Jaccard) is associated with a 2.22% increase in

loan spread. This is equivalent to an increase of about $US 10 million dollars for the average loan size of $US 440 million dollars and loan maturity of about four years (i.e., 50 months).

In terms of control variables, we find that firms with higher leverage and weaker credit ratings have higher loan spreads. Larger firms and firms with higher ROA, MB, Z-score, and loan size have lower loan spreads. These findings are consistent with prior studies (e.g., Sengupta, 1998; Hasan et al., 2014; Fang et al., 2016; Ertugrul et al., 2017; Balachandran et al., 2019; Chakraborty et al., 2022).

Our findings imply that bankers do not merely rely on the quantitative aspects of financial statements, but also consider the qualitative factors, like the 10-K textual changes while determining the cost of debt for the borrowing firms. Cohen et al. (2020) examine the response of shareholders to these financial statement dissimilarities. Their findings suggest that shareholders remain indifferent to financial statement dissimilarity for a considerable period. On the other hand, our findings suggest that bankers pay additional attention to these dissimilarities while determining the cost of debt for the borrowing firms. We believe, ex-ante screening of loans by banks and thereafter the process of monitoring, gives an edge to banks to collect additional information (Chakraborty et al., 2022). Collection of such information upon dissimilarity of 10-K texts guides banks to price their loans accordingly.

[INSERT TABLE 2 HERE]

3.3 Addressing Endogeneity Concerns

Despite a strong relationship between financial statement textual dissimilarity and the cost of bank loans, our results could be affected by the presence of potential omitted variables that simultaneously influence financial statement dissimilarity and bank loan contracting. We attempt to mitigate this issue by including various control variables and fixed effects to capture the potential influence of firm or loan-level characteristics on loan spreads. We further deal with this issue by performing several additional tests including the instrumental variable

regression, the impact threshold for confounding variable (ITCV) analysis, the coefficient stability test, and the propensity score matching (PSM) analysis.

3.3.1 Instrumental Variable (IV) Regression

We employ an instrumental variable regression using Lewbel's (2012) method of identifying instrumental variables as a function of the model's data. In this method, in the absence of suitable external instruments, identification is achieved by constructing variables that are uncorrelated with the product of heteroskedastic errors. This approach is specifically helpful in the absence of an appropriate external instrument and has been used in several recent studies (see, for example, Anderson & Core, 2018; Mavis et al., 2020; Hasan et al., 2021). In Table 3, our findings remain consistent using Lewbel's (2012) method, reporting a statistically significant relationship between the dissimilarity measures and the cost of bank loans. Financial statement textual dissimilarity increases the bank loan spreads. The Hansen J-test also confirms the validity of overidentifying restrictions.

[INSERT TABLE 3 HERE]

3.3.2 Impact Threshold for Confounding Variable (ITCV)

We also estimate the threshold value (the ITCV value) and the impact values for the control variables. A high threshold value indicates that the results are robust to potential omitted variables bias (Frank, 2000). Panel A of Table 4 reports the ITCV results for the Jaccard and Mineditdist measures. The ITCV value reveals the minimum correlation a confounding variable needs to have with the dissimilarity measure and the cost of bank loans to make the coefficient of the dissimilarity measure statistically insignificant at the 10% level. It enables us to determine how intensely an omitted variable has to be correlated with the dependent variable and the independent variable of interest to invalidate the effect of the independent variable of interest (Larcker & Rusticus, 2010). For the Jaccard measure, the impact threshold is 0.016, i.e., the product of the correlation between Jaccard and a

confounding variable and the correlation between the cost of bank loans and a confounding variable. Specifically, the minimum correlation that a confounding variable must have with Jaccard and the cost of bank loans is 0.13 (the square root of 0.016) to invalidate the effect of Jaccard on the cost of bank loans.

As a benchmark, we then compare the threshold value to the impact values of the control variables. These impact values are estimated as the product of the partial correlation between the independent variable of interest, i.e., Jaccard, and the control variables and the partial correlation between the dependent variable (i.e., the cost of bank loans) and the control variables. The impact values for the control variables are substantially lower than the threshold value, suggesting that a confounding variable is unlikely to overturn our results. For Mineditdist, the impact threshold is also 0.016, and none of the impact values of the control variables are higher than the threshold value. Thus, after including a wide range of control variables, our results for the ITCV analysis conclude that a confounding variable is unlikely to invalidate the effect of financial statement dissimilarity on the cost of bank loans.[2]

[INSERT TABLE 4 HERE]

3.3.3 Coefficient Stability Test

We further investigate the robustness of our results through the coefficient stability test of Oster (2019). The latter approach builds on the method proposed by Altonji et al. (2005) that the relationship between the treatment (i.e., the independent variable of interest) and potential confounding variables (unobservables) can be deduced from the information from coefficient movements and the change in R^2 when the observables are included. The Oster (2019) test estimates $\tilde{\delta}$ that captures the degree of selection on unobservables relative to observables and

[2] We further investigate how severe the endogeneity issue has to be to invalidate our reported effect of the dissimilarity measures on the cost of bank loans using percent bias analysis (Frank et al., 2013; Cinelli & Hazlett, 2020). Our results suggest that we may need to replace around 51 percent of observations for Jaccard and Mineditdist with observations for which the effect of dissimilarity on the bank loan spreads is zero, i.e., to overturn our baseline results.

helps in determining whether the documented effect of dissimilarity on the cost of bank loans suffers from the selection on unobservables (Mavis et al., 2020; Chen et al., 2023). A high value of $\tilde{\delta}$ indicates a lower likelihood of unobservables confounding the results. Following Oster (2019), we assume that unobservables and observable covariates have an equal effect on the treatment (δ=1) and the maximum R^2 with the addition of unobservables can be 1.3 times the estimated R^2 of the regression with observables (i.e., controlled regression).

We consider firm and loan-level characteristics as our set of confounders: firm size, leverage, tangibility of assets, cash holdings, ROA, MB, sales growth, cash flow volatility, Z-score, credit ratings, loan maturity, loan size, loan syndication, industry fixed effects, and year fixed effects (Degryse et al., 2023). Panel B of Table 4 provides the results of this analysis. We find that the confounding effects of unobservables are unlikely to drive our results. The bias-adjusted coefficient of Jaccard (i.e., 0.112) is less than the controlled regression coefficient of Jaccard (0.117) and the changes in the adjusted coefficient fall within the 95% confidence interval of the controlled regression coefficient. Our $\tilde{\delta}$ estimate reports that unobservables would require to have a higher effect on the cost of bank loans than our observables to overturn the results. For Jaccard, the estimated value of $\tilde{\delta}$ is 7.39, which is higher than the cut-off value of 1, suggesting that unobservables need to be at least 7.39 times as influential as the observables to invalidate the documented effect of Jaccard on the cost of bank loans (Altonji et al., 2005; Oster 2019). Our findings for Mineditdist also remain qualitatively similar. For Mineditdist, the estimated value of $\tilde{\delta}$ is 5.10, which is again higher than the cut-off value of 1, indicating that the confounding effects of unobservables are unlikely to drive our results.

3.3.4 Propensity Score Matching (PSM) Approach

One could argue that firms with high financial statement dissimilarity could fundamentally be different from firms that have low financial statement dissimilarity. Therefore, we also consider a propensity score matching (PSM) approach while determining

the relationship between financial statement dissimilarity and the cost of bank loans. Table 5 provides the PSM results to examine the impact of dissimilarity measures on loan spreads. Dissimilarity measures (Jaccard and Mineditdist) with values greater than the industry-year median are categorized as treatment observations and are matched with control observations within the same industry and year. Observations are matched based on the closest propensity scores calculated using the determinants, such as firm size, leverage, tangibility, cash holdings, ROA, MB, sales growth, cash flow volatility, Z-score, and credit ratings along with industry and year fixed effects.

Panel A of Table 5 reports the mean differences between the treatment and control observations for the matched sample. Treatment and control observations are statistically indistinguishable across almost all the firm-level characteristics. This finding suggests that there are no observable differences in treatment and control firms. Panel B of Table 5 reports the regression results for the relationship between our dissimilarity measures and the bank loan spreads for the matched sample. We essentially compare the loan spreads of firms with high financial statement dissimilarity (treatment firms) to those of firms with low financial statement dissimilarity (control firms) with comparable firm-level characteristics. Our PSM regression results also support a positive and statistically significant relationship between the respective dissimilarity measures and the cost of bank loans. These results are consistent with the baseline results in Table 2 and show that loan spreads increase with an increase in financial statement dissimilarity.

[INSERT TABLE 5 HERE]

3.4 Robustness

We show a positive association between financial statement dissimilarity and loan spreads in our baseline results. Prior studies show that other aspects of financial statements such as readability (Ertugrul et al., 2017) or comparability (Kim et al., 2013; Fang et al., 2016)

are also related to the cost of debt. Motivated by these findings, we examine whether the relationship between financial statement dissimilarity and loan costs is independent of the influence of financial statement readability (or comparability) on the loan costs.

We use three measures of financial statement readability: the natural logarithm of total file size (Loughran & McDonald, 2014); the Fog index (Li, 2008), and the Bog index (Bonsall et al., 2017). For financial statement comparability, we use the firm-level comparability measures of De Franco et al. (2011), i.e., *CompAcc4*, *CompAcc10,* and *CompAccInd*. These measures capture the pairwise comparability of firm *i's* financial statement with other firms *J* during period *t*. For instance, *CompAcc4* is the average comparability of firm *i*'s financial statement to the top four firms with the highest comparability to firm *i* during period *t*. *CompAcc10* is the average comparability of firm *i*'s financial statement to the top 10 firms with the highest comparability to firm *i* during period *t,* and *CompAccInd* is the median comparability of firm *i*'s financial statement to all firms within the same industry to firm *i* during period *t.*

The findings in Panel A of Table 6 show that firms with less readable financial statements (i.e., those with higher File Size, higher Fog Index, or higher Bog Index) have higher loan costs. These results are consistent with Ertugrul et al. (2017). Similar to Fang et al. (2016), we also find in Panel B of Table 6 that higher financial statement comparability is related to lower costs of loan financing. More importantly, the association between dissimilarity measures and loan costs remains intact even after the inclusion of financial report readability, or financial statement comparability measures. Overall, the findings in Table 6 show that financial statement dissimilarity conveys an important source of information, independent from financial statement readability or comparability, for banks when pricing loan contracts.

[INSERT TABLE 6 HERE]

We also undertake an array of sensitivity tests to gauge the robustness of our results. Table 7 reports the results of our robustness tests. For brevity, Table 7 reports only the coefficients of our main variables of interest. Other loan-level, as well as firm-level variables, are similar to those employed in the baseline regression model, in Table 2, along with fixed effects. In Model (1), we consider another alternative dissimilarity measure, i.e., Cosine, as our main variable of interest. The Cosine dissimilarity measure uses the same term frequency vectors as in the Jaccard measure (Cohen et al., 2020). It is based on the frequency of each word. We multiplied our Cosine measure with -1 to facilitate its interpretation alongside our other dissimilarity measures. The coefficient remains positive and statistically significant, implying that loan spreads increase with an increase in financial statement dissimilarity.

Further, in Model (2), we use two-way clustering at the firm and the lead bank level to account for the possibility that loans issued from the same lead bank might not be independent and thus bias the standard errors (Ivashina, 2009; Hasan et al., 2014). In Model (3), we employ the overall cost of borrowing as proposed by Berg et al. (2016) as the alternative dependent variable. The overall cost of borrowing includes the interest cost and various fee components.[3] Similar to Hasan et al. (2014), we use all loan facilities in the baseline regression model. In Model (4), we instead include only the largest loan facility within a loan package per year [following Anantharaman et al. (2014)]. We further exclude observations in the financial turmoil period (2007-2008) in Model (5) and exclude all loan observations granted after 2007 in Model (6) to control for the effect of the global financial crisis on our findings. Model (7) employs firm, year, loan type, and primary purpose-based fixed effects. To further mitigate the influence of outliers, we use a median regression with robust standard errors in Model (8).

[3] The total borrowing costs data are available from Tobias Berg's website https://sites.google.com/view/tobias-berg/startseite/data-and-code.

To account for reverse causality, i.e., borrowers managing financial statement dissimilarity in anticipation of borrowings from banks in the coming year, we also consider our financial statement textual dissimilarity measures two- and three years before loan origination in Model (9). In Model (10), we also control for the firm's information environment, i.e., analyst coverage and forecast dispersion, that could potentially influence financial statement dissimilarity and the debt contracts (Linnainmaa et al., 2016; Chen et al., 2020; Jeon et al., 2022). One could also argue that corporate transactions such as acquisitions could mechanically affect financial statement textual dissimilarity, thereby influencing the bank loan contracts in the next year. To account for corporate acquisition-induced changes in the textual features of financial statements, we exclude loans for borrowers with acquisition-related expenditures one year before the loan origination in Model (11). Our findings in these models report a statistically significant and positive relationship between financial statement dissimilarity and the cost of bank loans across all the model specifications. These findings suggest that loan costs are higher for borrowers with high financial statement dissimilarity in the fiscal year before the loan starting date.

[INSERT TABLE 7 HERE]

4. Potential Channels

After documenting a significant and robust relationship between financial statement dissimilarity and loan spreads, we explore the underlying channels. We argue that financial statement dissimilarity provides a signal about weak firm performance (Cohen et al., 2020), which in turn increases the cost of bank loans. For this analysis, we consider ROA, net income volatility, and default risk as our indicators of firm performance. Net income volatility is the five-quarter rolling standard deviation of net income as a proportion of total assets. Our variable 'default' is an indicator variable, i.e., equal to 1 if the Altman's Z-score values fall below 1.81 (distress level), and 0 otherwise.

Panel A of Table 8 provides the regression results for the impact of financial statement dissimilarity at *t-1* (one year before the loan origination year) on firm performance at *t*. Our findings suggest that firms with high financial statement dissimilarity have lower ROA, higher net income volatility, and a higher likelihood of default. These findings suggest that banks anticipate these negative signals arising from financial statement dissimilarity while determining the cost of bank loans in the loan origination year.

Since financial statement dissimilarity provides a signal about weak firm performance, Panel B of Table 8 also examines the relationship between financial statement dissimilarity at *t-1* (one year before our measures of the likelihood of covenants violation) and the likelihood of covenants violation at *t*. We expect that financial statement dissimilarity should be positively related to the likelihood of covenants violation, as it contains information about potentially weak firm performance. We gather data relating to covenants violation from Demerjian and Owens (2016).[4] The authors provide covenants violation data separately for performance and capital-based covenants. Performance covenants include (1) Minimum Cash Interest Coverage, (2) Minimum Debt Service Coverage, (3) Minimum EBITDA, (4) Minimum Fixed Charge Coverage, (5) Minimum Interest Coverage, (6) Maximum Debt-to-EBITDA, and (7) Maximum Senior Debt-to-EBITDA. On the other hand, capital covenants include (1) Minimum Quick Ratio, (2) Minimum Current Ratio, (3) Maximum Debt-to-Equity, (4) Maximum Debt-to-Tangible Net Worth, (5) Maximum Leverage, (6) Maximum Senior Leverage, (7) Minimum Net Worth, and (8) Minimum Tangible Net Worth. Supporting our argument that financial statement dissimilarity signals weak firm performance, we find that, rather than capital covenants violation, financial statement dissimilarity is positively related to the probability of performance covenants violation in the future.

[4] The covenant violations data are available from Peter Demerjian's website https://peterdemerjian.weebly.com.

Finally, we examine how financial statement dissimilarity is related to covenant restrictions. Covenants play an important role for lenders in monitoring the affairs of the borrowing firms (Diamond,1984; Jensen & Meckling, 1976; Myers, 1977; Smith & Warner, 1979; Rajan & Winton, 1995; Bradley & Roberts, 2015; Balachandran et al., 2019; Honigsberg et al., 2021). We conjecture that if financial statement dissimilarity raises concern over the borrower's performance, banks would protect themselves from this potential deterioration by imposing more covenants, especially performance-related covenants. We use two different measures for covenant restrictions: (1) the covenants intensity index developed by Bradley and Roberts (2015); and (2) the performance covenants ratio. The covenants intensity index is developed by counting the following covenants groups: (1) dividend restrictions, (2) security, (3) more than two financial covenants, (4) equity sweep, (5) debt sweep, and (6) asset sweep. The performance covenants ratio is the ratio of performance covenants to total covenants (including capital covenants) imposed on the borrowing firms.

Panel C of Table 8 provides regression results for covenant restrictions at *t*, and our dissimilarity measures at *t-1* (one year before the loan origination year). Our findings report that financial statement dissimilarity increases both the intensity of covenants and the performance covenants ratio for the borrowing firms. These findings are consistent with our conjecture that since financial statement dissimilarity provides a signal about weak firm performance and a higher future likelihood of performance covenant violations, banks increase the proportion of performance covenants (relative to capital covenants) and the covenants intensity for the borrowing firms.

[INSERT TABLE 8 HERE]

5. Cross-sectional Tests

In the next set of analyses, we examine the cross-sectional variations of the relationship between financial statement dissimilarity and loan costs. First, we examine the role of ambiguous tone in financial statements while examining the association between financial statement dissimilarity and the cost of bank loans. Prior studies have documented that ambiguous tone in financial statements, measured using uncertain and modal words, is related to firm's high perceived information risk and creditworthiness (Loughran & McDonald, 2011; Ertugrul et al., 2017). Therefore, based on the ambiguous tone in financial statements, we expect that the influence of financial statement dissimilarity on loan costs should be more pronounced for borrowers with high proportions of uncertain and low proportions of strong modal words in their financial statements.

Table 9 provides regression results for ambiguous tone in financial statements, dissimilarity measures and the cost of bank loans. We gather data for ambiguous tone, i.e., the proportion of uncertain and strong modal words in financial statements, from the SEC Analytics suite by WRDS. We classify borrowers as having a low (high) ambiguous tone if the proportion of uncertain words is in the bottom (top) quintile across industry-years. Similarly, we classify borrowers as having a low (high) ambiguous tone if the proportion of strong modal words is in the top (bottom) quintile across industry-years. Our findings are consistent with our argument that the influence of financial statement dissimilarity is restricted to firms with high proportions of uncertain and low proportions of strong modal words, i.e., for borrowers with a highly ambiguous tone in financial statements. This suggests that financial statement dissimilarity contains information about the borrowers' perceived information risk and creditworthiness, which is priced by creditors, especially so when the financial statement contains more uncertain and less strong modal words.

[INSERT TABLE 9 HERE]

Second, we study the impact of the firm's information environment on the relationship between financial statement dissimilarity and the cost of bank loans. We argue that since financial statement dissimilarity contains information about firm performance, such information is expected to be more valuable for banks when the borrowers have a more opaque information environment. We use the analyst coverage and analysts' forecast dispersion as our measures of borrowers' information environment (Linnainmaa et al., 2016; Chen et al., 2020; Jeon et al., 2022; Hallman et al., 2023). We classify borrowers as having a weak (strong) information environment if the analyst coverage is in the bottom (top) quintile across industry-years. Similarly, we classify borrowers as having a weak (strong) information environment if analysts' forecast dispersion is in the top (bottom) quintile across industry-years. Table 10 provides regression results for firms' information environment, dissimilarity measures, and the cost of bank loans. Consistent with our expectations, our findings report that the relationship between financial statement dissimilarity and loan costs is restricted to borrowers with low analyst coverage and high forecast dispersion, i.e., for borrowers with a weak information environment.

[INSERT TABLE 10 HERE]

Third, we also examine the effect of relationship lending on the association between financial statement dissimilarity and the cost of bank loans. Relationship lending facilitates banks' access to alternative sources of valuable information (Bharath et al., 2011). We argue that since financial statement dissimilarity contains information related to firm performance, such information should be more valuable if banks do not have an existing lending relationship with borrowers and thus have no access to borrowers' soft information through relationship lending. Following Bushman et al. (2017), we categorize each loan into relationship lending if the lead arranger's allocation share is more than 50% of the loan amount in the previous five years and remaining as non-relationship lending. Our results in Table 11 support that financial

statement dissimilarity increases bank loan spreads for the non-relationship lending group. Banks find the information contained in the dissimilarity of financial statements more peculiar when they do not have access to borrowers' other valuable soft information through relationship lending.

[INSERT TABLE 11 HERE]

Finally, we also explore the effect of firm-level corporate governance mechanisms on the relationship between financial statement dissimilarity and the cost of bank loans. Prior studies show that monitoring from institutional investors and board of directors play an important role in improving firm performance (Shleifer & Vishny, 1986; Coles et al., 2014; Borochin & Yang, 2017; Masulis & Zhang, 2019; Liu et al., 2020; Kim et al., 2022). Based on these studies, we expect that if financial statement dissimilarity is an indicator of weak firm performance, such concern will be more pronounced for borrowers with a weak level of monitoring from institutional investors and the board of directors. We consider institutional ownership and board co-option as our indicators of firm-level corporate governance mechanisms. Coles et al. (2014) document that directors appointed by the CEO share allegiance to the CEO which ultimately decreases their monitoring of firm affairs.

Table 12 provides regression results for firm-level corporate governance, dissimilarity measures, and the cost of bank loans. We classify borrowers as having a weak (strong) corporate governance if the institutional ownership is in the bottom (top) quintile across industry-years. Similarly, we classify borrowers as having a weak (strong) corporate governance if the level of co-opted board members is in the top (bottom) quintile across industry-years. Our findings support that the documented effect of financial statement dissimilarity on loan costs is concentrated among borrowers with low levels of institutional ownership and high levels of co-opted board members, i.e., for borrowers with a weak corporate governance. These findings suggest that if financial statement dissimilarity is an

indicator of weak firm performance, such concerns will be more pronounced for borrowers with a weak level of monitoring from institutional investors and the board of directors.

[INSERT TABLE 12 HERE]

6. Conclusion

Financial statements are the dominant source of information about firms' past and future performance for various stakeholders. Year-over-year textual dissimilarity of 10-K reports may convey vital information about the firm, but this information may not be well-understood by equity market participants (Cohen et al., 2020). We examine the consequences of dissimilar 10-K reports on the pricing and design of bank loan contracts. Our main results suggest that bankers consider the textual dissimilarity of 10-K filings while extending credit to the borrowing firms. Other things being equal, on average, firms that present dissimilar 10-K reports experience a higher cost of bank loans, compared to firms that present similar financial reports. The association between financial statement dissimilarity and loan spreads is not driven by the influence of financial statement readability and comparability on loan costs.

Further tests suggest that higher dissimilarity in the financial reports indicates lower firm performance and a higher probability of violation of performance covenants. Banks also impose stricter covenants, especially performance covenants, for firms with higher financial statement dissimilarity. From the cross-sectional perspective, our findings report that 10-K textual dissimilarity matters more for the loan costs of firms that have a more ambiguous tone of financial reports, that are more informationally opaque (those with lower analyst coverage and higher forecast dispersion), that do not share a lending relationship with banks, or subject to lower monitoring from institutional investors or board of directors.

Taken in their entirety, our results provide new evidence on the importance of financial statements in general, and financial statement dissimilarity in particular, in the lending process.

In particular, our findings address SEC's concerns over users' attention to non-repeated disclosures in financial statements, suggesting that, unlike equity market participants who may remain inattentive to non-repeated material information, banks are acutely aware of the relevance of financial statement textual dissimilarity for the borrower's performance when deciding the pricing and covenant structure of loan contracts.

References

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. *Journal of Human Resources*, *40*(4), 791-821.

Anantharaman, D., Fang, V. W., & Gong, G. (2014). Inside debt and the design of corporate debt contracts. *Management Science*, *60*(5), 1260-1280.

Anderson, J. D., & Core, J. E. (2018). Managerial incentives to increase risk provided by debt, stock, and options. *Management Science*, *64*(9), 4408-4432.

Balachandran, B., Duong, H. N., & Vu, V. H. (2019). Pension deficits and the design of private debt contracts. *Journal of Financial and Quantitative Analysis*, *54*(4), 1821-1854.

Berg, T., Saunders, A., & Steffen, S. (2016). The total cost of corporate borrowing in the loan market: Don't ignore the fees. *Journal of Finance*, *71*(3), 1357-1392.

Bharath, S. T., Sunder, J., & Sunder, S. V. (2008). Accounting quality and debt contracting. *The Accounting Review*, *83*(1), 1-28.

Bharath, S. T., Dahiya, S., Saunders, A., & Srinivasan, A. (2011). Lending relationships and loan contract terms. *Review of Financial Studies*, *24*(4), 1141-1203.

Blankespoor, E., deHaan, E., & Marinovic, I. (2020). Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics*, *70*(2-3), 101344.

Bonsall, S. B., Leone, A. J., Miller, B. P., & Rennekamp, K. (2017). A plain English measure of financial reporting readability. *Journal of Accounting and Economics*, *63*(2-3), 329-357.

Borochin, P., & Yang, J. (2017). The effects of institutional investor objectives on firm valuation and governance. *Journal of Financial Economics*, *126*(1), 171-199.

Bradley, M., & Roberts, M. R. (2015). The structure and pricing of corporate debt covenants. *Quarterly Journal of Finance*, *5*(2), 1550001.

Brown, S. V., & Tucker, J. W. (2011). Large-sample evidence on firms' year-over-year MD&A modifications. *Journal of Accounting Research*, *49*(2), 309-346.

Brown, N. C., Gale, B. T., & Grant, S. M. (2022). How Do Disclosure Repetition and Interactivity Influence Investors' Judgments? *Journal of Accounting Research*, *60*(5), 1775-1811.

Bushman, R. M., Williams, C. D., & Wittenberg-Moerman, R. (2017). The informational role of the media in private lending. *Journal of Accounting Research*, *55*(1), 115-152.

Cazier, R.A., & Pfeiffer, R. J., (2017). 10-K disclosure repetition and managerial reporting incentives. *Journal of Financial Reporting, 2*(1), 107-131.

Chakraborty, I., Leone, A. J., Minutti-meza, M., & Phillips, M. A. (2022). Financial statement complexity and bank lending. *The Accounting Review, 97*(3), 155-178.

Chava, S., & M. R. Roberts. 2008. How does financing impact investment? The role of debt covenants. *Journal of Finance, 63*(5), 2085–2121.

Chen, Y., Kelly, B., & Wu, W. (2020). Sophisticated investors and market efficiency: Evidence from a natural experiment. *Journal of Financial Economics*, *138*(2), 316-341.

Chen, Y., Ofosu, E., Veeraraghavan, M., & Zolotoy, L. (2023). Does CEO overconfidence affect workplace safety? *Journal of Corporate Finance*, 102430.

Cinelli, C., & Hazlett, C. (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, *82*(1), 39-67.

Cohen, L., Malloy, C., & Nguyen, Q. (2020). Lazy prices. *Journal of Finance*, *75*(3), 1371-1415.

Coles, J. L., Daniel, N. D., & Naveen, L. (2014). Co-opted boards. *Review of Financial Studies*, *27*(6), 1751-1796.

De Franco, G., Kothari, S. P., & Verdi, R. S. (2011). The benefits of financial statement comparability. *Journal of Accounting Research*, *49*(4), 895-931.

Degryse, H., Goncharenko, R., Theunisz, C., & Vadasz, T. (2023). When green meets green. *Journal of Corporate Finance*, 102355.

Demerjian, P. R., & Owens, E. L. (2016). Measuring the probability of financial covenant violation in private debt contracts. *Journal of Accounting and Economics*, *61*(2-3), 433-447.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *Review of Economic Studies*, *51*(3), 393-414.

Diamond, D. W. (1991). Monitoring and reputation: The choice between bank loans and directly placed debt. *Journal of Political Economy*, *99*(4), 689-721.

Donelson, D. C., R. Jennings, & J. McInnis. 2017. Financial statement quality and debt contracting: Evidence from a survey of commercial lenders. *Contemporary Accounting Research, 34*(4), 2051–2093.

Dyer, T., Lang, M., & Stice-Lawrence, L. (2017a). The evolution of 10-K textual disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics*, *64*(2-3), 221-245.

Dyer, T. A., Lang, M. H., & Stice-Lawrence, L. (2017b). Discussion of: 10-K Disclosure Repetition and Managerial Reporting Incentives: What Have We Learned and Where Do We Go with Textual Research?. *Journal of Financial Reporting*, *2*(1), 133-138.

Ertugrul, M., Lei, J., Qiu, J., & Wan, C. (2017). Annual report readability, tone ambiguity, and the cost of borrowing. *Journal of Financial and Quantitative Analysis*, *52*(2), 811-836.

Fama, E. F. (1985). What's different about banks?. *Journal of Monetary Economics*, *15*(1), 29-39.

Fang, X., Li, Y., Xin, B., & Zhang, W. (2016). Financial statement comparability and debt contracting: Evidence from the syndicated loan market. *Accounting Horizons*, *30*(2), 277-303.

Frank, K. A. (2000). Impact of a confounding variable on a regression coefficient. *Sociological Methods & Research*, *29*(2), 147-194.

Frank, K. A., Maroulis, S. J., Duong, M. Q., & Kelcey, B. M. (2013). What would it take to change an inference? Using Rubin's causal model to interpret the robustness of causal inferences. *Educational Evaluation and Policy Analysis*, *35*(4), 437-460.

Graham, J. R., Li, S., & Qiu, J. (2008). Corporate misreporting and bank loan contracting. *Journal of Financial Economics*, *89*(1), 44-61.

Hallman, N., Howe, J. S., & Wang, W. (2023). Analyst coverage and syndicated lending. *Review of Accounting Studies*, *28*(3), 1531-1569.

Hasan, I., Hoi, C. K. S., Wu, Q., & Zhang, H. (2014). Beauty is in the eye of the beholder: The effect of corporate tax avoidance on the cost of bank loans. *Journal of Financial Economics*, *113*(1), 109-130.

Hasan, M. M., Lobo, G. J., & Qiu, B. (2021). Organizational capital, corporate tax avoidance, and firm value. *Journal of Corporate Finance*, *70*, 102050.

Honigsberg, C., Katz, S. P., Mutlu, S., & Sadka, G. (2021). State contract law and the use of accounting information in debt contracts. *Review of Accounting Studies*, *26*, 124-171.

Hope, O. K., Porumb, V. A., Rusanescu, S., & Vyas, D. (2023). Private information and bank-loan pricing: The effect of upcoming corporate spinoffs. *Contemporary Accounting Research*. Forthcoming.

Ivashina, V. (2009). Asymmetric Information Effects on Loan Spreads. *Journal of Financial Economics, 92*(2), 300-319.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305-360.

Jeon, Y., McCurdy, T. H., & Zhao, X. (2022). News as sources of jumps in stock returns: Evidence from 21 million news articles for 9000 companies. *Journal of Financial Economics*, *145*(2), 1-17.

Kim, S., Kraft, P., & Ryan, S. G. (2013). Financial statement comparability and credit risk. *Review of Accounting Studies*, *18*, 783-823.

Kim, D., Wang, Q., & Wang, X. (2022). Geographic clustering of institutional investors. *Journal of Financial Economics*, *144*(2), 547-570.

Larcker, D. F., & Rusticus, T. O. (2010). On the use of instrumental variables in accounting research. *Journal of Accounting and Economics*, *49*(3), 186-205.

Lehavy, R., Li, F., & Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review*, *86*(3), 1087-1115.

Lewbel, A. (2012). Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, *30*(1), 67-80.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics, 45*, 221–247.

Li, H. (2019). Repetitive Disclosures in the MD&A. *Journal of Business Finance & Accounting*, *46*(9-10), 1063-1096.

Linnainmaa, J. T., Torous, W., & Yae, J. (2016). Reading the tea leaves: Model uncertainty, robust forecasts, and the autocorrelation of analysts' forecast errors. *Journal of Financial Economics*, *122*(1), 42-64.

Liu, C., Low, A., Masulis, R. W., & Zhang, L. (2020). Monitoring the monitor: Distracted institutional investors and board governance. *Review of Financial Studies*, *33*(10), 4489-4531.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance*, *66*(1), 35-65.

Loughran, T., & McDonald, B. (2014). Measuring readability in financial disclosures. *Journal of Finance*, *69*(4), 1643-1671.

Loughran, T., & Mcdonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, *54*(4), 1187-1231.

Ma, Z., Novoselov, K. E., Stice, D., & Zhang, Y. (2022). Firm innovation and covenant tightness. *Review of Accounting Studies*, 1-43.

Masulis, R. W., & Zhang, E. J. (2019). How valuable are independent directors? Evidence from external distractions. *Journal of Financial Economics, 132*(3), 226-256.

Mavis, C. P., McNamee, N. P., Petmezas, D., & Travlos, N. G. (2020). Selling to buy: Asset sales and acquisitions. *Journal of Corporate Finance*, *62*, 101587.

Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics*, *5*(2), 147-175.

Nini, G., D. C. Smith, & A. Sufi. 2012. Creditor control rights, corporate governance, and firm value. *Review of Financial Studies, 25*(6), 1713–1761.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187-204.

Rajan, R., & Winton, A. (1995). Covenants and collateral as incentives to monitor. *Journal of Finance*, *50*(4), 1113-1146.

Regenburg, K., & Seitz, M. N. B. (2021). Criminals, bankruptcy, and cost of debt. *Review of Accounting Studies*, *26*(3), 1004-1045.

Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, *39*(3), 437-485.

Roychowdhury, S., Shroff, N., & Verdi, R. S. (2019). The effects of financial reporting and disclosure on corporate investment: A review. *Journal of Accounting and Economics*, *68*(2-3), 101246.

Securities and Exchange Commission (SEC) (2018). Inline XBRL Filing of Tagged Data. Release No. 33–105142; 34–83551; IC-33139; File No. S7-03-17.

Securities and Exchange Commission (SEC) (2019). FAST Act Modernization and Simplification of Regulation S-K: Final Rule. Release No. 33–10618; 34–85381; IA-5206; IC-33426; File No. S708-17.

Securities and Exchange Commission (SEC) (2020). Management's Discussion and Analysis, Selected Financial Data, and Supplementary Financial Information. Release No. 33–10750; 34–88093; IC33795; File No. S7-01-20.

Sengupta, P. (1998). Corporate disclosure quality and the cost of debt. *The Accounting Review*, *73*(4), 459-474.

Shleifer, A., & Vishny, R. W. (1986). Large shareholders and corporate control. *Journal of Political Economy*, *94*(3), 461-488.

Smith, C. W., & Warner, J. B. (1979). On financial contracting: An analysis of bond covenants. *Journal of Financial Economics*, *7*(2), 117-161.

Vashishtha, R. (2014). The role of bank monitoring in borrowers′ discretionary disclosure: Evidence from covenant violations. *Journal of Accounting and Economics*, *57*(2-3), 176-195.

Appendix

Table A1: Variable Definitions

Variables	Definitions
Jaccard	Similarity measure obtained from the WRDS SEC analytics database, multiplied by -1 to ease its interpretation alongside Mineditdist
Mineditdist	Similarity measure obtained from the WRDS SEC analytics database
ln(Loan Spread)	Natural logarithm of the all-in-drawn spread in the Dealscan database
ln(Maturity)	Natural logarithm of loan maturity. Loan maturity is expressed in months in the Dealscan database
ln(Loan Size)	Natural logarithm of loan facility amount
Firm Size	Natural logarithm of total assets
Leverage	Total debt to total assets
Tangibility	Net plant, property, and equipment to total assets
Cash Holdings	Cash and short-term investments to total assets
ROA	Earnings before interest, taxes, depreciation, and amortization divided by total assets
MB	Market value to book value of assets
Sales Growth	Growth rate of sales from *t-1* to *t*
CF Vol.	Rolling standard deviation (5 quarters) of net cash flow from operating activities
Z-score	Altman Z-score is the ratio of 1.2*working capital + 1.4*retained earnings + 3.3*pre-tax income + 0.999*sales to total assets
Credit Rating	S&P Domestic Long-term issuer credit rating of firms. Larger number indicates lower credit rating.
Syndication	An indicator variable equal to 1 for syndicated loans and 0 otherwise
ln(File Size)	Natural logarithm of the file size for the 10-K filing
Fog Index	Financial statement readability measure obtained from the WRDS SEC analytics database
Bog Index	Financial statement readability measure obtained from Bonsall et al. (2017)
CompAcc4	Financial statement comparability measure obtained from De Franco et al. (2011)
CompAcc10	Financial statement comparability measure obtained from De Franco et al. (2011)
CompAccInd	Financial statement comparability measure obtained from De Franco et al. (2011)
NI Volatility	Standard deviation of net income as a proportion of total assets over the past five years
Default	An indicator variable equal to 1 if the Altman's Z-score value falls below 1.81 and 0 otherwise
All Covenants	Likelihood of violation of all covenants obtained from Demerjian and Owens (2016)
Performance Covenants	Likelihood of violation of performance covenants obtained from Demerjian and Owens (2016)
Capital Covenants	Likelihood of violation of capital covenants obtained from Demerjian and Owens (2016)
Covenants Intensity Index	Index reflecting intensity of the covenants, developed following Bradley and Roberts (2015)
Performance Covenants Ratio	Proportion of performance covenants to total covenants (including capital covenants)
Uncertain	Proportion of "uncertain" words in a financial statement obtained from the WRDS SEC analytics database
Strong Modal	Proportion of "strong modal" words in a financial statement obtained from the WRDS SEC analytics database
Analyst Coverage	Number of analysts following a particular firm

Forecast dispersion	Dispersion of analysts' forecasts
Relationship Lending	Loans when the lead arranger's allocation share is more than 50% of the loan amount in the previous five years
Institutional Ownership	Percentage of shares held by institutional investors
Board Co-option	Proportion of co-opted board members

Table 1: Descriptive Statistics

This table reports the descriptive statistics for the data used to examine the relationship between dissimilarity measures and bank loans. Financial and utility firms are excluded from the analysis. The sample period ranges from 1995 to 2017. All the variable definitions are provided in the appendix.

Variables	N	Mean	S.D.	p25	p50	p75
Jaccard	16,647	-0.66	0.19	-0.80	-0.70	-0.59
Mineditdist	16,647	0.43	0.25	0.23	0.37	0.60
Loan Spread	16,647	230.85	159.65	125.00	200.00	300.00
Maturity (Months)	16,647	49.73	22.44	36.00	60.00	60.00
Loan Size (in $US million)	16,647	440.00	1090.00	50.00	150.00	425.00
Firm Size (in $US million)	16,647	5991.30	18747.99	340.77	1230.08	4168.06
Leverage	16,647	0.31	0.23	0.15	0.28	0.43
Tangibility	16,647	0.29	0.23	0.11	0.22	0.43
Cash Holdings	16,647	0.10	0.12	0.02	0.05	0.13
ROA	16,647	0.13	0.09	0.09	0.13	0.17
MB	16,647	1.76	0.96	1.17	1.48	2.01
Sales Growth	16,647	0.13	0.32	-0.01	0.08	0.20
CF Vol.	16,647	184.49	430.87	12.16	43.00	142.75
Z-score	16,647	1.43	1.57	0.74	1.55	2.34
Credit Rating	16,647	16.11	6.06	11.00	15.00	22.00
Syndication	16,647	0.95	0.22	1.00	1.00	1.00

Table 2: Baseline Regression

This table provides regression results for financial statement dissimilarity and its impact on bank loan spreads. The dependent variable is the logarithm of all-in-drawn loan spread. Industry, year, loan type and primary purpose dummies are included in all the regression specifications. The test-statistics - standard errors clustered by firms – are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Variables	**Jaccard**		**Mineditdist**	
Dissimilarity	0.247***	0.117***	0.181***	0.0936***
	(6.37)	(3.99)	(6.16)	(4.00)
Firm Size		-0.0356***		-0.0356***
		(-3.86)		(-3.85)
Leverage		0.549***		0.548***
		(16.11)		(16.06)
Tangibility		-0.0335		-0.0340
		(-0.65)		(-0.66)
Cash Holdings		0.146**		0.147**
		(2.33)		(2.34)
ROA		-0.925***		-0.928***
		(-9.87)		(-9.88)
MB		-0.0921***		-0.0921***
		(-10.04)		(-10.03)
Sales Growth		0.0237		0.0240
		(1.50)		(1.52)
CF Vol.		-0.0001***		-0.0001***
		(-4.31)		(-4.32)
Z-score		-0.0470***		-0.0470***
		(-8.58)		(-8.59)
ln(Maturity)		-0.000462		-0.000380
		(-0.03)		(-0.03)
ln(Loan Size)		-0.121***		-0.121***
		(-13.68)		(-13.65)
Credit Rating		0.0148***		0.0148***
		(8.29)		(8.28)
Syndication		0.000497		-0.000307
		(0.02)		(-0.01)
Constant	5.346***	6.055***	5.105***	5.939***
	(197.16)	(69.40)	(315.97)	(67.27)
Observations	16,631	16,631	16,631	16,631
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.42	0.63	0.42	0.63

Table 3: Lewbel's (2012) IV Regression

This table provides the results of the instrumental variable (IV) regression approach using Lewbel's (2012) method of identifying instrumental variables. The dependent variable is the logarithm of all-in-drawn loan spreads. Industry, year, loan type and primary purpose dummies are included in all the regression specifications. The test-statistics - standard errors clustered by firms – are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Variables	**Jaccard**	**Mineditdist**
Dissimilarity	0.117***	0.0936***
	(4.01)	(4.01)
Controls	Yes	Yes
Observations	16,647	16,647
Industry FEs	Yes	Yes
Year FEs	Yes	Yes
Loan Type FEs	Yes	Yes
Primary Purpose FEs	Yes	Yes
Hansen J stats	157.13	154.72
[p-value]	[0.21]	[0.26]
R^2	0.38	0.38

Table 4: Impact Threshold for Confounding Variable and Coefficient Stability Tests

This table provides the results for the ITCV and coefficient stability tests. Panel A presents the results for the ITCV analysis. The ITCV values reveal the minimum correlation a confounding variable need to have with the dissimilarity measure and the cost of bank loans to make the coefficient of the dissimilarity measure statistically insignificant at the 10% level. Below ITCV values, the impact values for the control variables are reported. Panel B reports the Oster test results. Baseline effect excludes our set of confounders while regressing the cost of bank loans against the dissimilarity measures. Controlled effect includes all controls while regressing the cost of bank loans against the dissimilarity measures. Identified set is bounded below by the bias-adjusted coefficient of the respective dissimilarity measures and above by the controlled regression coefficient. $\tilde{\delta}$ is the estimated delta.

Panel A: ITCV Analysis

Variables	**Jaccard**	**Mineditdist**
	Partial Impact	**Partial Impact**
ITCV	0.0160	0.0160
Leverage	0.0078	0.0108
ln(Loan Size)	0.0061	0.0063
ROA	0.0041	0.0037
Z-score	-0.0002	-0.0016
Syndication	-0.0002	-0.0001
MB	-0.0014	-0.001
Tangibility	-0.0014	-0.0015
CF Vol.	-0.0020	-0.0016
Firm Size	-0.0022	-0.0022
Sales Growth	-0.0023	-0.0021
Cash Holdings	-0.0041	-0.0048
Credit Rating	-0.0063	-0.0058
ln(Maturity)	-0.0144	-0.0155

Panel B: Coefficient Stability Test

Variables	**Jaccard**	**Mineditdist**
Baseline Effect	0.124^{***}	0.107^{***}
[R^2]	[0.32]	[0.32]
Controlled Effect	0.117^{***}	0.094^{***}
[R^2]	[0.64]	[0.64]
Identified Set	0.112, 0.117	0.083, 0.094
Confidence Interval	0.059 to 0.174	0.048 to 0.140
$\tilde{\delta}$	7.39	5.10

Table 5: Propensity Score Matching

This table provides propensity score matching results to examine the impact of dissimilarity measures on the bank loan spreads. Dissimilarity measures (Jaccard and Mineditdist) with values greater than industry-year median are categorized as treatment observations and matched with control observations from the same industry and year. Observations are matched based on the closest propensity scores calculated using the determinants, such as firm size, leverage, tangibility, cash holdings, ROA, MB, sales growth, cash flow volatility, Z-score, and credit ratings along with industry and year fixed effects. Panel A reports the mean differences between the treatment and control observations for the matched sample. Panel B reports the regression results for the dissimilarity measures and the bank loan spreads. The dependent variable is the logarithm of all-in-drawn loan spread. Industry, year, loan type and primary purpose dummies are included in both the regression specifications. The test-statistics - standard errors clustered by firms – are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Mean Differences between Treatment and Control Observations

Variable	Jaccard				Mineditdist			
	Treatment	Control	Diff	t-stats	Treatment	Control	Diff	t-stats
Firm Size	7.167	7.092	0.075	1.41	7.057	7.038	0.019	0.36
Leverage	0.295	0.286	0.009	1.33	0.277	0.288	-0.011*	-1.80
Tangibility	0.276	0.276	0.000	-0.03	0.268	0.274	-0.006	-0.89
Cash Holdings	0.101	0.104	-0.003	-0.80	0.103	0.100	0.003	0.81
ROA	0.129	0.129	0.000	-0.21	0.128	0.129	-0.001	-0.58
MB	1.736	1.764	-0.029	-1.05	1.747	1.779	-0.032	-1.20
Sales Growth	0.105	0.115	-0.010	-1.31	0.106	0.113	-0.007	-0.93
CF Vol.	187.180	174.340	12.840	1.02	188.620	186.590	2.030	0.16
Z-score	1.421	1.452	-0.031	-0.72	1.482	1.450	0.032	0.78
Credit Rating	16.109	16.374	-0.265	-1.47	16.185	16.286	-0.101	-0.58

Panel B: Regression Results: Matched Sample

Variables	Jaccard		Mineditdist	
Dissimilarity	0.255***	0.158***	0.146***	0.086***
	(4.81)	(3.76)	(3.54)	(2.63)
Controls	No	Yes	No	Yes
Observations	4,857	4,857	5,218	5,218
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.43	0.64	0.41	0.64

Table 6: Dissimilarity Measures, Financial Statement Readability, and Comparability

This table reports the results of dissimilarity measures, financial statement readability, financial statement comparability, and the bank loan spreads. In Panel A, we include financial readability, i.e., logarithm of file size, fog index and bog index, as one of our control variables. In Panel B, we also include financial statement comparability measures in our regressions. The test-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Financial Statement Readability

Variables		**Jaccard**			**Mineditdist**	
Dissimilarity	0.0886***	0.1043***	0.1163***	0.0672***	0.0799***	0.0902***
	(2.97)	(3.56)	(3.89)	(2.81)	(3.39)	(3.77)
ln(File Size)	0.0465***			0.0458***		
	(4.80)			(4.71)		
Fog Index		0.0228***			0.0224***	
		(3.95)			(3.85)	
Bog Index			0.0075***			0.0075***
			(5.96)			(5.94)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,546	16,546	16,283	16,546	16,546	16,283
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.64	0.64	0.64	0.64	0.64	0.64

Panel B: Financial Statement Comparability

Variables		**Jaccard**			**Mineditdist**	
Dissimilarity	0.1203***	0.1208***	0.1190***	0.1055***	0.1055***	0.1040***
	(3.06)	(3.08)	(3.02)	(3.40)	(3.41)	(3.35)
CompAcct4	-0.0370***			-0.0369***		
	(-4.17)			(-4.16)		
CompAcct10		-0.0305***			-0.0304***	
		(-4.35)			(-4.34)	
CompAcctInd			-0.0211***			-0.0210***
			(-4.43)			(-4.42)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,165	9,165	9,165	9,165	9,165	9,165
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.66	0.66	0.66	0.66	0.66	0.66

Table 7: Robustness Tests

This table reports the results of several robustness tests performed on the bank loan spreads. For brevity, the table reports the coefficients of our main variables of interest. Other loan-level as well as firm-level variables are similar to those employed in the baseline regression along with fixed effects. Model (1) considers an alternative dissimilarity measure, i.e., Cosine as our main variable of interest. Model (2) uses two-way clustering at the firm and lead bank level. Model (3) uses the overall cost of borrowing including various fees, as in Berg et al. (2016), as the dependent variable. Model (4) includes only the largest loan facility within a loan package per year. Model (5) excludes observations in the financial turmoil period (2007-2008). Model (6) excludes all loans granted after 2007. Model (7) employs firm, year, loan type and primary purpose fixed effects. Model (8) uses the median regression with robust standard errors. Model (9) includes dissimilarity measures two- and three-years prior to loan origination. Model (10) also controls for the proxies of firm's information environment (analyst forecast and forecast dispersion). Model (11) excludes loans for borrowers with acquisition-related expenditures one-year before the loan origination. The test-statistics are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	Dissimilarity Measures			**Pseudo/Adjusted R^2**			**Observations**
	Jaccard	Mineditdist	Cosine	Jaccard	Mineditdist	Cosine	
(1) Cosine as a measure of dissimilarity			0.3279**			0.63	16,631
			(2.51)				
(2) Firm and lead bank two-way clustering	0.1153***	0.0908***		0.63	0.63		15,803
	(3.44)	(3.60)					
(3) Total borrowing cost (Berg et al., 2016)	0.1231***	0.0994***		0.63	0.63		16,631
	(4.17)	(4.22)					
(4) Include only largest loan facility	0.0812***	0.0795***		0.66	0.66		11,776
	(2.72)	(3.43)					
(5) Exclude financial crisis period (2007-2008)	0.1209***	0.0972***		0.64	0.64		14,948
	(3.95)	(3.98)					
(6) Exclude post-2007 period	0.1414***	0.1267***		0.65	0.65		9,993
	(3.77)	(4.20)					
(7) Firm fixed effects regression	0.0860***	0.0610**		0.75	0.75		15,831
	(2.84)	(2.53)					
(8) Median regression	0.1094***	0.0873***		0.40	0.40		16,647
	(6.52)	(6.42)					
(9) Financial statement dissimilarity:							
(a) Two-years prior to loan origination	0.1221***	0.0985***		0.64	0.64		13,363
	(3.69)	(3.78)					
(b) Three-years prior to loan origination	0.1226***	0.1007***		0.65	0.65		11,757
	(3.47)	(3.68)					
(10) Controlling for information environment:							
(a) Analyst Coverage	0.1004***	0.0804***		0.64	0.64		16,631
	(3.43)	(3.43)					
(b) Analyst Forecast Dispersion	0.1066***	0.0789***		0.66	0.66		10,629
	(2.92)	(2.78)					
(11) Excluding loans for borrowers with acquisition-related expenditures one-year prior to loan origination	0.1069***	0.0979***		0.61	0.61		7,948
	(2.70)	(3.08)					

Table 8: Firm Performance, Covenant Violations, and Covenant Restrictions

This table provides the regression results for the impact of dissimilarity measures at *t-1* (one year before the loan initiation) on firm performance, covenant violations, and covenant restrictions at *t*. Panel A reports results for firm performance, including return-on-assets (ROA), Net Income Volatility (NI Volatility) – net income as a proportion of total assets, and the likelihood of a default through a logit model. The variable 'default' is an indicator variable, i.e., equal to 1 if the Altman's Z-score values fall below 1.81 (distress level), and 0 otherwise. Panel B reports the likelihood of covenant violations. Data related to covenant violations are gathered from Demerjian and Owens (2016). Panel C presents the results for covenants intensity index and performance covenants ratio. The covenants intensity index is developed following Bradley and Roberts (2015). Performance covenants ratio is the proportion of performance covenants to total covenants (including capital covenants). Industry and year fixed effects are included in all the regression specifications. The test-statistics - standard errors clustered by firms – are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Firm Performance

Variables	ROA		NI Volatility		Default	
	Jaccard	**Mineditdist**	**Jaccard**	**Mineditdist**	**Jaccard**	**Mineditdist**
Dissimilarity	-0.0131***	-0.0114***	0.0109***	0.00718***	0.725**	0.629***
	(-2.61)	(-2.87)	(4.27)	(3.68)	(2.57)	(2.88)
Firm Size	0.00188*	0.00190*	-0.00551***	-0.00551***	-0.248***	-0.248***
	(1.68)	(1.69)	(-10.76)	(-10.75)	(-4.14)	(-4.13)
Leverage	0.0775***	0.0777***	-0.00327	-0.00330	4.214***	4.206***
	(10.80)	(10.83)	(-0.76)	(-0.77)	(14.93)	(14.84)
Tangibility	0.0464***	0.0464***	-0.0145***	-0.0146***	1.728***	1.728***
	(5.91)	(5.91)	(-3.11)	(-3.12)	(3.91)	(3.90)
Cash Holdings	-0.0837***	-0.0838***	0.0361***	0.0362***	2.238***	2.247***
	(-6.39)	(-6.40)	(5.86)	(5.86)	(4.37)	(4.40)
MB	0.0295***	0.0295***	0.00284***	0.00285***	0.169**	0.170**
	(17.33)	(17.34)	(3.54)	(3.55)	(2.22)	(2.23)
Sales Growth	-0.00625*	-0.00626*	0.000382	0.000481	-0.104	-0.103
	(-1.75)	(-1.75)	(0.18)	(0.23)	(-0.77)	(-0.76)
CF Vol.	-0.000001	-0.000001			0.0005***	0.0005***
	(-0.26)	(-0.27)			(2.67)	(2.66)
Z-Score	0.0275***	0.0275***	-0.0134***	-0.0134***		
	(20.76)	(20.77)	(-13.14)	(-13.15)		
Credit Rating	-0.001**	-0.001**	-0.00003	-0.00003	0.0328**	0.0327**
	(-2.26)	(-2.24)	(-0.18)	(-0.18)	(2.10)	(2.09)
ROA			0.0150	0.0145	-10.42***	-10.44***
			(1.18)	(1.15)	(-13.68)	(-13.73)
Constant	-0.0144	-0.00110	0.0924***	0.0822***	-4.391***	-5.124***
	(-1.11)	(-0.09)	(14.04)	(12.35)	(-3.02)	(-3.53)
Observations	15,771	15,771	16,089	16,089	15,054	15,054
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted/Pseudo R^2	0.32	0.32	0.31	0.31	0.35	0.36

Panel B: Probability of Covenant Violations

Variables	All Covenants		Performance Covenants		Capital Covenants	
	Jaccard	**Mineditdist**	**Jaccard**	**Mineditdist**	**Jaccard**	**Mineditdist**
Dissimilarity	0.0481	0.0347	0.0790***	0.0640***	-0.0232	-0.0184
	(1.64)	(1.55)	(2.68)	(2.85)	(-1.25)	(-1.27)
Firm Size	-0.0347***	-0.0348***	-0.0227***	-0.0228***	-0.0235***	-0.0234***
	(-4.22)	(-4.23)	(-2.74)	(-2.76)	(-3.94)	(-3.93)
Leverage	0.576***	0.575***	0.623***	0.622***	0.0784***	0.0786***
	(15.92)	(15.91)	(16.74)	(16.70)	(3.22)	(3.22)
Tangibility	-0.0565	-0.0571	-0.126***	-0.126***	0.104***	0.104***
	(-1.40)	(-1.42)	(-3.12)	(-3.14)	(3.40)	(3.40)
Cash Holdings	-0.0273	-0.0277	-0.0694	-0.0697	0.0578	0.0579
	(-0.46)	(-0.47)	(-1.14)	(-1.15)	(1.51)	(1.52)
ROA	-1.082***	-1.082***	-1.027***	-1.026***	-0.284***	-0.284***
	(-12.36)	(-12.36)	(-11.08)	(-11.08)	(-4.15)	(-4.15)
MB	-0.0412***	-0.0412***	-0.0487***	-0.0487***	0.00251	0.00251
	(-5.67)	(-5.68)	(-6.50)	(-6.51)	(0.49)	(0.49)
Sales growth	0.00167	0.00205	-0.0111	-0.0108	0.0254*	0.0253*
	(0.09)	(0.11)	(-0.56)	(-0.55)	(1.67)	(1.66)
CF Vol.	0.000003	0.000003	-0.00004	-0.00004	0.0001***	0.0001***
	(0.09)	(0.11)	(-1.62)	(-1.58)	(3.16)	(3.14)
Z-score	-0.00296	-0.00307	0.0100*	0.00986*	-0.00605	-0.00600
	(-0.54)	(-0.56)	(1.70)	(1.67)	(-1.50)	(-1.49)
ln(Loan Size)	-0.00807	-0.00811	-0.00249	-0.00256	-0.00767	-0.00765
	(-1.18)	(-1.18)	(-0.36)	(-0.37)	(-1.50)	(-1.49)
Credit Rating	0.00596***	0.00596***	0.00750***	0.00748***	-0.00177*	-0.00176*
	(3.96)	(3.96)	(5.15)	(5.14)	(-1.94)	(-1.94)
Constant	0.626***	0.580***	0.459***	0.381***	0.274***	0.296***
	(9.48)	(8.92)	(6.84)	(5.88)	(6.47)	(7.06)
Observations	5,594	5,594	5,594	5,594	5,594	5,594
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R^2	0.31	0.31	0.28	0.28	0.15	0.15

Panel C: Covenants Intensity Index and Performance Covenants Ratio

Variables	Covenants Intensity Index		Performance Covenants Ratio	
	Jaccard	Mineditdist	Jaccard	Mineditdist
Dissimilarity	0.0982**	0.0829***	0.0407**	0.0403***
	(2.52)	(2.62)	(2.12)	(2.66)
Firm Size	-0.164***	-0.165***	-0.0172**	-0.0173**
	(-11.66)	(-11.66)	(-2.52)	(-2.53)
Leverage	0.357***	0.356***	0.209***	0.208***
	(7.38)	(7.36)	(8.98)	(8.95)
Tangibility	-0.193***	-0.194***	-0.142***	-0.142***
	(-3.20)	(-3.22)	(-4.26)	(-4.27)
Cash Holdings	0.130	0.129	-0.104**	-0.104**
	(1.48)	(1.47)	(-2.30)	(-2.30)
ROA	0.358***	0.357***	0.286***	0.287***
	(3.31)	(3.30)	(4.32)	(4.34)
MB	-0.130***	-0.130***	-0.0211***	-0.0210***
	(-10.13)	(-10.14)	(-3.89)	(-3.88)
Sales growth	0.0593***	0.0595***	-0.0253*	-0.0255**
	(2.63)	(2.64)	(-1.96)	(-1.97)
CF Vol.	-0.000721***	-0.000720***	-0.000151***	-0.000151***
	(-4.54)	(-4.53)	(-4.02)	(-4.02)
Z-score	-0.0318***	-0.0318***	-0.00645*	-0.00644*
	(-4.33)	(-4.35)	(-1.66)	(-1.66)
ln(Loan Size)	0.244***	0.243***	0.0470***	0.0469***
	(20.24)	(20.23)	(8.24)	(8.24)
Credit Rating	0.0146***	0.0145***	0.00807***	0.00804***
	(5.46)	(5.45)	(5.45)	(5.44)
Constant	0.839***	0.740***	0.580***	0.537***
	(2.97)	(2.62)	(10.92)	(9.97)
Observations	7,442	7,442	7,342	7,342
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Pseudo/Adjusted R^2	0.09	0.09	0.25	0.25

Table 9: Ambiguous Tone, Dissimilarity Measures, and Bank Loan Spreads

This table provides regression results for the relationship between financial statement dissimilarity and bank loan spreads for borrowers with high versus low ambiguous tone of the financial statements. The ambiguous tone of financial statements are classified based on the proportion of uncertain or strong modal words. In Panel A, borrowers are classified as having a low (high) ambiguous tone in financial statements if the proportion of uncertain words in their financial statements is in the bottom (top) quintile across industry-years. In Panel B, borrowers are classified as having a low (high) ambiguous tone in financial statements if the proportion of strong modal words in their financial statements is in the top (bottom) quintile across industry-years. The dependent variable is the logarithm of all-in-drawn loan spread. Industry, year, loan type and primary purpose dummies are included in all the regression specifications. The test-statistics - standard errors clustered by firms – are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Ambiguous Tone - Uncertain

Variables	**Low**		**High**	
	Jaccard	**Mineditdist**	**Jaccard**	**Mineditdist**
Dissimilarity	0.0799	0.0733	0.160***	0.111**
	(1.04)	(1.29)	(2.70)	(2.31)
Controls	Yes	Yes	Yes	Yes
Observations	2,822	2,822	2,943	2,943
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.67	0.67	0.61	0.61

Panel B: Ambiguous Tone - Strong Modal

Variables	**Low**		**High**	
	Jaccard	**Mineditdist**	**Jaccard**	**Mineditdist**
Dissimilarity	0.0254	0.0130	0.264***	0.189***
	(0.41)	(0.26)	(4.20)	(3.90)
Controls	Yes	Yes	Yes	Yes
Observations	2,957	2,957	2,864	2,864
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.58	0.58	0.70	0.70

Table 10: Information Environment, Dissimilarity Measures, and Bank Loan Spreads

This table provides regression results for dissimilarity and its impact on bank loan spreads after considering the information environment of firms, i.e., analyst coverage and forecast dispersion. In Panel A, borrowers are classified as having a weak (strong) information environment if the analyst coverage is in the bottom (top) quintile across industry-years. In Panel B, borrowers are classified as having a weak (strong) information environment if analysts' forecast dispersion is in the top (bottom) quintile across industry-years. The dependent variable is the logarithm of all-in-drawn loan spread. Industry, year, loan type and primary purpose dummies are included in all the regression specifications. The test-statistics - standard errors clustered by firms – are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Information Environment - Analyst Coverage

Variables	**Weak**		**Strong**	
	Jaccard	**Mineditdist**	**Jaccard**	**Mineditdist**
Dissimilarity	0.166***	0.111**	0.0298	0.0148
	(2.73)	(2.23)	(0.45)	(0.27)
Controls	Yes	Yes	Yes	Yes
Observations	2,429	2,429	2,878	2,878
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.55	0.55	0.70	0.70

Panel B: Information Environment - Forecast Dispersion

Variables	**Weak**		**Strong**	
	Jaccard	**Mineditdist**	**Jaccard**	**Mineditdist**
Dissimilarity	0.186**	0.186***	0.134	0.109
	(2.32)	(3.18)	(1.57)	(1.54)
Controls	Yes	Yes	Yes	Yes
Observations	1,782	1,782	1,660	1,660
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.62	0.62	0.74	0.74

Table 11: Relationship Lending, Dissimilarity Measures, and Bank Loan Spreads

This table provides regression results for relationship lending, financial statement dissimilarity, and its impact on bank loan spreads. Following Bushman et al. (2017), we categorize each loan into relationship lending if the lead arranger's allocation share is more than 50% of the loan amount in the previous five years and non-relationship lending otherwise. The dependent variable is the logarithm of all-in-drawn loan spread. Industry, year, loan type and primary purpose dummies are included in all the regression specifications. The test-statistics - standard errors clustered by firms – are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Variables	**Relationship Lending**		**Non-Relationship Lending**	
	Jaccard	**Mineditdist**	**Jaccard**	**Mineditdist**
Dissimilarity	0.0597	0.0629	0.116***	0.0912***
	(0.83)	(1.11)	(3.75)	(3.71)
Controls	Yes	Yes	Yes	Yes
Observations	1,798	1,798	14,826	14,826
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.69	0.69	0.63	0.63

Table 12: Corporate Governance, Dissimilarity Measures, and Bank Loan Spreads

This table provides regression results for dissimilarity and its impact on bank loan spreads after considering the role of corporate governance, i.e., institutional ownership, and board co-option. In Panel A, borrowers are classified as having a weak (strong) corporate governance if the institutional ownership is in the bottom (top) quintile across industry-years. In Panel B, borrowers are classified as having a weak (strong) corporate governance if the level of co-opted board members is in the top (bottom) quintile across industry-years. The dependent variable is the logarithm of all-in-drawn loan spread. Industry, year, loan type and primary purpose dummies are included in all the regression specifications. The test-statistics - standard errors clustered by firms – are reported in the parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Corporate Governance - Institutional Ownership

Variables	**Weak**		**Strong**	
	Jaccard	**Mineditdist**	**Jaccard**	**Mineditdist**
Dissimilarity	0.208***	0.149***	0.0440	0.00421
	(3.43)	(2.95)	(0.77)	(0.09)
Controls	Yes	Yes	Yes	Yes
Observations	2,805	2,805	2,951	2,951
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.58	0.58	0.64	0.64

Panel B: Corporate Governance - Board Co-option

Variables	**Weak**		**Strong**	
	Jaccard	**Mineditdist**	**Jaccard**	**Mineditdist**
Dissimilarity	0.268**	0.215***	-0.00738	0.0213
	(2.55)	(2.78)	(-0.07)	(0.26)
Controls	Yes	Yes	Yes	Yes
Observations	1,164	1,164	1,060	1,060
Industry FEs	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Loan Type FEs	Yes	Yes	Yes	Yes
Primary Purpose FEs	Yes	Yes	Yes	Yes
Adjusted R^2	0.67	0.67	0.70	0.70