Managerial Learning from Decoding Noisy Stock Prices: New(s) Evidence*

Alan Kwan, Tse-Chun Lin, and Po-Yu Liu†

This Version: October 2022

Abstract

A long literature argues corporate managers learn from stock prices, but organizations' learning process is challenging to observe. We present a novel test using firm-level readership of financial media articles as a manifestation of managerial learning. We hypothesize that reading financial media helps managers with the interpretation of noisy signals in stock prices. We show that the classic Q-sensitivity of R&D expenditure increases by 26% when firms' reading of financial articles increases by one standard deviation. This relationship is mainly driven by reading from near the headquarters where managers are likely located and by articles likely more informative to managers.

Keywords: big data, managerial learning, market feedback effects, financial news, R&D.
JEL Classification Codes: C55, G11, G23, G24, Q01.

* We are grateful for helpful comments from Wei Jiang, Gaurav Kanakanhalli, Shiyang Huang, and seminar participants at the University of Hong Kong. All remaining errors are ours.
† Faculty of Business and Economics, The University of Hong Kong. E-mail address of Alan Kwan: apkwan@hku.hk. E-mail address of Tse-Chun Lin: tsechunlin@hku.hk. E-mail address of Po-Yu Liu: liupoyu@connect.hku.hk.

Learning is a hypothetical construct: It cannot be directly observed but only inferred from observable behavior.

Richard Gross (2010)

1. Introduction

Is the stock market a sideshow, or does the stock market have real effects on corporate decisions? If so, how? These questions have been the subject of vigorous debate in the finance literature for decades. Chen, Goldstein, and Jiang (2007) put forth the possibility that managers *learn* from stock prices. Since then, the literature has studied how investment sensitivity to Q increases with exogenous variations in price informativeness to indirectly argue that managers learn from the market (Bond, Edmans, and Goldstein (2012)). However, while existing evidence is in line with the managerial learning mechanism, learning is not observed directly in prior studies, making it empirically challenging to verify whether it takes place.

To that end, in this paper, we introduce a novel and highly granular dataset of firms' internet financial media readership. How does reading financial media factor into the theory of learning from stock prices? Suppose that, as in the canonical narrative of Chen, Goldstein, and Jiang (2007), managers rely on the stock price to help make investment decisions. However, price movements are noisy (Bakke and Whited (2007)), and managers need to interpret those movements to extract useful information outside their information set. We hypothesize that reading financial media can be one mechanism through which managers interpret stock price movements and learn from stock prices (see Figure 1 for our conceptual framework).

Financial journalists work to supply investors with information, conducting both primary and secondary research (e.g., Veldkamp (2006a, 2006b) and Engelberg and Parsons (2011)). If financial media serve to aggregate information, it stands to reason that managers may rely on the media to decode information in their stock prices. Thus, we would expect to see a greater investment-sensitivity of Q when managers more intensely acquire information about their own stock through the financial media, as such reading either brings new information unto itself or even if not, helps the manager to synthesize market views and interpret stock prices.

Before providing details, we offer two caveats. First, firms have many reasons to read the news, not just to decode stock prices. Second, reading the financial media is not exhaustive of all ways financial markets may provide information to managers and guide their decisions.[1] But, as long as organizational news consumption captures a substantial set of learning activities, or if financial media consumption is correlated with learning activity overall, then our novel empirical setting helps overcome a major gap in the managerial learning literature by providing a direct measure of learning behavior.

To start, we construct a dataset of firm-quarter-level readership of financial media articles. To obtain this dataset, we collaborate with a company ("the Data Partner") operating in the digital marketing space which focuses on developing metrics of what firms read about (and thus their potential purchasing interests) across the web. The Data Partner orchestrates a large partnership network of business media publishers, including, but not limited to, some of the world's largest financial media websites. As part of the Data Partner's core business, it has developed a robust method to associate website visitors with firms based on a variety of different association methods at the visitor level. Using this setting, we can link specific news articles to the firms that read the news if they are part of the Data Partner's publisher network. We identify 12 major financial media sites in this dataset, including four key household names in financial media, for which we observe content interactions by firms at the article level. This data is *not* sold publicly and is made specially available to us as part of academic research.

Our analysis proceeds in three steps. First, we construct detailed measures of managerial learning and document correlations between readership and firm characteristics to

[1] For example, managers could learn from the market via private interactions with investors, analysts, and bankers, as well as public interactions over conference calls.

better understand our measures. To do so, we merge our database with RavenPack News Analytics, a database of analytics of financial news articles tailored specifically for the financial industry, which includes (i) the companies mentioned in specific articles, (ii) the topic of the financial news article as relevant to a financial participant, (iii) the sentiment of this article, and (iv) a host of other rich features, which allow us a flexible decomposition of news into several types. In particular, we decompose the consumption of news into three categories, including (i) negative versus positive news, (ii) news about a focal manager's firm, her peers, or general news, and (iii) articles that are full-length or press releases.

We begin by understanding patterns in firm news reading to help assess why firms read the news. First, we plot the temporal distribution of reading. We find that on weekdays, reading climbs during work hours and falls at the end of the workday, and on weekends, there is no such climb. Second, we regress our main measure of reading - the number of reading for financial news articles scaled by firm assets - against firm characteristics. The recent literature suggests that various economic agents such as investors exhibit limited attention (e.g., Kacperczyk, Nieuwerburgh, and Veldkamp (2016)). If corporate managers' attention capacities are also limited, firms should pay attention to news of the variety that is most likely to impact their business.

Interesting and intuitive relationships emerge from this exercise. For example, firms read the news more if their stock prices are more volatile, consistent with volatility serving as a proxy for events that consume managers' attention (Gondhi (2022)). Firms that face more product market competition are more likely to read *peer* news. Such managers are more likely to benefit from reading about peer firm news in the face of more severe competition. We also find that profitable firms acquire more information (Charoenwong, Kimura, Kwan, and Tan (2022)). While firms may pay attention to the news for various reasons (Bond, Edmans, and

Goldstein (2012)), our findings suggest that incentives drive the types of articles read, implying information acquisition could be a predominant motive.

In the second (main) phase of our analysis, we test whether corporate investment is more sensitive to market valuation when managers learn more about their own firms from financial news media. In other words, we expect that the investment sensitivity to Tobin's Q should be more positive when reading activity is higher. We follow Bai, Phillipon, and Savov (2016) and use R&D expenditure as the main investment measure as there is more room for managerial learning given the uncertain nature of R&D. We find that a one-standard-deviation increase in log-ratio of financial news reading about own firm corresponds to an increase in Q-sensitivity of R&D by 26%. We find a qualitatively similar result of 16% when we measure investment by R&D plus CAPX. Our finding is robust to a large array of controls, such as firm fixed effects, quarter fixed effects, as well as control variables that could reflect private information by managers, such as analyst coverage, and managerial private information (Chen, Goldstein, and Jiang (2007)). Our results remain robust when we also control for news supply to mitigate the concern that news coverage mechanically represents the flow of news.

Our finding also indicates that managerial learning from the news is indeed weaker for capital expenditure than for R&D (and, as we will show later, SG&A). One major reason could be that the pay-off on spending on intangible capital may be relatively uncertain compared to pure capital expenditures (Bai, Phillipon, and Savov (2016), Peters and Taylor (2017), and Andrei, Mann, and Moyen (2019)). Hence, it is not surprising that firms learn more about these types of investments from the market. We also follow Andrei, Mann, and Moyen (2019) and consider the impact of the measurement error of Q on our results. However, we do not find evidence that measurement error would materially affect our inferences.[2]

[2] See Section 4.3 and Internet Appendix B.

To help pin down the mechanism, we decompose reading geographically. While we cannot personally identify the news "reader" in our dataset individually, we assume that managers and employees who are delegated by managers to collect information read from the headquarters metropolitan statistical area (MSA) of the firm.[3] In this case, we expect stronger results for corporate reading near firm headquarters. Partitioning reading in the headquarters MSA versus outside of it, we find that the learning effect is only significant when considering articles read at the headquarters. This finding corroborates our argument that the positive interaction term of reading and Q in our regressions represents organizational learning.

Existing studies in the managerial learning literature generally acknowledge a major alternative explanation for the price-feedback hypothesis: Q-sensitivity could be a proxy for financial constraints being alleviated (e.g., Hennessy, Levy, and Whited (2007)). All else being equal, higher firm valuations imply lower expected returns and investment hurdle rates, justifying higher investment levels. To examine the role of financial constraints in our findings, we examine whether managerial learning differs among firms that are financially constrained.[4] We find firms exhibiting similar managerial learning effects irrespective of their level of financial constraints. Hence, the documented relationship between corporate reading and investment sensitivity to Q is unlikely to be driven by this alternative explanation.

Having established the main result, we next examine the types of information managers acquire from financial news to adjudicate various theories of managerial learning. To do so, we rely on the rich features of RavenPack. We present two sets of these tests. First, we decompose news articles into articles about one's own firm, industry peers, or general news.

[3] The notion that the headquarters is the location of key decision-making activity is implicit within the literature of strategy, accounting, finance, and economics (e.g., Malenko (2019) and Dessein, Galleoti, and Santos (2016)). Large corporations often have analysts at the headquarters to track market conditions and make forecasts of market demand. We posit that such workers would likely work near managers as it is easier to report to managers.

[4] To measure financial constraints, we adopt a wide range of financial constraint proxies, including the size-age index (Hadlock and Pierce (2010)), the four-variable KZ index (Kaplan and Zingales (1997) Baker, Stein, and Wurgler (2003)), the WW index (Whited and Wu (2006)), whether a firm is a dividend payer, and a firm's market capitalization. We separately analyze the interaction of these proxies as well as two composite indices of financial constraints created from these five individual proxies.

Traditionally, theories of learning from stock prices focus on firm-specific news. However, recent contributions to the literature have explored the possibility that firms learn from peers as well (e.g., Foucault and Fresard (2014), Dessaint, Foucault, Fresard, and Matray (2021), and Yan (2021)). Moreover, another stream of literature also explores the role of aggregate versus idiosyncratic shocks.[5] Identifying which types of financial news managers learn from the market sheds light on the types of information comprising the feedback effect.

On a standalone basis, we find that reading about either a focal firm's own news or its industry peer news amplifies the investment sensitivity to Q, consistent with our main result. However, in a horse race, the strongest component of the learning effect comes from reading about one's own news, either we measure investment by R&D or R&D plus CAPX. Consistent with Foucault and Fresard (2014) and Bai, Phillipon, and Savov (2016), managers also learn from industry peer news when we measure investment by R&D expenditure alone. Moreover, when we evaluate the relative economic magnitudes of the effect, the firm-specific channel is twice as strong as the industry-peer channel. Managers seem not to learn from general news.

Next, we examine whether firms learn more from positive news or non-positive news. The feedback hypothesis suggests that managers learn more from the market when their stock valuation contains novel information outside their information set. Under this hypothesis, it is plausible that managers might learn more from neutral or negative news. This is perhaps because positive news may affirm one's priors instead of offering new information. In the data, we find that the effect is driven by negative or neutral news, consistent with our expectations.

Finally, we decompose news into articles that are full-length articles versus press releases and news flashes. Intuitively, a full-length article is more likely to convey an opinion or detailed analysis of a firm's activities, which is more likely to provide novel information

[5] See Foerster, Sarte, and Watson (2011), Acemoglu, Carvalho, Ozdaglar, and Tabbaz-Salehi (2012), Gabaix (2011), and Gondhi (2022).

that managers do not possess. In contrast, press releases or news flashes are less likely to provide managers with new information since the firm is presumably aware of its own press releases. Consistent with our conjecture, we find that when firms read news about their own firms, they learn more from full-length articles.

Our results provide novel and direct evidence of managerial learning, one of the most studied questions in the literature.[6] A key concern with managerial learning literature is that higher Q may alleviate financial constraints (Hennessy, Levy, and Whited (2007)), which could confound the interpretation of investment sensitivity to Q. Learning is not observed directly. Instead, prior work focuses more on how Q-sensitivity becomes higher when stock prices are likely to be more informative.[7] Crucially, our paper differs from these studies on price informativeness by capturing the managerial learning process in the form of reading the financial news which managers plausibly use to help them interpret movements in stock prices.

Moreover, we also show the *composition* of the information managers acquire from capital markets: both firm-specific and to a lesser extent peer firms' news. Overall, the main takeaway of our paper is that learning by reading helps managerial decisions on R&D expenditure, the riskier investment decisions where arguably managerial learning could be most beneficial. Finally, while not exhaustive of all kinds of learning, our paper suggests that financial media may be one channel where managers acquire information from the market to interpret prices. This then contributes also to the literature on news media in financial markets, which focuses on the role media plays in informing investors.[8] Incremental to this literature, we show that managers learn from the media as well.

---

[6] A handful of these papers include Chen, Goldstein, and Jiang (2007), Bakke and Whited (2010), Goldstein and Guembel (2008), Bond, Edmans, and Goldstein (2012), Foucault and Fresard (2012), Foucault and Fresard (2014), Peress (2014b), Edmans, Goldstein, and Jiang (2015), and Bai, Phillipon, and Savov (2016).

[7] For example, Goldstein, Yang, and Zuo (2020) use staggered implementations of the EDGAR system as a shock to information dissemination and show that better dissemination of corporate disclosures could crowd out investors' private information acquisition, thereby reducing managerial learning. Also see Fernandes and Ferreira (2009), Lin, Liu, and Sun (2019), Fox, Kim, and Schonberger (2021), and Brogaard, Shi, Wei, and You (2022).

[8] For example, see Peress (2008), Fang and Peress (2008), and Peress (2014a).

The rest of this paper proceeds as follows. Section 2 discusses our data. Section 3 analyzes our reading measures. Section 4 presents the main relation between investments, Q, and reading. Section 5 decomposes the types of reading firms do. Section 6 concludes and discusses potential economic implications of our findings.

2. Institutional Setting

2.1 Dataset of Financial News

To construct metrics of information acquisition, we need to observe a large set of firms and the subject of the content they acquire information about. Our proxy for information acquisition is internet content. To measure readership of internet content, we partner with a data analytics company ("the Data Partner") from the marketing technology space. Figure 2 provides a data diagram to describe our data merge process.

[Figure 2 about here]

To obtain the abovementioned data, the Data Partner maintains a large network of partnerships with online business content publishers. In partnering with these publishers, the Data Partner collects data on the visitors to partner websites, keeping track of the visitor and the content they consume. Overall, the platform aggregates around *one billion* content consumption events per day. From this large dataset, the Data Partner creates an analytics product that aims to quantify what specific business topics certain companies are reading by visiting websites. This data is primarily sold to companies to facilitate sales and marketing. By identifying companies with heightened research interest in specific business topics, suppliers in principle may be able to narrow down likelier customers for the product or service. The Data Partner's services are not sold to financial institutions for trading or any other purposes as far as we are aware. We have access to this data as part of academic collaboration.

Participating publishers contribute to the Data Partner's pooled dataset via a technology mechanism, which shares information about web content consumption, including the external IP address of the network originating the HTTP request and the URL of content accessed. Participating publishers receive some of the Data Partner analytics in return. The process is briefly described as follows.

First, the Data Partner maps visitors to publisher websites to the Company they work for when such an association can be inferred. For each visitor to a Data Partner website, the Data Partner creates a profile through the use of first and third-party cookies.[9] This profile enables the publisher, and in turn, the Data Partner, to observe when a visitor returns to a website. Over time, the Data Partner infers the association between the profile and the Company through a wide ensemble of industry-accepted methods. For example, user profiles are associated with a company when visitors use a work email to log into a member's website. Another example is through IP addresses. That is, if a profile consistently logs onto a publisher's website from a work-associated IP address, this trace gives a strong association the profile belongs to a particular company.

IP-based methods constitute about half of identity resolution events in the sample. IP-based methods used in industry practice differ from those used in prior academic literature.[10] Once a user has been associated with a company, they are associated with that company even if they appear on other IP addresses (i.e. mobile, home, etc) as first- and third-party cookies permit the website to identify the visitor coming back again, even if the visitor is at another location. Visitor-level identification is also provided by third-party sources. Once a cookie has

[9] Over the sample period, and even as of the time of writing, the vast majority of users consent to cookies even in countries that were spearheading GPDR.

[10] Speaking generally, and not of the Data Partner per se, various data collection methods allow for an association of IP addresses to company. Large companies tend to self-identify their IP addresses via domain name service records. This methodology has been used in prior research (e.g. Chen, Cohen, Gurun, Lou, and Malloy (2020)). For smaller companies, third-party databases infer the owner of an IP address through various underlying data sources beyond DNS records. For example, websites that send marketing emails to work emails would record the IP address when clicked. This is one example of methods used to identify IP addresses belonging to specific companies beyond DNS records.

been resolved by the third-party data provider, in theory, any subscriber of that third-party source can reference that third-party database for demographic details. Cookie-based identity resolution is particularly relied upon during the period of remote work and is generally considered higher fidelity overall. It has also led the Data Partner to claim that their data is not disrupted by the pandemic.[11]

Through its proprietary processes, the Data Partner assembles these various sources of data and determines whether a reliable association between a profile and a Company can be inferred, and when it can be, what that association is. From this dataset, we construct a metric of how much a reader reads about a particular company on a given day. To start, we obtain "event-level" data from the Data Partner from the period of Nov 2017 to July 2021. Event-level data describes an instance of an article being consumed, including the timestamp, company associated with the visitor, and the URL being read. The data are scrubbed of personally identifiable information and are accessed remotely.

From the event dataset, we focus on a little over a dozen major financial publishers. The names of the publishers cannot be disclosed but are among some of the largest financial publishers in the world, with daily event counts totaling well over 100 million observations across the sites. Although not exclusively, these platforms are primarily English-language and based in North America, some of which constitute household names among financial investors. The resources include both subscription-based and free websites.

Next, we merge with news data from RavenPack in order to identify the financial news topic, subject, and sentiment of articles in our dataset. Historically, the Data Partner has performed analytics on the content firms read, but it has not recorded the companies referenced in an article in the comprehensive manner required for our analysis. Therefore, we obtain these

[11] The Data Partner's core product actually became more product during the pandemic, which consists of data analytics to help companies identify buying interests of specific companies.

features from RavenPack because it is a leading data analytics provider for financial services, which is widely used by high-frequency trading firms, hedge funds, banks, and asset managers and has been relied upon in a variety of academic studies.

We access data from RavenPack 1.0, which includes RavenPack's most detailed offering inclusive of major financial publishers it licenses content from as well as open-access content across the blogs, social media posts, news sites, and regulatory filings. We describe our merge process in the Appendix. In principle, we aim to match the article from RavenPack that is most likely to be the article that is read by the end-user. We retain the matches we believe to be reliable. The Data Partner does not retain the raw content of each article at the time it is read. Thus, we rely on fuzzy matching based on a variety of data sources to find the article. Thus, our best efforts involve matching the URL of the article to the headline of RavenPack, a process that involves multiple intermediate steps.

Once matched to RavenPack, we leverage the rich features of RavenPack to inform our analysis. Most importantly, RavenPack identifies the subjects of each article, including any mentioned firms, as well as concepts from politics, government, the economy, commodities, products, or entities that might be relevant to a financial reader. In total, RavenPack has 140,000 different entities. Given that an article can mention many entities, and sometimes only tangentially, we rely on RavenPack to help us identify the most relevant entities in each article. We apply a cut-off relevance score of 90 (out of 100) to identify those firms which are most relevant within an article. A score of at least 90 implies that the firm is mentioned in the headline of an article, and relevance declines further down into the text that a firm is mentioned. While this filter substantially reduces the number of articles we could use, we have run the analysis using lower thresholds, such as 70, and found qualitatively similar but weaker results.[12] This result affirms our choice of a strict threshold.

In addition, RavenPack has a number of measures that we leverage in our decomposition analyses. First, we obtain their event sentiment scores. These sentiment scores allow us to classify whether an article delivers positive or negative news. Second, we use their classification of articles into articles which are full articles versus press releases. These two types of features are populated for the majority of articles in our sample, allowing us to perform comparisons across types of articles.

2.2 Other data

First, we obtain standard financial data: Compustat Quarterly on financial data, analyst coverage from the Institutional Broker Estimates System (I/B/E/S), and insider trading from the Refinitiv Insiders database.

Second, we obtain industry peers from the Hoberg and Phillips (Hoberg and Phillips (2016)) Textual Network Industry Classifications (TNIC) database, which is available through 2019 and describes the similarity between firms based on textual classifications as well as a set of industries implied by these product descriptions. TNIC's baseline version, the one used in this paper, aims at matching the granularity of the three-digit Standard Industrial Classification. For observations in 2020 and 2021, we extend their 2019 industry peers with the assumption that industry peers do not change drastically in these two years.

3. Empirical Design

3.1 Sample Selection and Panel Construction

[12] Moreover, in a horse-race regression where we compare the reading from articles with a relevance score of 90 to those with a relevance score below 90, it is the the reading of the highly relevant articles which drive our main results. For brevity, we do not tabulate these analyses.

Our firm-quarter panel is from U.S. firms between 2018Q1 and 2021Q4 in the Compustat Quarterly database, excluding firms with fewer than 14 quarters of non-zero lagged news reading, financial firms (SIC 6000-6999), utility firms (SIC 4900-4999), and observations with less than one million US dollars in assets or sales. The exclusion of financial firms is important as they have clearly different incentives to read the financial media than other firms. We choose 14 (of a total of 16) quarters to ensure that we pick firms whose reading behavior is consistently observed. Presumably, the financial media articles we study herein are amongst the largest financial media sites. But it is conceivable that some firms do not gather information from the media sites we study or are otherwise absent due to a lack of coverage in our dataset. Therefore, to be conservative, we opt to include only those firms in our dataset with consistent coverage.

Table 1 describes our sample selection process. Excluding financial and utility firms and other firms with assets or sales below one million dollars, there are 3,610 Compustat U.S. firms after 2018. Among the Compustat firms, 3,317 firms are tracked by the Data Partner, so we can observe their internet consumption. 3,024 firms visited financial media websites included in RavenPack at least once in our dataset. Finally, after requiring firms to have 14 quarters of non-zero financial media website reading, we are left with a total of 1,737 firms in our sample.

[Table 1 about here]

Appendix A.2 displays key financial variables (Tobin's Q, cash flows, assets, R&D, and R&D + CAPX) for the full Compustat universe as well as our selected firms. In general, the summary statistics for the key variables are similar. The median firm in our sample reports zero R&D as is the case for the Compustat sample, whereas the R&D expenditure over the total asset for the 75th percentile firm is 1.55% in our sample and 1.53% in the full sample.

Tobin's Q is also similar across both distributions, while our sample has a median of 1.72% and the Compustat sample has a median of 1.67%.

In our additional analysis reported in Table B.2 of Appendix B, we follow Peters and Taylor (2017) and construct a quarterly version of (i) total capital, (ii) total investment, and (iii) total Q, which more accurately measure a firm's intangible capital and serve as alternative measures for total assets, investment, and Tobin's Q, respectively. Specifically, the total capital, K^{tot}, is the sum of knowledge capital (perpetual inventory method on R&D with 15% effective annual depreciation rate), organizational capital (perpetual inventory method on 30% of SG&A with 20% effective annual depreciation rate), intangible assets, and PP&E from Compustat Quarterly. Total investment, ι^{tot}, is R&D plus 30% of SG&A plus CAPX scaled by lagged total capital. Total Q, Q^{tot}, is the market value of assets divided by total capital. Moreover, Peters and Taylor (2017) also define intangible investment, ι^{int}, as R&D plus 30% of SG&A scaled by lagged total capital.

3.2 Reading Measures

We provide summary statistics of the variables we create for our analyses in Table 2. News variables share a common form of $\log(1000 \times Count/atq + 1)$, where *Count* is the specific count for each type of news reading. We decompose reading by subject: $ReadNews^{Own}$ refers to how many times the firm reads news about its own firm with relevance greater than or equal to 90. A score of 90 suggests that the article mentions the company in the headline. $ReadNews^{Peer}$ describes how many times the firm reads news about peers defined by baseline TNIC with relevance greater than or equal to 90. $ReadNews^{General}$ is how many times the firm reads news about non-company entities only with relevance greater than or equal to 90. For *NewsSupply*, *Count* is the number of news articles about the firm. $ReadNews^{Own,Ktot}$ is the same as *ReadNews*^{Own}, except now the denominator is total capital (K^{tot}) from Peters and Taylor (2017) instead of assets.

[Table 2 about here]

While we cannot directly identify managers in our dataset, we can approximate whether reading is done by investment decision-makers based on location, presuming that these decision-makers are likely to be proximal to the headquarters. This assumption is implicit within the literature on strategy, accounting, finance, and economics (see, for example, Malenko (2019) and Dessein, Galleoti, and Santos (2016)). Accordingly, *ReadNews*^{Own} is decomposed by whether reading happens in the same metropolitan statistical area (MSA) of the headquarters' location or not, i.e., *ReadNews*^{Own,HQ} and *ReadNews*$^{Own,Non\text{-}HQ}$.

We also decompose articles by type. First, we decompose them by sentiment. In *ReadNews*$^{Own,Non\text{-}Pos}$ and *ReadNews*^{Own,Pos}, reading firms' own news (*ReadNews*^{Own}) is decomposed by RavenPack's sentiment score (event_sentiment_score) into negative (less than zero), neutral (equal to zero), and positive (greater than zero), where we group negative and neutral together as non-positive. Second, we decompose articles by whether they are full articles or not. Non-full articles are typically press releases and other short news flashes. This decomposition leverages the RavenPack field "news_type," where we classify a full article if this variable equals "full-article" into *ReadNews*$^{Own,Full\text{-}Article}$ and the remainder into *ReadNews*$^{Own,PR/Flash}$. We describe the remainder of the variables we construct in Appendix C.

3.3 Understanding the reading measures

Our goal in this section is to characterize why firms read the news. While firms may read the news for many reasons, we argue that the bulk of the reading measured in our dataset is work-related. We present two pieces of supporting evidence. First, in Figure 3, we plot the temporal distribution (measured in local time). On weekdays, we see that there is elevated reading during work hours, climbing at 8 am, plateauing from 10 am to 1 pm, and falling off toward the end of the day. On the weekend, we see no difference between work hours and evening hours. This pattern is consistent with our measure capturing activity that happens during work hours.

[Figure 3 about here]

Next, we examine how firm characteristics relate to the type of news read. If firms acquire information relevant to work, the type of information they acquire should be types of information that correspond to firm characteristics. In Table 3, we examine the relationships between reading and firm characteristics. These are panel regressions at the quarterly level with industry and quarter fixed effects, where the industry classification is defined at the 2-digit NAICS level. We have three outcome variables: $ReadNews^{Own}$, $ReadNews^{Peer}$, and $ReadNews^{General}$. A number of findings emerge that support the idea that firms pay attention to the news in order to consume the information relevant to them.

[Table 3 about here]

Our first variable of interest is *NewsSupply*, which is the log of the ratio between the number of articles about the firm and assets. The coefficient is positive in all specifications but is the largest for $ReadNews^{Own}$. Our preferred interpretation is that firms in the news most often tend to read more articles about themselves. Given that they read these news articles, the marginal cost of paying attention to other articles is also lower. Hence, there is a positive correlation between the articles about one's own firm and the reading of other articles.

Perhaps our most unambiguous piece of evidence is the finding that firms facing more product market competition (as measured by the variable *ProdMktFluidity*) are substantially more likely to read peer news, but not general news or own news. The coefficient on *ProdMktFluidity* is strongly positive but nonsignificant for one's own news, and marginally negative for macro news. Intuitively, when facing more competition, a firm reads more about peers' news and less about other news.

Next, motivated by a recent theory proposed by Gondhi (2022), we examine the role of stock volatility. In the narrative of Gondhi (2022), volatility is a proxy for events that capture significant uncertainty or events surrounding the firm. Such events tend to consume the manager's attention. Consistent with the model prediction, we find that firms that face more idiosyncratic volatility read more of their own news. On the other hand, firms pay less attention to general news when volatility is high. We do not find any effect of volatility on the reading of peers' news. Moreover, we also find that firms read less of their own news if recent returns are lower. This finding could be related to an ostrich effect whereby firms are reluctant to read the news of information that has recently been undesirable.

Finally, we also explore a number of financial characteristics in the interest of understanding firms' reading behavior. Interestingly, we find that firms with lower bid-ask spreads read more of their own stock. Similarly, we find that more analyst stock coverage and institutional ownership leads to more self-reading. This result is consistent with Baker and Wurgler (2006) who contend that such stocks, i.e., larger, high liquidity, more analyst coverage, and high institutional ownership, are likely to have more informative stock prices. Correspondingly, it is thus expected that these firms do more self-reading. Lastly, we find that larger firms read more articles, as do more profitable firms. That more profitable firms acquire more information is consistent with rational incentives to acquire information being greater for firms with greater productivity, as per Charoenwong, Kwan, Kimura, and Tan (2022).

4. Main Result

4.1 Main effect

In this section, we begin our main analysis by examining the relationship between news readership and investment sensitivity to Q. Following the literature on managerial learning, our key empirical specification is as follows:[13]

$$Expenditure_{it} = \beta_1 Q_{it-1} \times Read_{it-1} + \beta_2 Read_{it-1} + \beta_3 Q_{it-1} + Controls_{it-1} + FE + \epsilon_{it}$$

If $ReadNews^{Own}$ is a good proxy for managerial learning, we expect a higher investment sensitivity to Q when there is a higher readership of news. In other words, we expect β_1 to be positive. Given that we de-mean Q, β_3 describes the average firm's investment-to-price sensitivity. The expected sign for β_3 is also positive if (1) we do not capture all forms of learning, and there is learning from the stock price that occurs independently of our reading-based measures, or (2) beyond learning, there are other reasons why the investment-to-price sensitivity could still be positive (e.g., financial constraints).

Our outcome variables in this analysis are R&D and R&D + CAPX. We demean both *Q* and *Read* in the regression, so we can interpret β_2 as the investment-reading sensitivity for a firm with the average Q, and β_3 as the investment-Q sensitivity for a firm with average reading activity. The prior literature tends to examine the outcome of R&D and CAPX. We find stronger results for R&D. Our interpretation is that if managers are to "learn," they would most likely learn about difficult and uncertain types of investment such as R&D which may be more challenging to ascertain the value of, and thus might incentivize managerial learning (Bai, Phillipon, and Savov (2016)).

[Table 4 about here]

[13] See, for example, the recent work by Goldstein, Yang, and Zuo (2020).

Table 4 reports our main results. Columns 1 and 2 of Table 4 show the analysis for R&D. A one-unit increase in learning (with a standard deviation of 1.763) corresponds to an increase in investment sensitivity to Q of 1.763 * 0.0170 ~ 0.03. Relative to the base coefficient of 0.1144, this means that a one-standard-deviation increase in reading activity corresponds to a 26% increase in investment sensitivity to Q. The economic magnitude is quite meaningful. Column 3 reports the analysis for R&D and CAPX. Relative to the base term of 0.2228, a one-standard-deviation increase leads to a 16% increase in the investment sensitivity to Q. Columns 2 and 4 include redacted controls such as cash flow and the inverse of assets following Chen, Goldstein, and Jiang (2007).[14] Our estimates are resilient and largely unaffected by the addition of these controls.

4.2 Robustness and additional tests

In Table 5, we report our main robustness checks. First, one may be concerned that our reading measure might be a proxy of insider information, analyst information, or news supply. Following Chen, Goldstein, and Jiang (2007), we control for interactions of Q with insider trading and analyst coverage, which are proxies for insider information and analyst information, respectively. These are alternative sources of information other than the market feedback effect. In Panel A of Table 5, Columns 1 and 3 report our results for R&D and CAPX + R&D, respectively. In Panel B, we also repeat this analysis but control for the log of the ratio between the number of news articles about the firm and assets (*NewsSupply)* instead to rule out another alternative explanation that our reading measure is simply a proxy of news supply. The key message of Table 5 is that by controlling for these other sources of information, our main effect remains robust.

---

[14] It is worth noting that Chen, Goldstein and Jiang (2007) use a third control variable that is value-weighted market adjusted stock return for the next three years. However, given our short and recent sample, future stock return is unavailable for most of our observations. Hence, we drop this third control variable.

[Table 5 about here]

In Table B.1 of Appendix B, we examine whether the imputation of missing R&D to zero affects our results. Given that R&D is often missing (Koh and Reeb (2015)), instead of imputing a zero value, we also explore filling R&D with the industry average when missing. For firms with missing R&D, we fill their scaled R&D (R&D scaled by lagged assets) with the average scaled R&D from firms in the same 2-digit SIC and the same year-quarter. We also mark firms with missing R&D with a dummy variable. This additional variable ensures that our findings are robust to controlling for the possibility that some firms strategically redact R&D, which may skew our sample. We find that our results are largely unaffected after imputing missing R&D with the industry average. Finally, in Section 5.5, we discuss our approach to tackling measurement error in Q by re-defining Q by incorporating intangibles.

4.3 Geographic decomposition of reading – Headquarters versus other places

One concern might be that employees who read financial news at the firm about the firm's news are *not decision-makers* or gathering information for the decision-makers. While we cannot observe the learning behavior of decision-makers directly, we can approximate those readers who are likely to be or directly report to decision-makers by using the location of the anonymized readers. We split the reading variable into those who are located in the same *metropolitan statistical area (MSA)* as the headquarters of the firm and others.[15]

It is reasonable to assume that managers tend to work at the headquarters such that the main effect we documented is from the reading activities of the headquarters. Moreover, even if someone who is not the manager reads, they may be part of an internal team that collects

[15] It is worth noting how this location is derived. The location of a reader is inferred through a variety of industry-standard methods, such as the IP address or third-party sources common in the digital marketing industry used to identify the demographics of users. Typically, IP addresses are associated with particular cities. While these sources of information are not completely accurate, they are reported at the zip code level and likely accurate at a broader level such as MSA that we study.

information on behalf of the manager. Given it would be easier for these workers to coordinate with the manager at the headquarters, we argue this work arrangement is likelier at the headquarters [16] Alternatively, if the main effect was driven by regional rank-and-file employees, who might be learning of the firm's initiatives through the news but have no decision-making power, we might expect that the reading generated outside of the headquarters MSA drives the main learning result.

[Table 6 about here]

We run a horse-race regression whereby we split our main variable $ReadNews^{Own}$ into two buckets, one where the reading is *inside of the headquarters MSA* versus outside of it. Table 6 shows that for both of these variables, the main effect is driven by reading at the headquarters MSA. This finding would be what one expects if, on average, managers or other key decision-makers for corporate investment work at the headquarters.

4.4 Do financial constraints play a role in managerial learning?

The investment sensitivity to Q might also be explained by financial constraints, which is the major alternative explanation for managerial learning. *Ceteris paribus*, when valuations are higher, firm constraints are alleviated as the required return expected of the manager is lower. This explanation poses a challenge to the feedback literature because managerial learning behavior is not observed directly. In our context, given that we can observe organizational learning, we provide direct support to the feedback literature and address this criticism.

[16] In theory, "adaptive" tasks are usually located at the headquarters because coordination costs are lower and tasks are less likely to be routine (Dessein, Galleoti, and Santos (2016)). The importance of the HQ is also reflected in studies such as Malenko (2019).

Moreover, we explore how learning and financial constraints interact to better understand to what extent financial constraints play a role in our findings. Of course, even if a manager is financially constrained, this does not mean managers cannot learn from markets. Prior studies suggest managers endogenously gather information based on their incentives for information acquisition (e.g., Charoenwong, Kimura, Kwan, and Tan (2022)). Thus, it is arguable managers might have more incentive to learn when their firms are financially constrained such that they can better optimize their investments.

To examine how financial constraints factor into our findings, we construct a proxy for financial constraints based on a composite of five different indicators in the literature. Our first measure is simply the logarithm of market capitalization (minus MktCap). The intuition of this measure is that well-capitalized firms are likely to be able to issue equity as a small percentage of their total market capitalization. Second, when a firm has the ability to pay dividends (minus DivPos), it indicates that the firm has enough free cash flow to do so. Third, KZ4 is the four-variable KZ index from Kaplan and Zingales (1997) and Baker, Stein, and Wurgler (2003). Intuitively, the KZ4 index measures a firm's dependence on external financing. KZ4 is higher, i.e., the firm is more financially constrained, when it has lower cash holding, lower cash flow, lower dividend payout, and higher leverage. Fourth, we consider the WW index from Whited and Wu (2006), which modifies KZ with additional variables and is estimated via generalized methods of moments. Finally, we leverage the size-age index (SA index) from Hadlock and Pierce (2010). The intuition of this measure is that larger firms (by assets) have more assets to pledge, and thus it is easier for them to raise capital. By the same token, older firms have more assets in place and are more familiar to investors. Thus, their ability to raise capital is greater.

To create a composite indicator, we blend together the five indicators through two methods. The first one is the average of the z-scores of each of our financial constraint variables (AVG^z). Implicitly, this method assumes all indicators have equal weight, although some indicators generally exhibit less variation than others. They are constructed so that higher values mean tighter constraints. The second composite indicator is the first principal component of the five indicators (PCA^z), with a higher value indicating a firm being more financially constrained. This method allows the financial constraint index to have different weights according to each of the five indicators' partial contribution to overall variance. We also analyze each indicator separately.

[Table 7 about here]

Table 7 reports our results. We interact financial constraint measure (*Constraint*) with our measures of Q and learning. However, no matter what measure of constraints we use, interacting with these measures does little to affect our main finding. In fact, for our first two columns, there is no differential interaction between the effect of learning (the coefficient on $ReadNews^{Own}$ and *Q*) and financially constrained firms. Examining the individual indicators, only one of our measures shows a positive relationship between learning and financial constraints (the KZ4 index in Column 5) that is significant at the 90% confidence level. Overall, however, the majority of our tests do not show that the learning effect is substantially higher or lower for firms with financial constraints. These findings mitigate the concern that our reading measure does not capture managerial learning but rather financial constraints.

5. Additional Decompositions of Reading Activity and other Tests

5.1 Firm-specific news, industry news, or general news reading

In this section, we explore whether the learning effect is driven by the reading of firm-specific news or the reading of other news types. In these regressions, we take our main regression from Table 4 and add the reading of industry peers and reading of "general" news, where "general" simply refers to non-company-specific reading. For example, its title may include mentions of political events, laws, regulations, movements in commodity prices, or sovereign bonds, without mention of any specific company.

[Table 8 about here]

What we find in Table 8 is that the main effect is driven by corporate reading firm-specific news, and to a smaller extent, driven by industry peer news. There are two takeaways from this analysis. First, it suggests that we are not accidentally capturing the effect of news reading in general and the associated fundamentals. By and large, our main effect comes from corporate reading on firm-specific news.

Our second takeaway is that we also find some role for peer news, suggesting managers do learn not only from their own firm-specific news but also from their relation to others in the industry. This finding speaks to Foucault and Fresard (2014), who argue that managers learn from peer stock prices. Our setup allows us to quantify their relative extent: firm-specific news learning is approximately twice as large as industry peer news learning. Interestingly, we find no effect on general news. This result is consistent with the theory of Gondhi (2022), who indicates that aggregate news and attention serve as a distraction for the capital expenditure decisions of managers.

5.2 Feedback effect from negative, neutral, and positive news

Next, we examine the relationship between the content of the articles and the learning effect. One of the reason firms may pay attention to stock prices is that it reinforces the firm's priors (Bond, Edmans, and Goldstein (2012)), rather than providing new information. If one expects that managers simply read the news to confirm their own priors, we might expect that managers will invest more when they read *positive* news. In contrast, due to the nature of managerial overoptimism, managers might be able to learn more when they receive information that is against their prior, i.e., negative or neutral news. If this is the case, we would expect the learning effect to be mostly driven by non-positive news.

[Table 9 about here]

Table 9 reports the result of this analysis. Column 1 reports our regression using R&D as the dependent variable and splitting reading into non-positive news and positive news without control variables. Column 2 adds control variables. Columns 3 and 4 examine the same results but for R&D + CAPX. We find that the learning effect is driven by reading negative or neutral news but not by positive news. The result on positive news is essentially null, suggesting that managers do not find positive articles relevant for investment decisions. This further helps support the argument that firms' primary motivation in reading the news is to acquire useful information.

5.3 Types of articles

Finally, we decompose firm-specific news into articles that are full-length articles versus press releases/news flashes. The rationale behind this decomposition is that full-length articles about a firm may include subjective opinions from outsiders that insiders do not know (e.g., analysis or comments by industry experts), while press releases and news flashes about a firm are most likely objective facts that insiders already know (or insiders themselves supply such facts to news publishers).

[Table 10 about here]

Table 10 presents our findings. We find that by and large, the main reading effect on firm-specific news originates from reading full-length articles. This dichotomy finding suggests that firms do not benefit from reading news when the media reports breaking news but rather in-depth coverage. This result is sensible since firms presumably should be aware of basic facts about their own firm.

5.4 Other tests: Relative size to peers and M&A decisions

For completeness, we also explore the relationship between learning and relative-to-rival peers. For each firm, we calculate the relative size of a focal firm to its top 10 Hoberg and Phillips industry peers (defined as having the closest textual similarity of product descriptions). We report these results in Table B.2 of Appendix B. We find that when managers are larger than their rivals, on the margin, they learn less from their own news. This is consistent with the idea that leaders or firms with a large competitive edge have lower incentives to learn.

We also explore mergers and acquisitions decisions. Generally, the results are broadly supportive of the idea that managers learn from the market during the M&A process. But our sample is very recent and many announced deals are still pending. Thus, any positive results we have seen so far could be premature. Hence, we do not report them for brevity, and the results are available upon request.

5.5 Addressing Measurement Error in Q

Measurement error has been a focus of some prior work on investment sensitivity to Q. The impact of measurement error is mitigated in our context because our main goal is not to estimate the effect of *marginal Q* on investment, but rather to see how managerial learning alters this relationship. If measurement error in Q is not correlated with the managers' learning activity, the potential concerns are alleviated.

Nevertheless, we follow Andrei, Mann, and Moyen (2019) and aim to deal with the measurement error of Q in two ways. First, we corroborate our main result by examining the effect of managerial learning on intangible capital investment. We follow Peters and Taylor

(2017) to use intangible investment and total investment as the dependent variable and intangible stock to deflate financial ratios instead of assets. If the managers learn from reading financial news for their intangible investment, we expect results analogous to those in Table 4. Table B.3 of Appendix B reports the result. We indeed find a positive and significant interaction term between our reading measure and a firm's Total Q at *t–1* when we use either intangible investment or total investment as the dependent variable. This result is consistent with our argument of managerial learning from reading, which helps mitigate the concern of the potential estimation bias due to the measurement error of Tobin's Q.

Second, we explore the cumulant estimator of Erickson, Jiang, and Whited (2014) with code from Erickson, Parham, and Whited (2017). We offer a critical disclaimer: the cumulant estimator assumes uncorrelated measurement errors across the mismeasured variables. However, if our mismeasured variables are a base term and an interaction term, then that assumption is likely to be violated by definition. With this caveat in mind and to address the econometric issue, we conduct a split sample analysis following Andrei, Mann, and Moyen (2019). A split sample design mitigates the aforementioned issue by not introducing multiple correlated mismeasured variables. Specifically, we split our sample into high residualized reading (top quartile) vs. the low remainder. To residualize the reading measure, we partial out firm and quarter fixed effects from $\text{ReadOwn}^{\text{News}}$. We report the results in Table B.4 and find that in these specifications, the Q-sensitivity is 40% higher when the firm-quarter observation is in the upper quartile. When we run the same OLS specification, the increase in magnitude is 14%, suggesting that correcting measurement issues could potentially even enlarge our point estimates.[17]

[17] Meanwhile, we also employ the estimator on our main specification. In this exercise, we assume all three of our variables are mismeasured. In untabulated results, we find supportive evidence that investment-to-price sensitivity is 32.5% larger as opposed to 22% for R&D, and for CAPX +RND, the learning effect is 52% larger. Thus, based on this specification, the magnitude of amplification of investment-to-price sensitivity is actually larger than our main regression. However, the τ^2 and ρ^2 values plummet in this specification, implying potential issues with the model fit.

Overall, both additional analyses yield the same conclusion as our main specification, helping mitigate the concern of the measurement error of Tobin's Q.

6. Conclusion and discussion

In this paper, we furnish a novel setting to test the managerial learning hypothesis. This hypothesis is typically challenging to provide direct evidence because learning activities are not directly observed at the firm level in prior studies. To deliver these results, we examine *firms reading of* article-level data. We argue that managers who learn from the stock market rely on reading financial news to decode the noisy signals embedded in the stock prices. We find that a firm's investment sensitivity to Q is 24% larger when its employees read more of their own news. Consistent with the reading activity reflecting the decoding actions of managers, reading at the headquarters MSA matters for investment sensitivity to Q, but reading outside the headquarters MSA does not. Meanwhile, financial constraints do not play a significant role in our setup.

Our findings have several implications. First, omitted variables have been a key concern for interpreting the relationship between Q and investment. To the extent one finds our measure of managerial decoding signals in stock prices credible, this concern is significantly mitigated as we provide direct evidence managers do learn from the stock prices and use the financial media to interpret those price movements. Second, indirectly, our findings imply that learning is costly for firms. If not, then variation in information acquisition should not vary in response to firm characteristics, and variation in reading would not exist, and in turn, drive the Q-sensitivity. It also implies that managers and not just investors learn from the media.

The findings also lead to additional potential research questions. First, if managers learn from news media as well, models of investors learning about firms in information markets might also incorporate the effects of coverage on firms and their decisions, as managers might respond to the same information generated about firms and peers.[18] Second, there are multiple types of market feedback that might be important for firms' learning: social media, analyst reports, or private interactions with institutional investors. In addition, firms hire their own teams to do capital budgeting (Charoenwong, Kimura, Kwan, and Tan (2022)) irrespective of market feedback. It would be useful to quantify the relative contributions to information acquisition, learning, and corporate investment, as well as to characterize the dynamics such as whether these different types of information are complementary or substitutive. In this way, we can better understand the value firms' information acquisition has for corporate decisions and firm performance.

[18] For instance, the theory of Veldkamp (2006) focuses on how commonalities in information production across firms generates comovement in returns – one additional reason for excess comovement observed in the data might be that other firms take actions as a result of reading the news and obtaining market feedback.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5), 1977-2016.

Andrei, D., Mann, W., & Moyen, N. (2019). Why did the q theory of investment start working?. *Journal of Financial Economics*, 133(2), 251-272.

Bai, J., Philippon, T., & Savov, A. (2016). Have financial markets become more informative?. *Journal of Financial Economics*, 122(3), 625-654.

Baker, M., Stein, J. C., & Wurgler, J. (2003). When does the market matter? Stock prices and the investment of equity-dependent firms. *The Quarterly Journal of Economics*, 118(3), 969-1005.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The journal of Finance*, 61(4), 1645-1680.

Bakke, T. E., & Whited, T. M. (2010). Which firms follow the market? An analysis of corporate investment decisions. *The Review of Financial Studies*, 23(5), 1941-1980.

Bond, P., Edmans, A., & Goldstein, I. (2012). The real effects of financial markets. *Annual Review of Financial Economics*, 4(1), 339-360.

Brogaard, J., Shi, W., Wei, J., & You, H. (2022). Do Managers Learn from Analysts? *Working Paper*.

Charoenwong, B., Kimura, Y., Kwan, A., & Tan, E. (2022). Capital budgeting, uncertainty, and misallocation. *Working Paper*.

Chen, Q., Goldstein, I., & Jiang, W. (2007). Price informativeness and investment sensitivity to stock price. *The Review of Financial Studies*, 20(3), 619-650.

Chen, H., Cohen, L., Gurun, U., Lou, D., & Malloy, C. (2020). IQ from IP: Simplifying search in portfolio choice. Journal of Financial Economics, 138(1), 118-137.

Dessein, W., Galeotti, A., & Santos, T. (2016). Rational inattention and organizational focus. *American Economic Review*, 106(6), 1522-36.

Dessaint, O., Foucault, T., & Frésard, L. (2021). Does alternative data improve financial forecasting? the horizon effect. *Working Paper*.

Edmans, A., Goldstein, I., & Jiang, W. (2012). The real effects of financial markets: The impact of prices on takeovers. *The Journal of Finance*, 67(3), 933-971.

Edmans, A., Goldstein, I., & Jiang, W. (2015). Feedback effects, asymmetric trading, and the limits to arbitrage. *American Economic Review*, 105(12), 3766-3797.

Engelberg, J. E., & Parsons, C. A. (2011). The causal impact of media in financial markets. *The Journal of Finance*, 66(1), 67-97.

Erickson, T., Jiang, C. H., & Whited, T. M. (2014). Minimum distance estimation of the errors-in-variables model using linear cumulant equations. *Journal of Econometrics*, 183(2), 211-221.

Erickson, T., Parham, R., & Whited, T. M. (2017). Fitting the errors-in-variables model using high-order cumulants and moments. *The Stata Journal*, *17*(1), 116-129.

Fang, L., & Peress, J. (2009). Media coverage and the cross-section of stock returns. *The Journal of Finance*, 64(5), 2023-2052.

Fernandes, N., & Ferreira, M. A. (2009). Insider trading laws and stock price informativeness. *The Review of Financial Studies*, 22(5), 1845-1887.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.

Foerster, A. T., Sarte, P. D. G., & Watson, M. W. (2011). Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy,* 119(1), 1-38.

Foucault, T., & Frésard, L. (2012). Cross-listing, investment sensitivity to stock price, and the learning hypothesis. *The Review of Financial Studies,* 25(11), 3305-3350.

Foucault, T., & Fresard, L. (2014). Learning from peers' stock prices and corporate investment. *Journal of Financial Economics*, 111(3), 554-577.

Fox, Z. D., Kim, J., & Schonberger, B. (2021). Investment Opportunities, Market Feedback, and Voluntary Disclosure: Evidence from the Shale Oil Revolution. *Market Feedback, and Voluntary Disclosure: Evidence from the Shale Oil Revolution (February 4, 2021)*.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3), 733-772.

Goldstein, I., Yang, S., & Zuo, L. (2020). The real effects of modern information technologies (No. w27529). *National Bureau of Economic Research*.

Goldstein, I., & Guembel, A. (2008). Manipulation and the allocational role of prices. *The Review of Economic Studies,* 75(1), 133-164.

Goldstein, I., Liu, B., & Yang, L. (2021). Market feedback: Evidence from the horse's mouth. *Available at SSRN 3874756*.

Gondhi, N. (2022). Rational inattention, misallocation, and the aggregate economy. *Journal of Monetary Economics, forthcoming*.

Gross, R. (2010). Psychology: The science of mind and behaviour. Hodder Education Publishers, 6th edition.

Hadlock, C. J., & Pierce, J. R. (2010). New evidence on measuring financial constraints: Moving beyond the KZ index. *The Review of Financial Studies*, 23(5), 1909-1940.

Hennessy, C. A., Levy, A., & Whited, T. M. (2007). Testing Q theory with financing frictions. *Journal of Financial Economics*, 83(3), 691-717.

Hoberg, G., & Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5), 1423-1465.

Hoberg, G., Phillips, G., & Prabhala, N. (2014). Product market threats, payouts, and financial flexibility. *The Journal of Finance*, 69(1), 293-324.

Kacperczyk, M., Van Nieuwerburgh, S., & Veldkamp, L. (2016). A rational theory of mutual funds' attention allocation. *Econometrica*, 84(2), 571-626.

Kaplan, S. N., & Zingales, L. (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints?. *The Quarterly Journal of Economics*, 112(1), 169-215.

Kelley, E. K., & Tetlock, P. C. (2017). Retail short selling and stock prices. *The Review of Financial Studies*, 30(3), 801-834.

Lin, T. C., Liu, Q., & Sun, B. (2019). Contractual managerial incentives with stock price feedback. *American Economic Review*, 109(7), 2446-68.

Malenko, A. (2019). Optimal dynamic capital budgeting. *The Review of Economic Studies*, 86(4), 1747-1778.

Morck, R., Shleifer, A., & Vishny, R. W. (1990). The stock market and investment: is the market a sideshow?. *Brookings Papers on Economic Activity*, 1990(2), 157-215.

Peress, J. (2008). Media coverage and investors' attention to earnings announcement. Available at SSRN 2723916.

Peress, J. (2014). The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *the Journal of Finance*, 69(5), 2007-2043.

Peress, J. (2014). Learning from stock prices and economic growth. *The Review of Financial Studies*, 27(10), 2998-3059.

Peters, R. H., & Taylor, L. A. (2017). Intangible capital and the investment-q relation. *Journal of Financial Economics*, 123(2), 251-272.

Tetlock, P. C. (2015). The role of media in finance. *Handbook of Media Economics*, 1, 701-721.

Veldkamp, L. L. (2006a). Media frenzies in markets for financial information. *American Economic Review*, 96(3), 577-601.

Veldkamp, L. L. (2006b). Information markets and the comovement of asset prices. *The Review of Economic Studies*, 73(3), 823-845.

Whited, T. M., & Wu, G. (2006). Financial constraints risk. *The Review of Financial Studies*, 19(2), 531-559.

Yan, D. (2021). Do private firms (mis) learn from the stock market?. *Working Paper*.

Figure 1: Conceptual Framework

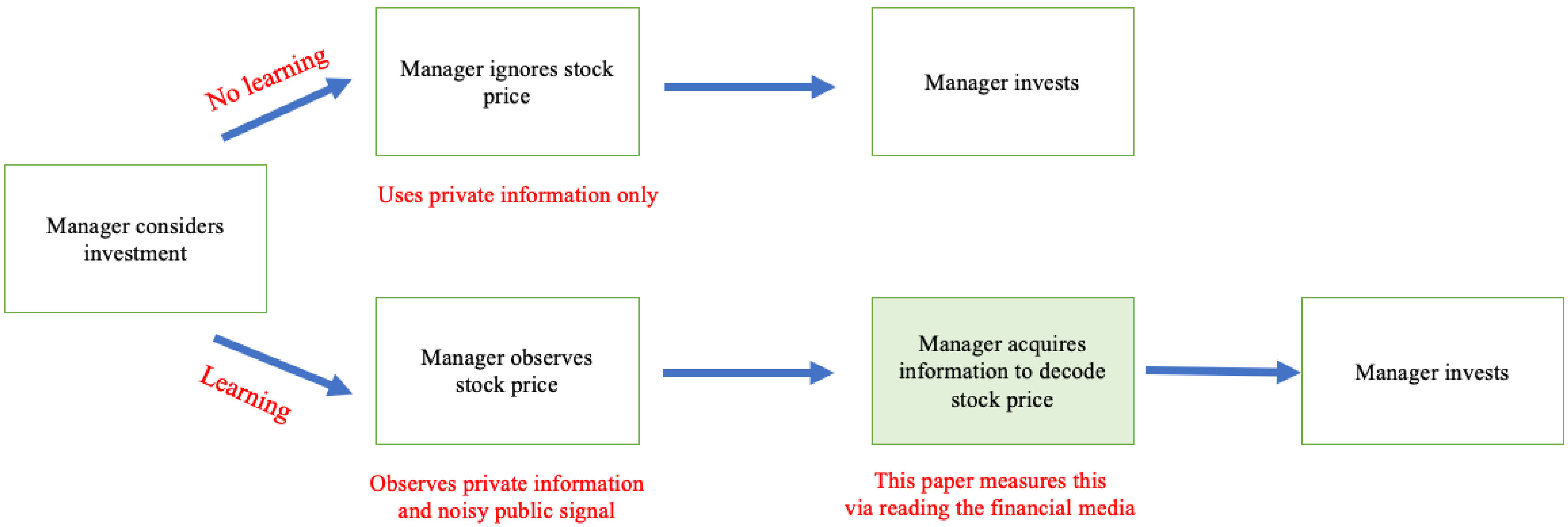

Figure 2: Data Diagram

In this diagram, we illustrate a scenario in which Publisher A and Publisher B agree to share data with the "Data Partner," whereas Publisher C does not.

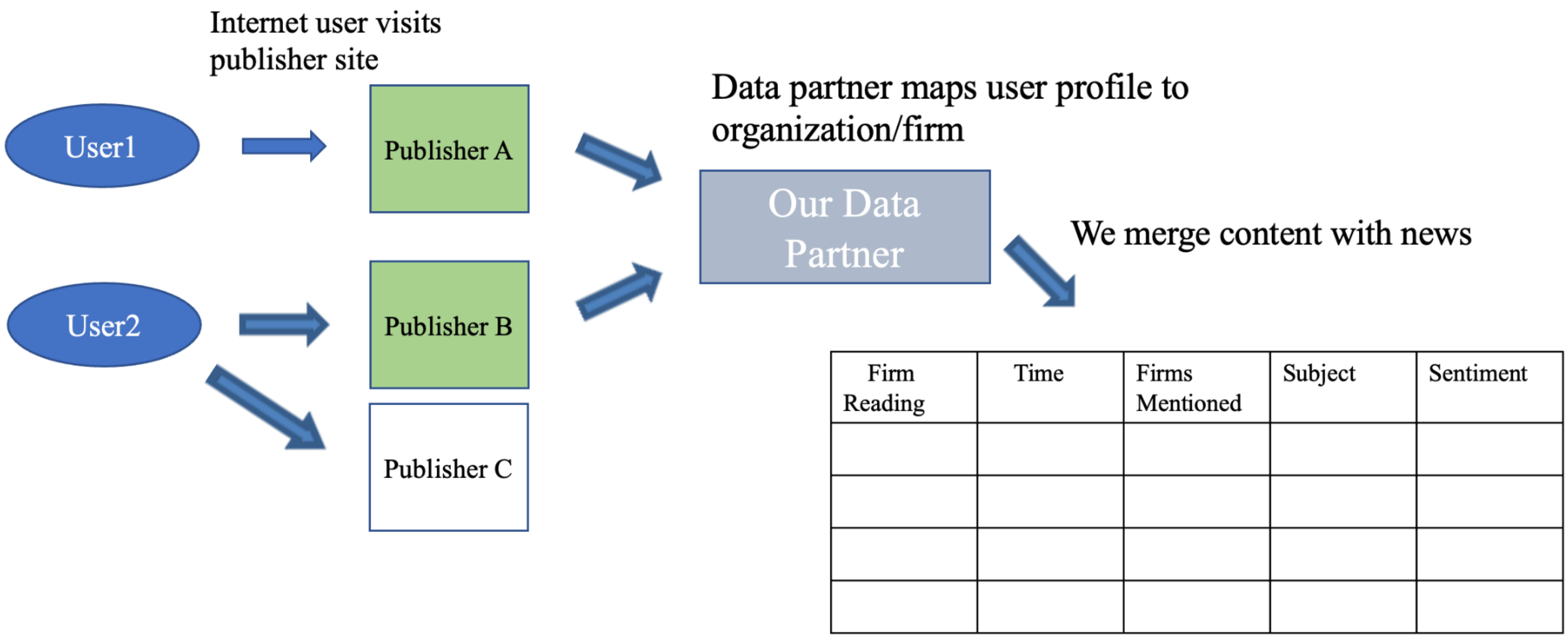

Figure 3: Does Financial Media Consumption Happen *During the Workday*?

We present the temporal distribution of articles read by the hour of the day for the publishers we observe in our dataset.

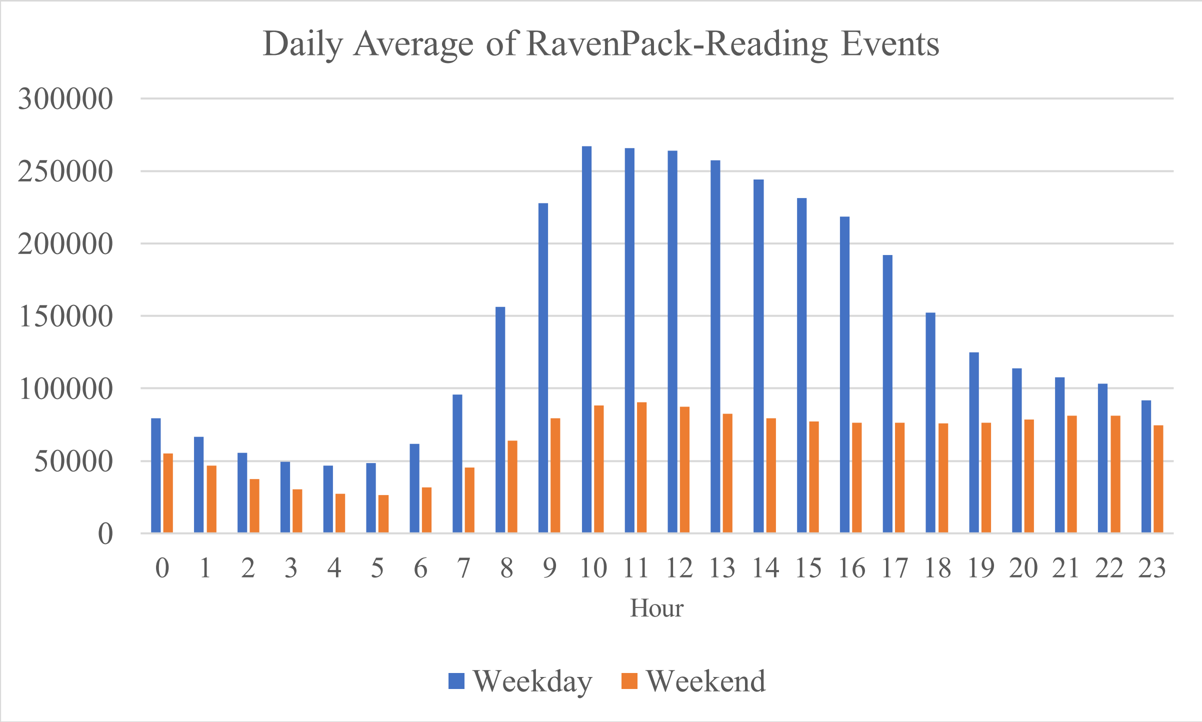

Table 1: Summary statistics

This table describes our sample firm selection process. Each step provides details of the filters applied, and the number of firms left after the filtering.

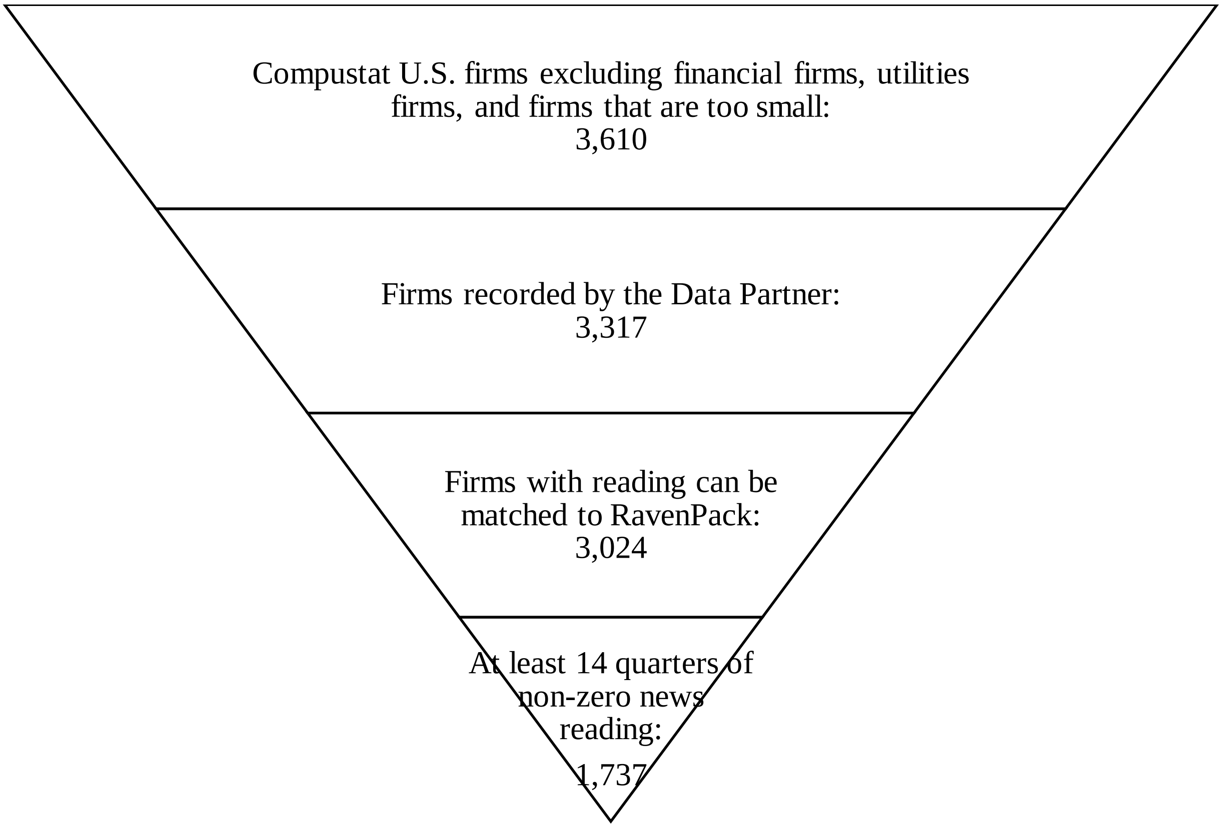

Table 2: Sample Summary Statistics

Panel A: News and news reading variables

This panel provides summary statistics for the news-related variables. News variables share a common form of $\log(1000 \times Count/atq + 1)$, where *Count* is the specific count for each type of news reading. We decompose reading by subject. *Count* in $ReadNews^{Own}$ refers to how many times the firm reads news about its own firm with relevance scores greater than or equal to 90. A score of 90 suggests that the article mentions the company in the headline or in the first paragraph. *Count* in $ReadNews^{Peer}$ describes how many times the firm reads news about peers defined by baseline TNIC with relevance scores greater than or equal to 90. For $ReadNews^{General}$, *Count* is how many times the firm reads news about non-company entities only with relevance scores greater than or equal to 90. For *NewsSupply*, *Count* is number of news articles about the firm. $ReadNews^{Own,\ Ktot}$ shares the same *Count* as $ReadNews^{Own}$, but scaled by total capital K^{tot} from Peters and Taylor (2017) instead of total assets *atq*.

	N	**Mean**	**Std. Dev.**	**P25**	**Median**	**P75**
$\mathrm{ReadNews}^{\mathrm{Own}}$	27,126	1.415	1.763	0	0.56	2.547
$\mathrm{ReadNews}^{\mathrm{Own,\ Ktot}}$	25,851	1.413	1.748	0	0.58	2.536
$\mathrm{ReadNews}^{\mathrm{Own,\ HQ}}$	26,204	0.889	1.465	0	0	1.355
$\mathrm{ReadNews}^{\mathrm{Own,\ Non\text{-}HQ}}$	26,204	0.7	1.214	0	0	0.986
$\mathrm{ReadNews}^{\mathrm{Own,\ Non\text{-}Pos}}$	27,126	0.788	1.311	0	0	1.204
$\mathrm{ReadNews}^{\mathrm{Own,\ Pos}}$	27,126	0.721	1.225	0	0	1.051
$\mathrm{ReadNews}^{\mathrm{Own,\ Full\text{-}Article}}$	27,126	1.122	1.543	0	0.114	1.973
$\mathrm{ReadNews}^{\mathrm{Own,\ PR\ /\ Flash}}$	27,126	0.691	1.312	0	0	0.803
$\mathrm{ReadNews}^{\mathrm{Peer}}$	23,645	1.503	1.868	0	0.569	2.703
$\mathrm{ReadNews}^{\mathrm{General}}$	27,493	6.964	1.663	6.088	7.113	8.061
NewsSupply	26,844	5.445	1.526	4.347	5.288	6.471

Panel B: Firm fundamentals

This panel provides summary statistics for common firm variables. Appendix C provides detailed definitions for each variable including descriptions of the data source.

	N	Mean	Std. Dev.	P25	Median	P75
Core						
R&D	27,515	1.204	2.222	0	0	1.548
R&D+CAPX	27,429	2.18	2.525	0.535	1.263	2.886
ι^{int}	26,194	2.646	2.747	0.804	1.936	3.442
ι^{tot}	26,131	3.559	2.984	1.715	2.76	4.282
Q	27,288	2.432	1.904	1.194	1.715	2.88
Q^{tot}	24,864	1.747	2.314	0.455	0.927	1.975
Controls						
CF	27,194	0.861	4.982	–0.031	1.781	3.185
CF^{Ktot}	25,904	1.201	4.979	0.002	1.69	3.075
1/AT	27,515	7.707	20.771	0.214	0.8	3.679
$1/K^{tot}$	26,194	6.283	16.004	0.205	0.795	3.314
Insider	27,515	0.015	0.051	0	0.001	0.007
Analyst	27,515	2.521	1.28	1.946	2.773	3.466
Reading Determinants						
Size	27,515	6.974	2.127	5.605	7.131	8.452
Leverage	26,620	44.433	37.737	16.736	40.24	61.393
Tangibility	27,488	25.246	22.824	8.321	16.951	35.724
Profitability	27,462	25.03	17.895	12.736	20.318	32.017
Volatility	25,536	0.028	0.018	0.016	0.023	0.034
Return4Q	25,180	0.223	0.693	–0.191	0.089	0.432
QuotedSpread	25,768	0.006	0.009	0.001	0.003	0.006
InstitutionOwn	27,515	0.388	0.404	0	0.236	0.816
ProdMktFluidity	14,114	5.866	3.243	3.518	5.128	7.287
Cross-sectional Traits						
FinConstraintAVG	25,470	–0.097	0.523	–0.457	–0.059	0.24
FinConstraintPCA	25,470	–0.158	1.221	–1.186	–0.131	0.712
$Sales^{RelToPeers}$	23,650	5.894	17.133	0.433	1.215	3.733

Table 3: Understanding why firms read the financial press

In this table, we present firm-quarter panel regressions relating financial news media reading intensity to firm characteristics. The dependent variable is reading intensity, defined as log of ratio between count of financial news media reading at *t* and assets at *t*, where we count reading when the news is relevant to its own firm ($ReadNews^{Own}$) in Columns 1 and 2, when the news is relevant to its baseline TNIC peers ($ReadNews^{Peer}$) in Columns 3 and 4, and when the news is relevant to non-company entities only ($ReadNews^{General}$) in Columns 5 and 6. Firm characteristics are measured at quarter *t*, and their definitions can be found in Appendix C. Even-numbered columns include Product Market Fluidity (*ProdMktFluidity*) which significantly shrinks our sample size. Regressions contain industry and year-quarter fixed effects. Standard errors reported in parentheses are clustered by firm.

	$\mathbf{ReadNews^{Own}}$		$\mathbf{ReadNews^{Peer}}$		$\mathbf{ReadNews^{General}}$	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
NewsSupply	0.8369***	1.0251***	0.6035***	0.7427***	0.4021***	0.5675***
	(0.0452)	(0.0583)	(0.0580)	(0.0700)	(0.0509)	(0.0623)
CashFlow	–0.0119*	–0.0123	–0.0248***	–0.0137*	–0.0076	–0.0138**
	(0.0064)	(0.0085)	(0.0072)	(0.0080)	(0.0048)	(0.0057)
Size	0.6486***	0.7982***	0.2452***	0.4176***	–0.0180	0.1259**
	(0.0401)	(0.0512)	(0.0524)	(0.0623)	(0.0457)	(0.0543)
Leverage	0.0013	0.0012	0.0016	0.0006	0.0013*	0.0004
	(0.0009)	(0.0011)	(0.0012)	(0.0012)	(0.0008)	(0.0009)
Tangibility	–0.0021	–0.0024	–0.0064***	–0.0060***	–0.0002	–0.0008
	(0.0015)	(0.0018)	(0.0018)	(0.0019)	(0.0020)	(0.0021)
Profitability	0.0077***	0.0091***	0.0021	0.0081***	0.0173***	0.0185***
	(0.0017)	(0.0020)	(0.0023)	(0.0023)	(0.0020)	(0.0023)
Volatility	12.6002***	12.8802***	2.5068	0.7107	–3.5393**	–4.6934**
	(1.5775)	(2.3988)	(1.8526)	(2.4474)	(1.4672)	(2.1226)
Return4Q	0.0815***	0.1127**	–0.0271	–0.0156	–0.0511**	–0.0265
	(0.0280)	(0.0545)	(0.0322)	(0.0533)	(0.0212)	(0.0423)
QuotedSpread	–15.7427***	–16.2034***	–19.9014***	–13.7233**	2.7640	4.3770
	(3.2485)	(4.1622)	(4.9239)	(5.3695)	(3.5966)	(4.1941)
Analyst	0.1366***	0.0869*	0.2636***	0.1400***	0.0089	–0.0130
	(0.0397)	(0.0483)	(0.0486)	(0.0505)	(0.0478)	(0.0544)
InstitutionOwn	0.2485***	0.1905*	0.2638**	0.3460***	0.3622***	0.3229***
	(0.0954)	(0.1102)	(0.1096)	(0.1196)	(0.1084)	(0.1215)
ProdMktFluidity		0.0060		0.1151***		–0.0214*
		(0.0114)		(0.0108)		(0.0113)
Industry FE	NAICS2	NAICS2	NAICS2	NAICS2	NAICS2	NAICS2
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	23,769	12,747	21,571	11,796	23,798	12,735
Adjusted R^2	0.2677	0.2699	0.2477	0.2783	0.3370	0.2792

Table 4: Main Result

In this table, we present firm-quarter panel regressions relating a firm's financial news media reading intensity at quarter *t–1* to financial spending recorded at *t* by following the regression model:

$$Expenditure_{it} = \beta_1 Q_{it-1} \times Read_{it-1} + \beta_2 Read_{it-1} + \beta_3 Q_{it-1} + Controls_{it-1} + FE + \epsilon_{it}.$$

The dependent variable *Expenditure* is R&D (Columns 1 and 2) and R&D+CAPX (Columns 3 and 4) spending at *t* scaled by assets at *t–1*. *Read* is log of ratio between count of financial news media reading at *t–1* and assets at *t–1*, where we count reading when the news is relevant to its own firm ($ReadNews^{Own}$). *Q* refers to a firm's Tobin's Q at *t–1*. Firm controls include *CF*, a firm's cash flow at *t–1* scaled by assets at *t–2*, and *1/AT*, the inverse ratio of a firm's assets at *t–1*. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Standard errors reported in parentheses are clustered by firm.

	R&D		**R&D+CAPX**	
	(1)	**(2)**	**(3)**	**(4)**
Q×ReadNews^{Own}	0.0170***	0.0162***	0.0202***	0.0191***
	(0.0042)	(0.0043)	(0.0052)	(0.0052)
Q	0.1144***	0.0941***	0.2228***	0.1994***
	(0.0193)	(0.0187)	(0.0226)	(0.0226)
ReadNews^{Own}	0.0104*	0.0108*	0.0037	0.0042
	(0.0063)	(0.0061)	(0.0096)	(0.0097)
CF		–0.0217***		–0.0088*
		(0.0035)		(0.0049)
1/AT		0.0244***		0.0259***
		(0.0069)		(0.0084)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	26,998	26,703	26,949	26,661
Adjusted R^2	0.8857	0.8904	0.7819	0.7843

Table 5: Robustness

In this table, we perform various robustness checks with firm-quarter panel regressions relating a firm's financial news media reading intensity at quarter *t–1* to financial spending recorded at *t*. The dependent variable is R&D (Columns 1 and 2) and R&D+CAPX (Columns 3 and 4) spending at *t* scaled by assets at *t–1*. Reading intensity is log of ratio between count of financial news media reading at *t–1* and assets at *t–1*, where we count reading when the news is relevant to its own firm ($ReadNews^{Own}$). *Q* refers to a firm's Tobin's Q at *t–1*. Firm controls include *CF*, a firm's cash flow at *t–1* scaled by assets at *t–2*, and *1/AT*, the inverse ratio of a firm's assets at *t–1*. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Panel A tests whether financial news media reading intensity is a proxy for insider or analyst information by including *Insider*, share of trading volume by firm insiders, and *Analyst*, log number of analysts covering the firm, and their interaction with *Q*. Panel B tests whether financial news media reading intensity is a proxy for news supply by including *NewsSupply*, log of ratio between number of news covering the firm at *t–1* and assets at *t–1*, and its interaction with *Q*. Standard errors reported in parentheses are clustered by firm.

Panel A: Is reading a proxy for insider or analyst information?

	R&D		**R&D+CAPX**	
	(1)	**(2)**	**(3)**	**(4)**
Q×ReadNews^{Own}	0.0167***	0.0142***	0.0211***	0.0181***
	(0.0042)	(0.0044)	(0.0053)	(0.0055)
Q×Insider	0.1525	0.1945**	0.1240	0.1689
	(0.0971)	(0.0975)	(0.1241)	(0.1250)
Q×Analyst	–0.0089	0.0214*	–0.0230	0.0109
	(0.0109)	(0.0111)	(0.0142)	(0.0142)
Q	0.1365***	0.0386	0.2792***	0.1682***
	(0.0310)	(0.0309)	(0.0415)	(0.0414)
ReadNews^{Own}	0.0104*	0.0121**	0.0028	0.0043
	(0.0063)	(0.0060)	(0.0096)	(0.0098)
Insider	–0.0007	0.0081	0.2299	0.2709
	(0.1138)	(0.1143)	(0.2249)	(0.2257)
Analyst	–0.1731***	–0.0945***	–0.0668	0.0325
	(0.0363)	(0.0346)	(0.0523)	(0.0542)
CF		–0.0216***		–0.0087*
		(0.0035)		(0.0048)
1/AT		0.0242***		0.0266***
		(0.0070)		(0.0086)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	26,998	26,703	26,949	26,661
Adjusted R^2	0.8863	0.8908	0.7821	0.7844

Panel B: Is reading a proxy for the news supply?

	R&D		R&D+CAPX	
	(1)	(2)	(3)	(4)
Q×$ReadNews^{Own}$	0.0116***	0.0124***	0.0159***	0.0164***
	(0.0041)	(0.0042)	(0.0052)	(0.0053)
Q×NewsSupply	0.0514***	0.0401***	0.0324***	0.0208**
	(0.0085)	(0.0078)	(0.0108)	(0.0100)
Q	0.0318*	0.0320*	0.1569***	0.1534***
	(0.0192)	(0.0185)	(0.0244)	(0.0244)
$ReadNews^{Own}$	–0.0018	0.0012	–0.0129	–0.0100
	(0.0063)	(0.0061)	(0.0097)	(0.0099)
NewsSupply	0.2086***	0.1639***	0.2989***	0.2597***
	(0.0223)	(0.0224)	(0.0309)	(0.0322)
CF		–0.0195***		–0.0061
		(0.0034)		(0.0049)
1/AT		0.0229***		0.0249***
		(0.0078)		(0.0096)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	26,676	26,385	26,631	26,346
Adjusted R^2	0.8881	0.8917	0.7844	0.7861

Table 6: Interpretation – Headquarters metro versus other

In this table, we compare learning by readers located in a firm's headquarters metropolitan statistical area (MSA) versus readers located in other places, with firm-quarter panel regressions relating a firm's financial news media reading intensity at quarter *t–1* to financial spending recorded at *t*. For each visitor in our dataset, we use the estimated location of the reader to decompose the location of the reader with respect to the headquarters MSA of the firm, versus those readings done by someone not in the headquarters MSA. The dependent variable is R&D (Columns 1 and 2) and R&D+CAPX (Columns 3 and 4) spending at *t* scaled by assets at *t–1*. Reading intensity is log of ratio between count of financial news media reading at *t–1* and assets at *t–1*, where we count reading when the news is relevant to its own firm and the reader is located in headquarters MSA ($ReadNews^{Own,HQ}$), and when the news is relevant to its own firm but the reader is located outside of headquarters MSA ($ReadNews^{Own,Non\text{-}HQ}$). *Q* refers to a firm's Tobin's Q at *t–1*. Firm controls include *CF*, a firm's cash flow at *t–1* scaled by assets at *t–2*, and *1/AT*, the inverse ratio of a firm's assets at *t–1*. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Standard errors reported in parentheses are clustered by firm.

	R&D		**R&D+CAPX**	
	(1)	**(2)**	**(3)**	**(4)**
$Q \times ReadNews^{Own, HQ}$	0.0162^{***}	0.0133^{**}	0.0149^{**}	0.0114^{*}
	(0.0053)	(0.0053)	(0.0066)	(0.0065)
$Q \times ReadNews^{Own, Non\text{-}HQ}$	0.0009	0.0035	0.0076	0.0106
	(0.0058)	(0.0055)	(0.0081)	(0.0079)
Q	0.1177^{***}	0.0974^{***}	0.2239^{***}	0.2005^{***}
	(0.0195)	(0.0189)	(0.0228)	(0.0228)
$ReadNews^{Own, HQ}$	0.0086	0.0098	0.0043	0.0063
	(0.0078)	(0.0075)	(0.0124)	(0.0122)
$ReadNews^{Own, Non\text{-}HQ}$	–0.0013	–0.0033	–0.0050	–0.0073
	(0.0096)	(0.0089)	(0.0136)	(0.0129)
CF		-0.0219^{***}		-0.0086^{*}
		(0.0036)		(0.0050)
1/AT		0.0248^{***}		0.0266^{***}
		(0.0071)		(0.0087)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	26,091	25,798	26,043	25,757
Adjusted R^2	0.8852	0.8901	0.7846	0.7872

Table 7: Interpretation - financially constrained versus unconstrained firms

In this table, we compare learning among firms with varying degrees of financial constraints with firm-quarter panel regressions relating a firm's financial news media reading intensity and financial constraints at quarter *t–1* to financial spending recorded at *t*. The dependent variable is R&D+CAPX spending at *t* scaled by assets at *t–1*. Reading intensity is log of ratio between count of financial news media reading at *t–1* and assets at *t–1*, where we count reading when the news is relevant to its own firm ($ReadNews^{Own}$). *Q* refers to a firm's Tobin's Q at *t–1*. Firm controls include *CF*, a firm's cash flow at *t–1* scaled by assets at *t–2*, and *1/AT*, the inverse ratio of a firm's assets at *t–1*. *Constraint* is one of our seven financial constraint measures. To measure financial constraints, we create two composite financial constraint indices combining the KZ4 index, WW index, the Size-Age index of Hadlock and Pierce (2010), whether the firm is a dividend payer, and the firm's log market capitalization, all measured at quarter *t–1*. *Constraint* is configured such that its value is higher when a firm is more financially constrained. In Column 1, *Constraint* is the average of the five measures' z-score in each quarter. In Column 2, *Constraint* is the first principle component of the five measures' z-score in each quarter. In Columns 3 and 4, *Constraint* is z-score in each quarter for minus *MktCap* and minus *DivPos*, respectively. In Columns 5 to 7, *Constraint* is z-score in each quarter for *KZ4*, *WW*, and *SA*, respectively. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Standard errors reported in parentheses are clustered by firm.

	R&D+CAPX						
	AVG^z	**PCA^z**	**-MktCap^z**	**-DivPos^z**	**KZ4^z**	**WW^z**	**SA^z**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Q×ReadNews^{Own}	0.0175***	0.0169***	0.0198***	0.0152***	0.0168***	0.0192***	0.0130**
	(0.0051)	(0.0050)	(0.0053)	(0.0050)	(0.0051)	(0.0051)	(0.0052)
Q×ReadNews^{Own}×**Constraint**	0.0015	0.0011	–0.0025	0.0085	0.0129*	0.0060	–0.0001
	(0.0064)	(0.0030)	(0.0023)	(0.0056)	(0.0067)	(0.0041)	(0.0059)
Q	0.1959***	0.1849***	0.2044***	0.1847***	0.2121***	0.1840***	0.1798***
	(0.0220)	(0.0218)	(0.0230)	(0.0229)	(0.0228)	(0.0217)	(0.0228)
Q×**Constraint**	0.0547***	0.0334***	–0.0006	0.0347*	0.0247**	0.0161	–0.0031
	(0.0184)	(0.0105)	(0.0058)	(0.0205)	(0.0106)	(0.0111)	(0.0253)
ReadNews^{Own}	0.0057	0.0037	0.0032	0.0018	0.0059	0.0060	0.0000
	(0.0093)	(0.0090)	(0.0097)	(0.0095)	(0.0095)	(0.0096)	(0.0098)
ReadNews^{Own}×**Constraint**	0.0369**	0.0138**	0.0168**	–0.0013	–0.0619*	–0.0048	0.0385***
	(0.0153)	(0.0063)	(0.0074)	(0.0079)	(0.0334)	(0.0068)	(0.0116)
Constraint	0.0765	0.0479	0.1588***	–0.0132	–0.0573	0.0055	2.4522***
	(0.0763)	(0.0327)	(0.0504)	(0.0542)	(0.0509)	(0.0186)	(0.3810)
CF	–0.0065	–0.0063	–0.0085*	–0.0084*	–0.0091*	–0.0063	–0.0050
	(0.0054)	(0.0054)	(0.0049)	(0.0048)	(0.0053)	(0.0049)	(0.0050)
1/AT	0.0299***	0.0296***	0.0257***	0.0257***	0.0318***	0.0259***	–0.0121
	(0.0069)	(0.0069)	(0.0084)	(0.0084)	(0.0069)	(0.0088)	(0.0109)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,060	25,060	26,661	26,661	25,362	25,967	26,661
Adjusted R^2	0.7878	0.7879	0.7845	0.7845	0.7860	0.7883	0.7887

Table 8: Firm-specific, industry, and general learning

In this table, we compare learning from articles about own firms, peer firms, and general issues, with firm-quarter panel regressions relating a firm's financial news media reading intensity at quarter *t–1* to financial spending recorded at *t*. The dependent variable is R&D (Columns 1 and 2) and R&D+CAPX (Columns 3 and 4) spending at *t* scaled by assets at *t–1*. Reading intensity is log of ratio between count of financial news media reading at *t–1* and assets at *t–1*, where we count reading when the news is relevant to its own firm ($ReadNews^{Own}$), when the news is relevant to its baseline TNIC peers ($ReadNews^{Peer}$), and when the news is relevant to non-company entities only ($ReadNews^{General}$). *Q* refers to a firm's Tobin's Q at *t–1*. Firm controls include *CF*, a firm's cash flow at *t–1* scaled by assets at *t–2*, and *1/AT*, the inverse ratio of a firm's assets at *t–1*. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Standard errors reported in parentheses are clustered by firm.

	R&D		**R&D+CAPX**	
	(1)	**(2)**	**(3)**	**(4)**
Q×ReadNews^{Own}	0.0122**	0.0115**	0.0198***	0.0185***
	(0.0048)	(0.0050)	(0.0061)	(0.0062)
Q×ReadNews^{Peer}	0.0140**	0.0167***	0.0082	0.0117*
	(0.0057)	(0.0054)	(0.0073)	(0.0070)
Q×ReadNews^{General}	0.0089	0.0050	–0.0033	–0.0080
	(0.0071)	(0.0058)	(0.0096)	(0.0081)
Q	0.1029***	0.0884***	0.2119***	0.1938***
	(0.0211)	(0.0200)	(0.0247)	(0.0244)
ReadNews^{Own}	–0.0055	0.0009	–0.0204*	–0.0137
	(0.0073)	(0.0071)	(0.0107)	(0.0108)
ReadNews^{Peer}	–0.0103	–0.0132	–0.0162	–0.0198*
	(0.0088)	(0.0082)	(0.0118)	(0.0117)
ReadNews^{General}	0.1178***	0.0868***	0.1531***	0.1256***
	(0.0177)	(0.0161)	(0.0251)	(0.0236)
CF		–0.0236***		–0.0078
		(0.0040)		(0.0055)
1/AT		0.0300**		0.0301**
		(0.0120)		(0.0143)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	23,425	23,188	23,416	23,179
Adjusted R^2	0.8889	0.8932	0.7853	0.7868

Table 9: Negative, neutral versus positive feedback

In this table, we compare learning from articles with positive, neutral, and negative sentiment, with firm-quarter panel regressions relating a firm's financial news media reading intensity at quarter *t–1* to financial spending recorded at *t*. The dependent variable is R&D (Columns 1 and 2) and R&D+CAPX (Columns 3 and 4) spending at *t* scaled by assets at *t–1*. Reading intensity is log of ratio between count of financial news media reading at *t–1* and assets at *t–1*, where we count reading when the news is relevant to its own firm and news sentiment (RavenPack field event_sentiment_score) is negative or neutral ($ReadNews^{Own,Non\text{-}Pos}$), and when the news is relevant to its own firm and news sentiment is positive ($ReadNews^{Own,Pos}$). *Q* refers to a firm's Tobin's Q at *t–1*. Firm controls include *CF*, a firm's cash flow at *t–1* scaled by assets at *t–2*, and *1/AT*, the inverse ratio of a firm's assets at *t–1*. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Standard errors reported in parentheses are clustered by firm.

	R&D		**R&D+CAPX**	
	(1)	**(2)**	**(3)**	**(4)**
Q×ReadNews$^{Own, Non\text{-}Pos}$	0.0178***	0.0214***	0.0262***	0.0304***
	(0.0067)	(0.0063)	(0.0076)	(0.0072)
Q×ReadNews$^{Own, Pos}$	0.0016	0.0023	–0.0011	–0.0007
	(0.0063)	(0.0061)	(0.0074)	(0.0071)
Q	0.1271***	0.1046***	0.2361***	0.2103***
	(0.0195)	(0.0189)	(0.0229)	(0.0226)
ReadNews$^{Own, Non\text{-}Pos}$	0.0139*	0.0144**	–0.0086	–0.0049
	(0.0083)	(0.0073)	(0.0111)	(0.0104)
ReadNews$^{Own, Pos}$	–0.0095	–0.0083	0.0091	0.0096
	(0.0084)	(0.0078)	(0.0112)	(0.0111)
CF		–0.0220***		–0.0092*
		(0.0035)		(0.0049)
1/AT		0.0247***		0.0262***
		(0.0069)		(0.0085)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	26,998	26,703	26,949	26,661
Adjusted R^2	0.8856	0.8906	0.7819	0.7846

Table 10: Types of articles

In this table, we compare learning from reading full articles versus reading press releases and news flashes, with firm-quarter panel regressions relating a firm's financial news media reading intensity at quarter *t–1* to financial spending recorded at *t*. The dependent variable is R&D (Columns 1 and 2) and R&D+CAPX (Columns 3 and 4) spending at *t* scaled by assets at *t–1*. Reading intensity is log of ratio between count of financial news media reading at *t–1* and assets at *t–1*, where we count reading when the news is relevant to its own firm and the news is a full article ($ReadNews^{Own,Full\text{-}Article}$), and when the news is relevant to its own firm and the news is not a full article (mostly press releases and news flashes, $ReadNews^{Own,PR/Flash}$). *Q* refers to a firm's Tobin's Q at *t–1*. Firm controls include *CF*, a firm's cash flow at *t–1* scaled by assets at *t–2*, and *1/AT*, the inverse ratio of a firm's assets at *t–1*. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Standard errors reported in parentheses are clustered by firm.

	R&D		**R&D+CAPX**	
	(1)	**(2)**	**(3)**	**(4)**
$Q \times ReadNews^{Own,\ Full\text{-}Article}$	0.0125***	0.0167***	0.0147**	0.0193***
	(0.0047)	(0.0046)	(0.0061)	(0.0059)
$Q \times ReadNews^{Own,\ PR\ /\ Flash}$	0.0130**	0.0090*	0.0163**	0.0120*
	(0.0053)	(0.0053)	(0.0066)	(0.0065)
Q	0.1148***	0.0947***	0.2227***	0.1994***
	(0.0194)	(0.0188)	(0.0227)	(0.0227)
$ReadNews^{Own,\ Full\text{-}Article}$	0.0028	0.0038	–0.0018	0.0011
	(0.0068)	(0.0064)	(0.0098)	(0.0097)
$ReadNews^{Own,\ PR\ /\ Flash}$	0.0111	0.0109	0.0026	0.0014
	(0.0096)	(0.0090)	(0.0143)	(0.0145)
CF		–0.0218***		–0.0089*
		(0.0035)		(0.0049)
1/AT		0.0246***		0.0261***
		(0.0069)		(0.0085)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	26,998	26,703	26,949	26,661
Adjusted R^2	0.8857	0.8906	0.7820	0.7845

Appendix A.1: Merge Procedure with RavenPack

In this section, we describe the merge between RavenPack and our URL-level dataset. This merge proceeds in three steps. First, we build a database of URLs and headlines. Our dataset is at the URL level while RavenPack is at the headline level. Therefore, to merge with RavenPack, we require an intermediary dataset as RavenPack does not contain the original URL of an article. Second, we have to merge the URL-headline reading event database with RavenPack through the headline. by exactly matching on headline and day the article was published. Third, we perform data cleaning steps to ensure that the RavenPack story that we link an article-readership event to is the most appropriate. While our goal is to ensure accuracy and minimize the potential systematic bias in our merge procedure, some of our design choices are informed by computational scale as we must merge several datasets of billions of rows.

Building a Headline-URL database: First, to develop a headline-URL dataset, obtain two sources of data: Global Database of Events, Language, and Tone (GDELT) database, and Tiingo. Tiingo is a financial analytics data provider that caters to financial institutions. Institutional clients range from large pension and hedge funds to independent registered investment advisers (RIAs). One of the products Tiingo's provides is a news feed that records both headlines and URLs for articles across a wide range of financial news sites.

The GDELT Project is an open-source project supported by Google Jigsaw and monitors the world's broadcast, print, and web news in over 100 languages. By their own description, their dataset *"identifies the people, locations, organizations, themes, sources, emotions, counts, quotes, images, and events driving our global society every second of every day."* They collect millions of news articles on a daily basis and also record the URL and title of every article. We also collect various headline-URL datasets made available on Kaggle, a platform where scholars and companies often post datasets for participants to practice machine learning techniques against. We combine these three datasets to form an amalgamated date-

URL-headline dataset. If an article appears in two datasets, we use the headline from Tiingo, then GDELT, and then Kaggle.

We Ignore Frontpage Articles We focus on non-frontpage articles. It is difficult to know what exact article that is present on a front page at any given point in time, given that front pages change often. Moreover, given that investors do not specifically choose to read an article on the frontpage (but rather to check the website itself), it is more difficult to interpret reading about a firm on the front page of a website as the investor intends to pay attention to or acquire information about the specific stock. It may be a sheer coincidence that the investor happens to read about the firm on the front page at that particular time.

Joining to RavenPack: After joining against our URL-level database, we perform a match to RavenPack. We proceed in two steps. First, we perform an exact date-headline match between RavenPack and the master date-URL-headline dataset. We are able to match over half of all non-frontpage reading events via exact match. However, a considerable fraction is not exactly matched and requires us to perform a fuzzy match between the headline in our headline-URL dataset and RavenPack.

There are a number of reasons fuzzy matching of headlines may be necessary. First, RavenPack may record the headline in an article slightly differently. For example, consider the headline “Breaking News: Stocks Slated to End the Quarter on a Historic Run-Up”. In one dataset, the term “Breaking News” might be omitted as “Stocks Slated to End the Quarter on a Historic Run-Up”. A second reason fuzzy matching may be necessary is that headlines change during the day. For example, if the headline is “Breaking News: Stocks Slated to End the Quarter on a Historic Run-Up”, this headline can change to “Stocks Slated to End the Quarter on a Historic Run-Up”, or then finally later “Stocks End the Quarter on a Historic Run-Up”. Hence two different datasets may parse a given text similarly, but headlines are somewhat mutable.

For all remaining URLs, we perform a fuzzy match between RavenPack and the amalgamated dataset using 4-gram matching. We choose 4-gram matching because of the availability of computationally efficient algorithms to compute this. Given that we must merge tens of millions of headlines in our amalgamated datasets with over 400 million articles scraped in RavenPack, other approaches are not feasible. We retain all articles above 66%, which means at least two-thirds of all possible 4-grams match. We perform extensive spot-checking and the results suggest that 66% 4-gram similarity is a reasonable indication the two articles have the same subject.

De-duplication

At this step, for each unique URL, we have all potential RavenPack stories which could be potential matches for this URL. Even in the case of a headline that is exactly matched, sometimes we may have two matches from RavenPack. The first reason is that RavenPack may record two entries for the same story with the same headline. The second is that an article may be reprinted across different websites. For example, articles from the Associated Press are often re-printed across many different websites. One of our publishers is not directly licensed by RavenPack but re-prints its content through partner publishers with a minor delay.

Therefore, for every of our 11 billion events, we find what we consider to be the best match article in consideration of when the article was read. In principle, we consider the article closest to the event that comes before the event. We consider the RavenPack article with the closest timestamp to the event, conditional on the RavenPack article coming before the event. Finally, we notice that a number of URLs are not articles but rather searches for a specific stock on a financial news site. To the extent it is a quote lookup, we retrieve the ticker embedded in the URL and re-enter it into our dataset.

Final dataset

In the end, we match around 85% of reading events of non-frontpage articles. The missing articles are a combination of the inability to find a headline in our master URL-headline database, as well as a corresponding article from RavenPack. Upon visual inspection of some of the unlinked articles, a substantial fraction related to Covid, political news such as the election, or other non-value relevant events. Therefore, we believe the effective match rate to be much higher.

Appendix A.2: Sample selection

	≥14 (a)						Compustat (b)					
	N	**Mean**	**Std. Dev.**	**P25**	**Median**	**P75**	**N**	**Mean**	**Std. Dev.**	**P25**	**Median**	**P75**
R&D	27,515	1.2	2.22	0	0	1.55	62,351	1.36	2.69	0	0	1.53
R&D+CAPX	27,429	2.18	2.53	0.54	1.26	2.89	61,950	2.4	3.04	0.52	1.3	3.02
Q	27,288	2.43	1.9	1.19	1.72	2.88	60,882	2.35	1.82	1.19	1.67	2.76
CF	27,194	0.86	4.98	–0.03	1.78	3.18	60,781	0.02	6.68	–0.66	1.56	3
Size	27,515	6.97	2.13	5.61	7.13	8.45	62,351	6.59	2.2	5.09	6.73	8.08

Appendix B: Additional results

Table B.1: Missing R&D firms

In this table, we analyze the sensitivity of our results to missing R&D with firm-quarter panel regressions relating a firm's financial news media reading intensity at quarter *t–1* to financial spending recorded at *t*. Given that R&D is often missing (Koh and Reeb, 2015), we fill their scaled R&D (R&D scaled by lagged assets) by the average scaled R&D from firms in the same 2-digit SIC and year-quarter. The dependent variable is $R\&D^{fillavg}$ (Columns 1 and 2) and $R\&D^{fillavg}$+CAPX (Columns 3 and 4) spending at *t* scaled by assets at *t–1*. We include a dummy variable *𝟙(MissingR&D)*, which equals one when R&D is missing and zero otherwise. Reading intensity is log of ratio between count of financial news media reading at *t–1* and assets at *t–1*, where we count reading when the news is relevant to its own firm ($ReadNews^{Own}$). *Q* refers to a firm's Tobin's Q at *t–1*. Firm controls include *CF*, a firm's cash flow at *t–1* scaled by assets at *t–2*, and *1/AT*, the inverse ratio of a firm's assets at *t–1*. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Standard errors reported in parentheses are clustered by firm.

	$R\&D^{fillavg}$		$R\&D^{fillavg}$+CAPX	
	(1)	(2)	(3)	(4)
Q×$ReadNews^{Own}$	0.0242***	0.0221***	0.0273***	0.0246***
	(0.0083)	(0.0078)	(0.0089)	(0.0083)
Q	0.1815***	0.1553***	0.2868***	0.2588***
	(0.0340)	(0.0322)	(0.0358)	(0.0344)
$ReadNews^{Own}$	0.0064	0.0092	0.0009	0.0042
	(0.0175)	(0.0176)	(0.0195)	(0.0198)
𝟙(MissingR&D)	3.8753***	3.8456***	3.7557***	3.7257***
	(0.4016)	(0.4032)	(0.4111)	(0.4128)
CF		–0.0284***		–0.0127
		(0.0067)		(0.0077)
1/AT		0.0335***		0.0324***
		(0.0111)		(0.0120)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	25,035	24,761	24,990	24,722
Adjusted R^2	0.6458	0.6473	0.6171	0.6180

Table B.2: Relative to Rival Size

In this table, we analyze the incentive to learn by calculating a firm's relative to rival size with firm-quarter panel regressions relating a firm's financial news media reading intensity at quarter *t–1* to financial spending recorded at *t*. Rivals are defined as the top 10 firms closest on the Hoberg and Phillips product description textual similarity database (the Textual Network Industrial Classification database). $\boldsymbol{Sales^{RelToPeers}}$ is the average ratio of the focal firm's sales the prior year relative to the sales of peers in that same year, assuming the peers have positive sales the prior year. Reading intensity is log of ratio between count of financial news media reading at *t–1* and assets at *t–1*, where we count reading when the news is relevant to its own firm ($ReadNews^{Own}$). *Q* refers to a firm's Tobin's Q at *t–1*. Firm controls include *CF*, a firm's cash flow at *t–1* scaled by assets at *t–2*, and *1/AT*, the inverse ratio of a firm's assets at *t–1*. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Standard errors reported in parentheses are clustered by firm.

	R&D		R&D+CAPX	
	(1)	**(2)**	**(3)**	**(4)**
Q×ReadNews^{Own}	0.0160***	0.0151***	0.0213***	0.0201***
	(0.0046)	(0.0047)	(0.0056)	(0.0057)
Q×ReadNews^{Own}×**Sales**$^{\mathbf{RelToPeers}}$	–0.0100**	–0.0102**	–0.0103*	–0.0104*
	(0.0050)	(0.0049)	(0.0055)	(0.0054)
Q	0.1150***	0.1004***	0.2207***	0.2024***
	(0.0206)	(0.0193)	(0.0243)	(0.0238)
Q×**Sales**$^{\mathbf{RelToPeers}}$	0.0249*	0.0273**	0.0229	0.0258*
	(0.0142)	(0.0134)	(0.0148)	(0.0144)
ReadNews^{Own}	0.0092	0.0108	–0.0019	0.0008
	(0.0069)	(0.0066)	(0.0105)	(0.0106)
ReadNews^{Own}×**Sales**$^{\mathbf{RelToPeers}}$	0.0062	0.0046	–0.0014	–0.0023
	(0.0087)	(0.0088)	(0.0104)	(0.0105)
Sales$^{\mathbf{RelToPeers}}$	0.0221	0.0219	0.0254	0.0247
	(0.0244)	(0.0236)	(0.0259)	(0.0256)
CF		–0.0227***		–0.0077
		(0.0039)		(0.0055)
1/AT		0.0323***		0.0328**
		(0.0120)		(0.0145)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	23,441	23,224	23,431	23,215
Adjusted R^2	0.8885	0.8930	0.7848	0.7867

Table B.3: Intangible Investment and Total Q

In this table, we run firm-quarter panel regression with either intangible investment or total investment at *t* as the dependent variable on Total Q at *t–1*, our reading measure at *t–1*, the interaction term of the two, and control variables. The variables construction follows Peters and Taylor (2017). Appendix C provides detailed definitions for each variable. The dependent variable is intangible investment, ι^{int}, for Columns 1 and 2 which is 30% of SG&A plus R&D at *t* divided by total capital at *t–1*, and total investment, ι^{tot}, for Columns 3 and 4 which is 30% of SG&A plus R&D plus CAPX at *t* divided by total capital at *t–1*. Reading intensity is the log of the ratio between count of financial news media reading at *t–1* and total capital at *t–1*, where we count reading when the news is relevant to its own firm ($ReadNews^{Own}$). Q^{tot} refers to a firm's total Q at *t–1* from Peters and Taylor (2017). Firm controls include CF^{Ktot}, a firm's cash flow at *t–1* scaled by total capital at *t–2*, and $1/K^{tot}$, the inverse ratio of a firm's total capital at *t–1*. Regressions contain firm and year-quarter fixed effects. *Q* and *Read* are de-meaned. Standard errors reported in parentheses are clustered by firm.

	ι^{int}		ι^{tot}	
	(1)	**(2)**	**(3)**	**(4)**
Q^{tot}×$ReadNews^{Own, Ktot}$	0.0158**	0.0143**	0.0192**	0.0169**
	(0.0067)	(0.0063)	(0.0081)	(0.0077)
Q^{tot}	0.3616***	0.3311***	0.4861***	0.4505***
	(0.0289)	(0.0269)	(0.0329)	(0.0312)
$ReadNews^{Own, Ktot}$	0.0357***	0.0320***	0.0392***	0.0337***
	(0.0092)	(0.0087)	(0.0115)	(0.0113)
CF^{Ktot}		–0.0226***		–0.0102
		(0.0055)		(0.0069)
$1/K^{tot}$		0.0744***		0.0785***
		(0.0128)		(0.0161)
Firm FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Observations	24,621	24,390	24,588	24,358
Adjusted R^2	0.8895	0.8965	0.8094	0.8134

Table B.4: Implementing Measurement Error Correction

In this table, we re-run our main result in Table 4 using the Erickson, Jiang, and Whited (2014) higher moment estimator. We assume that all three of our main variables are mismeasured ($Q \times ReadNews^{Own}$, Q, *and* $ReadNews^{Own}$) and assume up to five higher moments.

	(1)	(2)	(3)	(4)
		R&D + CAPX		
Q	1.0594***	1.0651***	1.3998***	1.4270***
	(0.1313)	(0.1366)	(0.1229)	(0.2275)
CF		–0.0138**		–0.0034
		(0.0057)		(0.0106)
1/AT		0.0160		0.0244
		(0.0118)		(0.0286)
Observations	18,527	18,368	6,213	6,137
ρ^2	0.0814	0.0923	0.161	0.164
τ^2	0.230	0.220	0.174	0.154

Appendix C: Variable definition

Variable	Definition	Source
Reading Measures		
Common form	$\log(1000 \times Count_t/atq_t + 1)$.	The Data Partner, RavenPack, Compustat Quarterly
ReadNews^{Own}	*Count* is how many times the firm reads news relevant to its own firm with relevance scores greater than or equal to 90.	
ReadNews^{Own,Ktot}	Same as *ReadNews^{Own}* above except scaled by total capital K^{tot} instead of assets *atq*.	Peters and Taylor Total Q
ReadNews^{Peer}	*Count* is how many times the firm reads news relevant to peers defined by baseline TNIC with relevance scores greater than or equal to 90.	Hoberg and Phillips Data Library
ReadNews^{General}	*Count* is how many times the firm reads news relevant to non-company entities only with relevance scores greater than or equal to 90.	
NewsSupply	*Count* is number of news articles about a firm in RavenPack database with relevance scores greater than or equal to 90.	
Firm Fundamentals		
R&D	$100 \times xrdq_t/atq_{t-1}$, filled with zero when missing.	Compustat Quarterly
R&D+CAPX	$100 \times (xrdq_t + capxq_t)/atq_{t-1}$, where $xrdq$ is filled with zero when missing, and $capxq$ is derived from year-to-date CAPEX measure $capxy$.	Compustat Quarterly
ι^{int}	Intangible investment in Peters and Taylor (2017). $\left(R\&D_t + (0.3 \times SG\&A_t)\right)/K_{t-1}^{tot}$.	Compustat Quarterly, Peters and Taylor Total Q
ι^{tot}	Total investment in Peters and Taylor (2017). $(R\&D_t + (0.3 \times SG\&A_t) + CAPX_t)/K_{t-1}^{tot}$.	Compustat Quarterly, Peters and Taylor Total Q
Q	Tobin's Q. Market value of assets divided by book value of assets. $(atq + mkvaltq - ceqq)/atq$.	Compustat Quarterly
Q^{tot}	Total Q. Market value of assets divided by total capital. $(cshoq \times prccq + dlttq + dlcq - actq)/K^{tot}$	Compustat Quarterly, Peters and Taylor Total Q
CF	Cash flow. Income before extraordinary items and depreciation scaled by lagged assets. $100 \times (ibq_t + dpq_t)/atq_{t-1}$.	Compustat Quarterly
CF^{tot}	Cash flow scaled by total capital. $100 \times (ibq_t + dpq_t)/K_{t-1}^{tot}$	Compustat Quarterly, Peters and Taylor Total Q
1/AT	Inverse of assets (measured in \$billions). $1000/atq$.	Compustat Quarterly
$1/K^{tot}$	Inverse of total capital (measured in \$billions). $1000/K^{tot}$, where total capital, K^{tot}, is the sum of knowledge capital (perpetual inventory method on R&D with 15% effective annual depreciation rate), organizational capital (perpetual inventory method on 30% of SG&A with 20% effective annual depreciation rate), as well as intangible assets and PP&E from Compustat Quarterly.	Compustat Quarterly, Peters and Taylor Total Q

Insider	Insider trading volume divided by total volume.	CRSP, Refinitiv Insiders
Analyst	Sum of 1. Number of analysts issuing recommendations. 2. Number of analysts issuing earnings forecast. 3. Number of analysts issuing price targets.	I/B/E/S
Size	Logarithm of assets. $\log(atq)$.	Compustat Quarterly
Leverage	Book leverage. Sum of short-term and long-term debt, divided by sum of short-term debt, long-term debt, and shareholders' equity. $100 \times (dlttq + dlcq)/(dlttq + dlcq + seqq)$.	Compustat Quarterly
Tangibility	PP&E scaled by assets. $100 \times ppentq/atq$.	Compustat Quarterly
Profitability	Gross profitability. $100 \times saleq/atq$.	Compustat Quarterly
Volatility	Standard deviation of daily idiosyncratic returns (e.g. residuals) in a quarter from Fama and French 5-factor model (Fama and French, 2015).	Fama and French Data Library
Return4Q	Return from the beginning of quarter $t - 3$ to the end of quarter t.	CRSP
QuotedSpread	Average of daily quoted spread, measured in %.	TAQ
InstitutionOwn	Percentage of institutional ownership.	Thomson-Reuters 13F
ProdMktFluidity	Product market fluidity from Hoberg, Phillips, and Prabhala (2014), a proxy for product market competition.	Hoberg and Phillips Data Library
Cross-sectional Traits		
MktCap	Logarithm of market capitalization.	Compustat Quarterly
DivPos	1 if the firm pays cash dividend to common or preferred shares, 0 otherwise. $1\ if\ (dvc + dvp) > 0, 0\ otherwise$.	Compustat Annual
KZ4	Four-variable KZ index from Kaplan and Zingales (1997) and Baker, Stein, and Wurgler (2003).	Compustat Quarterly
WW	WW index from Whited and Wu (2006).	Compustat Quarterly
SA	Size-Age index from Hadlock and Pierce (2010)	Compustat Quarterly
AVG^z	Average of z-scores in a quarter of minus *MktCap*, minus *DivPos*, *KZ4*, *WW*, and *SA*.	Compustat
PCA^z	The first principal component of z-scores in a quarter of *MktCap*, *DivPos*, *KZ4*, *WW*, and *SA*.	Compustat
$Sales^{RelToPeers}$	Median of $sale/sale^{peer}$ among the top 10 baseline TNIC peers.	Compustat Annual