Earnings Guidance and Price Informativeness: The Role of Media*

Yunus Topbas†

December 30, 2021

Abstract

This paper investigates how the contrasting trends in media coverage and earnings guidance have affected stock price informativeness over the past two decades. I develop a model to understand what trade-offs investors face when acquiring information through the media and how earnings guidance changes those trade-offs. In the model, the effects of media coverage and earnings guidance on price informativeness are ambiguous. To resolve this ambiguity, I empirically test which predictions of the model are supported by the data. I show that, contrary to the common belief, high media coverage can cause lower price informativeness. Motivated by the model, I propose a new empirical measure of media coverage, which better gauges the informational contents of news. I also find that the impact of earnings guidance on price informativeness depends on media coverage. Earnings guidance improves price informativeness only at high levels of media coverage.

Keywords: Media, Price Informativeness, Earnings Guidance

*I would like to thank Robert Korajczyk, Alessandro Pavan and Paola Sapienza for their continuous guidance and support. For helpful comments and suggestions, I am also grateful to Cassiano Alves, Daniel Barron, Onsel Gurel Bayrali, Matteo Camboni, Carlos da Costa, Ramazan Dababi, Joao Guerreiro, Arzu Isik, Annie Liang, Joel Mokyr, Laura Murphy, Charlie Nathanson, Michael Porcellacchia, Jeroen Swinkels, Mehmet Nazim Tamkoc, and seminar participants at Northwestern University.

† The Department of Economics, Northwestern University, 2211 Campus Drive, Evanston, IL 60208, YunusTopbas2020@u.northwestern.edu.
1 Introduction

By aggregating different sources of information, stock prices act as a public signal to investors, potentially influencing investors’ decisions. There is no consensus on whether the informational content of stock prices (price informativeness) has increased over time. The literature has focused on how trading strategies or stock characteristics affect stock price informativeness. For instance, Weller (2018) and Sammon (2021) provide evidence that increased algorithmic trading and passive ownership have decreased price informativeness over time, respectively, whereas Dávila and Parlatore (2018) and Farboodi et al. (2020) demonstrate that market capitalization increases price informativeness.

This paper contributes to the literature by considering how price informativeness is affected by the amount of information available to investors. From this perspective, the past two decades featured two contrasting trends. On the one hand, media coverage expanded dramatically: on a yearly average basis, a publicly traded US company appeared in financial news 100 times in 2000 compared to 1200 times in 2019.\(^1\) On the other hand, the fraction of publicly traded US companies issuing earnings guidance decreased from 32% in 2000 to 14% in 2019.\(^2\)

More media coverage is generally considered good for investors and stock markets. It is based on the idea that more media coverage means more information available in the media. This idea is correct, but incomplete. More media coverage also brings more replicas: the news-releases that contain very similar information. At least 70% of the increase in media coverage comes from replicas and there is a great heterogeneity in the number of replicas across stocks. That is, media coverage has two dimensions: information and replicas. Consider two stocks with similar levels of media coverage: one is dominated by replicas, one is dominated by informative news. Media coverage has different meanings for these two stocks. Therefore, both dimensions should be taken into account when measuring media coverage and analyzing its relation with stock prices and other information sources such as earnings guidance. This is the missing part in the literature.

\(^1\)I define media coverage as the total number of financial news-releases that appear in written and online official news providers and stored by my data source—Ravenpack Analytics (RPA). RPA tracks and monitors financial news that contains relevant information and delivers event data about 200,000 companies in a consistent, structured format. I exclude perfect duplicates: if the same article appears in both written and online sources, I include only one of them. Social media or unofficial blogs such as Reddit are not included.

\(^2\)Earnings guidance, a.k.a. forward-looking statements, is the information issued by the management of a publicly traded company regarding its expected future earnings and sales. I include firms that appear on I/B/E/S data source.
In this paper, I construct a new empirical measure of media coverage that explicitly takes into account both information and replicas in the media, and I use it to examine how the contrasting trends in information availability have affected stock price informativeness over time. I first develop a rich theoretical framework to understand what trade-offs investors face when acquiring information through the media in the presence of replicas and how earnings guidance changes those trade-offs. I then use the intuitions from the model to construct a new empirical measure of media coverage. Finally, I construct empirical analogues of price informativeness measures implied by the model and test the model’s predictions.

In my model, the media play a role in the information creation process. They report financial news that contains useful information to predict stock values. I define media coverage as the total number of news-releases in the media. Investors incur two different monetary costs to acquire information through the media. They need to pay a subscription fee to access news, and once accessed they have to choose how many news-releases to analyze. The media have both informative news and replicas. A replica contains very similar information with another informative new-release. It may not be ideal for an investor to analyze both an informative news-release and its replica(s). However, investors do not know which news-releases are replicas when selecting news-releases to analyze.\(^3\)

I assume that total information available in the media increases with media coverage. This assumption implies that in an unconstrained world in which every investor has enough budgets to analyze every piece of news, higher media coverage always provides (weakly) better underlying information to investors and hence it monotonically improves price informativeness. However, in my setting, investors have limited budgets for information acquisition. There are two types of investors: small-budget and large-budget. The latter have considerably more budgets to spend on information acquisition.

In the baseline model, with no earnings guidance, here is the main theoretical finding: media coverage has a hump-shaped impact on stock price informativeness under two conditions: (C1) as media coverage increases, the number of replicas increases more than the number of informative news, (C2) the fraction of small-budget investors in the economy is sufficiently high. If these two conditions hold, increasing media coverage improves price informativeness first, but after media coverage goes beyond a level, increasing media coverage

\(^3\)Information analysis cost accounts for the fact that having access to news outlets does not automatically provide information to investors in practice. They need to hire research analysts and provide them with equipment to analyze news and extract useful information.
reduces price informativeness.\footnote{I use price informativeness measures that are proposed by Sammon (2021). I will discuss these measures in detail in Section 3. These measures are positively correlated with Grossman and Stiglitz (1980)’s price informativeness definition— the inverse of conditional variance of fundamental given stock prices.}

The intuition is as follows. Stock price informativeness is a function of how many investors acquire information (quantity of information) and the accuracy of information acquired (quality of information). Large-budget investors have enough resources to analyze every piece of news. At low levels of media coverage, small-budget investors’ budget constraints are not binding. They can analyze as much news as they want, like large-budget investors. Because total information available in the media increases with media coverage, the quality of information that an unconstrained investor acquires in equilibrium also increases with media coverage.\footnote{In equilibrium, an unconstrained investor (an investor with a non-binding budget constraint) does not have to analyze every piece of news to improve her acquired information quality. As long as the rate of increase in total information in the media is sufficiently high relative to information analysis cost, the quality of information that she acquires in equilibrium and media coverage are positively related. I am focusing on these cases.} This dynamic incentivizes more investors to acquire information. Therefore, more media coverage increases both quantity and quality of information in the economy and hence it increases stock price informativeness.

In contrast, at high levels of media coverage, small-budget investors’ budget constraints are binding. They can analyze only a subset of news, and this subset contains more replicas as media coverage increases (due to C1). They pay the same amount of money but receive lower-quality signals. In this case, increasing media coverage discourages small-budget investors from acquiring information through the media. If the fraction of small-budget investors is sufficiently high in the economy (i.e., C2), then more media coverage reduces stock price informativeness.

I extend the baseline model by incorporating earnings guidance into the model to study how it modifies investors’ incentives to purchase and analyze news. Earnings guidance is an additional source for investors to learn about stock values. It affects price informativeness through two competing channels. First, because earnings guidance is available to investors free of charge, every investor uses it to update her belief about stock values. To this end, it directly injects more fundamental information into stock prices and hence improves price informativeness. Second, given that information acquisition through the media is costly, introducing a free alternative, i.e., earnings guidance, decreases investors’ incentives to purchase and analyze news. This crowding-out effect deteriorates price informativeness.
The net impact of earnings guidance on price informativeness depends on media coverage and distribution of investors’ types. In particular, under conditions C1 and C2, I find that earnings guidance improves price informativeness at high levels of media coverage, and reduces price informativeness at low levels of media coverage. The intuition is as follows. Sufficiently high media coverage induces small-budget investors to stop paying attention to the media. In this case, crowding-out effects are zero for them. There are still crowding-out effects for large-budget investors. But, if the economy has a sufficiently large number of small-budget investors, then earnings guidance improves price informativeness. In contrast, at low levels of media coverage, crowding-out effects exist for both small- and large-budget investors. Earnings guidance causes all investors to divert attention from the media: they either cease to purchase news or analyze less news. In this case, earnings guidance deteriorates price informativeness.

The analysis shows that the effects of media coverage and earnings guidance on price informativeness can be ambiguous. The results hold only under conditions C1 and C2. If one of them is violated, media coverage monotonically increases price informativeness, and earnings guidance monotonically increases or decreases price informativeness depending on the parameters of the model. The natural way to resolve this ambiguity is to calibrate the model to match the data. Neither information acquisition costs nor investors’ budgets, however, are readily observable. Therefore, I create empirical analogues of price informativeness measures implied by the model (proposed by Sammon (2021)) to test which predictions of the model are supported by the data. But first I use the intuitions from the model to construct a new empirical measure of media coverage that explicitly takes into account both information and replicas in the media.

There are two common measures of media coverage in the literature. The first measure counts the number of news-releases in a single outlet like Wall Street Journal (WSJ) or New York Times (NYT). But some stocks are not covered by those outlets at all. Therefore, this measure underestimates media coverage for some stocks. The second measure counts all news-releases in the media. However, across stocks, there is a great heterogeneity in the number of replicas.⁶ My new measure fixes these problems to a great extent by taking

---

into account three different dimensions of media coverage: cross-outlet coverage dispersion, replicas and information. First, I identify how many news outlets would be enough for an investor to get almost all the information in the media about a particular stock. These outlets can differ from one stock to another. Second, I identify replicas and distinguish between different types of replicas. Third, I calculate the fraction of non-replicas among all news-releases. Higher fraction increases the likelihood of informative news.

I use my new measure of media coverage to analyze the impact of media coverage on stock price informativeness. In cross-sectional regressions, I find that media coverage has a hump-shaped impact on price informativeness: as media coverage increases, price informativeness first increases and then declines. I employ the method proposed by Lind and Mehlum (2010) to test the existence of a hump-shaped relationship. I also find that earnings guidance improves price informativeness only at high levels of media coverage. These findings are consistent with the presence of conditions C1 and C2 in practice.

C1 states that as media coverage increases, the number of replicas increases more than the number of informative news-releases. This aspect is observable and indeed present in the data. C2 states that the economy contains sufficiently large number of small-budget investors. This aspect is not directly observable. I use assets under management as a proxy for investors’ budgets. The idea is that a large institutional investor in terms of assets under management can allocate more resources to extract information from news compared to a small institutional investor. I find that media coverage almost monotonically increases stock price informativeness for stocks traded mostly by large institutional investors, whereas the hump-shaped impact is preserved for small and medium institutional investors. It provides suggestive evidence to the hypothesis that budget constraints play roles in having a negative impact of media coverage on price informativeness.

Sammon (2021) demonstrates that average price informativeness has decreased over the past twenty years. My results from cross-sectional regressions show that media coverage and lack of earnings guidance are negatively correlated with stock price informativeness in high media coverage environments. These findings suggest the trends in information availability can contribute to lower price informativeness over the past two decades. To provide further evidence, I investigate how the relation between media coverage and price informativeness has changed over time by using recursive regressions.

Information technology has been improving since 2000. It has two opposing effects on
investors. On the one hand, improvements in information technology have been reducing the cost of acquiring and analyzing information for investors (e.g. Kacperczyk, Nosal, and Stevens (2019)). In this respect, technological improvements can allow investors to analyze more news utilizing the same amounts of budgets. On the other hand, improved technology has also made news producers more efficient. They have managed to significantly increase the number of news-releases using fully or semi-automated systems. For instance, Bloomberg News started to use some form of automated technology, called Cyborg, to assist reporters in churning out thousands of articles on company earnings reports each quarter.7 An increase in the number of news-releases can make it more difficult for investors to identify which news-releases contain new information.

If the increase in the media’s productivity is disproportionately greater than the increase in investors’ productivity, a technological improvement creates a productivity gap between investors and the media. In the model, under conditions C1 and C2, I find that higher productivity gaps reduce price informativeness by distorting small-budget investors’ incentives to analyze news. I run recursive regressions over different time periods to test this prediction of the model. The regressions illustrate that the hump-shaped impact of media coverage on price informativeness is preserved over time. However, the negative impact of media coverage on price informativeness became more prominent. Media coverage has been negatively affecting more and more stocks over time in terms of informativeness of their prices. That is, from 2000 to 2019, it became detrimental to stock price informativeness for more stocks. Furthermore, in recursive regressions, the impacts of earnings guidance on price informativeness are qualitatively similar to its impacts in the cross-sectional regressions: earnings guidance improves price informativeness only at sufficiently high levels of media coverage.

My paper sheds light on the role of the media and earnings guidance regarding price informativeness and shows that increased media coverage and lack of earnings guidance may lead to lower price informativeness in massive media coverage environments. My results are important in consideration of the fact that, since 2000, media coverage has been constantly increasing and the number of publicly traded firms issuing earnings guidance has been decreasing. I provide suggestive evidence that these two contrasting trends can play a role in the decreasing trends in stock price informativeness over time.

The findings also contribute to a recent debate regarding the role of earnings guidance in communicating information to the market and whether releasing quarterly earnings guidance

---

indeed provides better underlying information. This debate has called for discontinuing the practice of releasing quarterly earnings guidance because it is thought that earnings guidance has shifted the focus of companies to short-term tactics at the expense of long-term value. But what is the cost of reducing this communication to the market? My paper shows that for stocks with high media coverage, removing earnings guidance would be costly for investors since stock prices would become less informative. More information would be revealed on earnings dates, and accordingly there would be higher stock price jumps and volatility. Besides, the model indicates that it is more costly for small-budget investors. Hence, removal of earnings guidance can widen information asymmetries between small- and large-budget investors.

Related Literature

The Role of Media in Financial Markets. I make two contributions to this literature. First, I propose a new empirical measure of media coverage, which takes into account cross-outlet coverage dispersion as well as the heterogeneity in the distribution of replicas across stocks. There are two common measures in the literature. The first one is to include all news in a single news outlet: for instance, Davies and Canes (1978), Nofsinger (2001), Antweiler and Frank (2004), Tetlock (2007) focus on WSJ; Klibanoff et al. (1998), García (2013) analyze NYT. This measure does not take into account cross-outlet coverage dispersion: some stocks have pretty high coverage in large outlets such as Bloomberg Terminal or Dow Jones Newswire despite the fact that they receive little coverage in WSJ or NYT. The second common measure is to include all news in a newswire or news aggregator. For instance, Tetlock et al. (2008), Tetlock (2010), Rogers et al. (2016) analyze Dow Jones Newswires, Bushee et al. (2010) examine Factiva, Bonsall et al. (2020) include all news in Ravenpack Analytics. This measure does not consider the heterogeneity in the distribution of replicas across stocks, which is an important determinant of overall informativeness of financial news. My second contribution to this literature is documenting novel empirical results about how the media affect informational content of stock prices and present a theoretical framework that helps understand the mechanism.

Earnings Guidance and Investors. Whether releasing quarterly earnings guidance provides better underlying information has extensively been studied in the literature on voluntary disclosures. Ajinkya and Gift (1984), Coller and Yohn (1997), Verrecchia (2001) find that voluntary disclosures reduce information asymmetry between insiders and outside investors. Graham, Harvey, and Rajgopal (2005) demonstrate that earnings guidance reduces stock price volatility and surprises on earnings announcement dates. They also argue that
earnings guidance attracts investors that focus on short-term earnings. Houston, Lev, and Tucker (2010) and Kim, Su, and Zhu (2017) provide further support for investors’ short-termism resulting from earnings guidance. Call et al. (2014) find that earnings guidance decreases managerial incentives for earnings management. There are also papers examining the relation of earnings guidance with the media. Bae et al. (2016), Bae, Litjens, and Zeng (2019) find that higher dissemination of earnings guidance in the media encourages managers to continue issuing it. Twedt (2016) finds that newswire dissemination of earnings guidance causes larger initial price reactions and faster incorporation of guidance information into stock prices. My paper points out that whether suspending earnings guidance harms investors or not can be a function of media coverage. If a stock attracts too much attention from the media, earnings guidance can be useful to guide small-budget investors’ expectations.

Stock Price Informativeness. Bai, Philippon, and Savov (2016) and Dávila and Parlatore (2018) use cash flow variables as a proxy for stock fundamentals and document that price informativeness increased over the last three decades. In a structural framework, Farboodi et al. (2020) show that these increasing trends are applicable to only large/growth stocks. I use price informativeness measures proposed by Sammon (2021). It provides evidence that increased passive ownership over time has caused lower price informativeness since the 1990s. Using a similar measure, Weller (2018) shows that increased algorithmic trading is another contributing factor to lower price informativeness over time. My paper documents that increased media coverage and companies’ suspension of earnings guidance can also play crucial roles in the decreasing trends in price informativeness over time.

Technological Progress and Information Acquisition in Financial Markets. Peress (2005) shows that a low information cost discourages investors from participating in stock markets. Miher (2019) show financial innovations cause poorer investors not to invest in stock markets and hence increase wealth inequality. Kacperczyk et al. (2019) demonstrate that improvements in information processing technologies can result in greater capital income inequality by disproportionately benefiting the initially more skilled investors. Farboodi and Veldkamp (2020) find that a rise in information processing productivity can induce investors to focus more on extraction of information from order flows rather than conducting fundamental analyses of firms’ profitability. Malikov (2021) argues that falling information acquisition costs lead investors to invest more in passive funds. Pavan, Sundaresan, and Vives (2021) show that improvements in information technology eventually create an economy with excessive information acquisition and this inefficiency cannot be addressed by
policies based on stock prices and volumes without jeopardizing trade efficiency. My paper focuses on the consequences of technological improvements when they affect information producers and investors disproportionately.

**Public Information and Crowding-out.** It has been well-documented that public information crowds out usages of other information sources, e.g. Diamond (1985), Morris and Shin (2005), Amador and Weill (2010, 2012), Llosa and Venkateswaran (2012), Gao and Liang (2013). Colombo, Femminis, and Pavan (2014) show that more accurate public information crowds out endogenous private information acquisition regardless of whether actions are strategic complements or substitutes. Goldstein and Yang (2017) find that firms could prefer to disclose less due to crowding-out effects of public disclosures. Crane, Crotty, and Umar (2018) argue that if public information complements investors’ private information, it does not necessarily crowd out information acquisition. In my model, earnings guidance serves as public information and has similar crowding-out effects on investors’ information acquisitions behaviors. However, in my setting, public information can also remedy unintended consequences of technological improvements by providing small-budget investors with better underlying information.

**Organization.** The rest of the paper is organized as follows. Section 2 describes the model. Section 3 maps the model into data and introduces media coverage and price informativeness measures. Section 4 presents empirical results from cross-sectional regressions. Section 5 conducts recursive regressions to examine the relation of media coverage with price informativeness over time. Section 6 discusses the robustness of the results and Section 7 concludes.

## 2 Model

I first present a model without earnings guidance. It is useful to understand what trade-offs investors face when acquiring information through the media.

### 2.1 Key Model Ingredients and Timing

The economy has a single risky asset, which I call stock. It is homogeneous and perfectly divisible. The economy consists of information producers and investors. Information producers are news outlets that provide financial news-articles to investors in return for a subscription fee. The news-articles contain information that can potentially be used to predict the value
of the stock.

The model has three periods. At $t = 0$: (i) information producers make entry decisions in the information market, (ii) investors decide whether or not to subscribe to a news outlet. At $t = 1$: (i) those who subscribe (informed investors) privately observe news-articles, choose number of news-articles to analyze, and receive private signals accordingly, (ii) all investors submit their demand schedules. At $t = 2$: market clears, equilibrium price is formed, trades are executed, and payoffs are realized.

2.2 Information Market

The information market is similar to Perloff and Salop (1985)’s monopolistic competition model. There are $m = 1, ..., M$ news outlets that report financial news. Each investor attaches relative values to these outlets according to a preference vector $b_i = (b_{i1}, ..., b_{iM})$. $b_{im}$ can be interpreted as total benefits that investor $i$ gets through subscribing to news outlet $m$. An investor subscribes to the outlet among those available that maximizes her net surplus

$$b_{im} - p_m, \ m = 1, ..., M$$

where $p_m$ is the price that outlet $m$ charges to its subscribers (subscription price). For simplicity, preferences are assumed to be symmetric in the sense that aggregate preferences for each particular outlet $m$ are independent and identically distributed with a density function $g(b)$.

Each producer allocates a total resource of $R$ to report news about the stock. However, an entry into the information market requires a fixed cost. Each outlet decides whether to pay a fixed cost $\bar{R}$ to enter the market at $t = 0$. Once entered, she reports $N(R - \bar{R}, \theta_s)$ financial news-articles about the stock where $\theta_s$ is her productivity in state $s$, and $s$ is a state variable measuring state of technology.\footnote{I assume $N(R - \bar{R}, \theta_s)$ is exogenous. The qualitative results do not hinge on this assumption. It allows me to focus on the impacts of media coverage on investors’ acquisition behavior. I relax this assumption in the Online Appendix.} Besides, there is a free entry into the information market and news outlets compete over prices.

I assume $N_{R-\bar{R}} > 0$. In my setting, the fixed entry cost should not be interpreted as establishing a new firm. It is more of a transfer of resources. In practice, a news outlet reports news about many different stocks. She already exists in the news industry. However,
she can find it profitable to transfer a fixed resource $\bar{R}$ to report news about stock $A$ instead of stock $B$. The fixed entry cost in my model is closer to this type of transfer. I also assume $N_{\theta_i} > 0$. That is, a news outlet can report more news-articles using the same amount of resources when productivity is higher.

News outlet $m$’s profit is given by

$$\pi_m := (p_m - c)Q_m(p_m, p_{-m}) - R$$

(1)

where $p_{-m}$ is the vector of other outlets’ prices, $c$ is marginal cost of selling news, and $Q_m(p_m, p_{-m})$ is fraction of investors that $m$ is selling at $p_m$.

### 2.3 Information Acquisition

At $t = 0$, investor $i$ subscribes to 0 or $\bar{M}$ news outlets where $\bar{M}$ is an exogenous positive number. This simplification is needed to achieve tractability in the information market. In the Online Appendix, I will relax this assumption and endogenize the number of outlets that investors subscribe. I first analyze the model for $\bar{M} = 1$ and then generalize it. An investor $i$ is labeled (informed) (uninformed) investor if she chooses (to) (not to) to subscribe to a news outlet.

**Acquisition cost.** Once subscribed, an informed investor endogenously chooses $N_i \leq N$ news-articles to analyze, which requires $C(N_i, \theta_{is})$ dollars. $\theta_{is}$ is investor-specific productivity of analyzing news-articles in state $s$. I will call $C(N_i, \theta_{is})$ information processing cost or simply processing cost. I define acquisition cost ($c_{ims}$) as information processing cost net of surplus,

$$c_{ims} := C(N_i, \theta_{is}) + p_m - b_{im}$$

(2)

That is, acquisition cost is cost of analyzing news plus subscription fee net of benefits. Remember that $p_m$ is outlet $m$’s subscription price, $b_{im}$ represents total benefits of subscription, and $s$ is state of technology. I assume $C'_{N_i} > 0, C'_{\theta_{is}} < 0$.

**Capacity constraints.** In my framework, investors have finite information processing capacities. I am motivated by the fact that investors have limited time, attention and budget to analyze news. An investor $i$ cannot spend more than her initial resources, $\bar{c}_i$ on subscribing and processing information. That is, $c_{ims} - \bar{c}_i \leq 0$. There are two types of
investors. *High-type* investors have a productivity of \( \theta_{is} = \theta_{hs} \) and are indexed by \( i \in [0, 1] \). *Low-type* investors have a productivity of \( \theta_{is} = \theta_{ls} \) and are indexed by \( i \in [1, \mu] \) where \( \theta_{hs} > \theta_{ls} \) in each state.

**Private Signals.** If an investor \( i \) subscribes to a news outlet, she privately observes a noisy signal \( \tilde{s}_i \) of the stock value,

\[
\tilde{s}_i = \bar{v} + \tilde{\varepsilon}_i
\]

where \( \tilde{\varepsilon}_i \sim \mathcal{N}(0, \tau_{\tilde{\varepsilon}_i}^{-1}) \). The precision of private signal \( \tau_{\tilde{\varepsilon}_i} \) is determined by (i) number of news-articles processed—\( N_i \), and (ii) total number of news-articles in an outlet—\( N \). In particular,

\[
\tau_{\tilde{\varepsilon}_i} = f(N_i, N)
\]

\( N \) appears in the precision function to control information dispersion across news within an outlet. I will explicitly define the precision function after introducing the signal structure.

**Signal structure.** Suppose total number of news-articles in an outlet can be decomposed into \( K \) partitions and each partition contains \( n \geq 1 \) news—\( N = nK \). Each partition represents a payoff relevant factor that is independent of other factors. Let \( n_{jk} \) denote the \( j^{th} \) news-article in partition \( k \) and \( \tilde{\eta}_{jk} \) denote the signal induced by the news-article \( n_{jk} \). Suppose \( \tilde{\eta}_{jk} \sim \mathcal{N}(0, \sigma_{\tilde{\eta}}^2(N)) \) where \( \sigma_{\tilde{\eta}}^2(N) \) is a finite function of \( N \) and

\[
\text{Corr}(\tilde{\eta}_{j_1 k_1}, \tilde{\eta}_{j_2 k_2}) = \begin{cases} 
0 & \text{for each } j_1 \in k_1 \text{ and } j_2 \in k_2 \text{ if } k_1 \neq k_2 \\
\rho_N & \text{for each } j_1 \in k_1 \text{ and } j_2 \in k_2 \text{ if } k_1 = k_2 \text{ and } j_1 \neq j_2.
\end{cases}
\]

(5)

That is, news-articles in one partition are uncorrelated with news-articles in a different partition, and news-articles in the same partition have a correlation ratio of \( \rho_N \).\(^9\) Assuming signals from news-articles are additively separable, private noise can be re-written as

\[
\tilde{\varepsilon}_i = \tilde{\eta}_1 + \ldots + \tilde{\eta}_K + \tilde{\varepsilon}_{\eta}
\]

(6)

where \( \tilde{\eta}_k := \sum_j \tilde{\eta}_{jk} \) and \( \tilde{\varepsilon}_{\eta} \sim \mathcal{N}(0, \sigma_{\tilde{\varepsilon}_{\eta}}^2) \) is i.i.d. representing idiosyncratic disturbances.

---

\(^9\)I indeed do not need partitions. It is easy to show that this partition economy can be mapped into an equivalent economy where each news-article is correlated with each other at a ratio of \( \rho_N^E \). The partition economy provides a nicer characterization of private signal precision.
The variance of partition $k$ is

$$\text{Var}(\eta_k) = n\sigma^2_{\eta}(N) + n(n-1)\rho_N\sigma^2_{\eta}(N). \quad (7)$$

An investor can reduce the variance of $\eta_{jk}$ to $\sigma^2_{\eta}(N) - \delta(N)$ by processing news $n_{jk}$, that is, $\text{Var}(\tilde{\eta}_{jk} | n_{jk}) = \sigma^2_{\eta}(N) - \delta(N)$. If an investor processes $K_i$ partitions, she observes a private signal with variance

$$\tilde{\sigma}^2_{\varepsilon_i} := \text{Var}(\tilde{\varepsilon}_i | K_i) = \left(1 - \frac{\delta(N)}{\sigma^2_{\eta}(N)} \frac{N_i}{N}\right)(N\sigma^2_{\eta}(N)) \left[1 + (n-1)\rho_N\right] + \sigma^2_{\varepsilon_i} \quad (8)$$

where $N_i = nK_i$. The *precision of private signal* $f(N_i, N)$ is then $1/\tilde{\sigma}^2_{\varepsilon_i}$. I assume $\rho_N \in (\frac{1}{1-n}, 1]$ to ensure that $\partial f(N_i, N)/\partial N_i \geq 0$.

$\delta(N)/\sigma^2_{\eta}(N)$ takes values between zero and one. Hence, it can be interpreted as the probability of observing new information in the news. That is, when an investor processed $N_i$ news-articles, only $(\delta(N)/\sigma^2_{\eta}(N))N_i$ of them would provide new information. Using this interpretation, $((\delta(N)/\sigma^2_{\eta}(N))N_i)/N$ is similar to *success rate*. Let $S(N_i, N)$ denote the success rate: $S(N_i, N) := [(\delta(N)/\sigma^2_{\eta}(N))N_i]/N$. The term $(N\sigma^2_{\eta}(N))$ would be the total noise in the news if news-articles were independent of each other. The term $1 + (n-1)\rho_N$ corrects the total noise for correlation among news-articles. I will call the entire term $(N\sigma^2_{\eta}(N)) \left[1 + (n-1)\rho_N\right]$ *total noise* in the news.

By analyzing $N_i$ news-articles, an investor successfully eliminates $S(N_i, N)$ fraction of total noise. If $\delta = 0$, the success rate is zero for all $N_i$. If $\delta = \sigma^2_{\eta}$, then $S(N, N) = 1$: an investor who analyzes all news-articles is able to eliminate all the noise in the news. It is easy to show that the precision and success rate are positively related. In the Online Appendix, I formally show that for each $(\delta, \sigma^2_{\eta}, \rho_N)$ in my setup, there exists an equivalent economy for some $(\tilde{\delta}, \tilde{\sigma}^2_{\eta}, \tilde{\rho}_N)$ in which only a certain fraction of $N$ news-articles contain new information to an investor. Hence, success-rate interpretation is without loss of generality in terms of qualitativeness of the results.

This signal structure also implies an investor might choose to subscribe to a news outlet even if she does not analyze any news. That is, $f(0, N) > 0$. This lower bound of precision could be interpreted as the information extracted from top stories, quick facts or summarized data that are provided by news outlets.
2.4 Trading Environment

The trading environment is a canonical quadratic-Normal microstructure framework à la Grossman and Stiglitz (1980). There is a continuum of investors, indexed by \([0, 1 + \mu]\). As defined in the previous sub-section, \(\mu\) fraction of investors are low-types: they have lower productivity in information processing compared to high-types. Investors have constant absolute risk aversion preferences over final wealth \(w_{2,i}\),

\[ U(w_{2,i}) = -e^{-\gamma w_{2,i}} \]  

where \(\gamma\) is a positive scalar measuring risk aversion. Investor \(i\) initially has \(c_i\) dollars of resources and \(x_i\) units of the stock. Let \(x_i\) denote her demand for the stock. The final wealth is then:

\[ w_{2,i} = c_i + \tilde{v}(x_i + \bar{x}_i) - px_i - \mathbb{1}_{ims}c_{ims} \]  

where \(\tilde{v}\) is the stock payoff at the final date \((t = 2)\), and \(p\) is the price of the stock at \(t = 1\). The stock payoff \(\tilde{v}\) is uncertain as of time 1. It is normally distributed with a mean of 0 and a precision (reciprocal of variance) of \(\tau_v\), that is, \(\tilde{v} \sim \mathcal{N}(0, \tau_v^{-1})\), with \(\tau_v > 0\). \(\mathbb{1}_{ims}\) is an indicator function equal to 1 if investor \(i\) decides to subscribe outlet \(m\) in state \(s\) and \(c_{ims}\) is total cost of information acquisition defined in equation 2. There is also a third set of investors in the economy, namely noise traders. They submit an exogenous demand \(\bar{x}_{\text{noise}} \sim \mathcal{N}(0, \tau_x^{-1})\) at \(t = 1\), which prevents prices from being fully revealed.

2.5 Equilibrium

The equilibrium requires that (i) news producers optimally chooses information prices subject to participation constraint, (ii) informed and uninformed investors optimally choose demands for stock conditional on their respective information sets including stock price, (iii) investors optimally choose whether to subscribe or not, and if subscribed number of news-articles to analyze subject to capacity constraints (iv) information and stock markets clear, (v) producers and investors have rational expectations.

Formally, an equilibrium consists of information prices \(\{p_m\}_{m=1}^M\), stock price \(p\), demands for stock \(\{x_i\}_{i\in[0,1+\mu]}\), demands for information \(\{\mathbb{1}_{ims}\}_{m=1}^M, N_i\}_{i\in[0,1+\mu]}\) such that
1. given other information producers’ prices $p_m, p_{-m}$ solves producer $m$’s time−0 problem

$$\pi_m^* := \max_{p_m} \mathbb{E}_{t=0} \left[ (p_m - c)Q_m (p_m, p_{-m}) - R \right]$$

s.to $\pi_m^* \geq 0$,

2. given $p, x_i$ solves investor $i$’s time−1 problem conditional on her information set $I_i$

$$U_{1,i}^* := \max_{x_i} \mathbb{E}_{t=1} \left[ -\exp (\bar{c}_i + \bar{v}(x_i + \bar{x}_i) - px_i - 1_{ims}c_{ims}) \mid I_i \right]$$

where $I_i = \{ \bar{s}_i, p \}$ if informed and $I_i = \{ p \}$ if uninformed,

3. given $\{ p_m \}_{m=1}^M, (\{ 1_{ims} \}_{m=1}^M, N_i)$ solves investor $i$’s time−0 problem

$$U_{0,i}^* := \max_{\{ 1_{ims} \}_{m=1}^M, N_i} \mathbb{E}_{t=0} \left[ U_{1,i}^* \right]$$

s.to $c_{ims} \leq \bar{c}_i$,

4. information and stock markets clear in each state

$$\int_0^{1+\mu} 1_{ims} (p_m, p_{-m}) \, di = Q_m (p_m, p_{-m})$$

$$\int_0^{1+\mu} x_i \, di = \int_0^{1+\mu} \bar{x}_i \, di.$$  

2.6 Media Coverage

I define media coverage as the total number of news-articles per producer, $N(R - \bar{R}, \theta_s)$. Hence, an increase in media coverage stems from either an exogenous increase in producers’ resources or improved technology. I will focus on the latter. An improvement in technology makes information processing cheaper for investors. Keeping all else constant, they analyze more news-articles ($N_i$), face higher success rates ($S(N_i, N_i)$), and thus observe more accurate signals ($f(N_i, N)$). However, it also benefits producers. They are able to report more news-articles using the same amount of resources. If an increase in $N$ reduces the success rate, then an investor can observe less accurate signals in a state of the world with better technology. I will now present a parametric example to illustrate these trade-offs. I will focus on cases where an increase in media coverage does not add additional uncertainty to the economy, i.e., $\partial [N \sigma_{\eta}^2(N)] / \partial N \leq 0$. 

16
2.7 A Parametric Example

In this section, I will give a parametric example to demonstrate how higher media coverage affects investors. Suppose news-articles are independent of each other, i.e. $\rho_N = 0$. In this case, each news-article constitutes a partition. Let $\sigma^2_{\eta} = \bar{\sigma}_\eta^2 / N$ and $\delta = \bar{\delta} / N^\alpha$ where $\alpha \geq 0$ and $\bar{\sigma}_\eta^2$ is sufficiently high to ensure that $\delta \leq \sigma^2_{\eta}$. The total noise $N\sigma^2_{\eta}(N)$ is constant. The success rate is

$$S(N_i, N) = \frac{\bar{\delta} \frac{N_i}{N}}{\sigma^2_{\eta} N^\alpha}$$

The success rate is proportional to $N_i / N^\alpha$. $\alpha$ measures the degree to which information is dispersed across news-articles within a news outlet. Higher dispersion corresponds to lower success rates. Suppose $N > 1$. When $\alpha$ is equal to 0, information dispersion is minimized across news, i.e., the success rate is maximized for a given $N_i$. As $\alpha$ goes to infinity, information dispersion is maximized across news, i.e., the success rate is minimized for a given $N_i$. Indeed, the success rate is zero in this case. Analyzing news does not provide any new information to investors.

By processing $N_i$ news-articles, an investor observes a private signal with precision

$$f (N_i, N) = \frac{1}{[1 - S(N_i, N)]\bar{\sigma}^2_{\eta} + \sigma^2_{\bar{\eta}}}.$$  \hspace{1cm} (16)

$\alpha$ also measures total information in the media. To see this relation, define $f (N) := f (N, N)$. It is the precision of an investor that processes all news-articles. It is proportional to $N^{1-\alpha}$. If $\alpha = 1$, $\partial f (N) / \partial N = 0$: an increase in media coverage does not change total information in the media. If $\alpha < 1$, then $\partial f (N) / \partial N > 0$: higher media coverage increases total information in the media. Similarly, $\alpha > 1$ implies $\partial f (N) / \partial N < 0$: total information in the media reduces with media coverage.

Suppose that $C (N_i, \theta_{is}) = (N_i)^{1/(\hat{\theta}_{is})}$ and $N (R - \bar{R}, \theta_s) = (R - \bar{R})^{\hat{\theta}_s}$. If an investor’s resource constraint is binding, the best signal that she can observe has a precision of $f (\bar{N}_i, N)$ where

$$\bar{N}_i = (\bar{c}_i - p_m + b_{im})^{\hat{\theta}_s}.$$  \hspace{1cm} (17)

Now suppose that there is a positive technology shock. Then, $\partial f (\bar{N}_i, N) / \partial s > 0$ if and
only if

\[
\begin{align*}
\alpha \left( \frac{\bar{\theta}}{\theta_i} \right) &< \frac{\log (\bar{c}_i - p_m + b_{im})}{\log (R - \bar{R})} + \frac{-\partial p_m/\partial s}{\log (R - \bar{R}) \left[ (\bar{c}_i - p_m + b_{im}) / s \right]} \\
&\text{relative productivity} & \text{distr. of resources} & \text{crowding-in/out effect of tech.}
\end{align*}
\] (18)

The first term (excluding \(\alpha\)) measures productivity differences between producers and investors. A higher ratio implies that an improvement in technology benefits more to producers relative to investors. The second term measures distribution of resources in the economy. A lower ratio implies that the media is allocating more resources to report news-articles about a stock than the resources that an investor is allocating to analyze news-articles. The last term is the crowding-in (\(\partial p_m/\partial s < 0\)) or crowding-out (\(\partial p_m/\partial s > 0\)) effect of technology. A higher ratio indicates higher savings for investors due to improved technology.

If equation 18 does not hold, a resource-constrained investor gets less accurate information through the media after a positive technology shock. It is because she cannot sufficiently increase the number of news-articles she analyzes as media coverage increases. It implies that higher media coverage can make a resource-constrained investor worse off if she has low productivity relative to producers or resources are distributed more unequally among producers and investors. The last term in equation 18 is relatively small compared to the other two terms due to its large denominator.

2.8 Media Coverage and Learning Trade-offs

I will focus on cases where an increase in media coverage results from improvements in technology. In my parametric example, investors have lower information processing costs in a state of the world with higher media coverage. This feature is always held in my model since I assume that \(C'_{\theta_i} < 0\) and \((\theta_i)'_s > 0\). Lower processing costs allow investors to analyze more news-articles (\(\uparrow N_i\)) and thus to get more accurate underlying fundamental information about the stock (\(\uparrow f(N_i, N)\)).

However, an investor has more news-articles to process in a state of the world with higher media coverage. Consider the term \([1 - S(N_i, N)](N \sigma^2_\eta(N))\) in equation 8. The lower the term is, the more accurate the private signal is. For a given \(N_i\), consider the derivative of this term with respect to \(s\):

\[
\left( -\frac{\partial S(N_i, N)}{\partial N} \right) \left[ N \sigma^2_\eta(N) \right] + \left[ 1 - S(N_i, N) \right] \frac{\partial}{\partial N} \left[ N \sigma^2_\eta(N) \right] \frac{\partial N}{\partial s}
\] (19)
\[ \frac{\partial N}{\partial s} \] is strictly positive. I also focus on cases where increased media coverage does not increase overall uncertainty in the economy, i.e., \( \frac{\partial}{\partial N} \left[ N \sigma^2_N (N) \right] \leq 0 \). Hence, the second term in equation 19 is negative. However, increased media coverage can result in a lower success rate, i.e. \( \frac{\partial}{\partial N} S(N_i, N) < 0 \). In other words, information becomes more dispersed across news. If this derivative is sufficiently negative, the equation 19 becomes positive. It means that analyzing a fixed number of news-articles provides less accurate information when media coverage is higher. It is because identifying news-articles that contain new information relative to an investor’s information set becomes more difficult as media coverage increases.

In summary, increased media coverage (due to improved technology) causes a trade-off between lower information processing costs and higher information dispersion (equivalently, lower success rates) whenever equation 19 is positive.

### 2.9 Price Informativeness

Grossman and Stiglitz (1980) defines price informativeness as the inverse of conditional variance of fundamental given stock prices– \( \frac{1}{\text{Var}(\tilde{v} | p)} \). In equilibrium, all same-type informed investors have the same precision level, \( \tau^{l}_{\epsilon_i} = \tau^{l}_{\epsilon} \) and \( \tau^{h}_{\epsilon_i} = \tau^{h}_{\epsilon} \) where \( l (h) \) denotes investors with \( \theta_{ls} = \theta_{ls} (\theta_{hs} = \theta_{hs}) \). It is easy to show that price informativeness is positively related with both \( \tau^{l}_{\epsilon} \) and \( \tau^{h}_{\epsilon} \) as well as total fraction of informed investors.

I choose \( \theta_{hs} \) large enough that high-type investors never run out of budgets to analyze news. It ensures that they are able to get more accurate information through the media as media coverage increases. Thus, in an unconstrained world where each investor has sufficient budget to analyze news, higher media coverage (due to improved technology) monotonically increases price informativeness.

However, as I showed in my example, after a positive technology shock, some investors may not be able to sufficiently increase the number of news-articles they analyze due to resource constraints. In my setting, these investors are low-type investors. As media coverage increases, resource-constrained investors cannot reap all the benefits of lower information processing costs while incurring all the costs of higher information dispersion. Accordingly, they get less accurate information through the media as media coverage increases. It directly decreases price informativeness. Besides, because subscription to news is costly, acquiring less accurate information gives resource-constrained investors lower incentives to purchase news in the first place.
Hence, higher media coverage improves price informativeness through resource-abundant investors (high-type), and deteriorates it through resource-constrained investors (low-type). The net impact in equilibrium depends on the parameters of the model. It particularly depends on productivity differences between producers and investors, distribution of resources among producers and investors, and distribution of types.

2.10 Introducing Earnings Guidance to the Model

Earnings guidance is an additional source for investors to learn about the stock value. I assume zero purchasing and processing cost.\(^\text{10}\) It induces a signal about the stock:

\[
\tilde{s}_g = \tilde{v} + \tilde{\varepsilon}_g
\]

where \(\tilde{\varepsilon}_g \sim \mathcal{N}(0, \tau_g^{-1})\). Earnings guidance affects investors’ incentives to purchase and process news through two competing forces. First, unlike financial news, earnings guidance is available to investors free of charge. Hence, every investor uses it to update her belief about the stock value. This process directly injects more fundamental information into the stock price and hence improves price informativeness. Higher precision of earnings guidance (\(\tau_g\)) implies a stronger injection.

Second, in the presence of a free information source, investors have less incentives to spend resources on purchasing and analyzing costly news. This crowding-out effect deteriorates price informativeness. It is stronger when information prices and/or processing costs are higher. In equilibrium, which effect dominates depends on the parameter of the model. That is, like media coverage, earnings guidance has also an ambiguous effect on price informativeness.

2.11 Back to the Parametric Example

In this subsection, I will analyze the impacts of media coverage and earnings guidance on price informativeness using the example presented in 2.7. I choose the parameters to ensure that technological improvements sufficiently reduce information processing costs in order for higher media coverage (due to improved technology) to monotonically increase price informativeness in an unconstrained world. The Panel A of Figure 1 depicts this limiting case in

\(^{10}\)In practice, it is true that earnings guidance has zero purchasing cost. However, it is still costly to process it. I assume this cost is low relative to processing costs of high-scale news.
which almost no investor is constrained ($\mu \to 0$). As media coverage increases, each investor acquires better underlying fundamental information through the media and hence price informativeness monotonically increases.

The Panel B of Figure 1 provides another limiting case. Almost all investors are highly constrained: $\mu$ is very high and $\bar{c}_i$ is very low. In this case, investors’ resource constraints are binding even at very low levels of media coverage. As media coverage increases, they get less and less accurate information. This process also lowers their incentives to purchase news in the first place. These two dynamics result in a monotonic decrease in price informativeness.

In the Panel C of Figure 1, I present an intermediate case. I choose ($\mu = 2$) and a moderate level of $\bar{c}_i$. In this case, the impact of media coverage on price informativeness is an inverted U-shape. It is almost an unconstrained world for investors at low levels of media coverage. All investors are able to process as much news as they want without having any difficulty in finding resources. In equilibrium, some investors with low productivity would prefer not to subscribe to news. However, it is mainly because stock prices are informative enough that acquiring new information is not worth much for those investors. As media coverage keeps increasing due to improved technology, low-type investors’ information processing capabilities start to be constrained by their limited resources. Stock prices incorporate less information thereafter.
Figure 1: Effects of Media Coverage and Earnings Guidance on Price Informativeness. The y-axis is price informativeness. It is defined as conditional variance of fundamental given stock price \( \text{Var}(\tilde{v} | p) \). The x-axis is the state of technology. From left to right, \( s \) increases and hence media coverage increases. The common parameter values are: \( \tau_x = \tau_v = 1, \bar{x}_i = 1, \gamma = 0.5, \theta = 0.5 \).

Figure 1 also demonstrates that the impact of earnings guidance on price informativeness depends on media coverage. In particular, it depends on the degree to which higher media coverage limits low-type investors’ information processing capabilities. When this degree is high, they have lower incentives to acquire information through the media. In this case, media coverage itself has a large crowding out effect on their information acquisition. Issuing earnings guidance then does not cause a significant change in low-type investors’ information acquisition decisions. It still distorts high-type investors’ incentives and induces them to analyze less news. However, if the fraction of low-type investors is sufficiently high, earnings guidance improves price informativeness by providing better fundamental information to every investor. In Panel B, that degree is very high from the very beginning. In Panel C, it becomes high after media coverage goes beyond a point.

The reason why earnings guidance increases price informativeness at high media coverage in Panel A is different. No investor is resource-constrained in the economy in Panel A. In my model, there are complementarities in information demand. That is, increased demand for information causes information to be supplied to more investors at a lower price. This feature is a result of competitive information market, free entry and non-rivalry characteristic...
of information.\footnote{See Appendix A.1 in Veldkamp (2006b) for details.}

\section{Mapping the Model to Observables}

The model produces ambiguous effects of media coverage and earnings guidance on price informativeness. In the model, I defined price informativeness as the conditional variance of fundamental given stock price. In practice, it is hard to measure fundamentals. There are two approaches to measure price informativeness. The first approach measures price informativeness as forecasting power of prices for future cash flows. A cash flow measure such as EBIT/EBITDA is regressed on past stock prices plus controls. For instance, Field and Lowry (2009), Turley (2012), Li, Richardson, and Tuna (2014), Bai et al. (2016), Dávila and Parlatore (2018), Farboodi et al. (2020), Martineau (2018).

There is a more novel approach to measure informational content of stock prices. It focuses on price jump ratios. It divides stock price movements into two periods: movements on an information disclosure date and movements during a pre-disclosure period. If stock prices are not informative in pre-disclosure periods, prices jump more on disclosure dates. That is, higher jumps on disclosure dates indicate that information is not incorporated into prices until publicly revealed. Although this idea dates back to Morse [1981] and Meulbroek [1992], its usage is pretty new: see Weller (2018), Sammon (2021). I will focus on price jump ratios, proposed by Sammon (2021), where disclosures represent earnings announcements and apply the first approach in the Online Appendix to check robustness of my findings.

In the model, $t = 1$ represents pre-earnings announcement date, and $t = 2$ represents earnings announcement date. Time-2 stock price is equal to the fundamental value of the stock. As time-1 stock price gets more informative, it becomes closer to the fundamental value. Hence, it results in a smaller price jump from $t = 1$ to $t = 2$. Smaller jumps are associated with higher price informativeness. I will focus on two jump ratios: pre-earnings drift and pre-earnings volatility share.
3.1 Pre-Earnings Drift

Model. Pre-earnings drift measures jumps in stock returns. It is defined as

\[
\text{Drift} = \begin{cases} 
\frac{1+r_{(0,1)}}{1+r_{(0,2)}} & \text{if } r_2 > 0 \\
\frac{1+r_{(0,2)}}{1+r_{(0,1)}} & \text{if } r_2 < 0.
\end{cases}
\] (20)

\(r_{(t-n,t)}\) is cumulative stock return from \(t-n\) to \(t\), defined as \(r_{(t-n,t)} := \sum_{\tau=t-n}^{t} r_{\tau}\), where \(r_t := \frac{p_t-p_{t-1}}{p_t}\).\[^{12}\] Suppose \(r_2\) is positive. Higher \(r_2\) requires larger difference between time-1 and time-2 stock prices, i.e., higher stock price jumps. As \(r_2\) increases, the drift approaches zero. Hence, lower pre-earnings drift is associated with lower price informativeness. If \(r_2\) is negative, the measure needs to be inverted to measure this relationship correctly. In the model, time-1 stock price is a function of informed trading, informational accuracy of financial news and earnings guidance. Higher informed trading or informational accuracy, ceteris paribus, results in higher drifts. Similarly, all else equal, more accurate earnings guidance increases drifts.

Empirical Analogue. In the model, there is only one trading date. However, in practice, there are many trading dates between earnings announcements. Investors might prefer to spread out their trades over days or weeks before earnings announcement dates. I will take 22 trading days before an earnings announcement as my pre-announcement period. Empirical pre-earnings drift for stock \(j\) is then defined as the cumulative return from \(t-22\) to \(t-1\), divided by the cumulative return from \(t-22\) to \(t+a\), where \(t\) is earnings announcement date and \(a\) is a non-negative integer:

\[
\text{Drift}_{jt} = \begin{cases} 
\frac{1+r_{(t-22,t-1)}}{1+r_{(t-22,t+a)}} & \text{if } r_t > 0 \\
\frac{1+r_{(t-22,t+a)}}{1+r_{(t-22,t-1)}} & \text{if } r_t < 0.
\end{cases}
\] (21)

Suppose \(r_t > 0\). If more new information is disclosed on an earnings date, returns on that date and the following few days become larger, that is, higher \(r_t, r_{t+1}, \ldots, r_{t+a}\). It implies a higher \(r_{(t,t+a)}\) and hence a lower drift. Sammon (2021) documents that average pre-earnings drift has decreased between 1990 and 2018.

\[^{12}\]I calculate time-0 stock price such that it clears the stock market when every investor would be happy to keep her endowment.
3.2 Pre-Earnings Volatility Share

Model. Pre-earnings volatility share measures jumps in stock price volatility. It is defined as

\[
\text{Volatility} = \frac{r_t^2}{r_t^2 + r_{t-1}^2}
\]

where \( r_t = \frac{p_t - p_{t-1}}{p_{t-1}} \). The price at \( t = 2 \) is equal to the true value of the fundamental \( \tilde{v} \). The return at \( t = 2, r_2 \), is large when \( p_1 \) and \( p_2 \) are far from each other. In this case, earnings date volatility is high, which implies a lower pre-earnings volatility share. Hence, lower pre-earnings volatility share is associated with lower price informativeness. In the model, higher informed trading or informational accuracy, \textit{ceteris paribus}, causes higher pre-earnings volatility share. Similarly, \textit{all else equal}, more accurate earnings guidance increases pre-earnings volatility share.

Empirical Analogue. If investors learn new information about a stock’s value during a pre-earnings period, they want to trade based on the information before it becomes public on the next earnings date. Similarly, if a firm issues a strong earnings guidance, investors tend to trade its stock before the firm announces the actual earnings figure. These trading activities increase stock return volatility during pre-earnings announcement periods and hence cause a higher pre-earnings volatility share. Empirical pre-earnings volatility share is defined as

\[
\text{Volatility}_{jt} = \frac{\sum_{\tau=-22}^{t-1} r_{i,t+\tau}^2 / \sum_{\tau=-22}^a r_{i,t+\tau}^2}{a}
\]  

(22)

where \( t \) is earnings announcement date and \( a \) is a non-negative integer. Sammon (2021) documents that average pre-earnings volatility share has decreased by 19.4% between 1990 and 2018.

3.3 Data and Descriptive Statistics

I use Ravenpack Analytics (RPA) to get data on financial news. RPA analyzes unstructured news from thousands of sources to extract information and determine financially-relevant events about stocks. They assign a relevance score to each company in an event. It measures how strongly a company relates to the underlying news story. It takes values between 0 and 100 with higher values indicating greater relevance. I include only significantly relevant news (relevance score \( \geq 90 \)) into the sample.
The data on RPA starts from January 2000. However, it includes only 4 sources until January 2007: Dow Jones Financial Wires, Wall Street Journal and its regional editions, Barron’s and MarketWatch. RPA launched Web Edition in January 2007 and started to process hundreds of thousands of articles a day from other leading publishers and web aggregators. I conduct the same analysis for both Jan 2000 to Dec 2019 and Jan 2007 to Dec 2019. The results are qualitatively the same. Hence, I present the results from the larger sample: Jan 2000 to Dec 2019.

Daily stock data are retrieved from CRSP. I restrict to ordinary common shares traded on major exchanges (share codes 10 and 11, exchange codes 1 to 3). Company fundamentals are from Compustat. I use both Compustat and I/B/E/S to construct effective earnings announcement dates using the methodology of DellaVigna and Pollet (2009). If earnings are released before 4:00 PM eastern time on a trading day, the effective earnings date is equal to actual announcement date. If earnings are released after 4:00 PM eastern time or on a non-trading day, the effective earnings date is set to the next trading day.

Data on earnings guidance are retrieved from I/B/E/S Guidance. Data on institutional ownership are from Thomson/Refinitiv 13-F. I use CRSP-Compustat Link and IBES-CRSP Link suites in WRDS to merge CRSP, Compustat, and I/B/E/S. I use CRSP’s CUSIP or historical CUSIP to merge 13-F and RPA with the rest of the data.

3.4 Media Coverage Measures

I construct three variables to measure media coverage. The first two measures are commonly used in the literature and constructed for comparison purposes. The last one is the new measure that I propose. I will discuss the problems associated with the existing measures and show how my new measure fixes those problems to a great extent. Each measure counts the number of financial news-articles that appear in an official news outlet or multiple news outlets subject to certain criteria. Social media or unofficial blogs such as Reddit are not included. In construction of each variable, I include only significantly relevant news about a stock. If a stock’s name appears in a news-article and it is one of the main protagonists, the article is considered significantly relevant for that stock.
3.4.1 Total Number of Major News-Articles

It is defined as the total number of news-articles reported by a single major news outlet. I will use three different outlets: Wall Street Journal (WSJ), New York Times (NYT), and Washington Post (WP).

\[ MNEW S_{i,t} := \sum_{\tau = t_0 + n_0}^{t-1} NEW S_{i,\tau}(O) \]

where \( t \) is current earnings announcement date, \( t_0 \) is previous earnings announcement date, \( n_0 \in \{0,1,\ldots,5\} \), and \( NEW S_{i,\tau}(O) \) is total number of news-articles in outlet \( O \) on date \( \tau \), where \( O \in \{WSJ, NYT, WP\} \). This metric is the most common variable to measure media coverage in the literature: see Davies and Canes (1978), Klibanoff et al. (1998), Nofsinger (2001), Antweiler and Frank (2004), Tetlock (2007), García (2013) for instance.

3.4.2 Total Number of News-Articles

It is defined as the total number of financial news-articles reported in a newswire or news aggregator. I will use Dow Jones Newswire (DJ) and Ravenpack Analytics (RPA) as examples of newswire and news aggregator, respectively.

\[ TNEW S_{i,t} := \sum_{\tau = t_0 + n_0}^{t-1} NEW S_{i,\tau}(O) \]

where \( t \) is current earnings announcement date, \( t_0 \) is previous earnings announcement date, \( n_0 \in \{0,1,\ldots,5\} \), and \( NEW S_{i,\tau}(O) \) is total number of news-articles in outlet \( O \) on date \( \tau \), where \( O \in \{DJ, RPA\} \). This metric has been getting more common in the literature: see Tetlock et al. (2008), Tetlock (2010), Bushee et al. (2010), Rogers et al. (2016), Bonsall et al. (2020) for instance.

3.4.3 Problems with Measuring Media Coverage

My goal is to measure the set of relevant news-articles for investors. The model implies that as this set gets larger, investors with limited resources may find it difficult to identify which news-articles contain new information relative to their information sets. In this respect, a good measure should have two fundamental features at minimum. First, it should take into account information dispersion across different news outlets. Different stocks can receive coverage from different outlets. Second, a good measure should take into account the degree to which a news-article can contain new information compared to the recently reported
news-articles and eliminate articles that are easy to compare and conclude that they are almost perfect duplicates of each other. These kinds of news-articles are generally triggered by an event and reported one after another in a very short period of time.

Consider the following example. On July 11, 2019, Ford Motor Co. announced a new partnership with Volkswagen. My data source, Ravenpack Analytics (RPA), captured 19 news-articles within the first 24 hours in news outlets that are regarded as trustworthy. RPA considers an outlet trustworthy if it is fully accountable, reputable, and impartial. Here is a breakdown of 19 articles:

<table>
<thead>
<tr>
<th></th>
<th>WSJ</th>
<th>NYT</th>
<th>WP</th>
<th>DJ</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>14</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 1: Distribution of News-articles after Ford Motor Co. announced a new partnership with Volkswagen on July 11, 2019.

The first media coverage measure (MNEWS) is 0 or 1. It suggests that focusing on a single news outlet can underestimate the set of relevant news-articles for investors, e.g. media coverage would be measured as 0 if an econometrician focused on WSJ or WP as a proxy for media coverage to which investors pay attention. The second media coverage measure (TNEWS) is 4 or 19. It is very unlikely that an investor pays attention to all these articles. She could easily notice that these articles were almost duplicates of each other since they appeared in the news with similar headlines within a short period of time. Hence, she could understand that analyzing one or two of them would be enough for her to understand the nature of the event and update her beliefs.

Here is another example. On March 28, 2019, Equifax Inc. announced a new partnership with FICO. RPA captured 8 news-articles within the first 24 hours of this event in news outlets that are regarded as trustworthy. Here is a breakdown of 8 articles:

<table>
<thead>
<tr>
<th></th>
<th>WSJ</th>
<th>NYT</th>
<th>WP</th>
<th>DJ</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 2: Distribution of News-articles after Equifax Inc. announced a new partnership with FICO on March 28, 2019.

Similar problems arise in the second example too. An econometrician using NYT or WP as a proxy for media coverage would not include the articles triggered by this event. However, it does not imply that an investor did not pay attention to these articles. It is also
unlikely that she paid attention to all these 8 articles.

These examples are not exceptions. In the cross-section of financial news data, different stocks receive coverage from different outlets. Hence, cross-outlet dispersion should be taken into account when constructing a measure for media coverage. However, a newswire or news aggregator tends to cover the same information multiple times within a short period of time, which could easily be detected by investors.¹³ I will now construct a new measure that takes into account both cross-outlet dispersion and duplication of news-articles.

### 3.4.4 Total Number of Novel News-Articles

My new measure has two dimensions. First, I do not fix a news outlet across stocks. I use different outlets for different stocks. It is important to capture cross-sectional dispersion. I choose one or two outlets for each stock as its universe of media coverage where those outlets provide almost all genuine news-articles about the stock. In general, many stocks receive considerably good coverage from either Dow Jones Newswire (DJ) or Bloomberg Terminal (BM). Yet, there are some stocks that are covered by specialized outlets. Second, I include only news-articles with event novelty scores (ENS) of 100. RPA assigns a news-article to ENS of 100 when there is no similar news in the past 24 hours. This aspect eliminates news-articles that are almost perfect duplicates of each other and investors can easily notice their duplicative features.

\[
N_{\text{NEW S}}^{i,t} := \sum_{\tau=t_0+n_0}^{t-1} \text{NEW S}_i,\tau \ (\text{ENS}=100)
\]

where \( t \) is current earnings announcement date, \( t_0 \) is previous earnings announcement date, \( n_0 \in \{0, 1, ..., 5\} \), and \( \text{NEW S}_i,\tau \ (\text{DJ, ENS = 100}) \) is total number of news in DJ with ENS of 100 on date \( \tau \).

⁠¹³It is worth mentioning that the existing measures are better at capturing media coverage in some settings. For instance, the first measure is a good one to measure the set of financial news for retail investors. WSJ, NYT or WP tend to cover stocks that are popular among retail investors. It is also very likely that one of these outlets is a good representation of how much news retail investors pay attention. Similarly, the second measure, especially counting the number of news-articles in a news aggregator, is better at capturing the dissemination role of the media. By reaching a broader population of investors, higher coverage can alleviate informational frictions and affect security pricing even if it does not supply genuine information.
4 Empirical Results

I divide the sample into two sub-samples based on how persistently a company issues earnings guidance. The first sub-sample is persistent earnings guidance issuers. If a company issues guidance at least \( n_1 \) quarters without any lag, where \( n_1 \in [12, 20] \), it falls into this category. If a company never issues earning guidance for at least \( n_2 \) consecutive quarters, where \( n_2 \geq 12 \), it falls into no earnings guidance issuers category. If a company does not fall into either of these two categories, it is labeled occasional earnings guidance issuers. I will not focus on occasional issuers in this paper. For each sub-sample, I regress price informativeness measures on media coverage measures and control variables:

\[
\text{PriceInfo}_{i,t} = \alpha_i + \gamma_t + \beta_1 \text{MEDIA}_{i,t} + \beta_2 \text{MEDIA}_{i,t}^2 + \text{Controls} + \varepsilon_{i,t} \tag{23}
\]

where PriceInfo is pre-earnings drift (Drift\(_{it}\)) or pre-earnings volatility (Volatility\(_{it}\)), MEDIA is total number of news (TNEWS\(_{it}\)) or major number of news (MNEWS\(_{it}\)) or novel number of news (NNEWS\(_{it}\)). The controls include standard control variables in the literature: firm age, lagged market capitalization, returns from \( t - 12 \) to \( t - 2 \), lagged book-to-market ratio, total and change in institutional ownership, total volatility, idiosyncratic volatility, guidance accuracy (MAD or MADP) and other guidance characteristics.

4.1 Pre-earnings Drift

I run equation 23 for each sub-sample and each media coverage measure using pre-earnings drift as my price informativeness measure. The regression results are in Table 3. All coefficients of media are statistically significant. MEDIA has a positive coefficient, but its square has a negative coefficient. It implies that media has a hump-shaped impact on price informativeness. That is, price informativeness first increases with media coverage, but then it starts to decline as media coverage rises more. Consider the marginal effect of media on price informativeness: \( \partial(\text{PriceInfo})/\partial(\text{MEDIA}) = \beta_1 + 2\beta_2 \text{MEDIA} \).

Figure 4 in the Appendix plots this function. It indicates that excessive media coverage adversely affects pre-earnings drift. I will first analyze the sample with firms that do not issue guidance. Average total number of news increases by 260 from 2000 to 2019 in my sample. It causes around 0.015 decrease in pre-earnings drift, which is not a big impact in economic terms. Average decline in pre-earnings drift is 0.0161 in my sample. Hence, increasing trends in total number of news could explain only 4.7% of the change in pre-earnings drift over the past twenty years.
Average major number of news increases by 150 from 2000 to 2019 in my sample. It causes a decrease of 0.072 in pre-earnings drift, which is a big impact in economic terms. Temporal trends in major number of news-releases could explain 22.6% of the total change in pre-earnings drift in my sample. Average novel number of news rises by 80 from 2000 to 2019 in my sample. It decreases pre-earnings drift by around 0.040, which corresponds to 12.6% of the total temporal change in pre-earnings drift.

Averaging all three measures’ impacts on pre-earnings drifts, trends in financial news could explain around 13.3% of the decrease in average pre-earnings drift over the last twenty years. For firms that issue earnings guidance, that number is around 12.37%. It suggests that firms issuing earning guidance are slightly less responsive to media coverage, but it does not cause a major change in pre-earnings drifts’ responses to media coverage. Table 3 and Figure 4 also indicate that earnings guidance lowers pre-earnings drift at high media coverage.

Averaging all three measures’ impacts on pre-earnings drifts, trends in financial news could explain around 13.3% of the decrease in average pre-earnings drift over the last twenty years. For firms that issue earnings guidance, that number is around 12.37%. It suggests that firms issuing earning guidance are slightly less responsive to media coverage, but it does not cause a major change in pre-earnings drifts’ responses to media coverage. Table 3 and Figure 4 also indicate that earnings guidance lowers pre-earnings drift at high media coverage.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDIA</td>
<td>0.0163***</td>
<td>0.0136***</td>
<td>0.0413***</td>
<td>0.0411***</td>
<td>0.0529***</td>
<td>0.0432***</td>
</tr>
<tr>
<td></td>
<td>(0.000756)</td>
<td>(0.00187)</td>
<td>(0.00204)</td>
<td>(0.00549)</td>
<td>(0.00195)</td>
<td>(0.00495)</td>
</tr>
<tr>
<td>MEDIASQ</td>
<td>-0.00289***</td>
<td>-0.00244***</td>
<td>-0.0244***</td>
<td>-0.0234***</td>
<td>-0.0261***</td>
<td>-0.0190***</td>
</tr>
<tr>
<td></td>
<td>(0.000151)</td>
<td>(0.000337)</td>
<td>(0.00154)</td>
<td>(0.00401)</td>
<td>(0.00114)</td>
<td>(0.00271)</td>
</tr>
<tr>
<td>MADP</td>
<td>-2.39e-08</td>
<td>-1.81e-08</td>
<td>-2.84e-08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.74e-08)</td>
<td>(4.73e-08)</td>
<td>(4.74e-08)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>216502</td>
<td>27320</td>
<td>216502</td>
<td>27320</td>
<td>216502</td>
<td>27320</td>
</tr>
</tbody>
</table>

Table 3: Pre-earnings Drift and Media Coverage

* p < 0.10, ** p < 0.10, *** p < 0.10. Dependent variable is (pre-earnings drift)$_t$, defined as $\frac{1+\rho_i(t-\tau-t_{-1})}{1+\rho_i(t-\tau-t_{+\alpha})}$ where $t$ is earnings date, and $\alpha = 2$. (N) denotes the sample that includes firms issuing no guidance, (G) denotes sample that includes firms persistently issuing guidance. MEDIA measures number of (1-2): total news, (3-4): major news, (5-6): novel news. MEDIA is in 100s. MEDIASQ is the square of MEDIA. MADP is mean absolute percentage deviation of earnings guidance. Standard errors are clustered at the firm level.

4.2 Pre-earnings Volatility Share

I use pre-earnings volatility share as my price informativeness measure and run regression 23 for each sub-sample and each media measure. The regression results are in Table 4. All coefficients of media are statistically significant. As in the regressions of pre-earnings drift, media has a non-linear impact on price informativeness. That is, pre-earnings volatility share first increases with media coverage, but then it starts to decline as media coverage rises.
more. Figure 5 in the Appendix plots marginal effects of media on pre-earnings volatility share \((\beta_1 + 2\beta_2 MEDIA)\). Although media affects pre-earnings drift and volatility share in a similar way, the effects of media are more pronounced for the latter in economic terms.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility(N)</td>
<td>0.0423***</td>
<td>0.0123</td>
<td>0.0449***</td>
<td>0.0175</td>
<td>0.0784***</td>
<td>0.0308**</td>
</tr>
<tr>
<td>Volatility(G)</td>
<td>0.00989</td>
<td>0.00497</td>
<td>0.0147</td>
<td>0.0050</td>
<td>0.0127</td>
<td></td>
</tr>
<tr>
<td>MEDIA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEDIA</td>
<td>-0.00889***</td>
<td>-0.00578***</td>
<td>-0.0170***</td>
<td>-0.0105**</td>
<td>-0.0210***</td>
<td>-0.0109***</td>
</tr>
<tr>
<td>MEDIASQ</td>
<td>0.00799</td>
<td>0.00187</td>
<td>0.00183</td>
<td>0.00514</td>
<td>0.00147</td>
<td>0.00375</td>
</tr>
<tr>
<td>MADP</td>
<td>0.000000219</td>
<td>0.000000201</td>
<td>0.000000231</td>
<td>0.000000200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>216502</td>
<td>27320</td>
<td>216502</td>
<td>27320</td>
<td>216502</td>
<td>27320</td>
</tr>
</tbody>
</table>

Table 4: Pre-earnings Volatility Share and Media Coverage

For firms that do not issue guidance, average decline in pre-earnings volatility share is 0.8% in my sample. An increase in total number of news by 260 decreases pre-earnings volatility share by around 0.046. It corresponds to 30.3% of the decline in pre-earnings volatility share over the past twenty years. An increase in major number of news-releases by 150 causes a decrease of 0.0507 in pre-earnings volatility share, corresponding to 33.4% of the total temporal change in pre-earnings volatility share. An increase in novel number of news by 80 causes a decrease of 0.0327 in pre-earnings volatility share, which corresponds to 21.54% of the total temporal change in pre-earnings volatility share.

Averaging all three measures’ impacts on pre-earnings volatility share, trends in financial news could explain around 28.4% of the decrease in pre-earnings volatility share over the last twenty years. For firms that issue earnings guidance, that number is around 14.9%. It is in spite of the fact that average decline in pre-earnings volatility share for those firms is 1.28% in my sample. That is, relative to firms without earnings guidance, firms that persistently issue earnings guidance have experienced a larger decrease in pre-earnings volatility share over the past twenty years. However, pre-earnings volatility share of the latter is less sensitive to media coverage. Table 3 and Figure 4 also indicates that earnings guidance lowers pre-earnings drift at high media coverage.
4.3 Mapping Results Back to Model

The empirical results are consistent with Panel C in Figure 1. That is, investors’ information environment is moderately constrained. The economy is almost an unconstrained world for investors at low levels of media coverage and starts to become a constrained one as media coverage rises. The cross sectional regressions suggest that these dynamics might be happening in practice as well. The model provides two reasons why an investor gets resource-constrained as media coverage increases. First, information producers have higher productivity relative to her. This case is probably true in practice as well. A large news producer such as Bloomberg or Dow Jones is very likely to benefit more from technological improvements compared to small and mid-size investors. They can afford to purchase faster information technologies and hire more qualified people who can make better use of them. This situation may not be true for a small or mid-size investor. It is unlikely that she has enough resources to purchase the same technologies as large news producers.

Second, in the model, if producers allocate more resources to make news than the resources an investor allocates to process news, a resource-constrained investor gets less accurate information through the media after a positive technology shock. In practice, a large news producer can choose to allocate more resources to make news about a stock than a small or mid-size investor allocates to process news about that stock. It is because news is purchased by large institutional investors who can process large amount of information. It can be more important for a news provider to meet large investors’ information needs.

As in Panel C in Figure 1, the cross sectional regressions also suggest that earnings guidance improves price informativeness only at high media coverage. In the model, it is because high media coverage crowds out information acquisition in a constrained world. It might also be true in practice. A small or mid-size investor may not be able to have sufficient technology or expertise to extract information from news when media coverage is excessive. She could prefer to pay attention to only top news. Earnings guidance does not divert her attention much from top news for two reasons. First, amount of top news is small compared to total news. Thus, processing cost is not very high. Second, if she did not follow top news, she could miss a large price change, which would be very costly for her. However, when media coverage is moderate, a small or mid-size investor would prefer to process more news beyond top news. In this case, earnings guidance could make her stop processing additional news since earnings guidance is free and provides underlying fundamental information too.
5 Technological Evolution in Financial Markets

Improvements in information technology have revolutionized the way financial markets operate. Increased computerization over time has benefited both information producers and investors. Many investors now have improved access to information about securities. Machine learning (ML) and other applications of artificial intelligence (AI) allow traders to identify complex trading patterns, build new algorithmic trading systems and test them in real-time. Speech recognition and natural language processing tools save traders time since they do not need to go over every single note or conversation. Due to these techniques among many others, traders are now able to analyze large-scale information at a lower cost.

Financial news producers have also benefited from increased computerization. Some corporations disseminate press releases using newswires. Even if a corporation posts a press release in its own website, a newswire is easily able to retrieve it using automated computer programs. It makes information delivery faster, less expensive, and more widespread. Other examples are auto-journalism, BM Cyborg, AP Wordsmith.

All these improvements have increased total productivity of investors and information producers over time. In the example in 2.7, their productivity has two components: state of technology ($s$) and agent-specific productivity ($\tilde{\theta}_{is}$ for investors and $\tilde{\theta}$ for producers). Consider Panel C in Figure 1. An increase in technology causes a movement along a price informativeness curve. However, a change in agent-specific productivity parameter causes a shift of the curve.

If average investor-specific productivity raises, investors can process more news at a given technology and hence the curve shifts to right. The threshold after which media coverage reduces price informativeness increases. An increase in producer-specific productivity allows news producers to make more news using the same amount of resources at a given technology. The immediate effect is to have higher media coverage. Another effect is that it causes more investors to become resource-constrained and hence the curve shifts to left. The net impact on the threshold depends on the parameters.

Unlike my model in which there is only one stock, in practice, investors trade many stocks that producers report news about. The threshold in the model can be interpreted as the average threshold across many different stocks, and media coverage can represent average coverage across many stocks. In a multi-stock environment, if an investor has a higher
productivity ($\theta_{ts}$), she can choose to process more information about one particular stock, 
choose to include more stocks into her portfolio and use increased productivity to process 
information about new stocks, or choose to allocate equally among stocks that she is cur-
rently trading. Similarly, if a producer has a higher productivity ($\theta$), she has several options 
to use her new capacity. Hence, in this environment, in addition to the level of threshold, 
two variables are of interest: the fraction of stocks with media coverage above the threshold 
and the ratio of threshold to average media coverage.

Consider now a revised version of the main regression in 23:

$$\text{PriceInfo}^T_{i,t} = \alpha_i + \gamma_t + \beta_1 \text{MEDIA}_{i,t} + \beta_2 \text{MEDIA}_{i,t}^2 + \text{Controls} + \varepsilon_{i,t}$$  \hspace{1cm} (24)$$

where $T$ denotes the sample with years $\leq T$. I run this regression for $T = 2002, ..., 2019$ 
and for each sub-sample. Define optimal empirical threshold as media coverage level after 
which higher media coverage reduces price informativeness, i.e., $-(1/2)\beta_1/\beta_2$. For each T, I 
calculate thresholds and fraction of firm-quarters with media coverage above the thresholds 
to measure negatively-affected stocks.

Figure 6 shows optimal empirical thresholds for pre-earnings drift have increased over 
time. It suggests that on average investors are able to process more information over time. 
However, Figure 7 shows that the fraction of negatively affected firms increased over time too. 
It provides suggestive evidence that producers benefit more from technological improvements, 
but they mostly use improved capacity to cover more stocks or disproportionately increase 
coverage of certain stocks. The results are similar for pre-earnings volatility share. Besides, 
the main results of cross-sectional regressions continue to hold. In all regressions, media 
coverage has a hump-shaped impact on price informativeness and earnings guidance improves 
price informativeness only at high media coverage. Another way to check temporal dynamics 
of media coverage is adding interaction of year-dummies with MEDIA to regression 23. The 
results qualitatively remain the same.

6 Robustness Checks

Financial news contains information about events that affect a company’s fundamentals. 
Firms experiencing more fundamental events are more likely to attract more media attention. 
Some investors may not have the expertise to process high number events. They may
not know how to disentangle and disperse information from different events if the number of events is very high. In this case, less informed trading happens and stock prices contain less fundamental information (i.e., lower price informativeness). However, it would have nothing to do with media coverage itself. It would be about fundamental changes.

To take into account this endogeneity problem, I include variables that are proxy for fundamental events. I use three variables that are highly correlated with each other. The first one is the number of Key Developments of Capital IQ. Key Developments provide summaries of material events that may affect the value of a stock. The second variable is the number of press releases. Ravenpack Analytics provides detailed data on press releases. My last variable is the number of 8-K filings (a.k.a. current reports). In addition to Form 10-K and Form 10-Q, publicly traded firms have to file with the SEC to disclose certain material corporate events on a more current basis.

In my regressions, I include only one of the three variables because an important event generally triggers an 8-K filing accompanied by a press release and this event is captured by Capital IQ as a key development. Capital IQ also divides key developments into categories such as executive changes, acquisitions. Firms with more variety of developments may attract more attention and some investors may not be able to extract information from a large variety of events. I am adding frequency of different Key Developments to control this variation across stocks. The results are qualitatively similar after controlling number and variety of fundamental events.

Two additional threats to identification are increased algorithmic trading (Weller (2018)) and passive ownership (Sammon (2021)) over time. Weller (2018) show that algorithmic trading could reduce price informativeness through lowering investors’ incentives to acquire information. Sammon (2021) shows that passive ownership changes investors’ incentives to acquire information about stock-specific risks and systemic risk and finds that stocks with high passive ownership have smaller pre-earnings drifts and volatility share. I include algorithmic trading measures used in Weller (2018) and construct my passive ownership measure using the methodology of Sammon (2021).

Another concern for identification is that media coverage may be correlated with analyst coverage. Analyst and media coverage are negatively correlated. Hence, price informativeness might decrease due to lower coverage by analysts. Finally, I do the same analysis for only S&P 500 firms. Media coverage of these firms are systematically higher than non-index
firms. The results qualitatively remain similar after controlling these factors.

7 Conclusion

This paper examines how the contrasting trends in media coverage and earnings guidance have affected stock price informativeness over the past two decades. I first build a model to understand how and why media coverage and earnings guidance affect informational content of stock prices. I then use the model to construct a new empirical measure of media coverage, which takes into account cross-outlet coverage dispersion as well as the heterogeneity in the distribution of replicas of news across stocks. Finally, I empirically test the main predictions of the model.

In both cross-sectional and recursive regressions over different time periods, I show that, contrary to the common belief, high media coverage can lower price informativeness. In particular, media coverage has a hump-shaped impact on price informativeness. I also find that the impact of earnings guidance on price informativeness depends on media coverage. Earnings guidance improves price informativeness only at high levels of media coverage. The results provide suggestive evidence that increased media coverage and removal of earnings guidance can contribute to lower price informativeness over the past two decades.

The results are robust to using different measures of media coverage that are commonly used in the literature. However, both cross-sectional and recursive regressions suggest that my new empirical measure of media coverage is better at predicting the impact of financial news on stock price informativeness. Unlike the existing measures, my new measure takes into account coverage dispersion across different news outlets as well as the heterogeneity in the number of replicas across stocks.

The model helps us to understand the mechanism. In the model, investors have to spend resources on identifying to what extent one news-release contains new information relative to another news-release. In an unconstrained world in which every investor has enough resources to analyze every piece of news, higher media coverage always provides (weakly) better underlying information to investors and hence it improves price informativeness. However, in a constrained world where investors have limited and heterogeneous budgets to purchase and analyze news, as in my framework, this argument is not necessarily true. If higher media coverage made the identification of replicas more difficult and financially cumbersome, it could discourage investors with small budgets from paying attention to financial news. In
such an environment, earnings guidance helps them get better underlying information and hence it improves price informativeness.

In an article in the TechBulluion, it reads that “at the touch of a button, anyone can get access to information...relevant news about companies issuing stocks...the stock market is now full of better-informed traders”. This view is not particular to this article. Many other articles citing benefits of improved information technology for traders are providing similar views. Those views are partially correct. It is much easier now to access information compared to 20 years ago. However, my paper emphasizes that accessing more information alone does not necessarily make a trader better informed. She has to process the materials to extract information. Media coverage is a case in point. Subscribing to a financial news outlet is not very costly for an investor. However, extraction of information from news is not an easy process. It requires resources and expertise, and investors have different budgets and abilities to extract information. Thus, it is not necessarily true that easier access to massive media coverage makes investors better informed.

The results also shed light on a recent debate that calls for ending quarterly earnings guidance. Although the main focus of this debate is centered around earnings guidance’s effects on executives’ short-term focuses, its impact on investors should also be taken into account. In a 2006 study, McKinsey & Co. found that quarterly guidance does not reduce stock price volatility. My paper implies that this conclusion may not be immediate anymore, especially for companies with high media coverage.
References


Figure 2: **Trends in Financial News.** Financial News includes written or online number of unique news with contents appeared in official news providers. Social media or unofficial blogs such as Reddit are not included.
Figure 3: Trends in Earnings Guidance.
Figure 4: **Marginal Effects of Media Coverage on Price Informativeness.** Y-axis is marginal effect of media measure on pre-earnings drift: \( \frac{\partial (\text{PriceInfo})}{\partial (\text{MEDIA})} = \beta_1 + 2\beta_2 \text{MEDIA}. \) X-axis measures media (in 100s): total news, major news, novel news from left to right.
Figure 5: **Marginal Effects of Media Coverage on Price Informativeness.** Y-axis is marginal effect of media measure on pre-earnings drift: $\partial(\text{PriceInfo})/\partial(\text{MEDIA}) = \beta_1 + 2\beta_2\text{MEDIA}$. X-axis measures media (in 100s): total news, major news, novel news from left to right.
Figure 6: **Empirical Thresholds for Pre-earnings Drift over Time.** Y-axis is empirical thresholds for pre-earnings drift, defined as $-(1/2)\beta_1/\beta_2$ where $\beta$s of year T are obtained from regressing pre-earnings drift on media measures plus controls using the sample with years $\leq T$. X-axis is calendar year.
Figure 7: Fraction of Negatively-affected Stocks over Time. Y-axis is fraction of negatively-affected stocks, defined as fraction of firm-quarters with media coverage above empirical threshold for pre-earnings drift. X-axis is calendar year.
Figure 8: Trends in the past twenty years.
Table 5: Firms that Issue Guidance Persistently

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>p1</th>
<th>p50</th>
<th>p99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drift</td>
<td>27,966</td>
<td>0.94</td>
<td>0.07</td>
<td>0.71</td>
<td>0.96</td>
<td>1.00</td>
</tr>
<tr>
<td>Volatility</td>
<td>27,966</td>
<td>0.77</td>
<td>0.75</td>
<td>1.00</td>
<td>0.88</td>
<td>0.11</td>
</tr>
<tr>
<td>TNEWS</td>
<td>27,966</td>
<td>219</td>
<td>1271</td>
<td>1</td>
<td>95</td>
<td>1953</td>
</tr>
<tr>
<td>MNEWS</td>
<td>27,966</td>
<td>95</td>
<td>95</td>
<td>1</td>
<td>74</td>
<td>552</td>
</tr>
<tr>
<td>NNEWS</td>
<td>27,966</td>
<td>78</td>
<td>37</td>
<td>1</td>
<td>52</td>
<td>386</td>
</tr>
<tr>
<td>Book to Market</td>
<td>27,966</td>
<td>0.54</td>
<td>0.48</td>
<td>0.04</td>
<td>0.43</td>
<td>2.31</td>
</tr>
<tr>
<td>Market Cap (bn)</td>
<td>27,966</td>
<td>6.0</td>
<td>19.1</td>
<td>0.03</td>
<td>1.2</td>
<td>83.0</td>
</tr>
<tr>
<td>MADP</td>
<td>27,966</td>
<td>1531</td>
<td>311820</td>
<td>-10</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>AGE</td>
<td>27,966</td>
<td>19</td>
<td>13</td>
<td>5</td>
<td>16</td>
<td>48</td>
</tr>
<tr>
<td>INST</td>
<td>27,966</td>
<td>0.48</td>
<td>0.34</td>
<td>0.004</td>
<td>0.46</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>Std.Dev.</td>
<td>p1</td>
<td>p50</td>
<td>p99</td>
</tr>
<tr>
<td>--------------------------</td>
<td>---</td>
<td>------</td>
<td>----------</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>Drift</td>
<td>225,819</td>
<td>0.95</td>
<td>0.18</td>
<td>0.70</td>
<td>0.97</td>
<td>1.00</td>
</tr>
<tr>
<td>Volatility</td>
<td>225,819</td>
<td>0.83</td>
<td>0.78</td>
<td>1.00</td>
<td>0.93</td>
<td>0.13</td>
</tr>
<tr>
<td>TNEWS</td>
<td>225,819</td>
<td>188</td>
<td>1203</td>
<td>1</td>
<td>56</td>
<td>1786</td>
</tr>
<tr>
<td>MNEWS</td>
<td>225,819</td>
<td>86</td>
<td>252</td>
<td>1</td>
<td>48</td>
<td>481</td>
</tr>
<tr>
<td>NNEWS</td>
<td>225,819</td>
<td>54</td>
<td>68</td>
<td>1</td>
<td>32</td>
<td>320</td>
</tr>
<tr>
<td>Book to Market</td>
<td>225,819</td>
<td>0.76</td>
<td>15.17</td>
<td>0.03</td>
<td>0.55</td>
<td>3.75</td>
</tr>
<tr>
<td>Market Cap (bn)</td>
<td>225,819</td>
<td>4.2</td>
<td>21.8</td>
<td>0.01</td>
<td>0.4</td>
<td>73.3</td>
</tr>
<tr>
<td>AGE</td>
<td>225,819</td>
<td>19</td>
<td>13</td>
<td>4</td>
<td>16</td>
<td>47</td>
</tr>
<tr>
<td>INST</td>
<td>225,819</td>
<td>0.72</td>
<td>0.27</td>
<td>0.005</td>
<td>0.64</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 6: Firms that Does Not Issue Guidance
Table 7: Pre-earnings Drift and Media Coverage

<table>
<thead>
<tr>
<th></th>
<th>(1) Drift(N)</th>
<th>(2) Drift(G)</th>
<th>(3) Drift(N)</th>
<th>(4) Drift(G)</th>
<th>(5) Drift(N)</th>
<th>(6) Drift(G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDIA</td>
<td>0.0163***</td>
<td>0.0136***</td>
<td>0.0413***</td>
<td>0.0411***</td>
<td>0.0529***</td>
<td>0.0432***</td>
</tr>
<tr>
<td></td>
<td>(0.000756)</td>
<td>(0.00187)</td>
<td>(0.00204)</td>
<td>(0.00549)</td>
<td>(0.00195)</td>
<td>(0.00495)</td>
</tr>
<tr>
<td>MEDIASQ</td>
<td>-0.00289***</td>
<td>-0.00244***</td>
<td>-0.0244***</td>
<td>-0.0234***</td>
<td>-0.0261***</td>
<td>-0.0190***</td>
</tr>
<tr>
<td></td>
<td>(0.000151)</td>
<td>(0.000337)</td>
<td>(0.00154)</td>
<td>(0.00401)</td>
<td>(0.00114)</td>
<td>(0.00271)</td>
</tr>
<tr>
<td>AGE</td>
<td>-0.0000108</td>
<td>-0.000190</td>
<td>-0.000365</td>
<td>-0.000286</td>
<td>0.0000289</td>
<td>-0.000217</td>
</tr>
<tr>
<td></td>
<td>(0.000140)</td>
<td>(0.000228)</td>
<td>(0.000144)</td>
<td>(0.000212)</td>
<td>(0.000138)</td>
<td>(0.000196)</td>
</tr>
<tr>
<td>MCAP</td>
<td>0.00609***</td>
<td>0.00816***</td>
<td>0.00608***</td>
<td>0.00829***</td>
<td>0.00667***</td>
<td>0.00902***</td>
</tr>
<tr>
<td></td>
<td>(0.000337)</td>
<td>(0.00108)</td>
<td>(0.000339)</td>
<td>(0.00109)</td>
<td>(0.000339)</td>
<td>(0.00109)</td>
</tr>
<tr>
<td>BM</td>
<td>-0.00102**</td>
<td>0.00846***</td>
<td>-0.00107**</td>
<td>0.00845***</td>
<td>-0.000853*</td>
<td>0.00861***</td>
</tr>
<tr>
<td></td>
<td>(0.000473)</td>
<td>(0.00225)</td>
<td>(0.000474)</td>
<td>(0.00226)</td>
<td>(0.000471)</td>
<td>(0.00225)</td>
</tr>
<tr>
<td>IVOL</td>
<td>0.765***</td>
<td>0.614***</td>
<td>0.735***</td>
<td>0.582***</td>
<td>0.753***</td>
<td>0.629***</td>
</tr>
<tr>
<td></td>
<td>(0.0597)</td>
<td>(0.194)</td>
<td>(0.0595)</td>
<td>(0.194)</td>
<td>(0.0595)</td>
<td>(0.194)</td>
</tr>
<tr>
<td>TVOL</td>
<td>-1.139***</td>
<td>-1.278***</td>
<td>-1.097***</td>
<td>-1.241***</td>
<td>-1.119***</td>
<td>-1.286***</td>
</tr>
<tr>
<td></td>
<td>(0.0561)</td>
<td>(0.174)</td>
<td>(0.0559)</td>
<td>(0.173)</td>
<td>(0.0560)</td>
<td>(0.173)</td>
</tr>
<tr>
<td>MADP</td>
<td>-2.39e-08</td>
<td>-1.50e-08</td>
<td>-2.84e-08</td>
<td>-2.84e-08</td>
<td>-2.84e-08</td>
<td>-2.84e-08</td>
</tr>
<tr>
<td></td>
<td>(4.74e-08)</td>
<td>(4.76e-08)</td>
<td>(4.74e-08)</td>
<td>(4.76e-08)</td>
<td>(4.74e-08)</td>
<td>(4.76e-08)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses. *p < 0.10, **p < 0.10, ***p < 0.01. Dependent variable is (pre-earnings drift)$_{it}$, defined as $\frac{1 + r_i(t-22, t-1)}{1 + r_i(t-22, t-1)}$ where $t$ is earnings date, and $a = 2$. (N) denotes the sample that includes firms issuing no guidance, (G) denotes sample that includes firms persistently issuing guidance. MEDIA measures number of (1-2): total news, (3-4): major news, (5-6): novel news. MEDIA is in 100s. MEDIASQ is the square of MEDIA. MADP is mean absolute percentage deviation of earnings guidance. Standard errors are clustered at the firm level.
<table>
<thead>
<tr>
<th></th>
<th>(1) Volatility(N)</th>
<th>(2) Volatility(G)</th>
<th>(3) Volatility(N)</th>
<th>(4) Volatility(G)</th>
<th>(5) Volatility(N)</th>
<th>(6) Volatility(G)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDIA</td>
<td>0.0423**</td>
<td>0.0123</td>
<td>0.0449**</td>
<td>0.0175</td>
<td>0.0784**</td>
<td>0.0308**</td>
</tr>
<tr>
<td></td>
<td>(0.00350)</td>
<td>(0.00989)</td>
<td>(0.00497)</td>
<td>(0.0147)</td>
<td>(0.00450)</td>
<td>(0.0127)</td>
</tr>
<tr>
<td>MEDIASQ</td>
<td>-0.00889***</td>
<td>-0.00578***</td>
<td>-0.0170***</td>
<td>-0.0105**</td>
<td>-0.0210***</td>
<td>-0.0109***</td>
</tr>
<tr>
<td></td>
<td>(0.000799)</td>
<td>(0.00187)</td>
<td>(0.00183)</td>
<td>(0.00514)</td>
<td>(0.00147)</td>
<td>(0.00375)</td>
</tr>
<tr>
<td>AGE</td>
<td>-0.00139**</td>
<td>-0.000298</td>
<td>-0.00145**</td>
<td>-0.000362</td>
<td>-0.00126**</td>
<td>-0.000412</td>
</tr>
<tr>
<td></td>
<td>(0.000587)</td>
<td>(0.000744)</td>
<td>(0.000580)</td>
<td>(0.000640)</td>
<td>(0.000582)</td>
<td>(0.000600)</td>
</tr>
<tr>
<td>MCAP</td>
<td>-0.0120***</td>
<td>-0.00996**</td>
<td>-0.0132***</td>
<td>-0.00904*</td>
<td>-0.00990***</td>
<td>-0.00802*</td>
</tr>
<tr>
<td></td>
<td>(0.00113)</td>
<td>(0.00472)</td>
<td>(0.00114)</td>
<td>(0.00482)</td>
<td>(0.00114)</td>
<td>(0.00477)</td>
</tr>
<tr>
<td>BM</td>
<td>-0.00332**</td>
<td>0.0205**</td>
<td>-0.00401***</td>
<td>0.0216***</td>
<td>-0.00265*</td>
<td>0.0214***</td>
</tr>
<tr>
<td></td>
<td>(0.00151)</td>
<td>(0.00806)</td>
<td>(0.00152)</td>
<td>(0.00816)</td>
<td>(0.00150)</td>
<td>(0.00807)</td>
</tr>
<tr>
<td>IVOL</td>
<td>1.120***</td>
<td>1.647*</td>
<td>1.119***</td>
<td>1.803**</td>
<td>1.933***</td>
<td>1.822**</td>
</tr>
<tr>
<td></td>
<td>(0.208)</td>
<td>(0.874)</td>
<td>(0.209)</td>
<td>(0.880)</td>
<td>(0.208)</td>
<td>(0.877)</td>
</tr>
<tr>
<td>TVOL</td>
<td>-0.817***</td>
<td>-0.675</td>
<td>-0.815***</td>
<td>-0.785</td>
<td>-0.763***</td>
<td>-0.800</td>
</tr>
<tr>
<td></td>
<td>(0.198)</td>
<td>(0.792)</td>
<td>(0.199)</td>
<td>(0.799)</td>
<td>(0.197)</td>
<td>(0.795)</td>
</tr>
<tr>
<td>MADP</td>
<td>0.000000219</td>
<td>0.000000201</td>
<td>0.000000219</td>
<td>0.000000201</td>
<td>0.000000200</td>
<td>0.000000200</td>
</tr>
<tr>
<td></td>
<td>(0.000000233)</td>
<td>(0.000000233)</td>
<td>(0.000000233)</td>
<td>(0.000000233)</td>
<td>(0.000000233)</td>
<td>(0.000000233)</td>
</tr>
<tr>
<td>N</td>
<td>216502</td>
<td>27320</td>
<td>216502</td>
<td>27320</td>
<td>216502</td>
<td>27320</td>
</tr>
</tbody>
</table>

Table 8: Pre-earnings Volatility Share and Media Coverage

Standard errors in parentheses. * p < 0.10, ** p < 0.10, *** p < 0.10. Dependent variable is (pre-earnings volatility share)\(_t\), defined as \(\sum_{\tau=-22}^{t-1} r_{i,t+\tau}^2 / \sum_{\tau=-22}^{n} r_{i,t+\tau}^2\) where \(t\) is earnings date, and \(n = 2\). (N) denotes the sample that includes firms issuing no guidance, (G) denotes sample that includes firms persistently issuing guidance. MEDIA measures number of (1-2): total news, (3-4): major news, (5-6): novel news. MEDIA is in 100s. MEDIASQ is the square of MEDIA. MADP is mean absolute percentage deviation of earnings guidance. Standard errors are clustered at the firm level.