# News Selection and Asset Pricing

Charles Martineau and Jordi Mondria*

This draft: October 18, 2022

**Abstract**

This paper builds a theoretical framework to endogeneize the editorial decisions of media and analyze their asset pricing implications. The media outlet optimally reports man-bites-dog signals by choosing to report about the firm that generates more uncertainty to investors. There are three main implications of the model. First, the editorial choice is state-dependent and not only has asset pricing implications for reported firms, but also for non-reported firms. Second, the model generates an asymmetric response of asset prices to positive and negative news. Specifically, the asset price reaction is much stronger for negative news than for positive news. Third, public information does not necessarily crowd out the acquisition of private information. Failing to capture the information implications of editorial choices may lead the econometrician to estimate a misspecified asset pricing model. We provide empirical evidence inline with the model predictions.

*JEL Classification*: G10, G12, G14

*Keywords*: News, Media, Asset Prices, Information Acquisition, Public Information

*Martineau: University of Toronto, 105 St. George Street, Toronto ON, Canada, M5S 3E6 (charles.martineau@rotman.utoronto.ca). Mondria: University of Toronto, 150 St. George Street, Toronto ON, Canada, M5S 3G7 (jordi.mondria@utoronto.ca). We would like to thank Daniel Andrei, Vladimir Asriyan, Isaac Baley, Jesse Davis, Rahul Deb, Thierry Foucault, Naveen Gondhi (discussant), Marcin Kacperczyk, Jordan Martel, Kristoffer Nimark, Giang Nguyen (discussant), Jan Schneemeier, Laura Veldkamp, Xavier Vives, and seminar participants at CREi, Guelph, 2022 BFWG, 2022 CEA, 2022 EEA, 2022 ESW, 2022 NASMES. We are also grateful to Sedat Ersoy, Alykhan Jetha, and Luka Vulicevic for their excellent research assistance. The authors gratefully acknowledge the financial support from the Social Sciences and Humanities Research Council of Canada.

# 1 Introduction

The main objective of this paper is to study the role of media outlets in financial markets. An enormous amount of information is created every day, and each bit of information is potentially relevant for decision-making of households and firms. Thus, agents in the economy delegate their information choice to media outlets which have a better technology to monitor relevant events. Media outlets monitor worldwide news and select the most relevant events for publication. Most literature in finance takes the editorial decisions as given when studying the effects of media in financial markets. We aim to build a theoretical framework to endogeneize the editorial decisions of media and analyze their asset pricing implications. The decision to publish a story about a particular firm will not only provide information to investors about the firm selected for publication, but also will convey information about firms not selected for publication. As a consequence, the decision to select a firm to be reported in a media outlet will have asset pricing implications for both reported firms and non-reported firms. Failing to capture the information implications for both types of firms may lead the econometrician to estimate a misspecified asset pricing model.

Consider the following example: assume we could rank the priority of firms for news coverage based on firm characteristics. News about firms that are highly ranked would have priority over news of firms that have lower rank. Imagine now that the Financial Times publishes a story about a low-ranked company on the front-page. This story will, of course, have implications for the asset price of this low-ranked company (this is what has been studied in most of the literature). In addition, this editorial decision must also imply that on this particular day, there is no relevant news about high-ranked firms such as Microsoft or Apple, since any news story about those firms (no matter how small) would have made the front-page of the Financial Times, bumping out news from low-ranked firms. Hence, the front-page of the Financial Times may provide investors with information about firms not chosen to be reported. This argument would not be true if reversed. If the Financial Times were to publish a story about Microsoft, then the story would not convey any information about the low-ranked firms. There may be very relevant news about low-ranked firms that day, but news about Microsoft has priority in the Financial Times. Thus, modeling the selection of news by the media outlet is relevant to understand the information that investors have in each period.

We introduce editorial decisions to a multi-asset noisy rational expectations model. A key assumption in the model is that there is uncertainty about the risk-regime of each firm. Firms may be in a high volatility risk-regime or in a low volatility risk-regime. The media outlet has a monitoring advantage over investors about the risk regime and its choice consists of selecting one firm to publish a news story. The media outlet will choose to report about the firm that is generating more uncertainty to investors in a given period. Hence, the model endogeneizes the man-bites-dog signals. Investors know that when a news story gets reported, then tail events are more likely to occur. There are three main implications of the model. First, the editorial decision not only has implications for reported firms, but also for non-reported firms. Specifically, the editorial choice is state-dependent. The media outlet has an ex-ante ranking of publication priority. This ranking implies that if none of the top ranked firm are selected for publication, then it must be that these top ranked firms are in a low volatility risk-regime. Editorial decisions will lead to three different types of asset prices: i) there will be a specific asset price with public information about the firm that received media coverage; ii) there will be a specific asset price for non-reported firms that have a higher rank than those firms covered in the media that day. Investors know that there is no news about those firms that day; iii) there will be a specific asset price for non-reported firms that have a lower rank than those firms covered in the media that day. Investors will not know if there is news about those firms that day. Hence, editorial decisions will have asset prices implications for firms with news coverage and for firms with no coverage.

Second, the model generates an asymmetric response of asset prices to positive and negative news. In particular, the asset price reaction is much stronger for negative news than for positive news. Intuitively, a firm appears on the news when this firm is in a high volatility risk-regime, leading to an initial asset price decrease. Negative news generate an even stronger negative price reaction, while positive news generate a positive price reaction that counteracts the initial decrease in price due to the increased riskiness.

Third, an extension of the model, where investors are allowed to acquire private information, shows that public information does not necessarily crowd out the acquisition of private information. Since public information about a firm only appears on media outlets when the firm is in a high uncertainty regime, then investors have more incentives to process more private information when there are news about the firm.

In line with the theoretical predictions, we show empirically that the effect of media coverage on abnormal stock turnover and volatility (i.e., proxies for firm uncertainty) depends on expected media coverage. Previous literature has shown that low media coverage indicates more uncertainty (e.g., Fang & Peress, 2009). Our empirical results show that this result is driven by firms that have low expected media coverage. However, unlike conventional wisdom, for firms with high expected media coverage (e.g., large firms), we show empirically that lower than expected coverage leads to less uncertainty as the model predicts. Additionally, we find that stock returns are almost twice more responsive to negative news than positive news. Finally, in contrasts to Fang & Peress (2009), we find that media coverage is positively associated with higher expected stock returns when coverage exceeds expected coverage.

## 2 Literature Review

**Theoretical framework.** Our framework builds upon a standard rational expectations model of asset prices such as Grossman & Stiglitz (1980), Hellwig (1980) and Verrecchia (1982). These papers provide the foundation and the essential tools to build a theoretical framework of stock market trading, asset prices and information choices. The other key building block of the theoretical framework is the work of Admati & Pfleiderer (1986) and Admati & Pfleiderer (1987), where traders buy information from a monopolistic seller, which is subsequently used in a speculative market.

**Role of public information.** There is an extensive literature analyzing the role of public information on stock market trading and price discovery. Financial transparency has been a key aspect in improving the stability of our financial system. Investors need transparent financial statements to make informed investment decisions. Yet the literature analyzing the role of public information has challenged the conventional wisdom that more public information is welfare improving. Morris & Shin (2002) argue that public information may lead to too much coordination and overreaction to public information. Also, Amador & Weill (2012), Gao & Liang (2013), Han et al. (2016), Banerjee et al. (2018) and Goldstein & Yang (2019) study the impact of public information on the incentives to acquire information and real efficiency. Our paper will contribute to this strand of the literature by analyzing the role of editorial decisions on asset prices, portfolio decisions and information acquisition.

**Theory of media.** Our paper is closely related to the theory on state-dependent editorial behavior by Nimark & Pitschner (2019). In their paper, when reporting decisions are state-dependent, media outlets convey information not only via the contents of their news stories, but also via the editorial decision itself. In our paper, the media outlet choice of news to report is state-dependent as well and thus conveys information about non-reported firms. We introduce this state-dependent editorial behavior to a multi-asset noisy rational expectations model and study the financial market implications of these editorial decisions. In addition, our model extension with information acquisition is able to endogeneize the man-bites-dog signals of Nimark (2014), where events that generate more uncertainty are more likely to be reported. These man-bites-dog signals are also consistent with the survey evidence on financial journalists by Call et al. (2022). Financial journalists are more likely to report about firms and topics that are controversial and provocative.

Our project is also related to the existing theoretical literature that studies the editorial decisions of media in politics. This literature normally assumes that media outlets are concerned about their reputation as providers of political news stories such as Mullainathan & Shleifer (2005) and Gentzkow & Shapiro (2006). In these papers, media tends to bias their news stories to satisfy the beliefs of their readers. Alternatively, Perego & Yuksel (2018) focus on news provision of media outlets that are not partisan and show that media competition leads to information specialization. Instead, our project will focus on news stories about financial markets. The main difference with political news stories is that readers of financial news are able to trade on financial information released by media. Goldman et al. (2022) builds a theory of financial media, where journalists try to eliminate bias in the obfuscated announcements of firm managers. Our project abstracts from biases on media and announcements, and focuses on the editorial decision of which story should be reported based on the amount of information provided.

**Impact of media in financial markets.** There is a large empirical literature documenting a strong correlation between media and asset prices. A notorious case of the effect of media on financial markets was raised by Huberman & Regev (2001), where they document that a Sunday New York Times article on a potential cancer-curing drug caused the stock price of *EntreMed* to triple in a day, even though the potential breakthrough had been published in the journal Nature months before. Several papers have examined the implications of firm news to returns, volume and volatility (e.g., Chan, 2003; Fang & Peress, 2009; Tetlock, 2010). A common view of the implication of media coverage in financial markets is that a public news story decreases the information asymmetry between investors and result in lower stock returns and volatility. In contrast, we show that if we endogenize the decision of the media outlet to cover a specific firm, the relationship between media coverage and information asymmetry depends on whether a firm is generally well covered by the media. Recent literature is currently interested in addressing the causal relationship of media coverage in stock markets because a simple correlation may be just the result of omitted variables or reverse causality. Dougal et al. (2012) exploit exogenous rotation and writing style differences across Wall Street Journal columnists to identify the causal relation between financial reporting and stock market performance. Peress (2014) uses newspapers strikes to infer the causal impact of media on trading volume, returns and intraday volatility. Fisher et al. (2022) find that news coverage around macroeconomic news announcements increases in the amount uncertainty associated with announcements.

Our theoretical framework is consistent with the evidence presented by Schwenkler & Zheng (2022). Their paper shows empirically that media outlets provide a larger amount of information to their readers than just the reported current events and that financial media editors choose to report about stocks based on their risk characteristics. The asymmetric response of asset prices to good and bad news due to editorial decisions in our model resemble those implications by Campbell & Hentschel (1992), Veronesi (1999), and Calvet & Fisher (2007). In these models the sign of the news interact with the volatility of the market as in our paper.

## 3 Model Description

Let's consider an economy with three dates, $t = 0, 1, 2$. At $t = 1$, N+1 assets are traded: a riskless asset and N independent risky assets. The riskless asset has a constant value of 1 and is in unlimited supply. There are $N$ independent risky assets, and we assume N is a large number. Each risky asset $n \in N$ is traded at an endogenous price $p_n$, has a noisy supply of $\tilde{z}_n \sim N(\bar{z}, \tau_z^{-1})$, and pays an uncertain cash flow $\tilde{v}_n = \bar{\delta} + \tilde{\rho}_n\tilde{\delta}_n$ at date $t = 2$. Cash flows have three components: a constant benchmark cash flow $\bar{\delta}$, a firm-specific risk regime $\tilde{\rho}_n$ and a firm-specific risk factor $\tilde{\delta}_n$. The firm-specific regime $\tilde{\rho}_n$ consists of a binary random variable, with probability $\pi_n$ we have that $\tilde{\rho}_n = \rho_{h,n}$ and with probability $1-\pi_n$ we have that $\tilde{\rho}_n = \rho_{l,n}$, with $\rho_{h,n} > \rho_{l,n}$. This component captures that firms may be in a high volatility or low volatility regime. The firm-specific risk factor $\tilde{\delta}_n$ is a standard normally distributed random variable given by $\tilde{\delta}_n \sim N(0, \tau_\delta^{-1})$. All random variables $\tilde{z}_n$, $\tilde{\rho}_n$ and $\tilde{\delta}_n$ are mutually independent.

There are two types of agents in the economy: a media outlet and a continuum of investors of measure one. The media outlet reveals $\tilde{\rho}_{n^*}$ and provides a public signal $\tilde{y}_{n^*} = \tilde{\delta}_{n^*} + \tilde{\eta}_{n^*}$, where $\tilde{\eta}_{n^*} \sim N(0, \tau_\eta^{-1})$, about one and only one of the firms $n^* \in N$ at $t = 0$. We assume that the media outlet can only transmit one public signal. This assumption aims to capture the idea that media outlets have to choose one main topic for the front-page of the newspaper or main news story in a broadcast. The media outlet perfectly observes the realization of $\tilde{\rho}_n$ for all firms and can produce a signal $\tilde{y}_{n^*}$ about one of the firms. We assume that the media outlet perfectly transmits the risk regime $\tilde{\rho}_{n^*}$ of one firm for free with the headline of the front-page. With this assumption, we are trying to capture that a headline provides some information, i.e., in our case is the firm-specific regime, and the media outlet is not able to charge for just reading the headline. In addition, the media outlet transmits an imperfect public signal $\tilde{y}_{n^*}$ about one of the firms for a price with a pay-to-read news article. We follow Admati & Pfleiderer (1986) and Admati & Pfleiderer (1987) to determine the monopolistic media profits. The value of a private signal $\tilde{y}_{n^*}$ is the certainty equivalent of the information, which is determined by subtracting the level of ex-ante expected utility when only the price is observed from the ex-ante expected utility when the public signal is observed. This assumption ensures that all investors will choose to pay for the public signal.

There also exist a continuum of investors of measure one. Each investor $i$ has mean-variance preferences given by

$$EU_i = E_0\left[E_1[\tilde{W}_i \mid I_i] - \frac{\gamma}{2}V_1[\tilde{W}_i \mid I_i]\right], \tag{1}$$

where $E_t$ for $t = 0, 1$ represents the expected value with information available at time $t$, $V_t$ for $t = 1$ represents the variance conditional on information available at time $t$, $\gamma > 0$ is the coefficient of absolute risk aversion, $I_i$ is the information set of investor $i$ at $t = 1$, and $\tilde{W}_i$ is the final wealth. The investor has an initial endowment $W_{0i}$ of wealth that allocates between the $N + 1$ assets in the economy to maximize the investor's preferences subject to the following budget constraint

$$\tilde{W}_i = W_{0i} - \phi(\tilde{y}_{n^*}) + \sum_{n=1}^{N} D_{ni}(\tilde{v}_n - p_n), \tag{2}$$

where $D_{ni}$ are the asset holdings of risky asset $n$, and $\phi(\tilde{y}_{n^*})$ is the monetary value of the signal $\tilde{y}_{n^*}$ about firm $n^*$ released by the media outlet. Let us define $EU_{ni}$ as the contribution that each asset $n$ has in the expected utility of the investor $i$. For any firm $n$, $EU_{ni}$ is given by

$$EU_{ni} = E_1[D_{ni}(\tilde{v}_n - p_n)] - \frac{\gamma}{2}V_1[D_{ni}(\tilde{v}_n - p_n)]. \tag{3}$$

Hence, we can write total expected utility $EU_i$ as a sum of each asset's contribution:

$$EU_i = W_{0i} - \phi(\tilde{y}_{n^*}) + \sum_{n=1}^{N} E_0[EU_{ni}].$$

The timeline of the model is given by Figure 1.

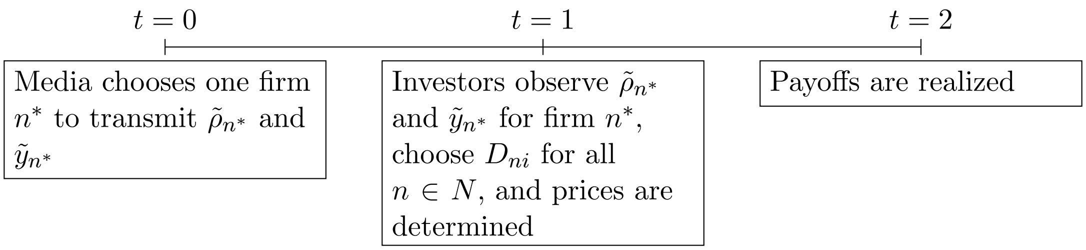


Figure 1: Timeline

The model is solved using backward induction. First each investor solves for the optimal portfolio when there is a media report and when there is no information. Then, given the optimal asset holdings under each information structure, the media outlet chooses to publish $\tilde{y}_{n^*}$ for one firm.

# 4 Investor's Problem

We first need to solve the investor's problem. Since we have mean-variance preferences and assets are independent, the holdings of each asset can be studied independently from each other. There are three scenarios to consider: $a$) the investor has no information about $\tilde{\rho}_n$ and $\tilde{\delta}_n$; $b$) the investor knows the realization of $\tilde{\rho}_n$, but has no information about $\tilde{\delta}_n$; $c$) the investor knows the realization of $\tilde{\rho}_n$, and has a public signal $\tilde{y}_n$ about $\tilde{\delta}_n$. Scenario $b$ will arise in equilibrium and it is a limiting case of scenario $c$ with a public signal $\tilde{y}_n$ that is completely uninformative. Thus, we solve for the scenario $a$ with no information and the scenario $c$ with a public signal $\tilde{y}_n$ about firm $n$.

We focus on symmetric equilibria, where all investors have the same information structure.

The only reason why we need a continuum of investors is to calculate the media fee and the discussion on information acquisition.

### 4.1 With No Information

If the investor has no information about cash flows, then the information set only includes the price $I_i = p_n$. Note that $p_n$ will not reveal any information about $\tilde{v}_n$ because no investor has any information about cash flows. In fact, $p_n$ will only provide (perfectly revealing) information about the noisy supply $\tilde{z}_n$.

In this scenario, cash flows $\tilde{v}_n$ do not follow a Normal Distribution and we cannot apply standard results from mean-variance preferences. Specifically, cash flows follow a mixture of two Normal distributions. For a given realization of $\tilde{\rho}_n$, cash flows do follow a Normal Distribution. If $\tilde{\rho}_n = \rho_{h,n}$, then cash follows $\tilde{v}_n|\rho_{h,n} \sim N(\bar{\delta}, \rho^2_{h,n}\tau_\delta^{-1})$. If instead $\tilde{\rho}_n = \rho_{l,n}$, then cash follows $\tilde{v}_n|\rho_{l,n} \sim N(\bar{\delta}, \rho^2_{l,n}\tau_\delta^{-1})$. Thus, the contribution of an asset $n$ to the total expected utility $EU_i$ when investors have no information about the asset is given by

$$\begin{aligned}
EU_{ni} &= E_1[D_{ni}(\tilde{v}_n - p_n) \mid p_n] - \frac{\gamma}{2}V_1[D_{ni}(\tilde{v}_n - p_n) \mid p_n] \\
&= \pi\left(E_1[D_{ni}(\tilde{v}_n - p_n) \mid p_n, \tilde{\rho}_n = \rho_{h,n}] - \frac{\gamma}{2}V_1[D_{ni}(\tilde{v}_n - p_n) \mid p_n, \tilde{\rho}_n = \rho_{h,n}]\right) + \\
&+ (1-\pi)\left(E_1[D_{ni}(\tilde{v}_n - p_n) \mid p_n, \tilde{\rho}_n = \rho_{l,n}] - \frac{\gamma}{2}V_1[D_{ni}(\tilde{v}_n - p_n) \mid p_n, \tilde{\rho}_n = \rho_{l,n}]\right). \quad (4)
\end{aligned}$$

The investor chooses the asset holdings of asset $n$ by maximizing (4) subject to (2). The optimal asset demand for asset $n$ when the investor has no information about cash flows is then given by

$$\begin{aligned}
D_{ni}(p_n) &= \frac{\pi E_1[\tilde{v}_n \mid \tilde{\rho}_n = \rho_{h,n}] + (1-\pi)E_1[\tilde{v}_n \mid \tilde{\rho}_n = \rho_{l,n}] - p_n}{\gamma(\pi V_1[\tilde{v}_n \mid \tilde{\rho}_n = \rho_{h,n}] + (1-\pi)V_1[\tilde{v}_n \mid \tilde{\rho}_n = \rho_{l,n}])} \\
&= \frac{(\bar{\delta} - p_n)\tau_\delta}{\gamma(\pi\rho^2_{h,n} + (1-\pi)\rho^2_{l,n})}. \quad (5)
\end{aligned}$$

We have removed $p_n$ from the information set because the price does not contain any information about the realization of $v_n$. Given the noisy supply of each asset is given by $\tilde{z}_n$, then the market clearing condition is given by $\int_0^1 D_{ni}di = \tilde{z}_n$ and asset prices are given by

$$p_n = \bar{\delta} - \frac{\gamma \tilde{z}_n(\pi \rho_{h,n}^2 + (1-\pi)\rho_{l,n}^2)}{\tau_\delta}. \tag{6}$$

The price perfectly reveals $\tilde{z}_n$, but contains no information about $\tilde{v}_n$.

### 4.2 With a public Signal

If investors receive a public signal $\tilde{y}_n$ about cash flows, then the realization of $\tilde{\rho}_n$ is also known. Recall that we assume that the media outlet freely and perfectly reveals the risk regime $\tilde{\rho}_n$ about one firm with a headline, but investors will have to pay for the signal $\tilde{y}_n$ about $\tilde{\delta}_n$. Let's assume for now (it will be true in equilibrium) that all investors are willing to pay for the signal. The information set of investor $i$ is now given by $I_i = \{p_n, \tilde{\rho}_n, \tilde{y}_n\}$. We conjecture a linear price function

$$p_n = a_{0n} + a_{yn}\tilde{y}_n + a_{zn}\tilde{z}_n,$$

where the $a$'s coefficients are endogenous. Note that the price will not reveal any additional information about cash flows, but it will reveal perfectly the realization of the noisy supply $\tilde{z}_n$.

The investor chooses the asset holdings of asset $n$ by maximizing (3) subject to (2). The optimal asset demand for asset $n$ when the investor has no information about cash flows is then given by

$$D_{ni}(p_n, \tilde{\rho}_n, \tilde{y}_n) = \frac{E_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, p_n] - p_n}{\gamma V_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, p_n]}, \tag{7}$$

where

$$E_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, p_n] = \bar{\delta} + \frac{\tilde{\rho}_n \tau_\eta \tilde{y}_n}{\tau_\delta + \tau_\eta},$$

and

$$V_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, p_n] = \frac{\tilde{\rho}_n^2}{\tau_\delta + \tau_\eta}.$$

If we plug the asset demand into the market clearing condition given by $\int_0^1 D_{ni} di = \tilde{z}_n$, then asset prices are given by

$$p_n = a_{0n} + a_{yn}\tilde{y}_n + a_{zn}\tilde{z}_n, \tag{8}$$

where

$$\begin{aligned} a_{0n} &= \bar{\delta}, \\ a_{yn} &= \frac{\tilde{\rho}_n \tau_\eta}{\tau_\delta + \tau_\eta}, \\ a_{zn} &= -\frac{\gamma \tilde{\rho}_n^2}{\tau_\delta + \tau_\eta}. \end{aligned} \tag{9}$$

It will be useful for the next section to derive asset prices when investors know the risk-regime $\tilde{\rho}_n$, but they do not receive any public information about the firm. In this case, we can take the $\lim_{\tau_\eta \to 0} p_n = a_{0n} + a_{yn}\tilde{y}_n + a_{zn}\tilde{z}_n$ in equation (8), which is given by

$$p_n = a_{0n} + a_{yn}\tilde{y}_n + a_{zn}\tilde{z}_n, \tag{10}$$

where

$$\begin{aligned} a_{0n} &= \bar{\delta}, \\ a_{yn} &= 0, \\ a_{zn} &= -\frac{\gamma \tilde{\rho}_n^2}{\tau_\delta}. \end{aligned} \tag{11}$$

## 5 Media Problem

The media outlet chooses to publish a news story about one firm to maximize their profits. We follow Admati & Pfleiderer (1986) and Admati & Pfleiderer (1987) to determine the monopolistic media profits. The value of a private signal $\tilde{y}_n$ is the certainty equivalent of the information, which is determined by subtracting the level of ex-ante expected utility when only the price is observed from the ex-ante expected utility when the public signal is observed. The media outlet observes the realization of $\tilde{\rho}_n$ for all $n \in N$, calculates the profits that each firm $n$ would generate if a signal were to be published, and chooses to publish a story about only one firm $n^* \in N$.

For any firm $n$, the media outlet profits for a given realization of $\tilde{\rho}_n$ are given by

$$\begin{aligned} Profit_n(\tilde{\rho}_n) = \phi(\tilde{y}_{n^*}) =& E_0\{E_1[D_{ni}(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n] - \frac{\gamma}{2} V_1[D_{ni}(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n] \mid \tilde{\rho}_n\} \\ & - E_0\{E_1[D_{ni}(\tilde{v}_n - p_n) \mid \tilde{\rho}_n, p_n] - \frac{\gamma}{2} V_1[D_{ni}(\tilde{v}_n - p_n) \mid \tilde{\rho}_n, p_n] \mid \tilde{\rho}_n\}. \end{aligned}$$

Note that the media outlet knows the realization of $\tilde{\rho}_n$ for all firms, and that investors when deciding whether they want to buy the signal, they will also know the $\tilde{\rho}_{n^*}$ of the published firm through a free headline (recall that we assume that $\tilde{\rho}_{n^*}$ is freely revealed by media). In the appendix, we show that, for a given $\tilde{\rho}_n$, media profits of firm $n$ can be written as

$$\begin{aligned} Profit_n(\tilde{\rho}_n) = \phi(\tilde{y}_n) &= \frac{1}{2\gamma} V[\tilde{v}_n - p_n \mid \tilde{\rho}_n] \left( \frac{1}{V[\tilde{v}_n \mid \tilde{\rho}_n, p_n, \tilde{y}_n]} - \frac{1}{V[\tilde{v}_n \mid \tilde{\rho}_n, p_n]} \right) \\ &= \frac{\gamma \tau_\eta \tilde{\rho}_n^2 (\gamma^2 \tilde{\rho}_n^2 + \tau_z(\tau_\delta + \tau_\eta))}{2\tau_z(\tau_\delta + \tau_\eta)^2 (\gamma^2 \tilde{\rho}_n^2 + \tau_\eta \tau_z)} \end{aligned} \tag{12}$$

Media profits for all firms have the same structure and only differ by the realization of $\tilde{\rho}_n$. Hence, the media outlet can just focus on the realization of $\tilde{\rho}_n$ to decide what story to publish.

**Lemma 1** *$Profit_n(\tilde{\rho}_n)$ is increasing in $\tilde{\rho}_n$. Thus, the media outlet will choose to provide a* *public signal about the firm with the highest realization of $\tilde{\rho}_n$.*

For any given firm, the media outlet is able to charge a higher fee when publishing news about risk regime $\rho_{h,n}$ than risk-regime $\rho_{l,n}$. From the lemma above, we know that $Profit_n(\rho_{h,n}) > Profit_n(\rho_{l,n})$. Hence, the media outlet can just focus on the high realizations $\rho_{h,n}$, and rank all firms by $\rho_{h,n}$, which is a sufficient statistic of $Profit_n(\rho_{h,n})$. The media outlet can rank all firms by $\rho_{h,n}$ setting firm $n = 1$ as the firm with the highest $\rho_{h,n}$ and $n = N$ as the firm with the lowest $\rho_{h,n}$. Let's assume that if two firms have the same level of profits, then the media outlet will randomly select one of them with equal probability. This result is in line with standard results from the literature on information acquisition, which state that the value of information is higher when there is more risk. We need the following two definitions to be able to state the additional results of the model.

**Definition 1** *Let us define $\check{n}$ as $\check{n} = \arg\max_n \{\rho_{l,n}\}_{n=1}^{N}$.*

The firm $\check{n}$ is the firm with the highest $\rho_{l,n}$. Note that having the highest $\rho_{l,n}$ is independent of how high is $\rho_{h,n}$.

**Definition 2** *Let us define $\hat{n}$ as the lowest $\hat{n}$ such that $\rho_{h,\hat{n}} < \rho_{l,\check{n}}$.*

The firm $\hat{n}$ is a firm for which their highest realization of $\tilde{\rho}_{\hat{n}}$ is smaller than $\rho_{l,\check{n}}$ for firm $\check{n}$. Hence, it is always more preferable for the media outlet to publish a story about $\check{n}$ (independently of the realization of $\tilde{\rho}_{\check{n}}$), than to publish a story of firm $\hat{n}$ with the highest realization of $\tilde{\rho}_n$. Thus, it will never be optimal for the media outlet to publish a story about firm $\hat{n}$. The next result shows that if the media outlet can make more profits by selling a public signal about firm $\check{n}$ with $\rho_{l,\check{n}}$ rather than publishing a story about any firm $n'$ with $\rho_{h,n'}$, then firm $n'$ will never see a story published in a media outlet.

**Lemma 2** *Any firm $n$ such that $n \geq \hat{n}$ will never get a news story on media.*

Instead, any firm below $\hat{n}$ will get their stories published in media sometimes.

**Lemma 3** *Any firm $n$ such that $n < \hat{n}$ will get a news story on media with positive probability.*

When media publishes a story about a firm $n^*$, the media is indirectly revealing the risk-regime state of higher ranked firms to the public. If the media outlet publishes a public signal about $n^*$, then it must be the case that for any firm $n$ such that $n < n^*$, the risk-regime factor is $\rho_{l,n}$. Intuitively, when a media outlet publishes a story about a firm $n^*$, then any higher ranked firms must be in the low volatility risk-regime since any of them would have been selected for publication before $n^*$ if they had a realization $\rho_{h,n}$. In this case, the published story provides information about higher ranked firms not selected for publication.

Using a similar argument, if the media outlet publishes a public signal about $n^*$, then it must be the case that for any firm $n$ such that $n > n^*$, the risk-regime factor is unknown. Intuitively, when a media outlet publishes a story about a firm $n^*$, then the risk regime of all the lower ranked firms is unknown since the media outlet would not have published a story about firm $n$ even if the firm was in the high risk-regime. In this case, the published story does not convey any information about lower ranked firms. The case where the media outlet publishes a story about firm $n^* = 1$ is the scenario that generates more uncertainty in the market. Intuitively, investors know that firm $n^* = 1$ is in a high uncertainty scenario and does not have any information about the risk regime of any other firm.

The next proposition discusses the asset pricing implications of editorial decisions. The publication of a news story about *one* firm will produce three different types of asset prices.

**Proposition 1** *When the media outlet publishes a signal $y_{n^*}$ about firm $n^*$ when $\tilde{\rho}_{n^*} = \rho_{h,n^*}$, then*

1. *Firm $n^*$ is in a high volatility risk-regime $\rho_{h,n^*}$ and asset prices are given by (8) with $\tilde{\rho}_{n^*} = \rho_{h,n^*}$.*

2. *Any firm $n$ such that $n < n^*$ is in a low volatility risk-regime $\rho_{l,n}$ with no public signal and asset prices are given by (10) with $\tilde{\rho}_n = \rho_{l,n}$.*

3. *Any firm $n$ such that $n > n^*$ is in an unknown risk-regime and asset prices are given by (6).*

The firm selected for publication will have asset pricing implications for both reported and non-reported firms. For any reported firm except for the special case of firm $\check{n}$, the model endogenously generates a man-bites-dog signal as in Nimark (2014). News stories get reported when the risk-regime is high and tail events are more likely to occur. For unreported firms ranked above the published firm, not being published means that these firms are in a low risk regime and they will have high asset prices. While for unreported firms ranked below the published firm, not being published means that investors are uncertain about their risk regime and will have low asset prices. The next corollary analyzes the asset prices in the case that the media outlet publishes a story about firm $\check{n}$ when this firm is in the low volatility risk-regime $\tilde{\rho}_{\check{n}} = \rho_{l,\check{n}}$.

**Corollary 1** *When the media outlet publishes a signal $y_{\check{n}}$ about firm $\check{n}$ when $\tilde{\rho}_{\check{n}} = \rho_{l,\check{n}}$, then*

1. *Firm $\check{n}$ is in a low volatility risk-regime $\rho_{l,\check{n}}$ and asset prices are given by (8) with $\tilde{\rho}_{\check{n}} = \rho_{l,\check{n}}$.*

2. *Any firm $n$ such that $n < \hat{n}$ is in a low volatility risk-regime $\rho_{l,n}$ with no public signal and asset prices are given by (10) with $\tilde{\rho}_n = \rho_{l,n}$.*

3. *Any firm $n$ such that $n \geq \hat{n}$ is in an unknown risk-regime and asset prices are given by (6).*

The next result states that a firm may have different asset prices even if there is no new information about that particular firm.

**Corollary 2** *A firm $n'$ with the exact same realizations of cash flows and noisy supply may have different asset prices depending on the story reported in the news.*

This corollary is able to explain why asset prices move even in the absence of relevant specific macro or micro information about the firm. The asset price moves because of information published about a completely unrelated firm. Hence, when analyzing asset prices, it is important to analyze the effect of editorial decisions on unreported firms. The next result shows that there are asymmetric effects to good and bad news.

**Proposition 2** *For $\bar{z} > 0$, when $\tilde{\rho}_{n^*} = \rho_{h,n^*}$ negative news have a stronger price reaction than positive news.*

Intuitively, when $N$ is large, the story reported about the media outlet will be about a firm with a high volatility risk-regime. The increased riskiness of the asset will lead to an initial drop on its expected price. This effect only occurs when $\bar{z}$ is positive.[1] Negative news (modeled as $\tilde{y}_{n^*} < 0$) will accentuate even more the decrease in asset prices, leading to a strong price reaction to negative news. Instead, positive news (modeled as $\tilde{y}_{n^*} > 0$) will lead to an increase in price that will counteract the decrease in price generated by the high volatility risk-regime. Hence, negative news lead to an unambiguous price decrease, while positive news generate an ambiguous effect on price depending on how the increased in riskiness is compensated by the positive realization of the signal.

[1] If $\bar{z} = 0$, then expected price would still be zero even if there is an increase in the riskiness of the asset.

# 6 Public Information Crowding Out

The objective of this section is to reconcile the apparent disconnect between the theoretical literature on information acquisition and the empirical literature on attention allocation. In the theoretical literature, a public signal decreases the traders' incentives to acquire information because a public signal decreases uncertainty about the asset. In other words, public information crowds out private information. Instead, the empirical literature on attention allocation finds that investors acquire more information when a firm appears in a media outlet. Our model is able to reconcile these two results. The editorial decision of a media outlet is to publish news when companies are in a high risk regime (when there is high uncertainty), hence investors will choose to acquire more information when a firm appears on the news.

Let us modify the theoretical framework to include private information. To this end, we extend the baseline model as follows: each investor $i$ receives a private signal $\tilde{s}_{ni} = \tilde{\delta}_n + \tilde{\varepsilon}_{ni}$ about each asset $n$, where $\tilde{\varepsilon}_{ni} \sim N(0, \tau_{\varepsilon ni}^{-1})$. Let us also assume that the cost of acquiring information is given by $C(\tau_{\varepsilon ni}) = \frac{1}{2}\tau_{\varepsilon ni}^2$. The budget constraint is then given by

$$\tilde{W}_i = W_{0i} - \phi(\tilde{y}_{n^*}) + \sum_{n=1}^{N} D_{ni}(\tilde{v}_n - p_n) - \sum_{n=1}^{N} C(\tau_{\varepsilon ni}). \tag{13}$$

The timeline of the model is now given by Figure 2.

| $t = 0$ | $t = 1$ | $t = 2$ | $t = 3$ |
|---|---|---|---|
| Media chooses one firm $n^*$ to transmit $\tilde{\rho}_{n^*}$ and $\tilde{y}_{n^*}$ | Investors observe $\tilde{\rho}_{n^*}$ and choose $\tau_{\varepsilon ni}$ for each asset $n$ | Investor $i$ observes $\tilde{y}_{n^*}$ and $\tilde{s}_{ni}$ for all $n$, chooses $D_{ni}$ for all $n$, and prices are determined | Payoffs are realized |

Figure 2: Timeline

The model is solved using backward induction. First each investor solves for the optimal portfolio when there is a media report and when there is no information. Second, given the optimal asset holdings under each information structure, investors acquire private information about each asset. Then, given asset demands and information acquisition choices, the media outlet chooses to publish $\tilde{y}_{n^*}$ for one firm.

### 6.1 Portfolio Choice and Asset Prices

In this section, we solve the portfolio choice and asset prices for firms with and without media reports.

**Firms with media report**

If investors receive a public signal $\tilde{y}_n$ about cash flows, then the realization of $\tilde{\rho}_n$ is also known. Let's assume for now (it will be true in equilibrium) that the cost of the signal is low enough so that every investor is willing to pay for the signal. We conjecture a linear price function

$$p_n = a_{0n} + a_{\delta n}\tilde{\delta}_n + a_{yn}\tilde{y}_n + a_{zn}\tilde{z}_n,$$

where the $a$'s coefficients are endogenous. Unlike the baseline model, the price will now reveal additional information about cash flows. The information contained in the price is equivalent to a signal $\tilde{s}_{pn}$:

$$\tilde{s}_{pn} = \frac{p_n - a_{0n} - a_{yn}\tilde{y}_n}{a_{\delta n}} = \tilde{\delta}_n + \alpha_n^{-1}\tilde{z}_n,$$

where $\alpha_n = a_{\delta n}/a_{zn}$. The information set of investor $i$ is now given by $I_i = \{\tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}\}$. We consider a symmetric equilibrium in which all investors choose the same amount of private information. Hence, we impose $\tau_{\varepsilon i,n} = \tau_{\varepsilon n}$.

The investor chooses the asset holdings of asset $n$ by maximizing (3) subject to (13). The optimal asset demand for asset $n$ is then given by

$$D_{ni}(p_n, \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}) = \frac{E_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}] - p_n}{\gamma V_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}]}, \tag{14}$$

where

$$E_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}] = \bar{\delta} + \tilde{\rho}_n \frac{\tau_{\varepsilon n}\tilde{s}_{ni} + \tau_\eta \tilde{y}_n + \alpha_n^2 \tau_z \tilde{s}_{pn}}{\tau_\delta + \tau_{\varepsilon n} + \tau_\eta + \alpha_n^2 \tau_z},$$

and

$$V_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}] = \frac{\tilde{\rho}_n^2}{\tau_\delta + \tau_{\varepsilon n} + \tau_\eta + \alpha_n^2 \tau_z}.$$

If we plug the asset demand into the market clearing condition given by $\int_0^1 D_{ni} di = \tilde{z}_n$, then asset prices are given by

$$p_n = a_{0n} + a_{\delta n}\tilde{\bar{\delta}}_n + a_{yn}\tilde{y}_n + a_{zn}\tilde{z}_n, \tag{15}$$

where

$$\begin{aligned} a_{0n} &= \bar{\delta}, \\ a_{\delta n} &= \frac{\tilde{\rho}_n(\tau_{\varepsilon n} + \alpha_n^2 \tau_z)}{\tau_\delta + \tau_{\varepsilon n} + \tau_\eta + \alpha_n^2 \tau_z}, \\ a_{yn} &= \frac{\tilde{\rho}_n \tau_\eta}{\tau_\delta + \tau_{\varepsilon n} + \tau_\eta + \alpha_n^2 \tau_z}, \\ a_{zn} &= \frac{\tilde{\rho}_n \alpha_n \tau_z - \gamma \tilde{\rho}_n^2}{\tau_\delta + \tau_{\varepsilon n} + \tau_\eta + \alpha_n^2 \tau_z}, \end{aligned} \tag{16}$$

where $\alpha_n = a_{\delta n}/a_{zn}$ is given by

$$\alpha_n = -\frac{\tau_{\varepsilon n}}{\tilde{\rho}_n \gamma}.$$

**Firms with known risk-regime but without media report**

In this section we solve for asset prices when investors know the risk-regime $\tilde{\rho}_n$, but they do not receive any public information about the firm. In this case, we can take the $\lim_{\tau_\eta \to 0} p_n = a_{0n} + a_{\delta n}\tilde{\bar{\delta}}_n + a_{yn}\tilde{y}_n + a_{zn}\tilde{z}_n$ in equation (15), which is given by

$$p_n = a_{0n} + a_{\delta n}\tilde{\bar{\delta}}_n + a_{yn}\tilde{y}_n + a_{zn}\tilde{z}_n, \tag{17}$$

where

$$
\begin{aligned}
a_{0n} &= \bar{\delta}, \\
a_{\delta n} &= \frac{\tilde{\rho}_n(\tau_{\varepsilon n} + \alpha_n^2 \tau_z)}{\tau_\delta + \tau_{\varepsilon n} + \alpha_n^2 \tau_z}, \\
a_{yn} &= 0, \\
a_{zn} &= \frac{\tilde{\rho}_n \alpha_n \tau_z - \gamma \tilde{\rho}_n^2}{\tau_\delta + \tau_{\varepsilon n} + \alpha_n^2 \tau_z},
\end{aligned} \tag{18}
$$

where $\alpha_n = a_{\delta n}/a_{zn}$ is given by

$$
\alpha_n = -\frac{\tau_{\varepsilon n}}{\tilde{\rho}_n \gamma}.
$$

**Firms with unknown risk-regime and no media report**

In this section, we solve for portfolio choice and asset prices when the risk-regime $\tilde{\rho}_n$ is unknown and there is no public signal. We conjecture a linear price function

$$
p_n = a_{0n} + a_{\delta n}\tilde{\delta}_n + a_{zn}\tilde{z}_n,
$$

where the $a$'s coefficients are endogenous. The information contained in the price is equivalent to a signal $\tilde{s}_{pn}$:

$$
\tilde{s}_{pn} = \frac{p_n - a_{0n}}{a_{\delta n}} = \tilde{\delta}_n + \alpha_n^{-1}\tilde{z}_n,
$$

where $\alpha_n = a_{\delta n}/a_{zn}$. We consider a symmetric equilibrium in which all investors choose the same amount of private information. Hence, we impose $\tau_{\varepsilon i,n} = \tau_{\varepsilon n}$. The investor chooses the asset holdings of asset $n$ by maximizing (3) subject to (13). The optimal asset demand for asset $n$ when the investor has no information about cash flows is then given by

$$
D_{ni}(p_n, \tilde{s}_{ni}) = \frac{\bar{\delta} + (\pi\rho_{h,n} + (1-\pi)\rho_{l,n})E_1[\tilde{\delta}_n \mid \tilde{s}_{ni}, \tilde{s}_{pn}] - p_n}{\gamma(\pi\rho_{h,n}^2 + (1-\pi)\rho_{l,n}^2)V_1[\tilde{\delta}_n \mid \tilde{s}_{ni}, \tilde{s}_{pn}]}, \tag{19}
$$

where

$$E_1[\tilde{\delta}_n \mid \tilde{s}_{ni}, \tilde{s}_{pn}] = \frac{\tau_{\varepsilon n}\tilde{s}_{ni} + \alpha_n^2 \tau_z \tilde{s}_{pn}}{\tau_\delta + \tau_{\varepsilon n} + \alpha_n^2 \tau_z},$$

and

$$V_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}] = \frac{1}{\tau_\delta + \tau_{\varepsilon n} + \alpha_n^2 \tau_z}.$$

Given the noisy supply of each asset is given by $\tilde{z}_n$, then the market clearing condition is given by $\int_0^1 D_{ni} di = \tilde{z}_n$ and asset prices are given by

$$p_n = \bar{\delta} + (\pi\rho_{h,n} + (1-\pi)\rho_{l,n})E_1[\tilde{\delta}_n \mid \tilde{s}_{ni}, \tilde{s}_{pn}] - \gamma(\pi\rho_{h,n}^2 + (1-\pi)\rho_{l,n}^2)V_1[\tilde{\delta}_n \mid \tilde{s}_{ni}, \tilde{s}_{pn}]\tilde{z}_n. \quad (20)$$

### 6.2 Information Acquisition

The information acquisition level for each asset $n$ under each scenario is determined by inserting the asset demand function $D_{ni}$ of each scenario solved in the previous section to the expected utility function (3) and maximizing with respect to $\tau_{\varepsilon ni}$. Then, we solve for a symmetric equilibrium in information acquisition levels by imposing $\tau_{\varepsilon ni} = \tau_{\varepsilon n}$ in the first-order conditions. For the case where firms know their risk-regime and receive a public signal, the maximization problem becomes:

$$\max_{\tau_{\varepsilon ni}} \frac{1}{2\gamma} \frac{V_0[\tilde{v}_n - p_n]}{V_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}]} - C(\tau_{\varepsilon ni}).$$

If $C(\tau_{\varepsilon ni}) = \frac{1}{2}\tau_{\varepsilon ni}^2$ and imposing that $\tau_{\varepsilon ni} = \tau_{\varepsilon n}$ in the first-order conditions, then $\tau_{\varepsilon n}$ is implicitly given by

$$\tau_{\varepsilon n} = \frac{1}{2\gamma} \frac{V_0[\tilde{v}_n - p_n]}{\tilde{\rho}_n^2}. \quad (21)$$

### 6.3 Editorial decision

The media outlet chooses to publish a news story about one firm to maximize their profits. As derived in section 5, for a given $\tilde{\rho}_n$, media profits of firm $n$ can be written as

$$Profit_n(\tilde{\rho}_n) = \frac{1}{2\gamma} V[\tilde{v}_n - p_n \mid \tilde{\rho}_n] \left( \frac{1}{V[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}]} - \frac{1}{V[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{s}_{ni}, \tilde{s}_{pn}]} \right)$$

It is important to note that if an investor wants to deviate from an equilibrium where everyone chooses to observe the public signal, then their information acquisition level for the deviated investor will still be the same as everyone else as derived in (21). Media profits for all firms have the same structure and only differ by the realization of $\tilde{\rho}_n$. Hence, the media outlet can just focus on the realization of $\tilde{\rho}_n$ to decide what story to publish.

**Lemma 4** *$Profit_n(\tilde{\rho}_n)$ is increasing in $\tilde{\rho}_n$. Thus, the media outlet will choose to provide a public signal about the firm with the highest realization of $\tilde{\rho}_n$.*

For any given firm, the media outlet is able to charge a higher fee when publishing news about risk regime $\rho_{h,n}$ than risk-regime $\rho_{l,n}$. In short, if $C(\tau_{\varepsilon ni}) = \frac{1}{2}\tau_{\varepsilon ni}^2$, then the editorial decisions of the media outlet will be the same with or without information acquisition. Hence, all the results derived before for the media outlet apply to the case with information acquisition.

### 6.4 Interaction of private and public information

This section shows that public information does not necessarily crowd out private information. Let us consider a firm $n$ such that $n \neq \check{n}$. Intuitively, when the media outlet publishes a story about a firm $n$, then firm $n$ is in a high volatility risk-regime. While, if the media outlet publishes a story about firm $n'$ such that $n' > n$, then firm $n$ is in a low volatility risk-regime. Since investors choose to acquire more information when there is higher uncertainty, if the public signal was quite uninformative, then investors would choose to acquire more information when the media outlet publishes a story about the firm than when the media outlet publishes a story about any firm $n'$.

**Proposition 3** *For sufficiently small $\tau_\eta$, public information does not crowd out private information.*

This result reconciles the apparent disconnect between the theoretical literature on information acquisition and the empirical literature on attention allocation. For low enough precision of the public signal, traders' incentives to acquire information increase when the media outlet publishes a story about a firm, consistent both with the empirical literature on attention allocation and the theoretical literature on information acquisition. It is consistent with the empirical literature on attention allocation because public information leads to more attention to the firm by investors choosing a higher level of information acquisition. It is also consistent with the theoretical literature on information acquisition as traders choose to acquire more information when there is more uncertainty about the payoffs. In this literature, the uncertainty about payoffs is held constant when public information is released, which leads to a decrease in the information acquired by traders. In contrast, in our model, a public signal about the firm implies that the firm will be in a high volatility risk-regime and will lead to an increase in information acquisition. The key feature is that the uncertainty about payoffs changes when a public signal is released.

# 7 Linking Theory to Empirical Predictions about Volume, Volatility, and Stock Return Response to News

The main implication of the model is that editorial decisions about one firm will impact non-reported firms. The publication ranking in the model is only based on the firm-specific risk regime. In other words, the uncertainty about the assets is what drives editorial decisions. The model is clearly a simplification of how editorial decisions are taken in reality. In the real world, there are many more drivers of news coverage, i.e., size of the firm among many others. Yet if we could rank firms by publication priority through an empirical analysis, then we could test the main implications of the model. This is what we do in the next empirical section. We first study the determinants of news coverage to ascertain the expected news coverage of a firm. Then, we analyze the asset pricing implications of receiving more or less than the expected news coverage. Firms with high expected news coverage are firms that are highly ranked in the publication priority. Hence, if these firms receive lower than expected news, we can then conclude that they are in a low volatility regime. While for firms with low expected new coverage that receive lower than expected news, we can then conclude that investors have high uncertainty about their firm-specific risk regime. Thus, we will use expected news coverage as the publication ranking and we will use below expected coverage as a measure of non-reported firms.

### 7.1 Empirical findings

We retrieve editorial articles from Ravenpack that include Wall Street Journal, Barron's, Dow Jones, and MarketWatch. Specifically, we select only full-length articles. The sample period is from January 2000 to December 2021. We sum the number of articles per month for each firm. We also retrieve the sentiment scores for each articles that is computed by RavenPack.[2] We then select US-traded stocks from CRSP (with share code 10 and 11 and exchange code 1,2, and 3) and retrieved monthly returns and trade volume. Price volatility and turnover are commonly associated with the level of uncertainty surrounding the stock (e.g., Zhang, 2006;

[2] We use the ESS as sentiment score from Ravenpack. We rescale the sentiment measure to be between -1 (negative sentiment) and 1 (positive sentiment).

Barinov, 2014).

We first examine the drivers of news coverage by running the following regression:

$$\begin{aligned} Ln\ article_{i,t} = \alpha + &\beta_1 Ln\ MCAP_{i,t} + \beta_2 \mathbb{1}_{\mathrm{EA},i,t} + \beta_3 Analyst_{i,t} + \beta_4 Turnover_{i,t} + \beta_5 IO_{i,t} \\ &+ \beta_6 (Ret_{i,t} - Ret_t^M) + \beta_7 IVOL_{i,t} + \beta_8 Age_{i,t} + \beta_9 IndFE_i + \beta_{10} TimeFE_t + \epsilon_{i,t}, \end{aligned} \tag{22}$$

where $Ln\ article$ corresponds to the natural logarithm of the 1+total number of RavenPack editorial articles in month $t$ for stock $i$, $Ln\ MCAP$ is the natural logarithm of firm market capitalization at the end of month $t$, $\mathbb{1}_{\mathrm{EA}}$ is an indicator variable equal to one if stock $i$ has an earnings announcement on month $t$ and zero otherwise, $Analyst$ is the number of analyst following from I/B/E/S, $Turnover$ is the monthly share turnover, $IO$ is the fraction of shares held by institutions in month $t$ during the quarter of the respective month, $(Ret - Ret^M)$ is the excess stock return over the CRSP value-weighted market return, $IVOL$ is the stock's idiosyncractic volatility computed as in Ang et al. (2006), and $Age$ is the number of years since appearance in CRSP. $IndFE$ and $TimeFE$ are the industry (GIC 2-digit sector code) and year-month fixed effects, respectively. $Analyst$, $(Ret - Ret^M)$, $IVOL$, and $Turnover$ are rescaled to have a standard deviation of one.

For various regression specifications reported in Table 1, firm's market capitalization and analyst coverage are the main drivers to news coverage. Column (5) shows that a one percent increase in firm size is associated with an increase of 0.156 (log) articles and a one standard deviation increase in analyst coverage, turnover, and volatility is associated with an increase of 0.073, 0.065, and 0.058 articles, respectively.

We next run Equation (22) on each year-month (excluding the time fixed effects) and retrieve the fitted values to measure "expected" news coverage. We next compute "unexpected" news coverage a firm receives over month $t + 1$ as the difference between the log number of articles in month $t+1$ and the expected number of articles estimated from the fitted values at month $t$. The model predicts firms with lower than expected coverage are *predicted* to have higher expected volatility and turnover due to higher information uncertainty.

We first report in Figure 3 the average abnormal turnover and abnormal idiosyncratic volatility over month $t+1$ by unexpected news quintiles in Panels A and B, respectively. We compute abnormal turnover (indiosyncratic volatility) as the difference between turnover (idiosyncratic volatility) minus a six-month rolling average in turnover (idiosyncratic volatility). The figure shows that abnormal turnover and volatility, i.e., uncertainty, increases in the amount of unexpected news coverage.

We next examine how abnormal turnover and idiosyncratic volatility differs in the amount of unexpected news after conditioning on expected news coverage, i.e., the firm news ranking. Figure 4 shows the average abnormal turnover and abnormal idiosyncratic volatility on month $t+1$ for stocks with high and low unexpected news by the quintile of expected news coverage on month $t+1$ in Panels A and B, respectively. We define stocks with high (low) unexpected news when the number of news articles in month $t+1$ is above (below) expected news computed on month $t$. Our main focus are firms with low unexpected news (dark bars). For these firms, Figure 4 shows a large drop in abnormal turnover and abnormal idiosyncratic volatility when the firm expected news ranking is high. Such patterns are inline with the model predictions. In other words, when a firm is generally well-covered by the media (i.e., high expected news) such as large firms, its uncertainty drops when it receives below than expected coverage. In contrasts to firms with high expected coverage, uncertainty is higher for stocks that are not well-covered (e.g., small firms) when its news coverage falls below expected coverage.

Another empirical prediction of the model is that stock returns are more responsive to negative news sentiment and more so for stocks with high abnormal news. To examine these predictions, we run the following regression

$$
\begin{aligned}
Ret_{i,t+1} = \alpha + & \beta_1 Sent_{i,t+1} + \beta_2 \mathbb{1}_{\text{Sent}^-} + \beta_3 Sent_{i,t+1} \times \mathbb{1}_{\text{Sent}^-} \\
& + \beta_4 Sent_{i,t+1} \times \mathbb{1}_{\text{Sent}^-} \times \mathbb{1}_{\text{High news}} + \beta_5 \mathbb{1}_{\text{Sent}^-} \times \mathbb{1}_{\text{High news}} \\
& + \beta_6 Sent_{i,t+1} \times \mathbb{1}_{\text{High news}} + \beta_7 \mathbb{1}_{\text{High news}} + \beta_8 Ret^M_{t+1} \\
& + \beta_9 IndFE_i + \beta_{10} TimeFE_t + \Gamma' Controls + \epsilon_{t+1},
\end{aligned}
$$

where $Ret$ and $Ret^M$ corresponds to the stock $i$'s excess return and the market excess return over the risk-free rate at month $t+1$, respectively. $Sent$ corresponds to the average news sentiment in Ravenpack in month $t+1$ minus its six-month rolling average in sentiment for stock $i$. $\mathbb{1}_{\text{Sent}^-}$ is an indicator variable equal to one if $Sent < 0$, zero otherwise. $\mathbb{1}_{\text{High news}}$ is an indicator variable equal to one if the firm $i$ coverage on month $t+1$ exceeds expected coverage, zero otherwise. The additional control variables are the Fama-French HML and SMB factors on month $t+1$ and two-month lags of the excess stock return and excess market returns. The regression analysis includes only stocks with reported news in Ravenpack on month $t+1$.

Table 2 reports the regression results. Column (1) reports a positive and statistically significant (at the 1% level) relationship between returns and sentiment. A one unit increase in sentiment increases return by 0.066. The results reported in columns (2) and (3) support our theoretical predictions that stock returns are more responsive to news when the news is negative. The positive and statistically significant loadings (at the 1% level) for $Sent_{t+1} \times \mathbb{1}_{\text{Sent}^-}$ of 0.037 and 0.039 in columns (2) and (3), respectively, indicate that stock returns are approximately 81% (e.g., 0.039/0.048) more sensitive to negative than positive news. The results reported in columns (5)-(6) indicate that the increase in sensitivity to negative news is generally only the case for stocks with higher than expected news. While the loadings for $Sent_{t+1} \times \mathbb{1}_{\text{Sent}^-} \times \mathbb{1}_{\text{High news}}$ in columns (3) and (4) are positive and statistically significant, the loadings for $Sent_{t+1} \times \mathbb{1}_{\text{Sent}^-}$ are approximately zero and not statistically significant.

A key implication of our model is that stocks with higher than expected news coverage have higher uncertainty. Consistent with the literature associating higher uncertainty with higher expected returns (e.g., Zhang, 2006; Ang et al., 2006), stocks with higher than expected news coverage should have higher expected returns. To examine this premise, we run the following regression

$$Ret_{i,t+2} = \alpha + \beta_1 \mathbb{E}_t[\text{News qnt}]_{i,t+1} + \beta_2 \mathbb{1}_{\text{High news}} + \beta_3 \mathbb{E}_t[\text{News qnt}]_{i,t+1} \times \mathbb{1}_{\text{High news}}$$
$$+ \beta_4 Ret^M_{t+2} + \beta_5 IndFE_i + \beta_6 TimeFE_t + \Gamma' Controls + \epsilon_{t+1},$$

where $Ret_{t+2}$ and $Ret^M_{t+2}$ correspond to the stock $i$'s excess return and the market excess return over the risk-free rate at month $t+2$, respectively. $\mathbb{E}_t[\text{News qnt}]$ corresponds to the expected news quintile at $t+1$ constructed from the quintile sort of the fitted values of regression (22) estimated at each time $t$ (excluding the time fixed-effects). $\mathbb{1}_{\text{High news}}$ is an indicator variable equal to one if the firm $i$ coverage on month $t+1$ exceeds expected coverage, zero otherwise. We define stocks with high (low) unexpected news when the number of news articles in month $t+1$ is above (below) expected news computed on month $t$. The additional control variables are the Fama-French HML and SMB factors on month $t+2$ and two-month lags of the excess stock return and excess market returns. We report the results in Table 3.

Consistent with Fang & Peress (2009), we find that stocks with higher expected media coverage have lower expected returns. A unit increase in the quintile of expected news coverage, $\mathbb{E}_t[\text{News qnt}]$, is associated with a decrease in expected returns of approximately 14 to 17 basis points depending on model specifications. Stocks with higher than expected news coverage, $\mathbb{1}_{\text{High news}}$, columns (2)-(3) report an increase in expected returns of 24 to 27 bps. Finally, the interacting term coefficients $\mathbb{E}_t$[News qnt] $\times$ $\mathbb{1}_{\text{High news}}$ reported in columns (4)-(5) indicate that stocks in top quintile of expected news coverage with higher than expected news coverage is associated with negative stock returns (e.g., -(0.0017+0.008)×5+0.0044=-0.0081). The effect of unexpected news coverage for stocks belonging to the bottom expected news quintile is associated with positive returns (e.g., -(0.0017+0.008)+0.0044=0.0019). The loadings on the interacting terms confirm the pattern shown in Figure 4; stocks with low expected coverage and with greater unexpected coverage are more volatile and is associated with higher expected returns. Overall, in contrast to Fang & Peress (2009), we find that more media coverage can be associated with higher expected returns.

## 8 Conclusion

This paper builds a theoretical framework to endogeneize the editorial decisions of media and analyze their asset pricing implications. The decision to publish a story about a particular firm does not only provide information to investors about the firm selected for publication (which is the focus of the literature), but also conveys information about non-reported firms. Specifically, the investor is able to distinguish the risk regime of non-reported firms with high expected news coverage from those with low expected news coverage. As a consequence, the decision to select a firm to be reported in a media outlet has asset pricing implications for reported firms, non-reported firms with high expected news coverage and non-reported firms with low expected news coverage. Failing to capture the information implications for all types of firms may lead the econometrician to estimate a misspecified asset pricing model.

Empirically, we show that firms which receive lower than expected media coverage in a given month (typically having higher media coverage) have abnormally low stock price volatility and turnover, i.e., less uncertainty. In contrast, a firm with typically low media coverage which receives lower than expected media coverage in a given month has its abnormal stock price volatility and turnover unaffected. Additionally, we also find that stock returns are more responsive to negative news sentiment and more so for stocks with high abnormal news. These findings are in line with the theoretical predictions.

## References


Admati, A. R. & Pfleiderer, P. (1986). A monopolistic market for information. *Journal of Economic Theory*, 39(2), 400 – 438.

Admati, A. R. & Pfleiderer, P. (1987). Viable allocations of information in financial markets. *Journal of Economic Theory*, 43(1), 76 – 115.

Amador, M. & Weill, P.-O. (2012). Learning from private and public observations of others actions. *Journal of Economic Theory*, 147(3), 910–940.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1), 259–299.

Banerjee, S., Davis, J., & Gondhi, N. (2018). When Transparency Improves, Must Prices Reflect Fundamentals Better? *The Review of Financial Studies*, 31(6), 2377–2414.

Barinov, A. (2014). Turnover: Liquidity or uncertainty? *Management Science*, 60(10), 2478–2495.

Call, A. C., Emett, S. A., Maksymov, E., & Sharp, N. Y. (2022). Meet the press: Survey evidence on financial journalists as information intermediaries. *Journal of Accounting and Economics*, 73(2), 101455.

Calvet, L. E. & Fisher, A. J. (2007). Multifrequency news and stock returns. *Journal of Financial Economics*, 86(1), 178–212.

Campbell, J. Y. & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, 31(3), 281–318.

Chan, W. S. (2003). Stock price reaction to news and no-news: Drift and reversal after headlines. *Journal of Financial Economics*, 70(2), 223–260.

Dougal, C., Engelberg, J., García, D., & Parsons, C. A. (2012). Journalists and the Stock Market. *The Review of Financial Studies*, 25(3), 639–679.

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

Fisher, A. J., Martineau, C., & Sheng, J. (2022). Macroeconomic attention and announcement risk premia, forthcoming. *The Review of Financial Studies*.

Gentzkow, M. & Shapiro, J. M. (2006). Media bias and reputation. *Journal of Political Economy*, 114(2), 280–316.

Gao, P. & Liang, P. J. (2013). Informational feedback, adverse selection, and optimal disclosure policy. *Journal of Accounting Research*, 51(5), 1133–1158.

Goldman, E., Martel, J., & Schneemeier, J. (2022). A theory of financial media. *Journal of Financial Economics*, 145, 239–258.

Goldstein, I. & Yang, L. (2019). Good disclosure, bad disclosure. *Journal of Financial Economics*, 131(1), 118–138.

Grossman, S. J. & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3), 393–408.

Han, B., Tang, Y., & Yang, L. (2016). Public information and uninformed trading: Implications for market liquidity and price efficiency. *Journal of Economic Theory*, 163, 604 – 643.

Hellwig, M. F. (1980). On the aggregation of information in competitive markets. *Journal of Economic Theory*, 22(3), 477 – 498.

Huberman, G. & Regev, T. (2001). Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *The Journal of Finance*, 56(1), 387–396.

Morris, S. & Shin, H. S. (2002). Social value of public information. *American Economic Review*, 92(5), 1521–1534.

Mullainathan, S. & Shleifer, A. (2005). The market for news. *American Economic Review*, 95(4), 1031–1053.

Nimark, K. P. (2014). Man-bites-dog business cycles. *American Economic Review*, 104(8), 2320–67.

Nimark, K. P. & Pitschner, S. (2019). News media and delegated information choice. *Journal of Economic Theory*, 181, 160–196.

Perego, J. & Yuksel, S. (2018). Media competition and social disagreement. *Working Paper*.

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

Schwenkler, G. & Zheng, H. (2022). Why does news coverage predict returns? Evidence from the underlying editor preferences for risky stocks. *Working Paper*.

Tetlock, P. C. (2010). Does public financial news resolve asymmetric information? *The Review of Financial Studies*, 23(9), 3520–3557.

Veronesi, P. (1999). Stock market overreaction to bad news in good times: A rational expectations equilibrium model. *The Review of Financial Studies*, 12(5), 975–1007.

Verrecchia, R. E. (1982). Information acquisition in a noisy rational expectations economy. *Econometrica*, 50(6), 1415–1430.

Zhang, X. F. (2006). Information uncertainty and stock returns. *The Journal of Finance*, 61(1), 105–137.

Figure 3: Turnover and Volatility Conditioned on Unexpected News

This figure shows the average abnormal turnover (in %) and abnormal volatility on month $t+1$ by the quintile of unexpected news coverage on month $t+1$. The unexpected news coverage is the difference between the number of news articles in RavenPack on month $t+1$ minus the expected level of news coverage calculated as the fitted values from Equation (22) on month $t$. We compute the monthly abnormal turnover (volatility) as the difference between turnover (volatility) and its six-month rolling average. The error bars correspond to the 95% confidence intervals.

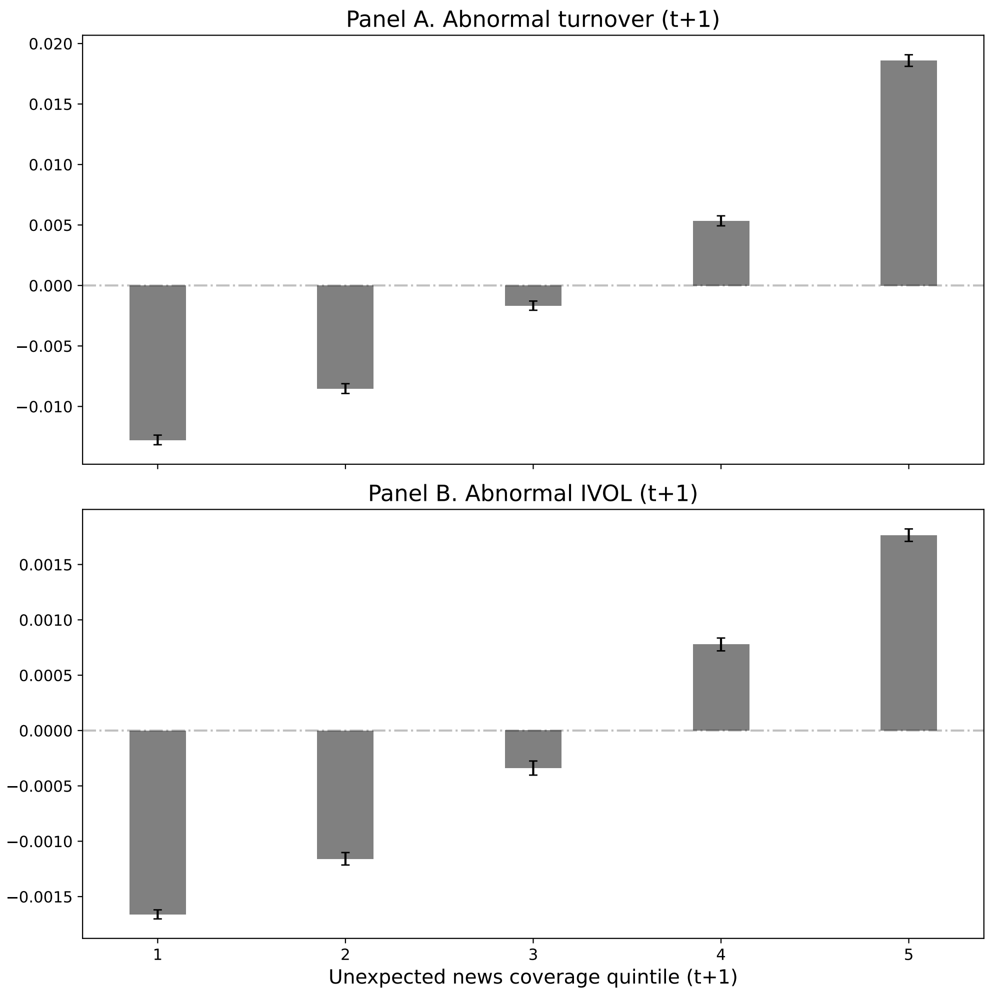

Figure 4: Turnover and Volatility Conditioned on Expected and Unexpected News

This figure shows the average abnormal turnover (in %), and abnormal volatility on month $t+1$ for stocks with high and low unexpected news by the quintile of expected news coverage on month $t+1$. The unexpected news coverage is the difference between the number of news articles in RavenPack on month $t+1$ minus the expected level of news coverage calculated as the fitted values from Equation (22) on month $t$. We define stocks with high (low) unexpected news when the number of news articles in month $t+1$ is above (below) expected news computed on month $t$. We compute the monthly abnormal turnover (volatility) as the difference between turnover (volatility) and its six-month rolling average. The error bars correspond to the 95% confidence intervals.

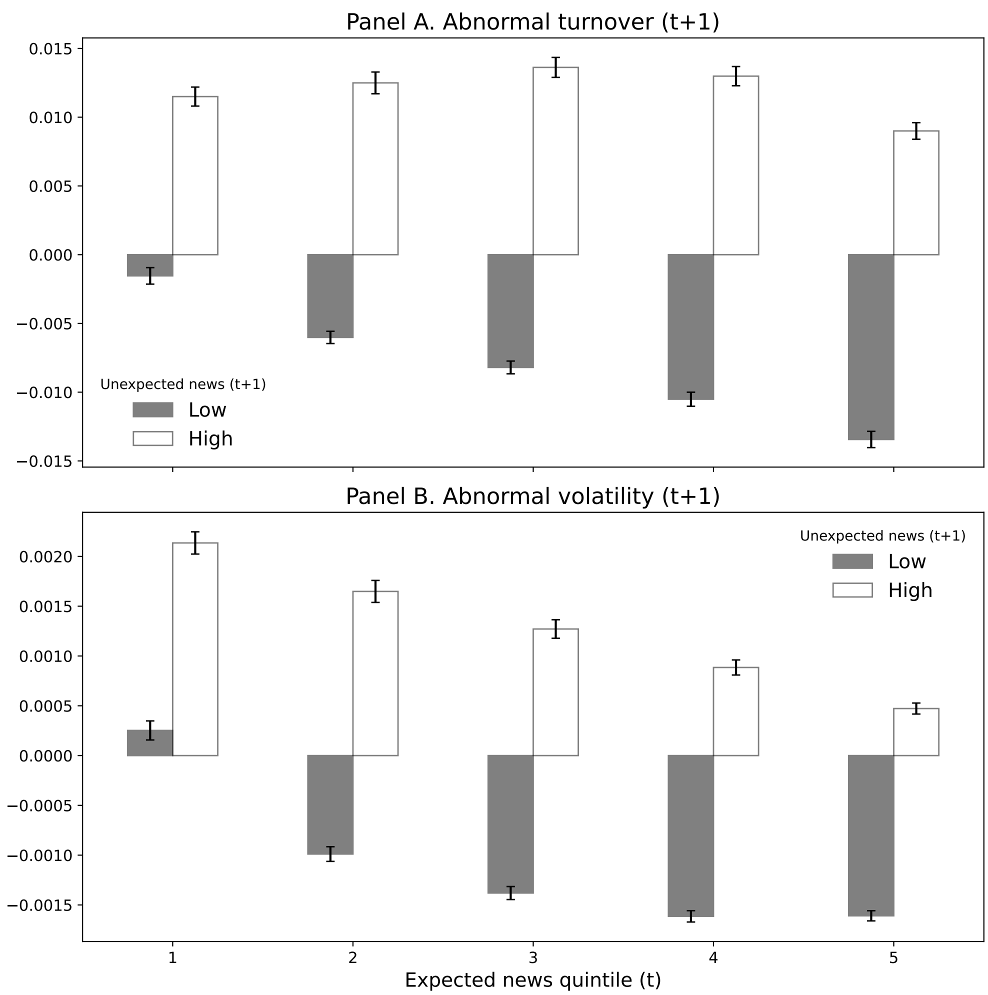

Table 1: News and Firm Characteristics

This table reports the coefficients of the following regressions

$$\begin{aligned} Ln\ article_{i,t} = \alpha + & \beta_1 Ln\ MCAP_{i,t} + \beta_2 \mathbb{1}_{\text{EA},i,t} + \beta_3 Analyst_{i,t} + \beta_4 Turnover_{i,t} + \beta_5 IO_{i,t} \\ & + \beta_6 (Ret_{i,t} - Ret^M_t) + \beta_7 IVOL_{i,t} + \beta_8 Age_{i,t} + \beta_9 IndFE_i + \beta_{10} TimeFE_t + \epsilon_{i,t}, \end{aligned}$$

where *Ln article* corresponds to the natural logarithm of the 1+total number of RavenPack editorial articles in month $t$ for stock $i$, $Ln\ MCAP$ is the natural logarithm of firm market capitalization at the end of month $t$, $\mathbb{1}_{\text{EA}}$ is an indicator variable equal to one if stock $i$ has an earnings announcement on month $t$ and zero otherwise, *Analyst* is the number of analyst following from I/B/E/S, $Turnover$ is the monthly share turnover, $IO$ is the fraction of shares held by institutions in month $t$ during the quarter of the respective month, $(Ret - Ret^M)$ is the stock return minus the market return, $IVOL$ is the stock's idiosyncractic volatility, and *Age* is the number of years since appearance in CRSP. $IndFE$ and $TimeFE$ are the industry (GIC 2-digit sector code) and year-month fixed effects, respectively. *Analyst*, $(Ret - Ret^M)$, $IVOL$, and $Turnover$ are rescaled to have a standard deviation of one. The standard errors are clustered at the industry level and year-month. $^{***}$, $^{**}$, and $^{*}$ denotes statistical significance at the 1, 5, and 10% levels. The sample period is from January 2000 to December 2021.

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Ln MCAP | 0.168*** | 0.140*** | 0.141*** | 0.153*** | 0.156*** |
| | (0.003) | (0.003) | (0.003) | (0.003) | (0.004) |
| $\mathbb{1}_{\text{EA}}$ | 0.220*** | 0.219*** | 0.209*** | 0.200*** | 0.202*** |
| | (0.011) | (0.011) | (0.011) | (0.011) | (0.011) |
| Analyst | | 0.098*** | 0.074*** | 0.070*** | 0.073*** |
| | | (0.004) | (0.004) | (0.004) | (0.004) |
| Turnover | | | 0.088*** | 0.068*** | 0.065*** |
| | | | (0.003) | (0.003) | (0.003) |
| IO | | | -0.005 | 0.004 | -0.005 |
| | | | (0.004) | (0.004) | (0.004) |
| Ret-Ret$^M$ | | | | -0.008*** | -0.007*** |
| | | | | (0.002) | (0.002) |
| IVOL | | | | 0.064*** | 0.058*** |
| | | | | (0.003) | (0.003) |
| Age | | | | 0.002*** | 0.002*** |
| | | | | (0.000) | (0.000) |
| $R^2$ | 0.197 | 0.205 | 0.215 | 0.219 | 0.218 |
| $N$ | 1,063,799 | 1,063,799 | 1,063,799 | 1,057,902 | 1,057,902 |
| Industry F.E. | N | N | N | N | Y |
| Year-Month F.E. | Y | Y | Y | Y | Y |

Table 2: Stock Returns' Asymmetric Response to Positive and Negative News

This table reports the coefficients of the following regression

$$\begin{aligned} Ret_{i,t+1} = \alpha + &\beta_1 Sent_{i,t+1} + \beta_2 \mathbb{1}_{\text{Sent}^-} + \beta_3 Sent_{i,t+1} \times \mathbb{1}_{\text{Sent}^-} \\ &+\beta_4 Sent_{i,t+1} \times \mathbb{1}_{\text{Sent}^-} \times \mathbb{1}_{\text{High news}} + \beta_5 \mathbb{1}_{\text{Sent}^-} \times \mathbb{1}_{\text{High news}} \\ &+\beta_6 Sent_{i,t+1} \times \mathbb{1}_{\text{High news}} + \beta_7 \mathbb{1}_{\text{High news}} + \beta_8 Ret^M_{t+1} \\ &+\beta_9 IndFE_i + \beta_{10} TimeFE_t + \Gamma' Controls + \epsilon_{t+1}, \end{aligned}$$

where $Ret$ and $Ret^M$ correspond to the stock $i$'s excess return and the market excess return over the risk-free rate at month $t+1$, respectively. $Sent$ corresponds to the average news sentiment in Ravenpack in month $t+1$ minus a six-month rolling average in sentiment. $\mathbb{1}_{\text{Sent}^-}$ is an indicator variable equal to one if $Sent < 0$, zero otherwise. $\mathbb{1}_{\text{High news}}$ is an indicator variable equal to one if the firm $i$ coverage on month $t+1$ exceeds expected coverage, zero otherwise. We define stocks with high (low) unexpected news when the number of news articles in month $t+1$ is above (below) expected news computed on month $t$. $IndFE$ and $TimeFE$ are the industry (GIC 2-digit sector code) and year-month fixed effects, respectively. The additional control variables are the Fama-French HML and SMB factors on month $t+1$, and two-month lags of the excess stock return and excess market returns. The sample period is from January 2000 to December 2021 and includes observations for stocks with reported news in Ravenpack on month $t+1$. The standard errors are clustered at the industry level and year-month. ***, **, and * denotes statistical significance at the 1, 5, and 10% levels.

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Sent | 0.0659*** | 0.0493*** | 0.0478*** | 0.0430*** | 0.0416*** |
| | (0.010) | (0.007) | (0.007) | (0.005) | (0.004) |
| Sent$\times \mathbb{1}_{Sent^-}$ | | 0.0370*** | 0.0391*** | -0.0025 | 0.0002 |
| | | (0.008) | (0.009) | (0.006) | (0.006) |
| $\mathbb{1}_{\text{Sent}^-}$ | | 0.0008 | 0.0011 | 0.0022*** | 0.0023*** |
| | | (0.002) | (0.002) | (0.001) | (0.001) |
| Sent$\times \mathbb{1}_{Sent^-} \times \mathbb{1}_{\text{High news}}$ | | | | 0.0366*** | 0.0357*** |
| | | | | (0.011) | (0.010) |
| $\mathbb{1}_{Sent^-} \times \mathbb{1}_{\text{High news}}$ | | | | -0.0014 | -0.0012 |
| | | | | (0.001) | (0.001) |
| Sent$\times \mathbb{1}_{\text{High news}}$ | | | | 0.0176*** | 0.0174*** |
| | | | | (0.007) | (0.007) |
| $\mathbb{1}_{\text{High news}}$ | | | | 0.0122*** | 0.0122*** |
| | | | | (0.003) | (0.003) |
| Ret$^M$ | 0.2093 | 0.2083 | 0.1978 | 0.2058 | 0.1951 |
| | (0.192) | (0.193) | (0.208) | (0.194) | (0.209) |
| Intercept | 0.0112*** | 0.0128*** | 0.0096 | 0.0021 | -0.0031 |
| | (0.001) | (0.001) | (0.067) | (0.002) | (0.066) |
| $R^2$ | 0.004 | 0.004 | 0.004 | 0.004 | 0.005 |
| $N$ | 457,892 | 457,892 | 450,437 | 457,892 | 450,437 |
| Controls | N | N | Y | N | Y |
| Industry F.E. | Y | Y | Y | Y | Y |
| Year-Month F.E. | Y | Y | Y | Y | Y |

Table 3: Expected and Unexpected Coverage and Stock Returns

This table reports the coefficients of the following regression

$$Ret_{i,t+2} = \alpha + \beta_1 \mathbb{E}_t[\text{News qnt}]_{i,t+1} + \beta_2 \mathbb{1}_{\text{High news}} + \beta_3 \mathbb{E}_t[\text{News qnt}]_{i,t+1} \times \mathbb{1}_{\text{High news}}$$
$$+\beta_4 Ret^M_{t+2} + \beta_5 IndFE_i + \beta_6 TimeFE_t + \ \Gamma' Controls + \epsilon_{t+1},$$

where $Ret$ and $Ret^M$ correspond to the stock $i$'s excess return and the market excess return over the risk-free rate at month $t+2$, respectively. $\mathbb{E}_t$[News qnt] corresponds to the expected news quintile at $t+1$ constructed from the quintile sort of the fitted values of regression (22) estimated at each time $t$ (excluding the time fixed-effects). $\mathbb{1}_{\text{High news}}$ is an indicator variable equal to one if the firm $i$ coverage on month $t+1$ exceeds expected coverage, zero otherwise. We define stocks with high (low) unexpected news when the number of news articles in month $t+1$ is above (below) expected news computed on month $t$. $IndFE$ and $TimeFE$ are the industry (GIC 2-digit sector code) and year-month fixed effects, respectively. The additional control variables are the Fama-French HML and SMB factors on month $t+2$, and two-month lags of the excess stock return and excess market returns. The standard errors are clustered at the industry level and year-month. ***, **, and * denotes statistical significance at the 1, 5, and 10% levels. The sample period is from January 2000 to December 2021.

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $\mathbb{E}$[News qnt] | -0.0016** | | -0.0017** | -0.0014* | -0.0017** |
| | (0.001) | | (0.001) | (0.001) | (0.001) |
| $\mathbb{1}_{\text{High news}}$ | | 0.0024*** | 0.0027*** | 0.0041*** | 0.0044*** |
| | | (0.001) | (0.001) | (0.001) | (0.001) |
| $\mathbb{E}$[News qnt] $\times$ $\mathbb{1}_{\text{High news}}$ | | | | -0.0007** | -0.0008*** |
| | | | | (0.000) | (0.000) |
| Ret$^M$ | 0.3737** | 0.3742** | 0.3728* | 0.3732** | 0.2987 |
| | (0.190) | (0.190) | (0.190) | (0.190) | (0.197) |
| Intercept | 0.0103*** | 0.0060*** | 0.0092*** | 0.0086*** | 0.0096*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| $R^2$ | 0.000 | 0.000 | 0.000 | 0.000 | 0.002 |
| $N$ | 1,031,174 | 1,031,174 | 1,031,174 | 1,031,174 | 991,645 |
| Controls | N | N | N | N | Y |
| Industry F.E. | Y | Y | Y | Y | Y |
| Year-Month F.E. | Y | Y | Y | Y | Y |

# A Proofs

## A.1 Proof of Lemma 1

Before proceeding to the proof of Result 1, we first show that the profit of the media outlet can be written as in (12). To do so, the first step requires the following calculation:

$$\begin{aligned}
&E_0\{E_1[D_{ni}(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n] - \frac{\gamma}{2}V_1[D_{ni}(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n] \mid \tilde{\rho}_n\} \\
&= E_0\{D_{ni}E_1[(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n] - \frac{\gamma}{2}D_{ni}^2 V_1[\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n] \mid \tilde{\rho}_n\} \\
&= E_0\left\{\frac{E_1[(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n]^2}{\gamma V_1[\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n]} - \frac{E_1[(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n]^2}{2\gamma V_1[\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n]} \mid \tilde{\rho}_n\right\} \\
&= E_0\left\{\frac{E_1[(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n]^2}{2\gamma V_1[\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n]} \mid \tilde{\rho}_n\right\} \\
&= E_0\left\{E_1[(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n]^2 \mid \tilde{\rho}_n\right\}\frac{1}{2\gamma V_1[\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n]} \\
&= \left[V_0\left\{E_1[(\tilde{v}_n - p_n) \mid \tilde{y}_n, \tilde{\rho}_n, p_n] \mid \tilde{\rho}_n\right\} + \left(E_0\left\{(\tilde{v}_n - p_n) \mid \tilde{\rho}_n\right\}\right)^2\right]\frac{1}{2\gamma V_1[\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n]} \qquad \text{(A.1)} \\
&= \left[V_0\left\{(\tilde{v}_n - p_n) \mid \tilde{\rho}_n\right\} - V_1\{\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n\}\right]\frac{1}{2\gamma V_1[\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n]} \\
&= \frac{V_0[(\tilde{v}_n - p_n) \mid \tilde{\rho}_n]}{2\gamma V_1[\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n]} - \frac{1}{2\gamma}, \qquad \text{(A.2)}
\end{aligned}$$

where the first equality follows from the fact that given $(\tilde{y}_n, \tilde{\rho}_n, p_n)$, $D_{ni}$ and $p_n$ are constant, the second follows from (7), the fourth follows from the fact that $V_1[\tilde{v}_n \mid \tilde{y}_n, \tilde{\rho}_n, p_n]$ is not a function of $\tilde{y}_n$ and $p_n$, the fifth follows from the definition of variance, the sixth one follows from the law of total variance and fact that $E_0\left\{(\tilde{v}_n - p_n) \mid \tilde{\rho}_n\right\} = 0$.

Similar calculations show that

$$E_0\{E_1[D_{ni}(\tilde{v}_n - p_n) \mid \tilde{\rho}_n, p_n] - \frac{\gamma}{2}V_1[D_{ni}(\tilde{v}_n - p_n) \mid \tilde{\rho}_n, p_n] \mid \tilde{\rho}_n\} = \frac{V_0[(\tilde{v}_n - p_n) \mid \tilde{\rho}_n]}{2\gamma V_1[\tilde{v}_n \mid \tilde{\rho}_n, p_n]} - \frac{1}{2\gamma}. \qquad \text{(A.3)}$$

Combining (A.2) and (A.3) yields

$$Profit_n(\tilde{\rho}_n) = \frac{1}{2\gamma}V_0[\tilde{v}_n - p_n \mid \tilde{\rho}_n]\left(\frac{1}{V_1[\tilde{v}_n \mid \tilde{\rho}_n, p_n, \tilde{y}_n]} - \frac{1}{V_1[\tilde{v}_n \mid \tilde{\rho}_n, p_n]}\right).$$

We now need to calculate these conditional variances. First, given that only risk-regime $\tilde{\rho}_n$ is observed the conditional variance of $\tilde{v}_n - p_n$, where $p_n$ is given by (8), can be written as follows:

$$\begin{aligned}
V(\tilde{v}_n - p_n \mid \tilde{\rho}_n) &= V(\tilde{\rho}_n\tilde{\delta}_n - a_y\tilde{\delta}_n - a_y\tilde{\eta}_n - a_z\tilde{z}_n \mid \tilde{\rho}_n) \\
&= (\tilde{\rho}_n - a_y)^2\tau_\delta^{-1} + a_y^2\tau_\eta^{-1} + a_z^2\tau_z^{-1} \\
&= \frac{\tau_\delta\tilde{\rho}_n^2}{(\tau_\delta + \tau_\eta)^2} + \frac{\tau_\eta\tilde{\rho}_n^2}{(\tau_\delta + \tau_\eta)^2} + \frac{\gamma^2\tilde{\rho}_n^4\tau_z^{-1}}{(\tau_\delta + \tau_\eta)^2} = \frac{\tilde{\rho}_n^2(\tau_\delta + \tau_\eta + \gamma^2\tilde{\rho}_n^2\tau_z^{-1})}{(\tau_\delta + \tau_\eta)^2}. \qquad \text{(A.4)}
\end{aligned}$$

Second, the conditional variance of $\tilde{v}_n$ when both $p_n$ and $\tilde{y}_n$ are observed as well as $\tilde{\rho}_n$ can be written as

$$V(\tilde{v}_n \mid \tilde{\rho}_n, p_n, \tilde{y}_n) = \frac{\tilde{\rho}_n^2}{\tau_\delta + \tau_\eta}. \tag{A.5}$$

Third, the conditional variance of $\tilde{v}_n$ when only $p_n$ and $\tilde{\rho}_n$ are observed can be written as

$$\begin{aligned} V(\tilde{v}_n \mid \tilde{\rho}_n, p_n) &= V(\bar{\delta} + \tilde{\rho}_n \tilde{\delta}_n \mid \tilde{\rho}_n, p_n) = \tilde{\rho}_n^2 V(\tilde{\delta}_n \mid \tilde{\rho}_n, p_n) \\ &= \tilde{\rho}_n^2 \left( \tau_\delta^{-1} - \frac{a_y^2 \tau_\delta^{-2}}{a_y^2(\tau_\delta^{-1} + \tau_\eta^{-1}) + a_z^2 \tau_z^{-1}} \right) \\ &= \frac{\tilde{\rho}_n^2 (\gamma^2 \tilde{\rho}_n^2 + \tau_\eta \tau_z)}{\gamma^2 \tau_\delta \tilde{\rho}_n^2 + \tau_\eta^2 \tau_z + \tau_\delta \tau_\eta \tau_z}. \end{aligned} \tag{A.6}$$

Taken together (A.4)-(A.6), the profit of the media outlet can be written as follows:

$$Profit_n(\tilde{\rho}_n) = \frac{\gamma \tau_\eta \tilde{\rho}_n^2 (\gamma^2 \tilde{\rho}_n^2 + \tau_z(\tau_\delta + \tau_\eta))}{2\tau_z(\tau_\delta + \tau_\eta)^2 (\gamma^2 \tilde{\rho}_n^2 + \tau_\eta \tau_z)}.$$

We are now ready to show that the profit is increasing in $\tilde{\rho}_n$. Taking the derivative of $Profit_n(\tilde{\rho}_n)$ with respect to $\tilde{\rho}_n$ yields

$$\frac{dProfit_n(\tilde{\rho}_n)}{d\tilde{\rho}_n} = \frac{\gamma \tau_\eta \tilde{\rho}_n (\gamma^4 \tilde{\rho}_n^4 + 2\gamma^2 \tau_\eta \tau_z \tilde{\rho}_n^2 + \tau_\eta^2 \tau_z^2 + \tau_\delta \tau_\eta \tau_z^2)}{\tau_z (\tau_\delta + \tau_\eta)^2 (\gamma^2 \tilde{\rho}_n^2 + \tau_\eta \tau_z)^2} > 0,$$

which is positive since both the numerator and denominator are positive. This is because $\gamma, \tau_\eta, \tau_\delta, \tau_z$, and $\tilde{\rho}_n$ are positive. Therefore, $Profit_n(\tilde{\rho}_n)$ is increasing in $\tilde{\rho}_n$.

### A.2 Proof of Lemma 2

Given that i) $\rho_{h,\check{n}} > \rho_{l,\check{n}}$ by assumption; ii) $\rho_{l,\check{n}} > \rho_{h,\hat{n}}$ by definition 2; and iii) all firms are ranked by $\rho_{h,n}$ in descending order by Result 1, then we have $\rho_{h,\check{n}} > \rho_{l,\check{n}} > \rho_{h,\hat{n}} \geq \rho_{h,n}$, $\forall n \geq \hat{n}$. Since the profit of the media outlet is increasing in $\rho_n$, the media outlet will always prefer to publish a story by firm $\check{n}$ than publishing a news story about firm $n$, $\forall n \geq \hat{n}$.

### A.3 Proof of Lemma 3

For any firm $n'$ such that $n' < \hat{n}$, we have that $max\{\rho_{l,1}\}_{n \neq \check{n}} \leq \rho_{l,\check{n}} \leq \rho_{h,n'}$ where the first inequality follows from definition 1 and the second inequality follows from definition 2. Consider now the following scenario: a firm $n'$ such that $n' < \hat{n}$ is in a high volatility regime $\rho_{h,n'}$, while all the other firms are in the low volatility regime $\rho_{l,n}$ for $n \neq n'$. This scenario may happen with a positive probability $\mathbb{P}(\tilde{\rho}_{n'} = \rho_{h,n'}) \prod_{n \neq n'} \mathbb{P}(\tilde{\rho}_n = \rho_{l,n}) > 0$,. Since profits are increasing in $\tilde{\rho}_n$ and we have that $max\{\rho_{l,1}\}_{n \neq \check{n}} \leq \rho_{l,\check{n}} \leq \rho_{h,n'}$, then firm $n'$ would be the firm selected for publication in this scenario with positive probability.

### A.4 Proof of Proposition 1

For the first part, if the media outlet publishes a signal $y_{n'}$ about firm $n'$ when $\tilde{\rho}_{n'} = \rho_{h,n'}$, then asset prices are given by (8) with $\tilde{\rho}_{n'} = \rho_{h,n'}$.

For the second and third parts, if the media outlet publishes a public signal about $n'$ when $\tilde{\rho}_{n'} = \rho_{h,n'}$, then it must be the case that i) for any firm $n$ such that $n < n'$, the risk-regime factor is $\rho_{l,n}$, and ii) for any firm $n$ such that $n > n'$, the risk-regime is unknown. Hence, any firm $n$ such that $n < n'$ is in a low volatility risk-regime $\rho_{l,n}$ with no public signal and asset prices are given by (10) with $\tilde{\rho}_n = \rho_{l,n}$, and any firm $n$ such that $n > n'$ is in an unknown risk-regime and asset prices are given by (6).

### A.5 Proof of Corollary 1

For the first part, if the media outlet publishes a signal $y_{\check{n}}$ about firm $\check{n}$ when $\tilde{\rho}_{\check{n}} = \rho_{l,\check{n}}$, then it immediately follows that firm $\check{n}$ is in a low volatility risk-regime $\rho_{l,\check{n}}$ and asset prices are given by (8) with $\tilde{\rho}_{\check{n}} = \rho_{l,\check{n}}$.

For the second and third parts, if the media outlet publishes a public signal about $\check{n}$ when $\tilde{\rho}_{\check{n}} = \rho_{l,\check{n}}$, then it must be the case that i) for any firm $n$ such that $n < \hat{n}$, the risk-regime factor is $\rho_{l,n}$, and ii) for any firm $n$ such that $n \geq \hat{n}$, the risk-regime is unknown. Hence, any firm $n$ such that $n < \hat{n}$ is in a low volatility risk-regime $\rho_{l,n}$ with no public signal and asset prices are given by (10) with $\tilde{\rho}_n = \rho_{l,n}$, and any firm $n$ such that $n \geq \hat{n}$ is in an unknown risk-regime and asset prices are given by (6).

### A.6 Proof of Corollary 2

Let's fix the realizations of the following random variables for any firm $n'$: $\tilde{\delta}_{n'} = \delta_{n'}, \tilde{z}_{n'} = z_{n'}$, and $\tilde{\rho}_{n'} = \rho_{l,n'}$. Consider the following two scenarios. In the first scenario, suppose that any firm $n$ such that $n \leq n'$ is in a low volatility risk-regime and the media outlet publishes a story about firm $n^*$, where $n^* > n'$. In this scenario, the asset price for firm $n'$ will be given by (10) with $\tilde{\rho}_{n'} = \rho_{l,n'}$. In the second scenario, the realizations for firm $n'$ are exactly the same as the first scenario, but suppose that the media publishes a story about firm $n^*$, where $n^* < n'$. In this scenario, the asset price for firm $n'$ will be given by (6). In these two scenarios, firm $n'$ has different asset prices although the realizations of cash flows and noisy supply for firm $n'$ are exactly the same, which completes the proof.

### A.7 Proof of Proposition 2

We interpret positive news as increases in $\tilde{y}_n$ and negative news as decreases in $\tilde{y}_n$. A news story has two effects on expected prices: i) the risk-regime is high $\tilde{\rho}_n = \rho_{h,n}$ and ii) investors receive a signal $\tilde{y}_n$. The expected price is given by

$$E(p_n \mid \tilde{\rho}_n, \tilde{y}_n) = \bar{\delta} + a_y \tilde{y}_n + a_z \bar{z}.$$

Effect i) has a negative effect on the expected price when $\bar{z} > 0$:

$$\frac{\partial E(p_n \mid \tilde{\rho}_n, \tilde{y}_n)}{\partial \tilde{\rho}_n} = \frac{1}{\tau_\delta + \tau_\eta}[-2\gamma \bar{z} \tilde{\rho}_n] < 0.$$

Effect ii) has a positive effect on the expected price for increases on $\tilde{y}_n$:

$$\frac{\partial E(p_n \mid \tilde{\rho}_n, \tilde{y}_n)}{\partial \tilde{y}_n} = a_y > 0.$$

Hence, effects i) and ii) go in opposite directions when news are positive and go in the same direction when news are negative.

## A.8 Proof of Lemma 4

If $\tilde{\rho}_n$ is known, the information acquisition problem becomes:

$$\max_{\tau_{\varepsilon ni}} \frac{1}{2\gamma} \frac{V_0[\tilde{v}_n - p_n]}{V_1[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}]} - C(\tau_{\varepsilon ni}).$$

Hence, $\tau_{\varepsilon n}$ is implicitly given by

$$\tau_{\varepsilon n} = \frac{1}{2\gamma} \frac{V_0[\tilde{v}_n - p_n]}{\tilde{\rho}_n^2}. \tag{A.7}$$

The information acquisition $\tau_{\varepsilon ni}$ is the same in both scenarios when the public signal is observed and when the investors chooses to ignore the public signal. Thus, the cost of acquiring information is the same in both cases and the profit function of the media outlet for any firm $n$ can be written as

$$\begin{aligned} Profit_n(\tilde{\rho}_n) &= \frac{1}{2\gamma} V[\tilde{v}_n - p_n \mid \tilde{\rho}_n] \left( \frac{1}{V[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{y}_n, \tilde{s}_{ni}, \tilde{s}_{pn}]} - \frac{1}{V[\tilde{v}_n \mid \tilde{\rho}_n, \tilde{s}_{ni}, \tilde{s}_{pn}]} \right), \\ &= \frac{1}{2\gamma} \frac{V[\tilde{v}_n - p_n \mid \tilde{\rho}_n] \tau_\eta}{\tilde{\rho}_n^2} = \tau_{\varepsilon n} \tau_\eta. \end{aligned} \tag{A.8}$$

Hence, $Profit_n(\tilde{\rho}_n)$ is increasing in $\tilde{\rho}_n$ if $\tau_{\varepsilon n}$ is increasing in $\tilde{\rho}_n$. Define $\Phi = \frac{V[\tilde{v}_n - p_n \mid \tilde{\rho}_n]}{\tilde{\rho}_n^2}$. From the information acquisition problem, we can derive

$$\frac{d\tau_{\varepsilon n}}{d\tilde{\rho}_n} = \frac{\frac{\partial \Phi}{\partial \tilde{\rho}_n}}{2\gamma - \frac{\partial \Phi}{\partial \tau_{\varepsilon n}}} > 0$$

This expression is positive because $\frac{\partial \Phi}{\partial \tilde{\rho}_n} > 0$ and $\frac{\partial \Phi}{\partial \tau_{\varepsilon n}} < 0$.

## A.9 Proof of Proposition 3

From the information acquisition problem when $\tilde{\rho}_n$ is known, $\tau_{\varepsilon n}$ is implicitly given by

$$\tau_{\varepsilon n} = \frac{1}{2\gamma} \frac{V_0[\tilde{v}_n - p_n]}{\tilde{\rho}_n^2}. \tag{A.9}$$

Let us consider any firm $n$ such that $n \neq \check{n}$. If the media outlet publishes a story about a firm $n$, then firm $n$ is in a high volatility risk-regime. While, if the media outlet publishes a story about firm $n'$ such that $n' > n$, then firm $n$ is in a low volatility risk-regime. Let me denote $\tau_{\varepsilon n}(n^* = n)$ as the information acquired about firm $n$ when the media outlet publishes a story about firm $n$ and $\tau_{\varepsilon n}(n^* = n')$ as the information acquired about firm $n$ when the media outlet publishes a story about firm $n'$. Thus, if $\lim\limits_{\tau_\eta \to 0} \tau_{\varepsilon n}(n^* = n) > \tau_{\varepsilon n}(n^* = n')$.