Corporate isomorphism and board interlock networks: Evidence from 10k Filings

Anindya S. Chakrabarti* Sudarshan Kumar† Kaustav Sen‡

September 24, 2025

Abstract

What do board networks do? The literature provides a wide array of answers ranging from corporate control to providing social embedding to companies to strategic cooptation. However, in his influential work, Mizruchi noted that board networks functionalities are more likely to be a communication mechanism than a control mechanism. Much of the literature that ensued builds on this theoretical insight. However, empirical demonstration was missing which some more recent work started to address. This paper provides an empirical analysis showing two things. One, corporate filings are becoming isomorphic to each other over time. Two, board interlocks positively influence this relationship and in effect, acts as a conduit of mimetic isomorphism for corporate strategies. We exploit a bi-layered mapping of the interlock network along with the network of textual similarities in 10k filings for US companies for a period of 16 years – 2005 to 2021. Post-2009 (GFC) the effect of the board interlocks on textual isomorphism has become stronger.

Keywords: Board interlocks, Networks, Natural Language Processing

*(corresponding author) anindyac@iima.ac.in. Economics area, Indian Institute of Management, Vastrapur, Ahmedabad, Gujarat-380015, India.

†Email: sudarshank@iima.ac.in. Finance & Control Area, Indian Institute of Management, Calcutta, Kolkata, West Bengal-700104

‡Email: ksen@pace.edu. Accounting Area, Lubin School of Business, Pace University, New York NY 10038, USA

1 Introduction

Corporate governance literature (Shleifer and Vishny, 1997; Daily et al., 2003) has long emphasized the critical role that boards of directors play in shaping the strategic and ethical direction of firms. Beyond their fundamental fiduciary responsibilities and managerial oversight, directors exert significant influence on organizational behavior and a firm's risk appetite (Fama, 1983; Zahra et al., 2000). Among the various mechanisms through which boards exert their influence, board interlocks, where directors serve on the boards of multiple firms, have emerged as a critical conduit that influences organizational behavior and decision-making (Mizruchi, 1996). These interlocks create intricate networks that facilitate the exchange of vital information, the diffusion of prevailing practices, and the transmission of established norms between otherwise independent firms, potentially leading to a notable convergence in firm behavior (Galaskiewicz and Burt, 1991; Gulati and Gargiulo, 1999).

Much of these analysis build on theory and coarse-grained aggregate data. This paper builds on this literature and explores the impact of board interlocks on the disclosure styles of corporate firms, specifically focusing on how shared directors influence the textual characteristics, readability, similarity, and sentiment of Securities and Exchange Commission (SEC) filings. The new feature we bring in is the matching between a pair of firms in terms of board interlocks and quantified similarity in filings inferred based on newly developed natural language processing tools.

The concept of board interlock networks, arising from the shared membership of a single individual across the boards of multiple firms, is a well-documented theme within the fields of sociology and organizational studies (Davis and Whitman, 2000; Granovetter, 1985). The genesis of this significant strand of research can be traced back to the influential works of DiMaggio and Powell (1983), which popularized the seminal idea of institutional isomorphism. Institutional isomorphism refers to the process by which organizations operating within the same field tend to become increasingly similar to each other over time, often irrespective of whether these changes demonstrably improve their operational efficiency or overall performance.

Here are the three primary types of institutional isomorphism. Coercive isomorphism is primarily driven by both formal and informal pressures exerted on organizations by government, regulations, industry-wide mandates, or even prevailing cultural norms that organizations feel compelled to adopt to maintain legitimacy and avoid sanctions (North, 1990; Scott, 2008). Mimetic isomorphism arises particularly when organizations face significant uncertainty or ambiguity in their environment. In such situations, they tend to model themselves on other organizations that they perceive as successful or legitimate. This imitation can involve replicating organizational structures, adopting similar strategies, or mirroring the practices of industry leaders as a seemingly rational response to uncertainty (Haunschild, 1993; Stinchcombe, 1965). The adoption of best practices, even without concrete evidence of their superiority, often falls under this category. Normative isomorphism stems primarily from the professionalization of occupations and the increasing influence of professional networks. Through formal education, professional associations, and the dissemination of shared values and norms, professionals within a field develop a relatively homogeneous set of practices and standards that they then promote within their respective organizations (Barley, 1983; Larson, 1977). The diffusion of these norms often leads to greater similarity in organizational practices across the field.

Subsequently, the seminal work of Mizruchi (1992, 1996) and numerous others (Koenig and Gogel, 1987; Burris, 2005) further explored the pivotal role of board interlocks as a conduit of the propagation of isomorphic tendencies. This foundational work has since spurred a substantial and rich body of literature dedicated to exploring the underlying motivations behind the formation of board interlocks and rigorously examining how these interlocks theoretically and empirically influence a wide range of firm behaviors (Carter et al., 2003; Westphal and Zajac, 1999).

Major theoretical lenses through which corporate network structures are explored include social network theory, which emphasizes the importance of relationships and network positions in shaping organizational outcomes (Wasserman and Faust, 1994), resource dependence theory, highlighting how interlocks can facilitate access to critical resources and reduce uncertainty (Pfeffer and Salancik, 2015; Johnson et al., 1996), and agency theory, which examines how interlocks might affect monitoring and control mechanisms within firms (Eisenhardt, 1989; Fama and Jensen, 1983). Institutional and network-based diffusion models have also been employed to study how behaviors spread across interconnected firms (Chamley et al., 2013; López-Pintado, 2008).

But in practice, how is the role of a director seen? EX SEC Chair Mary Jo White stated –

> *"Directors Are Essential Gatekeepers.... By law, it is ultimately the fiduciary responsibility of the board of directors to oversee the business and affairs of a company.... It is up to directors, along with senior management under the purview of the board, to set the all-important "tone at the top" for the entire company."*

This view is also seen empirically. The effect of board interlocks has been investigated across various organizational outcomes, including strategic decision-making

(Finkelstein, 1991), financial performance (Mizruchi and Koenig, 1989), innovation adoption (Haunschild, 1998), and corporate social responsibility (Hernández-Lara and Gonzales-Bustos, 2019). Furthermore, research has also explored the potential negative consequences of dense interlock networks, such as the propagation of negative shocks and the potential for collusive behavior (Mizruchi and Stearns, 2004).

We indeed want to capture this "tone at the top" through a numerical rather than descriptive analysis. We build on the extensive literature on institutional isomorphism and empirically examines its manifestation in the context of corporate disclosure. Corporate disclosure plays a fundamental role in promoting transparency and accountability within the financial ecosystem. Regulatory bodies such as the U.S. Securities and Exchange Commission (SEC) require public companies to disclose pertinent financial and qualitative information, ensuring that stakeholders have access to relevant data for decision-making. From an investor's perspective, disclosure is a crucial mechanism for reducing information asymmetry and mitigating agency conflicts (Healy and Palepu, 2001). Transparent and high-quality disclosure enhances a firm's reputation and can significantly influence its access to capital and its cost of financing (Core et al., 2015). Given the high stakes involved, firms often engage in strategic impression management through carefully crafted narratives in their reports (Leung et al., 2015).

Annual reports, typically exceeding 200 pages, include a combination of quantitative and qualitative information. While financial tables and statements account for approximately 30 pages, the majority of the report comprises unstructured textual content. These textual disclosures are critical as they provide interpretive context to the numerical data, offering insights into corporate strategy, operational risks, and future prospects. In this study, we utilize 25 natural language processing (NLP)-based indicators extracted from annual SEC filings, sourced from the WRDS platform. These indicators are categorized into four key dimensions: Text Characteristics, Similarity, Readability, and Sentiment.

To examine the influence of board interlocks on disclosure practices, we construct director interlock networks using BOARDex firm-board data. Director interlock networks two-mode (bipartite) structures comprising directors and firms as distinct node types (Valeeva et al., 2020). We transform this two-mode affiliation network into a one-mode firm-to-firm network by projecting edges between firms that share common directors. The number of shared directors between any two firms forms our primary independent variable of interest, referred to as network interlocks. We employ firm fixed-effects panel regression models to analyze the relationship between network interlocks and the suite of NLP-based disclosure measures. Our empirical results reveal a consistent pattern of convergence in disclosure styles among firms connected through shared directors, across all four categories of textual indicators.

Interestingly, we also observe a general trend toward convergence in disclosure style among all firms, regardless of interlock connections—suggesting the presence of broader field-level forces or other diffusion mechanisms, such as industry norms or regulatory expectations. For instance, convergence is more pronounced among firms operating within the same industry. Nevertheless, the presence of shared directors significantly accelerates the pace of convergence, even after accounting for other observable firm characteristics. Moreover, we find only limited evidence that independent directors contribute meaningfully to convergence in disclosure style. This suggests that isomorphic pressures through agency-based monitoring may be less influential than previously assumed.

Social network theory argues that not all ties exert equal influence—rather, the positional power of directors within the network can modulate their impact on firm behavior (Mizruchi, 1996; Haunschild, 1998). To explore this, we incorporate measures of director power, as degree centrality of a director within the network. Then We employ a weighted interlock variable, which combines the number of shared directors with their respective degree centralities. Our main findings remain robust under these alternative specifications, reinforcing the conclusion that network interlocks, particularly through influential directors, play a pivotal role in driving convergence in corporate disclosure styles.

The central tenet of this paper is that the established board linkages between firms, arising from shared directorships, lead to a discernible similarity in the disclosure styles adopted by these interconnected corporate entities (Chen et al., 2022). We posit that the information exchange and normative pressures transmitted through these interlocking directorates will manifest in the way firms communicate with their stakeholders through mandatory filings such as 10-K reports.

2 Data

We utilize data on individual networks of board members obtained from BoardEx. BoardEx offers comprehensive details on the board members of publicly listed and large private entities and is widely used in academic research. Individual network data from BoardEx includes the names of interconnected individuals, the organizations that connect them, and the duration of their association with each organization. Although data is accessible from 1999, it is notably sparse in the early years; consequently, we confine our study to the period from 2005 to 2021.

Initially, we expand the network data at the Firm.x-Firm.y-year-Director level. Our sample set consists of 22993 unique directors connecting 6924 unique firms over 17 years. Subsequently, we consolidate this data at the firm-firm-year level by totaling the number of shared directors, which will be referred to as Network.Interlock henceforth. Along with the Network.Interlock measure, we also have the following additional variables capturing the further granular and dynamicity of the board network

- **Indep** calculates the count of shared Independent directors serving on both firms' boards.
- **New** indicates the count of shared directors who joined the firm during that particular year.
- **Close** represents the count of shared directors who left any of these firms in that particular year.

Our central hypothesis is that shared board members influence a company's disclosure style in SEC filings. To examine this, we use 25 Natural Language Processing (NLP) indicators provided by the WRDS SEC Analytics suite. The NLP indicators in this study fall into four broad categories: Text Characteristics, Similarity, Readability, and Sentiment.

Text Characteristics measure document properties such as file size, word count, sentence structure, and paragraph density, offering insights into textual complexity.

Similarity measures (e.g., Cosine Similarity, Jaccard Index, LookBack) quantify how closely related two documents are based on word overlap and structural alignment.

Readability indices (e.g., Gunning Fog, Flesch-Kincaid, SMOG) assess how easily a document can be understood, estimating the education level required to comprehend the text. These indices are based on factors such as sentence length, syllable count, and complex word frequency.

Sentiment indicators, drawn from financial lexicons such as Loughran-McDonald (LM) and FinBERT, categorize text into positive, negative, litigious, and uncertain sentiments.

A brief description of the variables used in this study is provided in Table 1.

3 Methodology

3.1 Network construction

The primary objective of this study is to explore how director interlock networks influence firms' disclosure styles. Director interlock networks are two-mode networks with directors and firms as distinct node types, where connections can only exist between firms and directors (Valeeva et al., 2020). First, we project this two-node affiliation network into a one-node network with firms as nodes and common directors across the firms as connections. Simple representation of the projection of the two mode network into a one mode network is presented in figure 1

$Network.Interlock_{i,j,t}$ measured as a number of the common directors between the firm i and j at time t defines the weight of the connection and our main explanatory variable.

A visual representation of the projected director interlock network for a sample firm, IBM, derived from the WRDS. platform, is shown in figure 2.

3.2 Natural language processing based indicators of the SEC disclosure

We rely on 25 NLP-based indicators extracted from annual SEC filings provided by WRDS. These indicators are grouped into four categories: Text Characteristics, Similarity, Readability, and Sentiment. Figure 3 presents a correlation heatmap of these indicators. Many of the measures, especially those in the Readability and Text Characteristics categories, exhibit high correlation, with some exceeding 0.9. For brevity, we report results based on only eight NLP variables, two from each category. For robustness, we also include the first principal component (PC) of all the variables in each category in our baseline regression.

Text Characteristics: We use the natural log of file size and word count as proxies for text characteristics. File size is a well-documented proxy for effective communication of valuation-relevant information to investors. Loughran and McDonald (2014) argue that file size is a more reliable indicator of readability than some other popular NLP measures, such as the Gunning Fog index. However, file size can be influenced by factors like resolution and images, which is why we also include word count, which is unaffected by these factors. As mentioned earlier, we also incorporate the first principal component of all text characteristic measures, denoted as $PC1_{Text}$.

Readability: Readability is an important metric for assessing the intent of financial managers in disseminating information. Firms may have an incentive to use complex jargon to obscure information. This tendency has been flagged by regulators, such as the SEC, and investors, including Warren Buffet. Former SEC Chairman Christopher Coones advocated for using the Gunning Fog index to evaluate firms' disclosure compliance. Li (2008) employed Gunning Fog to investigate the relationship between readability and firm financial performance. We also include Gunning Fog, along with RIX and the first principal component of readability ($PC_{Readability}$), as measures in our analysis.

Similarity: Similarity measures from WRDS assess the similarity between the current year's filing and the previous year's report. One way that firms may obfuscate information is by reusing the same template year after year. We use two similarity measures—Jaccard M and Minimum Edit Distance—along with the first principal component of similarity ($PC1_{Similarity}$) as variables of interest.

Sentiment: Traditional sentiment measures, often derived from general-purpose linguistic dictionaries, fail to capture domain-specific nuances. Words like "liability" or "risk," which carry negative connotations in everyday language, may have neutral or even positive implications in financial contexts. To address this limitation, (Loughran and McDonald, 2011) developed a finance-specific sentiment dictionary categorizing words into positive, negative, uncertain, litigious, and modal tones tailored to financial disclosures. We utilize two indicators, LM Positive and LM Uncertainty, along with the first principal component ($PC1_{Sentiment}$) of all the sentiment indicators.

3.3 Extraction of the first principal component

The extraction of principal components is achieved through Eigenvalue Decomposition (EVD) on the covariance matrix Σ_X of the standardized NLP indicators matrix X:

$$\Sigma_X = V \Lambda V^T$$

where V is the matrix of eigenvectors and Λ is the diagonal matrix of corresponding eigenvalues. Z_1 is given by:

$$Z_1 = X v_1$$

where X is the standardized matrix of NLP indicators, and v_1 is the eigenvector corresponding to the largest eigenvalue of the covariance matrix Σ_X. This principal component captures the direction of maximum variance in the data and serves as the most informative linear combination of the original NLP indicators. For the text characteristics and similarity categories, first principal components explain more than 99% variation, while for the readability the variation explained by the first PC is 96%. However first PC of the sentiment measure only provide 69% explanation.

3.4 Calculation of the textual distance of the disclosure

Our primary variable of interest is the textual distance between the disclosure styles of two firms, measured as the absolute difference between firm-level NLP variables:

$$NLP_{i,j,t} = |NLP_{i,t} - NLP_{j,t}|$$

where $NLP_{i,t}$ and $NLP_{j,t}$ represent the NLP-based indicators for firms i and j at time t, respectively. This measure captures the divergence in corporate disclosure style between firms based on the specific NLP indicator.

Our primary hypothesis is that director interlock leads to the convergence of the disclosure styles between firms. We demonstrate this with an example presented in Figure 4, which shows the director network for the firms Adobe and Comcast. Each node represents a company, with lines indicating common directors. The Gunning Fog measure for each company is shown next to their labels. As demonstrated, Amy Banse establishes a new connection between Adobe and Comcast, and following this connection, the Gunning Fog measure quickly converges. However, this observation is correlational and does not, by itself, confirm a systematic convergence in NLP measures due to corporate connections. The remainder of the paper investigates this hypothesis further and demonstrates that the effect is indeed systematic.

4 Regression Framework and Results

We utilize a fixed-effect panel regression model to examine the impact of shared board members on textual distance, as indicated by the NLP metrics. The baseline regression is as follows:

$$\begin{aligned} NLP_{i,j,t} =& \alpha + \beta Network.Interlock_{i,j,t-1} + \gamma_1 NLP_{i,j,t-1} + \quad \gamma_2 Group.Dummy + \gamma_3 Indep_{t-1} \\ &+ f_i + g_j + \mu_t + \epsilon_{i,j,t} \end{aligned} \tag{1}$$

In this model, $NLP_{i,j,t}$ represents the absolute difference between the NLP indicators (such as the Gunning Fog index) of firms i and j at time t. $Network.Interlock_{i,j,t-1}$ is the count of shared directors between firms i and j at time $t-1$. We use the lagged value of this interlock measure to establish a causal link between it and textual distance. Additionally, we include the lagged value of the NLP distance measure to account for the persistence in these NLP characteristics. Since it is expected that firms within the same industry will exhibit similar disclosure styles, we control for this by adding an industry dummy variable (Group.Dummy), which takes the value of 1 if both firms belong to the same industry group according to the GICS (Global Industry Classification Standard), and 0 otherwise. Fogel et al. (2021) suggests that influential independent directors positively affect shareholder wealth by detecting and countering CEO missteps, which may also influence disclosure styles. Thus, we introduce the variable $Indep_{t-1}$, which represents the number of common independent directors, to capture this effect. The terms f_i, g_j, and μ_t represent the fixed effects for firm i, firm j, and time t, respectively.

4.1 Interlocks and Textual Distance

In this section, we demonstrate that director interlocks consistently reduce textual distance. We estimate equation 1 by regressing the absolute difference in NLP characteristics between two firms on their directorial interlocks. The results, shown in Table 5, indicate that director interlocks reduce the textual distance across all four NLP categories.

Two possible explanations exist for this convergence. First, shared directors may introduce best practices in disclosure styles from more transparent firms to less transparent ones, leading to similarities in text characteristics and readability. However, this convergence does not necessarily imply the spillover of good practices. Directors or companies might have perverse incentives, such as copying a style that involves less transparency or using technical jargon to obscure information (convergence in readability or text characteristics), or adopting template-based reporting (convergence in similarity). Another explanation for sentiment spillover could lie in the behavioral limitations of the shared director, who may struggle to assess each firm's financial situation in isolation.

All control variables in the regression show consistent signs. All NLP indicators, except for file size, exhibit high persistence, with positive and significant coefficients for $NLP_{i,j,t-1}$. The *Group.Dummy* variable is negative and significant, indicating that firms in the same industry group tend to have similar disclosure styles.

4.2 Influential Interlocking Directors and the Choice of Disclosure Style

In the previous section, we defined $Network.Interlock$ as the sum of common directors between two firms, assuming that each director has equal influence. However, existing literature highlights that powerful directors often have more influence in firm decisions. A common approach to measure director influence is through centrality measures in the director network. We follow this approach and use degree centrality to proxy for director influence. The new variable of interest is $Degree_{i,j,t-1}$, which is the weighted sum of common directors, with weights based on each director's degree centrality.

$$\begin{aligned} NLP_{i,j,t} =& \alpha + \beta Degree_{i,j,t-1} + \gamma_1 NLP_{i,j,t-1} + \\ & \gamma_2 Group.Dummy + \gamma_3 Indep_{t-1} + f_i + g_j + \mu_t + \epsilon_{i,j,t} \end{aligned} \tag{2}$$

Table 6 presents the results of this regression across all eight NLP indicators. The results are robust to alternative specifications of network weights, with the coefficient for $Degree_{i,j,t-1}$ being negative and significant across all NLP distance measures.

4.3 Effect of Directors Leaving or Joining a Firm

Next, we examine the granularity of this convergence, particularly focusing on the impact when a director joins a firm for the first time or when a common director leaves. The variable $New_{i,j,t-1}$ measures the number of directors who are working in both firms for the first time at time $t-1$, while $Close_{i,j,t-1}$ measures the number of common directors who left the firm in the previous year.

We do not observe consistent effects for these variables across all textual distance measures. A negative and significant coefficient for $New_{i,j,t-1}$ is observed in regressions with Gunning Fog (readability), JACCARD M, Minimum Edit Distance (similarity), and LM Uncertainty (sentiment), suggesting an additional convergence effect in the first year after network formation. Similarly, a negative and significant coefficient for $Close_{i,j,t-1}$ is found in regressions with RIX (readability), JACCARD M, Minimum Edit Distance (similarity), LM Positive, and LM Uncertainty (sentiment), indicating a slight decrease in convergence as directors leave a firm. However, the effects are not significant for file characteristics, suggesting that effect of director leaving and joining is not that strong in the immediate year.

4.4 Convergence in the disclosure style and firm financial performance

An alternative explanation for convergence in disclosure styles could be the similarity in financial characteristics between firms. Larger firms, with more resources, are expected to provide more detailed disclosures than smaller firms. On the other hand, highly leveraged firms may have an incentive to window-dress their disclosures. To test this alternative explanation, we modify equation 1 by including three firm-level financial variables as explanatory factors:

$$\begin{aligned} NLP_{i,j,t} =& \alpha + \beta Network.Interlock_{i,j,t-1} + \gamma_1 NLP_{i,j,t-1} + \gamma_2 Group.Dummy + \gamma_3 Indep_{t-1} \\ & + |lnasset_{i,j,t-1}| + |B/M_{i,j,t-1}| + |CapitalRatio_{i,j,t-1}| f_i + g_j + \mu_t + \epsilon_{i,j,t} \end{aligned} \tag{3}$$

Since our dependent variable is the textual distance, we also use the distance between firm-level financial variables. $|lnasset_{i,j,t-1}|$ is the absolute difference between the natural log of the assets of firm i and j at time $t-1$. Similarly, we have $|B/M_{i,j,t-1}|$ and $|CapitalRatio_{i,j,t-1}|$ as an absolute difference of Book by Market and capital ratio, respectively.

The significance of our primary variable of interest, $Network.Interlock_{i,j,t-1}$, holds for most NLP indicators, except for firm text characteristics. This suggests that the network effect is independent of firms' financial characteristics. However, the distance in size and capital structure between firms is positively related to disclosure distance, indicating that disclosure convergence may be partly explained by similarities in firm size and capital structure. Interestingly, $|B/M_{i,j,t-1}|$ is negatively related to textual distance, which is somewhat counterintuitive.

4.5 Heterogeneity around the Crisis-periods

Finally, coercive isomorphism is primarily driven by both formal and informal pressures exerted on organizations by governments and regulatory bodies. The 2008 financial crisis, among other factors, led to heightened regulatory scrutiny and increased demands for transparency from firms, exemplified by reforms such as the Dodd-Frank Act. To examine the impact of the crisis on disclosure convergence, we include a GFC dummy variable for years from 2009 onwards and interact it with the network interlock measures. Table 9 presents the results. The negative and significant coefficient of the interaction term across most measures (excluding word count) strongly indicates that the role of common directors in driving convergence in disclosure became more pronounced after the crisis.

5 Summary and Managerial Implications

In this paper, we analyze a classic question in the organizational literature. What do board interlocks do? – Mizruchi asked this question around three decades back (Mizruchi, 1996) and provided a series of tentative answers ranging from corporate control to providing social embedding to companies to strategic cooptation. The ensuing literature built on these theoretical insights. Cai et al. (2014) took an important step in quantifying the effects of board interlocks on corporate disclosure policies. They exploit the sudden changes in directorial presence in provision of quarterly earnings guidance in the annual reports. This clever design shows that disclosure policy contagion takes place through directors who have earlier been part of a decision to stop quarterly earnings guidance.

We begin where Cai et al. (2014) stop. What about more general mimetic contagion through the directors' network? Do the firms simply copy some very specific strategies like quarterly earnings guidance, or is it a much broader phenomenon? We exploit natural language processing to extract sentiments across firms from their yearly 10k filings. This allows us to capture a wide range of strategic behavior in terms of transparency, future outlook, risk-taking attitude to persistence.

We show that directorial linkages indeed induce strong effects of similarity in future corporate filings. This holds true across dimensions and across time. In particular, post the global financial crisis, the effect has become stronger.

What is the learning for managers here? One, director's network is indeed a conduit of policy spillover and hence, this is an externality. The recognition of the fact is important for the managers. Two, the effect on managerial control is unclear. Most likely, the spillover is in terms of practices and information rather than direct control. Thus the immediate effect on the firm performance may not be pronounced. Third, major financial shocks like the global financial crisis can create a shift in the effect of the board interlocks.

Finally, our study certainly has limitations. The most prominent feature of this analysis is that in order to generalize the range of possible sentiments and textual expositions in the corporate filings, we increased the sample size to the maximum as opposed to the narrow but more causal nature of analysis by Cai et al. (2014). However, we think that our approach in fact complements the analysis of Cai et al. (2014) rather than try to substitute it and hence, expands the scope of the possible effects. The second feature is that as different trade barriers come up (for example, the Trump tariff war), corporate management structures are rapidly shifting with major changes in the sourcing and selling decisions - especially for the multinationals. While this is beyond the scope of this paper, the behavior of multinational companies can be quite different from domestic companies in presence of trade uncertainties. Finally, while we explore heterogeneity in sentiments, future work may shed more light on the nature of the relationship between specific business strategies and the corresponding sentiments, and the influence of board interlocks on them.

References

Stephen R Barley. The alignment of technology and structure through roles and networks. *Administrative science quarterly*, 28(1):61–103, 1983.

Val Burris. Interlocking directorates and firm performance in corporate america: The impact of financial, social, and human capital. *American sociological review*, 70(1): 103–121, 2005.

Ye Cai, Dan S Dhaliwal, Yongtae Kim, and Carrie Pan. Board interlocks and the diffusion of disclosure policy. *Review of Accounting Studies*, 19:1086–1119, 2014.

David A Carter, Betty J Simkins, and W Glenn Simpson. Corporate governance and board diversity. *Journal of corporate finance*, 9(3):261–283, 2003.

Christophe Chamley, Anna Scaglione, and Lin Li. Models for the diffusion of beliefs in social networks: An overview. *IEEE Signal Processing Magazine*, 30(3):16–29, 2013.

Hongzhi Chen, Ajay Mehra, Stefano Tasselli, and Stephen P Borgatti. Network dynamics and organizations: A review and research agenda. *Journal of Management*, 48(6):1602–1660, 2022.

John E Core, Luzi Hail, and Rodrigo S Verdi. Mandatory disclosure quality, inside ownership, and cost of capital. *European Accounting Review*, 24(1):1–29, 2015.

Catherine M Daily, Dan R Dalton, and Albert A Cannella Jr. *Corporate governance: Theories, practices, and research*. Wiley-Blackwell, 2003.

Gerald F Davis and Charlan Jeanne Whitman. The network organization: A structural approach to management. *Research in organizational behavior*, 22:79–106, 2000.

Paul J DiMaggio and Walter W Powell. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American sociological review*, pages 147–160, 1983.

Kathleen M Eisenhardt. Agency theory: An assessment and review. *Academy of management review*, 14(1):57–74, 1989.

Eugene F Fama. Agency problems and the theory of the firm. *Journal of political economy*, 91(2):288–307, 1983.

Eugene F Fama and Michael C Jensen. Separation of ownership and control. *The journal of law and Economics*, 26(2):301–325, 1983.

Sydney Finkelstein. Intercorporate ties and managerial pay. *Academy of Management Journal*, 34(1):194–216, 1991.

Kathy Fogel, Liping Ma, and Randall Morck. Powerful independent directors. *Financial Management*, 50(4):935–983, 2021.

Joseph Galaskiewicz and Ronald S Burt. Interorganization contagion in corporate philanthropy. *Administrative science quarterly*, pages 88–105, 1991.

Mark Granovetter. Economic action and social structure: the problem of embeddedness. *American journal of sociology*, 91(3):481–510, 1985.

Ranjay Gulati and Martin Gargiulo. Where do interorganizational networks come from? *American journal of sociology*, 104(5):1439–1493, 1999.

Pamela R Haunschild. Interorganizational imitation: The impact of interlocks on acquisition activity in the us savings and loan industry. *Administrative science quarterly*, 38(4):564–592, 1993.

Pamela R Haunschild. Board network size, board composition, and corporate diversification. *Administrative science quarterly*, 43(3):680–706, 1998.

Paul M Healy and Krishna G Palepu. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of accounting and economics*, 31(1-3):405–440, 2001.

Ana B Hernández-Lara and Juan P Gonzales-Bustos. The impact of interlocking directorates on innovation: the effects of business and social ties. *Management Decision*, 57(10):2799–2815, 2019.

Jonathan L Johnson, Catherine M Daily, and Alan E Ellstrand. Boards of directors: A review and research agenda. *Journal of management*, 22(3):409–438, 1996.

Thomas Koenig and Robert Gogel. Interlocking directorates and the organization of business community. *American journal of sociology*, 92(5):1151–1179, 1987.

Magali Sarfatti Larson. *The rise of professionalism: A sociological analysis*. University of California Press, 1977.

Sidney Leung, Lee Parker, and John Courtis. Impression management through minimal narrative disclosure in annual reports. *The British Accounting Review*, 47(3):275–289, 2015. ISSN 0890-8389. doi: https://doi.org/10.1016/j.bar.2015.04.002. URL https://www.sciencedirect.com/science/article/pii/S0890838915000207.

Feng Li. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and economics*, 45(2-3):221–247, 2008.

Dunia López-Pintado. Diffusion in complex social networks. *Games and Economic Behavior*, 62(2):573–590, 2008.

Tim Loughran and Bill McDonald. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of finance*, 66(1):35–65, 2011.

Tim Loughran and Bill McDonald. Measuring readability in financial disclosures. *The Journal of Finance*, 69(4):1643–1671, 2014. doi: https://doi.org/10.1111/jofi.12162. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12162.

Mark S Mizruchi. *The structure of corporate political action: Interfirm relations and their consequences*. Harvard university press, 1992.

Mark S Mizruchi. What do interlocks do? an analysis, critique, and assessment of research on interlocking directorates. *Annual review of sociology*, 22(1):271–298, 1996.

Mark S Mizruchi and Thomas Koenig. Cohesion, competition, and the structure of the american corporate elite: The case of large manufacturing firms, 1904-1979. *American sociological review*, pages 878–903, 1989.

Mark S Mizruchi and Linda Brewster Stearns. The capitalism without owners thesis revisited: The effects of board interlocks on corporate financing decisions. *American sociological review*, 69(2):248–279, 2004.

Douglass C North. *Institutions, institutional change and economic performance*. Cambridge university press, 1990.

Jeffrey Pfeffer and Gerald Salancik. External control of organizations—resource dependence perspective. In *Organizational behavior 2*, pages 355–370. Routledge, 2015.

W Richard Scott. *Institutions and organizations: Ideas, interests, and identities.* SAGE Publications, 2008.

Andrei Shleifer and Robert W Vishny. A theory of corporate governance. *The Journal of Finance*, 52(2):737–788, 1997.

Arthur L Stinchcombe. Social structure and organizations. *Handbook of organizations*, 142:153, 1965.

Diliara Valeeva, Eelke M. Heemskerk, and Frank W. Takes. The duality of firms and directors in board interlock networks: A relational event modeling approach. *Social Networks*, 62:68–79, 2020. ISSN 0378-8733. doi: https://doi.org/10.1016/j.socnet.2020.02.009. URL https://www.sciencedirect.com/science/article/pii/S0378873320300186.

Stanley Wasserman and Katherine Faust. *Social network analysis: Methods and applications.* Cambridge university press, 1994.

James D Westphal and Edward J Zajac. Board games: Self-serving biases in director selection. *Administrative science quarterly*, 44(1):103–133, 1999.

Shaker A Zahra, John A Pearce, and Shaker A Zahra. Corporate entrepreneurship and shareholder wealth creation: The role of board-level strategic orientation. *Academy of Management Executive*, 14(4):47–59, 2000.

Appendix

A Figures and Tables

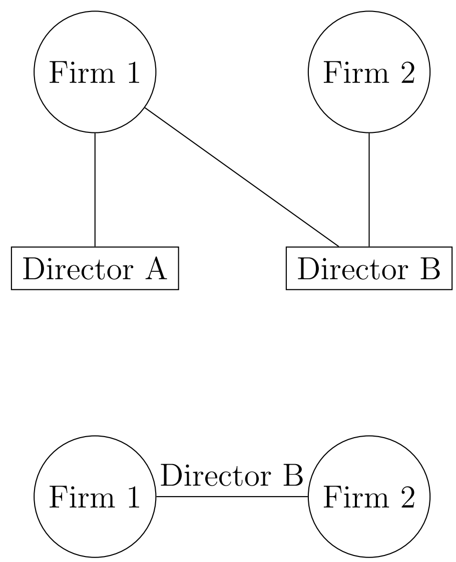

Figure 1: A Schematic of Board Interlock Networks: Two-mode Firm-Director Network (top) and its One-mode Projection (bottom)

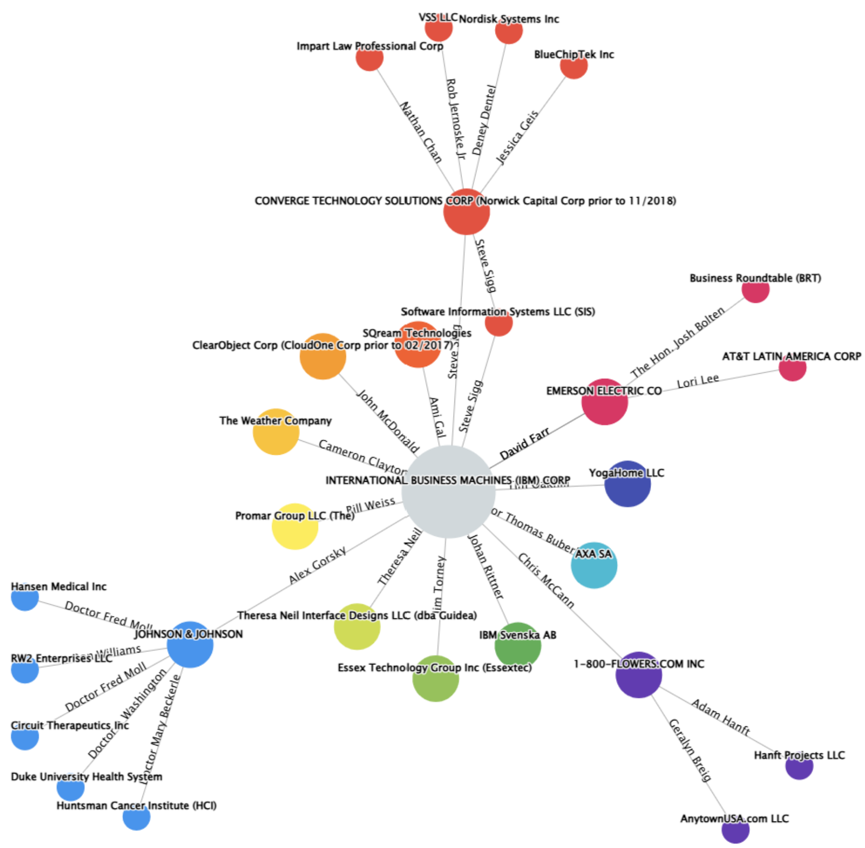

Figure 2: A network visualization of board interlocks. Source: WRDS BoardEx. This figure presents a sample from the BoardEx dataset, illustrating connections between firms and board members. IBM is shown as the central node, with edges representing shared directorships.

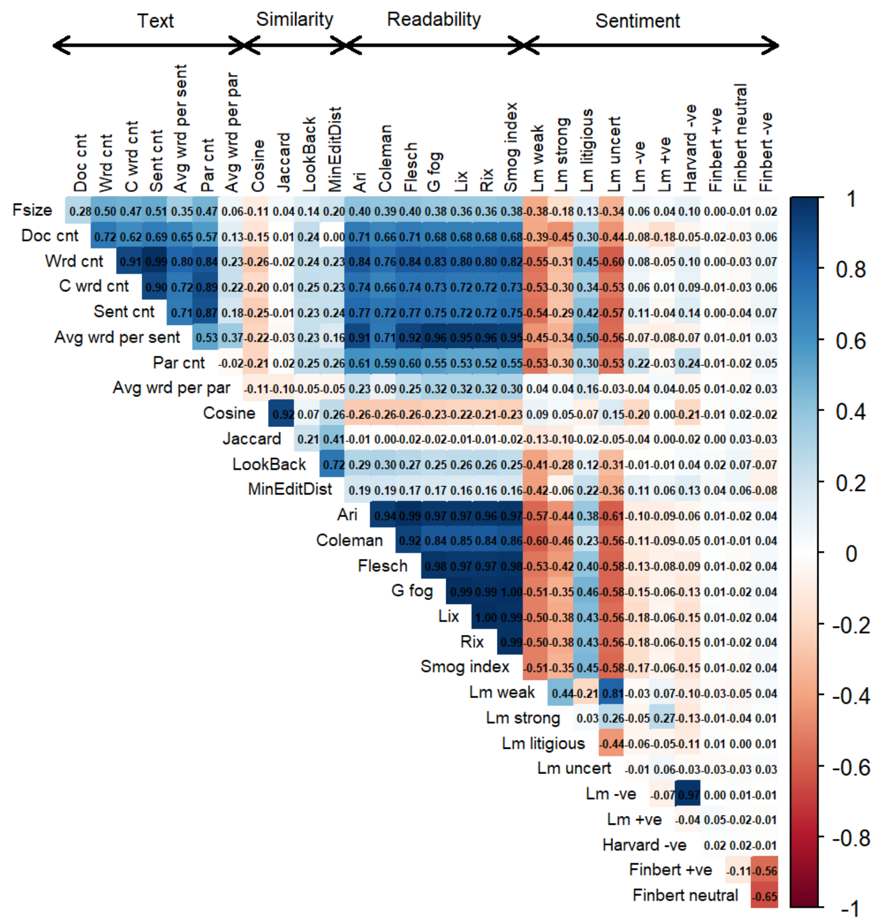

Figure 3: Correlation Heat Map of the NLP measures

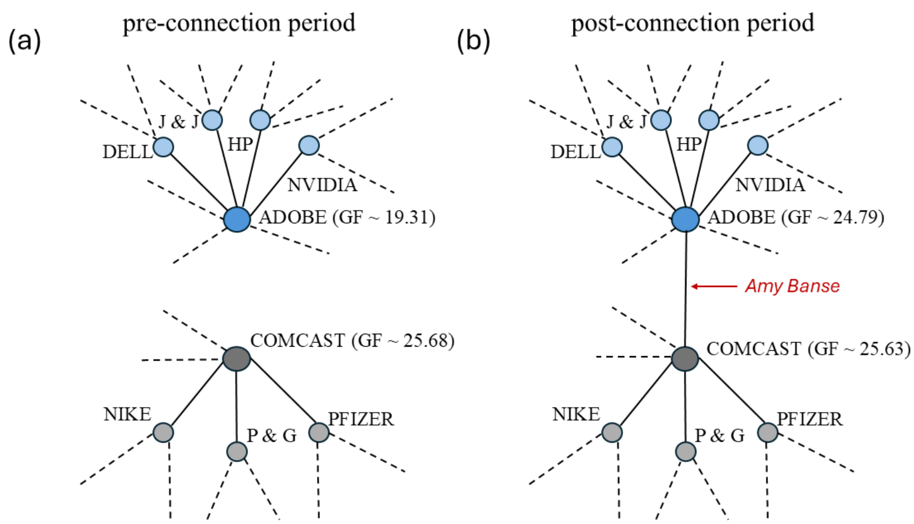

Figure 4: A schematic for the key result. Each node represents a company. Solid lines show connections between a pair of companies. Dashed lines indicate connection that are omitted for visual clarity. Panel (a) shows two companies – Adobe and Comcast, and some of their connections. We provide the Gunning-Fog measure for these two companies by the side of their labels. Panel (b) shows the same configuration except that *Amy Banse* connects now Adobe and Comcast. We see a rapid convergence in the Gunning-Fog measure post the connection materializes. This is of course mere correlation and by itself, one event not indicate systematic convergence in NLP measures due to corporate connections. The rest of the paper explores this basic hypothesis and shows that this is in fact systematic.

Table 1: Brief description of the NLP indicators

NLP Variable Name	**Description**
	Text Characteristics
File Size (fsize)	Observed file size in bytes.
Document Count	Number of documents in the dataset.
Word Count (wcnt)	Total number of words in sentences containing at least 5 words.
Sentence Count	Number of sentences that contain 5 words or more.
Average Words per Sentence	The total word count divided by the number of sentences.
Paragraph Count	Number of paragraphs with at least 10 words and 1 sentence.
Average Words per Paragraph	Number of words in all paragraphs divided by the paragraph count.
	Similarity
COSINE_eng	Cosine similarity between two text documents.
JACCARD_M_eng	Jaccard Modified Similarity, measuring word overlap.
LookBack	Measures similarity with previous documents.
MinEditDist_eng	Minimum edit distance required to transform one document into another.
	Readability
Automated Readability Index (ARI)	A readability measure based on word and sentence complexity: $ARI = 4.71 \times \frac{\text{characters}}{\text{words}} + 0.5 \times \frac{\text{words}}{\text{sentences}} - 21.43$
Coleman-Liau Index	A readability measure using letter count instead of syllables: $CLI = 0.0588 \times L - 0.296 \times S - 15.8$ where L is the average number of letters per 100 words, and S is the average number of sentences per 100 words.
Flesch-Kincaid Grade Level	Indicates readability in U.S. school grade level: $FKGL = 0.39 \times \frac{\text{words}}{\text{sentences}} + 11.8 \times \frac{\text{syllables}}{\text{words}} - 15.59$
Gunning Fog Index (Gfog Index)	Estimates years of formal education needed to understand the text: $GFI = 0.4 \times \left(\frac{\text{words}}{\text{sentences}} + 100 \times \frac{\text{complex words}}{\text{words}}\right)$ where complex words have three or more syllables.
Lix (Laesbarheds Index)	Measures readability based on sentence length and word complexity: $LIX = \frac{\text{words}}{\text{sentences}} + 100 \times \frac{\text{long words}\geq 7 \text{ letters)}}{\text{words}}$
Readability Index (RIX)	Defined as: $RIX = \frac{\text{long words}\geq 6 \text{ letters}}{\text{sentences}}$
SMOG Index	Estimates years of education needed to comprehend text: $SMOG = 1.0430 \times \sqrt{30 \times \frac{\text{polysyllabic words}}{\text{sentences}}} + 3.1291$
	Sentiment Measures Loughran and McDonald (2011)
LM Negative	$\frac{\text{Number of Loughran-McDonald financial negative words}}{\text{Total words in LM master dictionary}}$
LM Positive	$\frac{\text{Number of Loughran-McDonald financial positive words}}{\text{Total words in LM master dictionary}}$
LM Litigious	$\frac{\text{Number of Loughran-McDonald financial litigious words}}{\text{Total words in LM master dictionary}}$
LM Uncertainty	$\frac{\text{Number of Loughran-McDonald financial uncertainty words}}{\text{Total words in LM master dictionary}}$
MDA FinBERT Positive	Fraction of words classified as positive sentiment by FinBERT: $\frac{\text{Number of positive words}}{\text{Total words in document}}$
MDA FinBERT Neutral	Fraction of words classified as neutral sentiment by FinBERT: $\frac{\text{Number of neutral words}}{\text{Total words in document}}$
MDA FinBERT Negative	Fraction of words classified as negative sentiment by FinBERT: $\frac{\text{Number of negative words}}{\text{Total words in document}}$

Table 2: NLP Measures: Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Fsize	58,379	15.426	1.941	0.000	19.890
Doc cnt	58,379	3.361	1.153	0.000	6.528
Wrd cnt	58,379	11.804	0.829	6.708	14.752
C wrd cnt	58,379	49,968.080	37,673.480	174.000	599,601.000
Sent cnt	58,379	8.364	0.702	3.761	11.125
Avg wrd per sent	58,379	31.604	5.109	12.539	71.000
Par cnt	58,379	1,630.007	1,077.739	23.000	20,555.000
Avg wrd per par	58,379	106.345	53.719	26.092	3,462.627
Cosine	58,379	0.841	0.205	0.000	1.000
Jaccard	58,379	0.872	0.192	0.000	1.000
LookBack	58,379	7.757	4.479	0.000	25.500
MinEditDist	58,379	0.577	0.208	0.000	0.987
Ari	58,379	23.823	5.029	11.728	45.178
Coleman	58,379	20.004	3.749	3.342	36.115
Flesch	58,379	20.506	3.701	11.020	38.328
G fog	58,379	23.420	2.639	14.048	37.452
Lix	58,379	68.716	6.518	42.904	102.712
Rix	58,379	11.785	2.322	3.854	23.069
Smog index	58,379	19.763	1.821	12.786	26.512
Lm weak	58,379	0.004	0.002	0.0003	0.018
Lm strong	58,379	0.002	0.001	0.0003	0.023
Lm litigious	58,379	0.016	0.005	0.003	0.075
Lm uncert	58,379	0.011	0.003	0.001	0.028
Lm -ve	58,379	0.019	0.019	0.004	0.530
Lm +ve	58,379	0.008	0.002	0.002	0.022
Harvard -ve	58,379	0.044	0.019	0.012	0.542
Finbert +ve	58,379	0.120	0.197	0.000	0.961
Finbert neutral	58,379	0.113	0.218	0.000	0.975
Finbert -ve	58,379	0.751	0.291	0.000	0.957
year	58,379	2,012.830	4.989	2,005	2,021
ggroup	58,217	3,314.751	1,163.144	1,010	6,020

Table 3: Director Network Summary

Year	Mean (Degree Cent.)	SD (Degree Cent.)	Transitivity Global	Transitivity Local Average	Transitivity Random	APL	APL Random	Smallworld
2005	14.099	11.972	0.515	0.543	0.002	4.404	3.566	176.174
2006	14.343	11.987	0.524	0.542	0.002	4.431	3.574	188.909
2007	14.323	11.749	0.512	0.543	0.002	4.429	3.584	189.012
2008	13.323	10.492	0.504	0.541	0.002	4.498	3.640	184.923
2009	12.468	10.052	0.506	0.540	0.002	4.574	3.677	176.101
2010	12.427	9.793	0.511	0.539	0.002	4.581	3.672	173.808
2011	13.266	10.450	0.527	0.548	0.003	4.491	3.593	168.210
2012	13.135	10.023	0.510	0.546	0.002	4.513	3.604	163.439
2013	13.039	9.716	0.510	0.534	0.002	4.505	3.620	168.914
2014	13.899	10.751	0.511	0.531	0.002	4.441	3.568	169.020
2015	14.349	10.577	0.499	0.526	0.002	4.385	3.551	169.969
2016	13.998	10.112	0.486	0.525	0.002	4.413	3.575	167.325
2017	13.877	9.919	0.496	0.526	0.002	4.439	3.579	168.750
2018	14.971	11.531	0.526	0.529	0.002	4.355	3.508	172.760
2019	15.660	11.624	0.516	0.533	0.002	4.294	3.478	171.651
2020	16.096	11.972	0.502	0.533	0.002	4.260	3.470	173.637
2021	17.940	13.368	0.501	0.528	0.002	4.154	3.412	183.162

Table 4: Firm Network Summary

Year	Mean (Degree Cent.)	SD (Degree Cent.)	Transitivity Global	Transitivity Local Average	Transitivity Random	APL	APL Random	Smallworld
2005	5.550	5.216	0.126	0.222	0.001	5.050	4.965	83.567
2006	5.640	5.167	0.124	0.204	0.001	5.056	4.943	83.161
2007	5.644	5.124	0.120	0.196	0.001	5.044	4.940	80.524
2008	5.403	4.876	0.121	0.194	0.001	5.103	5.019	80.335
2009	5.041	4.563	0.114	0.190	0.001	5.223	5.166	75.537
2010	5.053	4.476	0.119	0.183	0.002	5.223	5.135	74.895
2011	5.152	4.544	0.112	0.174	0.002	5.130	5.053	66.469
2012	5.196	4.592	0.103	0.166	0.002	5.161	5.020	59.051
2013	5.372	4.648	0.103	0.159	0.002	5.102	4.935	57.020
2014	5.727	5.005	0.108	0.164	0.002	5.030	4.795	57.877
2015	6.054	5.203	0.109	0.157	0.002	4.926	4.673	55.792
2016	5.980	5.191	0.110	0.156	0.002	4.964	4.695	55.894
2017	5.959	5.052	0.108	0.159	0.002	4.950	4.690	54.407
2018	6.253	5.417	0.108	0.156	0.002	4.838	4.586	51.949
2019	6.582	5.658	0.109	0.160	0.002	4.732	4.476	50.193
2020	6.887	6.009	0.111	0.160	0.002	4.678	4.404	50.593
2021	7.644	6.778	0.109	0.153	0.002	4.518	4.243	48.203

Table 5: NLP Distance vs Network Interlock

	Text Characteristics			*Readability*		
	File Size	*Word Count*	$PC1_{Text}$	*GFog Index*	*RIX*	$PC1_{Readability}$
	(1)	(2)	(3)	(4)	(5)	(6)
$Network.Interlock_{i,j,t-1}$	-0.021^{***}	-0.004^{***}	-0.013^{***}	-0.015^{***}	-0.015^{***}	-0.020^{***}
	(0.006)	(0.001)	(0.003)	(0.004)	(0.004)	(0.003)
$\lvert NLP_{i,j,t-1}\rvert$	-0.146^{***}	0.252^{***}	0.227^{***}	0.216^{***}	0.289^{***}	0.232^{***}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$Indep_{t-1}$	0.010	-0.002	0.001	0.001	0.005	0.003
	(0.009)	(0.002)	(0.005)	(0.006)	(0.005)	(0.005)
Group.Dummy	-0.063^{***}	-0.014^{***}	-0.042^{***}	-0.057^{***}	-0.049^{***}	-0.067^{***}
	(0.001)	(0.0002)	(0.0005)	(0.001)	(0.001)	(0.001)
Observations	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927
R^2	0.278	0.497	0.487	0.254	0.305	0.295
Adj R^2	0.278	0.497	0.487	0.254	0.304	0.295

	Similarity			*Sentiment*		
	Jaccard M	*MinEditDist*	$PC1_{Similarity}$	*Lm Positive*	*LM Uncertainty*	$PC1_{Sentiment}$
	(1)	(2)	(3)	(4)	(5)	(6)
$Network.Interlock_{i,j,t-1}$	-0.002^{***}	-0.003^{***}	-0.021^{***}	-0.0001^{***}	-0.00005^{***}	-0.018^{***}
	(0.0001)	(0.0003)	(0.001)	(0.00000)	(0.00001)	(0.006)
$\lvert NLP_{i,j,t-1}\rvert$	0.034^{***}	0.495^{***}	0.087^{***}	0.562^{***}	0.335^{***}	0.012^{***}
	(0.00003)	(0.0001)	(0.00004)	(0.0001)	(0.0001)	(0.0001)
$Indep_{t-1}$	-0.0003	-0.001^{**}	-0.010^{***}	0.00001	-0.00002^{**}	-0.003
	(0.0002)	(0.001)	(0.002)	(0.00001)	(0.00001)	(0.009)
Group.Dummy	-0.002^{***}	-0.005^{***}	-0.025^{***}	-0.0001^{***}	-0.0002^{***}	-0.033^{***}
	(0.00002)	(0.0001)	(0.0002)	(0.00000)	(0.00000)	(0.001)
Observations	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927
R^2	0.319	0.625	0.375	0.528	0.396	0.341
Adj. R^2	0.319	0.625	0.374	0.528	0.396	0.341

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 6: NLP absolute distance vs Degree Centrality of the Common director

	Dependent variable:							
	File Size	Word Count	GFog	RIX	JACCARD M	MinEditDist	LM Positive	LM uncertainty
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Degree_{i,j,t-1}$	−0.001***	−0.0001	−0.0002*	−0.0003**	−0.00004***	−0.0001***	−0.00000***	−0.00000***
	(0.0002)	(0.00004)	(0.0001)	(0.0001)	(0.00000)	(0.00001)	(0.00000)	(0.00000)
$\lvert NLP_{i,j,t-1}\rvert$	−0.146***	0.252***	0.216***	0.289***	0.034***	0.495***	0.562***	0.335***
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.00003)	(0.0001)	(0.0001)	(0.0001)
$Indep_{t-1}$	0.008	−0.005***	−0.011**	−0.005	−0.001***	−0.002***	−0.00001***	−0.00004***
	(0.007)	(0.002)	(0.005)	(0.004)	(0.0002)	(0.0004)	(0.00000)	(0.00001)
Group.Dummy	−0.063***	−0.014***	−0.057***	−0.049***	−0.002***	−0.005***	−0.0001***	−0.0002***
	(0.001)	(0.0002)	(0.001)	(0.001)	(0.00002)	(0.0001)	(0.00000)	(0.00000)
Observations	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927
R^2	0.278	0.497	0.254	0.305	0.319	0.625	0.528	0.396
Adj. R^2	0.278	0.497	0.254	0.304	0.319	0.625	0.528	0.396

Note: *p<0.1; **p<0.05; ***p<0.01

Table 7: NLP Absolute:Director leaving or joining the firm

	Dependent variable:							
	File Size	Word Count	GFog	RIX	JACCARD M	MinEditDist	LM Positive	LM uncertainty
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Network.Interlock_{i,j,t-1}$	−0.019**	−0.002	−0.014***	−0.017***	−0.001***	−0.002***	−0.0001***	−0.00004***
	(0.008)	(0.002)	(0.005)	(0.005)	(0.0002)	(0.0004)	(0.00000)	(0.00001)
$New_{i,j,t-1}$	0.003	0.00003	−0.018**	−0.012	−0.003***	−0.008***	−0.00000	−0.0001***
	(0.013)	(0.003)	(0.009)	(0.008)	(0.0003)	(0.001)	(0.00001)	(0.00001)
$Close_{i,j,t-1}$	−0.007	−0.004*	0.008	0.012**	0.001***	0.002***	0.00002***	0.00003***
	(0.010)	(0.002)	(0.007)	(0.006)	(0.0002)	(0.001)	(0.00001)	(0.00001)
$\|NLP_{i,j,t-1}\|$	−0.142***	0.255***	0.218***	0.292***	0.035***	0.499***	0.564***	0.337***
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.00003)	(0.0001)	(0.0001)	(0.0001)
$Indep_{t-1}$	0.011	−0.002	−0.0001	0.006	−0.0004**	−0.001**	0.00001	−0.00002**
	(0.009)	(0.002)	(0.006)	(0.005)	(0.0002)	(0.001)	(0.00001)	(0.00001)
Group.Dummy	−0.061***	−0.014***	−0.057***	−0.049***	−0.002***	−0.005***	−0.0001***	−0.0002***
	(0.001)	(0.0002)	(0.001)	(0.001)	(0.00002)	(0.0001)	(0.00000)	(0.00000)
Observations	49,684,553	49,684,553	49,684,553	49,684,553	49,684,553	49,684,553	49,684,553	49,684,553
R^2	0.274	0.498	0.255	0.306	0.319	0.626	0.529	0.397
Adjusted R^2	0.274	0.498	0.254	0.305	0.319	0.626	0.529	0.397

Note: *p<0.1; **p<0.05; ***p<0.01

Table 8: NLP Absolute Distance: Firm Characteristics and NLP Absolute: Firm

	Dependent variable:				*Dependent variable:*			
	File Size	Word Count	GFog	RIX	JACCARD M	MinEditDist	LM Positive	LM Uncertainty
	(1)	(3)	(5)	(7)	(1)	(3)	(5)	(7)
$Network.Interlock_{i,j,t-1}$	−0.009	−0.002	−0.012**	−0.012***	−0.002***	−0.003***	−0.0001***	−0.00004***
	(0.009)	(0.001)	(0.005)	(0.005)	(0.0002)	(0.0004)	(0.00000)	(0.00001)
$\lvert NLP_{i,j,t-1}\rvert$	−0.209***	0.239***	0.235***	0.306***	0.054***	0.608***	0.604***	0.381***
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.00005)	(0.0001)	(0.0001)	(0.0002)
$\lvert lnasset_{x,y,t-1}\rvert$	0.146***	0.034***	0.044***	0.052***	0.001***	−0.0001***	0.00001***	0.00001***
	(0.001)	(0.0001)	(0.0004)	(0.0004)	(0.00001)	(0.00003)	(0.00000)	(0.00000)
$\lvert B/M_{x,y,t-1}\rvert$	−0.00002***	0.00000***	−0.0001***	−0.0001***	−0.00000***	−0.00000***	−0.00000***	−0.00000***
	(0.00001)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.000)	(0.000)
$\lvert CapitalRatio_{x,y,t-1}\rvert$	0.101***	0.025***	0.021***	0.019***	0.0002***	−0.003***	−0.00001***	0.00005***
	(0.002)	(0.0003)	(0.001)	(0.001)	(0.00004)	(0.0001)	(0.00000)	(0.00000)
$Indep_{t-1}$	0.017	0.002	0.010	0.013*	−0.0001	−0.0002	0.00001*	−0.00001
	(0.013)	(0.002)	(0.007)	(0.007)	(0.0003)	(0.001)	(0.00001)	(0.00001)
Group.Dummy	−0.077***	−0.014***	−0.054***	−0.050***	−0.003***	−0.004***	−0.0001***	−0.0002***
	(0.001)	(0.0002)	(0.001)	(0.001)	(0.00003)	(0.0001)	(0.00000)	(0.00000)
Observations	28,848,473	28,848,473	28,848,473	28,848,473	28,848,473	28,848,473	28,848,473	28,848,473
R^2	0.299	0.594	0.329	0.364	0.299	0.594	0.329	0.364
Adjusted R^2	0.299	0.594	0.329	0.363	0.299	0.594	0.329	0.363

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 9: NLP Absolute Distance: Post Crisis

	Dependent variable:							
					formula_1			
	File Size	Word Count	GFog	RIX	JACCARD M	MinEditDist	LM Positive	LM Uncertainty
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Network.Interlock_{i,j,t-1}$	0.001	-0.005^{**}	0.008	0.015^{**}	0.0001	0.002^{***}	-0.00004^{***}	-0.00000
	(0.011)	(0.002)	(0.008)	(0.007)	(0.0002)	(0.001)	(0.00001)	(0.00001)
$Network.Interlock_{i,j,t-1}$:GFC	-0.027^{**}	0.002	-0.028^{***}	-0.036^{***}	-0.002^{***}	-0.006^{***}	-0.00004^{***}	-0.0001^{***}
	(0.011)	(0.002)	(0.008)	(0.007)	(0.0002)	(0.001)	(0.00001)	(0.00001)
$\|NLP_{i,j,t-1}\|$	-0.146^{***}	0.252^{***}	0.216^{***}	0.289^{***}	0.034^{***}	0.495^{***}	0.562^{***}	0.335^{***}
	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.00003)	(0.0001)	(0.0001)	(0.0001)
$Indep_{t-1}$	0.010	-0.002	0.0002	0.004	-0.0004^{*}	-0.001^{**}	0.00000	-0.00002^{**}
	(0.009)	(0.002)	(0.006)	(0.005)	(0.0002)	(0.001)	(0.00001)	(0.00001)
Group.Dummy	-0.063^{***}	-0.014^{***}	-0.057^{***}	-0.049^{***}	-0.002^{***}	-0.005^{***}	-0.0001^{***}	-0.0002^{***}
	(0.001)	(0.0002)	(0.001)	(0.001)	(0.00002)	(0.0001)	(0.00000)	(0.00000)
Observations	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927	50,263,927
R^2	0.278	0.497	0.254	0.305	0.319	0.625	0.528	0.396
Adj. R^2	0.278	0.497	0.254	0.304	0.319	0.625	0.528	0.396
Note:								*p<0.1; **p<0.05; ***p<0.01

Table 10: NLP Change:1:

	Dependent variable:							
	File Size		Word Count		Word Count	Gfog Index	RIX	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lag1	−0.026***	−0.064***	−0.278***	−0.295***	−0.373***	−0.290***	−0.300***	−0.246***
	(0.007)	(0.011)	(0.010)	(0.010)	(0.014)	(0.018)	(0.011)	(0.012)
xlessy	−0.394***	−0.305***	−0.050***	−0.029***	−0.221***	−0.213***	−0.143***	−0.154***
	(0.030)	(0.028)	(0.005)	(0.004)	(0.018)	(0.023)	(0.014)	(0.017)
bm.x.y.l		0.0001		0.00005**		0.0002		0.0002
		(0.0001)		(0.00002)		(0.0001)		(0.0001)
lnassets.x.y.l		0.094***		0.040***		0.044***		0.038***
		(0.005)		(0.001)		(0.002)		(0.002)
capital_ratio.x.y.l		0.050**		0.029***		0.106***		0.091***
		(0.024)		(0.005)		(0.022)		(0.019)
group_d	0.081***	0.039***	0.016***	−0.0002	0.020***	0.003	0.018***	0.001
	(0.009)	(0.012)	(0.002)	(0.002)	(0.007)	(0.009)	(0.006)	(0.008)
lag1:xlessy	−0.942***	−1.013***	0.053***	−0.003	0.030*	−0.145***	−0.023	−0.148***
	(0.049)	(0.055)	(0.018)	(0.014)	(0.018)	(0.022)	(0.014)	(0.017)
Observations	180,568	108,702	180,568	108,702	180,568	108,702	180,568	108,702
R^2	0.510	0.534	0.296	0.403	0.427	0.512	0.428	0.496
Adjusted R^2	0.493	0.514	0.272	0.378	0.409	0.491	0.410	0.475

	Dependent variable:							
	JACCARD M		Min Edit Dist		LM Positive		lm uncertainty	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lag1	−0.012***	0.004***	−0.035***	−0.023***	−0.240***	−0.213***	−0.345***	−0.369***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.010)	(0.012)	(0.009)	(0.011)
xlessy	−0.034***	−0.031***	−0.017***	−0.009***	−0.00002	−0.00003	−0.0002***	−0.0002***
	(0.0005)	(0.001)	(0.001)	(0.001)	(0.00001)	(0.00002)	(0.00002)	(0.00002)
bm.x.y.l		0.00001		0.00000		0.00000		−0.00000*
		(0.00001)		(0.00001)		(0.00000)		(0.00000)
lnassets.x.y.l		−0.001***		0.003***		0.00002***		−0.00004***
		(0.0001)		(0.0002)		(0.00000)		(0.00000)
capital_ratio.x.y.l		0.001		0.003**		−0.00001		−0.0001***
		(0.001)		(0.001)		(0.00001)		(0.00003)
group_d	−0.0001	−0.0005	0.004***	0.002***	−0.00002***	−0.00002**	−0.0001***	−0.0001***
	(0.0005)	(0.001)	(0.001)	(0.001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
lag1:xlessy	−0.988***	−1.048***	−0.349***	−0.299***	0.154***	0.131***	0.141***	0.222***
	(0.007)	(0.010)	(0.010)	(0.012)	(0.011)	(0.012)	(0.011)	(0.013)
Observations	180,568	108,702	180,568	108,702	180,568	108,702	180,568	108,702
R^2	0.881	0.906	0.426	0.465	0.308	0.403	0.311	0.383
Adjusted R^2	0.877	0.902	0.407	0.442	0.285	0.378	0.288	0.357

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$