(In)Consistent Internal-External Communication? How Internal Communication Technology Affects Voluntary Disclosure

Ehsan Azarmsa, Lisa Yao Liu, and Suzie Noh*

March 2023

Abstract

We investigate how improvements in organizations' internal communication technology affect their voluntary disclosure, a form of external communication. By developing a model with a headquarters manager and several divisional managers, we formalize two competing economic forces—information learning and free riding—that shape the headquarters manager's information precision and, consequently, voluntary disclosure. While improved internal communication technology helps the headquarters manager to collect information from divisional managers, it reduces divisional managers' incentives to acquire information because these divisional managers anticipate the other divisional managers' information acquisition. As a result, these two forces jointly produce an inverse U-shape relation between internal and external communications. We empirically document robust inverse U-shapes for both public and private firms, and show evidence consistent with the two economic forces. Collectively, our paper furthers our understanding of voluntary disclosure from the perspective of internal agency frictions and sheds light on firms' internal-external communications.

Keywords: organizational communication, centralization, voluntary disclosure, guidance, management forecasts, information technology

JEL Classification: D80, D83, L20, M15, M21, M41

---

*Ehsan Azarmsa: Price College of Business, University of Oklahoma (azarmsa@ou.edu). Lisa Yao Liu: Columbia Business School, Columbia University (yl4689@gsb.columbia.edu). Suzie Noh: Stanford Graduate School of Business, Stanford University (suzienoh@stanford.edu). We appreciate helpful comments from Rishabh Aggarwal, Tim Baldenius, Phil Berger, Anne Beyer, Thomas Bourveau, Matthias Breuer, Jungho Choi, Ki-soon Choi, Jon Glover, Kyle Jung, Chongho Kim, Jung Min Kim, Sehwa Kim, Iván Marinovic, Xiumin Martin, Rodrigo Verdi, Amir Ziv, and participants at CCSRG at the University of Chicago and Columbia University. We thank Andy (Jiarui) Su, Ziqing Wang, and Sunho Yoo for their research assistance. We gratefully acknowledge financial support from our respective schools.

1 Introduction

The past decades have witnessed an increased use of technology in the workplace, which has led to faster and more effective collection, documentation, and reporting of information within firms. Conventional wisdom suggests that enhanced internal communication technology improves external communication as a result of direct information transfers. Specifically, firms that better collect information from divisions can improve internal information environments (e.g., Berger et al., 2022; Gallemore and Labro, 2015), and subsequently, external communication to stakeholders (e.g., Chen et al., 2018; Samuels, 2021).[1] However, if we endogenize information acquisition by divisional managers, it becomes less clear how firms' internal information environment and external communication will change. To gain a more comprehensive understanding of the relation between firms' internal communication technologies and their external communications, we develop a model that takes into account the endogenous nature of information acquisition and then use this model to guide our empirical analysis.

Our theoretical framework contrasts two economic forces through which improved internal communication technology shapes corporate voluntary disclosure: (1) information learning: divisional managers (or "DM") acquire information that is communicated to the headquarters manager (hereafter, "HQ manager"); and (2) free riding: increased internal communication reduces the marginal impact of each DM's acquired information on the HQ manager's decision-making, creating free-riding incentives and hence, diminishing their incentives to exert effort to acquire information. Collectively, as technology improves internal communication, the two competing economic forces jointly determine a non-monotonicity in the provision of voluntary disclosure.[2]

In our model (detailed in Section 2), there is a single HQ manager and several divisional managers. At the divisional level, the optimal decision depends on the aggregate state and some division-specific information. Each divisional manager can incur costs to acquire such information.

[1] In the paper, we use stakeholders to broadly refer to external parties with interest in the fundamental value of the firm (e.g., equity holders).

[2] As we focus on internal agency friction, we abstract away from external agency friction. We assume that, conditional on disclosing, the HQ manager truthfully reports her information to stakeholders. Prior literature shows stakeholders demand disclosure from both public and private firms (Breuer et al., 2020). While stakeholder demand for disclosure and external agency frictions may vary across firms with different ownership types, we assume that internal agency frictions are unrelated to these external factors.

For example, divisional managers can learn about customers or local investment opportunities (e.g., Berger et al., 2022; Chen et al., 2018). The extent of internal communication is pre-determined for reasons unrelated to divisional managers' incentives to produce information.[3] When internal communication is feasible, the divisional manager can report the acquired information to the HQ manager, if they choose to acquire information. When the HQ manager makes decisions, she makes division-level decisions based on two information sources: the information collected from communicable divisional managers (i.e., divisions with internal communication technology) and the HQ manager's private information about the aggregate state. For those non-communicable divisions (i.e., divisions without internal communication technology), the HQ manager pre-commits to delegating the decision-making to those divisional managers.

We endogenize divisional managers' information acquisition decisions and study the dynamics of the HQ manager's information set and, consequently, the HQ manager's voluntary disclosure provision. When there are only a few communicable divisions, these DMs have strong incentives to acquire information as their information can influence the HQ manager's decision toward their preferred action. By aggregating information from divisional managers, the HQ manager becomes more informed. Therefore, the HQ manager's information precision and the probability of voluntary disclosure increase.

However, as more divisional managers acquire information, the marginal impact of divisional managers' information on decisions made for their own division becomes smaller because the other divisional managers also share their acquired information with the HQ manager. Consequently, as the proportion of divisions communicating with the HQ manager exceeds a threshold value, divisional managers' incentives to acquire information decrease. If this free-riding force dominates the information learning force, then the HQ manager's information precision and the probability of voluntary disclosure decline.

We use our theoretical model to guide the empirical analyses. An important feature of our model is that it focuses on internal agency frictions that are not specific to a certain ownership type,

[3]In Section 2.1, we discuss possible reasons for implementing internal communication technology, such as resource allocation and security concerns. We assume that internal communication technologies are adopted for certain divisions for reasons described in Bloom et al. (2014), such as making decisions directly for divisions that may be economically or strategically important.

so we study both public and private firms.[4] We first assess whether there is a non-linear, inverse U-shape relation between internal communication with divisional managers facilitated by technology and external communication with stakeholders. We then cross-sectionally test whether the inverse U-shape is muted when the HQ manager has more prior information. Finally, we investigate whether firms with more (less) communicable divisions subject to similar industry conditions have more free-riding incentives, thereby showing an inverse U-shaped (positive linear) relation. The latter two analyses, respectively, substantiate the two endogenously generated economic forces predicted by our theory that produce the inverse U-shape: information learning and free riding.

To test the theoretical predictions, we use firms' intranet usage to proxy for advances in technology that facilitate internal communication. The intranet does not allow access to anyone outside their network and allows employees to store and share proprietary data securely. Therefore, it facilitates intra-firm communication and information sharing. Following Bloom et al. (2014), we measure firms' use of the intranet using the Harte-Hanks Ci Technology database (CiDB), which contains information on hardware, software, and personnel at the site level.[5] We construct a measure of intra-firm intranet intensity, $Intranet$, by calculating the proportion of divisions within a firm-year that adopt the intranet technology, weighted by the number of employees at each division following Bloom et al. (2014). This weighting also accounts for potential variations in the economic and informational importance of divisions to the organization. We obtain both public and private firm data from the CiDB.

We use two proxies for external communication, management EPS forecasts and corporate websites. The literature on voluntary disclosure tends to use management forecasts as a primary measure of firms' voluntary disclosure, because they are pervasive, informative, and represent a broad spectrum of voluntary disclosure (e.g., Beyer et al., 2010). Following the literature and our theoretical predictions, we use the frequency of management EPS forecasts to measure firms' provision of voluntary disclosure.[6] Because EPS forecasts are only available for public firms, we

[4]As public firms are subject to more capital market forces that may influence firms' voluntary disclosure, we study private firms to assess the degree to which capital market forces and external monitoring may drive our results.

[5]Due to the CiDB's broad coverage and high accuracy, it has been widely used by academics (e.g., Bresnahan et al., 2002; Beaudry et al., 2010; Forman et al., 2012; Bloom et al., 2016, 2012; Charoenwong et al., 2022).

[6]e.g., Skinner (1994); Balakrishnan et al. (2014); Boone and White (2015); Glaeser (2018); Heinle et al. (2022).

use corporate websites as our second measure of external communication. This newly developed measure of voluntary disclosure is available to both public and private firms. Prior research (e.g., Boulland et al., 2021; Lynch and Taylor, 2022) validates that corporate website disclosure is positively associated with established measures of firms' voluntary disclosure but not subsumed by those measures. We use this alternative dependent variable as a broad proxy for firms' disclosures about their business and products.

The main sample consists of 2,946 unique public firms and 60,501 unique private firms during the sample period of 2001 to 2015. Following Samuels et al. (2021) and Kim et al. (2021), we do not include firm fixed effects in the empirical specification, because non-linearity is better captured in the large cross-section of firms and can hardly be captured in limited within-firm variation. This choice of the empirical specification is also consistent with our theory that compares the degree of internal communication across different firms instead of within-firm changes.

Our first set of analyses documents a non-monotonicity relation between internal communication technology and external communication. We first graphically examine the shape of the relation. In particular, we sort firms into quintiles based on $Intranet$, our proxy for internal communication technology. We then plot average values of EPS forecasts and corporate website length, our two proxies for external communication, for each quintile of $Intranet$. We find consistent inverse U-shape patterns for both public and private firms (see Figure 8). We then conduct regression analyses, including the second-order polynomial of $Intranet$. Across different samples and disclosure proxies, we find a negative and statistically significant coefficient on this quadratic term, suggesting an inverse U-shape. Our inverse U-shape finding for private firms, which arguably face greater internal agency frictions than external agency frictions, is particularly useful, as it is consistent with the inverse U-shape driven by internal agency frictions.

We next provide empirical evidence of the two endogenously generated economic forces for the inverse U-shape between $Intranet$ and voluntary disclosure. With respect to the information learning channel, the theory predicts that when the HQ manager's prior information is more precise, internal communication contributes less to the HQ manager's information set, and the resulting relation between internal and external communications becomes flat. Therefore, we test whether the inverse U-shape between internal and external communications is muted (pronounced) when the HQ managers' prior informedness is high (low), as predicted by theory. We use the employee-weighted average distance between divisional sites and the headquarters to proxy for HQ managers' ex ante informedness. When the headquarters office is located closer to divisional sites, the HQ manager may easily gather information herself, resulting in more precise information regarding divisions in the absence of internal communication. Consistent with our prediction, we find that the inverse U-shape is muted (pronounced) in the subsamples of public firms with nearby (geographically dispersed) divisions.

As for the free-riding channel, the theory predicts that, when the common component in the information structure (i.e., the aggregate state in the model) is important across all communicable divisions, the role of other divisional managers' acquired information in the HQ manager's informedness is large. As a result, the free-riding incentive of divisional managers becomes prominent. To empirically substantiate the free-riding channel, we construct a measure of similarity between divisions based on whether or not their industries have the same first two-digit of the SIC industry code. These divisions likely face similar industry conditions, so each divisional manager may have incentives to free-ride on the others' information acquisition to inform the HQ manager. We predict and find that when the similarity measure is smaller, the free-riding incentive is less prominent, and hence, the negative force contributing to the inverse U-shape is weakened. When the similarity measure is greater and the free-riding incentive is stronger, the inverse U-shape, especially the negative slope, becomes pronounced.

Collectively, we find that the theoretical predictions of the inverse U-shape between internal communication technology and external communication hold across different settings, including public and private firms, and using different measures of voluntary disclosure, including management forecasts and corporate websites. Additionally, we show evidence consistent with the inverse U-shape relation being driven by two competing economic forces—"information learning" and "free riding." We further document that the non-monotonicity relation also holds when we account for ad hoc nonlinear relationships between voluntary disclosure and our control variables.[7]

[7] In our Online Appendix, we use the EPS forecast accuracy of firms that are always forecasters as an alternative measure of voluntary disclosure where the disclosure's informativeness is discretionary (conditional on disclosing).

This paper makes several contributions. To start, our paper adds to the literature on voluntary disclosure in two ways. First, prior research mainly takes the HQ manager's information as given, whereas we endogenize the internal information production process. We examine a model with many divisional managers, whose information production decisions interact. By so doing, it presents the interaction of two endogenously generated economic forces: information learning and free riding. We theoretically and empirically document an inverse U-shape between internal communication technology and external communication, and provide a seemingly counter-intuitive result that better internal communication technology could worsen external communication for some firms. As such, our paper furthers our understanding of voluntary disclosure from the perspective of internal agency frictions and, in doing so, features distinct economic forces from extant literature that focuses on external agency frictions (see reviews in Verrecchia (2001), Healy and Palepu (2001), and Beyer et al. (2010)).[8] Second, internal agency frictions are also prevalent in private firms, so this paper furthers our understanding of U.S. private firms' transparency and internal communication. Private firms are an integral part of the economy, but less is known about the determinants of their voluntary disclosure (Minnis and Shroff, 2017).[9] Our paper provides theoretical and empirical evidence on the relation between their internal communication technology and external communication.

Furthermore, internal communication also relates to managerial accounting, particularly relating to within-firm information-sharing and organizational structure. Ittner and Larcker (2001) advocate for more research into the integration of financial and managerial accounting research, which can improve our understanding of the choice and performance implications of internal and external accounting and control systems. By studying how internal and external communications are related, our paper adds to the branch of the literature on interdependencies between manage-

For a much smaller number of firms that always provide forecasts, we show a similar inverse U-shape. To the extent that forecast accuracy proxies for the HQ manager's information set, this result is also consistent with our theoretical prediction in Section 2.3.1.

[8]By combining disclosure costs (Verrecchia, 1983) and managers' private information (Dye, 1985; Jung and Kwon, 1988), Kim et al. (2021) theoretically and empirically show that the two forces jointly generate a non-monotonicity in voluntary disclosure. Relatedly, Richarson (2001) extends Verrecchia (1983) in which the disclosure cost rises as the manager's private information becomes more precise and posits the possibility of a unimodal relation. Recent literature has also looked at different economic forces such as stakeholder demand (e.g., Breuer et al., 2020) and contracting with supply chain partners (e.g., Bourveau et al., 2022).

[9]The literature has studied the determinants of private firm disclosures using the European setting, where private limited liability firms are required to disclose financial statements (e.g., Dedman and Lennox, 2009; Gassen and Muhn, 2018; Breuer et al., 2020).

ment accounting systems and external financial reporting (e.g., Kaplan, 1984; Hemmer and Labro, 2008; Zimmerman, 2009; Dichev et al., 2013; Ittner and Michels, 2017; Samuels, 2021).

Lastly, this paper adds to the emerging accounting literature that integrates theory with empirics (e.g., Bertomeu et al., 2016; Chen et al., 2016) and contributes to the theoretical literature on the interaction between the allocation of control and information (e.g., Aghion and Tirole, 1997; Dessein, 2002; Acemoglu et al., 2007). We develop a formal theory model that makes our economic forces transparent and offers a framework for making empirical predictions (Bertomeu et al., 2016). In our model, the divisional manager's information structure is similar to that of Aghion and Tirole (1997), in which the divisional manager is incentivized to acquire information on economic uncertainty when delegated. However, we have multiple divisional managers, who act as the HQ manager's agents in information acquisition and control. As the HQ manager obtains information from more divisional managers through internal communication, the marginal impact of any divisional manager on decision-making declines. It's a mechanism absent in the prior literature but crucial in generating non-monotonicity in the HQ manager's information set and, thus, their disclosure.

Our paper is subject to a few caveats. Our tests do not have an exogenous shock and therefore, there may be correlated omitted variable concerns. However, all of the tests that use cross-sectional variations are consistent with our theory that compares organizations with varying degrees of internal communication. We also implement the "Bartik instrument"-style analysis to purge the treatment of some endogenous variation and show similar results (e.g., Breuer, 2022). Furthermore, our empirical identification hinges on the inverse U-shape. This identification approach is akin to "identification by functional form" (e.g., Lewbel, 2019; Samuels et al., 2021). Alternative explanations would need to collectively explain our main non-monotonic findings for voluntary disclosure, and cross-sectional results on information learning and free riding. The final caveat is that our theory and empirical tests focus on the information implications of internal communication technology. Our results cannot speak to the optimal level of the adoption of internal communication technology, as the decision to adopt such technology may be influenced by factors other than information benefit.

2 A Model of Internal Information Production and Communication

2.1 Setup

There are three periods, $t = 0$, 0.5, and 1. A firm is run by the headquarters manager or HQ (female). The firm has $N > 1$ divisions, indexed by $i = 1, \ldots, N$. Each division is managed by a divisional manager or DM, which we denote by DM_i (male). At $t = 0.5$, a decision should be made for each division, which is represented by $a_i \in \mathbb{R}$. The decision can be interpreted broadly. For instance, it could represent the scale of production of a product or the provision of a service in each division. The decision depends on an aggregate state, represented by y, and a division-specific component, represented by Δ_i.

Divisional managers are concerned only about the decision in their own divisions.[10] In particular, for division $i \in \{1, \ldots, N\}$, DM_i has unimodal preferences over action a_i, maximized at $y + \Delta_i$:

$$U_{DM_i}(a_i; y, \Delta_i) = -(a_i - y - \Delta_i)^2 - c_i \mathbb{I}_{effort}. \tag{1}$$

The optimal action for each divisional manager depends on an aggregate (y) and a division-specific component (Δ_i). y is normally distributed with mean zero and variance σ_y^2. Moreover, each DM can research and perfectly learn about his optimal action ($y + \Delta_i$) at a random cost $c_i \in [\underline{c}, \bar{c}] \subset \mathbb{R}_+$, which has cumulative distribution function $G(\cdot)$. For simplicity, we assume that the distribution has no mass point in its domain. c_i is privately observed by DM_i, so it is not known by HQ and other DMs.

We also assume that $\bar{c} > \sigma_\Delta^2$; that is, the information acquisition is not always so low that DMs always find it beneficial to acquire information. Moreover, DMs have no other information source that enables them to distinguish the aggregate component from the division-specific component. Without learning, Δ_i ($i = 1, \ldots, N$) is normally distributed with a zero mean and a variance of σ_Δ^2. Δ_i's are distributed independently from each other and from the aggregate state y.

[10]This assumption is not crucial in driving our predictions. If the decision in each division affected all DMs, then the DMs would be more willing to produce information because better information production in each division improves the quality of decision-making in all divisions. However, it would not impact the key qualitative results.

HQ privately observes a noisy signal of y, specifically $\hat{y}_0 = y + \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. σ_ε^2 denotes the variance of HQ's signal without using any other source of information. We allow for the presence of a misalignment of incentives between HQ and DMs. HQ assigns weight $\beta \in [0,1]$ to the idiosyncratic component:

$$U_{HQ}(a_1, \ldots, a_N; y, \Delta_1, \ldots, \Delta_N) = -\frac{1}{N}\sum_{i=1}^{N}(a_i - y - \beta\Delta_i)^2. \tag{2}$$

For simplicity, we assume β is the same across all divisions.[11] The misalignment in preferences can arise for a number of reasons. First, DMs might overweight the importance of the division-specific component in the optimal action for the firm. Second, HQ might face internal constraints in treating the divisions differently (e.g., Scharfstein and Stein, 2000). In (2), β captures the extent of alignment in the preferences between HQ and DMs. We assume that HQ's information reporting preferences are aligned with those of the stakeholders, consistent with Chen et al. (2018), Samuels (2021), and Verrecchia (1990), who we broadly define as external parties with interest in the fundamental value of the firm. Thus, we abstract away from external agency frictions between stakeholders and HQ, whose role in external communication is examined in the literature (e.g., Fischer and Verrecchia, 2000). This assumption implies that HQ truthfully reports her information about the aggregate state (y) to stakeholders.[12] As such, stakeholders have a passive role in our model in the sense that they only receive information from HQ and make no decisions.

HQ can communicate with $1 \leq I < N$ DMs and learn their optimal action ($y + \Delta_i$) perfectly. We refer to these I DMs as the "communicable" DMs, and the rest as "non-communicable." $x \equiv \frac{I}{N}$, the fraction of communicable DMs, is a key quantity in our analysis. This fraction represents the extent of internal communication facilitated by technology, and, in Section 4, we empirically

[11]We show that our results are not driven by the degree of preference misalignment between DMs and HQ, captured by β, as the key results regarding the internal information production hold even under the case with a perfect alignment in preferences, i.e., $\beta = 1$. Moreover, the assumption that the βs are the same across all divisions is not crucial.

[12]In the presence of agency friction between HQ and stakeholders, HQ shares her information with a bias. The bias does not affect our results qualitatively as long as it is orthogonal to internal agency frictions. For instance, in Fischer and Verrecchia (2000), HQ adds a noise term to her report, which is independent of HQ's signal. For simplicity, we assume that stakeholders are only interested in learning about the aggregate state, and not the division-specific variables. However, were stakeholders interested in the Δ_i's, the precision of HQ's report for those variables would be equal to the precision of her information about $\hat{y}$ if HQ perfectly observes $y + \Delta_i$. If $y + \Delta_i$ is not available, then HQ has no information to share.

document its relationship with firms' external communication.

A key assumption in our baseline model is that the set of communicable DMs, I, is predetermined (before our model timeline) for reasons unrelated to DMs' information production incentive.[13] For example, internal communication technology is adopted to improve cyber security or efficient resource allocations across divisions. We assume that I is adopted for reasons discussed in Bloom et al. (2014), who shows that an important reason for HQ to implement internal communication technologies for certain divisions is to directly give instructions and orders to those divisions. For instance, if those divisions are economically or strategically important, HQ may prefer to rely on internal communication technology to make decisions directly instead of letting DMs make decisions.[14] Consistent with this notion, we assume that, HQ pre-commits to a decision-making policy by which she makes decisions herself for communicable divisions and delegates the decision to non-communicable divisions (Mookherjee, 2006). In Section 2.4, we show that our results hold for any type of pre-commitment on how DMs' information is implemented, including pre-committing to selecting DMs' desired action.

Some other assumptions are as follows. We assume that DMs report their optimal action truthfully and costlessly. Therefore, we abstract away from strategic communication within corporations, which is examined by Aghion and Tirole (1997), Dessein (2002), and Harris and Raviv (2005), among others. However, in Section 2.4, we relax this assumption by allowing DMs to misreport. We find our results hold in equilibrium. Additionally, as we endogenize DMs' information acquisition decisions, we do not allow HQ to force DMs to acquire information. That is, the level of effort for information acquisition is not contractible or enforceable. The implication is that, even if we allowed HQ to force DMs to acquire information, the information would still be less accurate than HQ's optimal level. Furthermore, when a communicable DM acquires information, he is aware that he commits to sharing his information with HQ. Figure 1 illustrates the structure of internal communication. Figure 2 provides the timeline (recall that HQ pre-commits to a decision-making

[13] In Section 2.5, we show that our economic forces are preserved even when we endogenize I, although the model becomes too complex. We make I exogenous in the main model to shed light on the implications of the internal information environment for a given set of communicable DMs.

[14] In the Online Appendix Table OA.1, we show that intranet adoption is not positively correlated with some empirical proxies for internal information asymmetry, such as the distance between HQ and DM, but is positively associated with the site's economic importance, such as the site's revenue and employees.

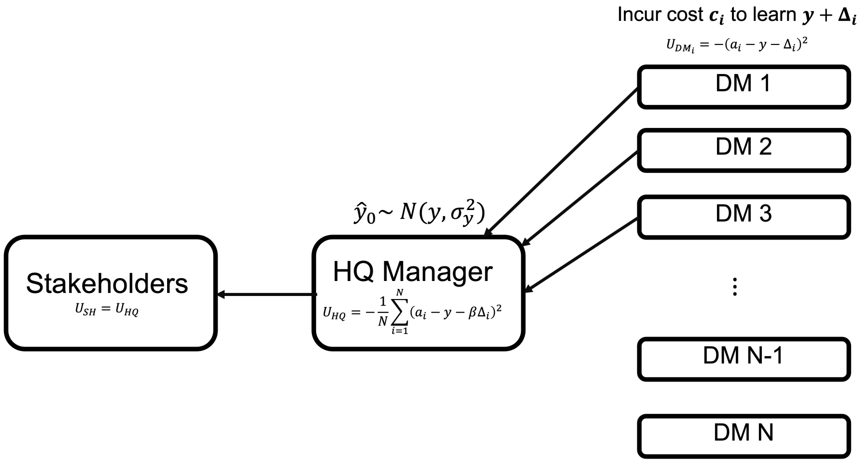


Figure 1: Internal information production and communication

policy by which she makes decisions herself for communicable divisions and delegates the decision to non-communicable divisions).

After collecting all available signals, HQ's posterior information about the aggregate state is characterized by normal distribution $y \sim \mathcal{N}(\hat{y}, \hat{\sigma}^2)$. Thus, $\hat{y}$ and $\hat{\sigma}^2$ are sufficient variables for HQ to fully share her information about y with the stakeholders. Note that in our model, the precision of information communicated by divisional managers is random, as it depends on the cost of information acquisition by communicable DMs.

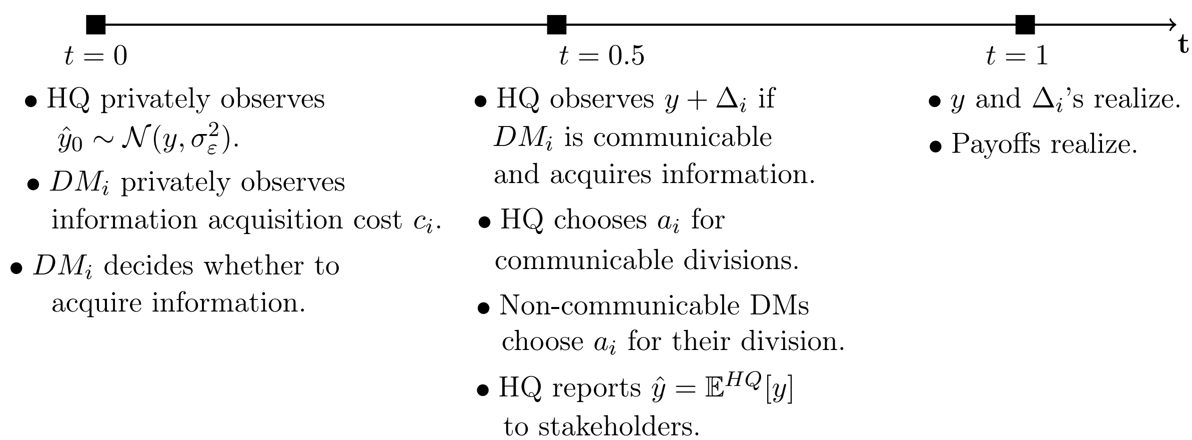


Figure 2: Timeline of the game

If the first k (which will be endogenized later) DMs produce information and communicate to HQ, the estimated value shared with stakeholders is:

$$\hat{y} = \mathbb{E}[y|\hat{y}_0, y + \Delta_1, \ldots, y + \Delta_k]. \tag{3}$$

Due to the normality and independence of the random variables, one can show:

$$\hat{y} = \gamma_k^{HQ} \hat{y}_0 + \gamma_k^{DM} \bar{y}, \tag{4}$$

where $\bar{y}$ is the average signal HQ obtains from the DMs:

$$\bar{y} = y + \bar{\Delta}, \qquad \bar{\Delta} = \frac{1}{k} \sum_{i=1}^{k} \Delta_i. \tag{5}$$

Moreover,

$$\gamma_k^{HQ} = \frac{\sigma_\Delta^2}{k\sigma_\varepsilon^2 + \sigma_\Delta^2 + \sigma_y^{-2}\sigma_\varepsilon^2\sigma_\Delta^2}, \qquad \gamma_k^{DM} = \frac{k\sigma_\varepsilon^2}{k\sigma_\varepsilon^2 + \sigma_\Delta^2 + \sigma_y^{-2}\sigma_\varepsilon^2\sigma_\Delta^2}. \tag{6}$$

Appendix A.1 provides the details of the derivations. Equation 4 illustrates how HQ weights her private information against DMs' information in forming the estimated value of y. As suggested by (6), HQ assigns a larger weight to DMs' information when her prior knowledge about the aggregate state is less accurate (i.e., higher σ_ε^2), or receives more signals from the DMs (i.e., higher k). Note that this result does not depend on the set of communicable DMs, or the set of DMs that acquire information.

The variance of HQ's estimation of y is:

$$\hat{\sigma}_k^2 = Var[y|\hat{y}_0, y + \Delta_1, \ldots, y + \Delta_k] = \frac{\sigma_\varepsilon^2\sigma_\Delta^2}{k\sigma_\varepsilon^2 + \sigma_\Delta^2 + \sigma_y^{-2}\sigma_\varepsilon^2\sigma_\Delta^2}. \tag{7}$$

We see that HQ provides more accurate information to stakeholders as she obtains information from more DMs. However, information acquisition is an endogenous decision in our model, which implies that HQ does not necessarily obtain more information with an increase in the fraction of communicable DMs.

HQ's optimal action for communicable divisions is:

$$a_i^{HQ*} = \begin{cases} (1-\beta)\hat{y} + \beta(y+\Delta_i) & \text{Info for division } i \text{ is available} \\ \hat{y} & \text{Info for division } i \text{ is not available.} \end{cases} \tag{8}$$

The details of the derivation are available in Section A.2. We see that when HQ has no information about a division, her optimal action would only be based on her forecast about the aggregate state ($\hat{y}$). If information about a division is available, HQ assigns weight β to the value reported by the corresponding DM (i.e., $y + \Delta_i$), and assigns weight $1-\beta$ to her estimation of y, which also depends on the reported value. Recall that DMs would assign weight one on their reported value and zero to $\hat{y}$. In Section 2.4, we consider a flexible set of decision rules and allow the DM to report his information strategically. We abstract from those complications here as they do not impact the key insights of the model.

Non-communicable DMs acquire information if their cost is not more than $\sigma_y^2 + \sigma_\Delta^2$. In this case, they choose $a_i^{DM*} = y + \Delta_i$. If the cost is large, then $a_i^{DM*} = \mathbb{E}[y + \Delta_i] = 0$.

2.2 Information Acquisition by Divisional Managers (DMs)

In this subsection, we analyze the decision to acquire information by DMs. For non-communicable DMs, they acquire information if their cost does not exceed $\sigma_y^2 + \sigma_\Delta^2$. For communicable DMs, their decision depends on how impactful their information is on HQ's decision for their division. This inter-dependence in information acquisition introduces a "free-riding motive" in information acquisition: If a DM believes that a large fraction of the other DMs acquires information, he knows that HQ has precise information about the aggregate state y. It crowds out the DM's incentive to acquire information about his optimal action ($y + \Delta_i$), which depends on both aggregate and division-specific states. Lemma 1 presents how the marginal gain from information acquisition with the number of communicable DMs that acquire information.

Lemma 1. *If DM_i is communicable and $k-1$ other communicable DMs acquire information, the expected net gain from acquiring information is $\pi_k - c_i$, where*

$$\pi_k = \hat{\sigma}_{k-1}^2 - (1-\beta)^2 \hat{\sigma}_k^2 + (2\beta - \beta^2)\sigma_\Delta^2. \tag{9}$$

The expression for $\hat{\sigma}_k^2$ is provided by Equation 7. Moreover, π_k is strictly decreasing in k and converges to $(2\beta - \beta^2)\sigma_\Delta^2$ from above as k goes to infinity.

Lemma 1 demonstrates that the expected gain from information acquisition is strictly decreasing in the number of other communicable DMs that acquire information. Note that the statement holds even for $\beta = 1$, when there is no misalignment in the preferences between HQ and DMs. That is, the free-riding mechanism identified here is present even when the preferences are perfectly aligned. Now, we use this result to study the endogenous information production behavior of communicable DMs. Proposition 1 reports the results.

Proposition 1. ***a)*** *All communicable DMs produce information iff*

$$\pi_{\bar{I}} \geq \bar{c}. \tag{10}$$

b) *No communicable DM produces information iff*

$$\pi_1 \leq \underline{c}. \tag{11}$$

c) *Suppose neither of conditions 10 nor 11 holds. Moreover, let $c^* \in [\underline{c}, \bar{c}]$ be such that:*

$$\sum_{k=1}^{I} \binom{I-1}{k-1} G(c^*)^{k-1}(1 - G(c^*))^{I-k}(\pi_k - c^*) = 0. \tag{12}$$

If DM_i is communicable, then he produces information iff $c_i \leq c^*$.

Proposition 1 describes how the fraction of communicable DMs impacts their information acquisition decision. To understand the proposition, recall that $\pi_k - c_i$ is the expected gain from information acquisition. The expected gain decreases as more communicable DMs acquire and share information with HQ, because HQ aggregates all of these decisions for decision-making, as shown by Lemma 1. Therefore, when there are only a few communicable DMs, their information is influential enough that the benefit from information acquisition outweighs its cost. However, when the number of communicable DMs exceeds some threshold value, communicable DMs only produce information

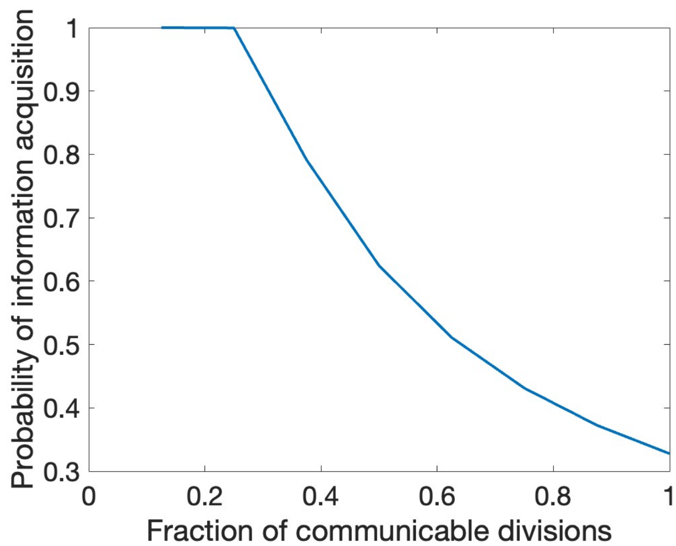


Figure 3: The relationship between the number of communicable DMs and the probability of information acquisition by communicable DMs.

when the information acquisition cost is sufficiently small. This relationship is illustrated in Figure 3.

In this regard, Proposition 2 in particular states that the probability of information acquisition by communicable DMs strictly decreases with the number of communicable DMs (I) after some threshold value.

Proposition 2. *Suppose $\bar{I}$ is the greatest positive integer for which $\pi_I \geq \bar{c}$. Then, for $I \leq \bar{I}$, all communicable DMs acquire information with probability one. For $I > \bar{I}$, the probability of information acquisition ($G(c^*)$) decreases in I.*

2.3 Implications

2.3.1 Internal communication and the precision of the HQ manager's information set

As stated in Equation 7, HQ can generate a more precise estimate of the aggregate state (y) when she receives more signals. Moreover, we see that the relationship is convex, which means that the information precision could be hurt by the uncertainty in the number of communicable DMs that acquire information (because of the stochastic information acquisition cost).[15] In particular, counter-intuitively, the probability of only a small number of DMs acquiring information for HQ

[15] Recall that HQ does not observe DMs' cost of information acquisition. Therefore, she is uncertain about the amount of information she will receive from DMs.

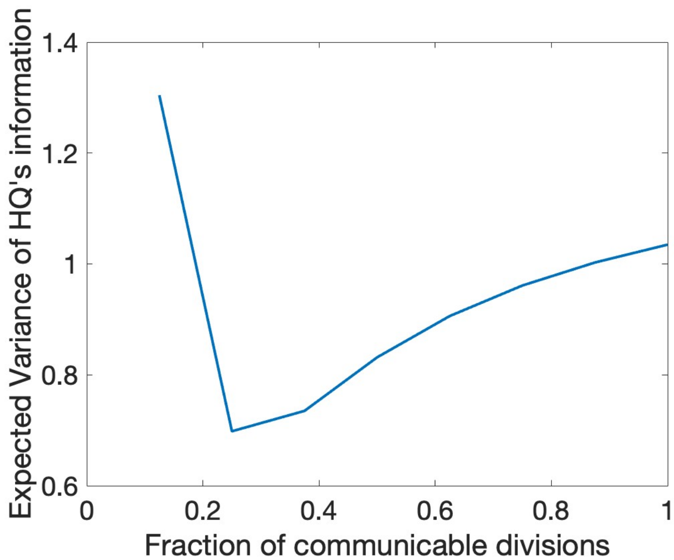


Figure 4: The average variance of the information about the aggregate state (y) as a function of the fraction of communicable DMs.

goes up with the fraction of communicable DMs. For instance, the probability that HQ does not receive any signals from DMs is zero when $I \leq \bar{I}$ (i.e., when the fraction of communicable DMs is less than the threshold value in Proposition 2), because DMs' information is sufficiently impactful to justify the information acquisition cost. However, for $I > \bar{I}$, there is a positive probability that no communicable DM acquires information due to the random nature of information acquisition cost and the fact that each DM believes that he has a small impact on HQ's decisions.[16] As a result, the average variance of HQ's belief about y (i.e., $\mathbb{E}[\hat{\sigma}_k^2]$) could be U-shaped, as illustrated in Figure 4.

The intuition for the U-shaped relationship is as follows: As stated earlier, when $I \leq \bar{I}$, all communicable DMs acquire information.[17] Therefore, within the range $I \leq \bar{I}$, $\bar{\sigma}_{HQ}^2$ unambiguously decreases with I, which represents the downward part of Figure 4. However, for $I \geq \bar{I}$, the uncertainty in the number of signals received from communicable DMs increases because each communicable DM is less likely to acquire information, according to Proposition 2. Therefore, the probability of having too few signals about y may increase with the fraction of communicable DMs after this threshold, which generates a high value of $\bar{\sigma}_{HQ}^2$. This explains the upward part of Figure

[16]Recall that DMs cannot observe each other's cost of information acquisition. Otherwise, DMs who know that they have the lowest information acquisition cost would acquire information.

[17]$\bar{I}$ depends on the highest possible cost of information production ($\bar{c}$), along with other model parameters.

4.

An important takeaway from this result is that, due to the endogenous nature of information acquisition, HQ with better internal communication infrastructure does not necessarily learn more from her divisions.

2.3.2 Internal communication and voluntary disclosure

Proposition 2 implies that the precision of HQ's information is non-monotone in the fraction of communicable DMs. To build a basis for our empirical analysis, we consider the implication for HQ's provision (frequency) of external disclosure. As previous theoretical studies show, a positive relationship exists between the frequency of voluntary disclosure and the precision of HQ's information (Verrecchia, 1990; Richarson, 2001; Kim et al., 2021). As a result, if we allow HQ to choose whether to disclose information to stakeholders, the non-monotonicity result for HQ's information precision can carry over to her probability of voluntary disclosure.

To formalize this point, consider the following extension of the model. Suppose HQ has the option to withhold her information about the aggregate state y from the stakeholders. Moreover, HQ's objective in her disclosure decision is to maximize $\mathbb{E}^{SH}[y]$, namely, the expected value of the aggregate state given the stakeholders' information set. For instance, HQ might aim to maximize the share price, and the price is monotonically related to $\mathbb{E}^{SH}[y]$. HQ discloses the precision of her information ($\hat{\sigma}^2$) only when she discloses $\hat{y}$. Following the voluntary disclosure literature, the disclosure is associated with a fixed cost, denoted by $m > 0$. The cost could represent a loss to stakeholders due to revealing proprietary information to competitors. Therefore, HQ discloses her forecast of y ($\hat{y}$) when

$$\hat{y} - m \geq y^*(I), \tag{13}$$

where $y^*(I)$ is the stakeholder's expected value of y without the disclosure. Note that Equation 13 implies a threshold rule for the disclosure; that is, HQ discloses $\hat{y}$ when it is above endogenous threshold $y^*(I) + m$. The threshold value is set such that HQ is indifferent between disclosing and not disclosing when $\hat{y} = y^*(I)$. In other words,

$$y^*(I) - m = \mathbb{E}^{SH}[\hat{y}|\hat{y} \leq y^*(I)]. \tag{14}$$

In (14), $\mathbb{E}^{SH}[\cdot]$ represents stakeholders' expectations when they receive no information from HQ other than what they learn from the lack of disclosure. Note that stakeholders are uncertain not only about HQ's forecast of y, but also about the precision of her information. In Appendix A.6, we show that $y^*(I)$ is the solution to the following equation:

$$
\begin{gathered}
y^*(I) - m + \sum_{k=0}^{I} \binom{I}{k} G(c^*)^k (1 - G(c^*))^{I-k} \frac{\Sigma_k \phi(\Sigma_k^{-1} y^*(I))}{\Phi(\Sigma_k^{-1} y^*(I))} = 0 \\
\Sigma_k = \sqrt{\frac{\sigma_y^2 (k\sigma_\varepsilon^2 + \sigma_\Delta^2)}{k\sigma_\varepsilon^2 + \sigma_\Delta^2 + \sigma_y^{-2}\sigma_\varepsilon^2\sigma_\Delta^2}}.
\end{gathered} \tag{15}
$$

Note that in (15), $G(c^*) = G(\bar{c}) = 1$ for $I \leq \bar{I}$. Therefore, there is no uncertainty about $\hat{\sigma}^2$, and consequently, the expression boils down to:

$$
y^*(I) - m + \frac{\Sigma_I \phi(\Sigma_I^{-1} y^*(I))}{\Phi(\Sigma_I^{-1} y^*(I))} = 0 \qquad (I \leq \bar{I}). \tag{16}
$$

We can obtain the probability of voluntary disclosure given $y^*(I)$. Verrecchia (1990) shows that the probability of voluntary disclosure goes up with the precision of HQ's information. This result applies for $I \leq \bar{I}$, where all communicable DMs acquire information. In this region, the probability of voluntary disclosure increases with the fraction of communicable DMs. However, it is not the case for $I \geq \bar{I}$ because, as discussed before, HQ does not necessarily become more informed when she can communicate with more DMs. Therefore, the relationship between the probability of voluntary disclosure and the fraction of communicable DMs can be inverse U-shaped, as illustrated in Figure 5.

2.3.3 The role of HQ's private information

It is intuitive that information production by DMs is crucial for external communication to the extent that they contribute to HQ's information set. Put differently, in the extreme case that HQ's private information is highly precise, which corresponds to a low value of σ_ε^2, there should be less connection between internal communication and the precision of HQ's information set, and subsequently HQ's external disclosure. Figure 6 illustrates this point visually.

Our empirical analysis in Section 4.2.1 provides evidence consistent with this observation: The

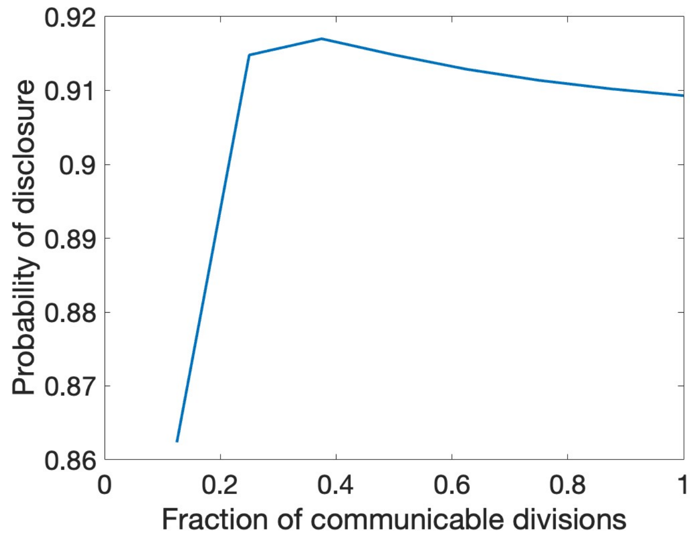


Figure 5: This figure displays the relationship between the probability of voluntary disclosure and the fraction of communicable DMs.

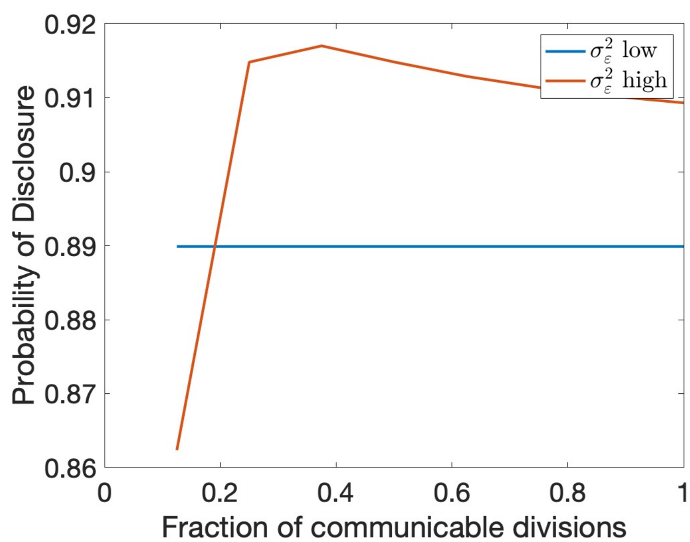


Figure 6: The relationship between the average variance of the information about y as a function of the fraction of communicable DMs for a large and a low value of σ_ε^2.

inverse U-shaped relationship between external and internal communication is more pronounced for firms where HQ is more dependent on communication with DMs in learning about fundamentals.

2.4 Pre-commitment and strategic reporting in internal communication

In this subsection, we demonstrate that the main results hold true under broader conditions, such as any form of pre-committed decision rule by HQ as well as DMs' strategic communication with HQ.

To this end, let us consider the following extension of the baseline model: Suppose HQ commits to the following decision rule based on the report she receives from communicable DMs, which does not need to be truthful, for some $b_1, b_2 \in [0,1]$:

$$a_i = \begin{cases} b_1\hat{y} + b_2 r_i & \text{If report } r_i \text{ is received from } DM_i \\ \hat{y} & \text{If no report is received from } DM_i, \end{cases} \tag{17}$$

where $\hat{y}$ is defined in Equation 4. Note that the baseline model corresponds to the case that $b_1 = 1 - \beta$, $b_2 = \beta$, and the DMs truthfully report, i.e., $r_i = y + \Delta_i$. Moreover, the case with $b_1 = 0$ and $b_2 = 1$ represents a decentralized structure of decision-making since HQ directly implements DMs' report for their division. Therefore, this extension covers a highly flexible set of possibilities to allocate control between HQ and DMs.

First, we show that communicable DMs scale their signals by some positive constant $\delta > 0$ when they report to HQ, i.e., $r_i = \delta(y + \Delta_i)$. This result ensures that HQ perfectly recovers the original signals (i.e., $y + \Delta_i$), even though DMs have an incentive to manipulate HQ's decision toward their desired action (Milgrom and Roberts, 1988).

Lemma 2. *Suppose k communicable DMs, including DM_i, acquire and report information to HQ. Then, DM_i reports $r_i = \delta(y + \Delta_i)$, where:*

$$\delta = b_2^{-1}\{1 - \frac{(\gamma_k^{HQ} + \frac{k-1}{k}\gamma_k^{DM})\sigma_y^2}{\sigma_y^2 + \sigma_\Delta^2} b_1 - \frac{\gamma_k^{DM}}{k} b_1\}. \tag{18}$$

The values of γ_k^{HQ} and γ_k^{DM} are provided in Equation 6.

To understand Equation 18, note that communicable DMs, conditional on reporting, report r_i so that the implemented action has the least expected squared deviation from their optimal action, $y + \Delta_i$, due to their quadratic preferences. As such, r_i is such that:

$$y + \Delta_i = \mathbb{E}[b_1\hat{y} + b_2 r_i | y + \Delta_i] \Rightarrow b_2 r_i = y + \Delta_i - b_1\mathbb{E}[\hat{y}|y + \Delta_i]. \tag{19}$$

From Equation 19, we learn that the value of b_2 has no impact on outcomes as DMs can fully counteract its effect by scaling their report. Moreover, the report function has no intercept since $\mathbb{E}[\hat{y}|y + \Delta_i]$ is proportional to $y + \Delta_i$. It reflects the fact that DMs cannot tell the sign of the bias between their information $(y + \Delta_i)$ and the aggregate state (y).

To verify Proposition 2, we only need to derive the expected net gain π_k for this more general framework, and show that it is decreasing in k. By so doing, we demonstrate that the free-riding mechanism does not depend on the assumptions regarding the allocation of control and DMs' truthful communication.

$$\begin{aligned} \pi_k &= -(-\mathbb{E}[(\hat{y}_{k-1} - y - \Delta_k)^2]) - \mathbb{E}[(b_1\hat{y}_k + (b_2\delta - 1)(y + \Delta_k))^2] \\ &= \hat{\sigma}^2_{k-1} + \sigma^2_\Delta - b_1^2 Var(\hat{y}|y + \Delta_i, k) \end{aligned} \tag{20}$$

Note that in the equation above, the first term on the first line represents the disutility from not acquiring information, which is equal to $\hat{\sigma}^2_{k-1} + \sigma^2_\Delta$, decreasing in k. The intuition is that as the number of reporting DMs increases, HQ obtains more precise information about y, which benefits all DMs. The second term in (20) reflects the utility of acquiring and reporting information. This term is also decreasing in k. To see the intuition, note that a DM impacts HQ's decision for his division (a_i) both directly through his report (r_i) and indirectly through affecting HQ's belief about y. As the number of reporting DMs (k) increases, each DM makes a smaller impact on HQ's belief about y, while their direct impact through their report is constant. Therefore, as more DMs report to HQ, the report of each DM becomes less impactful, and each DM loses less by not acquiring information. It implies that π_k is strictly decreasing in k. By exploiting this point, we can show that the probability of acquiring information goes down when the number of communicable DMs exceeds a threshold, as stated by Proposition 3.

Proposition 3. ***a)*** *π_k is decreasing in k for any value of b_1 and b_2. Moreover, π_k converges to $\underline{\pi} \equiv \sigma_\Delta^2 - \frac{b_1^2}{\sigma_y^{-2}+\sigma_\Delta^{-2}}$ from above.*

b) *Suppose $\pi_1 > \bar{c}$, and let $\bar{I}$ be the smallest positive integer for which $\pi_I < \bar{c}$.*[18] *Then, each communicable DM acquires information only if his information acquisition cost is less than some threshold c^*. c^* is the solution to the following equation:*

$$\sum_{k=1}^{I} \binom{I-1}{k-1} G(c^*)^{k-1}(1-G(c^*))^{I-k}(\pi_k - c^*) = 0. \tag{21}$$

c) *The probability of acquiring information, i.e., $G(c^*)$, is decreasing in the number of communicable DMs (I).*

2.5 Endogenizing internal communication

Thus far, we show that HQ does not necessarily learn more by communicating with more DMs because of their free-riding motive in their information acquisition. Note that in the absence of this strategic motive, it would be optimal for HQ to communicate with as many DMs as possible to have more precise information. In a sense, the optimal internal communication trades off the benefit from control centralization with the cost being less precise information about the aggregate state (y).

In this subsection, we analyze how the two economic forces change if we endogenize I. Similar to the baseline model, we assume that communicable divisions truthfully report to HQ if they acquire information, and HQ makes her optimal decision for all communicable divisions based on her information set. Moreover, for non-communicable divisions, the decision-making is delegated to their corresponding DMs. We discuss the optimal internal communication I chosen by HQ.[19]

According to Equation 2, HQ chooses I, the number of communicable DMs, to maximize the

[18] Such $\bar{I}$ exists since $\bar{c} > \sigma_\Delta^2$ by assumption.

[19] We use the term "optimal" narrowly by only considering information benefits. This need not be the optimal level when all other factors are considered and therefore is unlikely to be the level of internal communication chosen by the HQ manager that we see empirically.

following objective function:

$$\begin{aligned}
&\max_{I} \quad -\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[(a_i - y - \beta\Delta_i)^2] \\
\iff &\max_{I} \quad -\mathbb{E}[(I-k)(\hat{\sigma}_k^2 + \beta^2\sigma_\Delta^2) + k(1-\beta)^2\hat{\sigma}_k^2] \\
&-(N-I)\big\{(1-\beta)^2 G(\sigma_y^2+\sigma_\Delta^2)\sigma_\Delta^2 + (1-G(\sigma_y^2+\sigma_\Delta^2))(\sigma_y^2+\beta^2\sigma_\Delta^2)\big\}.
\end{aligned} \tag{22}$$

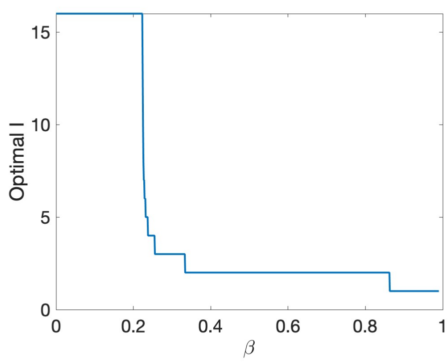
The first line of (22) is HQ's objective function, specified in Equation 2. The second line presents HQ's expected payoff from communicable divisions. HQ's expected payoff for a communicable division is $-(1-\beta)^2\hat{\sigma}_k^2$ if the corresponding DM acquires information, and is $-\hat{\sigma}_k^2 - \beta^2\sigma_\Delta^2$ otherwise. Note that the distribution of k is still determined endogenously, as discussed in Proposition 2. The third line in (22) presents HQ's expected payoff from non-communicable DMs. A non-communicable DM acquires information if the cost does not exceed $\sigma_y^2 + \sigma_\Delta^2$. Therefore, he acquires information with probability $G(\sigma_y^2 + \sigma_\Delta^2)$. HQ's expected payoff is $-(1-\beta)^2\sigma_\Delta^2$ if DM acquires information, and it is $-\sigma_y^2 + \beta^2\sigma_\Delta^2$ otherwise.

From Equation (22), we learn that perfect centralization ($I = N$) is optimal when

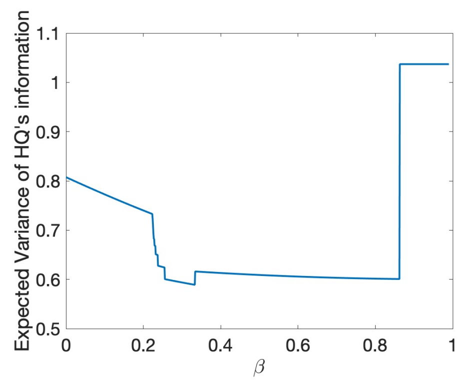
$$(1-\beta)^2\sigma_\Delta^2 > \hat{\sigma}_0^2 + \beta^2\sigma_\Delta^2. \tag{23}$$

The inequality above states that when the degree of misalignment in preferences is large (small β) and HQ's information is sufficiently precise ($Var(y|\hat{y}_0) = \hat{\sigma}_0^2$ is small), HQ strictly prefers centralization even if this results in less information production by DMs. Moreover, when β is large (low misalignment in internal preferences between HQ and DM) and $\bar{c} \leq \sigma_y^2 + \sigma_\Delta^2$ (that is, when the cost of information acquisition is low enough that non-communicable DMs always acquire information), then pure decentralization, i.e., $I = 0$, is optimal.

Figure 7a plots the relationship between the level of alignment in preferences between HQ and DMs (β) and the optimal level of internal communication (I). We see that the optimal I decreases with β, consistent with Aghion and Tirole (1997) and Dessein (2002) that find that delegation to DMs is more likely when internal agency friction is low.



(a) Optimal internal communication



(b) The expected variance of HQ's belief

Figure 7: The relationship between the level of alignment in preferences between HQ and DMs (β) and the optimal level of internal communication (I), and the expected variance of HQ's belief about the aggregate state (y) under the optimal internal communication.

Figure 7b displays the relationship between β and $\mathbb{E}[\hat{\sigma}_k^2]$, i.e., the expected variance of HQ's belief about the aggregate state. We see that the U-shaped relationship is preserved even when I is determined endogenously. This observation implies that the non-monotonic relationship between internal and external communication is present even when we endogenize internal communication. An extensive theoretical analysis of the determinants of optimal internal communication entails characterizing the optimal I in Equation (22), which adds complexity to the model but may not provide additional insights beyond what is discussed.

3 Data and Sample Selection

The data on firms' use of the intranet are available from the Harte-Hanks Ci Technology database (CiDB). The CiDB provides yearly site-level data on information and communications technology (ICT). The data provider Harte-Hanks is a multi-national firm that collects detailed information on hardware and software (e.g., computers, IT budget and staff, servers, and computing capacity) installed in millions of establishments worldwide. The CiDB contains ICT records of over 12 million public and private firms and nearly 15 million sites across 46 countries. Harte-Hanks sells their data to large ICT firms such as IBS, Cisco, and Dell.[20] To ensure a high level of accuracy, Harte-Hanks (re-)collects data every year and performs extensive quality checks. Due to their broad

[20]The name of the data provider for CiDB changed from Harte-Hanks to Aberdeen, and then to Spiceworks Ziff Davis (SWZD) because Aberdeen was divested from Harte-Hanks in 2015 and acquired by SWZD in 2020. We purchased the data from SWZD but refer to the data provider as Harte-Hanks to be consistent with the literature.

coverage and high quality, the data have been widely used by academics to measure firms' software or hardware technology.[21] We use their data from 2001, the first year in which data were collected on intranet-related technology. Data after 2015 are not used because their survey format changed significantly, and there is no substantial variation in firms' use of the intranet post-2015. Therefore, our sample period spans 2001–2015.

The intranet does not allow access to anyone outside its private network and creates a platform where employees can securely store and share proprietary data. Therefore, the use of the intranet facilitates intra-firm communication and information sharing (e.g., Bloom et al., 2014). Panel A of Appendix C illustrates the differences among the intranet, extranet, and internet. Following Bloom et al. (2014), we use the proportion of divisions within a firm-year that adopt the intranet technology, weighted by the number of employees in each division, as an empirical proxy for the extent of the firm's internal communication. We determine whether a site uses the intranet based on whether it has installed the technologies used to connect different sites—namely frame relays, leased lines, a synchronous optical network (SONET), digital subscriber line (xDSL), asynchronous transfer mode (ATM) technology, and virtual private networks (VPN).[22]

Panel B of Appendix C illustrates the role of the intranet in facilitating communication between divisional offices and the headquarters office. The figure shows a hypothetical case where two of the three divisional offices are connected to the headquarters office through the intranet, in which case our measure of the intranet takes the value of 2/3, assuming that each divisional office has the same number of employees. Other data from the CiDB are used as control variables, such as firms' use of the internet and PCs, as well as their employee and site counts, because they can be correlated with firms' information access, which could affect their external disclosure decisions. We aggregate the site-level CiDB data to the firm level because external communication is defined at the firm level.

We consider two forms of external communication: management EPS forecasts and website

[21] e.g., Bresnahan et al. (2002), Brynjolfsson and Hitt (2003), Beaudry et al. (2010), Forman et al. (2012), Bloom et al. (2016), Bloom et al. (2012), Bloom et al. (2014), and Charoenwong et al. (2022).

[22] These telecommunication technologies or equipment connect local area networks (LANs) or private telephone systems and allow transmission of data between them. Some of these technologies may be used to operate the internet. Therefore, we control for firms' use of the internet. The Harte-Hanks has information on site–level internet uses because their survey directly asks about internet access at each site.

disclosures. Management forecasts have been extensively shown to reduce information asymmetry between corporate insiders and outsiders (e.g., Skinner, 1994; Berger, 2011; Kasznik and Lev, 1995; Balakrishnan et al., 2014; Frankel et al., 1995; Guay et al., 2016; Noh et al., 2019; Seo, 2021; Lu and Skinner, 2020; Rickmann, 2022).[23] Our second measure of external communication is motivated by recent papers including Boulland et al. (2021) and Lynch and Taylor (2022), which have documented that corporate websites provide value-relevant information to investors and financial intermediaries. These papers show that the information on corporate websites is incremental to established measures of voluntary disclosure, such as 10-K, 8-K, and management forecasts. One notable benefit of this alternative disclosure proxy is that it can be measured for both public and private firms.

To construct our corporate website-based disclosure proxy, we obtain complete history records of company websites from Wayback Machine, a digital archive of websites. As discussed in Boulland et al. (2021), the Wayback Machine provides user-friendly access to archive data via an Application Programming Interface (API) that enables the retrieval of a URL's historical structure. We obtain the URLs of public companies' websites first from Compustat and then from the CiDB data. We obtain URLs of private companies' websites from the CiDB. We supplement missing URLs by conducting a Google keyword search of a company name and checking the first three search results with Python.[24]

For each firm-year, we collect all website records between January 1 and December 31 and calculate the average length of the homepage for all records, omitting those with 404 or "not found" failures. The length of the homepage refers to the size of the entire homepage in bytes. Given a large number of public and private firms in our sample, we measure the length of the homepage to proxy for the total content of all pages of each corporate website used in Boulland et al. (2021) and Lynch and Taylor (2022). To validate this empirical choice, we collect the content of all website pages—the homepage and interior pages—for 150 randomly selected firms and find a correlation coefficient of 0.91.

To construct our public firm sample, we merge the CiDB data with the intersection of Com-

[23]For example, Beyer et al. (2010) document that management forecasts explain the most variation of quarterly stock returns among earnings (pre-)announcements, management forecasts, analysts forecasts, and SEC filings.

[24]We do not consider URLs from *LinkedIn*, *Twitter*, and *Bloomberg*.

pustat and CRSP. The CiDB provides data not only on firms' use of the intranet but also on their employees, revenues, number of sites, industry classification, among other information. We further merge it with managements' and analysts' forecast data obtained from I/B/E/S, institutional ownership data obtained from Thomson Reuters, and corporate website disclosure data obtained from Wayback. Our private firm data primarily come from the CiDB. We merge the private firm data from the CiDB and website disclosure data from Wayback to construct our private firm sample. From both our public and private firm samples, we remove firms that do not have divisional offices and whose headquarters locations are zero distance from all sites, because the effect of the intranet on facilitating intra-firm communication is likely to be small for such firms. Firms in the utility and financial industries are also removed as they are highly regulated. Throughout our analyses, we use management EPS forecasts and corporate website length as proxies for public firms' external communication, and corporate website length as a proxy for private firms' external communication.

Table 1 presents descriptive statistics for the key variables used in our empirical tests. The main sample is comprised of 2,946 unique public firms and 60,501 unique private firms during the sample period of 2001 to 2015. In total, we have 20,778 public firm-year observations for EPS forecast analysis, 17,600 public firm-year observations for website length analysis, and 225,425 private firm-year observations for website length analysis. For public firms, the average value of *Intranet* is 0.49. For private firms, the average value of *Intranet* is 0.42. This descriptive is consistent with public firms on average having better internal communication systems. We show that our public firm sample has, on average, noticeably greater *Revenue*, *Employees*, and *Sites* than our private firm sample. We find that public firms, on average, provide 2.71 EPS forecasts per year. We also find that the average values of public and private firms' website lengths are similar. Particularly, the average value of the website length for public firms is 5,409 bytes, and that for private firms is 5,969 bytes.

4 Main Results

4.1 Association between the Intranet and EPS Forecasts: Public Firms

We begin by graphically presenting the shape of the relation between internal communication technology and external communication. Following Samuels et al. (2021) and Kim et al. (2021), we sort firms into quintiles based on $Intranet$, our proxy for internal communication technology, and plot average values of public firms' EPS forecasts for each quintile in Panel A of Figure 8. We find that public firms' EPS forecasts and website length first increase as the quintiles of intranet increase and then subsequently decline, thereby showing inverse U-shapes. This visualizes the nature of our findings well and provide a strong basis for our regression analysis.[25]

We then conduct regression analyses to examine the association between improvements in internal communication technology and the extent of external communication. Specifically, we investigate how improvements in internal communication technology, proxied by firms' use of the intranet, vary with voluntary disclosure, proxied by the frequency of EPS forecasts for public firms. We estimate the following regression, including both linear and second-order polynomial terms of $Intranet$:

$$\begin{aligned} Voluntary\ Disclosure_{i,t+1} =& \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} \\ &+ \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}, \end{aligned} \tag{24}$$

where the key independent variable $Intranet_{i,t}$ is our measure of intra-firm intranet intensity. It captures the fraction of a firm's divisions adopting the intranet, calculated as the weighted average site-level intranet with sites' employee counts used as the weights. Each site's intranet takes the value of 1 if it has installed the intranet, and 0 otherwise. The dependent variable is the frequency of management EPS forecasts measured between the earnings announcement corresponding to fiscal year t and the earnings announcement for the subsequent fiscal year $t + 1$.[26] Because the CiDB provides data on firms' use of the intranet by calendar year, we measure $Intranet_{i,t}$ as of the last calendar year ending before fiscal year t. Therefore, the measurement windows for voluntary disclosure always come after that for the intranet. $Controls_{i,t}$ is a vector of control variables measured for each firm i's fiscal year t, and include characteristics that could be associated with firms' vol-

[25]In our Online Appendix, we show that linear regressions of voluntary disclosure on $Intranet$ suggest an insignificant or negative relation. This result underscores the importance of using a formal model in making empirical predictions.

[26]Our regression design choices follow the prior literature on management forecasts (e.g., Noh et al., 2019; Glaeser, 2018). All of our analyses produce qualitatively and quantitatively similar results if we use a non-log-transformed EPS forecast count, as opposed to the log of 1 plus an EPS forecast count, following Guay et al. (2016).

untary disclosure decisions (e.g., Chen et al., 2018). These control variables include ROA, BTM, $ln(MVE)$, $R\&D$, $Loss$, $ln(1+analysts\ following)$, $Earnings\ Volatility$, $ln(business\ segments)$, $ln(geographical\ segments)$, $Related$, and $Institutional\ ownership$ obtained from Compustat, CRSP, I/B/E/S, and Thomson Reuters as well as $ln(employees)$, $ln(1+revenues)$, $ln(sites)$, $Internet$, and $Number\ of\ PCs\ per\ employee$ obtained from the CiDB.

We use industry × year fixed effects to account for time-varying industry characteristics. Following Samuels et al. (2021) and Kim et al. (2021), we do not include firm fixed effects because non-linearity is better captured in the large cross-section of firms and can hardly be captured in limited within-firm variation. This choice of empirical specification is also consistent with our theory, which compares the degree of internal communication across firms as opposed to comparing time-series changes within a firm. If the shape of the relation between internal communication technology and external communication is unimodal, as our theory predicts, we expect to find that β_2 in Equation 24 (i.e., the coefficient on the second-order polynomial term of $Intranet$) is statistically significant and negative.

Table 2 presents the results from the regression estimated for public firms using their EPS forecasts as a proxy for voluntary disclosure. We find that β_2 is significantly negative, suggesting that $Intranet$ and $ln(1+EPS\ forecasts)$ have an inverse U-shaped relation, which is consistent with the two economic forces — "information learning" and "free riding" — having countervailing effects, which our theory predicts produce an inverse U-relation between internal and external communications.

4.2 Evidence of the Economic Forces

4.2.1 Information learning channel

In this section, we provide empirical evidence of the two economic forces we illustrate in our theory. To corroborate the information learning channel, we explore how the link between the intranet and voluntary disclosure varies with the HQ manager's ex ante information about the firm. As we predicted in our theory, when the HQ manager already knows well about their divisions, internal communication does not facilitate her learning. Furthermore, it does not change divisional managers' free-riding incentives, which change with the HQ manager's informedness, thereby making the relation between internal communication technology and disclosure flat.

We empirically measure the HQ manager's prior knowledge (i.e., the inverse of σ_y^2) using the average distance between the HQ office and divisional offices. An assumption underlying the distance measure is that managers of firms that are geographically nearby have more comprehensive information about the firm's operations in the absence of internal communication. The marginal effect of improved internal communication technology on HQ managers' information set is likely less in firms that are less geographically dispersed.[27]

To examine the role of HQ managers' prior information, we estimate the main regression Equation 24 separately for firms with above-median and below-median distance to divisional offices within the two-digit SIC industry. The distance is measured as the average distance between the firm's HQ office and divisional offices weighted by employees at each divisional site. In these tests, we lose a small percentage of our observations due to increases in singletons within each fixed-effect group after we divide the sample into two groups.

Table 3 presents results for public firms when using EPS forecasts as a proxy for voluntary disclosure. We find that the coefficient on $Intranet^2$ is statistically significant and negative only for the subsample that has above-median distance between the headquarters and divisional offices but not for the subsample with below-median distance. In Table 3, the coefficients on $Intranet^2$ in column (1) and column (2) are significantly different with the p-value <0.05.[28] These contrasting results for firms with below- and above-median distances are consistent with Figure 6 discussed in Section 2.3.3.

Altogether, we find that the inverse U-shape relation between internal and external communications is concentrated (muted) in firms where the information learning force is pronounced (weakened). Our results are consistent with HQ managers' information learning channel being one force driving the inverse U-shape relation.

[27] In our Online Appendix, we do not find a significant association between the average distance and firms' intranet adoption.

[28] In our Online Appendix, we show these cross-sectional results are robust to controlling for potential nonlinear relationships between voluntary disclosure and our control variables.

4.2.2 Free riding channel

As discussed in Section 2, the increased free-riding incentive by the divisional manager leading to the HQ manager's worse information set, together with her information learning, shapes the unimodal relation between internal communication technology and external communication. Based on our theoretical framework, when the common component in the information structure is more important across all communicable divisions than the division-specific component (i.e., $\frac{\sigma_y^2}{\sigma_\Delta^2}$ is higher), the influence of the other divisional managers' acquired information on the HQ manager's information set is larger. Therefore, the free-riding incentive of divisional managers becomes more prominent.

To substantiate the free-riding channel, we empirically examine whether the negative relation between internal communication technology and voluntary disclosure is absent (present) in the subsample of firms with greater (smaller) similarity among their divisional sites that have internal communication technology. The idea is that, when there is greater similarity across divisional sites, a divisional manager's information acquisition helps the HQ manager learn about other divisions, which reduces other divisional managers' incentives to incur costs to acquire information.

We assess the similarity among a firm's sites based on whether they are classified in the same two-digit industry code. Specifically, we use the sum of the square of the share of each 2-digit SIC industry that the firm's sites with intranet technology operate in, denoted as *Site Similarity*.[29] We then estimate the main regression Equation 24 separately for firms with above- and below-median *Site Similarity* measured for the prior year (i.e., we measure *Site Similarity* in the year before we measure *Intranet*). The number of observations included in these tests is smaller because we drop first firm-year observations, for which we cannot measure *Site Similarity* for the prior year, and drop firm-years with no sites that have the intranet.

Table 4 shows the results for these cross-sectional tests. Columns (1) and (3) show the results of regressions excluding the quadratic term of *Intranet* estimated for below- and above-median

[29]For example, if a firm has two sites with the intranet operating in one 2-digit SIC industry and three sites with the intranet operating in another 2-digit SIC industry, then this firm's *Site Similarity* is $(2/5)^2 + (3/5)^2 = 0.52$. This measure is analogous to the Herfindahl-Hirschman Index (HHI) and takes a higher value if a firm has sites in many similar industries. As an alternative measure, we use the maximal similarity ratio that takes the maximal value of the proportion of similar-industry divisions by the total number of communicable divisions, and find similar results.

Site Similarity, respectively. The subsample of firms with smaller similarity across divisional sites show a significantly positive linear relation between *Intranet* and EPS forecasts, whereas firms with greater site-level similarity do not show a significant linear relation. The coefficients on *Intranet* in columns (1) and (3) are statistically different with p-value $<$0.01. Columns (2) and (4) show the results including the quadratic term of *Intranet*. Firms with below-median *Site Similarity* do not reveal a statistically significant inverse U-relation between *Intranet* and EPS forecasts in Column (2), but those with above-median *Site Similarity* show a significant inverse U-shape in Column (4), as predicted. These results together suggest that the free-riding force contributes to the unimodal relation between internal communication technology and external communication.[30]

4.3 Association between the Intranet and Website Disclosure: Public and Private Firms

An important distinguishing feature of internal agency frictions is that they are pervasive not just in publicly listed firms but also in private firms. Arguably, the setting of private firms may be less subject to external agency frictions (e.g., capital market incentives), and internal agency frictions are likely more pronounced. We, therefore, examine whether the unimodal relation holds using website disclosures as a proxy for external communication, which are available for both public and private firms.

Panel B and Panel C of Figure 8 plot average website disclosure length for each quintile of *Intranet* for public and private firms, respectively. Similar to our findings for public firms' EPS forecasts in Panel A of Figure 8, we find that website length first increases as the quintiles of intranet increase and then subsequently decline. To further document this unimodal relation in a regression, we estimate the regression Equation 24 using $ln(Website\ length)_{i,t+1}$ as the dependent variable defined as the size of the firm's corporate website homepage. For public firms, we measure $ln(Website\ length)_{i,t+1}$ during the first calendar year after fiscal year t. For private firms, we measure $ln(Website\ size)_{i,t+1}$ for calendar year $t+1$. Because the CiDB provides data on firms' use of the intranet by calendar year, we measure $Intranet_{i,t}$ as of the last calendar year ending before fiscal year t for public firms, and for calendaer year t for private firms. Therefore, the

[30] In our Online Appendix, we show these cross-sectional results are robust to controlling for potential nonlinear relationships between voluntary disclosure and our control variables.

measurement window for website disclosure always precedes that for the intranet. We use the same control variables as Table 2 for our public firm analysis and use a limited set of control variables for private firm analysis due to limited data availability. For private firm analysis, we include as controls $ln(employees)$, $ln(1+revenues)$, $ln(sites)$, $Internet$, and $Number\ of\ PCs\ per\ employee$ available from the CiDB.

The results of the regressions are shown in Table 5. In Panel A and Panel B, we continue to find an inverse U-relation using website length as an alternative proxy for public and private firms' external communication, as manifested by the negative and significant coefficient on $Intranet^2$. Collectively, we empirically document robust inverse U-shape relations between internal communication facilitated by the use of the intranet and the extent of external communication proxied by the amount of voluntary disclosure. These results hold for both public and private firms and are consistent with our theoretical predictions (e.g., see Figure 5). Our results for private firms, in particular, reinforce the idea that the inverse U-shape is driven mainly by internal agency frictions rather than by external agency frictions.

4.4 Robustness of the Inverse U-Shape and Endogeneity Discussions

We have documented that the theoretical predictions of the inverse U-shape hold across different settings and different measures of voluntary disclosure, including management forecasts (only for public firms) and corporate websites (for public and private firms). In this section, we test the robustness of the non-linear, inverse U-shape.

We re-estimate the regression Equation 24 after including the quadratic terms of all control variables to account for ad hoc nonlinear relationships between voluntary disclosure and our control variables. We do so for both our public and private firm samples, using EPS forecasts and website length as two proxies for public firms' external communication (Panel A and Panel B) and website length as a proxy for private firms' external communication (Panel C). Consistent with our main results, across Panel A-C of Table 6, we continue to find a statistically significant and negative coefficient on the second-order polynomial term for intranet, $Intranet^2$. We note that the coefficient on $Intranet^2$ estimated after including the quadratic terms of control variables is generally similar in magnitude and statistical significance to the coefficient on $Intranet^2$ estimated without those.

For instance, the coefficient on $Intranet^2$ estimated using public firms' EPS forecasts presented in Table 2 is -0.307 (t-statistic of -2.30) and the corresponding coefficient in Table 6 Panel A is -0.319 (t-statistic of -2.37).

While all of the tests are guided by a formal theory and we use the cross-sectional variation in the study, there might be correlated omitted variable concerns. For instance, firms with more communicable divisions may also have more complex organizational structures, which may result in a poor internal information environment and voluntary disclosures. We account for organizational structure-related variables in the regression. Additionally, the measurement we use for the free-riding channel can help mitigate this concern. Insofar as more different divisions may reflect organizational complexity, the alternative interpretation may not adequately explain the result of a less pronounced negative slope when the similarity across communicable divisions is low.

To further alleviate some endogeneity concerns, we use the "Bartik Instrument" to determine the extent to which these endogeneity concerns may influence our results (e.g., Breuer, 2022). To this end, we decompose the intranet adoption rate variable into two components: the time-varying intranet adoption rate at the national industry level (excluding the focal firm) and the predetermined industry share at the divisional level for a given firm. The "Bartik instrument" uses the differential effect of national trends on firms with a predetermined industry share at the company level, which purges the treatment of some endogenous variation. We use this "Bartik instrument" and re-run our analysis. In untabulated tests, we find similar results, which provides us with some confidence that endogeneity is not the primary driver of our results.

Overall, we find robust evidence of the inverse U-shape relation between improvements in internal communication technology and voluntary disclosure. These results suggest that the inverse U-shapes we document are not attributable to non-linear relations between voluntary disclosure and our control variables.

5 Conclusion

We theoretically and empirically investigate how improvements in organizations' internal communication technology affect their voluntary disclosure, a form of external communication. Our theoretical model endogenizes information acquisition in a setting with a headquarters manager and multiple divisional managers. We formalize two economic forces: information learning and free riding. We find that these two economic forces jointly determine an inverse U-shape relation between internal and external communications.

Unlike prior studies that center on external agency frictions, our work focuses on internal agency frictions. We find that the theoretical predictions of the inverse U-shape hold true across a variety of empirical specifications and settings, including public and private companies. In doing so, we add novel insights and findings to the extant literature studying factors that affect firms' voluntary disclosure decisions.

Our paper is subject to several caveats. First, our baseline model assumes that the number of divisions with which HQ can communicate is determined by many other factors than the divisional managers' incentive to produce information. Second, we do not have an exogenous shock for our empirical analysis, and we use cross-sectional variation among firms with different intranet usages, although we implement the "Bartik instrument"-style analysis and find similar results. Our empirical identification heavily relies on the inverse U-shape.

While the use of the intranet has become somewhat prevalent, innovations and new technologies will continue to facilitate communication and information sharing. To the extent that such progress improves managers' gathering and processing of organizational knowledge, our empirical findings extend beyond the intranet and speak to the potential implications of other modern information technologies.

References

Acemoglu, D., Aghion, P., Lelarge, C., Van Reenen, J., and Zilibotti, F. (2007). Technology, information, and the decentralization of the firm. The Quarterly Journal of Economics, 122(4):1759–1799.

Aghion, P. and Tirole, J. (1997). Formal and real authority in organizations. Journal of Political Economy, 105(1):1–29.

Balakrishnan, K., Billings, M. B., Kelly, B., and Ljungqvist, A. (2014). Shaping liquidity: On the causal effects of voluntary disclosure. The Journal of Finance, 69(5):2237–2278.

Beaudry, P., Doms, M., and Lewis, E. (2010). Should the personal computer be considered a technological revolution? evidence from u.s. metropolitan areas. Journal of Political Economy, 118(5):988–1036.

Berger, P. G. (2011). Challenges and opportunities in disclosure research—a discussion of 'the financial reporting environment: Review of the recent literature'. Journal of Accounting and Economics, 51(1-2):204–218.

Berger, P. G., Li, F., Liu, L. Y., and Wong, M. (2022). The role of managerial reporting quality in investment efficiency. Available at SSRN 4286813.

Bertomeu, J., Beyer, A., Taylor, D. J., et al. (2016). From casual to causal inference in accounting research: The need for theoretical foundations. Foundations and Trends in Accounting, 10(2-4):262–313.

Beyer, A., Cohen, D. A., Lys, T. Z., and Walther, B. R. (2010). The financial reporting environment: Review of the recent literature. Journal of Accounting and Economics, 50(2):296–343.

Bloom, N., Draca, M., and Reenen, J. V. (2016). Trade induced technical change? the impact of chinese imports on innovation, it and productivity. The Review of Economic Studies, 83(1):87–117.

Bloom, N., Garicano, L., Sadun, R., and Van Reenen, J. (2014). The distinct effects of information technology and communication technology on firm organization. Management Science, 60(12):2859–2885.

Bloom, N., Sadun, R., and Van Reenen, J. (2012). Americans do it better: Us multinationals and the productivity miracle. American Economic Review, 102(1):167–201.

Boone, A. L. and White, J. (2015). The effect of institutional ownership on firm transparency and information production. Journal of Financial Economics, 117(3):508–533.

Boulland, R., Bourveau, T., and Breuer, M. (2021). Corporate websites: A new measure of voluntary disclosure. Available at SSRN 3816623.

Bourveau, T., Kepler, J. D., She, G., and Wang, L. L. (2022). Firm boundaries and voluntary disclosure. Available at SSRN 4019607.

Bresnahan, T. F., Brynjolfsson, E., and Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. The Quarterly Journal of Economics, 117(1):339–376.

Breuer, M. (2022). Bartik instruments: An applied introduction. Journal of Financial Reporting, 7(1):49–67.

Breuer, M., Hombach, K., and Müller, M. A. (2020). The economics of firms' public disclosure: Theory and evidence. Available at SSRN 3037002.

Brynjolfsson, E. and Hitt, L. M. (2003). Computing productivity: Firm-level evidence. Review of Economics and Statistics, 85:793–808.

Charoenwong, B., Kowaleski, Z. T., Kwan, A., and Sutherland, A. (2022). Regtech. Available at SSRN 4000016.

Chen, C., Martin, X., Roychowdhury, S., Wang, X., and Billett, M. T. (2018). Clarity begins at home: Internal information asymmetry and external communication quality. The Accounting Review, 93(1):71–101.

Chen, Q., Gerakos, J., Glode, V., and Taylor, D. J. (2016). Thoughts on the divide between theoretical and empirical research in accounting. Journal of Financial Reporting, 1(2):47–58.

Dedman, E. and Lennox, C. (2009). Perceived competition, profitability and the withholding of information about sales and the cost of sales. Journal of Accounting and Economics, 48(2-3):210–230.

Dessein, W. (2002). Authority and communication in organizations. The Review of Economic Studies, 69(4):811–838.

Dichev, I. D., Graham, J. R., Harvey, C. R., and Rajgopal, S. (2013). Earnings quality: Evidence from the field. Journal of Accounting and Economics, 56(2, Supplement 1):1–33. Conference Issue on Accounting Research on Classic and Contemporary Issues.

Dye, R. A. (1985). Disclosure of nonproprietary information. Journal of Accounting Research, pages 123–145.

Fischer, P. E. and Verrecchia, R. E. (2000). Reporting bias. The Accounting Review, 75(2):229–245.

Forman, C., Goldfarb, A., and Greenstein, S. (2012). The internet and local wages: A puzzle. The American Economic Review, 102(1):556–575.

Frankel, R., McNichols, M., and Wilson, G. P. (1995). Discretionary disclosure and external financing. The Accounting Review, 70(1):135–150.

Gallemore, J. and Labro, E. (2015). The importance of the internal information environment for tax avoidance. Journal of Accounting and Economics, 60(1):149–167.

Gassen, J. and Muhn, M. (2018). Financial transparency of private firms: Evidence from a randomized field experiment. Available at SSRN 3290710.

Glaeser, S. (2018). The effects of proprietary information on corporate disclosure and transparency: Evidence from trade secrets. Journal of Accounting and Economics, 66(1):163–193.

Guay, W., Samuels, D., and Taylor, D. (2016). Guiding through the fog: Financial statement complexity and voluntary disclosure. Journal of Accounting and Economics, 62(2-3):234–269.

Harris, M. and Raviv, A. (2005). Allocation of decision-making authority. Review of Finance, 9(3):353–383.

Healy, P. M. and Palepu, K. G. (2001). 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.

Heinle, M., Samuels, D., and Taylo, D. (2022). Disclosure substitution. Management Science.

Hemmer, T. and Labro, E. (2008). On the optimal relation between the properties of managerial and financial reporting systems. Journal of Accounting Research, 46(5):1209–1240.

Ittner, C. D. and Larcker, D. F. (2001). Assessing empirical research in managerial accounting: a value-based management perspective. Journal of Accounting and Economics, 32(1-3):349–410.

Ittner, C. D. and Michels, J. (2017). Risk-based forecasting and planning and management earnings forecasts. Review of Accounting Studies, 22(3):1005–1047.

Jung, W.-O. and Kwon, Y. K. (1988). Disclosure when the market is unsure of information endowment of managers. Journal of Accounting research, pages 146–153.

Kaplan, R. S. (1984). The evolution of management accounting. The Accounting Review, 59(3):390–418.

Kasznik, R. and Lev, B. (1995). To warn or not to warn: Management disclosures in the face of an earnings surprise. The Accounting Review, 70:113–134.

Kim, J. M., Taylor, D. J., and Verrecchia, R. E. (2021). Voluntary disclosure when private information and disclosure costs are jointly determined. Review of Accounting Studies, 26(3):971–1001.

Lewbel, A. (2019). The identification zoo: Meanings of identification in econometrics. Journal of Economic Literature, 57(4):835–903.

Lu, Y. and Skinner, D. J. (2020). Moving forward: Management guidance and earnings announcement returns. Available at SSRN 3687764.

Lynch, B. and Taylor, D. (2022). The information content of corporate websites. Available at SSRN 3791474.

Milgrom, P. and Roberts, J. (1988). An economic approach to influence activities in organizations. American Journal of sociology, 94:S154–S179.

Minnis, M. and Shroff, N. (2017). Why regulate private firm disclosure and auditing? Accounting and Business Research, 47(5):473–502.

Mookherjee, D. (2006). Decentralization, hierarchies, and incentives: A mechanism design perspective. Journal of Economic Literature, 44(2):367–390.

Noh, S., So, E. C., and Weber, J. P. (2019). Voluntary and mandatory disclosures: Do managers view them as substitutes? Journal of Accounting and Economics, 68(1):101243.

Richarson, S. (2001). Discretionary disclosure: A note. Abacus, 37(2):233–247.

Rickmann, G. A. (2022). The effect of market transparency on corporate disclosure: Evidence from the observability of bond prices and trading. The Accounting Review, 97(4):371–397.

Samuels, D. (2021). Government procurement and changes in firm transparency. The Accounting Review, 96(1):401–430.

Samuels, D., Taylor, D. J., and Verrecchia, R. E. (2021). The economics of misreporting and the role of public scrutiny. Journal of Accounting and Economics, 71(1):101340.

Scharfstein, D. S. and Stein, J. C. (2000). The dark side of internal capital markets: Divisional rent-seeking and inefficient investment. The Journal of Finance, 55(6):2537–2564.

Seo, H. (2021). Peer effects in corporate disclosure decisions. Journal of Accounting and Economics, 71(1):101364.

Skinner, D. J. (1994). Why firms voluntarily disclose bad news. Journal of Accounting Research, 32(1):38–60.

Verrecchia, R. E. (1983). Discretionary disclosure. Journal of Accounting and Economics, 5:179–194.

Verrecchia, R. E. (1990). Information quality and discretionary disclosure. Journal of accounting and Economics, 12(4):365–380.

Verrecchia, R. E. (2001). Essays on disclosure. Journal of Accounting and Economics, 32(1-3):97–180.

Zimmerman, J. L. (2009). Accounting for decision making and control, volume Seventh Edition. Irwin/McGraw Hill Publishing Company.

Appendix A. Proofs

A.1 Derivation of Equations 4 and 7

Due to the symmetry among Δ_i's, it is intuitive that all signals obtained from DMs should receive the same weight. As a result, define the vector of observables for HQ as $Y = (\hat{y}_0, \bar{y})'$. Let $\gamma = (\gamma_k^{HQ}, \gamma_k^{DM})'$ be the vector of coefficients so that $\gamma' Y$ is the best linear predictor of y. In fact, γ is the solution to the following optimization problem:

$$\max_{\tilde{\gamma}} -\frac{1}{2}\mathbb{E}[(y - \tilde{\gamma}' Y)^2]. \tag{25}$$

The first-order condition implies that the solution to (25) is:

$$\begin{aligned}\gamma &= \underbrace{\begin{bmatrix} \sigma_y^2 + \sigma_\varepsilon^2 & \sigma_y^2 \\ \sigma_y^2 & \sigma_y^2 + \frac{\sigma_\Delta^2}{k} \end{bmatrix}^{-1}}_{\mathbb{E}[YY']^{-1}} \underbrace{\begin{bmatrix} \sigma_y^2 \\ \sigma_y^2 \end{bmatrix}}_{\mathbb{E}[Yy]} = \frac{\sigma_y^2}{(\sigma_y^2 + \sigma_\varepsilon^2)(\sigma_y^2 + \frac{\sigma_\Delta^2}{k}) - \sigma_y^4} \begin{bmatrix} \frac{\sigma_\Delta^2}{k} \\ \sigma_\varepsilon^2 \end{bmatrix} \\ &= \frac{1}{\sigma_\varepsilon^2 + \frac{\sigma_\Delta^2}{k} + \sigma_y^{-2}\sigma_\varepsilon^2\frac{\sigma_\Delta^2}{k}} \begin{bmatrix} \frac{\sigma_\Delta^2}{k} \\ \sigma_\varepsilon^2 \end{bmatrix}\end{aligned} \tag{26}$$

Moreover, the conditional variance can be obtained as follows:

$$\hat{\sigma}^2 = \mathbb{E}[(y - \gamma' Y)^2] = \sigma_y^2 - 2\gamma' \mathbb{E}[Yy] + \gamma' \mathbb{E}[YY']\gamma. \tag{27}$$

As shown in (26), $\mathbb{E}[YY']\gamma = \mathbb{E}[Yy]$. Thus, we can rewrite the expression above as follows:

$$\hat{\sigma}^2 = \sigma_y^2 - \gamma' \mathbb{E}[Yy] = \sigma_y^2 \{1 - \frac{\sigma_\varepsilon^2 + \frac{\sigma_\Delta^2}{k}}{\sigma_\varepsilon^2 + \frac{\sigma_\Delta^2}{k} + \sigma_y^{-2}\sigma_\varepsilon^2\frac{\sigma_\Delta^2}{k}}\} = \frac{\sigma_\varepsilon^2 \frac{\sigma_\Delta^2}{k}}{\sigma_\varepsilon^2 + \frac{\sigma_\Delta^2}{k} + \sigma_y^{-2}\sigma_\varepsilon^2\frac{\sigma_\Delta^2}{k}}. \tag{28}$$

A.2 Derivation of Equation 8

Note that HQ sets action a_i to maximize $-\mathbb{E}^{HQ}[(a_i - y - \beta\Delta_i)^2]$ for communicable divisions. The optimal action is given by:

$$\mathbb{E}^{HQ}[y + \beta\Delta_i] = \mathbb{E}^{HQ}[y] + \beta\mathbb{E}^{HQ}[\Delta_i] = \hat{y} + \beta\mathbb{E}^{HQ}[\Delta_i]. \tag{29}$$

For communicable DMs that acquire information, we have:

$$\mathbb{E}^{HQ}[\Delta_i] = \mathbb{E}[(y + \Delta_i) - y|\hat{y}_0, \{y + \Delta_j\}_{j \text{ is communicable}}] = y + \Delta_i - \hat{y}. \tag{30}$$

For other DMs, $\mathbb{E}^{HQ}[\Delta_i] = 0$ as no information is available. By substituting these values in $a_i^{HQ*} = \hat{y} + \beta\mathbb{E}^{HQ}[\Delta_i]$, we can verify Equation 8.

A.3 Proof of Lemma 1

π_k is the expected net gain from information acquisition for a communicable DM when $k-1$ other communicable DMs acquire information. Without loss, assume that $DM_1, \ldots, DM_k$ are communicable. Thus, π_k is the expected gain DM_k gets from information acquisition given the first $k-1$ DMs acquire information. As there is no delegation to communicable DMs, we have:

$$\pi_k = -\mathbb{E}[(a_i(\text{acq. info}) - y - \Delta_k)^2] + \mathbb{E}[(a_i(\text{no info}) - y - \Delta_k)^2], \tag{31}$$

where (According to Equation 8):

$$\begin{aligned} a_i(\text{no info}) &= \mathbb{E}[y|\hat{y}_0, y + \Delta_1, \ldots, y + \Delta_{k-1}] \\ a_i(\text{acq. info}) &= (1-\beta)\mathbb{E}[y|\hat{y}_0, y + \Delta_1, \ldots, y + \Delta_k] + \beta(y + \Delta_k). \end{aligned} \tag{32}$$

By substituting these values into (31), we get:

$$\begin{aligned} \pi_k &= \mathbb{E}[(\mathbb{E}[y|\hat{y}_0, y + \Delta_1, \ldots, y + \Delta_{k-1}] - y - \Delta_k)^2] - (1-\beta)^2\mathbb{E}[(\mathbb{E}[y|\hat{y}_0, y + \Delta_1, \ldots, y + \Delta_k] - y)^2] \\ &= \hat{\sigma}_{k-1}^2 + \sigma_\Delta^2 - (1-\beta)^2(\hat{\sigma}_k^2 + \sigma_\Delta^2) = \hat{\sigma}_{k-1}^2 - (1-\beta)^2\hat{\sigma}_k^2 + (2\beta - \beta^2)\sigma_\Delta^2. \end{aligned} \tag{33}$$

To show that π_k is decreasing in k, we only need to prove $\pi_{k+1} - \pi_k < 0$. Note that we can write $\hat{\sigma}_s^2 = \frac{A}{sB+C}$ for positive integer s, where $A = \sigma_\varepsilon^2 \sigma_\Delta^2$, $B = \sigma_\varepsilon^2$, and $C = \sigma_\Delta^2 + \sigma_y^{-2} \sigma_\varepsilon^2 \sigma_\Delta^2$.

$$
\begin{aligned}
\pi_{k+1} - \pi_k &= \left(\frac{A}{kB+C} - \frac{A}{(k-1)B+C}\right) - (1-\beta)^2 \left(\frac{A}{(k+1)B+C} - \frac{A}{kB+C}\right) \\
&= \frac{(1-\beta)^2 AB}{(kB+C)((k+1)B+C)} - \frac{AB}{(kB+C)((k-1)B+C)} \\
&\leq \frac{AB}{(kB+C)((k+1)B+C)} - \frac{AB}{(kB+C)((k-1)B+C)} < 0
\end{aligned} \tag{34}
$$

Lastly, according to Equation 7, $\hat{\sigma}_k^2$ converges to zero as k goes to infinity. Therefore, π_k converges to $(2\beta - \beta^2)\sigma_\Delta^2$. The convergence is from above since:

$$
\hat{\sigma}_{k-1}^2 - (1-\beta)^2 \hat{\sigma}_k^2 \geq \hat{\sigma}_{k-1}^2 - \hat{\sigma}_k^2 > 0. \tag{35}
$$

A.4 Proof of Proposition 1

If $\pi_I \geq \bar{c}$, all communicable DMs find it beneficial to produce information; conversely, if $\pi_1 \leq \underline{c}$, no communicable DM has the incentive to produce information. When neither of these conditions holds, only communicable DMs with a sufficiently low cost of information acquisition produce information. The threshold level c^* should be set in a way that makes communicable DMs indifferent between acquiring and not acquiring information when their cost of information acquisition is c^*. Therefore, the expected gain for information acquisition should be zero, which yields the following condition:

$$
\begin{aligned}
&\sum_{i=0}^{I-1} Prob(c_i \leq c^* \text{ for } i \text{ other communicable DMs})(\pi_{i+1} - c^*) = 0 \\
&\Rightarrow \sum_{i=0}^{I-1} \binom{I-1}{i} G(c^*)^i (1 - G(c^*))^{I-1-i} (\pi_{i+1} - c^*) = 0.
\end{aligned} \tag{36}
$$

It is equivalent with Condition 12. To prove the existence of c^*, we can evaluate the expression above at the extreme values, i.e., $\underline{c}$ and $\bar{c}$:

$$\sum_{i=0}^{I-1} \binom{I-1}{i} G(\underline{c})^i (1-G(\underline{c}))^{I-1-i} (\pi_{i+1} - \underline{c}) = \pi_1 - \underline{c} > 0, \tag{37}$$

$$\sum_{i=0}^{I-1} \binom{I-1}{i} G(\bar{c})^i (1-G(\bar{c}))^{I-1-i} (\pi_{i+1} - \bar{c}) = \pi_I - \bar{c} < 0. \tag{38}$$

We see that Expression 36 changes sign from $c^* = \underline{c}$ to $c^* = \bar{c}$. Since $G(\cdot)$ is continuous, the equation should have an interior solution.

A.5 Proof of Proposition 2

Lemma 1 states that π_k is decreasing in k. An implication of this finding is that if $\bar{I}$ is the greatest positive integer for which $\pi_{\bar{I}} \geq \bar{c}$, then $\pi_I \geq \bar{c}$ for all $I \leq \bar{I}$, which implies that all communicable DMs acquire information for this range of I, meaning that $\bar{\sigma}_{SH} = \frac{\sigma_\varepsilon^2 \sigma_\Delta^2}{I\sigma_\varepsilon^2 + \sigma_\Delta^2 + \sigma_y^{-2} \sigma_\varepsilon^2 \sigma_\Delta^2}$, according to (7), which is decreasing in I.

When $I \geq \bar{I}$, Equation 12 implies that we need to have the following condition for $p_I = G(c_I^*)$, where c_I^* is the level of cost of information acquisition that makes a DM indifferent between acquiring and not acquiring information when there are I communicable DMs. Therefore,

$$H(p_I; I) = 0 \tag{39}$$

,where

$$H(p; I) = \sum_{k=0}^{I-1} \binom{I-1}{k} p^k (1-p)^{I-k-1} (\pi_{k+1} - c) = 0, \qquad p = G(c). \tag{40}$$

Note that $\binom{I-1}{k} p^k (1-p)^{I-k-1}$ is the probability that k other DMs acquire information when each DM acquires information with probability p. When a new communicable DM is added, then larger values of k become probable, which are associated with strictly lower values of π_k. It implies $H(p; I) > H(p; I+1)$ for all values of $p \in (0, 1)$. As a result:

$$H(p_I; I+1) < H(p_I; I) = 0. \tag{41}$$

Moreover, as demonstrated by equation below, we have $\frac{\partial}{\partial p}H(p;I+1)|_{p=p_I} < 0$:

$$\begin{aligned}
&\frac{\partial}{\partial p}H(p,I+1) = \sum_{k=0}^{I}\binom{I}{k}((I-k)p+k(1-p))p^{k-1}(1-p)^{I-k-1}(\pi_{k+1}-c) \\
&= \sum_{k=0}^{I-1}\underbrace{\binom{I}{k}(I-k)}_{\binom{I-1}{k}I}p^{k}(1-p)^{I-k-1}(\pi_{k+1}-c) + \sum_{k=1}^{I}\underbrace{\binom{I}{k}k}_{\binom{I-1}{k-1}I}p^{k-1}(1-p)^{I-k}(\pi_{k+1}-c) \\
&= \underbrace{I\sum_{k=0}^{I-1}\binom{I-1}{k}p^{k}(1-p)^{I-k-1}(\pi_{k+1}-c)}_{H(p;I)} + \underbrace{I\sum_{k=0}^{I-1}\binom{I-1}{k}p^{k}(1-p)^{I-k-1}(\pi_{k+2}-c)}_{<\sum_{k=0}^{I-1}\binom{I-1}{k}p^{k}(1-p)^{I-k-1}(\pi_{k+1}-c)=H(p;I)} \\
&\Rightarrow \frac{\partial}{\partial p}H(p;I+1)|_{p=p_I} < 2IH(p_I;I) = 0.
\end{aligned} \tag{42}$$

By combining (41) and (42), we conclude that p_{I+1}, the solution to $H(p:I+1)=0$ should be smaller than p_I, which implies that the probability of information acquisition decreases with the number of communicable DMs.

A.6 Derivation of Equation 15

Recall that HQ's forecast about y is obtained by $\hat{y} = \gamma_k^{HQ}\hat{y}_0 + \gamma_k^{DM}\bar{y} = (\gamma_k^{HQ}+\gamma_k^{DM})y + \gamma_k^{HQ}\varepsilon + \gamma_k^{DM}\bar{\Delta}$. Therefore, for a given value of k (i.e., HQ's number of signals), $\hat{y}$ has a normal distribution with mean zero and variance:

$$Var(\hat{y}|k \text{ signals from DMs}) = (\gamma_k^{HQ}+\gamma_k^{DM})^2\sigma_y^2 + {\gamma_k^{HQ}}^2\sigma_\varepsilon^2 + {\gamma_k^{DM}}^2\frac{\bar{\Delta}^2}{k}, \tag{43}$$

which simplifies to Σ_k^2, where Σ_k is provided in (15). Let $\mu(k) \equiv \binom{I}{k}G(c^*)^k(1-G(c^*))^{I-k}$ be the probability that HQ receives exactly k signals from the DMs. The expected value for $\hat{y}$ when HQ does not disclose her information is:

$$\begin{aligned}
&\mathbb{E}^{ST}[\hat{y}|\hat{y}\le y^*] = \sum_{k=0}^{I}\mu(k)\frac{1}{\Phi(\Sigma_k^{-1}y^*)}\int_{-\infty}^{y^*}\frac{\hat{y}}{\sqrt{2\pi}\Sigma_k}exp(-\frac{\hat{y}^2}{2\Sigma_k^2})d\hat{y} \\
&= -\sum_{k=0}^{I}\mu(k)\frac{1}{\Phi(\Sigma_k^{-1}y^*)}\frac{\Sigma_k}{\sqrt{2\pi}}exp(-\frac{{y^*}^2}{2\Sigma_k^2}) = -\sum_{k=0}^{I}\mu(k)\frac{\Sigma_k\phi(\Sigma_k^{-1}y^*)}{\Phi(\Sigma_k^{-1}y^*)}.
\end{aligned} \tag{44}$$

In the equation above, $\phi(\cdot)$ and $\Phi(\cdot)$ respectively denote the PDF and CDF of the normal distribution with mean zero and unit variance. Equation 15 is obtained by plugging the expression above in (14).

A.7 Proof of Lemma 2

Without loss of generality, suppose the communicable DMs are $DM_1, DM_2, \ldots, DM_k$. Each communicable DM solves the following optimization problem in choosing their report:

$$\max_{r_i} -\mathbb{E}[(b_1\hat{y} + b_2 r_i - y - \Delta_i)^2 | y + \Delta_i]. \tag{45}$$

By examining the first-order condition, we find that the optimal report should be such that:

$$b_2 r_i = \mathbb{E}[y + \Delta_i - b_1\hat{y} | y + \Delta_i] = y + \Delta_i - b_1\mathbb{E}[\hat{y}|y + \Delta_i]. \tag{46}$$

From (4), we know:

$$\hat{y} = \gamma_k^{HQ}\hat{y}_0 + \gamma_k^{DM}(y + \frac{1}{k}\sum_{j=1}^{k}\Delta_j) = (\gamma_k^{HQ} + \frac{k-1}{k}\gamma_k^{DM})y + \gamma_k^{HQ}\varepsilon + \frac{1}{k}\gamma_k^{DM}\sum_{j\neq i, j\leq k}\Delta_j + \frac{\gamma_k^{DM}}{k}(y + \Delta_i). \tag{47}$$

Furthermore:

$$\begin{gathered} \mathbb{E}[y|y + \Delta_i] = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_\Delta^2}(y + \Delta_i), \quad \mathbb{E}[\varepsilon|y + \Delta_i] = 0 \\ \mathbb{E}[\Delta_j|y + \Delta_i] = 0 \quad j \neq i. \end{gathered} \tag{48}$$

We can verify Equation 18 by plugging the expressions in (47) and (48) in (46).

A.8 Proof of Proposition 3

Based on Equation 20, we only need to find $Var(\hat{y}|y+\Delta_i)$ to derive π_k. To this end, without loss of generality, suppose the first k DMs are communicable and $1 \leq i \leq k$. Note that:

$$\begin{aligned}
Var(\hat{y}|y+\Delta_i) &= Var(\mathbb{E}[y|\hat{y}_0, \{y+\Delta_j\}_{1\leq j\leq k}]|y+\Delta_i) \\
&= \mathbb{E}[(\mathbb{E}[y|\hat{y}_0, \{y+\Delta_j\}_{1\leq j\leq k}] - \mathbb{E}[y|y+\Delta_i])^2] \\
&= \mathbb{E}[((y-\mathbb{E}[y|\hat{y}_0, \{y+\Delta_j\}_{1\leq j\leq k}]) - (y-\mathbb{E}[y|y+\Delta_i]))^2] \\
&= Var(y|y+\Delta_i) - Var(y|\hat{y}_0, \{y+\Delta_j\}_{1\leq j\leq k}) = \frac{1}{\sigma_y^{-2}+\sigma_\Delta^{-2}} - \hat{\sigma}_k^2.
\end{aligned} \tag{49}$$

By substituting the expression above in (20), we get:

$$\pi_k = \hat{\sigma}_{k-1}^2 + b_1^2\hat{\sigma}_k^2 + \sigma_\Delta^2 - \frac{b_1^2}{\sigma_y^{-2}+\sigma_\Delta^{-2}}. \tag{50}$$

Since $\hat{\sigma}_k^2$ is decreasing in k, so is π_k. Moreover, $\hat{\sigma}_k^2$ converges to zero as k goes to infinity, which proves that π_k converges to $\sigma_\Delta^2 - \frac{b_1^2}{\sigma_y^{-2}+\sigma_\Delta^{-2}}$. The converges is from above since $\hat{\sigma}_k^2 > 0$. The proof for parts (b) and (c) are similar to that for Proposition 2.

Appendix B. Variable Definitions

Variable	Definition
ln(1+EPS forecasts)	natural log of 1 plus the number of management EPS forecasts issued during 1 year after the fiscal year-end, based on I/B/E/S. We count EPS forecasts issued after and on the earnings announcement date corresponding to the current fiscal year and before the earnings announcement for the subsequent fiscal year.
ln(Website length)	average size (in bytes) of a corporate official homepage in a year based on Wayback Machine, following Boulland et al. (2021).
EPS Forecast Accuracy	the absolute difference of the management EPS forecast and actual realized EPS, based on I/B/E/S. We consider EPS forecasts issued after and on the earnings announcement date corresponding to the current fiscal year and before the earnings announcement for the subsequent fiscal year. The average of absolute differences is used when a firm provides more than one EPS forecast during the period. This variable is used in the Online Appendix.
Intranet	the firm's use of the intranet which serves as a proxy for the extent of internal communication following Bloom et al. (2014). It is calculated as the average site-level intranet in a year weighted by employees at each site. Each site is assigned 1 if it has installed the intranet, and 0 otherwise, based on the Harte-Hanks Ci Technology database (CiDB). We infer the installment of intranets based on the presence of technologies used to connect offices or product sites, including frame relays, leased lines, a synchronous optical network (SONET), digital subscriber line (xDSL), asynchronous transfer mode (ATM) technology, and virtual private networks (VPN).
Site-level Intranet	1 if the divisional site has installed the intranet, and 0 otherwise. See the definition of $Intranet$ for details.
Internet	the firm's use of the internet calculated as the average site-level internet in a year weighted by employees at each site. Each site is assigned 1 if it has installed the internet, and 0 otherwise, based on the Harte-Hanks Ci Technology database (CiDB).
Site-level Internet	1 if the divisional site has installed the internet, and 0 otherwise, based on the Harte-Hanks Ci Technology database (CiDB).
Number of PCs per employee	the number of PCs over the number of employees within the firm in a year, based on the Harte-Hanks Ci Technology database (CiDB).
Number of PCs at site per employee	the number of PCs over the number of employees within the divisional site in a year, based on the Harte-Hanks Ci Technology database (CiDB).
ROA	return on assets measured as net income over average total assets between the last and current fiscal years.
BTM	book value of equity scaled by the market value of equity at the end of the fiscal year.
ln(MVE)	natural log of price per share $\times$ number of shares outstanding at the end of the fiscal year. MVE is measured in millions of USD.
R&D	research and development expenditures divided by revenues during the fiscal year.
Loss Indicator	1 if the firm reports losses for the fiscal year.
ln(1+analysts following)	natural log of 1 plus the number of analysts following the firm during the fiscal year, obtained from I/B/E/S.
Earnings Volatility	standard deviation of quarterly earnings over 12 quarters ending in the current fiscal year, divided by the average quarterly asset value of these quarters.
ln(employees)	natural log of the firm's total employees in a year based on the Harte-Hanks Ci Technology database (CiDB).
ln(site-level employees)	natural log of the divisional site's total employees in a year based on the Harte-Hanks Ci Technology database (CiDB).
ln(1+revenues)	natural log of 1 plus the firm's revenues in a year based on the Harte-Hanks Ci Technology database (CiDB). Revenues are measured in millions of USD.
ln(1+site-level revenues)	natural log of 1 plus the divisional site's revenues in a year based on the Harte-Hanks Ci Technology database (CiDB). Revenues are measured in millions of USD.
ln(sites)	natural log of the total number of sites operated by the firm in a year based on the Harte-Hanks Ci Technology database (CiDB).

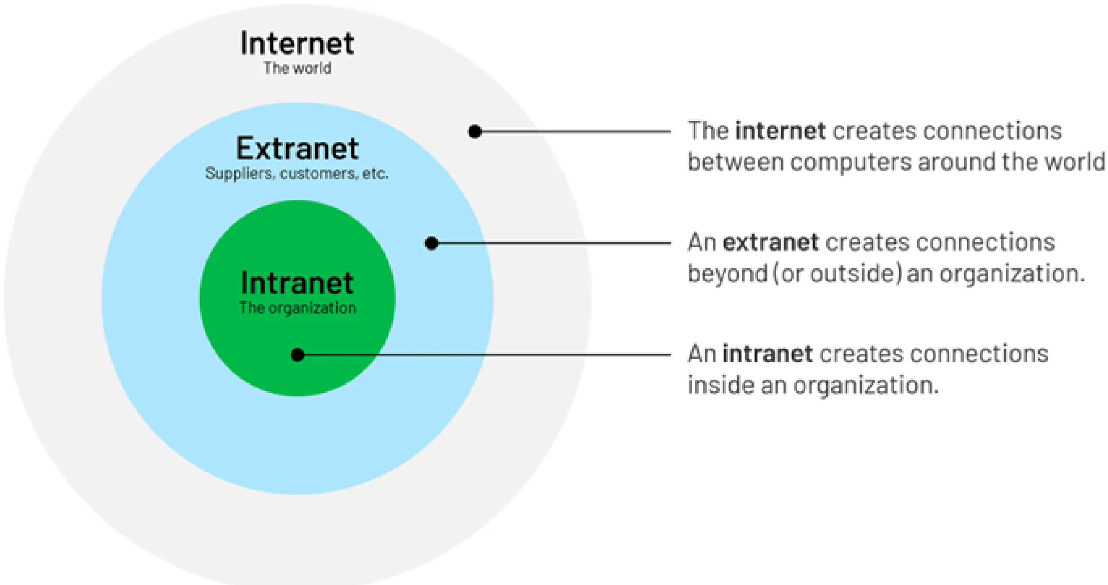
Appendix B (cont'd). Variable Definitions

Variable	Definitions
ln(business segments)	natural log of 1 plus the number of business segments as of the fiscal year-end, measured using Compustat Segments.
ln(geographical segments)	natural log of 1 plus the number of geographical segments as of the fiscal year-end, measured using Compustat Segments.
Related	the number of related business segments divided by the total number of business segments as of the fiscal year-end. The number of related segments is the difference between the total number of segments reported for the firm and the number of segments with different two-digit SIC codes. We assign Related =1 if all of the firm's business segments have the same two-digit SIC code.
Institutional ownership	percent of institutional investors at the fiscal year-end, obtained from Thomson Reuters.
Distance	average distance between the firm's HQ office and divisional offices weighted by employees at each divisional site, based on the Harte-Hanks Ci Technology database (CiDB). The distance is based on the zipcode of each office and the latitude/longitude coordinate if the zipcode is missing. It is measured in miles.
Distance between HQ and site	distance between the firm's HQ office and the divisional site, based on the Harte-Hanks Ci Technology database (CiDB). The distance is based on the zipcode and the latitude/-longitude coordinate if the zipcode is missing. It is measured in miles.
Site Similarity	sum of the square of the share of each 2-digit SIC industry code that the firm's sites with intranet operate in. If a firm has two sites with the intranet operating in one 2-digit SIC industry and three sites with the intranet operating in another 2-digit SIC industry, then this firm's *Site Similarity* is $(2/5)^2 + (3/5)^2 = 0.52$.
% of Revenue at HQ	ratio of revenue generated by the HQ office over total revenues of the firm in a year, based on the Harte-Hanks Ci Technology database (CiDB).
% of Employees at HQ	ratio of employees in the HQ office over total employees of the firm in a year, based on the Harte-Hanks Ci Technology database (CiDB).

Appendix C. The Intranet

Panel A illustrates the differences between the internet, extranet, and intranet. Panel B illustrates the role of the intranet in facilitating telecommunication between regional offices via a private network. Source for Panel A: https://jesseokeya.medium.com/building-blocks-of-the-internet-3dc6c39fbd75.

Panel A: Definition of the Intranet



Panel B: Illustration of Intra-firm communication via the Intranet

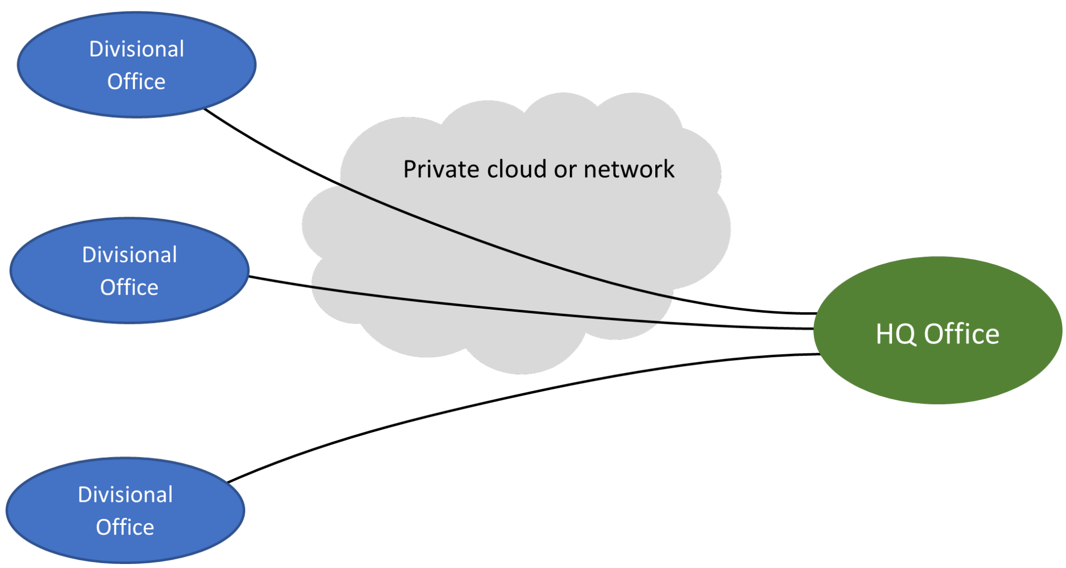
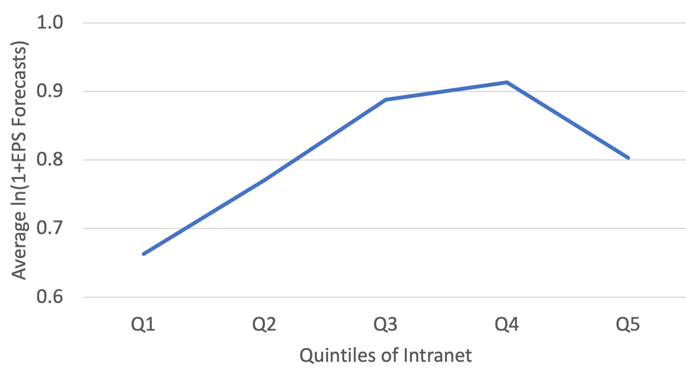


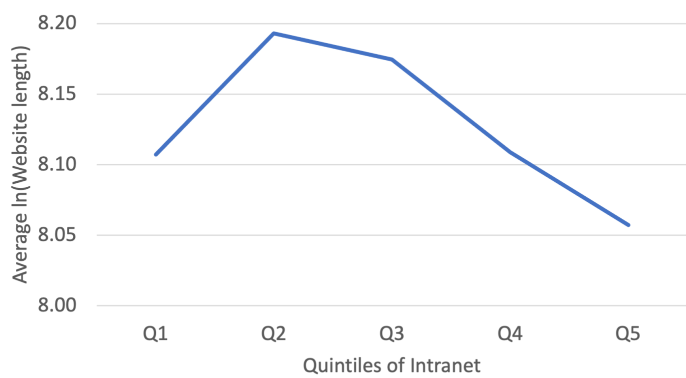
Fig. 8: Averages of Voluntary Disclosure by Quintile of Intranet

This figure plots average voluntary disclosure for each quintile of firms' use of the intranet ($Intranet$). Panel A presents average $ln(1 + EPS\ forecasts)$ for public firms by quintile of $Intranet$. Panel B and Panel C show average $ln(Website\ length)$ by quintile of $Intranet$ for public firms and private firms, respectively. All variables are defined in Appendix B.

Panel A: Public Firms' EPS Forecasts and Intranet



Panel B: Public Firms' Website Disclosure and Intranet



Panel C: Private Firms' Website Disclosure and Intranet

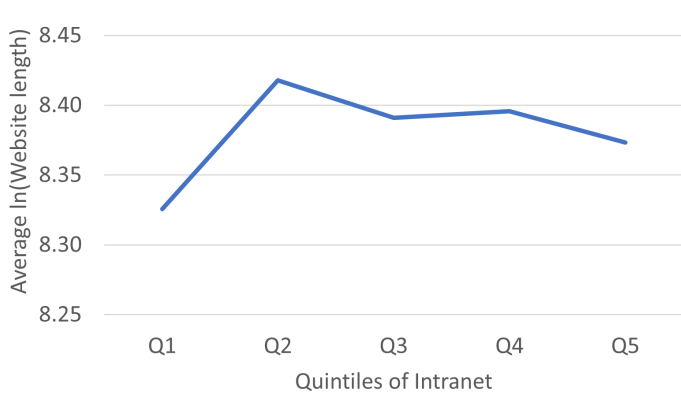


Table 1: Summary Statistics

The table contains summary statistics of some key variables used in our analyses. See Appendix B for variable definitions. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers.

Variable	**Mean**	**SD**	**p25**	**Median**	**p75**	***N***
Public Firm-Year Variables						
Intranet	0.49	0.33	0.18	0.51	0.80	20,778
EPS forecasts	2.71	3.95	0.00	0.00	5.00	20,778
Website length (in bytes)	5,409.1	5,227.7	1,991.6	4,165.7	6,836.0	17,600
ROA	0.03	0.12	0.00	0.04	0.08	20,778
BTM	0.57	0.56	0.28	0.47	0.75	20,778
MVE	6,709.1	18,874.8	242.6	982.3	3,865.7	20,778
R&D	0.04	0.09	0.00	0.00	0.04	20,778
Analysts following	6.76	7.32	1.00	4.00	11.00	20,778
Earnings Volatility	0.22	0.23	0.07	0.15	0.31	20,778
Employees	4,946.3	10,060.2	459.0	1,435.0	4,421.0	20,778
Revenues	1,437.1	3,338.4	97.0	327.0	1,125.0	20,778
Sites	39.48	88.22	4.00	10.00	30.00	20,778
Internet	0.92	0.19	0.97	1.00	1.00	20,778
Number of PCs per employee	0.64	0.37	0.41	0.58	0.81	20,778
Business segments	8.46	5.71	3.00	9.00	12.00	20,778
Geographical segments	8.30	7.25	3.00	6.00	12.00	20,778
Related	0.78	0.22	0.67	0.80	1.00	20,778
Institutional ownership	0.57	0.35	0.25	0.67	0.87	20,778
Loss Indicator	0.25	0.43	0.00	0.00	0.00	20,778
Distance	517.6	406.4	193.7	441.7	747.0	20,778
Site Similarity	0.68	0.29	0.44	0.63	1.00	14,832
% of Revenue at HQ	0.30	0.30	0.04	0.17	0.51	19,305
% of Employees at HQ	0.31	0.30	0.05	0.20	0.52	19,351
Private Firm-Year Variables						
Intranet	0.42	0.38	0.00	0.41	0.80	225,425
Website length (in bytes)	5,969.4	5,019.7	2,768.3	4,691.8	7,620.7	225,425
% of Revenue at HQ	0.45	0.32	0.16	0.46	0.71	166,099
% of Employees at HQ	0.47	0.29	0.21	0.48	0.70	170,922
Distance	144.6	271.3	3.6	18.4	138.9	225,425
Site Similarity	0.86	0.24	0.72	1.00	1.00	102,065
Number of PCs per employee	0.99	0.55	0.63	0.84	1.16	225,425
Internet	0.99	0.04	1.00	1.00	1.00	225,425
Employees	685.8	1710.3	47.0	135.0	461.0	225,425
Revenues	122.4	332.1	6.0	19.0	74.0	225,425
Sites	7.7	14.2	2.0	3.0	7.0	225,425

Table 2: Non-Linear Association between Intranet and Forecast

This table reports estimates from the following firm-year-level regression: $ln(1 + EPS\ forecasts)_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	ln(1+EPS forecasts)$_{t+1}$	
	(1)	(2)
Intranet	1.039***	0.276**
	(7.63)	(2.20)
Intranet^2	-0.839***	-0.307**
	(-5.77)	(-2.30)
ROA		0.008
		(0.06)
BTM		-0.047**
		(-2.28)
ln(MVE)		0.045***
		(3.36)
R&D		0.161
		(0.56)
ln(1+analysts following)		0.142***
		(4.65)
Earnings Volatility		-0.178***
		(-4.31)
ln(employees)		0.069
		(1.45)
ln(1+revenues)		-0.007
		(-0.17)
ln(sites)		-0.006
		(-0.25)
Internet		0.013
		(0.20)
Number of PCs per employee		0.030
		(0.67)
ln(business segments)		0.083***
		(3.53)
ln(geographical segments)		0.037**
		(2.45)
Related		0.042
		(0.79)
Institutional ownership		0.308***
		(3.85)
Loss Indicator		-0.176***
		(-7.67)
Industry × Year FE	Yes	Yes
N	20,778	20,778
R-squared	0.084	0.261
Clustering	Industry	

Table 3: Role of Information Learning

This table reports results from the following firm-year-level regressions estimated separately for firm-years above and below the median cutoff: $ln(1 + EPS\ forecasts)_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. Columns (1)-(2) split the sample based on median *Distance*. *Distance* is the average distance between the HQ office and divisional offices weighted by employees at each divisional site. Controls include: ROA, BTM, ln(MVE), R&D, ln(1+analysts following), Earnings Volatility, ln(employees), ln(1+revenues), ln(sites), Internet, Number of PCs per employee, ln(business segments), ln(geographical segments), Related, Institutional ownership, and Loss Indicator. All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. *t*-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	$\ln(1+\text{EPS forecasts})_{t+1}$	
	Below-Median Distance (1)	Above-Median Distance (2)
Intranet	0.035	0.560***
	(0.21)	(3.52)
Intranet^2	-0.035	-0.576***
	(-0.19)	(-3.19)
Controls	Yes	Yes
Industry × Year FE	Yes	Yes
N	10,299	10,371
R-squared	0.274	0.255
Clustering	Industry	

Table 4: Role of Free Riding

This table reports results from the following firm-year-level regressions estimated separately for firm-years below and above the median cutoff: $ln(1+EPS\ forecasts)_{i,t+1} = \beta_1 Intranet_{i,t}\ (+\beta_2 Intranet^2_{i,t}) + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. Columns (1)-(2) and columns (3)-(4) split the sample based on median $Site\ Similarity_{i,t-1}$. $Site\ Similarity_{i,t-1}$ captures the extent of similarity across a firm's divisional sites with intranet calculated by the sum of the square of the share of each 2-digit SIC industry. Controls include: ROA, BTM, ln(MVE), R&D, ln(1+analysts following), Earnings Volatility, ln(employees), ln(1+revenues), ln(sites), Internet, Number of PCs per employee, ln(business segments), ln(geographical segments), Related, Institutional ownership, and Loss Indicator. All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. *t*-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	$\ln(1+\text{EPS forecasts})_{t+1}$			
	Below-Median Site Similarity		Above-Median Site Similarity	
	(1)	(2)	(3)	(4)
Intranet	0.231***	0.672*	-0.093	0.566**
	(3.29)	(1.65)	(-1.42)	(2.19)
Intranet^2		-0.390		-0.644**
		(-1.15)		(-2.55)
Controls	Yes	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
N	7,103	7,103	7,729	7,729
R-squared	0.249	0.249	0.260	0.262
Clustering	Industry			

Table 5: Non-Linear Association between Intranet and Website Disclosure

Panel A and Panel B report results from the following firm-year-level regressions estimated for public firms and private firms, respectively: $ln(Website\ length)_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Public Firms		
	ln(Website length)$_{t+1}$	
	(1)	(2)
Intranet	0.223	0.427***
	(1.34)	(2.77)
Intranet^2	-0.353**	-0.570***
	(-2.15)	(-3.26)
ROA		-0.090
		(-0.71)
BTM		-0.071***
		(-3.05)
ln(MVE)		-0.017
		(-1.09)
R&D		0.465***
		(4.87)
ln(1+analysts following)		0.044*
		(1.85)
Earnings Volatility		-0.143**
		(-2.62)
ln(employees)		0.050
		(1.25)
ln(1+revenues)		-0.012
		(-0.30)
ln(sites)		-0.065**
		(-2.19)
Internet		-0.047
		(-0.57)
Number of PCs per employee		0.066
		(1.14)
ln(business segments)		-0.003
		(-0.14)
ln(geographical segments)		0.002
		(0.10)
Related		-0.023
		(-0.40)
Institutional ownership		0.045
		(1.06)
Loss Indicator		0.004
		(0.11)
Industry × Year FE	Yes	Yes
N	17,600	17,600
R-squared	0.058	0.066
Clustering	Industry	

Table 5: Non-Linear Association between Intranet and Website Disclosure (cont'd)

Panel B: Private Firms		
	ln(Website length)$_{t+1}$	
	(1)	(2)
Intranet	0.247***	0.160**
	(5.08)	(2.60)
Intranet^2	-0.179***	-0.159**
	(-3.77)	(-2.51)
ln(employees)		0.039***
		(5.52)
ln(1+revenues)		0.009
		(1.12)
ln(sites)		-0.014
		(-0.94)
Internet		-0.169***
		(-2.75)
Number of PCs per employee		0.051**
		(2.10)
Industry × Year FE	Yes	Yes
N	225,425	225,425
R-squared	0.078	0.081
Clustering	Industry	

Table 6: Robustness Test: Controlling for Arbitrary Non-Linearities

Panel A reports results from the following firm-year-level regression estimated for public firms: $ln(1 + EPS\ forecasts)_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \gamma Controls^2_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. Panel B reports results from the following firm-year-level regression estimated for public firms: $ln(Website\ length)_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \gamma Controls^2_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. Panel C reports results from the following firm-year-level regression estimated for private firms: $ln(Website\ length)_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \gamma Controls^2_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. All variables are defined in Appendix B. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: EPS Forecasts for Public Firms

		ln(1+EPS forecasts)$_{t+1}$	
Intranet		0.310**	
		(2.56)	
Intranet^2		-0.319**	
		(-2.37)	
ROA	0.053	ROA^2	0.299
	(0.29)		(0.81)
BTM	0.003	BTM^2	-0.011
	(0.11)		(-0.91)
ln(MVE)	0.232***	ln(MVE)^2	-0.016***
	(5.05)		(-4.43)
R&D	1.341*	R&D^2	-2.603**
	(1.68)		(-2.18)
ln(1+analysts following)	0.116**	ln(1+analysts following)^2	0.025
	(2.40)		(1.26)
Earnings Volatility	-0.504***	Earnings Volatility^2	0.333***
	(-5.18)		(3.87)
ln(employees)	-0.314***	ln(employees)^2	0.028***
	(-2.91)		(3.43)
ln(1+revenues)	0.043	ln(1+revenues)^2	-0.004
	(0.64)		(-0.64)
ln(sites)	0.036	ln(sites)^2	-0.009
	(0.76)		(-1.12)
Internet	0.277	Internet^2	-0.207
	(1.19)		(-1.04)
Number of PCs per employee	0.044	Number of PCs per employee^2	-0.015
	(0.44)		(-0.31)
ln(business segments)	0.506***	ln(business segments)^2	-0.118***
	(5.12)		(-3.80)
ln(geographical segments)	0.127***	ln(geographical segments)^2	-0.033**
	(2.81)		(-2.37)
Related	-0.054	Related^2	0.161
	(-0.22)		(0.76)
Institutional ownership	-0.710***	Institutional ownership^2	0.929***
	(-3.95)		(5.62)
Loss Indicator	-0.170***		
	(-6.60)		
Industry × Year FE		Yes	
N		20,778	
R-squared		0.285	
Clustering		Industry	

Table 6: Robustness Test: Controlling for Arbitrary Non-Linearities (cont'd)

Panel B: Website Length for Public Firms			
		ln(Website Length)$_{t+1}$	
Intranet		0.371**	
		(2.19)	
Intranet^2		-0.512***	
		(-2.75)	
ROA	-0.255	ROA^2	-0.365
	(-1.19)		(-0.73)
BTM	-0.050	BTM^2	-0.005
	(-1.48)		(-0.33)
ln(MVE)	0.113*	ln(MVE)^2	-0.010**
	(1.68)		(-2.16)
R&D	1.253***	R&D^2	-1.550***
	(5.07)		(-3.92)
ln(1+analysts following)	0.073	ln(1+analysts following)^2	-0.004
	(0.95)		(-0.15)
Earnings Volatility	-0.296**	Earnings Volatility^2	0.173
	(-2.19)		(1.46)
ln(employees)	0.015	ln(employees)^2	0.003
	(0.14)		(0.36)
ln(1+revenues)	0.044	ln(1+revenues)^2	-0.005
	(0.42)		(-0.54)
ln(sites)	-0.079*	ln(sites)^2	0.004
	(-1.94)		(0.45)
Internet	0.158	Internet^2	-0.185
	(0.48)		(-0.72)
Number of PCs per employee	0.162	Number of PCs per employee^2	-0.051
	(1.30)		(-1.08)
ln(business segments)	-0.025	ln(business segments)^2	0.008
	(-0.41)		(0.47)
ln(geographical segments)	0.040	ln(geographical segments)^2	-0.015
	(1.35)		(-1.39)
Related	-0.047	Related^2	0.016
	(-0.20)		(0.07)
Institutional ownership	-0.126	Institutional ownership^2	0.067
	(-0.60)		(0.36)
Loss Indicator	-0.009		
	(-0.23)		
Industry × Year FE		Yes	
N		17,600	
R-squared		0.069	
Clustering		Industry	

Table 6: Robustness Test: Controlling for Arbitrary Non-linearities (cont'd)

Panel C: Website Length for Private Firms			
		$\ln(\text{Website Length})_{t+1}$	
Intranet		0.126**	
		(2.38)	
Intranet^2		-0.122**	
		(-2.12)	
ln(employees)	0.021	$\ln(\text{employees})^2$	0.002
	(0.34)		(0.27)
ln(1+revenues)	0.015	$\ln(1+\text{revenues})^2$	-0.001
	(0.69)		(-0.26)
ln(sites)	0.095***	$\ln(\text{sites})^2$	-0.026***
	(3.89)		(-3.91)
Internet	-1.533	Internet^2	0.783
	(-0.87)		(0.74)
Number of PCs per employee	0.184***	Number of PCs per employee^2	-0.047***
	(4.08)		(-3.84)
Industry × Year FE		Yes	
N		225,425	
R-squared		0.082	
Clustering		Industry	

Online Appendix for '(In)Consistent Internal-External Communication? How Internal Communication Technology Affects Voluntary Disclosure'

Ehsan Azarmsa, Lisa Yao Liu, and Suzie Noh*

March 2023

Abstract

This Online Appendix contains supplementary results and discussion not included in the main manuscript. Additional tests presented herein study the determinants of intranet adoption, explore linear relations between intra-firm intranet intensity (*Intranet*) and voluntary disclosure, show cross-sectional results using website disclosure as an alternative proxy for voluntary disclosure, and assess the robustness of cross-sectional results to controlling for non-linear relations between voluntary disclosure and control variables. Additionally, we document similar non-linear, inverse U-shape relations between intra-firm intranet intensity (*Intranet*) and an alternative measure of voluntary disclosure, EPS forecast accuracy. Lastly, we validate the assumption of decision centralization for intranet adoption.

*Ehsan Azarmsa: Price College of Business, University of Oklahoma (azarmsa@ou.edu). Lisa Yao Liu: Columbia Business School, Columbia University (yl4689@gsb.columbia.edu). Suzie Noh: Stanford Graduate School of Business, Stanford University (suzienoh@stanford.edu).

1 Firm-level and Site-level Determinants of Intranet Adoption

In this section, we study associations between various firm and site-level characteristics and the intranet to shed light on the determinants of intranet adoption. First, we estimate the following firm-year-level regression for public and private firms, respectively:

$$Intranet_{i,t+1} = \theta Determinants_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} \ (+ \Sigma\rho_i Firm_i) + \epsilon_{i,t}, \qquad (1)$$

where the dependent variable $Intranet_{i,t}$ is our measure of intra-firm intranet intensity. It captures the fraction of a firm's divisions adopting the intranet, calculated as the weighted average site-level intranet with sites' employee counts used as the weights. Each site's intranet takes the value of 1 if it has installed the intranet and 0 otherwise.

The results are tabulated in Panel A and Panel B of Table OA.1. In Panel A, we find in columns (1)-(3) that public firms with higher ROA, greater geographic segment dispersion, and greater information technologies—such as the internet and the number PCs—have higher intranet intensity. We also find firms that are larger in size, proxied by the number of employees and revenues, are more likely to install the intranet in their divisional sites. Yet, firms that have a greater number of sites have lower intranet intensity as shown by the negative coefficient on $ln(sites)$, which is consistent with the costs associated with installing the intranet in each site. Our within-firm analysis including firm fixed effects shown in columns (4)-(6) suggests that firms become more likely to have the intranet in their divisional offices when they invest in information technologies and grow their employees or revenues. Across columns (1)-(6), we do not find a significant association between intranet intensity and the average distance between the firm's HQ office and divisional offices.

Panel B shows the results for private firms, which have fewer determinant variables due to limited data availability for private firms. Similar to our findings for public firms, we show that the internet, PCs, employees, and revenues positively predict firms' intranet adoption and that the number of sites negatively predicts it. While we show firms that have divisional sites farther away from the headquarters' office have lower intranet intensity (i.e., columns (1)-(3)), this relation disappears in our within-firm analysis, which include firm fixed effects (i.e., columns (4)-(6)).

Second, to explore the drivers of intranet we run the following site-year-level regression estimated for public and private firms, respectively:

$$\begin{aligned} \textit{Site-level Intranet}_{s,t+1} = & \theta \textit{Site-level Determinants}_{s,t} + \Sigma\rho_s Site_s \\ & + \Sigma\beta_{i,t} Firm \times Year_{i,t} + \epsilon_{s,t}. \end{aligned} \tag{2}$$

where the dependent variable *Site-level Intranet*$_{s,t+1}$ is 1 if the divisional site has installed the intranet, and 0 otherwise.

The results are presented in Panel C and Panel D of Table OA.1. In Panel C, we find that, within a public firm, the divisional site that has greater employees and revenues is more likely to have the intranet, as shown by the positive coefficient on *ln*(*site-level employees*) and *ln*(1 + *site-level revenues*). We also find that the site that has other information technologies is more likely to have the intranet, as proxied by the positive coefficients on *Site-level Internet* and *Number of PCs at site per employee*. Importantly, we do not find that site's likelihood of having the intranet is associated with the distance from the headquarters to the site, but is positively and significantly associated with the economic importance of the site. This result is consistent with our theoretical assumption that internal information asymmetry is unlikely to predict the installment of the intranet at the site but the site's economic importance.

The results for private firms shown in Panel B are similar. Within a private firm, the divisional sites that have greater PCs per employee and greater employees or revenues are more likely to install the intranet. We also continue to find no association between the distance and the intranet at private firms' sites.

2 Linear Relation between Intranet and Voluntary Disclosure

Our theory predicts a nonlinear, inverse U-shaped relation between internal communication facilitated by communication technology and external communication. To highlight the importance of our formal theory and its predictions in non-linear terms, we empirically document the findings of linear regressions of voluntary disclosure on internal communication technology. We estimate the following linear regression for both public and private firms:

$$\begin{aligned} Voluntary\ Disclosure_{i,t+1} = &\beta_1 Intranet_{i,t} + \theta Controls_{i,t} \\ &+ \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}, \end{aligned} \tag{3}$$

where the key independent variable $Intranet_{i,t}$ is our measure of intra-firm intranet intensity. It captures the fraction of a firm's divisions adopting the intranet, calculated as the weighted average site-level intranet with sites' employee counts used as the weights. Each site's intranet takes the value of 1 if it has installed the intranet and 0 otherwise.

For public firms, we have two dependent variables: $ln(1 + EPS\ forecasts)_{i,t+1}$ and $ln(Website\ length)_{i,t+1}$. The dependent variable $ln(1 + EPS\ forecasts)_{i,t+1}$ is the frequency of management EPS forecasts measured between the earnings announcement corresponding to fiscal year t and the earnings announcement for the subsequent fiscal year $t + 1$. The dependent variable $ln(Website\ length)_{i,t+1}$ is the size of the firm's corporate website homepage measured during the first calendar year after fiscal year t. Because the CiDB provides data on firms' use of the intranet by calendar year, we measure $Intranet_{i,t}$ as of the last calendar year ending before fiscal year t. Therefore, the measurement windows for voluntary disclosure always come after that for the intranet. For public firm analysis, $Controls_{i,t}$ is a vector of control variables measured for each firm i's fiscal year t, and include characteristics that could be associated with firms' voluntary disclosure decisions (e.g., Chen et al., 2018). These control variables include ROA, BTM, $ln(MVE)$, *R&D*, *Loss*, $ln(1+ analysts\ following)$, $Earnings\ Volatility$, $ln(business\ segments)$, $ln(geographical\ segments)$, *Related*, and *Institutional ownership* obtained from Compustat, CRSP, I/B/E/S, and Thomson Reuters as well as $ln(employees)$, $ln(1 + revenues)$, $ln(sites)$, *Internet*, and *Number of PCs per employee* obtained from the CiDB.

To present the linear relations estimated, we follow the format used in our main manuscript. Table OA.2 is similar to Table 2 of the main manuscript, except we exclude the second-order polynomial term of *Intranet* in estimating the regression. Similarly, Table OA.3 is a version of Table 5 in our manuscript, excluding the second-order polynomial term of *Intranet*. In Table

OA.2, we document an insignificant linear relation between $EPS\ forecasts$ and $Intranet$ for public firms when we include control variables. This no-result persists independent of excluding firm fixed effects (column (2)) or including them (column (4)). In Panel A of OA.3, we find a significantly negative relation between $ln(Website\ length)$ and $Intranet$ for public firms when we exclude firm fixed effects (column (2)) and an insignificant linear relation when we include them (column (4)). In Panel B of OA.3, we find a significant positive linear relation between $ln(Website\ length)$ and $Intranet$ for private firms if we do not include control variables and firm fixed effects, but find an insignificant linear relation if we include firm fixed effects.

These results underscore the limitation of using linear regression models to understand the relation between internal communication technology and external communication, and highlight the importance of formally modeling economic forces that lead to a non-linear prediction.

3 Role of Information Learning: Website Disclosure

Our main manuscript shows that one of the two economic forces shaping the inverse U-shape is "information learning": the HQ manager increases her information precision by acquiring information from divisional managers and therefore increases the provision of voluntary disclosure. Consistent with this channel, our theory further predicts that when the HQ manager's prior information is more precise, internal communication contributes less to the HQ manager's learning from divisional managers, and therefore the relation between internal communication technology and voluntary disclosure becomes flat (e.g., see Figure 6 of our main manuscript).

To provide empirical evidence corroborating this information learning channel, we show that the inverse U-shape is concentrated (muted) in the subsamples of public and private firms where HQ managers likely have less (more) precise prior information. Specifically, using the distance between the headquarters and divisional managers as a proxy for HQ managers' ex ante informedness, we re-estimate Equation 24 of the main manuscript after splitting the sample into two groups, above-median and below-median distance to divisional offices. In our main manuscript, we show that the inverse U-shape is pronounced (muted) in firms with above-median (below-median) distance, where the information learning force is pronounced (weakened).

In this section, we document similar results using website length as an alternative proxy for voluntary disclosure. Table OA.4 shows results for public firms (Panel A) and private firms (Panel B). For both Panel A and Panel B, the coefficients on $Intranet^2$ in column (1) and column (2) are significantly different with the p-value <0.05.

4 Robustness Test on the Role of Information Learning

In this section of the Online Appendix, we confirm the robustness of the cross-sectional test results for the information learning channel. We re-estimate the regression Equation 24 of the main manuscript for both public and private firms after including the quadratic terms of all control variables to account for ad hoc nonlinear relationships between voluntary disclosure and our control variables. Across Table OA.5 and Table OA.6, we find significant coefficients on $Intranet^2$ only when firms have above-median distance. We confirm that the coefficients on $Intranet^2$ are statistically different between column (1) and column (2) with p-values of <0.05, <0.01, and <0.10, respectively, for Table OA.5 and Table OA.6 Panel A - Panel B.

5 Robustness Test on the Role of Free Riding

We further confirm the robustness of cross-sectional test results for the free riding channel. We re-estimate the regression Equation 24 of the main manuscript for public firms after including the quadratic terms of all control variables. Table OA.7 shows the results. Similar to our main manuscript, we find a significantly positive linear relation between *Intranet* and EPS forecasts for firms with below-median *Site Similarity* (column (1)) and a significant inverse U-shape for firms with above-median *Site Similarity* (column (4)). The coefficients on *Intranet* in columns (1) and (3) are statistically different with p-value <0.01.

6 Unimodal Relation between Intranet and Forecast Accuracy

In this section, we use the EPS forecast accuracy of firms that are always forecasters as two proxies. First, we use it as a proxy for managers' information precision to test the theoretical prediction in Section 2.3.1. Second, we use it to proxy for the informativeness of disclosure, conditional on firms always providing disclosures. The informativeness of disclosure can be considered as an alternative measure of voluntary disclosure (intensive margin). In the theoretical predictions on voluntary disclosure in Section 2.3.2, we predict that, if firms always disclose, the HQ manager truthfully reports her information to stakeholders, but the discretionary component in the disclosure is its informativeness. Thus, this measurement is consistent with this theoretical prediction.

We estimate the regression Equation 24 in the main manuscript using EPS forecast accuracy as the dependent variable. Due to the availability of EPS forecasts, our analysis here is limited to public firms. Furthermore, this analysis requires limiting the sample to firms that provide forecasts which reduces our sample significantly. To avoid a look-ahead bias, we use firms that provided EPS forecasts every year during the last three years, 1998-2000, before the beginning of our sample period, 2001. This gives the final sample of 2,152 firm-years.[1] We measure EPS forecast accuracy as the absolute difference of the management EPS forecast and actual realized EPS, based on I/B/E/S. We consider EPS forecasts issued after and on the earnings announcement date corresponding to the current fiscal year and before the earnings announcement for the subsequent fiscal year. Therefore, the measurement window for EPS forecast accuracy precedes that for *Intranet*, which is measured as of the last calendar year ending before the current fiscal year. The average of absolute differences is used when a firm provides more than one EPS forecast during the period.

In Table OA.8, we find a significant coefficient on $Intranet^2$ suggesting an inverse U-shape between internal communication technology and forecasting accuracy. This finding is consistent with our main findings in Table 2 of the main manuscript, as well as our theoretical predictions in Section 2.3.1. This finding supports our assumption that, conditional on HQ managers' disclosure, her disclosure is based on her information set.

7 Validating the Decision Centralization of Intranet Adoption

To validate the decision centralization channel, we empirically examine whether the organizational structure becomes more centralized as internal communication improves.

We further use the CiDB data to construct measures of centralization for both public and private firms. With the exception of those using proprietary and/or survey-based data, prior papers have primarily used firm size, foreign assets over total assets, geographic segments, or geographic

[1] We find quantitatively and qualitatively similar results if we limit our sample to firms that provided EPS forecasts every year during the last 5 years before 2001. However, this limits the sample to 583 firm-years.

proximity as a relatively crude proxy for firm (de)centralization due to data limitations (e.g., Beck et al., 2019; Robinson et al., 2010; Garrett et al., 2014).[2] The site-level CiDB data allows us to measure centralization in a slightly less crude manner. We measure centralization in two ways: first, as the proportion of revenue generated by the HQ office over the total revenue of the firm; and, second, as the proportion of employees in the HQ office over the total employees of the firm. This approach is similar to Campbell et al. (2009), which measures centralization as the proportion of corporate and supervisory staff relative to the number of store-level employees, using the Environmental Systems Research Institute (ESRI) Business Location Data.

We estimate the following regression separately for public and private firms:

$$Centralization_{i,t+1} = \beta_1 Intranet_{i,t} + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}, \tag{4}$$

where $Centralization_{i,t}$ is proxied by *% of Revenue at HQ* and *% of Employees at HQ* obtained from the site-level CiDB data. *% of Revenue at HQ* is defined as the ratio of revenue generated by the HQ office over the firm's total revenues in the year. *% of Employees at HQ* is the ratio of employees in the HQ office over the total employees of the firm in the year. These proxies roughly follow Campbell et al. (2009), which measures centralization as the proportion of corporate and supervisory staff relative to the number of store-level employees, using business location data. Our measures of centralization are based on the idea that increases in revenue or employees at the HQ office are driven by increases in the decision-making authority of the HQ office.[3] All other variables are defined the same way as Equation 24. We include the industry $\times$ year fixed effects to account for time-varying industry characteristics. The coefficient of interest is β_1. Based on the theoretical prediction, we expect $\beta_1 > 0$.

Table OA.9 Panel A and Panel B present results for public and private firms, respectively.

[2] Robinson and Stocken (2013) use a firm's declared functional currency for each of its foreign subsidiaries to measure a multinational firm's centralization of decision rights.

[3] Unless they have proprietary and/or survey-based data, prior papers have used firm size, foreign assets over total assets, geographic segments, and geographic proximity as proxies for firm (de)centralization (e.g., Beck et al., 2019; Robinson et al., 2010; Garrett et al., 2014). One exception is Robinson and Stocken (2013) who use a multinational firm's declared functional currency for their subsidiaries to measure their centralization of decision rights. Our access to the CiDB allows us to construct less noisy proxies for centralization, but we recognize and caveat the crudeness of these proxies.

Our sample sizes for both public and private firms in Panel A-B are reduced in Table OA.9 because we drop firms with missing division-level revenue and employee data. Our findings in Panel A and Panel B are consistent with our prediction as well as with the findings of Bloom et al. (2014). We find $\beta_1 > 0$ in both columns (1) and (2) across different measures of centralization. The results suggest that an improvement in internal communication technology results in more centralized decision-making, which serves as supporting evidence for our assumption of HQ managers intranet adoption.

References

Beck, M. J., Gunn, J. L., and Hallman, N. (2019). The geographic decentralization of audit firms and audit quality. Journal of Accounting and Economics, 68(1):101234.

Bloom, N., Garicano, L., Sadun, R., and Van Reenen, J. (2014). The distinct effects of information technology and communication technology on firm organization. Management Science, 60(12):2859–2885.

Campbell, D., Datar, S. M., and Sandino, T. (2009). Organizational design and control across multiple markets: The case of franchising in the convenience store industry. The Accounting Review, 84(6):1749–1779.

Chen, C., Martin, X., Roychowdhury, S., Wang, X., and Billett, M. T. (2018). Clarity begins at home: Internal information asymmetry and external communication quality. The Accounting Review, 93(1):71–101.

Garrett, J., Hoitash, R., and Prawitt, D. F. (2014). Trust and financial reporting quality. Journal of Accounting Research, 52(5):1087–1125.

Robinson, J. R., Sikes, S. A., and Weaver, C. D. (2010). Performance measurement of corporate tax departments. The Accounting Review, 85(3):1035–1064.

Robinson, L. and Stocken, P. (2013). Location of decision rights within multinational firms. Journal of Accounting Research, 51:1261–1297.

Table OA.1: Determinants of Intranet

Panel A and Panel B report results from the following firm-year-level regression estimated for public and private firms, respectively: $Intranet_{i,t+1} = \theta Determinants_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} (+\Sigma\rho_i Firm_i) + \epsilon_{i,t}$. Panel C and Panel D report results from the following firm-site-year-level regression estimated for public and private firms, respectively: $Site\text{-}level\ Intranet_{s,t+1} = \theta Site\text{-}level\ Determinants_{s,t} + \Sigma\rho_s Sites_s + \Sigma\beta_{i,t} Firm \times Year_{i,t} + \epsilon_{s,t}$. All variables are defined in Appendix B of the main manuscript. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. *t*-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Public Firm-level Determinants						
	Intranet$_{t+1}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	-0.001	-0.004	-0.002	-0.001	-0.001	-0.001
	(-0.24)	(-1.06)	(-0.42)	(-0.22)	(-0.25)	(-0.25)
ROA	0.123***	0.115**	0.120***	0.020	0.020	0.020
	(2.80)	(2.61)	(2.73)	(0.72)	(0.72)	(0.74)
BTM	0.002	0.005	0.002	-0.012	-0.011	-0.012
	(0.35)	(0.77)	(0.33)	(-1.65)	(-1.57)	(-1.66)
ln(MVE)	-0.009**	-0.006*	-0.010***	-0.003	-0.003	-0.003
	(-2.36)	(-1.78)	(-2.76)	(-0.74)	(-0.59)	(-0.74)
R&D	-0.107**	-0.086	-0.095*	-0.004	0.000	-0.002
	(-2.00)	(-1.27)	(-1.68)	(-0.08)	(0.00)	(-0.05)
ln(1+analysts following)	-0.014*	-0.014*	-0.014*	-0.002	-0.002	-0.002
	(-1.87)	(-1.88)	(-1.89)	(-0.23)	(-0.24)	(-0.22)
Earnings Volatility	0.006	0.007	0.005	-0.008	-0.008	-0.008
	(0.37)	(0.42)	(0.27)	(-0.66)	(-0.68)	(-0.69)
ln(sites)	-0.075***	-0.058***	-0.074***	-0.019*	-0.015	-0.019*
	(-9.79)	(-8.00)	(-9.76)	(-1.78)	(-1.57)	(-1.82)
Internet	0.210***	0.212***	0.210***	0.131***	0.131***	0.131***
	(5.69)	(5.95)	(5.72)	(3.92)	(3.97)	(3.93)
Number of PCs per employee	0.077***	0.065***	0.076***	0.030***	0.028***	0.030***
	(9.25)	(7.63)	(9.04)	(3.68)	(3.53)	(3.67)
ln(business segments)	0.008	0.007	0.007	0.018*	0.018*	0.018*
	(0.85)	(0.67)	(0.78)	(1.97)	(1.98)	(1.98)
ln(geographical segments)	0.011**	0.012**	0.011**	-0.004	-0.004	-0.004
	(2.24)	(2.33)	(2.30)	(-0.44)	(-0.40)	(-0.42)
Related	0.023	0.030	0.024	0.020	0.019	0.020
	(1.13)	(1.55)	(1.23)	(1.42)	(1.31)	(1.36)
Institutional ownership	0.026	0.034	0.027	-0.024	-0.024	-0.024
	(1.17)	(1.59)	(1.25)	(-0.99)	(-0.97)	(-0.98)
Loss Indicator	0.001	0.003	0.001	-0.002	-0.001	-0.002
	(0.10)	(0.23)	(0.07)	(-0.46)	(-0.31)	(-0.40)
ln(employees)	0.111***		0.078***	0.046***		0.027
	(17.01)		(7.65)	(5.09)		(1.41)
ln(1+revenues)		0.092***	0.033***		0.039***	0.019
		(15.62)	(3.44)		(5.10)	(1.17)
Firm FE	No	No	No	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	17,524	17,524	17,524	17,240	17,240	17,240
R-squared	0.345	0.339	0.347	0.749	0.749	0.749
Clustering	Industry					

Table OA.1 (cont'd): Determinants of Intranet (cont'd)

Panel B: Private Firm-level Determinants						
	Intranet$_{t+1}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance	-0.004***	-0.012***	-0.004***	0.001	-0.001	0.001
	(-3.23)	(-4.00)	(-3.55)	(0.31)	(-0.64)	(0.31)
ln(sites)	-0.123***	-0.058***	-0.125***	-0.030***	-0.011**	-0.029***
	(-11.58)	(-2.94)	(-11.50)	(-4.82)	(-2.51)	(-4.82)
Internet	0.432***	0.493***	0.431***	0.150***	0.155***	0.150***
	(10.72)	(12.68)	(10.87)	(4.33)	(4.53)	(4.32)
Number of PCs per employee	0.136***	0.083***	0.138***	0.015***	0.010**	0.016***
	(6.96)	(6.98)	(8.18)	(2.82)	(2.04)	(2.89)
ln(employees)	0.173***		0.149***	0.025***		0.027***
	(15.52)		(20.27)	(5.64)		(5.81)
ln(1+revenues)		0.127***	0.026***		0.003*	-0.002
		(8.38)	(3.05)		(1.92)	(-1.45)
Firm FE	No	No	No	Yes	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	155,116	155,116	155,116	148,649	148,649	148,649
R-squared	0.339	0.298	0.341	0.901	0.901	0.901
Clustering	Industry					

Panel C: Within-Public-Firm Site-Level Determinants			
	Site-level Intranet$_{t+1}$		
	(1)	(2)	(3)
Distance between HQ and site	-0.001	-0.001	-0.001
	(-1.06)	(-1.07)	(-1.06)
Site-level Internet	0.038***	0.040***	0.038***
	(4.92)	(5.06)	(4.91)
Number of PCs at site per employee	0.019***	0.014***	0.019***
	(5.90)	(5.10)	(6.00)
ln(site-level employees)	0.017***		0.019***
	(4.62)		(5.14)
ln(1+site-level revenues)		0.003*	-0.002*
		(1.81)	(-1.89)
Site FE	Yes	Yes	Yes
Firm × Year FE	Yes	Yes	Yes
N	667.230	667.230	667.230
R-squared	0.903	0.902	0.903
Clustering	Industry		

Table OA.1 (cont'd): Determinants of Intranet (cont'd)

Panel D: Within-Private-Firm Site-Level Determinants			
	Site-level Intranet$_{t+1}$		
	(1)	(2)	(3)
Distance between HQ and site	0.000	0.000	0.000
	(0.16)	(0.16)	(0.16)
Site-level Internet	0.012	0.013	0.012
	(1.39)	(1.44)	(1.39)
Number of PCs at site per employee	0.010***	0.008***	0.010***
	(5.29)	(4.47)	(5.27)
ln(site-level employees)	0.011***		0.011***
	(4.50)		(4.48)
ln(1+site-level revenues)		0.002***	0.001**
		(5.98)	(2.02)
Site FE	Yes	Yes	Yes
Firm × Year FE	Yes	Yes	Yes
N	1,306,200	1,306,168	1,306,168
R-squared	0.933	0.933	0.933
Clustering		Industry	

Table OA.2: Linear Association between Intranet and Forecast Frequency

This table reports estimates from the following firm-year-level regression: $ln(1 + EPS\ forecasts)_{i,t+1} = \beta_1 Intranet_{i,t} + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t}(+\Sigma\rho_i Firm_i) + \epsilon_{i,t}$. All variables are defined in Appendix B of the main manuscript. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	ln(1+EPS forecasts)			
	(1)	(2)	(3)	(4)
Intranet	0.258***	-0.013	-0.012	-0.014
	(5.55)	(-0.44)	(-0.26)	(-0.28)
ROA		0.004		0.055
		(0.03)		(0.67)
BTM		-0.046**		0.017
		(-2.28)		(0.91)
ln(MVE)		0.044***		0.116***
		(3.27)		(6.17)
R&D		0.157		-0.164
		(0.56)		(-0.87)
ln(1+analysts following)		0.142***		0.062***
		(4.61)		(3.05)
Earnings Volatility		-0.181***		-0.075**
		(-4.37)		(-2.38)
ln(employees)		0.066		0.014
		(1.39)		(0.42)
ln(1+revenues)		-0.005		-0.020
		(-0.13)		(-0.76)
ln(sites)		0.004		-0.003
		(0.16)		(-0.16)
Internet		0.021		-0.119**
		(0.31)		(-2.35)
Number of PCs per employee		0.031		-0.003
		(0.68)		(-0.17)
ln(business segments)		0.084***		0.180***
		(3.57)		(6.42)
ln(geographical segments)		0.038**		0.118***
		(2.49)		(4.62)
Related		0.037		-0.007
		(0.69)		(-0.22)
Institutional ownership		0.310***		-0.057
		(3.87)		(-1.02)
Loss Indicator		-0.176***		-0.084***
		(-7.68)		(-4.15)
Firm FE	No	No	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
N	20,778	20,778	20,396	20,396
R-squared	0.079	0.261	0.709	0.726
Clustering	Industry			

Table OA.3: Linear Association between Intranet and Website Disclosure

Panel A and Panel B report results from the following firm-year-level regressions estimated for public firms and private firms, respectively: $ln(Website\ length)_{i,t+1} = \beta_1 Intranet_{i,t} + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} (+\Sigma\rho_i Firm_i) + \epsilon_{i,t}$. All variables are defined in Appendix B of the main manuscript. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Public Firms				
	ln(Website length)$_{t+1}$			
	(1)	(2)	(3)	(4)
Intranet	-0.109**	-0.114**	-0.096	-0.111
	(-2.01)	(-2.12)	(-1.26)	(-1.46)
ROA		-0.097		-0.117
		(-0.76)		(-0.98)
BTM		-0.070***		-0.015
		(-3.02)		(-0.51)
ln(MVE)		-0.020		0.048
		(-1.26)		(1.60)
R&D		0.455***		0.552**
		(4.68)		(2.07)
ln(1+analysts following)		0.042*		0.057**
		(1.79)		(2.01)
Earnings Volatility		-0.150***		-0.045
		(-2.75)		(-0.98)
ln(employees)		0.044		0.058
		(1.08)		(0.99)
ln(1+revenues)		-0.009		0.042
		(-0.25)		(0.96)
ln(sites)		-0.047		-0.112***
		(-1.66)		(-2.86)
Internet		-0.030		-0.079
		(-0.37)		(-0.69)
Number of PCs per employee		0.067		0.039
		(1.17)		(0.46)
ln(business segments)		-0.001		-0.015
		(-0.03)		(-0.36)
ln(geographical segments)		0.003		0.011
		(0.20)		(0.33)
Related		-0.032		-0.088
		(-0.56)		(-1.37)
Institutional ownership		0.050		-0.023
		(1.17)		(-0.28)
Loss Indicator		0.004		0.020
		(0.13)		(0.89)
Firm FE	No	No	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
N	17,600	17,600	17,204	17,204
R-squared	0.057	0.064	0.472	0.475
Clustering	Industry			

Table OA.3 (cont'd): Linear Association between Intranet and Website Disclosure (cont'd)

Panel B: Private Firms				
	$\ln(\text{Website length})_{t+1}$			
	(1)	(2)	(3)	(4)
Intranet	0.087***	0.018	0.008	0.008
	(5.93)	(1.51)	(0.39)	(0.39)
ln(employees)		0.035***		0.003
		(4.74)		(0.32)
ln(1+revenues)		0.010		-0.005*
		(1.27)		(-1.72)
ln(sites)		-0.007		-0.007
		(-0.53)		(-0.41)
Internet		-0.164***		0.069
		(-2.69)		(0.90)
Number of PCs per employee		0.048*		0.008
		(1.96)		(0.98)
Firm FE	No	No	Yes	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
N	225,425	225,425	211,689	211,689
R-squared	0.078	0.081	0.662	0.662
Clustering	Industry			

Table OA.4: Role of Information Learning: Website Disclosure

This table reports results from the following firm-year-level regressions estimated separately for firm-years above and below the median cutoff: $ln(Website\ length)_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. Panel A and Panel B report results for public firms and private firms, respectively. In Panel A-B, columns (1)-(2) split the sample based on median *Distance*. *Distance* is the average distance between the HQ office and divisional offices weighted by employees at each divisional site. In Panel A, controls include: ROA, BTM, ln(MVE), R&D, ln(1+analysts following), Earnings Volatility, ln(employees), ln(1+revenues), ln(sites), Internet, Number of PCs per employee, ln(business segments), ln(geographical segments), Related, Institutional ownership, and Loss Indicator. In Panel B, controls include: ln(employees), ln(1+revenues), ln(sites), Internet, and Number of PCs per employee. All variables are defined in Appendix B of the main manuscript. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. *t*-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Public Firms

	ln(Website length)$_{t+1}$	
	Below-Median Distance	Above-Median Distance
	(1)	(2)
Intranet	0.348	0.375*
	(1.60)	(1.64)
Intranet^2	-0.378	-0.738***
	(-1.56)	(-2.74)
Controls	Yes	Yes
Industry $\times$ Year FE	Yes	Yes
N	8,704	8,747
R-squared	0.067	0.071
Clustering	Industry	

Panel B: Private Firms

	ln(Website length)$_{t+1}$	
	Below-Median Distance	Above-Median Distance
	(1)	(2)
Intranet	0.107	0.228***
	(1.61)	(3.62)
Intranet^2	-0.100	-0.241***
	(-1.40)	(-3.66)
Controls	Yes	Yes
Industry $\times$ Year FE	Yes	Yes
N	112,521	112,778
R-squared	0.077	0.084
Clustering	Industry	

Table OA.5: Robustness Test on the Role of Information Learning: Forecasts

This table reports results from the following firm-year-level regressions estimated separately for firm-years above and below the median cutoff: $ln(1 + EPS\ forecasts)_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \gamma Controls^2_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. Columns (1)-(2) split the sample based on median *Distance*. Distance is the average distance between the HQ office and divisional offices weighted by employees at each divisional site. Controls include: ROA, BTM, ln(MVE), R&D, ln(1+analysts following), Earnings Volatility, ln(employees), ln(1+revenues), ln(sites), Internet, Number of PCs per employee, ln(business segments), ln(geographical segments), Related, Institutional ownership, and Loss Indicator. All variables are defined in Appendix B of the main manuscript. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	ln(1+EPS forecasts)$_{t+1}$	
	Below-Median Distance (1)	Above-Median Distance (2)
Intranet	0.035	0.615***
	(0.22)	(3.99)
Intranet^2	-0.024	-0.581***
	(-0.14)	(-3.29)
Controls	Yes	Yes
Controls^2	Yes	Yes
Industry × Year FE	Yes	Yes
N	10,299	10,371
R-squared	0.295	0.284
Clustering	Industry	

Table OA.6: Robustness Test on the Role of Information Learning: Website Disclosure

This table reports results from the following firm-year-level regressions estimated separately for firm-years above and below the median cutoff: $ln(Website\ length)_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \gamma Controls^2_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. Panel A and Panel B report results for public firms and private firms, respectively. In Panel A-B, columns (1)-(2) split the sample based on median *Distance*. *Distance* is the average distance between the HQ office and divisional offices weighted by employees at each divisional site. In Panel A, controls include: ROA, BTM, ln(MVE), R&D, ln(1+analysts following), Earnings Volatility, ln(employees), ln(1+revenues), ln(sites), Internet, Number of PCs per employee, ln(business segments), ln(geographical segments), Related, Institutional ownership, and Loss Indicator. In Panel B, controls include: ln(employees), ln(1+revenues), ln(sites), Internet, and Number of PCs per employee. All variables are defined in Appendix B of the main manuscript. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Public Firms

	ln(Website length)$_{t+1}$	
	Below-Median Distance (1)	Above-Median Distance (2)
Intranet	0.313	0.244
	(1.44)	(1.03)
Intranet^2	-0.33	-0.625**
	(-1.39)	(-2.40)
Controls	Yes	Yes
Controls^2	Yes	Yes
Industry × Year FE	Yes	Yes
N	8,704	8,747
R-squared	0.074	0.074
Clustering	Industry	

Panel B: Private Firms

	ln(Website length)$_{t+1}$	
	Below-Median Distance (1)	Above-Median Distance (2)
Intranet	0.094	0.185***
	(1.58)	(3.48)
Intranet^2	-0.086	-0.194***
	(-1.30)	(-3.39)
Controls	Yes	Yes
Controls^2	Yes	Yes
Industry × Year FE	Yes	Yes
N	112,521	112,778
R-squared	0.078	0.085
Clustering	Industry	

Table OA.7: Robustness Test on the Role of Free Riding: Forecasts

This table reports results from the following firm-year-level regressions estimated separately for firm-years above and below the median cutoff: $ln(1 + EPS\ forecasts)_{i,t+1} = \beta_1 Intranet_{i,t}\ (+\beta_2 Intranet^2_{i,t}) + \theta Controls_{i,t}\ (+\gamma Controls^2_{i,t}) + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. Columns (1)-(2) and columns (3)-(4) split the sample based on median $Site\ Similarity_{i,t-1}$. $Site\ Similarity_{i,t-1}$ captures the extent of similarity across a firm's divisional sites calculated by the sum of the square of the share of each 2-digit SIC industry. Controls include: ROA, BTM, ln(MVE), R&D, ln(1+analysts following), Earnings Volatility, ln(employees), ln(1+revenues), ln(sites), Internet, Number of PCs per employee, ln(business segments), ln(geographical segments), Related, Institutional ownership, and Loss Indicator. All variables are defined in Appendix B of the main manuscript. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. *t*-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	$\ln(1+\text{EPS forecasts})_{t+1}$			
	Below-Median Site Similarity		Below-Median Site Similarity	
	(1)	(2)	(3)	(4)
Intranet	0.231***	0.514	-0.093	0.618***
	(3.29)	(1.39)	(-1.42)	(2.72)
Intranet^2		-0.229		-0.706***
		(-0.75)		(-3.19)
Controls	Yes	Yes	Yes	Yes
Controls^2	No	Yes	No	Yes
Industry × Year FE	Yes	Yes	Yes	Yes
N	7,103	7,103	7,729	7,729
R-squared	0.249	0.271	0.260	0.295
Clustering	Industry			

Table OA.8: Non-linear Association between Intranet and Forecast Accuracy

This table reports estimates from the following firm-year-level regression: $EPS\ Forecast\ Accuracy_{i,t+1} = \beta_1 Intranet_{i,t} + \beta_2 Intranet^2_{i,t} + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. *EPS Forecast Accuracy* is the absolute difference of the management EPS forecast and actual realized EPS, based on I/B/E/S. We consider EPS forecasts issued after and on the earnings announcement date corresponding to the current fiscal year and before the earnings announcement for the subsequent fiscal year. The average of absolute differences is used when a firm provides more than one EPS forecast during the period. All variables are defined in Appendix B of the main manuscript. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	EPS Forecast Accuracy$_{t+1}$	
	(1)	(2)
Intranet	1.491**	1.116*
	(2.30)	(1.87)
Intranet^2	-1.429**	-1.051**
	(-2.40)	(-1.98)
ROA		-1.843***
		(-2.74)
BTM		0.145
		(1.21)
ln(MVE)		0.089**
		(2.46)
R&D		0.414
		(0.52)
ln(1+analysts following)		-0.094**
		(-2.05)
Earnings Volatility		0.033
		(0.23)
ln(employees)		0.000
		(0.00)
ln(1+revenues)		-0.024
		(-0.21)
ln(sites)		0.067
		(0.82)
Internet		0.045
		(0.17)
Number of PCs per employee		0.133
		(1.27)
ln(business segments)		-0.099
		(-1.34)
ln(geographical segments)		0.130**
		(2.30)
Related		-0.209
		(-0.69)
Institutional ownership		0.029
		(0.27)
Loss Indicator		-0.104
		(-0.91)
Industry × Year FE	Yes	Yes
N	2,152	2,152
R-squared	0.024	0.068
Clustering	Industry	

Table OA.9: Relations between Intranet Adoption and Decision Centralization

Panel A and Panel B report results from the following firm-year-level regression estimated for public and private firms, respectively: $Centralization_{i,t+1} = \beta_1 Intranet_{i,t} + \theta Controls_{i,t} + \Sigma\beta_{j,t} Industry \times Year_{j,t} + \epsilon_{i,t}$. $Centralization$ is measured as $\%\ of\ Revenue\ at\ HQ$ in column (1) and as $\%\ of\ Employees\ at\ HQ$ in column (2). $\%\ of\ Revenue\ at\ HQ$ is the ratio of revenue generated by the HQ office over the total revenues of the firm. $\%\ of\ Employees\ at\ HQ$ is the ratio of employees in the HQ office over the total employees of the firm. All variables are defined in Appendix B of the main manuscript. All continuous variables are winsorized at the 1% and 99% levels to limit the influence of outliers. t-statistics based on standard errors clustered by industry are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Public Firms		
	% of Revenue at HQ_{t+1} (1)	% of Employees at HQ_{t+1} (2)
Intranet	0.054***	0.050***
	(2.93)	(2.91)
ROA	0.363***	0.355***
	(7.97)	(8.88)
BTM	-0.017	-0.019*
	(-1.51)	(-1.74)
ln(MVE)	-0.037***	-0.038***
	(-7.38)	(-7.76)
R&D	0.897***	0.908***
	(7.60)	(8.88)
ln(1+analysts following)	0.016*	0.016*
	(1.68)	(1.75)
Earnings Volatility	-0.022	-0.017
	(-1.65)	(-1.34)
Institutional ownership	-0.030	-0.030
	(-1.32)	(-1.40)
Internet	-0.256***	-0.259***
	(-8.37)	(-8.39)
Number of PCs per employee	0.133***	0.136***
	(6.41)	(6.44)
Loss Indicator	0.021***	0.015*
	(2.78)	(1.99)
Industry × Year FE	Yes	
N	19,305	19,351
R-squared	0.208	0.245
Clustering	Industry	

Panel B: Private Firms		
	% of Revenue at HQ_{t+1} (1)	% of Employees at HQ_{t+1} (2)
Intranet	0.069***	0.066***
	(6.61)	(5.26)
ln(employees)	0.023**	0.041***
	(2.43)	(2.75)
ln(1+revenues)	0.037***	0.019
	(3.90)	(1.10)
ln(sites)	-0.248***	-0.256***
	(-25.92)	(-31.22)
Internet	-0.121*	-0.157***
	(-1.96)	(-2.68)
Number of PCs per employee	0.056	0.063
	(1.16)	(1.27)
Industry × Year FE	Yes	
N	166,099	170,922
R-squared	0.340	0.401
Clustering	Industry	